Inference for multiple proportions

Stat 120

May 15 2023

Tests for One Categorical Variable

Goodness-of-fit test

- Test a claim about the distribution of one categorical variable
- E.g. are 6 M&M colors equally likely?
- E.g. is Biden's approval rating 50%?

Tests for One Categorical Variable

Seen single proportion tests before

- ***Example : Test if the proportion of Reese's Pieces that are orange is different from 1/3.***

$$H_0 : p = 1/3$$
$$H_a : p \neq 1/3$$

What if we want to test proportions for several categories at once?

- **Example: Are the three colors (orange, yellow, brown) of Reese's Pieces equally likely?**

H_0 **specifies a proportion,** p_i **, for each category.**

Test for one categorical variable

The proportions do not have to be the same. e.g: Grade distribution

$$
\begin{aligned}
&H_0: p_A = 0.2, \quad p_B = 0.3, \quad p_C = 0.3, \quad p_D = 0.1, \quad p_F = 0.1 \\
&H_a: \text{ at least one proportion different}
\end{aligned}
$$

Rock-Paper-Scissors

ROCK	PAPER	SCISSORS	TOTAL
36	12	37	85

How would we test whether all of these categories are equally likely?

Conduct a hypothesis test

- *State Hypothesis*
- *Calculate a test statistic, based on your sample data*
- *Create a distribution of this statistic, as it would be observed if the null hypothesis were true*
- *Measure how extreme your test statistic is, as compared to the distribution generated under null*

Test Statistic

Why can't we use the familiar formula to get the test statistic?

$$\frac{\text{sample statistic - null value}}{\text{SE}}$$

- More than one sample statistic
- More than one null value

We need something a bit more complicated ...

Observed Counts

The observed counts are the actual counts observed in the study

ROCK	PAPER	SCISSORS	TOTAL
36	12	37	85

- **The expected counts are the expected counts if the null hypothesis were true**
- **For each cell, the expected count is the sample size n times the null proportion, p_o**

	ROCK	PAPER	SCISSORS	TOTAL
Observed	36	12	37	85
Expected	28.33	28.33	28.33	85

Chi-Square Statistic

- A test statistic is one number, computed from the data, which we can use to assess the null hypothesis
- The chi-square statistic is a test statistic for categorical variables:

$$\chi^2 = \sum \frac{(observed - expected)^2}{expected} = \sum \frac{(O - E)^2}{E}$$

Rock-Paper-Scissors

	ROCK	PAPER	SCISSORS	TOTAL
Observed	36	12	37	85
Expected	28.33	28.33	28.33	85

$$\chi^2 = \frac{(36 - 28.33)^2}{28.33} + \cdots + \cdots$$
$$\approx 2.076 + \cdots + \cdots$$

What next?

We have a test statistic. What else do we need to perform the hypothesis test?

- *A distribution of the test statistic assuming H_0 is true*

How do we get this? Two options:

1. *Simulation*
2. *Theoretical Distribution*

Simulation

1. Take 3 scraps of paper and label them Rock, Paper, Scissors. Fold or crumple them so they are indistinguishable. Choose one at random and record the result.
2. Repeat a number of times to match the original sample size and get a table of observed counts.
3. Calculate the χ^2-statistic.
4. Repeat this many times to get a randomization distribution of many χ^2-statistics.
5. How extreme is the actual test statistic in this randomization distribution?

Statkey: Chi-Square Distribution

StatKey Chi-square Goodness-of-Fit

Custom Dataset | Show Data Table | Edit Data

Generate 1 Sample | Generate 10 Samples | Generate 100 Samples | Generate 1000 Samples | Reset Plot

