The harmful effect from alcohol by analyzing average drinks people had during a week in 1989

Xinjing Guo 1005086620

August 22, 2021

Abstract

Due to drinking threatens people's health but only a few are aware of it (2018), it is crucial to focus on the total number of drinks people have and the harmful effect of alcohol. The research questions are finding average drinks people have during a week, and what is the relationship between the total amount of alcohol consumed in a week and the damage of alcohol to their lives. Harmful effects from alcohol during lifetime and total drink over the week are chosen from the dataset national alcohol and drug survey in 1989 in website ODESI. In addition, samples are independent and they will be studied in this analysis. Using Maximum likelihood estimator, bootstrap 95% confidence interval and Bayesian 95% credible interval to find the parameter which the true mean of the number of drinks people have during a week in 1989. Using hypothesis test to observe whether the parameter is less than 9.5 which is based on the article written by Sarah Boesveld. Using a sample linear regression model to show the relationship between the total drinks people consume and the harmful effect of alcohol for people. The parameter which the true average number of drinks people consume in a week is 6.9069504 from MLE, and through bootstrap 95% confidence interval and Bayesian 95% credible interval, the results are (6.669651, 7.162786) and (6.822754, 6.977682) respectively. By hypothesis test, there has very strong evidence to support that the true average of drinks people have during a week is less than 9.5. Sample linear regression model concludes there has a string evidence that as harmful effect of alcohol increase a stage, the total number of drinks people consume during a week increases by 1.9574. These methods are quite significant as they provide the useful result and it lets people know more about drinking and the harmful of it.

Introduction

According to Global News, it reports (2018) that alcohol is the most important factor that makes people whose age is between 15 to 49 disease and premature death all over the world. Moreover, it also points out (2018) that because of alcohol, indirectly causes Canada economic losses about $14.6 billion per year. Furthermore, people do not have well understood the dangers of alcohol which only less than 20 percent of people know that drinking can directly cause seven different types of cancer (2018). Thus, since people are lack awareness and do not know the damage of drinking, studying and research data on alcohol consumption and the effect of drinking on the body are significant. Based on above, we can find variables interested from ODESI. It is a website that has more than 5600 datasets covers a number of areas. The variables are collected from the national alcohol and drug survey, 1989. The name of two variables are total drink over the week and harmful effects from alcohol during lifetime.

Based on the variables select, in this paper, there are **two research questions**. The first one is what is the average number of drinks people had during a week in 1989 in Canada and the second one is what is the relationship between the total drinks people had over a week and the harmful effect of alcohol during their lifetime. After researching the two questions, it helps people have a better understanding of alcohol and know how it effects on people's health during the lifetime. Assume we only consider people who at least drink one alcohol during a week. Besides, assuming that the data in 1989 has not changed much from the recent years. This research uses 3 sections to analyze variables which are data, method, and result. For the data section,

table and plot would be represented, they can show some meaningful aspects of variable. For the method part, five different methods would be used which are **Maximum likelihood estimator**, **bootstrap 95% confidence interval**, **Bayesian 95% credible interval**, **hypothesis test** and **sample linear regression model**. Through this part, the features, assumptions, and parameters would be explained in detail. More specifically, expect sample linear regression model, others are all finding the parameter which the true mean of drinks people had during a week in 1989 in Canada. From the method sample linear regression model, it can show the relationship between the total drinks people had during a week and the harmful effect of alcohol during their lifetime. The third section is the result. Based on the method section, we have a basic understanding of each one. Then, this part would present the result for each statistical analysis and shows what is the meaning of them.

