Hypothesis Testing

November 20, 2023

[37]:
```
import numpy as np
from scipy.stats import chi2_contingency
import scipy.stats as stats
```

1 Alcohol and Substance Abuse: Chi Squared Tests (Income, Region, Race, Gender)

1.1 Race

[8]:
```
# Create a contingency table
observed_data = np.array([[142,23890],
                          [146,118733],
                          [383,176839],
                          [1001,474737],
                          [1130,581671],
                          [4004,1314526]])
```

[15]:
```
# Perform the chi-squared test
chi2, p, dof, expected = chi2_contingency(observed_data)
```

[17]:
```
chi2
```

[17]:
```
448.2708136665917
```

[18]:
```
p
```

[18]:
```
1.1594774961384505e-94
```

[12]:
```
# Check the p-value to determine statistical significance
alpha = 0.05  # Set your chosen significance level
if p < alpha:
    print("Reject the null hypothesis: The number of alcohol abuse cases is␣
 ↪dependent on race.")
else:
    print("Fail to reject the null hypothesis: The number of alcohol abuse␣
 ↪cases is independent of race.")
```

```
Reject the null hypothesis: The number of alcohol abuse cases is dependent on
race.
```

1.2 Gender

[19]:
```
# Create a contingency table
observed_data = np.array([[4488,1354413],
                          [2708,1544041],
                          ])
```

[20]:
```
# Perform the chi-squared test
chi2, p, dof, expected = chi2_contingency(observed_data)
```

[22]:
```
chi2
```

[22]:
```
704.5858230765883
```

[23]:
```
p
```

[23]:
```
3.0094796136033507e-155
```

[21]:
```
# Check the p-value to determine statistical significance
alpha = 0.05  # Set your chosen significance level
if p < alpha:
    print("Reject the null hypothesis: The number of alcohol abuse cases is␣
 ↪dependent on gender.")
else:
    print("Fail to reject the null hypothesis: The number of alcohol abuse␣
 ↪cases is independent of gender.")
```

```
Reject the null hypothesis: The number of alcohol abuse cases is dependent on
gender.
```

1.3 Region

[27]:
```
# Create a contingency table
observed_data = np.array([[1467,662619],
                          [2602,1172837],
                          [1257,388398],
                          [1871,674599]])
```

[28]:
```
# Perform the chi-squared test
chi2, p, dof, expected = chi2_contingency(observed_data)
```

[29]:
```
chi2
```

[29]:
```
163.5906715202513
```

[30]:

```
p
```

[30]:

```
3.077423050012854e-35
```

[31]:

```
# Check the p-value to determine statistical significance
alpha = 0.05  # Set your chosen significance level
if p < alpha:
    print("Reject the null hypothesis: The number of alcohol abuse cases is␣
 ↪dependent on region.")
else:
    print("Fail to reject the null hypothesis: The number of alcohol abuse␣
 ↪cases is independent of region.")
```

```
Reject the null hypothesis: The number of alcohol abuse cases is dependent on
region.
```

1.4 Income

[32]:

```
# Create a contingency table
observed_data = np.array([[1540,557327],
                          [1918,777093],
                          [1690,678509],
                          [2049,885524]])
```

[33]:

```
# Perform the chi-squared test
chi2, p, dof, expected = chi2_contingency(observed_data)
```

[34]:

```
chi2
```

[34]:

```
27.833423644850136
```

[35]:

```
p
```

[35]:

```
3.936499874097809e-06
```

[36]:

```
# Check the p-value to determine statistical significance
alpha = 0.05  # Set your chosen significance level
if p < alpha:
    print("Reject the null hypothesis: The number of alcohol abuse cases is␣
 ↪dependent on income.")
else:
    print("Fail to reject the null hypothesis: The number of alcohol abuse␣
 ↪cases is independent of income.")
```

```
Reject the null hypothesis: The number of alcohol abuse cases is dependent on
income.
```

1.5 Age : Correlation Test

[43]:

```python
countofasa_age =␣
 ↪[296,24,19,18,13,15,14,17,17,33,20,23,65,84,181,370,691,707,1086,1739,1765]

# Calculate the Pearson correlation coefficient
correlation_coefficient, p_value = stats.pearsonr(age_alcohol_abuse,␣
 ↪countofasa_age)

# Output the results
print(f"Pearson correlation coefficient: {correlation_coefficient:.2f}")
print(f"P-value: {p_value:.2f}")

# Interpret the results
if p_value < 0.05:  # You can choose your significance level
    print("There is a significant correlation between age and the count of␣
 ↪cases.")
else:
    print("There is no significant correlation between age and the count of␣
 ↪cases.")
```

```
Pearson correlation coefficient: 0.73
P-value: 0.00
There is a significant correlation between age and the count of cases.
```

