Chapter 10: Contingency tables I

Note made by: Timothy Hanson
Instructor: Peijie Hou

Department of Statistics, University of South Carolina

Stat 205: Elementary Statistics for the Biological and Life Sciences

Two-sample binary data

- In Chapter 9 we looked at one sample & looked at observed vs. "expected under H_0."
- Now we consider two populations and will want to compare two population proportions p_1 and p_2.
- In population 1, we observed y_1 out of n_1 successes; in population 2 we observed y_2 out of n_2 successes.
- This information can be placed in a contingency table

		Group	
		1	2
Outcome	Success	y_1	y_2
	Failure	$n_1 - y_1$	$n_2 - y_2$
	Total	n_1	n_2

- $\hat{p}_1 = y_1/n_1$ estimates p_1 & $\hat{p}_2 = y_2/n_2$ estimates p_2.

Example 10.1.1 Migraine headache

- Migraine headache patients took part in a double-blind clinical trial to assess experimental surgery.
- 75 patients were randomly assigned to real surgery on migraine trigger sites ($n_1 = 49$) or sham surgery ($n_2 = 26$) in which an incision was made but nothing else.
- The surgeons hoped that patients would experience "a substantial reduction in migraine headaches," which we will label as success.

Example 10.1.1 Migraine headache

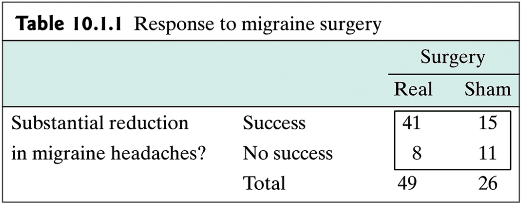

Table 10.1.1 Response to migraine surgery

		Surgery	
		Real	Sham
Substantial reduction	Success	41	15
in migraine headaches?	No success	8	11
	Total	49	26

- $\hat{p}_1 = 41/49 = 83.7\%$ for real surgeries.
- $\hat{p}_2 = 15/26 = 57.7\%$ for sham surgeries.
- Real appears to be better than sham, but is this difference significant?

Example 10.1.2 HIV testing

A random sample of 120 college students found that 9 of the 61 women in the sample had taken an HIV test, compared to 8 of the 59 men.

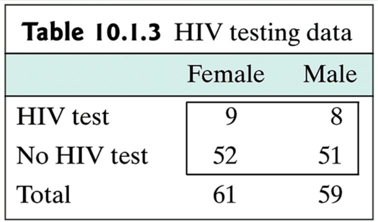

Table 10.1.3 HIV testing data

	Female	Male
HIV test	9	8
No HIV test	52	51
Total	61	59

- $\hat{p}_1 = 9/61 = 14.8\%$ tested among women.
- $\hat{p}_2 = 8/59 = 13.6\%$ tested among men.
- These are pretty close.

Conditional probabilities

- p_1 and p_2 are *conditional probabilities*. Remember way back in Section 3.3?
- For the migraine data, $p_1 = \text{pr}\{\text{success}|\text{real}\}$ and $p_2 = \text{pr}\{\text{success}|\text{sham}\}$. $\hat{p}_1 = 0.84$ and $\hat{p}_2 = 0.58$ estimate these conditional probabilities.
- For the HIV testing data, $p_1 = \text{pr}\{\text{tested}|\text{female}\}$ and $p_2 = \text{pr}\{\text{tested}|\text{male}\}$. $\hat{p}_1 = 0.15$ and $\hat{p}_2 = 0.14$ estimate these conditional probabilities.

χ^2 test for independence

- There is *no difference* between groups when $H_0 : p_1 = p_2$.
- That is, $H_0 : \Pr\{\text{success}|\text{group 1}\} = \Pr\{\text{success}|\text{group 2}\}$.
- If H_0 is true then the outcome (migraine reduction, being test for HIV, etc.) is *independent* of the group.
- This is tested using the chi-square statistic

  $$\chi^2_S = \sum_{i=1}^{4} \frac{(o_i - e_i)^2}{e_i},$$

  where $i = 1, 2, 3, 4$ are the four cells in the middle of the contingency table.
- The o_i are the observed counts and the e_i are what's expected if $p_1 = p_2$.

Computing e_i

- If $H_0 : p_1 = p_2$ is true then we can estimate the common probability $p = p_1 = p_2$ by $\hat{p} = (y_1 + y_2)/(n_1 + n_2)$. This is $\hat{p} = 56/75 = 0.747$ for migraine data.
- In the upper left corner we'd expect to see $\hat{p}n_1 = 0.747(49) = 36.59$ successes in the real surgery group, and so $49 - 36.59 = 12.41$ failures in the lower left.
- In the upper right corner we'd expect to see $\hat{p}n_2 = 0.747(26) = 19.41$ successes in the sham surgery group, and so $26 - 19.41 = 6.59$ failures in the lower right.

