

### Three-Way Contingency Tables

Three-way contingency tables involve three binary or categorical variables. I will stick mostly to the binary case to keep things simple, but we can have three-way tables with any number of categories with each variable. Of course, higher dimensions are also possible, but they are uncommon in practice and there are few commonly available statistical tests for them.

#### Notation

We described two-way contingency tables earlier as having  $I$  rows and  $J$  columns, forming an  $I \times J$  matrix. If a third variable is included, the three-way contingency table is described as  $I \times J \times K$ , where  $I$  is the number of rows,  $J$  is the number of columns, and  $K$  is the number of superordinate columns each level containing a  $2 \times 2$  matrix. The indexes  $i, j$ , and  $k$  are used to enumerate the rows or columns for each dimension. Each within-cell count or joint proportion has three subscripts now,  $n_{ijk}$  and  $p_{ijk}$ . There are two general types of marginal counts (or proportions), with three specific marginals for each type. The first type of marginal count combines one of the three dimensions. Using the + notation from two-way tables, the marginals combining counts across either  $I, J$ , or  $K$  dimensions are  $n_{+jk}$ ,  $n_{i+k}$ , or  $n_{ij+}$ , respectively. The second general type of marginal collapses two dimensions and is given as  $n_{i++}$ ,  $n_{+j+}$ , or  $n_{++k}$ . The total count (sample size) is  $n_{+++}$  (or sometimes just  $n$ ). The simplest case is a  $2 \times 2 \times 2$  shown below to illustrate the notation.

$k = 1$			$k = 2$		
$n_{111}$	$n_{121}$	$n_{1+1}$	$n_{112}$	$n_{122}$	$n_{1+2}$
$n_{211}$	$n_{221}$	$n_{2+1}$	$n_{212}$	$n_{222}$	$n_{2+2}$
$n_{+11}$	$n_{+21}$	$n_{++1}$	$n_{+12}$	$n_{+22}$	$n_{++2}$

Note: marginals collapsing across levels of  $K$  are not shown.

In the Azen and Walker text as well as elsewhere, the three variables are referred to by letter names  $X, Y$ , and  $Z$ .<sup>1</sup> Although we do not need to assign explanatory (independent) and response (dependent) roles to the variables for all hypotheses that might be tested, it can be easier to put the table into more familiar terms by imagining one response variable  $Y$  (with its  $J$  number of levels) and two explanatory variables,  $X$  and  $Z$  (with  $I$  and  $K$  number of levels, respectively). Thinking in this way, we now have the conceptual equivalent of a ANOVA design, where  $Y$  is the (binary) dependent variable and  $X$  and  $Z$  are two independent variables. One common hypothesis of interest is to examine the  $X$  effect on  $Y$  at different levels of  $Z$ , which is the same question asked in the test of the two-way factorial ANOVA (e.g., is the  $X$  group difference on  $Y$  the same for  $Z_1$  and  $Z_2$ ?). It is important to note that this is the real analogy to ANOVA for at least one possible or common hypothesis that can be tested with three-way contingency table. Azen and Walker use the three-way ANOVA as an analogy throughout the chapter for a different reason. Their intention is to use the three-way ANOVA analogy for the three-way contingency table to describe the process by which the researcher should proceed in analysis steps (e.g., examining the three-way table first, and then depending on the results pursuing two-dimensional or one-dimensional comparisons). So, be careful not to confuse their process analogy to ANOVA with the statistical test analogy to ANOVA.

