The Effect of Alcohol Consumption on Life Expectancy from 2000 to 2015 over 193 Countries

Kevin Ding, Cindy Fang, Hyesun Jun, Annie Seo

Emory College of Arts & Sciences

QTM 220: Regression Analysis

Dr. David Hirshberg

15 Dec. 2021

Total Word Count: 2,144

Abstract

Alcohol consumption is a major risk factor for death and disability responsible for approximately 139 million disability-adjusted life-years globally according to the World Health Organization. (WHO) Our study seeks to determine how alcohol consumption contributes to life expectancy rates using a Life Expectancy dataset from the World Health Organization (WHO) and United Nations website collected with the help of Deeksha Russell and Duan Wang. This dataset documents information including life expectancy, alcohol consumption, population, and more for 193 countries over the course of 2000-2015. We propose four possible models to fit the data: linear model, quadratic model, logarithmic model, and quadratic model with interaction variable, BMI. By fitting four models into our data, we found that the logarithmic model fits the data better than the first two, but it needs to be further tested on people who have alcohol consumption more than 20. The interaction model seems to suggest that alcohol consumption does very little to affect one's life expectancy.

Keywords: life expectancy, alcohol consumption, public health, causal analysis, community health, epidemiology

Introduction and Literature Review

Alcohol abuse has always been an on-going world-wide public health issue. According to the data from 2011 to 2015 collected from CDC, alcohol abuse caused 95,000 deaths and 2.8 million years of potential life lost each year in the United States. In addition, 1 out of 10 deaths among adults from 20-64 years old was due to excessive drinking (CDC). Over-consuming alcohol can also lead to the development of chronic diseases such as hypertension, cardiovascular disease, immunity-weakening, and growth of tumors. The development of these chronic diseases will also influence an individual's life expectancy (DuGoff et.al, 2014). However, there are also studies that show that moderate amounts of alcohol consumption with specific kinds of drink lowered the risk for coronary heart diseases (Chadwick and Goode, 1998). Therefore, we are interested to learn whether there is a causal relationship between alcohol consumption and life expectancy. Based on current research, the causal question we aim to investigate is: what would the average life expectancy(exp) be if people consume different amounts of alcohol(alc)?

There are many clinical research and epidemiological studies investigating the area of alcohol consumption and life expectancy. Most of the studies characterized the theoretical mechanism of how the compound in alcohol affects our body metabolism system and therefore leads to the consequence of death. Some of the remaining investigate the correlation between them (Emberson and Bennett, 2006). They all didn't validate the causal significance. Therefore, we think it will be useful to perform a causal inference test on alcohol consumption and life expectancy. Our dataset is obtained from the World Health Organization and the United Nations website which contains 193 countries over the year of 2000 to 2015. Our methods planned for this study are: 1) determine the assumptions required to answer our causal questions; 2) propose models that can fit our dataset; 3) evaluate the performance of the model and analyze the errors existing in our models.

Theory, Hypotheses, Data, and Methods

Dataset

Our dataset is the Life Expectancy dataset from the World Health Organization (WHO) and United Nations website (Russell and Wang, 2017). The dataset contains information such as life expectancy, alcohol consumption, population, and more for 193 countries over the course of 2000-2015. The data was collected from WHO and the United Nations website with the help of Deeksha Russell and Duan Wang.

Hypothesis

Our hypothesis is that a higher alcohol consumption will lead to a lower life expectancy because the CDC cites that excessive alcohol could lead to heart disease, cancer, and a weakened immune system. (CDC)

Our graphs aim to predict how the life expectancy changes based on alcohol consumption recorded per capita in liters of pure alcohol.

Methods

Our approach is to try various models and combinations of regressors to find a model that best fits our data. First, we fit a simple linear model into our data to see if it fits our data well. Then we fit a quadratic model and a logarithmic model to do the same thing. Finally, we suspect that people with different BMI might have very different life expectancy. Thus, we divide our data into two parts based on the cutoff value 25, because that is the cutoff value for determining if a person is overweight. Then, we run the interaction models on people with healthy BMI and people without healthy BMI to see if alcohol consumption affects life expectancy of people with different BMI differently.

We wanted to find the modeling error as well, but the methods described in class assume that we know m(w), which is our "real world", and in our case, we do not know the "real world" function or the distribution of the data so we are not able to compute the modeling error.