Table 1: The summary of the total drinks over the week in 1989 in Canada

n	mean	sd	max	min	range	Q1	Q3	median	Small_Outliers	Large_Outliers
4417	6.90695	8.66787	140	1	139	2	8	4	0	323

Data

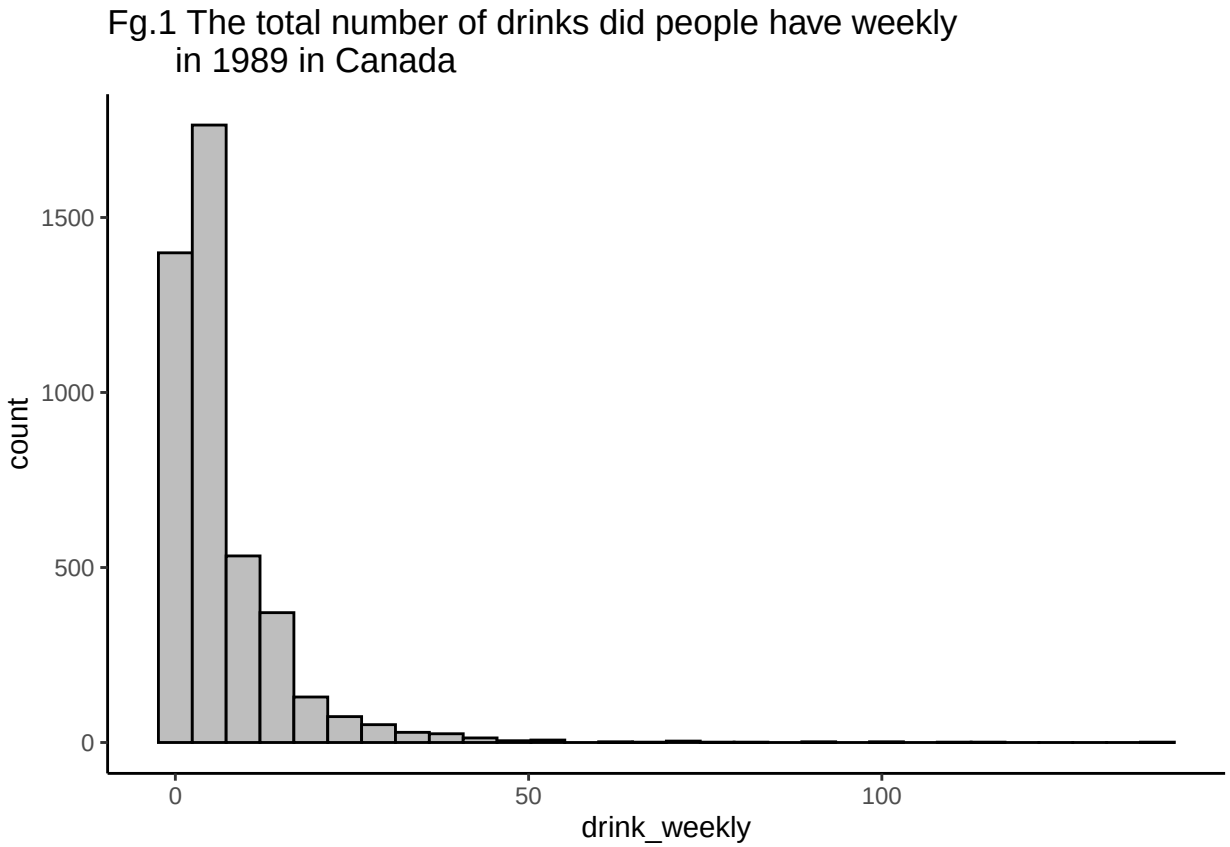
Data collection process:
Drinking has a serious effect on health and more people lose their lives because of drinking. It is important to research the number of drinks people had during a week and the harmful effect of alcohol during their lifetime. There has the dataset on the website ODESI is about the national alcohol and drug Survey in 1989. Under this dataset, there have thousands of categories to choose from. Then, select the variables interested which the total drinks people had during a week in 1898 in Canada and the harmful effect of alcohol during lifetime. This section, using plots and figures shows the variables' characters.

cleaning of data:
Since the variable has the missing value (NA), to analyze it, removing them. So, there are 4417 data available left.
DV24TOT means data that the total drink people have over the week and DV36_2_3 means the data that the stage of harm of drinking during lifetime. Rename DV24TOT to drink_weekly and DV36_2_3 to drink_harm, then, readers will have a better understanding of the data. In addition, the dataset has 11,634 rows, in order to have a clear view, select the two rows drink_weekly and drink_harm.
Moreover, since the research question is consider the number of drinks people have each week, so the 0 drink should not be considered.