#### Partial Tables

There are two  $2 \times 2$  tables above that make up the three-way table—an  $X \times Y$  table within  $Z_1$  and an  $X \times Y$  table within  $Z_2$ . These are known as partial tables. Although partial tables can be constructed for levels of the  $X$  or levels the  $Y$  variable also, it is most common to set up three-way tables in this manner and discuss the  $X \times Y$  partial tables within levels (or strata) of  $Z$ . Marginals for the two-way partial tables can be expressed as conditional proportions, similar to the simple conditional proportion in the two-way case. The notation  $p_{ijk}$  denotes one type of conditional proportion, which is the within-cell count relative to the

<sup>1</sup> Many authors use  $A, B$ , and  $C$ , but this is easily confused with the fourway table notation.

number of cases within one level of  $Z$ ,  $p_{ijk} = n_{ijk}/n_{++k}$ . Of course, we could also compute other conditional proportions, such as  $p_{ij|k}$  or  $p_{jk|i}$  and so on.

If the two  $I \times J$  contingency tables are considered separately, we could compute odds ratios as before, adding to the notation to indicate that the odds ratio is specific to one level of the  $Z$  variable. Following the notation in your text, we use  $X$ ,  $Y$ , and  $Z$  and the Greek letter  $\theta$  ("theta") for the odds ratio.

$$\theta_{XY|Z_1} = \frac{\text{odds}_{X=1|YZ_1}}{\text{odds}_{X=0|YZ_1}} = \frac{n_{211} / n_{221}}{n_{111} / n_{121}} = \frac{n_{111}n_{221}}{n_{211}n_{121}} = \frac{P_{X=1|YZ_1}}{P_{X=0|YZ_1}}$$

If we were to calculate these odds ratios for each level of  $Z$  and then average them taking into account the different number of cases in each stratum of  $Z$ , we get a common odds ratio called the *Mantel-Haenszel* (MH) odds ratio.

$$\theta_{MH} = \frac{\sum_{k=1}^K (n_{11k}n_{22k} / n_{++k})}{\sum_{k=1}^K (n_{12k}n_{21k} / n_{++k})}$$

The Mantel-Haenszel is the *common odds ratio* and assumes that the odds ratio does not differ across the levels of  $Z$ , so it may not be of interest itself in many circumstances. It can be used, however, if the test of whether the  $X$  and  $Y$  are associated overall by assessing the weighted average of this relationship over the levels of  $Z$  is of interest and the odds do not differ substantially across the levels.

### Breslow-Day Test

Perhaps the most useful test is a test of the hypothesis that the  $X$ - $Y$  association differs across levels of  $Z$ , known as the Breslow-Day test (Breslow & Day, 1980). Later modified by Tarone (1985), it is sometimes referred to as the Breslow-Day-Tarone test. If there is no difference across levels of  $Z$ , the  $X$ - $Y$  relationship is homogeneous. Another way to state this hypothesis test is as an evaluation of whether the partial table odds ratios are equal across the levels of  $Z$ , that  $\theta_{XY|Z_1} = \theta_{XY|Z_2}$ .

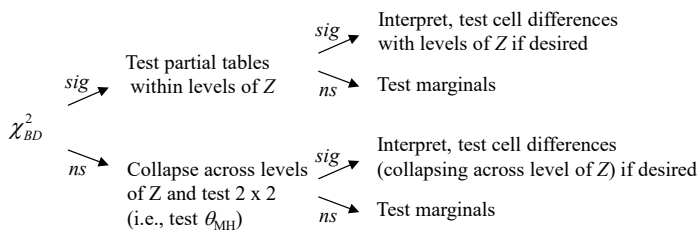
$$\chi_{BD}^2 = \sum_{k=1}^K \frac{[O_{Y_{1k}} - E(Y_{1k} | \theta_{MH})]^2}{V(Y_{1k} | \theta_{MH})}$$

