Ch 20: Inference for comparing two proportions

Corinne Riddell

November 1, 2020

### Learning objectives for today

1. Learn about two methods to make confidence intervals for the difference between two proportions: i) the large sample methods, and ii) the plus four method
2. Conduct a hypothesis test for the difference between two proportions using a z-test
3. Learn how to use prop.test(), a function introduced last class, to conduct the hypothesis test in R.

### Comparing two proportions (Chapter 20)

* Two SRS from independent populations

Notation:

|  |  |  |  |
| --- | --- | --- | --- |
| Population | Population proportion | Sample size | Sample proportion |
| 1 |  |  |  |
| 2 |  |  |  |

### Large-sample confidence interval for the difference of two proportions

* Use when the number of observed successes and failures are > 10 for both samples
* Just like for the difference between two means, the SE of the difference is the square root of the sum of the variances.
* This large-sample interval often has low **coverage**. That is, if you repeated the method 100 times, fewer than 95 of the 100 created intervals would contain the true value for the difference between the proportions for a 95% CI.
* This is the same issue as the large sample method for one proportion.

### Example using the large sample method

Patients in a randomized controlled trial who were severely immobilized were randomly assigned to receive either Fragamin (to prevent blood clots) or a placebo. The number of patients experiencing deep vein thrombosis (DVT) was recorded:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | DVT | no DVT | Total |  |
| Fragamin | 42 | 1476 | 1518 |  |
| Placebo | 73 | 1400 | 1473 |  |

Check the conditions:

* We can apply the large study method because the sample sizes are large
* The number of observed successes and failures are larger than 10 (i.e., 42, 73, 1476, and 1400 all larger than 10).
* The estimate of the difference between the two proportions is
* We can use the large sample method to make a confidence interval for this difference

### Example using the large sample method

= 0.008 to 0.0358

Thus, the 95% confidence interval for the difference goes from 0.8% to 3.58%

### When dataset is small: Plus 4 method for the comparison of two proportions

* When the assumptions of the large sample method are not satisfied, we use the plus four method.
* When you have two samples this method says: add 4 observations total (2 to each sample),and 1 success and 1 failure to each of the two samples:
* Use when the sample size is at least five, with any counts of success and failure (can even use when number of successes or failures = 0)
* Much more accurate when the sample sizes are small
* May be conservative (has higher coverage then suggested by the confidence level)

### Example using the Plus Four Method

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Flu | no Flu | Total |  |
| Vaccine | 4 | 96 | 100 | 0.04 |
| Placebo | 11 | 89 | 100 | 0.11 |

Here, we don’t have 10 “successes” (flu) in both groups, so we cannot use the large sample method.

### Example using the Plus Four Method

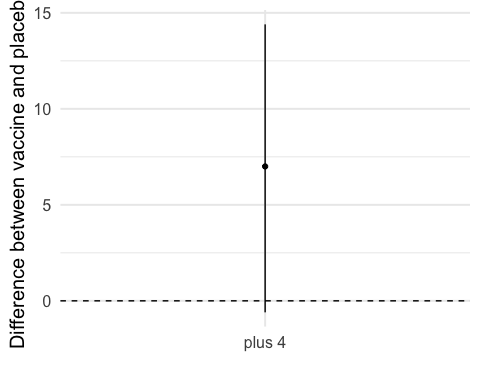
Filling in , and ==100 gives:

= -0.6% to 14.4%

The 95% CI of the difference ranges from -0.6 percentage points to 14.4% percentage points. While this CI contains 0 (the null hypothesized value for no difference) most of the values contained within the CI are positive, perhaps suggesting support for the alternative hypothesis. In this case, we might want to collect more data to create a more precise CI.

### Example using the Plus Four Method

The 95% CI of the difference ranges from -0.6 percentage points to 14.4% percentage points. While this CI contains 0 (the null hypothesized value for no difference) most of the values contained within it are positive, perhaps suggesting support for the alternative hypothesis. In this case, we might want to collect more data to create a more precise CI.