Table 1 shows there are 4417 numbers available. The mean of drinks people had is 6.9 and the median value is 4. Since the mean is larger than the median, the histogram should be right-skewed. The first and third quantile which shows the total number of drinks people had are 2 and 8 separately. It means that 50% of the data are between 2 and 8. In addition, as the number of large outliers is 323 which is quite large, it still can conclude the histogram is a serious right-skewed. The range of data is 139 which is the difference between minimum (1) and maximum (140). 0 drinks are not considered in the research. NO small outliers indicate there is no extreme small number. Thus, the number of the minimum value is quite large. In another word, the number of people who consume 1 alcohol a week is large. The standard deviation of the data is 8.66787 which is not large. Then, the means that the range of confidence interval calculated is not large. The number in the summary table is quite reasonable. For people who really love drinking, it is possible that had 140 drinks per week which is about 20 drinks per day.

Table 2: The summary of the harmful effect of alcohol during lifetime

n	mean	sd	max	min	range	Q1	Q3	median
4417	0.5666742	1.157503	6	0	6	0	1	0



Fg.1 shows the total number of drinks people had weekly in 1989 in Canada. It is easy to notice histogram is seriously right-skewed. The range of the graph is between 1 and 140. The largest number is 140 and the smallest number is 1. In this study, assuming 0 drink is not be researched. The most number concentrate on the range between 1 to 10 roundly. After the number of drinks people had over 10, the index decreases. The small outliers are none because there are a lot of people who have one drink during a week. There are a lot of large outliers especially after 50. It is quite reasonable because, for people who really love drinking, they would drink too much during a week which is possible to have 140 drinks. Most are not alcoholics, so the mode of this graph is between 1 to 10. Moreover, since there are a lot of large outliers, the mean value must be greater than the median.

Table 2 shows the harmful effect of alcohol during people's lifetimes. There are seven stages to represent the harm of alcohol. it from 0 to 6. 0 means alcohol takes people least harm during their life. 6 means drinking causes people the most serious harm. From the table, it shows that the mean is 0.5666742 and the median is 0. Thus, it represents that most people have the least harmful from drinking. It is reasonable since table 1, shows the mean drinks people have is 6.9 which means most people have 6.9 drinks a week which is not too much, then, the effect of drinking is not serious. The standard deviation is 1.157503 which is quite small. It means the range of the data is small. In addition, the range is 6 since the lowest level of alcohol harm is 0 and the highest level is 6. By observing the Q1 and Q3, it shows that 50% of numbers in data are from 0 to 1. It means half the data is located in this range. It is reasonable because, figure 1, shows the mode is 2 to 8 drinks which most drinks people have are in this range. Thus, as most people do not drink too much, the level of alcohol harm for them is small.