In the numerator, we have an alteration of the usual chi-squared statistics in which of the expected values are subtracted from the observed values, except that the observed and the expected values are for partial tables for each level of  $Z$  (indexed by subscript  $k$  for each stratum of  $Z$ ), with the expected values taking into account the Mantel-Haenszel odds ratio,  $\theta_{MH}$ . The expected values,  $E(Y_{1k} | \theta_{MH})$ , and variance estimate,  $V(Y_{1k} | \theta_{MH})$ , in the denominator are considerably more complicated than for the two-way contingency table.<sup>2</sup> The Breslow-Day test is only for  $2 \times 2 \times k$  tables, so  $X$  and  $Y$  must be binary. The  $\chi_{BD}^2$  can be evaluated against the chi-squared distribution with  $df = K - 1$  if the Tarone modification is used. Significance implies  $\theta_{XY|Z_1} \neq \theta_{XY|Z_2}$  but also that  $\theta_{XZ|Y_1} \neq \theta_{XZ|Y_2}$  and  $\theta_{YZ|X_1} \neq \theta_{YZ|X_2}$  for the  $2 \times 2 \times 2$  case. If there are many strata for  $Z$  (i.e.,  $K$  is large) and data are sparse, then an exponential weighted average of the common odds ratio (Woolf, 1955) is more efficient.

<sup>2</sup> Since you are not afraid to look, here is the computations for these two quantities (Hosmer & Lemeshow, 2000):

$$E(Y_{1k} | \theta_{MH}) = \left[ \frac{1}{2(\theta_{MH} - 1)} \right] \left\{ \theta_{MH} [(n_{11k} + n_{12k}) + (n_{11k} + n_{21k})] + [(n_{21k} + n_{22k}) - (n_{11k} + n_{21k})] \right\} \pm \sqrt{ \left\{ \theta_{MH} [(n_{11k} + n_{12k}) + (n_{11k} + n_{21k})] + [(n_{21k} + n_{22k}) - (n_{11k} + n_{21k})] \right\}^2 - \{ 4(\theta_{MH} - 1)\theta_{MH}(n_{11k} + n_{12k})(n_{11k} + n_{21k}) \} } \\ \text{and } V[E(Y_{1k} | \theta_{MH})] = \left[ \frac{1}{E(Y_{1k} | \theta_{MH})} + \frac{1}{(n_{11k} + n_{12k}) - E(Y_{1k} | \theta_{MH})} + \frac{1}{(n_{11k} + n_{21k}) - E(Y_{1k} | \theta_{MH})} + \frac{1}{(n_{21k} + n_{22k}) - (n_{11k} + n_{21k}) + E(Y_{1k} | \theta_{MH})} \right]^{-1}$$

I suspect that most social scientist are interested in testing the hypothesis for which the Breslow-Day test is designed, that the odds ratio at each level of  $Z$  is equal. If this is the question of interest, here is an overview of the process you might use:



### Cochran-Mantel-Haenszel

A special case of the homogeneity of association is the Cochran-Mantel-Haenszel (CMH) test that the odds ratios for all partial tables are equal to 1, which is referred to as a test of *conditional independence*. For the  $2 \times 2 \times 2$  case, it is a test of the null hypothesis that  $\theta_{XY|Z_1} = \theta_{XY|Z_2} = 1$ , but also whether  $\theta_{XZ|Y_1} = \theta_{XZ|Y_2} = 1$  and  $\theta_{YZ|X_1} = \theta_{YZ|X_2} = 1$ . If the test is significant, it may be due to any one of the partial tables with an odds ratio above or below 1.0. The computation of the CMH follows the general Pearson chi-squared equation and can be stated as a likelihood ratio,  $G^2$ , as well. The computation of the expected values is considerably simpler than for the Breslow-Day test.

$$\chi^2_{CMH} = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K \frac{(O_{ijk} - E_{ijk})^2}{E_{ijk}} \text{ or } G^2_{CMH} = -2 \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K O_{ijk} \ln \left( \frac{E_{ijk}}{O_{ijk}} \right)$$

where  $E_{ijk} = \frac{n_{i++} n_{+j+} n_{++k}}{n_{+++}^2}$

Both CMH tests are evaluated against the chi-squared distribution with  $df = IJK - I - J - K + 2$ .

