Friend or Foe?

Adapted from Allan Rossman and Beth Chance

Introduction

Objectives

In recent classes, we have laid the foundations for using data to draw conclusions about the world:

1. The principles of **experimental design**: we need to conduct carefully designed experiments to be able to draw conclusions about causal relationships between explanatory and response variables.
2. The **sampling distribution** is the distribution of values we could obtain for a sample statistic, across all possible samples from the population of a given size.

Today we will put these ideas together with results from a published experiment to get a first look at how we can quantify the strength of evidence that an experiment provides about a causal association between two variables.

Experiment description

Do children who are less than a year old recognize the difference between nice, friendly behavior as opposed to mean, unhelpful behavior? Do they make choices based on such behavior? In a study reported in the November 2007 issue of Nature, researchers investigated whether infants take into account an individual's actions towards others in evaluating that individual as appealing or aversive (Hamlin, Wynn, and Bloom, 2007).

In one component of the study, 10-month-old infants were shown a "climber" character (a piece of wood with "googly" eyes glued onto it) that could not make it up a hill in two tries. Then they were alternately shown two scenarios for the climber's next try, one where the climber was pushed to the top of the hill by another character ("friend") and one where the climber was pushed back down the hill by another character ("foe"). The infant was alternately shown these two scenarios several times. Then the child was presented with both pieces of wood (representing the friend and the foe) and asked to pick one to play with. Videos of this study are available at websites for the UBC Center for Infant Cognition Lab (http://cic.psych.ubc.ca/example-stimuli/) and the Yale Infant Cognition Center (https://campuspress.yale.edu/infantlab/).

Section 1: Thinking about the study

a) Identify the observational units and the variables in this study. Is each variable categorical or quantitative? We don't have a specific data frame available to look at - you'll have to identify these based on the study description above.

b) The Methodology section states that for the 10-month-olds, the climber was a yellow triangle; helper and hinderer were a red square and a blue circle. Which of the square and the circle was the helper and which was the hinderer was determined randomly for each baby, with half of the babies assigned to each treatment. Also randomized were which event (helping or hindering) they observed first and the positions of helper and hinderer when presented to the infants to play with (on left or right). Why are these important considerations?

c) The researchers found that 14 of the 16 infants in the study selected the nice toy. Is this a sample statistic or a population parameter?

d) If infants have no preference, how many of the 16 would you have expected to select the nice toy? Would you always expect to see exactly that many of the 16 infants select the nice toy?

e) In your judgment, how many infants, out of the 16, would have to select the nice toy in order for you to be fairly well convinced that the researchers' conjecture is correct, that infants really do have a tendency to prefer the nice toy?

Stop here! We'll do Section 2 together as a class.

Section 2: Specifying Hypotheses

f) What is the population parameter being examined in this study? What symbol will we use for this parameter?

g) What is our sample statistic? What symbol will we use for this statistic?

h) The null hypothesis is a statement about the value of the population parameter that expresses the idea that "nothing interesting is going on." What is the null hypothesis in this example, in words and in terms of the symbol we chose in part f?

i) The alternative hypothesis is a statement about the value of the population parameter that expresses the idea that "something interesting is going on." What is the alternative hypothesis in this example, in words and in terms of the symbol we chose in part f?

Section 3: A Simulation Study (by hand, then automated)

The key question here is what results could reasonably occur under the assumption that infants actually have no preference. Rephrased, we'd like to understand what the **sampling distribution** of our sample statistic would be if the null hypothesis was true.

This will let us see how often we would observe results like the ones we got in this study if the null hypothesis was true, and therefore determine whether or not the data we observed are consistent with the null hypothesis.

One way to get at the sampling distribution of our test statistic if the null hypothesis is true is by simulating (artificially re-creating) the selection process of 16 infants over and over, assuming that infants actually have no genuine preference.

j) Flip a coin 16 times. Record the number of heads that you obtain, which represents how of 16 hypothetical infants choose the nice toy if they are picking at random without any preference for the helpful toy.

k) Combine your result with your classmates by putting a dot on the plot above the number you obtained in part k.

l) Where is the distribution of number of heads in 16 tosses centered? Explain why this makes sense.

m) Looking at this dotplot, does it seem that the result obtained by the researchers would have been surprising if in fact the infants had no preference? What does this suggest about whether the researchers' result provides much evidence that the infants do genuinely prefer the nice toy?

Automating the Simulations using R

We really need to simulate this random selection process hundreds, preferably thousands of times. This would be very tedious and time-consuming with coins, so we'll turn to technology. I have provided code to this for you on Gryd (in Lab 05). Go ahead and run that code, then come back here to answer the questions below.

Section 4: Interpreting the Results

n) Describe the shape of the histogram you generated in R, and comment on whether it is centered where you expected.

o) Based on your simulation results, would you say that it would be very surprising, if infants actually have no genuine preference, that 14 out of 16 infants in the study would have chosen the nice toy just by chance? Explain.

p) Report how many of your 1000 repetitions (under the condition that the null hypothesis is true) produced 14 or more infants choosing the helpful toy. Also determine the proportion of these 1000 repetitions that produced such an extreme result.

The proportion you calculated in part p is called an approximate p-value. A p-value is the probability of obtaining a result as extreme as the one observed, assuming that there is no genuine preference/difference. A small p-value casts doubt on the null model/hypothesis used to perform the calculation (in this case, that infants have no genuine preference).

- A p-value of .10 or less is generally considered to be some evidence against the null model/hypothesis.
- A p-value of .05 or less is generally considered to be fairly strong evidence against the null model/hypothesis.
- A p-value of .01 or less is generally considered to be very strong evidence against the null model/hypothesis.
- A p-value of .001 or less is generally considered to be extremely strong evidence against the null model/hypothesis.

The quantity you calculated in part q is an "approximate" p-value because you obtained it by performing a random simulation rather than an exact probability calculation.

q) Would you say that the experimental data obtained by the researchers provide strong evidence that infants in general have a genuine preference for the friend toy over the foe toy?