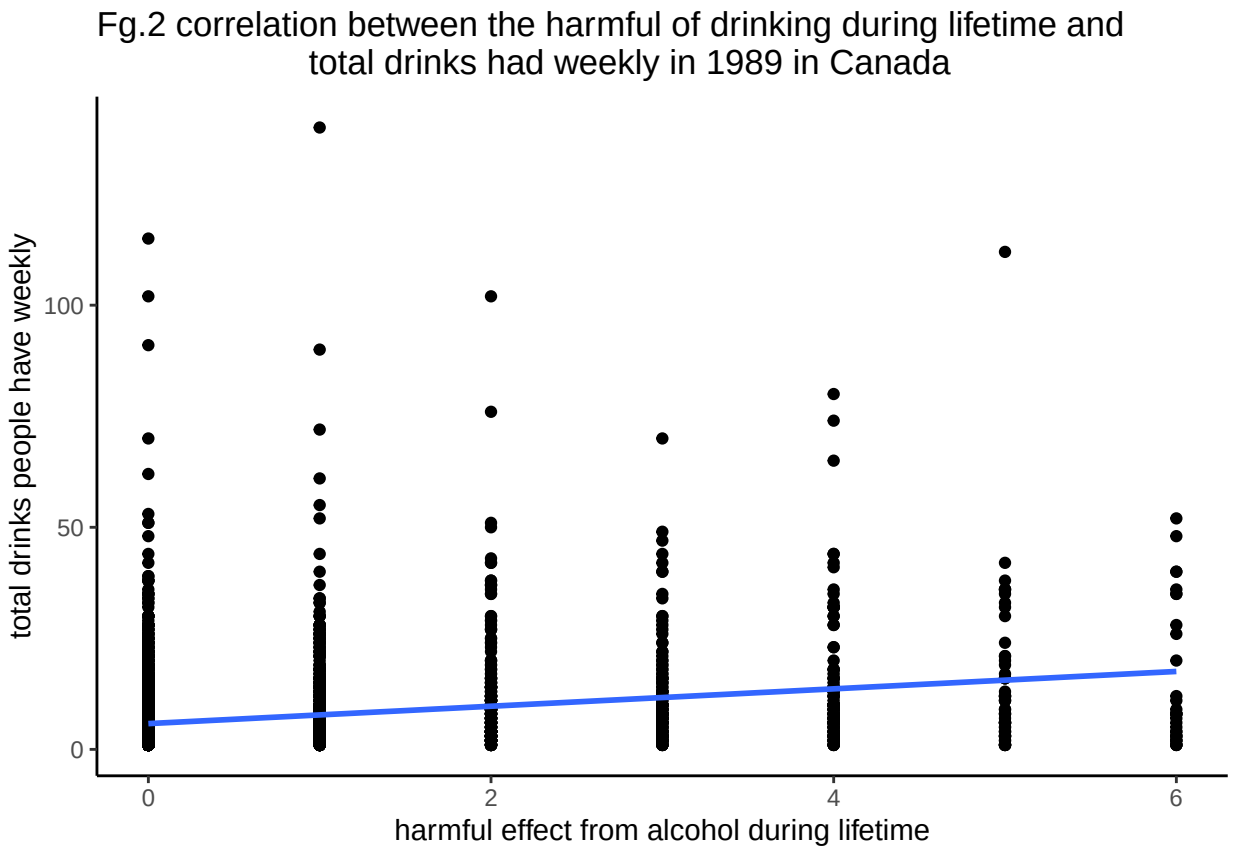


Fig.2 shows the relationship between the harmful of drinking during the lifetime and total drinks had weekly in 1989 in Canada. The harmful effect of alcohol has seven levels which are from 0 to 6. 0 means the least harmful effect of drinking and 6 means the most harmful effect the alcohol brings to people during their lifetime. In this scatter plot, the x-axis represents the stage of harmful effects from alcohol during a lifetime. y-axis represents the total drinks people have weekly. Since the values on the x-axis are positive integers, points have this special distribution.
This plot has a positive direction means as the harmful of drinking increase a unit, the number of drinks people have in a week also increases. It is reasonable as more drinking must lead to worse health. Most outliers are located when the total drinks people had is greater than 50. In addition, when the harmful effect of alcohol at 0 and 1 stages, they have more outliers than others. This plot is reasonable as people drink more, more serious of alcohol affects health.

Overall, this section using plots and table show the distribution and basic information of variable which how many drinks people had during a week in 1989 in Canada. Also, using the linear regression model analyzes the relationship between the harmful effect of alcohol during a lifetime and total drinks people had over a week. In the next section, I would analyze the variables carefully with 5 methods and show the similarities and differences between each other.

Method

Assumption: As the topic of interest is the number of drinks people had during a week, it would follow the Poisson distribution. It is because Poisson distribution means the given number of events occurring in a fixed time. The parameter λ in the Poisson distribution means the true average number of drinks do people had weekly in 1989 in Canada. In addition, only in Bayesian 95% credible interval, λ is random. In the hypothesis test, μ means the true mean of the drinks people had weekly.β in Posterior distribution means the number is close to the sample mean of the number of drinks people had during a week in the sample. H_0 means the null hypothesis which is given by the article. H_a means the alternative hypothesis. In this research, the alternative hypothesis is less than the the null hypothesis.