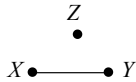
### Logistic Regression

Logistic regression can be used to test the same hypotheses as those investigated with the Mantel-Haenszel odds ratio and the Breslow-Day tests, and the logistic regression approach is probably the more commonly employed. As you may have guessed, the corresponding hypothesis in logistic regression is investigated by testing the  $X*Z$  interaction in predicting  $Y$ . Such tests are not exactly equal even though they address the same hypothesis (see Hosmer & Lemeshow, 2000, for a discussion). The association tests discussed above do not require assigning explanatory (independent) and response (dependent) roles to the variables per se. Logistic regression models assume that one of the three variables is a response (usually  $Y$  in our examples above) and the other two ( $X$  and  $Z$  in our examples above) are explanatory. The maximum likelihood estimation as well as the more general framework of the logistic regression modeling set the logistic approach apart somewhat from the Mantel-Haenszel odds ratio and Breslow-Day tests. We will return to how to test these hypotheses with logistic regression in a later section of the course.

### Other Three-Way Table Hypotheses

The Breslow-Day and the Cochran-Mantel-Haenszel are the most commonly employed hypothesis tests, but, with a three-way contingency table, there are many possibilities. With the properly derived expected frequencies, a variety of tests can investigate whether variables  $X$ ,  $Y$ , and  $Z$  are independent from one another. Wickens (1989) discusses association tests can be formulated with loglinear models. Using bracket notation, different tests can be distinguished by which variables are independent from one another. For example,  $[X][Y][XZ]$  indicates that each of the three variables are independent from one

another, whereas  $[XY][Z]$  indicates that  $Z$  is unrelated to  $X$  and  $Y$ . A nodal diagram can also be used to depict these relations, with the diagram below reflecting the independence implied by  $[XY][Z]$ .



Loglinear modeling is a general and very flexible framework which will be introduced in the next section of the course.

### Software Examples

To illustrate statistical tests of three-way contingency tables, I used data from the 2020 Quinnipiac presidential poll conducted in Ohio and Georgia.<sup>3</sup> The Breslow-Day test examines whether independents were more likely to favor Biden or Trump, comparing this association across the two states.

### SPSS

```
get file='c:\jason\spsswin\cdaclass\quinnipiac 2020 3way.sav'.

output comment text="OH".
temporary.
select if state eq 0.
crosstabs /tables=ind by response
         /statistics=none
         /cells=count row.

output comment text="GA".
temporary.
select if state eq 1.
crosstabs /tables=ind by response
         /statistics=none
         /cells=count row.

crosstabs /tables=ind by response by state
         /statistics=cmh
         /cells=none.
```

OH

**ind \* response Crosstabulation**

		response			
		0 Trump	1 Biden	Total	
ind	.00 party affiliate	Count	308	401	709
		% within ind	43.4%	56.6%	100.0%
	1.00 independent	Count	132	191	323
		% within ind	40.9%	59.1%	100.0%
Total		Count	440	592	1032
		% within ind	42.6%	57.4%	100.0%

GA

**ind \* response Crosstabulation**

		response			
		0 Trump	1 Biden	Total	
ind	.00 party affiliate	Count	323	374	697
		% within ind	46.3%	53.7%	100.0%
	1.00 independent	Count	103	128	231
		% within ind	44.6%	55.4%	100.0%
Total		Count	426	502	928
		% within ind	45.9%	54.1%	100.0%

<sup>3</sup> Data source: <https://poll.qu.edu/georgia/release-detail?ReleaseID=3679>. Note that the data extrapolated cell sample sizes and used some rounding, so the results should be taken as only approximate.

### Tests of Homogeneity of the Odds Ratio

	Chi-Squared	df	Asymptotic Significance (2-sided)
Breslow-Day	.029	1	.864
Tarone's	.029	1	.864

### Tests of Conditional Independence

	Chi-Squared	df	Asymptotic Significance (2-sided)
Cochran's	.787	1	.375
Mantel-Haenszel	.699	1	.403

Under the conditional independence assumption, Cochran's statistic is asymptotically distributed as a 1 df chi-squared distribution, only if the number of strata is fixed, while the Mantel-Haenszel statistic is always asymptotically distributed as a 1 df chi-squared distribution. Note that the continuity correction is removed from the Mantel-Haenszel statistic when the sum of the differences between the observed and the expected is 0.

