Dayne Andersen-Davis
Econ 470 — Professor John Krieg
Final Project Paper

Per Capita Alcohol Consumption and Suicide Rates

An Updated Examination of US rates through 2014

There has long been a link between depression and suicide rates in the general population. And because of the generally known link between alcohol consumption and increased rates of depression [1] there is the potential for alcohol to affect the overall suicide rate. In this paper I will attempt to answer two main questions with forecasting methods, the first will be what effect does the rate of alcohol consumption, including various forms of alcohol consumption, have on the suicide rate per state, the second will be what effect does the overall unemployment rate, and alcohol consumption rate in conjunction have on the suicide rate for the country as a whole. If either of these datasets result in statistical validity I will forecast the future values for these models. All of this data and analysis will be conducted on datasets describing the United States either as a whole or on a state by state granularity level.
(If time allows I will also do an examination of overall suicide rates by men and women both on a state level and on a country level. If we are to believe the overall trends that men are more likely to drink heavily [? (come up with reference)] then this should result in a more positive relationship between men's drinking rates and their suicide rates.)

PREVIOUS WORK

This is not a new issue and there have been a number of papers addressing this exact issue[4,6], however the previous work done on this is relatively old at this date as in excludes much of the data from our modern media enhanced age. Therefore I would suggest that this topic is in need of an updated examination, for this purpose I will follow many of the same methods discussed in these previous works in an attempt to see if any differences exist. The results of

Dayne Andersen-Davis
Econ 470 — Professor John Krieg
Final Project Paper

Time series analysis of alcohol consumption and suicide mortality in the United States, 1934-1987 [4] was that of a significant relationship between alcohol consumption and male suicide rates — for those under 60 years of age — but only after the inclusion of unemployment rates as a covariate predictor variable. The results of Per capita alcohol consumption and suicide rates in the US, 1950–2002 [6] was fairly limited for men showing non-significant values for prediction of their suicide rates from alcohol consumption. There was also a statistically significant positive effect on alcohol consumption rate for men based upon an increase in the unemployment rate (.008 at 95%). For women the estimated effect of total alcohol consumption was only significant at the 10% level, but spirits was significant at the 5% level and had a positive effect of 16% for each additional liter of spirits consumed per capita. So, according to the author of the second paper there may very well be some confounding issue that arose in changing timer period after WWII that may have changed the aggregate data. He also suggest that this could very well be an issue with using aggregate data in the first place and suggest additional studies with an increased level of granularity perhaps on the state level. This is where my paper will step in both to see if there has been an additional change in the new data to return us to that original model and to see if there is any difference in the state level data which I have aggregated.

MODEL

Since I am attempting to replicate the results of the stated previous models I will use a similar model. I will attempt both a linear regression of the data formatted in the following manner: LinReg(log(Crude suicide rate) ~ alcohol consumption per capita (14 years old+) + log(unemployment_rate). I will also attempt a simple arima model using built in packages in R

that will show the effectiveness of this model. Should either of these produce reasonable output I will then proceed to use them in a forecast.

DATA

My data for this consisted of a number of data points from various governmental agencies, the suicide rates came from the wonder database query tool provided by the CDC — for this value there were male, female, and total numbers for 1968-2016. The population figures came from census data collected or estimated by July 1st of each year — these figures were already collated by the CDC in their mortality data. The unemployment rate used here comes from the figures put out by the Bureau of Labor Statistics — again for this model I have figures running from 1968-2016. The final values I used were those of total alcohol consumption split up by type: beer, wine, spirits, total — for this I have values running from 1970-2015. There were three models used to measure death rates during this time, for 1968-1978 it was ICD-8 for which I used codes E950-E959 as they were the most directly applicable to suicide, for 1979-1998 I used codes where injury intent was marked as suicide, and for 1999-2016 I again used codes where injury intent was marked as suicide. For my models I used suicide rate as defined by Crude Rate = Count / Population * 100,000. The reasons this data make sense for modeling my process is due to the aggregated method of all parts. The thought here is that due to the consistent time sampling of each figure there will be a lack of seasonality and due to all counts included being based off of total population there should be no issues with a growing or moving average.