Results

With this first simple model we proposed, the only regressor is the alcohol consumption per capita ($exp = \beta_0 + \beta_1 alc$). This linear model does not seem to fit well as alcohol consumption per capita increases as the data plots show that life expectancy stagnates once we get to a certain point of alcohol consumption (Figure 1). With this model, it shows that life expectancy and alcohol consumption have a positive relationship, which means the more alcohol people consume the longer they live, which does not sound to be true based on the health warnings from the CDC.

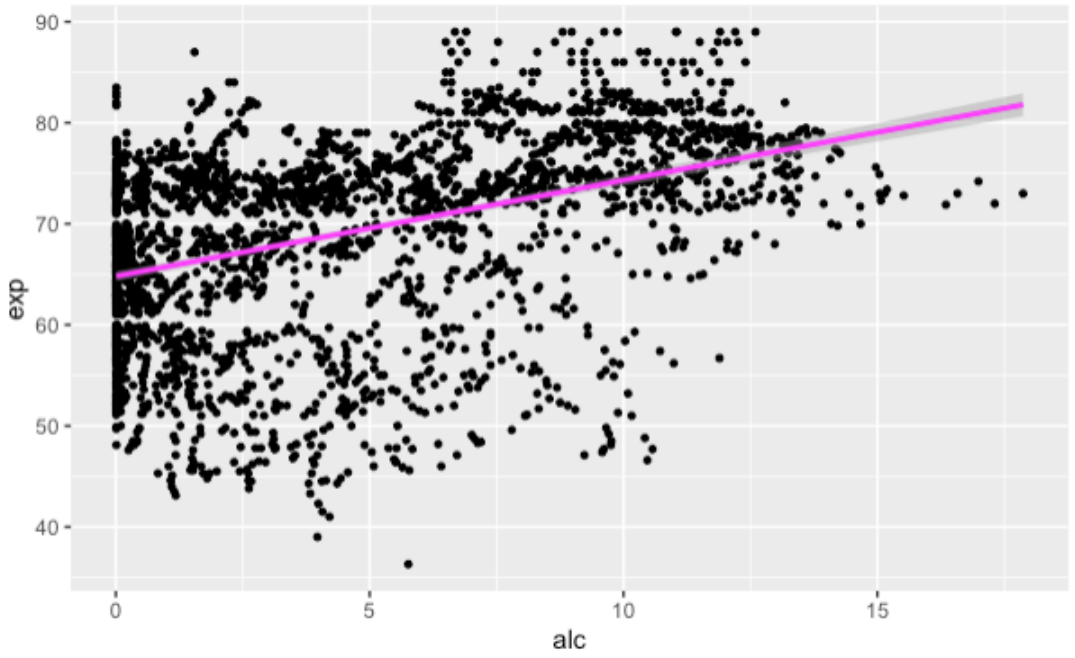

Figure 1. Simple linear predictive model of alcohol consumption on life expectancy ($exp = \beta_0 + \beta_1 alc$)

With the second model we proposed, the regressors are the alcohol consumption per capita and alcohol consumption per capita squared ($exp = \beta_0 + \beta_1 alc + \beta_2 alc^2$). This quadratic

model does not seem to fit well as alcohol consumption per capita increases as the data plots show that life expectancy stagnates once we get to a certain point of alcohol consumption (Figure 2). With this model, it shows that life expectancy and alcohol consumption have a positive relationship, which means the more alcohol people consume, they live exponentially longer, which does not sound to be true based on the health warnings from the CDC.

