Gun Violence

Data Analysis

Corey Hannum

Zahid Aziz

Ben Cooper

Computational Statistics (CSIT 395/495)

Professor Zhu

Abstract

Gun control in the United States has been a major political issue for many years and that issue only seems to grow with each year. Contradicting laws from state to state make it difficult to analyse the country as a whole. Using some data from the CDC and looking at individual states, our team has done statistical analysis to determine if certain laws have a statistically significant impact on the rate of gun violence per state. In this study we used linear regression, chi-square, and one and two way proportion testing to determine the statistical significance of certain laws effect on gun violence.

Dataset

Our team gathered much of the data for our analysis from the CDC Firearm Mortality by State 2014. This data set contains information on the death rates by firearms for each individual state. We also gathered some data through the US 2010 census for data on state populations.

Experiments

One test that we performed was a chi-square test of independence. Our team performed two chi-square tests, one to determine if there is a relationship between laws requiring background checks on private sales, and the region of the US that a state is in (North East, South, Midwest, and West), and another to determine if there is a relationship between the level of permit required for concealed carry and the region of the US that a state is in.



Figure 1

The first chi-square test done was finding a relationship between background checks for private sales and region of the US. The data in figure 1 shows the number of states in each region with or without required background checks.

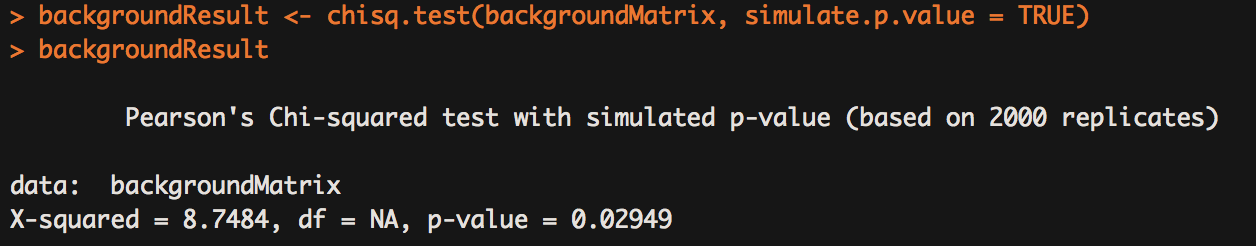


Figure 2

Figure 2 above shows the chi-square test run in R which outputs a p-value of 0.02949. This p-value being less than 0.05 tells us to reject the null hypothesis that there is no relationship between region of the US and the laws requiring background checks for private sales. Figure 3 below shows the expected values if there were no dependence on region. As is visible, those values are significantly different than those in Figure 1.

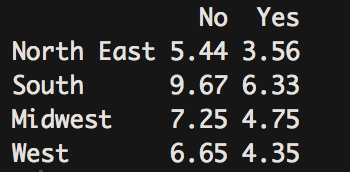


Figure 3

The next chi-square test that we performed was to determine if there is a relationship between states with unrestricted, shall-issue, and may-issue licenses for concealed carry compared to the region of the US that the state is in. Unrestricted means that there are no laws requiring a permit for concealed carry, shall-issue means that one only needs to meet specified requirements in order to receive a concealed carry license, and may-issue means that on must meet specified requirements and local authority must approve the license. Figure 4 below shows the observed number of states in each category (top), and the values that would be expected given that there is no relationship between the categories(bottom).

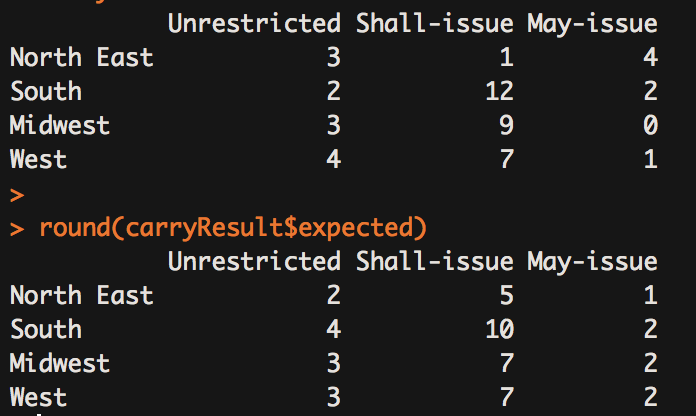


Figure 4

When we ran this chi-square test, a p-value of .02299 was output. This p-value being less than 0.05 shows that we should reject the null hypothesis that there is no relationship between the region of the US and the laws requiring certain permits for concealed carry of firearms. Figure 5 below shows the results of this chi-square test.