Randomization Dotplot of χ^2 , Null Hypothesis

Original Sample

Show Details

$n = 85, \chi^2 = 14.142$

	Count
Rock	36
Paper	12
Scissors	37

Randomization Sample

Show Details

$n = 85, \chi^2 = 2.989$

	Count
Rock	22
Paper	28
Scissors	35

Chi-Square Distribution

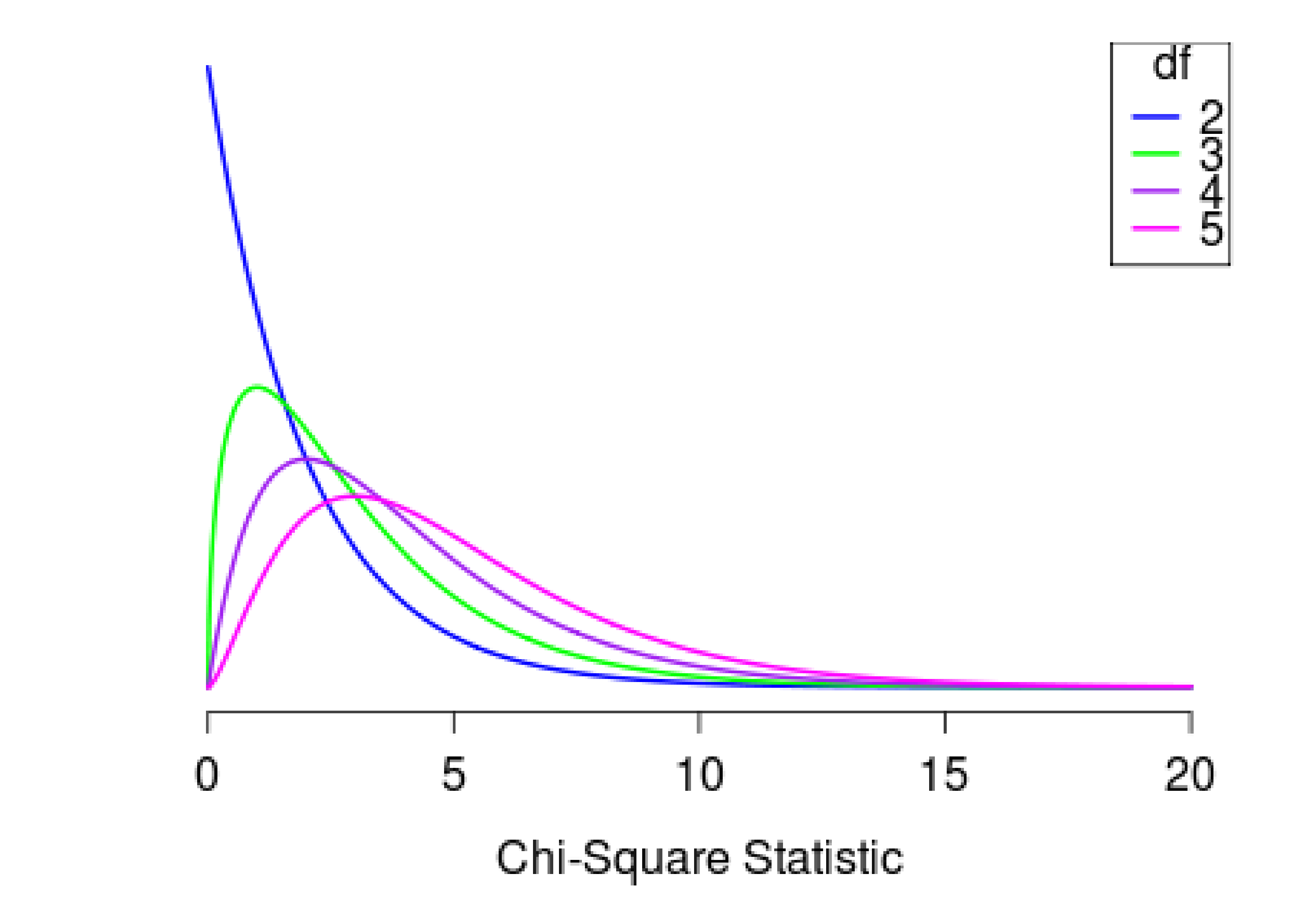
If each of the expected counts are at least 5, AND if the null hypothesis is true, then the χ^2 statistic follows a χ^2 distribution, with degrees of freedom equal to