Dayne Andersen-Davis
Econ 470 — Professor John Krieg
Final Project Paper

FINDINGS

I was able to eek some reasonable results out of these tests. However I was unable to determine an effective way to compare values from the many different combinations of variables at play here. So, I chose a limited subset of these possibilities that I thought would provide relevant and interesting output based upon previous considerations. My choices here was to run a linear regression on the log of the crude suicide rates for Maryland for women as my dependant variable, and as my predictor variables I chose consumption of spirits in each year along with the unemployment rate as measured in december of each year. I chose these values as previous studies have determine Maryland to have some of the highest rates of alcohol consumption per capita and study [6] showed a relatively strong relationship between spirits and suicide in women. The result of this regression was the following —

Log(Crude suicide rate — Dots) mapped against LinReg(Log(Crude suicide rate) — Line)

Dayne Andersen-Davis
Econ 470 — Professor John Krieg
Final Project Paper

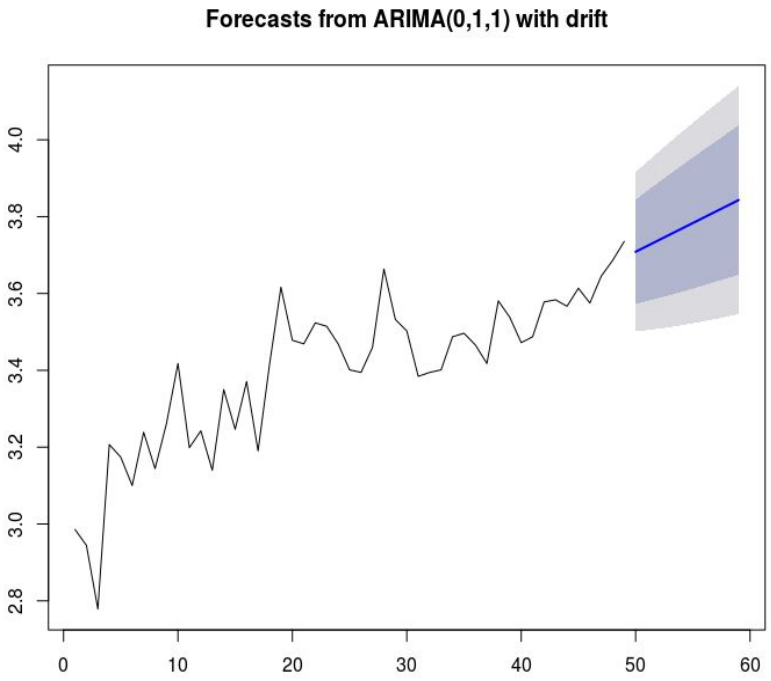
Results Linear reg. Maryland Women Spirits

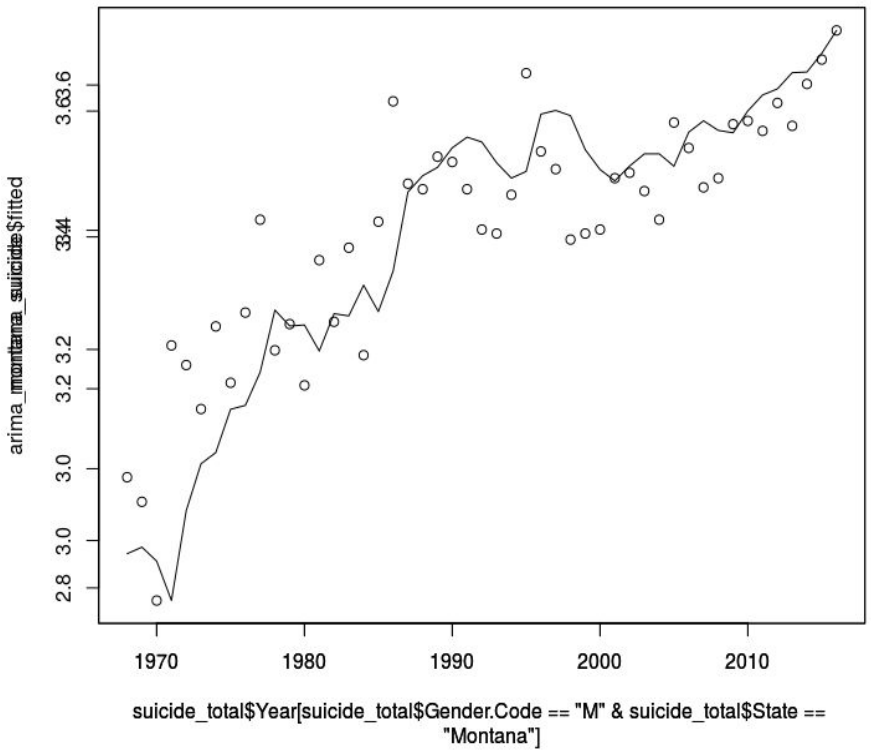
Residual standard error: 0.1226	42 degrees of freedom
Multiple R-squared: 0.7515,	Adjusted R-squared: 0.7396
F-statistic: 63.5 on 2 and 42 DF,	p-value: 2.009e-13