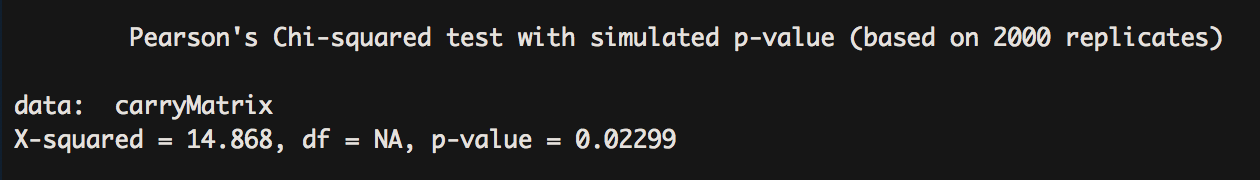
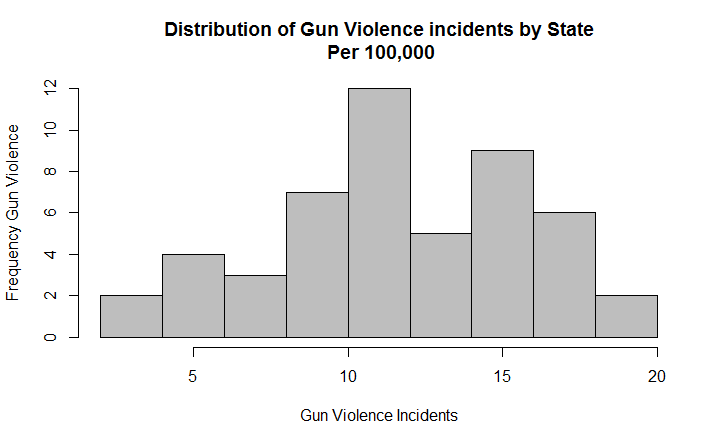


Figure 5

Histogram

The figure below shows the Histogram for gun-related deaths by state per 100,000 population. The Histogram shows the distribution is roughly unimodal and symmetric.

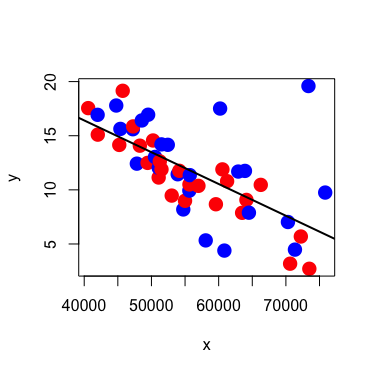


Below are the R script commands to generate this histogram.



Linear Regression

In this portion of our experimentation, we attempted to find a correlation between the number of gun-related deaths and the median household income per state. We indeed discovered a negative correlation with a p-value of 3.37e-7 between the amount of gun-related deaths and median household income. Although outliers were present, we found on average that states with a higher median household income also had less gun-related deaths.



We also factored in whether or not a state allowed open gun carry, but with a p-value of 0.013, we found this to have no real significant correlation to the amount of gun-related deaths in each state.

One Way ANOVA test

For the One Way ANOVA test, we used the average homicides per 100,000 people in the United States for each decade from the 1960s to 2010. Our goal was to discover if there are any differences for each decade, and our finding verified that this is the case. As is shown in the results below, our test has a ‘\*\*\*’ significance level which means that the differences are very significant. Looking at the bar chart below it is also visible that the 1960’s and the 2000’s have significantly lower homicide by firearm rates compared to the 1970’s 1980’s and 1990’s. These results again help to prove that there is a very significant difference in homicide rates between the decades from the 1960’s to the 2000’s.

Df Sum Sq Mean Sq F value Pr(>F)

Decade 4 120.4 30.100 36.32 1.48e-13 \*\*\*

Residuals 45 37.3 0.829

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