$$df = \text{number of categories} - 1$$

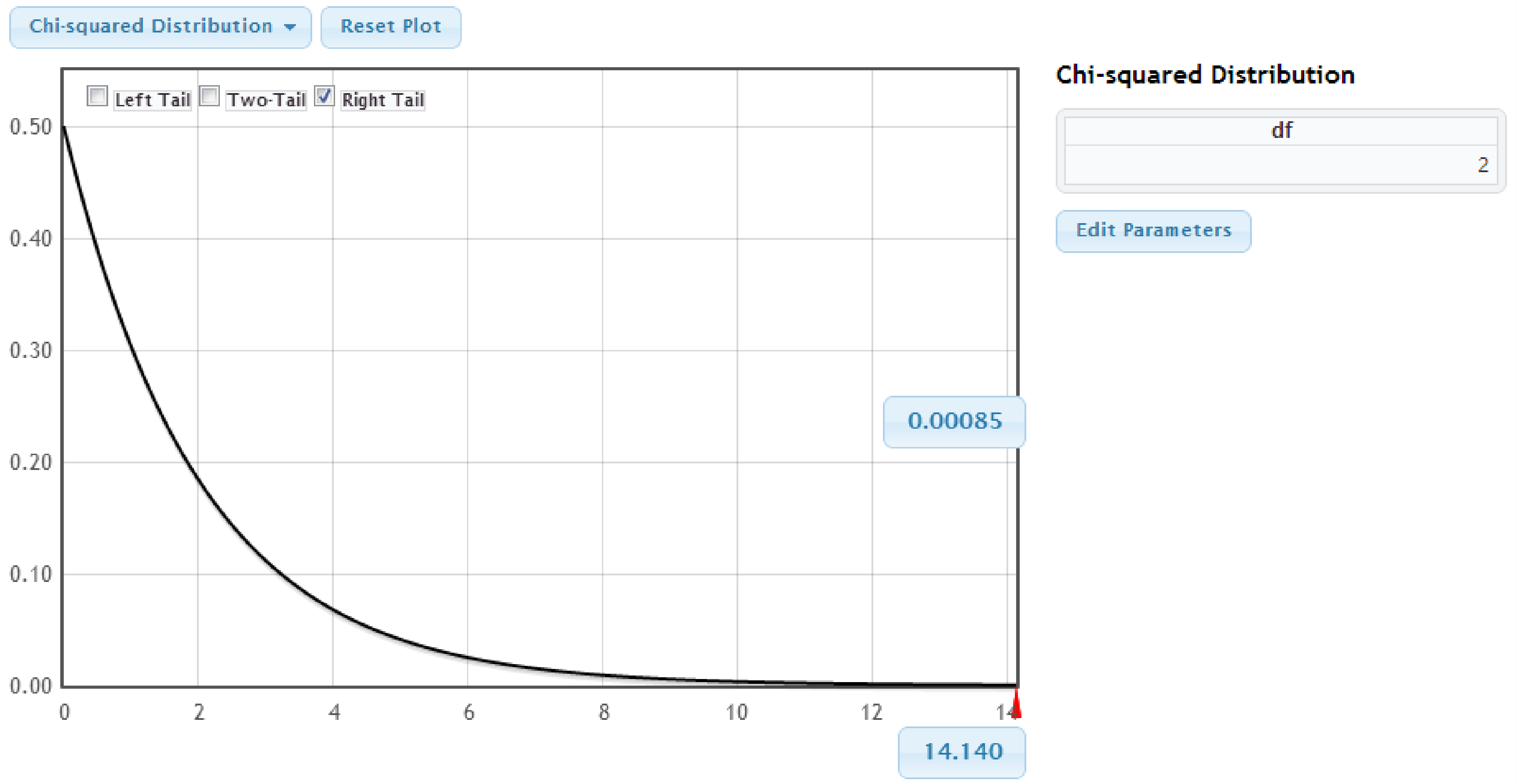
Rock-Paper-Scissors:

```
df = 3 - 1 = 2 # degrees of freedom
```

Chi-Square Distribution



Statkey: p-value using Chi-square distribution



Goodness of Fit

- A χ^2 test for goodness of fit test determines whether the distribution of a categorical variable is the same as some null hypothesized distribution
- The null hypothesized proportions for each category do not have to be the same

Chi-Square Test for Goodness of Fit

- State null hypothesized proportions for each category, pi. Alternative is that at least one of the proportions is different than specified in the null.
- Calculate the expected counts for each cell as $n \cdot p_i$. Make sure they are all greater than 5 to proceed.
- Calculate the χ^2 statistic: $\chi^2 = \sum \frac{(observed - expected)^2}{expected}$
- Compute the p-value as the area in the tail above the χ^2 statistic, for a χ^2 distribution with $df = (\text{number of categories - 1})$.
- Interpret the p-value in context.

