|  |  |
| --- | --- |
|  | 🞂Categorical Data Analysis  Ideology |
|  |  |
|  | **Ye In Jeon** 🞂STAT 2183 🞂 2015-10-15 |
|  |  |

Categorical Data Analysis

# Introduction

People have different ideology. There can be various factors that can affect the political preference of a person. To figure it out the factors, data, which is derived from a subset of the 2008 American National Election Study, is used. With this data, it can possibly find the trend of ideology by gender, marriage, and ethnic.

## Method and Results

Before testing the relationships, each portion of ideology can be found for overall study.



Figure . Pie Chart

Through Figure 1, it can be found that the population of moderate, slightly liberal, or slightly conservative are 56.78% of the all population. Also, the portion of population, who are either liberal or extremely liberal, is 13.42%. Moreover, the 29.8% of the population is either conservative of extremely conservative.

To find the relationships between ideology and other factors, the chi-square test is used. So it can be said the null hypothesis is ‘The rows and columns are independent’ and the alternative hypothesis is ‘The rows and columns are dependent.’ First, the relationship between gender and ideology are tested.

Table . Gender by Ideology

| **Table of gender by ideology** | | | |
| --- | --- | --- | --- |
| **gender(gender)** | **ideology(ideology)** | | |
| **Frequency Row Pct** | **1 Republican** | **2 Other** | **Total** |
| **Male** | 224 65.69 | 117 34.31 | 341 |
| **Female** | 248 57.14 | 186 42.86 | 434 |
| **Total** | 472 | 303 | 775 |

In Table 1, the rows indicate gender and the columns shows ideology. Also, it can be found that both male and female tend to be Republican.

|  |
| --- |
| ***Statistics for Table of gender by ideology*** |

Table . Chi-square Table

| **Statistic** | **DF** | **Value** | **Prob** |
| --- | --- | --- | --- |
| **Chi-Square** | 1 | 5.8576 | 0.0155 |
| **Likelihood Ratio Chi-Square** | 1 | 5.8855 | 0.0153 |
| **Continuity Adj. Chi-Square** | 1 | 5.5041 | 0.0190 |
| **Mantel-Haenszel Chi-Square** | 1 | 5.8500 | 0.0156 |
| **Phi Coefficient** |  | 0.0869 |  |
| **Contingency Coefficient** |  | 0.0866 |  |
| **Cramer's V** |  | 0.0869 |  |

In Table 2, p-value of Chi-Square is noted. P-value, which is 0.0155, is smaller than the value of alpha, 0.05. Therefore, the null hypothesis can be rejected, which means they are dependent. So, it can be said that there is a relationship between gender and ideology.

**Table 3. Fisher's Test Table**

| **Fisher's Exact Test** | |
| --- | --- |
| **Cell (1,1) Frequency (F)** | 224 |
| **Left-sided Pr <= F** | 0.9938 |
| **Right-sided Pr >= F** | 0.0094 |
|  |  |
| **Table Probability (P)** | 0.0032 |
| **Two-sided Pr <= P** | 0.0176 |

In Table 3, it can be seen that p-value of the fisher test is 0. 0176. Therefore, the null hypothesis, which the preference of male and female toward republican are equal, can be rejected because p-value is smaller than value of alpha, 0.05. So, it can be said that men are more likely than women to identify with the Republican Party.

Table . Confidence Interval Table

| **Estimates of the Relative Risk (Row1/Row2)** | | | |
| --- | --- | --- | --- |
| **Type of Study** | **Value** | **95% Confidence Limits** | |
| **Case-Control (Odds Ratio)** | 1.4359 | 1.0707 | 1.9256 |
| **Cohort (Col1 Risk)** | 1.1496 | 1.0279 | 1.2857 |
| **Cohort (Col2 Risk)** | 0.8006 | 0.6669 | 0.9610 |

In Table 4, the confidence intervals are noted. Since the ratio confidence interval does not include 1, it can be said that 95%confidence that there is a substantial evidence that the gender is related to the ideology.