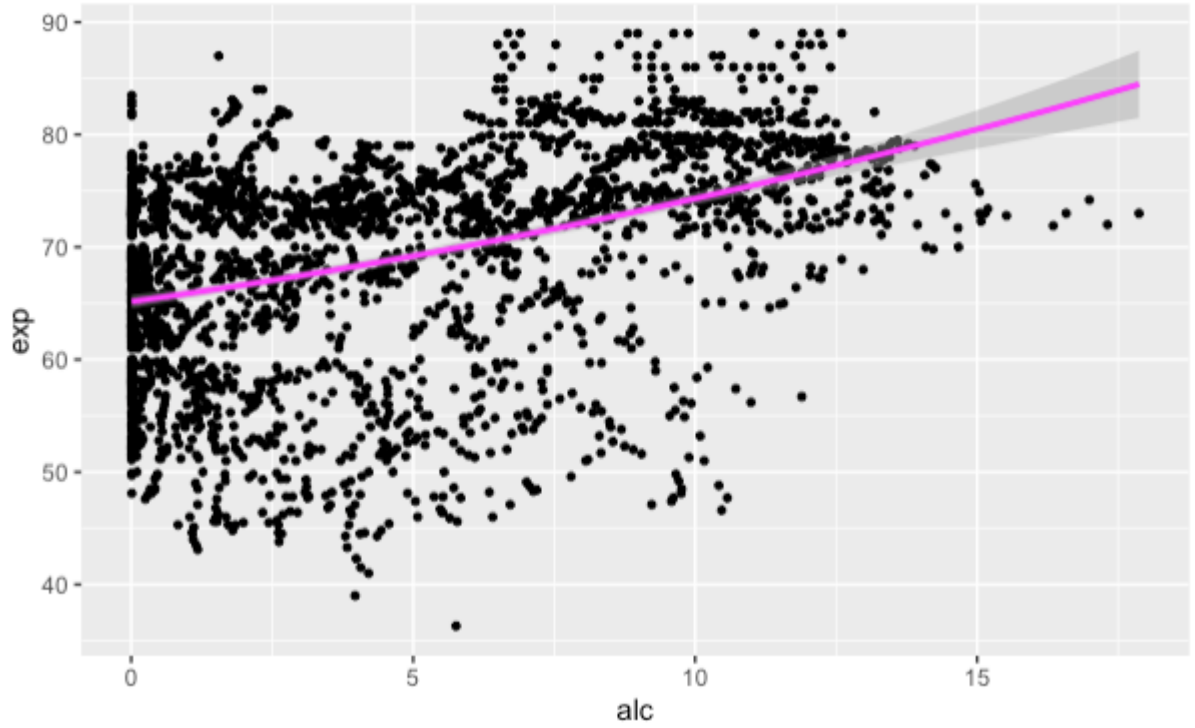

Figure 2. Quadratic predictive model of alcohol consumption on life expectancy ($exp = \beta_0 + \beta_1 alc + \beta_2 alc^2$)

With the third model we proposed, the regressor is the log of alcohol consumption per capita($exp = \beta_0 + \beta_1 log(alc)$). We chose this model to better show the data points stagnating off once alcohol consumption becomes higher than 10 (Figure 3). The logarithmic model seems to fit better than prior models because it takes into consideration this stagnation. With this model, it shows that life expectancy and alcohol consumption have a positive relationship, which means the more alcohol people consume, they live logarithmically longer, which could be plausible if the alcohol consumption is low enough to not cause any health problems. However, we feel that if our data had more information of those who have an

alcohol consumption per capita to be higher than 20, we would be able to get a better sense of how much alcohol consumption needs to occur to lower the life expectancy.

