LABORATORY 2

Field Study Research: Naturalistic Observations of Bumper Stickers (Bumperstickerology)

Purpose

- to introduce the concepts of **operational definitions**, **hypothesis testing, mutually exclusive categories**, **exhaustive categories**, and **sampling schedules**;
- to give you experience in collecting data using naturalistic observations (in this case, of bumper stickers);
- to give you practice **testing the significance of a proportion**;
- to illustrate how to test a hypothesis about the relative prevalence of a phenomenon, both empirically and statistically.

Introduction

This lab introduces *scientific inquiry* by means of a study of bumper stickers (*bumperstickerology*). As an area of investigation, bumperstickerology has no scientific literature (that I know of), which means you do not need to read articles describing other people's *hypotheses* (plural of "*hypothesis*") about bumper stickers and who uses them. However, you probably already have hypotheses about bumper stickers. In this lab, you will test two hypotheses. Specifically, you will use observations of bumper stickers to test hypotheses about answers to the following 2 questions:

Question 1. What is the prevalence of vehicles with bumper stickers?
Question 2. What is the prevalence of different kinds of stickers?

These questions are *empirical* questions, meaning that they can be answered more or less directly from observations. As you will learn, even a simple empirical answer involves assumptions about the relation of an observed sample to an unobserved population.

Getting Started

Definition of a bumper sticker

Before you begin collecting observations, you first need to develop an *operational definition* of a bumper sticker. An operational definition is a definition that can used in practice. Often, but not always, it involves a physical operation of some sort (for example, a measurement procedure or a specific activity). More generally, it is a definition that allows us to carry out a measurement or observational task. Consequently, the definition has to be clear about any features that are relevant to taking a measurement or making an observation, because a critical ingredient for doing science is that researchers agree on their basic terms.

For the purposes of this study, we need to agree on the definition of a bumper sticker. Although this may sound like a simple task, several questions come to mind that : Do window decals count? Does a bumper sticker have to be on a bumper? Does size matter? Does content matter? Do parking permits or AAA decals count? What kind of

vehicle can have a bumper sticker?

These questions have no right or wrong answers. Rather, the researchers simply have to come to a consensus about a definition and then agree to use the same definition.

Formulating testable hypotheses

After you have a definition, formulate testable versions of your hypotheses about what you are likely to observe. One hypothesis to test is your prediction about the prevalence of bumper stickers. That is, what proportion (or percentage) of vehicles do you expect to have bumper stickers? Answering this question requires that you have in mind a population of vehicles from which you can take a reasonable sample. After you make your prediction, write down your description of the population of vehicles about which the prediction was made. You should also have reasons for making this prediction.

1. Hypothesis about predicted prevalence:

I predict that a proportion of .____ (or _____ %) of vehicles will have bumper stickers. The population of vehicles for which this is true is: _____
_____.
The reason I predict this is: _____
_____.

The second hypothesis to test is your prediction about what kinds of bumper stickers are relatively common and what kinds are relatively uncommon. In this case, you should also have reasons to justify your hypothesis. State your hypothesis in terms of two kinds of stickers, one more common than the other.

For instance, you might hypothesize that "Sports stickers are more common than political stickers", and give the reason that "many professional teams are popular in the Delaware Valley but few political issues are popular". Or, you might hypothesize that "Sports stickers are more common than humorous stickers" and give the reason that "sports fans like to promote their teams whereas only a few people like to promote funny sayings".

2. Hypothesis about predicted relative prevalence:

I predict that _____ stickers will be more common than _____ stickers. The reason is: _____
_____.

Having stated your hypotheses, you are almost ready to collect data. First, though, you need to make a number of decisions about data collection, which are described in the next section.

Collecting Data

Decisions about data collection

Before making any observations, you need to decide how detailed they will be and how they will be categorized or coded. You also need to decide what sampling schedule you will use.

1) Descriptive detail - How much information will you record about each car observed? You need to record the number of stickers on each car and classify each sticker. You may also want to record details such as the type of car the sticker is on and the age and sex of the driver.

2) Classification scheme - Your classification scheme for stickers may be more or less detailed. For instance, you could have categories such as "sports", "politics", and "miscellaneous". If you wanted more detail, you could divide sport stickers into those supporting professional teams and those supporting high school or college teams, or into stickers for baseball, football, etc. You might separate out other categories from "miscellaneous". One test that your classification system should pass (if it is a good system) is the "Minimal Miscellany Test": the "miscellaneous" category should be one of the smallest categories. Whatever scheme you devise, the categories must be *mutually exclusive*, meaning that a sticker can only be put in one category, and *exhaustive*, meaning that every sticker is put in some category.
Note: Your second hypothesis, about what kinds of stickers are more common and what kinds are less common, determines two categories.

2) Sampling schedule - Determine in advance of recording your data how and where will you make your observations.
Important:
(1) If you plan to observe stickers while driving, recruit an assistant to drive. Do not try to drive and collect data.
(2) Take samples that represent the population you have hypothesized about. E.g., If your population is based on shoppers at a mall, don't limit your sample to vehicles outside a single store or near a single exit.

With these decisions made, you can begin observing in a scientific manner. Collect observations on at least 200 vehicles (not all of which will have bumper stickers). Once your data are collected, you can analyze them and determine answers to Questions 1 and 2.

Data Analysis

Step 1. Calculating the prevalence of bumper stickers

Question 1. What is the prevalence of vehicles with bumper stickers?

Question 1 can be answered in terms of the proportion (or percentage) of cars with stickers (percentage = proportion x 100%). After you answer the first question, state what population, if any, your sample represents. That is, does your proportion reflect the frequency of bumper stickers among all cars in the country, among all cars in this area, or only among all cars for some group of people?