Under the data from the ODESI, using **maximum likelihood estimator** firstly to estimate the parameter (λ). MLE method means finding a parameter that can max likelihood under the given data. The **assumption** of the MLE method is sample is independent, the average drinks people had weekly follows a Poisson distribution with estimator λ. And λ is constant. These assumptions are all fit the topic I choose. The **parameter** λ means the true average number of drinks do people have weekly in 1989 in Canada. There are steps to calculate the MLE. Firstly, write down the Likelihood function under the given data. When the likelihood function gets, take the log and take the first derivative of the function to get the maximum likelihood. It means when put the sample data to the equation, the result is parameter λ. In this study, we have $\lambda = \bar{X}$. It represents that the maximum likelihood estimate is the average of all given data which is the average number of total drinks people had over the week.

As the result of bootstrap 95% confidence interval is an interval, it can represent the true population means more accurately, then, use this method would be better. **Assume** samples are independent. The larger confidence, the wider of interval. The **parameter** λ means the true average number of drinks do people have weekly in 1989 in Canada. For this study, using **bootstrap 95% confidence interval** can measure the range of value for the parameter which the average drinks people had weekly (the true population mean). Here are steps to calculate the bootstrap 95% confidence interval. The first step is taking the bootstrap sample of the data with replacement is true. And make sure the number of observations is the same as the original data. Then, estimates the parameter which the average number of drinks people had during a week. Then, simulate thousand times to get the distribution of the bootstrap statistics. As the interval is calculated by a confidence interval, there has 95% confidence the true population mean is contained.

If the **parameter** λ is random, the method **Bayesian 95% credible interval** should be used. The **assumption** of is the number of drinks people had weekly follows the Poisson distribution with a random parameter λ. Using Bayesian 95% credible interval can get an interval which represents there is 95% probability that the average drinks people had weekly at in Canada in 1989 (true population mean). Assume the distribution of λ follows exponential distribution with $Exp(\beta = 7)$ which assume β is close to the $\bar{X}$. Also, the likelihood function follows Poisson distribution. It is because we search the number of drinks people had during a week, it follows Poisson distribution. Then, based on these information, the posterior of λ would follow the Gamma distribution $(n\bar{X} + 1, (n + 1/\beta)^{-1})$. The shape is $n\bar{X} + 1 = 30881.97345$ and scale is $n + 1/\beta$=4471.142857^{-1}.

According to Sarah Boesveld (2015), he points out that the annual drink Canadians had is about 502 drinks a year. On average, it means that about 9.5 drinks weekly that people had. Assume the number drinks people consumed in 1989 is similar to the recent year. Then, the **hypothesis test** is the best method to test whether the true average drink people had is less than 9.5. The assumption of the hypothesis test is the sample is independent and the sample size is large enough. The sample size of the given data is 4417 which is quite large and the sample mean is approximate to normal distribution. Following the steps to calculate the p-value. The first step is to set up the null hypothesis and alternative hypothesis. Since the article mentions that the average drinks weekly people had are 9.5, the null-hypotheses is 9.5 ($H_0 : \mu = 9.5$). As this test wants to show whether the true average drinks people had during a week is less than 9.5, the alternative hypothesis is $H_a : \mu < 9.5$. In this step, μ means the true population mean of the drinks people had weekly. The second step is under the equation $t = \frac{(\bar{X}-\mu_0)}{s/(\sqrt{n})}$ to find out the p-value. When the p-value is determined, if it is less than 0.05, then it can be concluded that we have evidence to show that the true average of drinks people had is greater than 9.5.

There are some factors connect to the drinks. One of the factors interested is the harmful effect of alcohol during the lifetime. **Simple linear regression model** can represent the relationship of two numerical data. The **assumption** of this model is U_i follows normal distribution which is $U_i \sim N(0, \sigma^2)$. And the sample is independent. The **parameter** λ means the true average drinks people consume per week in 1989 in Canada. This research will show the relationship between the number of drinks people had over the week in 1989 in Canada and the harm of alcohol. The harmful effect for the alcohol has seven stages which are from 0 to 6. 0 means during the lifetime, the alcohol does not bring people harm. As the number increases, the more harmful the alcohol will bring to people. When the index is 6, it means the alcohol makes that person had the most serious harm in their lifetime.
The equation that can represent their relationship is:

$$Y_i = \alpha + \beta x_i + U_i$$

- $i = 1, \ldots, n$:n means the number of observations (4417).
- Y_i :the number of drinks people had over a week in i^{th}.
- α:the intercept of the model.
- β : Slope in the model. It means the increase in average number of drinks people had over week when the harmful effects of alcohol during lifetime increases a stage (there are 6 stages of the harmful effects of alcohol).
- x_i :The i^{th} stage of harmful effects of alcohol during lifetime.
- U_i is the i^{th} error term.
- E(U_i)=0

The relationship between the average number of drinks people had weekly and the harmful effect of alcohol during lifetime is positive. It means it is the positive correlation which as the harmful effects of alcohol increases a stage, people will drink more.