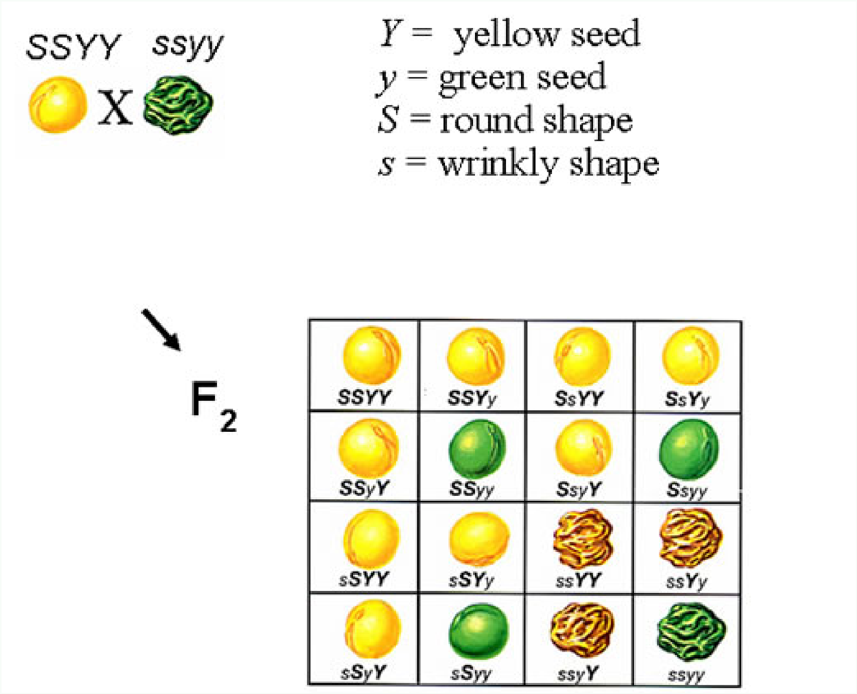
Mendel’s Pea Experiment

Between 1856 and 1863 Gregor Mendel cultivated and tested roughly 29,000 pea plants. This study showed that one in four pea plants was pure-bred recessive, two out of four were hybrid and one out of four were pure-bred dominant.

Mendel's work was rejected at first and was not widely accepted until after he died. He is now known as the “father of modern genetics

Source: Mendel, Gregor. (1866). Versuche ber Pflanzen-Hybriden. Verh. Naturforsch. Ver. Brünn 4: 3–47 (in English in 1901, Experiments in Plant Hybridization, J. R. Hortic. Soc. 26: 1–32)

Mendel's Pea Experiment



S, Y: **Dominant** and s, y: **Recessive**

Mate SSYY with ssyy: 1st Generation: all Ss Yy

Mate 1st Generation: → 2nd Generation

Phenotype	Theoretical Proportion
Round, Yellow	9/16
Round, Green	3/16
Wrinkled, Yellow	3/16
Wrinkled, Green	1/16

Mendel's Pea Experiment

Phenotype	Theoretical Proportion	Observed Counts
Round, Yellow	9/16	315
Round, Green	3/16	101
Wrinkled, Yellow	3/16	108
Wrinkled, Green	1/16	32

Let's test this data against the null hypothesis of each p_i equal to the theoretical value, based on genetics

$$H_0 : p_1 = 9/16, p_2 = 3/16, p_3 = 3/16, p_4 = 1/16$$

$$H_a : \text{ At least one } p_i \text{ is not as specified in } H_0$$

Mendel’s Pea Experiment

What is the expected count for round green peas?

A. 0.182

B. 108

C. 104.25

D. 139

Phenotype	Theoretical Proportion	Observed Counts
Round, Yellow	9/16	315
Round, Green	3/16	101
Wrinkled, Yellow	3/16	108
Wrinkled, Green	1/16	32

▶ Click for answer

Mendel's Pea Experiment

Phenotype	Theoretical Proportion	Observed Counts
Round, Yellow	9/16	315
Round, Green	3/16	101
Wrinkled, Yellow	3/16	108
Wrinkled, Green	1/16	32

The χ^2 statistic is the sum of 4 components (because there are 4 categories). What is the contribution to the χ^2 statistic from the "Round, Yellow" category?