Table . Odd Ratio Table

| **Odds Ratio (Case-Control Study)** | |
| --- | --- |
| **Odds Ratio** | 1.4359 |
|  |  |
| **Asymptotic Conf Limits** |  |
| **95% Lower Conf Limit** | 1.0707 |
| **95% Upper Conf Limit** | 1.9256 |
|  |  |
| **Exact Conf Limits** |  |
| **95% Lower Conf Limit** | 1.0597 |
| **95% Upper Conf Limit** | 1.9477 |

|  |
| --- |
| ***Sample Size = 775*** |

Table 5 shows the odds ratio. This indicates that males are 1.4359 times more likely than women to be identified as Republican. It is because odds of male republican is 65.69/34.31=1.91 and odds of female republican 57.14/42.86=1.33.



Figure . Barchart of Ideology by gender

Through Figure 2, it can be seen that both male and female are likely to be republicans.

Then the relationship between marriage and ideology is tested.

**Table 6. Marriage stat by Ideology**

| **Table of marstat by ideology** | | | |
| --- | --- | --- | --- |
| **marstat(marstat)** | **ideology(ideology)** | | |
| **Frequency Row Pct** | **1 Republican** | **2 Other** | **Total** |
| **1 Married** | 326 66.80 | 162 33.20 | 488 |
| **2 Not Married** | 137 52.29 | 125 47.71 | 262 |
| **Total** | 463 | 287 | 750 |

In Table 6, the rows indicate rather the person is married or not and the columns shows ideology. Also, it can be found that both group who married and not married more likely to be Republican.

|  |
| --- |
| ***Statistics for Table of marstat by ideology*** |

Table . Chi-square test Table

| **Statistic** | **DF** | **Value** | **Prob** |
| --- | --- | --- | --- |
| **Chi-Square** | 1 | 15.2001 | <.0001 |
| **Likelihood Ratio Chi-Square** | 1 | 15.0629 | 0.0001 |
| **Continuity Adj. Chi-Square** | 1 | 14.5919 | 0.0001 |
| **Mantel-Haenszel Chi-Square** | 1 | 15.1798 | <.0001 |
| **Phi Coefficient** |  | 0.1424 |  |
| **Contingency Coefficient** |  | 0.1409 |  |
| **Cramer's V** |  | 0.1424 |  |

I

n Table 7, p-value of Chi-Square is noted. P-value, which is 0.0001, is smaller than the value of alpha, 0.05. Therefore, the null hypothesis can be rejected, which means they are dependent. So, it can be said that there is a relationship between marriage and ideology.

Table . Fisher test Table

| **Fisher's Exact Test** | |
| --- | --- |
| **Cell (1,1) Frequency (F)** | 326 |
| **Left-sided Pr <= F** | 1.0000 |
| **Right-sided Pr >= F** | 7.119E-05 |
|  |  |
| **Table Probability (P)** | 3.380E-05 |
| **Two-sided Pr <= P** | 1.120E-04 |

In Table 8, it can be seen that p-value of the fisher test is 1.120E-04. Therefore, the null hypothesis, which the preference of married people and not married people toward republican are equal, can be rejected because p-value is smaller than value of alpha, 0.05. So, it can be said that married respondents are more likely than others to identify with the Republican Party.

Table . Confidence Interval Table

| **Estimates of the Relative Risk (Row1/Row2)** | | | |
| --- | --- | --- | --- |
| **Type of Study** | **Value** | **95% Confidence Limits** | |
| **Case-Control (Odds Ratio)** | 1.8361 | 1.3507 | 2.4959 |
| **Cohort (Col1 Risk)** | 1.2776 | 1.1201 | 1.4571 |
| **Cohort (Col2 Risk)** | 0.6958 | 0.5820 | 0.8319 |

In Table 9, the confidence intervals are noted. It is found that 95%confidence that there is substantial evidence that the marriage is related to the ideology because the interval does not include 1.

Table . Odd Ratio Table

| **Odds Ratio (Case-Control Study)** | |
| --- | --- |
| **Odds Ratio** | 1.8361 |
|  |  |
| **Asymptotic Conf Limits** |  |
| **95% Lower Conf Limit** | 1.3507 |
| **95% Upper Conf Limit** | 2.4959 |
|  |  |
| **Exact Conf Limits** |  |
| **95% Lower Conf Limit** | 1.3341 |
| **95% Upper Conf Limit** | 2.5247 |

|  |
| --- |
| ***Sample Size = 750*** |

Table 10 shows the odds ratio. This indicates that married people are 1.8361 times more likely than not married people to be identified with the Republican Party. It is because odds of married republican is 66.80/33.20=2.01 and odds of not married republican is 52.29/47.71=1.10.



