## Task 1 – Examine the research question

In 2002, a study was included in the October issue of the Archives of Pediatrics and Adolescent Medicine. In it, researchers reported the results of their experiment to determine if over-the-hardware-counter duct tape is a more effective, less painful alternative to liquid nitrogen, which is used to freeze warts.

In the study conducted at the Madigan Army Medical Center near Tacoma, Washington, 26 patients wore duct tape over their warts for six days. Then they removed the tape, soaked the area in water and used an emery board or pumice stone to scrape the spot. The tape was reapplied the next morning. The treatment continued for a maximum of two months or until the wart went away.

At the same time, 25 patients with similar warts had repeated visits to the center where their warts were treated with cryotherapy where liquid nitrogen was applied to the warts by a medical professional.

Of the 26 patients treated with duct tape, 22 got rid of their warts compared with 15 out of the 25 patients who received the freezing treatment. All patients were between the ages of 3 and 22.

Researchers believe the duct tape irritated the warts, and that apparently caused an immune system reaction that attacked the growths. They did not test other kinds of tape, and so they cannot say whether there is anything special about the gray, heavy-duty, fabric-backed tape.   
  
Researchers did not test the duct tape on older adults and also did not study whether warts recurred.

Organize the information from the study in the contingency table below:

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | Wart Gone | Wart Remains | Total |
| Duct Tape |  |  |  |
| Cryotherapy |  |  |  |
| Total |  |  |  |

Calculate the difference between the proportion of patients in the Duct Tape group whose warts were removed and the proportion of patients in the Cryotherapy group whose warts were removed, that is, .

We would like to know if this difference is convincing evidence that duct tape therapy is more effective than cryotherapy.

Our observed difference of suggests that the duct tape therapy is effective. However, we cannot yet be sure if this observed difference represents a real treatment effect or is just from chance. Generally there is a little bit of fluctuation in sample data and we wouldn’t expect the sample proportions to be exactly the same, even if the truth was that duct tape was no better than cryotherapy at removing warts.

We label two competing claims:

**H0: Null hypothesis.**

**HA: Alternative hypothesis:**

Let’s investigate the null hypothesis further:

Suppose that Duct Tape and Cryotherapy are equally effective, that is, the type of therapy and the results are independent. What would the independence model look like?

First we need to calculate the proportion of patients overall whose warts were successfully removed, then fill in the table below for the independence model:

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | Wart Gone | Wart Remains | Total |
| Duct Tape |  |  |  |
| Cryotherapy |  |  |  |
| Total |  |  |  |

Under the independence model, the difference is 0.01, very close to zero. Thus, given that the null hypothesis is true, any sample data we collect should give us an observed difference that varies slightly from 0. We’ll next investigate how much it ought to vary.

## Task 2 – Simulate the experiment by hand

We are now going to simulate the experiment under the assumption that the independence model is true. To do this, follow the steps below.

1. In your activity materials you will find 51 craft sticks, one for each of the patients in the study. 26 of the sticks should be plain wood and will represent patients who received duct tape therapy. The other 25 sticks should be colored and will represent patients who received cryotherapy. Verify that this true. Add or remove sticks if necessary—your instructor has extras.
2. Shuffle the sticks well.
3. Randomly draw 37 sticks from the bag. This will represent the 72.5% of patients whose warts were removed. Divide the number of plain sticks in this pile by 26. This gives you
4. Count the number of colored sticks out of the 37 you draw from the bag. Divide this number by 25 to calculate .
5. Calculate the difference in your results from steps 3 and 4,  
     
    = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
6. Write your difference on a post-it note and place it on the board at the front of the classroom when and as instructed. If your difference is larger than 0.246, circle your number.
7. If you have time, repeat steps 2-6. The more simulations we can do, the better we will be able to visualize the distribution of differences under the null model. Your instructor also has extra post-it notes.

Calculate the proportion of the class results that were 0.246 or larger and record it here:

## Task 3 – Use a computer to simulate results

Go to the website: <https://shiny.stt.msu.edu/fairbour/TwoProportion/>

Make sure the “Duct Tape Therapy” preset is selected. Press the reset button just to be safe.

Where the app says “Shuffle how many times?” enter the same number of simulations as your class was able to do and record with the post-it notes. Press the “Shuffle” button.

Compare the distribution of dots on the app to the results of your class. Are they similar? Different? What proportion of computer simulated results were 0.246 or larger?

1. Increase the number of shuffles to 300 and press the “Shuffle” button again. Under the dotplot, enter the value 0.246 in the box and make sure “greater than” is selected. Now what is the proportion of results that are 0.246 or larger?
2. Continue to press the Shuffle button to increase the number of simulations. As the number of simulations increases, where does the proportion of results that are 0.246 or larger seem to settle?

**The observed difference of 24.6% is a rare event if there really is no difference in the therapies.** We have two possible interpretations for the study result:

**H0: Null hypothesis:** The two therapies are equally effective at removing warts, and we observed a difference that is so large it would only happen rarely.

**HA: Alternative hypothesis:** Duct tape is a more effective treatment for warts.

When we conduct formal studies, we reject the independence model if our data strongly conflict with it. From our simulation, we determined that there was approximately a 4% probability of obtaining a sample where ≥ 24.6% more patients with the duct tape therapy were successful at removing their warts. In this case, we reject the null hypothesis in favor of the alternative.

## Task 4 – Try another example

The California Department of Public Health defines an overcrowded household as one where there is more than one person per room in the home. They explain that residential crowding has been linked to an increased risk of infection from communicable diseases, a higher prevalence of respiratory ailments, and greater vulnerability to homelessness among the poor.   
  
In an effort to determine the extent of the problem of household overcrowding, suppose public health officials take a random sample of 400 households in Madera County, California and find that 45 of the households in the sample are overcrowded. They take a random sample of 400 households in nearby Kings County and find that 34 of the households in the sample are overcrowded.

Use the table below to organize this information.

|  |  |  |  |
| --- | --- | --- | --- |
| County | Overcrowded | Not Overcrowded | Total |
| Madera |  |  |  |
| Kings |  |  |  |
| Total |  |  | 800 |

Calculate the observed difference = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**H0: Null hypothesis.** The variables **county** and **overcrowded** are independent. They have no relationship, and the observed difference of 0.027 was due to chance.

**HA: Alternative hypothesis:** The variables **county** and **overcrowded** are *not* independent. The difference in the overcrowded rate of 0.027 was not due to chance, and Madera County is more overcrowded than Kings County.

While it would be possible to take 800 cards (or craft sticks) and perform a simulation by hand, we can much more easily simulate this on a computer. Enter the values from your table into the app. Press the reset button. Count the samples greater than 0.027 and shuffle at least 1000 times. Record your result here.

Based on your result, which interpretation would you choose? The null hypothesis, or the alternative? State which and briefly explain your answer.