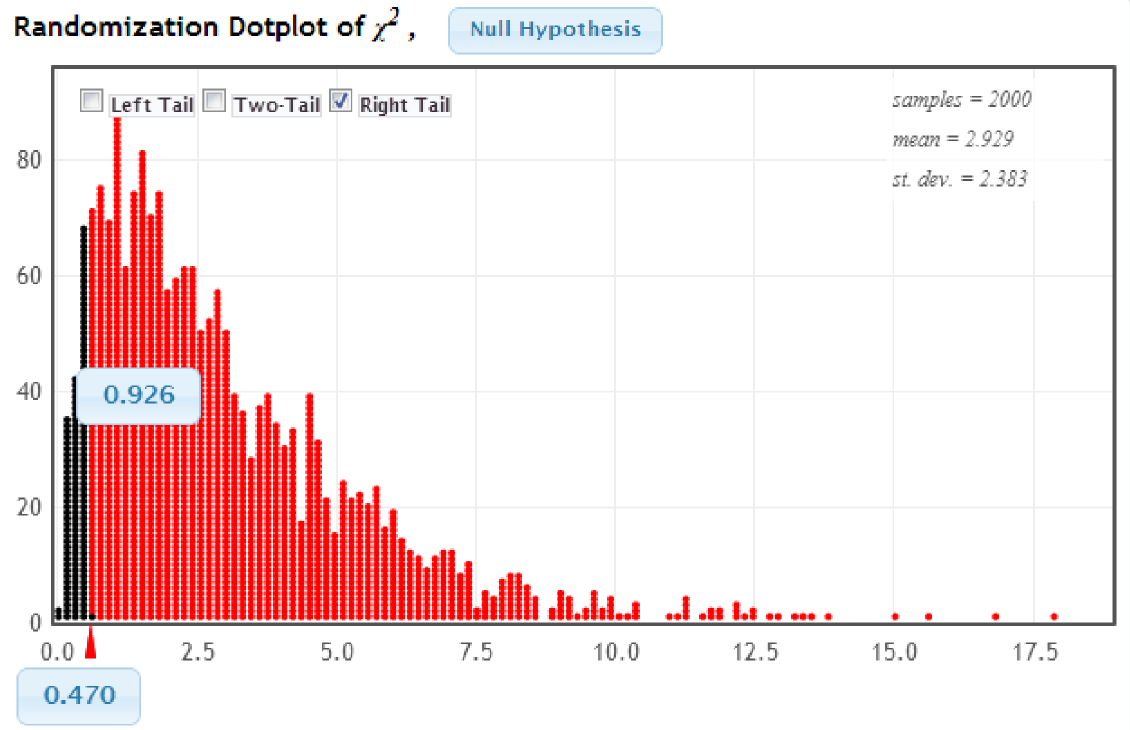
1. 0.016
2. 1.05
3. 5.21
4. 107.2

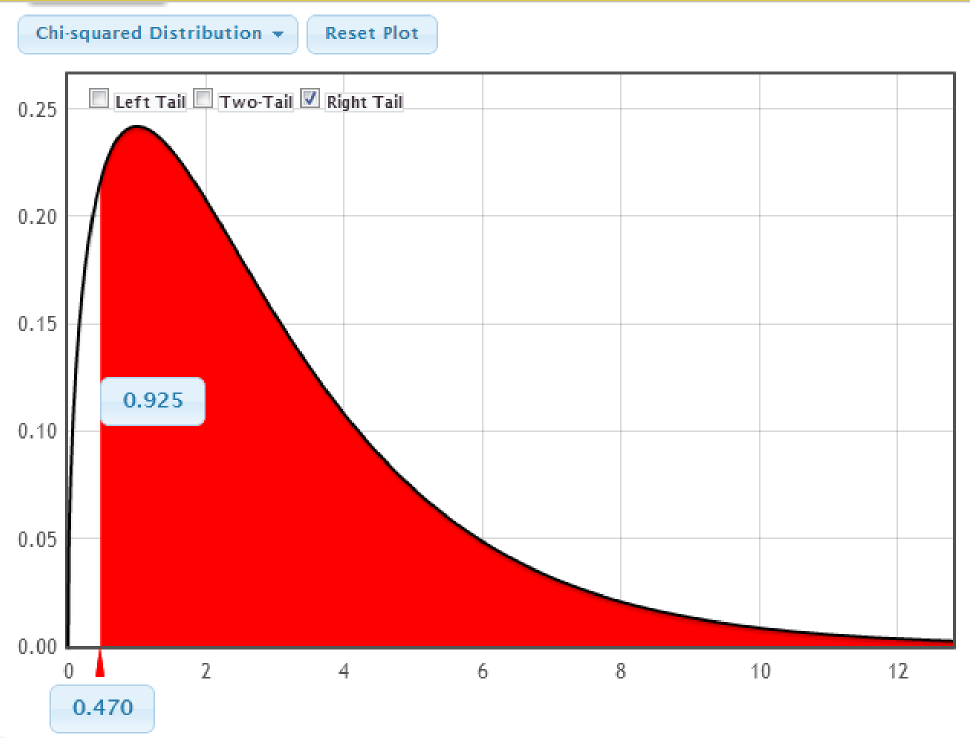
▶ Click for answer

Mendel's Pea Experiment: $\chi^2 = 0.47$

Two options:

- ***Simulate a randomization distribution***
- ***Compare to a χ^2 distribution with $df = 4 - 1$***





