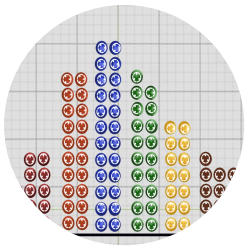

Math, Models, and M&Ms

Hypothesis Testing for Proportions

Ryan Murphy, Lydia Baker, Kelsey Willems
Published by *K20 Center*

Grade Level	11th – Secondary Grade	**Time Frame**	100 minutes
Subject	Mathematics	**Duration**	2–3 Class Periods
Course	AP Statistics, Statistics		

Essential Question

How can we use the sample proportions of M&M colors in a bag to test the manufacturer's claim?

Summary

In this lesson, students will investigate the proportion of one of the colors of M&Ms in fun-sized bags and test a hypothesis regarding the claimed distribution of M&M colors using statistical methods. Students will collect data to test the distribution of one color of M&Ms by creating confidence intervals, testing hypotheses, and calculating p-values from binomial and normal distributions.

Snapshot

Engage

Students compare the color distribution in a fun-sized bag of M&Ms to the company's claim and calculate a p-value using the binomial distribution for the color they find most unusual.

Explore

Students write hypotheses for the proportion of orange M&Ms, select a significance level, check test conditions, and determine if the sample size is sufficient for 1-proportion Z-test.

Explain

Students combine class data to create a larger sample, allowing them to use a normal approximation to conduct a 1-Proportion Z-test and interpret the results.

Extend

Students collect a second sample of data, combine trials, re-test the hypothesis, and construct a confidence interval for the true proportion of orange M&Ms.

Evaluate

Students graph cumulative relative frequency and margin of error versus sample size to determine if and when the sample proportion becomes statistically significant.

Standards

AP Statistics Course and Exam Description (Course at a Glance)

6.2: Constructing a Confidence Interval for a Population Proportion
6.3: Justifying a Claim Based on a Confidence Interval for a Population Proportion
6.4: Setting Up a Test for a 1 Population Proportion
6.5: Interpreting p-Values

Attachments

- Color by Numbers—Math, Models, and M&Ms.docx
- Color by Numbers—Math, Models, and M&Ms.pdf
- Lesson Slides—Math, Models, and M&Ms.pptx
- Testing M&Ms Claim—Math, Models, and M&Ms.xlsx
- Visualize Cumulative Data (Teacher Guide)—Math, Models, and M&Ms.docx
- Visualize Cumulative Data (Teacher Guide)—Math, Models, and M&Ms.pdf

Materials

- Lesson Slides (attached)
- Color by Number handout (attached; one per student)
- Visualize Cumulative Data (Teacher Guide) (attached)
- Fun-size bags of candy-coated chocolate (e.g., M&Ms) (at least 3 fun-size bags per student)
- Testing M&Ms Claim spreadsheet (see preparation note)
- Calculators or statistical software (e.g., Desmos, TI-84, or Excel/Google Sheets)
- Graphing tools (paper or digital)
- Pens/pencils
- Napkins, paper plates, or paper towels

10 minutes

Preparation

Preparing for Data Analysis

1. Download the attached **Testing M&Ms Claim** spreadsheet and make it available to students by uploading to a class LMS or to Google Sheets.
2. Duplicate the “Class Data” tab on the spreadsheet so that there is one for each group in the class, and rename each tab to be unique to each group.
3. Within each tab, replace the letters under the “student” column with either the student's name or number. If numbers, make sure to assign a number to each student.
4. Add the link or QR code to **slide 12** for students to have access.

In their tab, students will record the frequency of orange M&Ms and total M&Ms per fun-size bag in the table, and the spreadsheet will auto-populate a scatter plot, cumulative total, cumulative frequency, and p-value. Review the “Example” tab for a completed trial with data to see how the spreadsheet functions.

Data from the first trial will be used in the Engage through Explain sections. We will gather and analyze the data from subsequent trials in the Extend of the lesson. Feel free to adjust the number of trials your class collects to accommodate the time and the number of students in your class.

10 minutes

Engage

Use the attached **Lesson Slides** to guide the lesson. Start by posing the essential question and explaining the objectives using **slides 3–4**.

Move to **slide 5** and distribute a fun-size bag of M&Ms and a copy of the attached **Color By Numbers** handout to each student. Have students sort the M&Ms and record their sample distribution in the frequency table for question 1a on the handout.

Move to **slide 6** and ask the students what seemed the most unusual about their sample (i.e., "Half my M&Ms were blue," or "I didn't have any blue M&Ms"). After a brief discussion, have students identify the sample proportion that seemed most unusual to them for question 1b and give them time to fill in their handout. Then, transition to **slide 7** and reveal the actual claimed color distribution for chocolate M&Ms as it was in 2008, provided by the manufacturer:

- 24% blue
- 20% orange
- 16% green
- 14% yellow
- 13% red
- 13% brown

Ask students to write these distributions on their handout, using the provided lines under the corresponding colored M&Ms.

Then direct them to calculate the p-value based on the assumption that distribution provided by the manufacturer is true for question 1c.

Teacher's Note: Optional Review

If your students have prior knowledge of using binomial distribution to set up and calculate a p-value, move to hidden **slide 8** and ask students to choose the color proportion that is most unusual for them (or ask all students to choose orange) and have students calculate the p-value of that proportion using the binomial model. Have them use graphing technology to calculate and interpret their p-value with the binomial (some may accidentally use a normal model).