Observed and expected under H_0

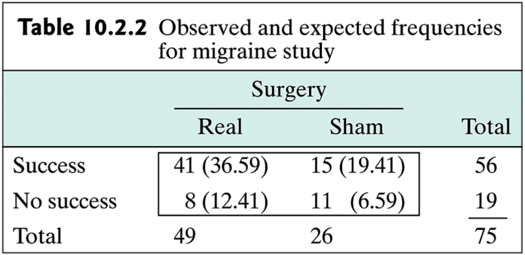

Table 10.2.2 Observed and expected frequencies for migraine study

	Surgery		
	Real	Sham	Total
Success	41 (36.59)	15 (19.41)	56
No success	8 (12.41)	11 (6.59)	19
Total	49	26	75

$$\chi_S^2 = \frac{(41-36.59)^2}{36.59} + \frac{(15-19.41)^2}{19.41} + \frac{(8-12.41)^2}{12.41} + \frac{(11-6.59)^2}{6.59} = 6.06.$$

The P-value

- When $H_0 : p_1 = p_2$ is true, χ^2_S has a χ^2_1 distribution, chi-square with 1 degree of freedom.
- The P-value is the tail probability of a chi-square density with 1 *df* greater than what we saw χ^2_S. The P-value is the probability of seeing $\hat{p}_1$ and $\hat{p}_2$ *even further away from each other* than what we saw.
- We can get the P-value out of R using `chisq.test`, but now we need to put in a contingency table in the form of a matrix to get our P-value.

Obtaining surgery data P-value in R

- Need to create a 2 × 2 matrix of values first

```
> surgery=matrix(c(41,8,15,11),nrow=2)
> colnames(surgery)=c("Real","Sham")
> rownames(surgery)=c("Success","No success")
> surgery
           Real Sham
Success      41   15
No success    8   11
```

- The default `chisq.test(surgery)` uses $\chi^2_Y = \sum_{i=1}^4 \frac{(|o_i - e_i| - 0.5)^2}{e_i}$. Called "Yates continuity correction" & gives more accurate P-values in small samples.

```
> chisq.test(surgery)

        Pearson's Chi-squared test with Yates' continuity correction

data:  surgery
X-squared = 4.7661, df = 1, p-value = 0.02902
```

Obtaining surgery data P-value in R

- To get the statistic and P-value in your book, we have to turn the Yates correction "off" using `chisq.test(surgery,correct=FALSE)`.

```
> chisq.test(surgery,correct=FALSE)

        Pearson's Chi-squared test

data:  surgery
X-squared = 6.0619, df = 1, p-value = 0.01381
```

- We reject $H_0 : p_1 = p_2$ at the 5% level. The surgery significantly reduces migraines.

10.3 Two ways to collect data

- There are two ways to collect 2×2 contingency table data.
- **Cross-sectional** data is collected by randomly sampling n individuals and cross-classifying them on two variables.
- Example Ask $n = 143$ random individuals two questions: salary high/low and education high-school/college.
- The row and column totals are random.
- **Product binomial** data is collected when a fixed number from one group is sampled, and a fixed number from another group is sampled.
- Example: Real vs. sham surgery for migraine.

10.4 Fishers exact test

- For the chi-square test to be valid, we cannot have very small sample sizes, say less than 5 in any cell.
- For small sample sizes there is an exact test, called Fisher's exact test for testing $H_0 : p_1 = p_2$.
- Fisher's test computes all possible 2 × 2 tables with the same number of successes and failures (56 successes and 19 failures for the migraine study) that make $\hat{p}_1$ and $\hat{p}_2$ even further apart than what we saw, and adds up the probability of seeing each table. Your book has details if you are interested on pp. 381–383.
- An alternative, that also works for small sample sizes, is the equivalent of the permutation test of Section 7.1, only for binary data, given by `chisq.test(surgery,simulate.p.value=TRUE)`.

Example 10.4.5 Flu shots

A random sample of college students found that 13 of them had gotten a flu shot at the beginning of the winter and 28 had not. Of the 13 who had a flu shot, 3 got the flu during the winter. Of the 28 who did not get a flu shot, 15 got the flu.

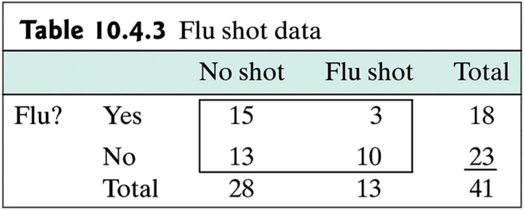

Table 10.4.3 Flu shot data

		No shot	Flu shot	Total
Flu?	Yes	15	3	18
	No	13	10	23
	Total	28	13	41

Want to test $H_0 : p_1 = p_2$ vs. $H_0 : p_1 > p_2$ where p_1 is probability of getting flu among those without shots and p_2 is probability of getting flu among those that got shots.

