Chapter 21

Tests of Hypotheses for Comparing Two Proportions

In this chapter we will cover the following topic:

- Significance tests for comparing two proportions using the R-function **prop.test**

Significance Tests for Comparing Proportions

To test the difference between two proportions we type

- `prop.test(c(y1,y2),c(n1,n2))`

  where $y1$ and $y2$ are the number of successes in the two samples, and $n1$ and $n2$ are their respective sample sizes.

- To turn off the continuity correction, we add the argument **correct=FALSE** as in

  ```
  prop.test(c(y1,y2),c(n1,n2),correct=FALSE)
  ```

Let Y_1 be a simple random sample of size n_1 from a binomial distribution with unknown success probability p_1. Let Y_2 be an independent simple random sample of size n_2 from another binomial distribution with unknown success probability p_2. Recall that Y_1 and Y_2 are the number of successes. The two sample proportions are given by $\hat{p}_1 = \frac{Y_1}{n_1}$ and $\hat{p}_2 = \frac{Y_2}{n_2}$. If the sample sizes are sufficiently large, it follows from the Central limit theorem that the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is approximately normal with mean $p_1 - p_2$ and standard deviation

$$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}.$$

Since p_1 and p_2 are unknown, we can replace them by the sample proportions $\hat{p}_1$ and $\hat{p}_2$, respectively. We obtain the standard error of $\hat{p}_1 - \hat{p}_2$ which is given by

$$SE = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}. \tag{0.1}$$

To test the hypothesis that the two proportions are equal, that is, $H_0 : p_1 = p_2$, we use the observed value z of the test statistics

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{SE}.$$

This statistics is approximately standard normal when the null hypothesis is true. Some statisticians choose to replace both $\hat{p}_1$ and $\hat{p}_2$ in the denominator of Z by a pooled sample proportion. The pooled sample proportion is given by

$$\hat{p} = \frac{\text{total number of successes in both samples}}{\text{sum of sample sizes of both samples}} = \frac{Y_1 + Y_2}{n_1 + n_2}.$$

The test statistics is then given by

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

This test should only be used when the number of successes and failures are at least 5 in both samples [3]. The numerical results are about the same for these two tests statistics.
The alternative hypotheses are $H_a : (p - p_0) > 0$ with p-value $P(Z \geq z)$; $H_a : (p - p_0) < 0$ with p-value $P(Z \leq z)$, and $H_a : (p - p_0) \neq 0$ with p-value $2P(Z \geq |z|)$.

An approximate C confidence interval for the difference in proportion $p_1 - p_2$ is given by

$$p_1 - p_2 \pm z^* SE, \tag{0.2}$$

where SE is the standard error of $\hat{p}_1 - \hat{p}_2$ given in (0.1), and the critical value z^* is chosen such that the area under the Normal density curve is C between $-z^*$ and z^*. This formula should only be used when the number of successes and failures are at least 10 in both samples [3].

Problem. The Trial Urban District Assessment (TUDA) is a study sponsored by the government of student achievement in large urban school district. In 2009, 1311 of a random sample of 1900 eighth-graders from Houston performed at or above the basic level in mathematics [1]. In 2011, 1440 of a random sample of 2000 eighth-graders from Houston performed at or above the basic level [2]. (The study reports the proportions).

(A)Is there an increase in the proportion of eighth-graders who performed at or above the basic level in mathematics from 2009 to 2011 at the 5% significance level?
(B) Compute the 95% confidence interval for the difference in proportion of eighth-graders who performed at or above the basic level in mathematics from 2009 to 2011

Solution to part (a). Let p_1 and p_2 be the proportions of eighth-graders that performed at or above the basic level in mathematics in 2011 and 2009, respectively. We wish to test

$$H_0 : p_1 = p_2 \text{ against } H_a : p_1 > p_2.$$

We obtain:

```
> prop.test(c(1440,1311),c(2000,1900),alternative="greater",correct=FALSE)

        2-sample test for equality of proportions without continuity
        correction

data:  c(1440, 1311) out of c(2000, 1900)
X-squared = 4.2197, df = 1, p-value = 0.01998
alternative hypothesis: greater
95 percent confidence interval:
```

```
 0.005972807 1.000000000
sample estimates:
prop 1 prop 2
  0.72   0.69
```

The p-value=0.02 $<$ 0.05 so we reject H_0. Thus, there is evidence that there is an increase from 2009 to 2011 in the proportion of eighth-graders who performed at or above the basic level at the 5% significance level.

Solution to part (b). We obtain

```
 > prop.test(c(1440,1311),c(2000,1900),correct=FALSE)

        2-sample test for equality of proportions without continuity
        correction

data:  c(1440, 1311) out of c(2000, 1900)
X-squared = 4.2197, df = 1, p-value = 0.03996
alternative hypothesis: two.sided
95 percent confidence interval:
 0.001369833 0.058630167
sample estimates:
prop 1 prop 2
  0.72   0.69
```

Thus, we are 95% confident that the percent of eighth-graders who performed at or above the basic level in mathematics in 2011 is between 0.14% and 5.86% higher than in 2009.

Explanation. The code can be explained as follows:

- The **prop.test** in R computes a chi-squared statistic, X-squared, (see a later chapter for the definition of chi-squared statistics). This test is equivalent to the z-test when you take its square root.
- We added the argument **correct=FALSE** to turn off the continuity correction.

Problem. The use of helmet among recreational alpine skiers and snowboarders are generally low. A study from Norway [4] wanted to examine if helmet use reduces the risk of head injury. In the study, they compared the helmet use among skiers and snowboarders that was injured with a control group. The control group consisted of skiers and snowboarders that was uninjured. 96 of 578 people with head injuries used a helmet and 656 of 2992 people in the uninjured group used a helmet.
Is helmet use lower among skiers and snowboarders who had head injuries?

Solution. Let p_1 be the proportion of helmet use among injured skiers and snowboarders. Let p_2 be the proportion of helmet use among uninjured skiers and snowboarders. We wish to test

$$H_0 : p_1 = p_2 \text{ against } H_a : p_1 < p_2.$$

We obtain:

```
> prop.test(c(96,656),c(578,2992),alternative=c("less"),correct=FALSE)

        2-sample test for equality of proportions without continuity
        correction

data:  c(96, 656) out of c(578, 2992)
X-squared = 8.2336, df = 1, p-value = 0.002056
alternative hypothesis: less
95 percent confidence interval:
 -1.00000000 -0.02482216
sample estimates:
   prop 1    prop 2
0.1660900 0.2192513
```

The p-value$= 0.0021 < 0.01$ so we have strong evidence that helmet use is lower among skiers and snowboarders who had head injuries compared to uninjured skiers and snowboarders.

References

[1] National Center for Education Statistics, *The Nation's Report Card, Mathematics 2009, Trial Urban District Assessment, Results at grades 4 and 8*, Institute of Education Sciences, U.S. Department of Education, NCES 2010-452

[2] National Center for Education Statistics, *The Nation's Report Card, Mathematics 2011, Trial Urban District Assessment, Results at grades 4 and 8*, Institute of Education Sciences, U.S. Department of Education, NCES 2012-452

[3] D. S. Moore, W. I. Notz, M. A. Fligner, R. Scoot Linder. *The Basic Practice of Statistics.* W. F. Freeman and Company, New York, 2013.

[4] S. Sulheim, I. Holme, A. Ekeland, R. Bahr. *Helmet use and risk of head injuries in alpine skiers and snowboarders.* Journal of the American Medical Association, 295, pp 919-924, 2006.

00