### Mantel-Haenszel Common Odds Ratio Estimate

Estimate	1.094		
ln(Estimate)	.090		
Standard Error of ln(Estimate)	.102		
Asymptotic Significance (2-sided)	.375		
Asymptotic 95% Confidence Interval	Common Odds Ratio	Lower Bound	.897
		Upper Bound	1.335
	ln(Common Odds Ratio)	Lower Bound	-.109
		Upper Bound	.289

The Mantel-Haenszel common odds ratio estimate is asymptotically normally distributed under the common odds ratio of 1.000 assumption. So is the natural log of the estimate.

```
R
> library(haven)
> d = read_sav("c:/jason/spsswin/cdaclass/quinnpiac 2020 3way.sav")
>
> library(lessR)
> #if need to convert data types in order to compute a correlation in R--lessR function
> #d <- Transform(response = as.numeric(response))
>
> d1 <- Subset(state=0)
>
> #this lessR BarChart function produces a chi-square test by default
> BarChart (response, by=ind, horiz = FALSE, stat = "proportion", beside = TRUE, xlab="Response: Trump=0, Biden=1", data=d1)
```

### Joint and Marginal Frequencies

-----

```
response
ind    0    1  Sum
  0  631  775 1406
  1  235  319  554
  Sum 866 1094 1960
```

Cramer's V (phi): 0.022

Chi-square Test: Chisq = 0.975, df = 1, p-value = 0.323

Cell Proportions within Each Column

```
-----  
response  
ind      0      1  
0  0.729 0.708  
1  0.271 0.292  
Sum 1.000 1.000
```

```
> d2 <- Subset(state=1)
```

```
> #this lessR BarChart function produces a chi-square test by default  
> BarChart (response, by=ind, horiz = FALSE, stat = "proportion", beside = TRUE, xlab="Response: Trump=0,  
Biden=1", data=d1)
```

```
Joint and Marginal Frequencies  
-----
```

```
response  
ind      0      1  Sum  
0  631  775 1406  
1  235  319  554  
Sum 866 1094 1960
```

```
Cramer's V (phi): 0.022
```

```
Chi-square Test:  Chisq = 0.975, df = 1, p-value = 0.323
```

```
Cell Proportions within Each Column  
-----
```

```
response  
ind      0      1  
0  0.729 0.708  
1  0.271 0.292  
Sum 1.000 1.000
```

```
> counts <-array(  
+   c(308,401,132,191,323,374,103,128),  
+   dim=c(2, 2, 2),  
+   dimnames=list(state=c("OH", "GA"),  
+                     ind =c("party aff", "independent"),  
+                     response =c("Biden", "trump"))  
+ )
```

```
> library(DescTools)  
> BreslowDayTest(counts,correct=FALSE)
```

```
      Breslow-Day test on Homogeneity of Odds Ratios
```

```
data: counts  
X-squared = 0.02913, df = 1, p-value = 0.8645
```

```
> mantelhaen.test(counts)
```

```
      Mantel-Haenszel chi-squared test with continuity correction
```

```
data: counts  
Mantel-Haenszel X-squared = 0.69909, df = 1, p-value = 0.4031  
alternative hypothesis: true common odds ratio is not equal to 1  
95 percent confidence interval:  
 0.896715 1.335499  
sample estimates:  
common odds ratio  
 1.094332
```

**SAS**

```
proc import datafile="c:\jason\spsswin\cdaclass\quinnipiac 2020 3way.sav" out=one
dbms = sav replace;
run;
```

```
proc freq data=one;
tables state*ind*response /cmh ;
title 'Test of 3-Way Table';
```

run;

