Using Adjusted Standardized Residuals for Interpreting Contingency Tables

Eric Naioti and Erika Mudrak

1 Introduction: chi-squared test for independence

A Pearson's chi-squared test for independence is used to test for an association between two variables in a two-way contingency table. When a significant association is detected in a large table, interpretations can be difficult. In this document, we will present a method for performing post-hoc tests on a contingency table that is easy to interpret.

Consider the following data reported by Snee on hair and eye color (Snee 1974):

Table 1.1: Contingency table for hair and eye color from Snee 1974.

	Brown	Blue	Hazel	Green	Total
Black hair	68	20	15	5	108
Brown hair	119	84	54	29	286
Red hair	26	17	14	14	71
Blond hair	7	94	10	16	127
Total	220	215	93	64	592

The null hypothesis of the chi-squared test is that there is no association between the two categorical variables:

$$H_0:\text{Hair color and eye color are independent}$$

The test also requires three assumptions:

1. Both variables must be categorical and mutually exclusive, with each subject falling into exactly one cell in the table.
2. The observations must be independent and come from a random sample.
3. The value of the expected cell counts should be 5 or more in at least 80% of the cells, and no cell should have an expected count of less than one (Bewick 2004).

The chi-squared statistic (χ^2) that is used for testing this association is calculated using the following formula where O_{ij}, E_{ij} are, respectively, the observed and expected cell counts under the null hypothesis:

$$\chi^2 = \sum_i \sum_J r_{ij}^2,$$

where $r_{ij} = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}}}$.

Given an I by J contingency table, the test statistic (χ^2) follows a chi-squared distribution with degrees of freedom equal to $(I-1)(J-1)$. In our example: $\chi^2 = 138.29$, $df = 9$, and $p < 0.001$. (For a more detailed description of how to calculate χ^2 by hand, see the Appendix.)

2   Adjusted Pearson residuals

The significant p-value indicates that hair and eye color are associated. One might then be interested in finding what values in this table are driving this significance. We may want to know e.g. if the number of people with blue eyes and blond hair are greater than expected. To answer these questions one might decide to analyze all the 2 by 2 sub-tables, which can become quite cumbersome and will not give a very good overall picture. Instead, we will answer this using residuals.

The quantities $O_{ij} - E_{ij}$ are the residuals. The cells with the largest residuals might contribute the most to the chi-square statistic. However cells with larger cell counts will also have larger residuals. To make the residuals more comparable, they are standardized by dividing by $\sqrt{E_{ij}}$. The r_{ij} above represent the standardized residuals, also called the **Pearson residuals**. Crude "rules of thumb" have been used to assess the relative contribution of each cell to the chi-squared statistic. For example a Pearson residual whose absolute value is greater than 2 or 3 has a significant deviation from expectancy.

To obtain a more appropriate way to compare cells, the Pearson residuals can be further divided by the standard deviation of *all the residuals*. This is called the **adjusted Pearson residuals** and can be calculated as follows:

$$\tilde{r}_{ij} = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}(1 - m_i/N)(1 - n_j/N)}}$$

with m_i being the row total, n_j the column total and N the total number of observations. We can calculate $\tilde{r}_{ij}$ for all 16 cells in our example:

Table 2.1: Adjusted Pearson residuals

	Brown	Blue	Hazel	Green
Black hair	6.14	-4.25	-0.58	-2.29
Brown hair	2.16	-3.40	2.05	-0.51
Red hair	-0.10	-2.31	0.99	2.58
Blond hair	-8.33	9.97	-2.74	0.73

The big advantage of the adjusted Pearson residual is that it follows a *standard normal distribution* $N(0,1)$. This enables us to run post-hoc hypotheses tests on any number of the standardized residuals we are interested in. Suppose that in this example we are really only interested in determining whether or not the number of individuals with blue eyes and blond hair is significantly different than expected at the $\alpha = 0.05$ level.

Since the adjusted Pearson residuals are normally distributed, those cells with absolute values greater than the critical value $N(0,1)_{1-\alpha/2} = 1.96$ will have raw p-values of less than 0.05 (for a

two-sided test). The adjusted Pearson residual for those with blue eyes and blond hair is 9.97, and thus significant at the $\alpha = 0.05$ level.

If we were interested in testing the adjusted Pearson residuals of multiple cells or even every cell in the table, we would need to adjust the cutoffs for multiple comparisons.

One method would be to use a Bonferroni correction, we would adjust the alpha level by dividing it by the number of tests and comparing the absolute value of the adjusted residual to a new critical value.

To test all 16 of the cells in our table, the new alpha level is $\alpha_{\text{Bon}} = \frac{0.05}{16} = 0.003125$ with corresponding critical value $N(0,1)_{1-\alpha_{\text{Bon}}/2} = 2.95$.

With this Bonferroni corrected critical value, five of the cells have Pearson residuals significantly different than expected under the null hypothesis of no assocation: Brown eyes with black hair, brown eyes with blond hair, blue eyes with black hair, blue eyes with brown hair, and blue eyes with blond hair.

Code to calculate adjusted Pearson residuals is given in the Appendix.