Result

Based on the five methods mentioned in last section, we have know how to use different methods to get the result want. And know the differences between them. This part would show the results and what is the meaning for each of them.

Maximum likelihood estimator (MLE):

The result of the **maximum likelihood estimator** is $\hat{\lambda}_{MLE} = 6.9069504$.
By calculating, the parameter $\hat{\lambda}_{MLE}$ equals to $\bar{X}$ (the mathematical process will be shown in the appendix), $\bar{X}$ represents the maximized likelihood. Thus, under the **maximum likelihood estimator**, the true population number of drinks people had in 1989 in Canada is $\hat{\lambda}_{MLE} = \bar{X} = 6.9069504$. The result is quite reasonable because only fewer people drink heavily, most people drink at a range between 2 to 8, thus, 6.9 drinks people have during a week is quite normal.

Bootstrap 95% confidence interval:

The answer of **bootstrap 95% confidence interval** is (6.669651, 7.162786).
For this method, repeat 1000 simulations and get the mean of each simulation, then, calculate the 0.025 quantile and 0.975 quantile to get the answer.
The result is (6.669651, 7.162786). It means that there is 95% confidence that the true average of the number of drinks people have each week is between 6.663284 and 7.155881 in 1989 in Canada. It is reasonable the range is (6.663284, 7.155881) because table 1 shows the standard deviation is 8 which is not large, then, it means the range of interval of the drink of people had is small which is consistent with the answer in this method. Moreover, since the range is small, the uncertainty of the estimator is small. This answer is reasonable because compared to the result of maximum likelihood estimator, $\bar{X}$ is in this interval. It means that they all estimate the parameter quite correctly which are both close to the true mean.

Bayesian 95% credible interval:

The result of **Bayesian 95% credible interval** is (6.822754, 6.977682).
In order to get the answer, finding the prior function and likelihood functions are essential. The prior function follows Exponential distribution $Exp(\beta = 7)$ and the likelihood function follows Poisson distribution. Then, the posterior function follows Gamma$(n\bar{X} + 1, (n + 1/\beta))^{-1}$ distribution. The answer is (6.822754, 6.977682). It means there is 95% probability that the population average of the number of drinks people had during a week in 1989 in Canada is between 6.822754 and 6.977682. The result is reasonable as the answer of this method is quite similar to the bootstrap 95% confidence interval. In addition, for most people, they are not very love to drink, it is normal they have drinks between this range.

Table 3: Linear Regression Model

term	estimate	std.error	statistic	p.value
(Intercept)	5.797769	0.1401827	41.35867	0
drink_harm	1.957353	0.1087822	17.99332	0

Hypothesis test of the mean:

The result of the Hypothesis test of the mean is $H_0 : \mu = 9.5$ and $H_a : \mu < 9.5$ (notation from the method section), the p-value is 1.285638×10^{-84}. According to the article, Sarah Boesveld (2015) mentions the average number of drinks people had each week is 9.5 ($H_0 : \mu = 9.5$). And in order to estimate whether the true average of drinks people had is less than 9.5, the alternative hypothesis is $H_a : \mu < 9.5$. According to the test statistic t(4417)=-19.8820951, the p-value get is 1.285638×10^{-84} which is less than 0.001. Thus, there is very strong evidence against the null hypothesis (H_0) which has very strong evidence to support the alternative hypothesis (H_a). In another word, the result is the average number of drinks people had during a week is less than 9.5. Therefore, the average in Sarah Boesveld's article should less than 9.5 drinks people had during a week. This result is reasonable because the number article given is larger than the 9.5 but based on table 1, the mean of data is 6.9 which means the average drink people have is 6.9. Then, it clearly shows it is incorrect. Thus, against null hypothesis is reasonable.