[ ]:

Alcohol & Substance Abuse ANOVA & Kruskal-Wallis

November 20, 2023

1 ANOVA

[22]:

```
import pandas as pd
import scipy.stats as stats
import statsmodels.api as sm
from statsmodels.formula.api import ols
import scikit_posthocs as sp
```

[4]:

```
data_anova = pd.read_excel('/Users/yashgupta/Downloads/BSAN Project Files␣
  ↪Edited/Final Datasets/Alcohol and Drug Abuse Subset.xlsx')
```

[5]:

```
data_anova.isna().sum()
```

[5]:

```
HOSP_REGION_DES    0
RACE_DES           0
TOTCHG             0
dtype: int64
```

[6]:

```
# Check for non-numeric values in the 'TOTCHG' column
non_numeric = data_anova['TOTCHG'].apply(lambda x: isinstance(x, str))
missing_data = data_anova['TOTCHG'].isnull()

# Remove rows with non-numeric 'TOTCHG' values
clean_data_anova = data_anova[~(non_numeric | missing_data)].copy()

# Convert 'TOTCHG' to numeric
clean_data_anova['TOTCHG'] = pd.to_numeric(clean_data_anova['TOTCHG'])

# Perform the Two-Way ANOVA
model_anova_clean = ols('TOTCHG ~ C(HOSP_REGION_DES) + C(RACE_DES) +␣
  ↪C(HOSP_REGION_DES):C(RACE_DES)', data=clean_data_anova).fit()
anova_results_clean = sm.stats.anova_lm(model_anova_clean, typ=2)

# Output the results
print(anova_results_clean)
```

```
                                      sum_sq      df          F        PR(>F)
C(HOSP_REGION_DES)              8.250780e+11     3.0  78.920493  4.357724e-50
```

```
C(RACE_DES)                     2.184381e+11     5.0  12.536444  3.914456e-12
C(HOSP_REGION_DES):C(RACE_DES)  1.026040e+11    15.0   1.962858  1.426386e-02
Residual                        2.134819e+13  6126.0        NaN           NaN
```

2 Kruskal-Wallis

[14]:

```
group_sizes = clean_data_anova.groupby(['HOSP_REGION_DES', 'RACE_DES']).size()
print(group_sizes)
```

```
HOSP_REGION_DES  RACE_DES
Midwest          Asian or Pacific Islander      23
                 Black                         232
                 Hispanic                      105
                 Native American                57
                 Other                          59
                 White                        1154
Northeast        Asian or Pacific Islander      27
                 Black                         167
                 Hispanic                      193
                 Native American                 2
                 Other                         116
                 White                         493
South            Asian or Pacific Islander      17
                 Black                         458
                 Hispanic                      332
                 Native American                16
                 Other                          93
                 White                        1366
West             Asian or Pacific Islander      73
                 Black                          82
                 Hispanic                      397
                 Native American                59
                 Other                          76
                 White                         553
dtype: int64
```

[16]:

```
# Filter groups with at least 3 observations
filtered_data = clean_data_anova.groupby(['HOSP_REGION_DES', 'RACE_DES']).
 ↪filter(lambda x: len(x) >= 3)
```

[18]:

```
# Group by each factor and check if there are at least two groups
groups_by_region = filtered_data.groupby('HOSP_REGION_DES')
groups_by_race = filtered_data.groupby('RACE_DES')

# Kruskal-Wallis Test for 'HOSP_REGION_DES' if there are at least two groups
if len(groups_by_region) >= 2:
    kruskal_results_region = stats.kruskal(*[group['TOTCHG'] for name, group in␣
 ↪groups_by_region])
    print('Kruskal-Wallis Test for HOSP_REGION_DES:', kruskal_results_region)
else:
    print('Not enough groups for Kruskal-Wallis Test on HOSP_REGION_DES')

# Kruskal-Wallis Test for 'RACE_DES' if there are at least two groups
if len(groups_by_race) >= 2:
    kruskal_results_race = stats.kruskal(*[group['TOTCHG'] for name, group in␣
 ↪groups_by_race])
    print('Kruskal-Wallis Test for RACE_DES:', kruskal_results_race)
else:
    print('Not enough groups for Kruskal-Wallis Test on RACE_DES')
```

```
Kruskal-Wallis Test for HOSP_REGION_DES:
KruskalResult(statistic=413.96141641337977, pvalue=2.0935623622457854e-89)
Kruskal-Wallis Test for RACE_DES: KruskalResult(statistic=204.29404777163893,
pvalue=3.4251160817379375e-42)
```