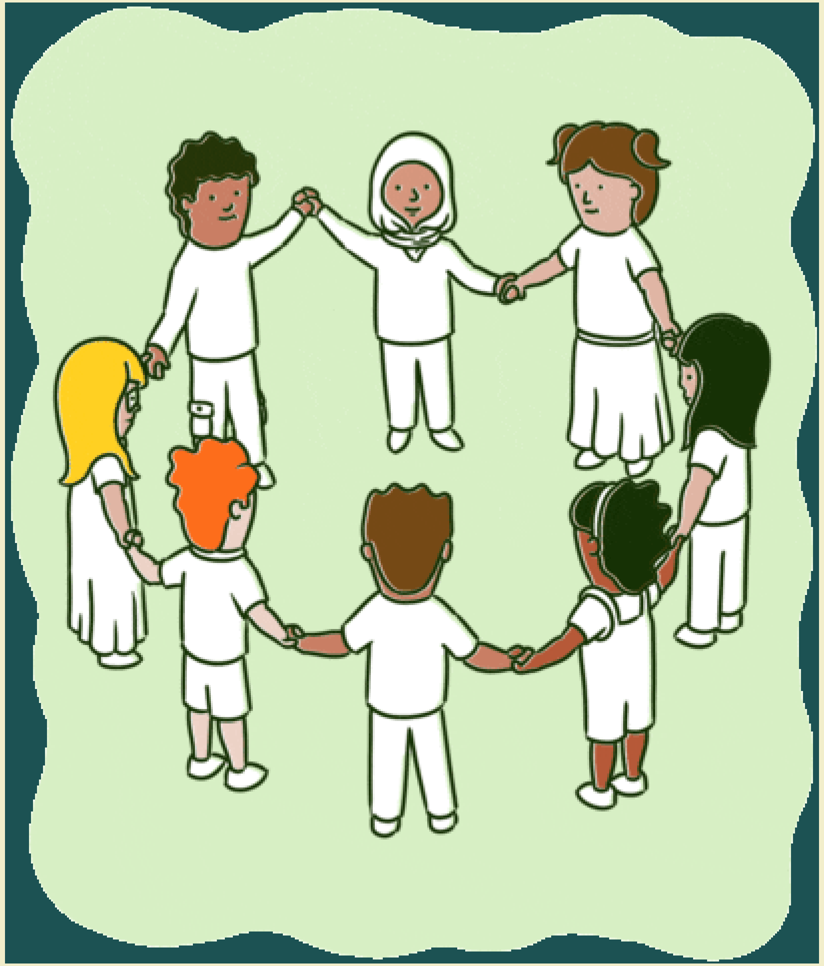
Does this prove Mendel's theory of genetics? Or at least prove that his theoretical proportions for pea phenotypes were correct?

1. Yes
2. No

▶ Click for answer

YOUR TURN 1

05:00



- Go over to the in class activity file
- Complete the activity as much as possible