MA256 Lesson 3 - Significance - Strength of Evidence (1.1-1.2)

Defintions review

Statistical Significance

Statistical Significance indicates the strength of evidence that the observed result was unlikely to have occurred by random chance

What is the 3S Strategy?

Statistic, Simulate, Strength of Evidence

What is the difference between a **parameter** and a **statistic**?

A parameter is the long-run behavior of a random process and it is not observable. A statistic is the summary of observed results and is used to infer about the parameter.

Define:

H_0 :

Null Hypothesis, the *extit*Random Chance model for the long-run behavior of the parameter.

H_a : Alternate Hypothesis, the *extit*Something More model for the long-run behavior of the parameter.

π : The proportion parameter, the long-run behavior of a random process. It is unobservable.

$\hat{p}$: The observed proportion statistic from data, used to argue or infer about the parameter}

n : The sample size; The number of observations within each sample to construct the statistic

p-value :

The probability of observing an observation at least as extreme as } $\hat{p}$ assuming the null hypothesis is true.}

Classify the strength of evidence for each range of p-values.

0.1	$< p$			Weak Evidence against the null
0.05	$< p \leq$	0.1		Moderate Evidence against the null
0.01	$< p \leq$	0.05		Strong Evidence against the null
	$p \leq$	0.01		Very Strong Evidence against the null

Chips Ahoy vs Keebler Original Chocolate Chip Cookies

Chips Ahoy! conducted a blind taste-test experiment with 50 randomly selected consumers and recorded whether they preferred Chips Ahoy! over Keebler's original chocolate chip cookies. Of those 50 consumers, 45 (90%) of them preferred Chips Ahoy! Chips Ahoy! now claims that 90% of consumers prefer their cookies over Keebler's.

Step 1: Ask a Research Question: Restate from the given information.

Do people prefer Chips Ahoy! over Keebler's chocolate chip cookies?

Step 2: Design a Study & Collect Data: Explain what type of study was conducted.

(single) blind study with 50 randomly selected consumers. Each participant was asked if they preferred Chips Ahoy! or Keebler's

Step 3: Explore the Data: What could we do with the data to help us understand it better?

(single) blind study with 50 randomly selected consumers. Each participant was asked if they preferred Chips Ahoy! or Keebler's

Step 4: Draw Inferences: Significance (today's lesson!)

To draw inferences for this or any other research question, we must set the framework by answering the following questions. Do this at boards in teams of 2-3:

Q1) What are the observational units?

The randomly selected consumers who participated in the blind taste test experiment.

Q2) What is the sample size and number of samples?

The sample size (n) = 50. There was only 1 sample of 50 consumers.

Q3) Identify & classify the variable collected.

The variable is whether a consumer preferred Chips Ahoy! It's a binary categorical variable (Yes or No)

Q4) What is the symbol, definition & value of the sample statistic?

$\hat{p}$ (p-hat) represents the short-run proportion of consumers from the experiment who preferred Chips Ahoy's original chocolate chip cookies over Keebler's original chocolate chip cookies.

NOTE: We can ALWAYS calculate this value from the data using R if not given. $\hat{p} = 45/50 = 90\%$ *or* 0.9.

Q5) What is the symbol, definition, & value of the population parameter?

π represents the long-run proportion of consumers who prefer Chips Ahoy's original chocolate chip cookies over Keebler's original chocolate chip cookies.

NOTE: This is an intuitive random by-chance value based on the context of the problem. For example, if people are just randomly choosing between two cookies, the long-run probability of them choosing the Chips Ahoy! cookie is 1 out of 2, or 50%, so $\pi = 1/2 = 50\%$ or 0.5.

Q6) What are the null & alternate hypotheses in words and symbol notation?

Null Hypothesis: The long run proportion of consumers who prefer Chips Ahoys original chocolate chip cookie over Keeblers original chocolate chip cookie is 50%.
$H_0 : \pi = 0.5$

Alternate Hypothesis: The long run proportion of consumers who prefer Chips Ahoy's original chocolate chip cookie over Keebler's original chocolate chip cookie is greater than 50%.
$H_a : \pi > 0.5$

Q7) With the framework established, we can use the 3S strategy and *One Proportion* applet (http://www.isi-stats.com/isi2nd/ISIapplets2021.html) to estimate the likelihood that Chips Ahoy got the result they did if consumers could not taste a difference and were randomly choosing one of the two cookies.

1. **Statistic:** Write out the statistic: $\hat{p} = 45/50 = 90\%$ *or* 0.9

2. **Simulate:** Run the simulation using the *One Proportion* applet.

Getting the Null Distribution:

Probability of choosing Chips Ahoy! (success/heads) = π (Parameter) = 0.5

Number of tosses = n (Sample size) = 50

Number of repetitions = 1,000

Select "Draw Samples" button.

3. **Strength of Evidence:** How common/uncommon is our observation? Report the p-value. Select the "Count samples" "Count" button to calculate the P-value for this simulation. (what number should you use for this value?)

We can see subjectively that our statistic (red vertical line) is very uncommon, so there is very strong evidence that our statistic is extremely unlikely to have occurred by random chance. the reported p-value should be around 0.

Step 4: Draw Inferences:

We are extremely confident that the long run proportion of consumers who prefer Chips Ahoy! original chocolate chip cookies over Keebler's original chocolate chip cookies is greater than 50% but how much greater do people prefer Chips Ahoy over Keebler? In lesson 8, we will calculate confidence intervals to estimate the range of values within which the true long run proportion lies based on 90%, 95% or 99% confidence.

Step 5: Formulate Conclusions: Can you generalize to other brands? Can you identify any bias in the study?

Consumers seem to prefer Chips Ahoy's original chocolate chip cookies over Keebler's original chocolate chip cookies based on this single experiment, but can we generalize to the entire Chips Ahoy brand versus the entire Keebler brand? We will learn more about generalization in lessons 6 & 7, but to answer this question – we should not be comfortable generalizing the results to the entire brand. It's possible that Keebler's Chewy or Chocolate Chunk cookies are better than Chips Ahoy's.

\textcolor{blue}{Is the cookie the reason people preferred one over the other, or could there have been bias in the research? Examples of intentional or unintentional bias could be: 1. If researchers lead the experiment by saying, "Folks seem to prefer cookie #1 (Chips Ahoy) over #2 (Keebler), what do you think?", then consumers may be more likely to group-think and prefer cookie #1. 2. If cookie #1 is served on a different/better plate than cookie #2, it may subconsciously influence people to prefer that cookie (because of the plate). 3. Telling people "Cookie #1" vs "Cookie #2" could influence people to select cookie #1, just because we typically think of #1 as first place or the best. 4. Perhaps after taking a bite of the first cookie, people's palate isn't as sensitive to the flavors. Did researchers clean the palate before consumers tasted the 2nd cookie? If not, perhaps everyone thinks the 1st cookie tasted is better due to palate sensitivity to flavors. We will explore what conditions allow us to determine cause and effect in lesson 09.}

Step 6: Look Back and Ahead: Identify limitations & suggest ways to overcome in new studies.

Structure the experiment as double-blind, and/or in additional ways to avoid bias examples above. Since Chips Ahoy! has something to gain from the results of the experiment, they are obviously biased. A neutral 3rd party should repeat the experiment to see if they get similar results. In fact, let's do that now.

...