Was your first hypothesis correct? Testing your hypothesis means that you must compare the observed proportion to your predicted proportion and determine if any difference between the two is *statistically significant*. "Statistically significant" means that the difference is greater than can be reasonably expected by chance.

What test is appropriate? In this case, the appropriate test is the *z-test for the significance of a proportion*. In symbols, you must calculate a value of *z*, using the formula:

$$z = (p - P) / SQRT[P * (1 - P) / N]$$

In this formula, p = the observed proportion
P = the predicted proportion
N = the total number of observations
SQRT = the square root function in Excel.

Plug in your numbers and calculate a value for z. z = _____.

Step 2. Interpreting your statistical result

To interpret your answer, you must first determine whether your value of z is likely to happen by chance (assuming events occurring at least 95% of the time are "likely") or unlikely to happen by chance (i.e., assuming events occurring less than 5% of the time are "unlikely"). Determining this requires knowing whether z falls between – 1.96 and +1.96 (which is likely to happen by chance) or falls outside that range (which is not likely to happen by chance, and is therefore a statistically significant outcome).

Once you know whether your *z* value is or is not statistically significant, determine whether or not your hypothesis was correct.

Step 3. Comparing the relative prevalence of particular kinds of stickers

Question 2. What is the prevalence of different kinds of stickers?

Question 2 can be answered either in terms of a simple count or in terms of proportions (or percentages). Again, test your hypothesis to determine the statistical significance of any differences between two kinds of stickers you predicted would be most and least common (that is, you want to know if the differences are probably due to real differences or are just an artifact of the variability of sampling).

In this case, an appropriate test is a *z-test for the significance of the difference between two proportions*. In symbols, you must calculate a value of *z*, using the formula:

$$z = p_1 - p_2 / \text{SQRT} [P_{12} * (1 - P_{12}) / n]$$

In this formula, p_1 = the observed proportion in group 1
p_2 = the observed proportion in group 2
n = the total number of observations
P_{12} = $(p_1 + p_2) / 2$.

A potential complication here has to do with the fact that the hypothesis can be phrased in two ways. One way is to ask about the relative proportion of bumper stickers, *in the sample of all bumper stickers.* The other way is to ask about the relative proportion of bumper stickers, *in the sample of all cars*. For this lab, asking the second question will be easier. To get p_1, first count all the cars that had a sticker of the kind you thought would be more popular. To get p_2, count all the cars that had a sticker of the kind you thought would be less popular. If a car has both kinds of stickers, discard it from the data analysis. The total number of observations, N, equals your original total (e.g., 200), minus any discards. The proportions, p_1 and p_2, equal your counts divided by this total.

Step 4. Interpreting the statistical results

To interpret your answer, follow an argument similar to the one you used for the first test. That is, when you are finished with your calculation, determine whether the *z*-value is or is not statistically significant. In this case, in contrast to your answer in Step 2, a significant z-score will confirm your hypothesis.

Lab Report

Your report of this lab will consist of a title page, no abstract, an introduction, a method section, a results, and a discussion, no references. Follow the Guidelines for Psychology Research Reports (Chapter 1 of this manual) for the structure and syntax of a report. Your report will probably be 3-5 pages long, not including the title page or any tables.

Please remember that this report is your report and is to be written in your own words. **DO NOT REPEAT INFORMATION FROM THESE INSTRUCTIONS! DO NOT USE THE TITLE OF THE LAB AS THE TITLE OF YOUR REPORT.**

The **title** should be descriptive of your purpose in collecting your data.

The **introduction** should state each hypothesi along and provide at least one reason that the hypothesis is a sensible one. The introduction should also indicate how you will test each hypothesis, indicating, for example, the kind of data you will collect.

The **method** should contain subheadings for Definition, Time and Place, Sampling Scheme, and Classification of Stickers. You need to present the definition of a bumper sticker, the time and location for your observations, how you collected your sample, and how you classified stickers.

The **results** section should state the prevalence of cars with stickers and report the statistical test of the first hypothesis, using the APA guidelines for how to report statistical tests. It should describe the kinds of stickers, state which kind was more prevalent and which was less prevalent, and report your test of the second hypothesis, also using the APA guidelines.

The **discussion** section should indicate which, if any, hypothesis was confirmed and which, if any, was not. The discussion section should also review possible reasons for any disconfirmation as well as possible limitations of any confirmation. The reasons can range from problems with the basis for the original hypothesis to problems of sampling.

Bumperstickerology: Basic Data

Vehicle #	Sticker Descriptions	Nmbr Stckr
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		

Bumperstickerology: Basic Data

Vehicle #	Sticker Descriptions	Nmbr Stckr
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		

Bumperstickerology: Basic Data

Vehicle #	Sticker Descriptions	Nmbr Stckr
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		
101		
102		

Bumperstickerology: Basic Data

Vehicle #	Sticker Descriptions	Nmbr Stckr
103		
104		
105		
106		
107		
108		
109		
110		
111		
112		
113		
114		
115		
116		
117		
118		
119		
120		
121		
122		
123		
124		
125		
126		
127		
128		
129		
130		
131		
132		
133		
134		
135		
136		

Bumperstickerology: Basic Data

Vehicle #	Sticker Descriptions	Nmbr Stckr
137		
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		

Bumperstickerology: Basic Data

Vehicle #	Sticker Descriptions	Nmbr Stckr
171		
172		
173		
174		
175		
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		
195		
196		
197		
198		
199		
200		