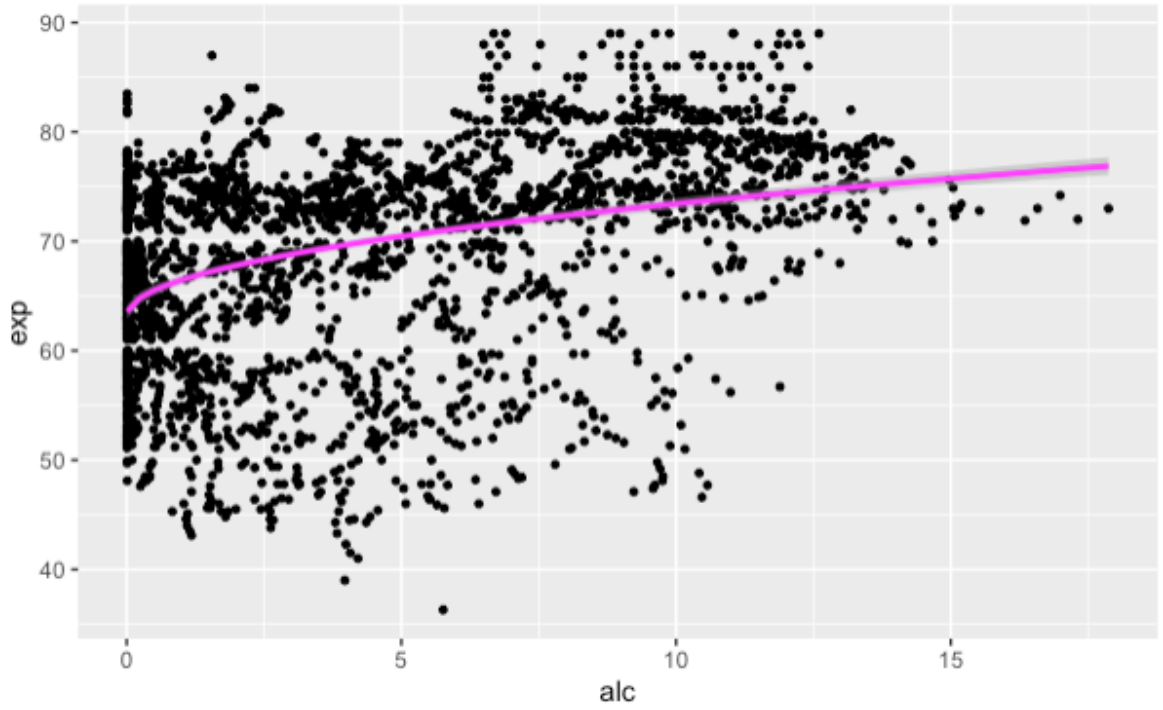

Figure 3. Logarithmic predictive model of alcohol consumption on life expectancy ($exp = \beta_0 + \beta_1 log(alc)$)

With the fourth model we proposed, the regressors are the alcohol consumption per capita and an interaction term, the log of alcohol consumption per capita multiplied by a BMI($exp = \beta_0 + \beta_1 alc + \beta_2\, log(alc) \times BMI$). If the BMI is less than 25, it is indicated by 0 and the red plots, and if it's 25 or greater, it's indicated by 1 and the blue plots. We chose the BMI threshold to be 25 because people with a BMI of 25 or higher will be considered overweight or obese by the CDC (Figure 4). For those with a BMI less than 25, the regression line is almost horizontal and flat. For those with a BMI of 25 or higher, the regression line is parabolic with a curve downwards (Figure 4). This means that of those who have a healthy BMI, alcohol consumption does little to affect their life expectancy. For those who don't have a healthy BMI, too little alcohol consumption or too much alcohol consumption can lower their life expectancy. It seems that the ideal amount of alcohol consumption for the highest life expectancy would be around 12 liters. This model provides more information than others

in that we can sort out those who are of healthy weight and those who aren't. This is useful because those who are not of healthy weight may have underlying conditions that may lower their life expectancy, and does not have to do anything with their alcohol consumption. Since the regression line with a healthy BMI shows to be horizontal, we can assume that alcohol consumption per capita does little to affect one's life expectancy.