[23]:

```
# Dunn's Test for 'HOSP_REGION_DES'
dunn_test_region = sp.posthoc_dunn(filtered_data, val_col='TOTCHG',␣
 ↪group_col='HOSP_REGION_DES', p_adjust='bonferroni')
print('Dunn\'s Test for HOSP_REGION_DES:\n', dunn_test_region)

# Dunn's Test for 'RACE_DES'
dunn_test_race = sp.posthoc_dunn(filtered_data, val_col='TOTCHG',␣
 ↪group_col='RACE_DES', p_adjust='bonferroni')
print('Dunn\'s Test for RACE_DES:\n', dunn_test_race)
```

```
Dunn's Test for HOSP_REGION_DES:
                  Midwest     Northeast         South          West
Midwest    1.000000e+00  2.442636e-50  2.389712e-14  3.770388e-71
Northeast  2.442636e-50  1.000000e+00  1.805234e-19  5.715719e-01
South      2.389712e-14  1.805234e-19  1.000000e+00  4.644086e-32
West       3.770388e-71  5.715719e-01  4.644086e-32  1.000000e+00
Dunn's Test for RACE_DES:
                             Asian or Pacific Islander         Black  \
Asian or Pacific Islander                 1.000000e+00  1.000000e+00
Black                                     1.000000e+00  1.000000e+00
Hispanic                                  1.000000e+00  1.967826e-02
Native American                           7.042033e-08  4.744021e-09
Other                                     1.000000e+00  4.865932e-02
White                                     3.572575e-04  1.258790e-09

                               Hispanic  Native American         Other  \
Asian or Pacific Islander  1.000000e+00     7.042033e-08  1.000000e+00
Black                      1.967826e-02     4.744021e-09  4.865932e-02
Hispanic                   1.000000e+00     4.417334e-14  1.000000e+00
```

```
Native American           4.417334e-14     1.000000e+00  8.019632e-13
Other                     1.000000e+00     8.019632e-13  1.000000e+00
White                     4.042140e-26     1.384044e-03  9.237443e-13

                                 White
Asian or Pacific Islander  3.572575e-04
Black                      1.258790e-09
Hispanic                   4.042140e-26
Native American            1.384044e-03
Other                      9.237443e-13
White                      1.000000e+00
```

These results from the Kruskal-Wallis tests for 'HOSP_REGION_DES' and 'RACE_DES' show significant differences in the 'TOTCHG' variable across the different levels of these categorical variables. Let's interpret these results:

Kruskal-Wallis Test for 'HOSP_REGION_DES':

Statistic: 413.96413.96 P-value: 2.09 ×10^{-89}2.09×10 −89

Interpretation: This extremely low p-value suggests that there are significant differences in the 'TOTCHG' values across the different hospital regions. The high test statistic value further indicates strong evidence against the null hypothesis of identical distributions of 'TOTCHG' across different hospital regions.

Kruskal-Wallis Test for 'RACE_DES':

Statistic: 204.29204.29 P-value: 3.43 ×10^{-42}3.43×10 −42

Interpretation: Similarly, this result indicates significant differences in the 'TOTCHG' values across different racial groups. The low p-value rejects the null hypothesis, suggesting that at least one racial group has a different distribution of 'TOTCHG' compared to others.

Interpretation of Dunn's Test Results:

For 'HOSP_REGION_DES':

The p-values are shown for each pair of regions. A p-value less than 0.05 typically indicates a statistically significant difference. For example, the p-value between Midwest and Northeast is approximately 2.44 ×10^{-50}2.44×10 −50 , indicating a statistically significant difference in 'TOTCHG' between these two regions. Similarly, significant differences are observed between several other pairs of regions, as indicated by the very low p-values (e.g., Midwest and West, South and West).