The FREQ Procedure

Table 1 of ind by response  
 Controlling for state=OH

ind(ind)	response(response)		
	Trump	Biden	Total
party affiliate	308	401	709
Percent	29.84	38.86	68.70
Row Pct	43.44	56.56	
Col Pct	70.00	67.74	
independent	132	191	323
Percent	12.79	18.51	31.30
Row Pct	40.87	59.13	
Col Pct	30.00	32.26	
Total	440	592	1032
Percent	42.64	57.36	100.00

Table 2 of ind by response  
 Controlling for state=GA

ind(ind)	response(response)		
	Trump	Biden	Total
party affiliate	323	374	697
Percent	34.81	40.30	75.11
Row Pct	46.34	53.66	
Col Pct	75.82	74.50	
independent	103	128	231
Percent	11.10	13.79	24.89
Row Pct	44.59	55.41	
Col Pct	24.18	25.50	
Total	426	502	928
Percent	45.91	54.09	100.00

Test of 3-Way Table  
 13:38 Wednesday, April 7, 2021 4

The FREQ Procedure

Summary Statistics for ind by response  
 Controlling for state

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	0.7864	0.3752
2	Row Mean Scores Differ	1	0.7864	0.3752
3	General Association	1	0.7864	0.3752

Common Odds Ratio and Relative Risks

Statistic	Method	Value	95% Confidence Limits	
Odds Ratio	Mantel-Haenszel	1.0943	0.8967	1.3355
	Logit	1.0943	0.8967	1.3355
Relative Risk (Column 1)	Mantel-Haenszel	1.0521	0.9395	1.1781
	Logit	1.0517	0.9393	1.1777
Relative Risk (Column 2)	Mantel-Haenszel	0.9615	0.8824	1.0477
	Logit	0.9613	0.8823	1.0474

Breslow-Day Test for  
 Homogeneity of Odds Ratios

Chi-Square	0.0291
DF	1
Pr > ChiSq	0.8645

Total Sample Size = 1960

### Sample Write-Up

A Breslow-Day test was used to investigate whether the association between identification as an independent (identified with a major party vs. independent) and candidate preference (Biden vs. Trump) was the same in Ohio and Georgia. Among Ohioans, 56.56% of respondents identifying with a major party favored Biden, whereas 59.13% of respondents identifying as independent favored Biden. Among Georgians, 53.66% of respondents identifying with a major party favored Biden, whereas 55.41% of respondents identifying with a major party favored Biden. The Breslow-Day test indicated that the association between party identification and the candidate preference did not differ significantly between Ohio and Georgia,  $\chi^2_{BD}(1) = .029$ ,  $p = .86$ . The common odds ratio indicated no overall difference in candidate preference by party identification,  $OR_{MH} = 1.09$ , 95% CI = .90, 1.34. Follow-up analyses [not shown above] indicated that there was no association between independent identification and the candidate favored within state, Pearson  $\chi^2(1) = .60$ ,  $p = .44$ , for Ohio, and Pearson  $\chi^2(1) = .21$ ,  $p = .64$ , for Georgia.<sup>4</sup>

<sup>4</sup> It might be desirable to go one step further and test the marginals to determine whether Biden or Trump was favored overall.



## References and Further Readings

- Breslow, N. E., & Day, N. E. (1980). *Statistical methods in cancer research. Vol. 1. The analysis of case-control studies* (Vol. 1, No. 32). Distributed for IARC by WHO, Geneva, Switzerland.
- Tarone, R. E. (1985). On heterogeneity tests based on efficient scores. *Biometrika*, 72(1), 91-95.
- Wickens, T. D. (1989). *Multway contingency tables analysis for the social sciences*. New York: Erlbaum.
- Woolf, B. (1955). On estimating the relation between blood group and disease. *Annals of Human Genetics*, 19(4), 251-253.