Simple linear regression:

The equation of Simple linear regression is:(more details and information from the method section)

$$Y_i = \alpha + \beta x_i + U_i$$

The result of **Simple linear regression** is $\widehat{drinks}_i = 5.7977689 + 1.957353 \times harmful.effect_i$.
From the table 3, it shows the relationship between the number of drinks people had during a week in 1989 in Canada and the harmful effect from the alcohol during lifetime. The intercept 5.7977689 and the slope 1.957353. The p-value is 0. Then, these have a very strong evidence support more drinking would affect health. The result $\widehat{drinks}_i = 5.7977689 + 1.957353 \times harmful.effect_i$ means they have the positive correlation. As the harmful effect of alcohol increase a unit, the total number of drinks people had during a week increases by 1.9574. This result is reasonable, it is because the health would be damaged as people drink too much. Also, the p-value in the table 3 is reasonable, it is because we truly know that number of drinks people have is proportional to the harmful effect of alcohol.

Conclusion

After the analysis, from the **maximum likelihood estimator**, the result got is 6.9069504. It means the population average number of drinks people had during a week is 6.9069504 ($\hat{\lambda}_{MLE}$). As the estimation cannot be guaranteed to be accurate. Using **bootstrap 95% confidence interval** to get the range (6.669651, 7.162786). It means 95% confidence that population average of drinks people had over a week in 1989 in Canada is between 6.669651 and 7.162786. On the other hand, the parameter λ can be random, then, **Bayesian 95% credible interval** can be used to calculate the interval. The result is (6.822754, 6.977682). It represents there is a 95% probability that the true average of drinks people had over a week in 1989 in Canada is (6.822754, 6.977682). Based on the article (2015), points out the average drink people had is 9.5. By using the **Hypothesis test of the mean**, set $H_0 : \mu = 9.5$ and $H_a : \mu < 9.5$ to calculate the p-value. The final answer of the p-value is 1.285638×10^{-84}. Thus, it can be concluded that there has very strong evidence to support the alternative hypothesis. In other words, the average number of drinks people had during a week in 1989 in Canada is less than 9.5. Lastly, based on data which the total drink people had over a week, there is a factor that relates to it which is the harmful effect of alcohol during their lifetime. In order to get the relationship between them, **Simple linear regression:** is used. The result is $\widehat{drinks}_i = 5.7977689 + 1.957353 \times harmful.effect_i$. It means that the harm of alcohol and the total number of drinks people had over a week are positively correlated. The harm of alcohol increases at one stage in a peoples' life, the total number of drinks people had would increase 1.957353. These methods methods represent data very well as the parameter calculated is similar. It means they really approximate to the true average alcohol people consume during a week in 1989. Moreover, it let more people have a deeper understanding on drink. Also, as the harmful effect of drinking has the positive correlation with the total drinks people consume, it warn people decrease the number of drinks.

The limitation of the research is the harmful effect of alcohol is only has 7 stages which are relatively small. When make a scatter plot, the value of the x-axis is only from 0 to 6 which is quite limited. So the plot is not very easy to analyze. If increase more stages of the harm of alcohol, a better view of the graph would be represented. Then, it would have a better analysis. Another limitation is the data I select is from 1989 which is quite old. The number of drinks people had will be changed as time goes by. Thus, when calculating the Hypothesis test of the mean, the article used is in 2005. Therefore, the average number of drinks people had in 2005 would have the difference in different years. Hence, in order to have a better estimation, the data is better to close to the year 2021. For future research, I would pay more attention to the number of drinks people had in recent years, and compare whether people consume more than 1989.

Bibliography

- David Robinson, Alex Hayes and Simon Couch (2021). broom: Convert Statistical Objects into Tidy Tibbles. R package version 0.7.4. https://CRAN.R-project.org/package=broom

- Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2021). dplyr: A Grammar of Data Manipulation. R package version 1.0.7. https://CRAN.R-project.org/package=dplyr

- Hao Zhu (2021). kableExtra: Construct Complex Table with 'kable' and Pipe Syntax. R package version 1.3.4. https://CRAN.R-project.org/package=kableExtra

- H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

- JJ Allaire and Yihui Xie and Jonathan McPherson and Javier Luraschi and Kevin Ushey and Aron Atkins and Hadley Wickham and Joe Cheng and Winston Chang and Richard Iannone (2020). rmarkdown: Dynamic Documents for R. R package version 2.6. URL https://rmarkdown.rstudio.com.