For 'RACE_DES':

The table shows p-values for pairwise comparisons between different racial groups. Many pairs show significant differences. For instance, the p-value between Hispanic and Native American is about 4.42 ×10^{-14}4.42×10 −14 , suggesting a significant difference in 'TOTCHG' between these two racial groups. However, some comparisons do not show significant differences (p-values close to 1), such as between Asian or Pacific Islander and Black.

[ ]:

Neonatal ANOVA & Kruskal-Wallis

November 20, 2023

```
[1]: import pandas as pd
     import scipy.stats as stats
     import statsmodels.api as sm
     from statsmodels.formula.api import ols
     import scikit_posthocs as sp
```

```
[2]: data_anova = pd.read_excel('/Users/yashgupta/Downloads/BSAN Project Files␣
       ↪Edited/Final Datasets/Neonatal ANOVA.xlsx')
```

```
[3]: data_anova.isna().sum()
```

```
[3]: Unnamed: 0                     0
     HOSP_KID                       0
     RECNUM                         0
     HOSP_LOCTEACH                  0
     H_CONTRL                       0
     HOSP_REGION                    0
     HOSP_BEDSIZE                   0
     AMONTH                         0
     AWEEKEND                       0
     DQTR                           0
     DQTR_DES                       0
     ELECTIVE                       0
     ELECTIVE_DES                   0
     APRDRG                         0
     APRDRG_FULL                    0
     APRDRG_Risk_Mortality          0
     APRDRG_Risk_Mortality_FULL     0
     APRDRG_Severity                0
     APRDRG_Severity_FULL           0
     AGE                            0
     AGE_NEONATE                    0
     DISPUNIFORM                    0
     DISPUNIFORM_DES                0
     DIED                           0
     FEMALE                         0
     FEMALE_DES                     0
     HOSP_REGION_DES                0
```

```
PAY1                          0
PAY1_DES                      0
RACE                          0
RACE_DES                      0
TOTCHG                        0
ZIPINC_QRTL                   0
Median Income                 0
LOS                           0
PCLASS_ORPROC                 0
PL_NCHS                       0
PL_NCHS_DES                   0
HOSP_LOCTEACH_FULL            0
H_CONTRL_FULL                 0
dtype: int64
```

1 ANOVA

[4]:

```
# Check for non-numeric values in the 'TOTCHG' column
non_numeric = data_anova['TOTCHG'].apply(lambda x: isinstance(x, str))
missing_data = data_anova['TOTCHG'].isnull()

# Remove rows with non-numeric 'TOTCHG' values
clean_data_anova = data_anova[~(non_numeric | missing_data)].copy()

# Convert 'TOTCHG' to numeric
clean_data_anova['TOTCHG'] = pd.to_numeric(clean_data_anova['TOTCHG'])

# Perform the Two-Way ANOVA
model_anova_clean = ols('TOTCHG ~ C(HOSP_REGION_DES) + C(RACE_DES) +␣
 ↪C(HOSP_REGION_DES):C(RACE_DES)', data=clean_data_anova).fit()
anova_results_clean = sm.stats.anova_lm(model_anova_clean, typ=2)

# Output the results
print(anova_results_clean)
```

```
                                    sum_sq      df         F    PR(>F)
C(HOSP_REGION_DES)            1.355665e+13     3.0  8.594334  0.000011
C(RACE_DES)                   8.833760e+12     6.0  2.800113  0.010109
C(HOSP_REGION_DES):C(RACE_DES)  1.549081e+13    18.0  1.636752  0.043352
Residual                      3.481834e+15  6622.0       NaN       NaN
```

2 Kruskal Wallis

[5]:

```
group_sizes = clean_data_anova.groupby(['HOSP_REGION_DES', 'RACE_DES']).size()
print(group_sizes)
```

```
HOSP_REGION_DES  RACE_DES
```

```
Midwest           Asian or Pacific Islander      40
                  Black                         305
                  Hispanic                       65
                  Missing                       312
                  Native American                 5
                  Other                          62
                  White                         638
Northeast         Asian or Pacific Islander      34
                  Black                         191
                  Hispanic                      129
                  Missing                       121
                  Native American                 2
                  Other                         157
                  White                         229
South             Asian or Pacific Islander      66
                  Black                        1042
                  Hispanic                      515
                  Missing                       253
                  Native American                13
                  Other                         275
                  White                        1117
West              Asian or Pacific Islander      74
                  Black                          76
                  Hispanic                      368
                  Missing                       127
                  Native American                 9
                  Other                          96
                  White                         329
dtype: int64
```

