**Chapter 10.1: Comparing Two Proportions**

*The sampling distribution of a difference between two proportions*

Why might we want to compare two proportions?

We start with two different samples:

|  |  |  |  |
| --- | --- | --- | --- |
| **Population or treatment** | **Parameter** | **Statistic** | **Sample size** |
| 1 |  |  |  |
| 2 |  |  |  |

The sampling distribution of

|  |  |
| --- | --- |
| Choose an SRS of size from population 1 with a proportion of successes and an independent SRS of size from population 2 with a proportion of successes . | |
| **Shape** |  |
| **Center** |  |
| **Spread** |  |

Draw the picture of the sampling distribution of (Figure 10.2)

*Example: Who Does More Homework?*

**Finding probabilities with the sampling distribution**

Suppose that there are two large high schools, each with more than 2000 students, in a certain town. At School 1, 70% of students did their homework last night. Only 50% of the students at School 2 did their homework last night. The counselor at School 1 takes an SRS of 100 students and records the proportion that did homework. School 2’s counselor takes an SRS of 200 students and records the proportion  that did homework. School 1’s counselor and School 2’s counselor meet to discuss the results of their homework surveys. After the meeting, they both report to their principals that  = 0.10.

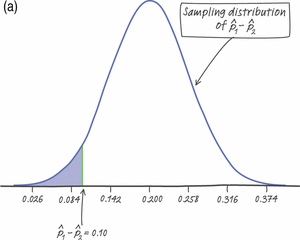
**PROBLEM:**

**(a) Describe the shape, center, and spread of the sampling distribution of .**

**Because *n*1*p*1 = 100(0.7) = 70, *n*1 (1 − *p*1) = 100(0.3) = 30, *n*2*p*2 = 200(0.5) = 100, and *n*2 (1 − *p*2) = 200(0.5) = 100 are all at least 10, the sampling distribution of** **is approximately Normal. Its mean is *p*1 − *p*2 = 0.70 − 0.50 = 0.20 and its standard deviation is http://ebooks.bfwpub.com/tps4e/pics/resized_ch10_ineqn607-01.jpg**

**(b) Find the probability of getting a difference in sample proportions** **of 0.10 or less from the two surveys. Show your work.**

[**Figure 10.3(a)**](javascript:top.OpenSupp('figure','10','3a'))**shows the sampling distribution of** **with the desired probability (area) shaded. To find this probability, we use the methods of**[**Chapter 2**](javascript:top.JumpToChapter('2'))**.**

[](javascript:top.OpenSupp('figure',10,'3a'))

[**Figure 10.3a**](javascript:top.OpenSupp('figure',10,'3a'))  (a) Normal curve that approximates the sampling distribution of  for the homework surveys. We want to find *P*( ≤ 0.10).

*Standardize.* When  = 0.10,

http://ebooks.bfwpub.com/tps4e/pics/resized_ch10_eqn608-01.jpg

*Use*[***Table A***](http://ebooks.bfwpub.com/tps4e/frontmatter/TableA.pdf)*.* The area to the left of *z* = −1.72 under the standard Normal curve is 0.0427. This is the probability we seek. We could use technology to get the desired probability. Using normalcdf(0,0.i,.2,.058) gives 0.0421.

[**Figure 10.3(b)**](javascript:top.OpenSupp('figure','10','3b')) shows this result.

[](javascript:top.OpenSupp('figure',10,'3b'))

[**Figure 10.3b**](javascript:top.OpenSupp('figure',10,'3b'))  (b) The desired probability can also be expressed as *P*(*z* < −1.72), which represents an area under the standard Normal curve.

**(c) Does the result in part (b) give us reason to doubt the counselors’ reported value? Explain.**

There is only about a 4% chance of getting a difference in sample proportions as small as or smaller than the value of 0.10 reported by the counselors. This does seem suspicious!

**CHECK YOUR UNDERSTANDING**

Your teacher brings two bags of colored goldfish crackers to class. She tells you that Bag 1 has 25% red crackers and Bag 2 has 35% red crackers. Each bag contains more than 500 crackers. Using a paper cup, your teacher takes an SRS of 50 crackers from Bag 1 and a separate SRS of 40 crackers from Bag 2. Let  be the difference in the sample proportions of red crackers.

**1.** What is the shape of the sampling distribution of ? Why?

*Hint: check sample size condition*

**2.** Find the mean and standard deviation of the sampling distribution. Show your work.

*Hint: use the formulas for mean and standard deviation in your table on the first page of notes.*

**3.** Find the probability that  is less than or equal to −0.02. Show your work.

*Hint: standardize the difference between two proportions using , then find the area under the standard normal curve using Table A or normcdf( ) on the calculator.*

**4.** Based on your answer to Question 3, would you be surprised if the difference in the proportion of red crackers in the two samples was  = −0.02? Explain.

*Hint: what is the probability you found in #3? So small you find it hard to believe the claim?*