These values are relatively interesting as they are very close to be significant in the R^2 value and with such a high F-value it would suggest that these are reasonable values to include in our regression. However, it would also seem that there is something from my regression as I am missing a fair portion of my correlation. I think this could very well be the lagged effects of depression which the arima process models well, however I did not include this process as it seemed excessive at the time. However, looking back on this it very should be the case that this value is included in our forecasting of the model.

After this I decided to see if I could get an effective model from a simple arima process model of suicide rates. Based upon this idea I used Montana as my baseline because of their high male suicide rates. I also isolated male suicide rates as they seemed more likely to be significant. The following graphs are a result of my mapping the arima process to the available data, and then plotting a forecast for ex ante values with both an 80% confidence interval (purple) and 95% (light purple). My prediction for this model was derived from the auto.arima function in R, this function minimizes error values until it has found a model close to ideal. For my model here it produced an arima(0,1,1) process with drift. I continue my examination of this data and graph after the following figures.

Dayne Andersen-Davis
Econ 470 — Professor John Krieg
Final Project Paper





Dayne Andersen-Davis
Econ 470 — Professor John Krieg
Final Project Paper

Ex-ante forecast (Suicides per 100,000 people)

Time	Point (logged val) : e^point	Low 95%	High 95%
2017	3.708524 : 40.7935507882	3.501195 : 33.1550485781	3.915854 : 50.1919171106
2018	3.723472 : 41.4079130897	3.504395 : 33.2613146686	3.942549 : 51.5498345009
2019	3.753368 : 42.6645344425	3.508193 : 33.3878813396	3.968647 : 52.9128912446
2020	3.768316 : 42.6645344425	3.512507 : 33.5322277912	3.994228 : 54.2839172594
2021	3.783263 : 43.9592469921	3.517271 : 33.692356448	4.019360 : 55.6654685242

Coefficients:

	ma1	drift
	-0.6586	0.0149
s.e.	0.1223	0.0054

sigma^2 estimated as 0.01119: log likelihood=40.45
AIC=-74.91 AICc=-74.36 BIC=-69.29

A Ljung-Box test, with lags = 1, of the fitted models residuals produces a test statistic of .19 which being greater than .05 does suggest 0 autocorrelation in the first lag. And considering with further testing the value only grows it suggest that we have accounted for all autocorrelation in this dataset.

Dayne Andersen-Davis
Econ 470 — Professor John Krieg
Final Project Paper

CONCLUSION

I was unable to find any truly convincing correlation values from my processes, but I will also argue that there very well could be an element that I am missing from my predictor variables. So, in terms of policy changes that could be made to assist with the issues presented I would examine the growing consumption of alcohol among women in the form of spirits. For men there seems to be a recent trends of growing suicide rates in men. This suggests to me that we should start focusing more on the mental health of our male population as they may be feeling disenfranchised or hopeless.

References

1. L. Sher; Depression and alcoholism, *QJM: An International Journal of Medicine*, Volume 97, Issue 4, 1 April 2004, Pages 237–240, https://doi.org/10.1093/qjmed/hch045
2. Alcohol and Alcoholism, Volume 41, Issue 5, 1 September 2006, Pages 473–478, https://doi.org/10.1093/alcalc/agl060
3. Haughwout, S.P., and Slater, M.E., 2017. Surveillance Report #108: Apparent Per Capita Alcohol Consumption: National, State, and Regional Trends, 1977—2015. Bethesda, MD: NIAAA, Alcohol Epidemiologic Data System.
4. Time series analysis of alcohol consumption and suicide mortality in the United States, 1934-1987. Caces F, Harford T J Stud Alcohol. 1998 Jul; 59(4):455-61.
5. https://www.thebalance.com/unemployment-rate-by-year-3305506
6. Landberg J. Per capita alcohol consumption and suicide rates in the US, 1950–2002. Suicide Life-Threat. Behav. 2009;39:452–460. doi: 10.1521/suli.2009.39.4.452.