[6]:

```
# Filter groups with at least 3 observations
filtered_data = clean_data_anova.groupby(['HOSP_REGION_DES', 'RACE_DES']).
 ↪filter(lambda x: len(x) >= 3)
```

[7]:

```
# Group by each factor and check if there are at least two groups
groups_by_region = filtered_data.groupby('HOSP_REGION_DES')
groups_by_race = filtered_data.groupby('RACE_DES')

# Kruskal-Wallis Test for 'HOSP_REGION_DES' if there are at least two groups
if len(groups_by_region) >= 2:
    kruskal_results_region = stats.kruskal(*[group['TOTCHG'] for name, group in␣
 ↪groups_by_region])
    print('Kruskal-Wallis Test for HOSP_REGION_DES:', kruskal_results_region)
else:
    print('Not enough groups for Kruskal-Wallis Test on HOSP_REGION_DES')

# Kruskal-Wallis Test for 'RACE_DES' if there are at least two groups
```

```
if len(groups_by_race) >= 2:
    kruskal_results_race = stats.kruskal(*[group['TOTCHG'] for name, group in␣
 ↪groups_by_race])
    print('Kruskal-Wallis Test for RACE_DES:', kruskal_results_race)
else:
    print('Not enough groups for Kruskal-Wallis Test on RACE_DES')
```

```
Kruskal-Wallis Test for HOSP_REGION_DES:
KruskalResult(statistic=35.538455797703996, pvalue=9.3746200866394e-08)
Kruskal-Wallis Test for RACE_DES: KruskalResult(statistic=64.05097955034579,
pvalue=6.738652359806292e-12)
```

[8]:
```
# Dunn's Test for 'HOSP_REGION_DES'
dunn_test_region = sp.posthoc_dunn(filtered_data, val_col='TOTCHG',␣
 ↪group_col='HOSP_REGION_DES', p_adjust='bonferroni')
print('Dunn\'s Test for HOSP_REGION_DES:\n', dunn_test_region)

# Dunn's Test for 'RACE_DES'
dunn_test_race = sp.posthoc_dunn(filtered_data, val_col='TOTCHG',␣
 ↪group_col='RACE_DES', p_adjust='bonferroni')
print('Dunn\'s Test for RACE_DES:\n', dunn_test_race)
```

```
Dunn's Test for HOSP_REGION_DES:
                  Midwest  Northeast         South          West
Midwest    1.000000e+00   0.072759  1.000000e+00  8.501805e-07
Northeast  7.275934e-02   1.000000  2.144338e-01  1.365091e-01
South      1.000000e+00   0.214434  1.000000e+00  8.787442e-07
West       8.501805e-07   0.136509  8.787442e-07  1.000000e+00
Dunn's Test for RACE_DES:
                            Asian or Pacific Islander     Black      Hispanic  \
Asian or Pacific Islander                   1.000000  1.000000  1.000000e+00
Black                                       1.000000  1.000000  3.054186e-03
Hispanic                                    1.000000  0.003054  1.000000e+00
Missing                                     0.245207  0.180464  3.377957e-07
Native American                             0.284367  0.610279  6.873750e-02
Other                                       1.000000  0.000462  1.000000e+00
White                                       1.000000  0.007327  1.000000e+00

                                Missing  Native American         Other  \
Asian or Pacific Islander  2.452072e-01         0.284367  1.000000e+00
Black                      1.804643e-01         0.610279  4.617213e-04
Hispanic                   3.377957e-07         0.068737  1.000000e+00
Missing                    1.000000e+00         1.000000  9.452361e-08
Native American            1.000000e+00         1.000000  2.992629e-02
Other                      9.452361e-08         0.029926  1.000000e+00
White                      4.092657e-07         0.111623  1.000000e+00

                                  White
```

```
Asian or Pacific Islander  1.000000e+00
Black                      7.327373e-03
Hispanic                   1.000000e+00
Missing                    4.092657e-07
Native American            1.116230e-01
Other                      1.000000e+00
White                      1.000000e+00
```

<font : color = 'red'> Interpretation for 'HOSP_REGION_DES': Midwest vs. Other Regions:

Midwest vs. Northeast: No significant difference (p 0.0728). Midwest vs. South: No significant difference (p = 1.0). Midwest vs. West: Significant difference (p 8.50 × 10^{-7}).