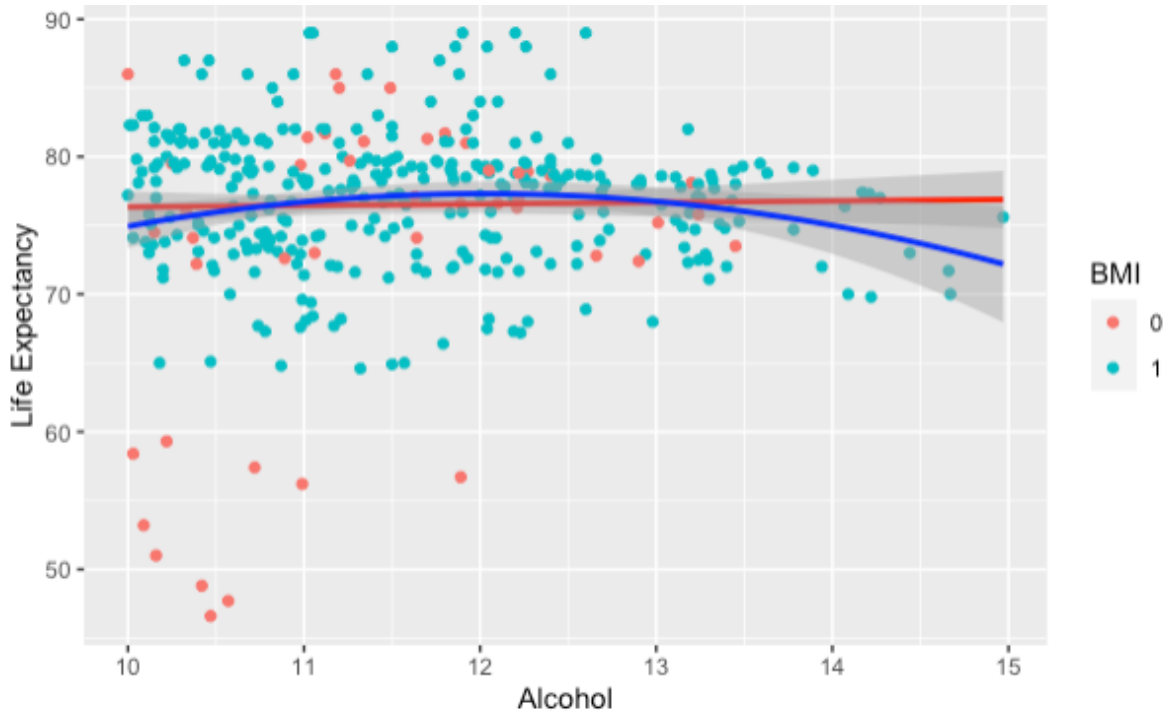

Figure 4. Interaction model of alcohol consumption on life expectancy with BMI ($exp = \beta_0 + \beta_1 alc + \beta_2 log(alc) \times BMI$)

Discussion

Overall, from our various models we can see that alcohol consumption causes insignificant differences in one's life expectancy. Our last model shows that life expectancy does not change at all despite the amount of alcohol consumed per capita amongst people with healthy BMI (Figure 4). The models seem to show that the more one may drink, the longer they may live, but we believe that this is due to the fact that the alcohol consumption cut off within our data is limited to 20 liters. We believe that if we had more data beyond the

20 liter cutoff, we will be able to see the holistic picture in that extreme amounts of alcohol causes lower life expectancy, which follows the warnings raised about alcohol consumption by the CDC. In addition, our last model also agrees with previous research on moderate amounts of alcohol consumption may reduce the risk of cardiovascular disease because of the bioactive compounds presented in specific drinks such as wine (Alcohol Res Health, 2000; Streppel et al., 2009).

We would have liked to calculate the modeling error, but the methods explained in class are done with the assumption that we know our distribution and the "real world" or m(w). We do not know this information with our dataset, so we were unable to calculate this numerically. However, we are able to estimate how well our model is fitting by seeing how well our data points fit to our regression lines. Our study contributes to the debate of the dangers of alcohol by providing data that certain amounts of alcohol will not be deadly or cause a lower life expectancy.

For future studies, we would like to do more research with bigger data to see what that cutoff is. For example, how much alcohol would need to be consumed for it to become a health concern and affect one's life expectancy. We hope the results from this study can help change the culture around drinking amongst college students and adults, to always drink in moderation and not in excessive amounts.

References

"Global Status Report on Alcohol and Health 2018." *World Health Organization*, World Health Organization, 27 Sept. 2018, https://www.who.int/publications/i/item/9789241565639.

"One in 10 Deaths among Working-Age Adults Due to Excessive Drinking." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 26 June 2014, https://www.cdc.gov/media/releases/2014/p0626-excessive-drinking.html.

Chadwick DJ, Goode JA, editors. *Alcohol and Cardiovascular Diseases: Novartis Foundation Symposium 216.* New York: John Wiley & Sons; 1998.

DuGoff, Eva H et al. "Multiple chronic conditions and life expectancy: a life table analysis." *Medical care* vol. 52,8 (2014): 688-94. doi:10.1097/MLR.0000000000000166

Emberson, Jonathan R, and Derrick A Bennett. "Effect of alcohol on risk of coronary heart disease and stroke: causality, bias, or a bit of both?." *Vascular health and risk management* vol. 2,3 (2006): 239-49. doi:10.2147/vhrm.2006.2.3.239

Health risks and benefits of alcohol consumption. *Alcohol Res Health*. 2000;24(1):5-11.

Russel, Deeksha, Duan Wang, WHO, and United Nations. Life Expectancy (WHO): Statistical Analysis on Factors Influencing Life Expectancy [.csv]. Retrieved from https://www.kaggle.com/kumarajarshi/life-expectancy-who

Streppel, M.T., M.C. Ocké, H.C. Boshuizen, F. J. Kok, and D Kromhout. "Long-term wine consumption is related to cardiovascular mortality and life expectancy independently of moderate alcohol intake: the Zutphen Study ". J. Epidemiol Community Health (2009); 63: 534-540. doi: 10.1136/jech.2008.082198

"Drinking Too Much Alcohol Can Harm Your Health. Learn the Facts." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 11 May 2021, https://www.cdc.gov/alcohol/fact-sheets/alcohol-use.htm.