For example: If,

$$H_o : p = 0.20$$

then the probability of observing 8 orange M&Ms out of a sample of 16 is 0.007.

This would be statistically significant at a significance level of 5% to allow one to reject the null hypothesis.

Consider using this Binomial Distribution Applet/Calculator to allow students to visualize the binomial distribution.

Teacher's Note: Original Data Source

Refer to Rick Wicklin's article "The distribution of colors for plain M&M candies" for more information about the data being used in this lesson.

Explore

Explain that these proportions were once listed on the website but have since been removed. Let's investigate if the population proportion is still 20% orange.

Explain to students that we don't yet know how to test the distribution of all 6 colors at once so we are going to focus on the proportion of orange M&Ms. Have students get with a partner. As you work through **slides 9–10,** have students use the Elbow Partners strategy to develop answers to each prompt for question #2 of their handout.

- **Null Hypothesis**: $H_o : p = 0.20$
- **Alternative Hypothesis**: $H_A : p \neq 0.20$
- **Select an alpha level**: $\alpha = .05$

- **Random Condition**: Was the sample randomly selected?
- **10% Condition**: Is the sample size <10% of the population?
- **Success/Failure Condition**: $np \geq 10$ and $n(1-p) \geq 10$

Teacher's Note: Optional Lesson Guidance

If your students are in a place where this additional information would be helpful, consider unhiding **slide 11** and provide an explanation or additional details about "Success/Failure Condition."

30 minutes

Explain

Using **slide 12**, have students navigate to the linked spreadsheet where they will record their sample data. First, ask students to input their data from the first bag of M&Ms as Trial 1.

Once all data is in the spreadsheet, move to **slide 13** and have students return to their Color by Number handout. Instruct students to analyze the data by answering all of question #3 on their handout with their elbow partner.

Use **slides 14–20** to lead a discussion on the statistical methods applied in the Explore phase concerning the following topics:

- Hypothesis testing and p-values
- Checking assumptions for normality
- Confidence interval interpretation, 1-proportion Z-test, and margin of error
- Hypothesis testing steps include setting up null and alternative hypotheses, calculating test statistics, and alpha levels
- Using binomial model for small samples
- Interpreting the p-value and making conclusions about the distribution of M&M colors

10 minutes

Extend

Transition to **slide 21.** Give each student at least 1 fun-size bag of M&Ms and have them record the frequency of orange M&Ms and the sample size in the spreadsheet for trial 2. Use the class spreadsheet to track the cumulative relative frequency of orange M&Ms as the class sorts more M&Ms.

Once all data is in the spreadsheet, have students return to their Color by Number handout. Instruct students to analyze the data by answering question #4 on their handout with their elbow partner.

Use the attached **Visualize Cumulative Data (Teacher Guide)** handout to help students visualize how cumulative data narrows the margin of error and changes the p-value, potentially crossing the significance threshold.

Teacher's Note: Optional Extension

Increase the power or sensitivity of the test by collecting more data. Give each student one (or more) additional fun-size bag of M&Ms and have them count three more samples for trial 3. Use the class spreadsheet to track the cumulative relative frequency of orange M&Ms as the class sorts more M&Ms. Track the p-values and graph the margin of error to see if at any point the sample proportion crossed the significance level.

15 minutes

Evaluate

Transition to **slide 22.** Let's look deeper into the scatterplot of cumulative relative frequency vs. sample size. In order to graph the margin of error we will need to input our trial into our calculator. Introduce students to the concepts of confidence intervals or show them how to use the classroom graphing technology to graph and calculate the margin of error.

Move to **slide 23** and conclude with the Law of Large Numbers while students record their answers for question #5 of their handouts.

Teacher's Note: Optional Speaking Activity

If you would like students to practice verbally explaining the math they've been working on, consider unhiding **slides 25–26** and introduce students to the Elevator Speech strategy.

Invite students to use the information on the slide and their Color by Number handout to create a script for their Elevator Speech. Each speech should be a maximum of thirty seconds long, which means students should take time to edit their answers for clarity and conciseness. Remind students to connect their thoughts with transition words as if they were talking to someone and that they should practice before filming themselves. The point of the elevator speeches is to get students to answer the question "why does this matter" so maybe try to get students to think about how statistics can help (dis)prove statements and what effects that may have.

You can have students present their findings in class or record themselves and post to your class LMS.

Speech Prompts:

1. State your conclusion regarding the color distribution of M&Ms compared to the 2008 distribution.
2. Does our sample proportion appear to be converging on .2 as we opened more bags of candy?

Resources

- Bognar, M. (2025). Binomial distribution applet/calculator. Department of Statistics and Actuarial Science, University of Iowa. https://homepage.divms.uiowa.edu/~mbognar/applets/bin.html
- K20 Center. (n.d.). Elbow partners. Strategies. https://learn.k20center.ou.edu/strategy/116
- K20 Center. (n.d.). Elevator speech. Strategies. https://learn.k20center.ou.edu/strategy/57
- Wicklin , R. (2017, February 20). The distribution of colors for plain M&M candies. The DO Loop. https://blogs.sas.com/content/iml/2017/02/20/proportion-of-colors-mandms.html