### Hypothesis testing when you have two samples and binary data

, which can also be written as

* Two-sided: or
* One-sided upper tail: or
* One-sided lower tail: or

### What does it mean to assume the null is true?

* If the null hypothesis is true, then is truly equal to . In this case, our best estimate of the underlying proportion that they are both equal to is
* Also, our best guess of the SE for is calculated by taking the formula for the SE for the difference between the two proportions and substituing in for and :

### Hypothesis testing when you have two samples and binary data

Using the information from the previous slide, we can create the z-test for the difference between two proportions as:

Use this z-test when the counts of successes and failures are 5 or larger in both samples

### Example of hypothesis testing when you have two samples and binary data

Recall the RCT data on the occurrence of deep vein thrombosis between Fragamin vs. placebo groups:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | DVT | no DVT | Total |  |
| Fragamin | 42 | 1476 | 1518 | 0.0277 |
| Placebo | 73 | 1400 | 1473 | 0.0496 |

, or that the proportion of DVT is the same between Fragamin and placebo groups.

Suppose you’re interested in knowing whether these two groups had different proportions of DVT. Then,

### Example of hypothesis testing when you have two samples and binary data

1. Compute
2. Compute the SE:
3. Compute the test statistic:
4. Calculate the p-value

pnorm(q = 3.112881, lower.tail = F)\*2

## [1] 0.001852707

The p-value is equal to 0.19%. Under then null hypothesis of no difference between the proportions, there is a 0.19% chance of observing the difference we saw (or more extreme) which provides evidence in favor of the alternative hypothesis that these proportions are different.

### Example of hypothesis testing when you have two samples and binary data

Here is how to conduct the hypothesis test using R’s prop.test():

prop.test(x = c(42, 73), # x is a vector of number of successes  
 n = c(1518, 1473),   
 correct = F) # n is a vector of sample sizes

##   
## 2-sample test for equality of proportions without continuity  
## correction  
##   
## data: c(42, 73) out of c(1518, 1473)  
## X-squared = 9.69, df = 1, p-value = 0.001853  
## alternative hypothesis: two.sided  
## 95 percent confidence interval:  
## -0.035708093 -0.008073386  
## sample estimates:  
## prop 1 prop 2   
## 0.02766798 0.04955872

Interpreting this output:

* Focus on the p-value output
* R is actually doing a chi-squared test (which we learn soon!). This test provides identical findings to performing a z-test for the equality of two proportions.
* In the output X-squared = 9.69 is the test statistic for the Chi-squared test (where Chi is the greek letter ). The X-squared test statistic is the z-test statistic squared. That is ,, which is what we found on previous slide.

### Example of hypothesis testing when you have two samples and binary data

Here is how to conduct the hypothesis test using R’s prop.test():

prop.test(x = c(42, 73), # x is a vector of number of successes  
 n = c(1518, 1473),   
 correct = T) # n is a vector of sample sizes

##   
## 2-sample test for equality of proportions with continuity correction  
##   
## data: c(42, 73) out of c(1518, 1473)  
## X-squared = 9.107, df = 1, p-value = 0.002546  
## alternative hypothesis: two.sided  
## 95 percent confidence interval:  
## -0.036376917 -0.007404562  
## sample estimates:  
## prop 1 prop 2   
## 0.02766798 0.04955872

* Typically, we would set correct = T to have R perform a continuity correction, which makes its test a bit better than the one we did by hand or without the correction.
* correct=T is the default option and the one we will use. Only set correct=F if you want to check work you did by hand

### Comparing treatments for UTIs

* In a study of urinary tract infections, patients were randomly assigned to one of two treatment regimes:
  + Treatment 1: trimethoprim / sulfamethoxazole or,
  + Treatment 2: fosfomycin / trometamol
* 92 of the 100 patients assigned to treatment 1 showed bacteriological cure while 61 of 100 assigned to treatment 2 showed bacteriological cure.

### Comparing treatments for UTIs

1. What is the estimate of the difference in proportions?
2. Perform a test to provide evidence whether or not this difference in proportions reflects a true difference between treatments.

### Comparing treatments for UTIs

1. What is the estimate of the difference in proportions?

.92 - .61 = .31. There is a 31 percentage point different between the two proportions.

* In online media “percentage point” is often abbreviated as “point”. So a “2 point difference” means a difference of 2 percentage points. Note how this is not the same as saying “a difference of 2 percent”!

### Comparing treatments for UTIs

1. Perform a test to provide evidence whether or not this difference in proportions reflects a true difference between treatments.

Large sample z-test of the form , where and

Without using R, we know the corresponding p-value will be very small.

pnorm(5.169899, lower.tail = F)\*2

## [1] 2.342205e-07

Interpretation of the p-value: Assuming the null hypothesis of no difference between the proportions, there is less than a 0.0001% chance of seeing a difference of the magnitude that we saw (or larger). This provides evidence in favor of the alternative hypothesis of a difference between the proportions