Northeast vs. Other Regions:

Northeast vs. South: No significant difference (p 0.2144). Northeast vs. West: No significant difference (p 0.1365). South vs. West: Significant difference (p 8.79 × 10^{-7}).

<font : color = 'blue'>Interpretation for 'RACE_DES':

Asian or Pacific Islander vs. Other Races: No significant differences observed against any race.

Black vs. Other Races:

Black vs. Hispanic: Significant difference (p 0.0031). Black vs. Other: Significant difference (p 0.0005). Hispanic vs. Other Races:

Hispanic vs. Missing: Significant difference (p 3.38 × 10^{-7}). Missing vs. Other Races:

Missing vs. Other: Significant difference (p 9.45 × 10^{-8}). Missing vs. White: Significant difference (p 4.09 × 10^{-7}). Native American vs. Other: Significant difference observed against 'Other' (p 0.0299).

White vs. Other Races: No significant differences observed against any race except for a significant difference with 'Black' (p 0.0073).

Summary: In terms of hospital regions, significant differences in 'TOTCHG' are observed between the Midwest and West and between the South and West. Regarding race, several significant differences are observed, notably between Black and Hispanic, Black and Other, and between the Missing category and Hispanic and White. Where p-values are high (close to 1), it suggests no significant difference between those groups.

```
[ ]:
```

Hypothesis Testing - Age Neonatal

November 20, 2023

[1]:
```
import numpy as np
from scipy.stats import chi2_contingency
import scipy.stats as stats
```

1 Neonatal Deaths: Chi Squared Tests (Income, Region, Race)

1.1 Race

[2]:
```
# Create a contingency table
observed_data = np.array([[297,118582],
                          [2100,473638],
                          [1470,581331],
                          [48,23984],
                          [793,176429],
                          [3070,1315460]])
```

[3]:
```
# Perform the chi-squared test
chi2, p, dof, expected = chi2_contingency(observed_data)
```

[4]:
```
chi2
```

[4]:
```
724.1498538735071
```

[5]:
```
p
```

[5]:
```
2.9458906253470425e-154
```

[8]:
```
# Check the p-value to determine statistical significance
alpha = 0.05  # Set your chosen significance level
if p < alpha:
    print("Reject the null hypothesis: The number of neonatal death cases is␣
 ↪dependent on race.")
else:
    print("Fail to reject the null hypothesis: The number of neonatal death␣
 ↪cases is independent of race.")
```

```
Reject the null hypothesis: The number of neonatal death cases is dependent on
race.
```

1.2 Region

[9]:

```
# Create a contingency table
observed_data = np.array([[1903,662183],
                          [4118,1171321],
                          [910,388745],
                          [2051,674419]])
```

[10]:

```
# Perform the chi-squared test
chi2, p, dof, expected = chi2_contingency(observed_data)
```

[11]:

```
chi2
```

[11]:

```
148.76816473980756
```

[12]:

```
p
```

[12]:

```
4.858351583892842e-32
```

[13]:

```
# Check the p-value to determine statistical significance
alpha = 0.05  # Set your chosen significance level
if p < alpha:
    print("Reject the null hypothesis: The number of neonatal death cases is␣
 ↪dependent on region.")
else:
    print("Fail to reject the null hypothesis: The number of neonatal death␣
 ↪cases is independent of region.")
```

```
Reject the null hypothesis: The number of alcohol abuse cases is dependent on
region.
```

1.3 Income

[14]:

```
# Create a contingency table
observed_data = np.array([[1343,557524],
                          [2200,776811],
                          [2264,677935],
                          [3175,884398]])
```

[15]:

```
# Perform the chi-squared test
chi2, p, dof, expected = chi2_contingency(observed_data)
```

[16]:

```
chi2
```

[16]:

```
184.35151289054085
```

[17]:

```
p
```

[17]:

```
1.0131597907224045e-39
```

[19]:

```python
# Check the p-value to determine statistical significance
alpha = 0.05  # Set your chosen significance level
if p < alpha:
    print("Reject the null hypothesis: The number of neonatal death cases is dependent on income.")
else:
    print("Fail to reject the null hypothesis: The number of neonatal death cases is independent of income.")
```

```
Reject the null hypothesis: The number of neonatal death cases is dependent on
income.
```