**M&M’s®**

M&M® candies have been around since 1941. The original candy colors were red, yellow, brown, green, and violet. Today, the milk chocolate M&M® colors are brown, yellow, red, green, blue, and orange, with blue being the most recent color to be added.

Let’s say that you are interested in finding out the distribution of the color proportions that are produced. If you were to try to search the M&M’s® website for the population proportions of these colors, it is quite hard to find. Instead, you decide to use a statistical estimation technique that will help you get an estimate for the population proportions. In particular, since blue is the ‘newest’ color to be added, you are interested in estimating the population proportion of blue M&M’s®.

1. Write down the research question of interest.

Instructions: You will be given a fun-sized M&M® bag.

1. Count the number of blue M&M’s® in that fun-sized bag and calculate the proportion that are blue.

**Discuss the Following Questions**

1. Based on the data collected from your fun-sized M&M® bag, provide a point estimate to offer an answer to the research question.
2. Consider other fun-sized M&M® bag carried out using the same methodology, but using a different sample of candies. Would you necessarily obtain the same answer to the research question as you did in question 1? Explain why or why not.

Record your point estimate for the proportion of blue M&M’s® in your fun-sized bag on the board.

1. Are all of the proportions identical? Are they similar?
2. Based on your answer in question 4, is this what you expected?
3. Describe the shape, center, and variability of the sample proportions of blue M&M’s®.

**CONSTRUCTING A BOOTSTRAP CONFIDENCE INTERVAL**

1. How can you physically construct one bootstrap sample?
2. What would you record as your bootstrap statistic?

One group will volunteer to name the M&M® colors from their fun-sized bag.

Write down on slips of paper the color of each of those candies. Place the slips of paper into a bag.

Using the process you described in question 8, carry out 5 bootstrap samples, each time recording the bootstrap statistic that you described in question 9.

|  |  |
| --- | --- |
| Bootstrap Sample Number | Bootstrap Statistic: |
| 1 |  |
| 2 |  |
| 3 |  |
| 4 |  |
| 5 |  |

Record your proportions on the dotplot on the board.

1. Sketch the plot below.
2. Where is this distribution centered? Explain why it makes sense that the distribution is centered at this value.

To obtain an estimate of the population proportion of blue M&M’s®, you will be using a technique called the bootstrap percentile interval technique. If a 95% bootstrap confidence interval is of interest, then you can just chop off the lowest 2.5% and the highest 2.5% of the bootstrap statistics to produce the interval.

1. Use the bootstrap percentile interval approach to find the middle 95% of the bootstrap distribution.
2. Provide an answer to the research question that you stated in question 1 using the interval estimate you found in question 12.
3. Why should you obtain an interval estimate instead of using the point estimate to answer the research question?
4. Describe the process you used to create the bootstrap confidence interval.