P-value for flu shot data

Tables where p_1 and p_2 are even further apart in the direction of $H_A : p_1 > p_2$

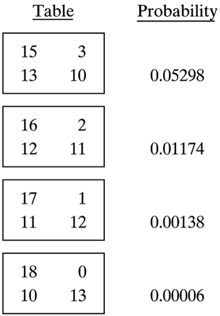

Table	Probability
15 3 13 10	0.05298
16 2 12 11	0.01174
17 1 11 12	0.00138
18 0 10 13	0.00006

Figure 10.4.1

$$\text{P-value} = 0.05298 + 0.01174 + 0.00138 + 0.00006 = 0.06616.$$

Fisher's exact test

The probability of each table is given by the *hypergeometric distribution* and is beyond the scope of this course, although your book does a nice job of explaining if you are interested. For the flu shot data to carry out Fisher's test we type

```
> flu=matrix(c(15,13,3,10),nrow=2)
> fisher.test(flu,alternative="greater")

        Fisher's Exact Test for Count Data

data:  flu
p-value = 0.06617
alternative hypothesis: true odds ratio is greater than 1
sample estimates:
odds ratio
  3.721944
```

We'll discuss what an odds ratio is next time. For now, we accept $H_0 : p_1 = p_2$ at the 5% level. There is not statistically significant evidence that getting a flu shot decreases the probability of getting the flu.

Directional alternatives

- Using `fisher.test` we can test $H_0 : p_1 = p_2$ versus one of (a) $H_A : p_1 \neq p_2$, (b) $H_A : p_1 < p_2$, or (c) $H_A : p_1 > p_2$.
- Use `alternative="two.sided"` (the default) or `alternative="less"` or `alternative="greater"`.
- Fisher's test is better than the chi-square test; just use the Fisher test in your homework.
- You will use `chisq.test` for tables larger than 2 × 2 instead, our next topic...

10.7 confidence interval for $p_1 - p_2$

- Recall, in population 1, we observe y_1 out of n_1 successes; in population 2 we observe y_2 out of n_2 successes, placed in a contingency table

		Group	
		1	2
Outcome	Success	y_1	y_2
	Failure	$n_1 - y_1$	$n_2 - y_2$
	Total	n_1	n_2

- $\hat{p}_1 = y_1/n_1$ estimates p_1 & $\hat{p}_2 = y_2/n_2$ estimates p_2.
- We want to compute a 95% confidence interval for $p_1 - p_2$.

Confidence interval for $p_1 - p_2$

- This interval is slightly different than your book's.
- The estimate of $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2$.
- The standard error is

$$SE_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}.$$

- At 95% confidence interval for $p_1 - p_2$ is

$$\hat{p}_1 - \hat{p}_2 \pm 1.96 SE_{\hat{p}_1 - \hat{p}_2}.$$

- This is given in R by `prop.test(success,total)` where `success` is a list of the number of **successes** in the two groups and `total` is a list of the **total number sampled** in each group.

Example 10.7.1 Migraine headache data

- Migraine headache patients took part in a double-blind clinical trial to assess experimental surgery.
- 75 patients were assigned real surgery ($n_1 = 49$) or sham surgery ($n_2 = 26$) so `total=c(49,26)`.
- There were $y_1 = 41$ successes among real surgery and $y_2 = 15$ successes among sham so `success=c(41,15)`.
- $\hat{p}_1 = 41/49 = 83.7\%$ & $\hat{p}_2 = 15/26 = 57.7\%$ so $\hat{p}_1 - \hat{p}_2 = 0.260$.
- The standard error of the difference is

$$SE_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{0.837(0.163)}{49} + \frac{0.577(0.423)}{26}} = 0.110.$$

- 95% confidence interval is $0.260 \pm 1.96(0.110) = (0.0444, 0.476)$.

R code for migraine headache data

Use `correct=FALSE` to get "old fashioned" confidence interval.

```
> total=c(49,26)
> success=c(41,15)
> prop.test(success,total,correct=FALSE)

        2-sample test for equality of proportions without continuity correction

data:  success out of total
X-squared = 6.0619, df = 1, p-value = 0.01381
alternative hypothesis: two.sided
95 percent confidence interval:
 0.04354173 0.47608150
sample estimates:
   prop 1    prop 2
0.8367347 0.5769231
```

We are 95% confident that real surgery reduces the probability of migraines by 4.3% to 47.6%.

R code for migraine headache data

Allowing the continuity correction changes the confidence interval a bit.

```
> prop.test(success,total)

        2-sample test for equality of proportions with continuity correction

data:  success out of total
X-squared = 4.7661, df = 1, p-value = 0.02902
alternative hypothesis: two.sided
95 percent confidence interval:
 0.01410688 0.50551635
sample estimates:
   prop 1    prop 2
0.8367347 0.5769231
```

We are 95% confident that real surgery reduces the probability of migraines by 1.4% to 50.6%. This interval is larger than the one on the previous slide.