Assuming we were indeed interested in all combinations in this table we could write the conclusion to this analysis as follows:

> In our sample of 592 individuals a chi-squared test showed that there is a significant association between hair color and eye color ($\chi^2 = 138.29$, df=9, $p < 0.001$). When examining all possible combinations of eye color and hair color, it was seen that those with blue eyes and blond hair are observed significantly more than expected while individuals with brown eyes and blond hair are observed significantly less than expected.

3   References

Snee, R. D. (1974) Graphical Display of Two-Way Contingency Tables. The American Statistician Vol. 28, No. 1 (Feb., 1974), pp. 9-12.

Agresti, A. (2007), An Introduction to Categorical Data Analysis, 2nd Edition, New York: John Wiley & Sons.

Bewick V, Cheek L, Ball J. (2004). Statistics review 8: Qualitative data—tests of association. Crit Care. 2004;8:46–53.

4   Appendix

4.1   Calculation of Pearson and adjusted Pearson residuals

The chi-squared statistic is calculated as the sum of the squared Pearson residuals:

$$\chi^2 = \sum_i \sum_J r_{ij}^2,$$

where $r_{ij} = \frac{O_{ij}-E_{ij}}{\sqrt{E_{ij}}}$.

In this formula O_{ij} is the number of observations in ith row and jth column of the table. E_{ij} is the expected number of observations in the ith row and jth column of the table. E_{ij} is calculated by multiplying the total from row i (O_{i+}) and the total from column j (O_{+j}) and dividing the product by the total number of observations (O_{++}).

Let's calculate the expected observations for those with blue eyes and blond hair ($E_{4,2}$). We have $O_{4,2} = 94$,

$$E_{4,2} = \frac{O_{4+} \times O_{+2}}{O_{++}} = \frac{127 \times 215}{592} \approx 46.12$$

The Pearson residuals for those with blue eyes and blond hair ($r_{4,2}$) would then be:

$$r_{4,2} = \frac{O_{4,2} - E_{4,2}}{\sqrt{E_{4,2}}} = \frac{94 - 46.12}{\sqrt{46.12}} \approx 7.05$$

We could do this a total of 16 times to find all of the Pearson residuals:

Table 4.1: Pearson residuals

	Brown	Blue	Hazel	Green
Black hair	4.40	-3.07	-0.48	-1.95
Brown hair	1.23	-1.95	1.35	-0.35
Red hair	-0.07	-1.73	0.85	2.28
Blond hair	-5.85	7.05	-2.23	0.61

The Chi-squared statistic is then calculated from the sum of all those residual values squared:

$$\chi^2 = \sum_j \sum_i r_{ij}^2 = 138.29$$

Recall that the adjusted Pearson residuals are calculated for a two-way table using the following formula (Agresti 2007):

$$\tilde{r}_{ij} = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}(1 - m_i/N)(1 - n_j/N)}}$$

where m_i is the total in the ith row, n_j is the total in the jth column and N the total number of observations.

The adjusted residual for those with blue eyes and blond hair, $\tilde{r}_{4,2}$ is therefore

$$\tilde{r}_{4,2} = \frac{O_{4,2} - E_{4,2}}{E_{4,2}(1 - m_4/N)(1 - n_2/N)} = \frac{94 - 46.12}{\sqrt{46.12 \times (1 - 127/592)(1 - 215/592)}} \approx 9.97$$

Performing the same calculation for each cell leads to the results in Table 2.1.

4.2   Code

The sample data set "HairEyeColor.csv" can be found here. For each program we will start with a file containing the long form of the data, with 592 observations and categorical variables "Hair" and "Eye" that each fall into 4 categories. Often we will need to aggregate this data into some other form to perform the analysis.

4.2.1 SAS

```
PROC IMPORT OUT= WORK.HairEye
            DATAFILE= "~\HairEyeColor.csv"
            DBMS=CSV REPLACE;
     GETNAMES=YES;
     DATAROW=2;
RUN; * Read in the Data;

proc freq data=HairEye;
table Hair*Eye/out=Counts; run; * Reorganize into Count data;

proc genmod data=Counts;
class Hair Eye ;
model Count =  Hair Eye /DIST=POISSON LINK=LOG;
OUTPUT OUT=resid predicted=PRED reschi=PEARS stdreschi=STDRED;
run; * Perform Test and Output Residuals;

proc print data=resid; run;
run;
```

4.2.2 SPSS

SPSS provides standardized residuals via the crosstab function. The following is the code that is run after reading in the CSV file into DataSet1.

```
DATASET ACTIVATE DataSet1.
CROSSTABS
  /TABLES=Hair BY Eye
  /FORMAT=AVALUE TABLES
  /STATISTICS=CHISQ
  /CELLS= ASRESID
  /COUNT ROUND CELL.
```

4.2.3 R

We can find the standardized Pearson residuals by first making a table from the data, then running a Chi-squared test and extracting the standardized residuals from the results of that test.

```
long.data <- read.csv("HairEyeColor.csv")
freq.table <- table(long.data$Hair, long.data$Eye)
Chi2 <- chisq.test(freq.table)
Chi2$stdres
```

4.2.4 Stata

To do this analysis in Stata, we will first install the program tab_chi, then run the Chi-squared test and request the adjusted standardized residuals.

```
ssc install tab_chi
tabchi hair eye, adj
```

Created December 2018. Last updated April 2022.