Figure . Barchart of Ideology by Marriage

Through Figure 3, it can be seen that both married and not married people are tend to be republicans.

Next, the relationship between ethnic and ideology is tested.

Table . Table of Ethnic by Ideology

| **Table of ethnic by ideology** | | | |
| --- | --- | --- | --- |
| **ethnic(ethnic)** | **ideology(ideology)** | | |
| **Frequency Row Pct** | **1 Republican** | **2 Other** | **Total** |
| **1 Other** | 65 56.03 | 51 43.97 | 116 |
| **2 White** | 405 61.83 | 250 38.17 | 655 |
| **Total** | 470 | 301 | 771 |

In Table 11, ‘Other’ variable in column can be considered as ‘democrats.’ The rows indicate ethnic and the columns shows ideology. Also, it can be found that both white and other races more tend to prefer Republican Party instead of Democratic Party.

|  |
| --- |
| ***Statistics for Table of ethnic by ideology*** |

Table . Chi-square test Table

| **Statistic** | **DF** | **Value** | **Prob** |
| --- | --- | --- | --- |
| **Chi-Square** | 1 | 1.3918 | 0.2381 |
| **Likelihood Ratio Chi-Square** | 1 | 1.3768 | 0.2406 |
| **Continuity Adj. Chi-Square** | 1 | 1.1589 | 0.2817 |
| **Mantel-Haenszel Chi-Square** | 1 | 1.3900 | 0.2384 |
| **Phi Coefficient** |  | -0.0425 |  |
| **Contingency Coefficient** |  | 0.0424 |  |
| **Cramer's V** |  | -0.0425 |  |

In Table 12, p-value of Chi-Square is noted. P-value, which is 0.2381, is bigger than the value of alpha, 0.05. Therefore, the null hypothesis cannot be rejected, which means they are independent. So, it can be said that there is no relationship between ethnic and ideology.

Table . Fisher test Table

| **Fisher's Exact Test** | |
| --- | --- |
| **Cell (1,1) Frequency (F)** | 65 |
| **Left-sided Pr <= F** | 0.1410 |
| **Right-sided Pr >= F** | 0.8998 |
|  |  |
| **Table Probability (P)** | 0.0408 |
| **Two-sided Pr <= P** | 0.2566 |

In Table 13, it can be seen that p-value of the fisher test is 0.2566. Therefore, the null hypothesis, which the preference of white group and other race group toward republican are equal, cannot be rejected because p-value is bigger than value of alpha, 0.05. So, it can be said that there is no statistical evidence of relationship between race and political preference.

Table . Confidence Interval Table

| **Estimates of the Relative Risk (Row1/Row2)** | | | |
| --- | --- | --- | --- |
| **Type of Study** | **Value** | **95% Confidence Limits** | |
| **Case-Control (Odds Ratio)** | 0.7867 | 0.5278 | 1.1726 |
| **Cohort (Col1 Risk)** | 0.9062 | 0.7630 | 1.0764 |
| **Cohort (Col2 Risk)** | 1.1519 | 0.9176 | 1.4460 |

In Table 14, the confidence intervals are noted. The ratio confidence interval includes 1. So, it can be said that there is no substantial evidence that the ethnic is related to the ideology.

Table . Odd Ratio Table

| **Odds Ratio (Case-Control Study)** | |
| --- | --- |
| **Odds Ratio** | 0.7867 |
|  |  |
| **Asymptotic Conf Limits** |  |
| **95% Lower Conf Limit** | 0.5278 |
| **95% Upper Conf Limit** | 1.1726 |
|  |  |
| **Exact Conf Limits** |  |
| **95% Lower Conf Limit** | 0.5183 |
| **95% Upper Conf Limit** | 1.1995 |

|  |
| --- |
| ***Sample Size = 771*** |

Table 15 shows the odds ratio. This indicates that white group is 0.7867 times more likely than other race group to be Republican. It is because odds of white republican is 61.83/38.17=1.62 and odds of other race republican 56.03/43.97=1.27.