- Laura Hensley(2018). Global News. Alcohol is killing Canadians, so why are we still drinking? https://globalnews.ca/news/4634194/harms-of-drinking-alcohol/

- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

- Sarah Boesveld(2005). How we drink: Here's everything you need to know about Canadians' overall boozy habits.National Post. https://nationalpost.com/life/how-we-drink-from-how-much-to-how-often-heres-everything-you-need-to-know-about-canadians-boozy-habits

- Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

- Yihui Xie and J.J. Allaire and Garrett Grolemund (2018). R Markdown: The Definitive Guide. Chapman and Hall/CRC. ISBN 9781138359338. URL https://bookdown.org/yihui/rmarkdown.

- Yihui Xie and Christophe Dervieux and Emily Riederer (2020). R Markdown Cookbook. Chapman and Hall/CRC. ISBN 9780367563837. URL https://bookdown.org/yihui/rmarkdown-cookbook.

- Dataset: National Alchol and Drug Survey, 1989. Nesstar WebView.date access:August 23, 2021. http://odesi2.scholarsportal.info/webview/

Appendix

MLE:

As the topic of interest is the number of drinks people had during a week, it would follow the Poisson distribution.
- Assumption:
Sample is independent. The average drinks people had weekly follows a Poisson distribution with estimator λ. And λ is constant.
- Parameter:
λ: the true average number of drinks do people have weekly in 1989 in Canada.
X_i : the i^{th} number of drinks people had during a week.

$$Y_i \sim \text{poiss}(\lambda)$$

The first step in finding the estimator is to state the likelihood function. The likelihood function is a tool to find the parameter λ by observing the data sample.

$$\begin{aligned} L(\lambda) &= p\left(y_1, \cdots y_n \mid \lambda\right) \\ &= \prod_{i=1}^{n} p\left(y_i \mid \lambda\right) \\ &= \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{y_i}}{y_i!} \\ &= \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^{n} y_i}}{\sum_{i=1}^{n} y_i!} \end{aligned}$$

The second step is to take the log of the likelihood function. It is due to the fact that calculating the equation is quite difficult. Then, an easier way to calculate the derivative is by taking the natural logarithm of the equation. And it can not have any effect on calculating the location of maximum because the natural logarithm is a monotonically increasing function which as y increases, the result would also increase. Therefore, it is ensured that the maximum likelihood can be find out.
Here is the step:

$$\begin{aligned} l(\lambda) &= \log(L(\lambda)) \\ &= log(\frac{e^{-n\lambda} \lambda^{\sum_{i=1}^{n} y_i}}{\sum_{i=1}^{n} y_i!}) \\ &= log(e^{-n\lambda}) + log(\lambda^{\sum_{i=1}^{n} y_i}) - log(\sum_{i=1}^{n} y_i!) \\ &= -n\lambda + (\sum_{i=1}^{n} y_i) \log(\lambda) - \log(\sum_{i=1}^{n} y_i!) \\ &= -n\lambda + (\sum_{i=1}^{n} y_i) \log(\lambda) - \sum_{i=1}^{n} \log\left(y_i!\right) \end{aligned}$$

As the natural logarithm of the equation is calculated, the final step is deviating this function whcich let l'(λ)=0. Then, rearranging for $\hat{\lambda}$ to get the answer of it. The final answer is $\hat{\lambda} = \bar{y}$.
Thus, by using the MLE method, the result of the estimator which has the largest likelihood. The final answer is $\hat{\lambda} = \bar{y}$.
Here is the step:

$l'(\lambda)$=0

$$l'(\lambda) = \frac{d(l(\lambda))}{d\lambda} = -n + (\sum_{i=1}^{n} y_i)\frac{1}{\lambda} - 0 = 0$$

$$\frac{\sum_{i=1}^{n} y_i}{\lambda} = n$$

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} y_i}{n} = \bar{y}$$