Figure . Barchart of Ideology by Ethnic

In Figure 4, we can say ‘Other’ as ‘Democrats’ an. Through Figure 4, it can be seen that both white race group and other race group are likely to be Republican Party than Democratic Party.

Conclusion & Results

To find the relationship between various factors, such as gender, marriage, and race, Chi-square test and Fisher test are used. Also, odds confidence interval and odds ratio are calculated for more evidence. Before the research, it was expected that there are no relationship between gender, marriage and ideology. Also, it was expected that white people tend to prefer Republican Party than other race group.

However, the result of the tests above indicates was different from the expectation which is made before research. There is a relationship between gender and ideology. Moreover, it can be said that men are more likely than women to identify with the Republican Party. Also, there is a relationship between marriage and ideology. As a result, it can be said that married respondents are more likely than others to identify with the Republican Party. In contrast, there is no relationship between ethnic and ideology. So, it cannot be said that whites are less likely than others to identify with the Democratic Party.

Through this research it can be found that Republican Party has to put their effort on married people and male to gain more supporters

Appendix

SAS Codes

ods rtf file="Project\_Analysis2.rtf";

**proc** **format**;

value idformatpie

**0**,**8**,**9** = "3 Drop"

**1** = "Extremely Liberal"

**2** = "Liberal"

**3** = "Slightly Liberal"

**4** = "Moderate"

**5** = "Slightly Conservative"

**6** = "Conservative"

**7** = "Extremely Conservative"

;

value idformat

**5**,**6**,**7** = "1 Republican"

**1**,**2**,**3**,**4** ="2 Other"

**0**,**8**,**9** = "3 Drop"

;

value marformat

**1** = "1 Married"

**2**,**3**,**4**,**5** = "2 Not Married"

**6**,**8**,**9**,**7**,**0** = "3 Drop"

;

value genderformat

**1**= "Male"

**2**= "Female"

;

value ethnicformat

**1**,**2**,**3**,**4**,**6** = "1 Other"

**5** = "2 White"

**7**,**8**,**9**,**.** = "3 Drop"

;

**run**;

**proc** **gchart** data=Project2;

pie ideology/ percent=arrow discrete noheading value=inside;

format ideology idformatpie.;

where (put(ideology,idformatpie.) ne '3 Drop');

**run**;

/\*Relationship between gender and Ideology\*/

**proc** **sort** data=Project2;

by descending ideology;

**run**;

**proc** **freq** data=Project2 order=data;

table gender \* ideology /chisq nocol nopercent;

exact or;

format gender genderformat. ideology idformat.;

where (put(ideology,idformat.) ne '3 Drop');

**run**;

**proc** **gchart** data=Project2;

vbar ideology / group=gender type=pct g100 discrete

patternid=midpoint;

format gender genderformat. ideology idformat.;

where (put(ideology,idformat.) ne '3 Drop');

**run**;

/\*Relationship between Marstat and Ideology\*/

**proc** **sort** data=Project2;

by descending ideology;

**run**;

**proc** **freq** data=Project2 order=data;

table marstat \* ideology /chisq nocol nopercent;

exact or;

format marstat marformat. ideology idformat.;

where (put(ideology,idformat.) ne '3 Drop')and (put(marstat,marformat.) ne '3 Drop');

**run**;

**proc** **gchart** data=Project2;

vbar ideology / group=marstat type=pct g100 discrete

patternid=midpoint;

format marstat marformat. ideology idformat.;

where (put(ideology,idformat.) ne '3 Drop')and (put(marstat,marformat.) ne '3 Drop');

**run**;

/\*Relationship between Ethnic and Ideology\*/

**proc** **sort** data=Project2;

by descending ideology;

**run**;

**proc** **freq** data=Project2 order=data;

table ethnic \* ideology /chisq nocol nopercent;

exact or;

format ethnic ethnicformat. ideology idformat.;

where (put(ideology,idformat.) ne '3 Drop')and (put(ethnic,ethnicformat.) ne '3 Drop');

**run**;

**proc** **gchart** data=Project2;

vbar ideology / group= ethnic type=pct g100 discrete

patternid=midpoint;

format ethnic ethnicformat. ideology idformat.;

where (put(ideology,idformat.) ne '3 Drop')and (put(ethnic,ethnicformat.) ne '3 Drop');

**run**;

ods rtf close;