**Stat 301 – Day 25**

**Simulations for comparing two proportions (Section 3.1)**

**Example 1:** The study described in Investigation 3.1 compared the proportion of teens with “some level of hearing loss” between the 1998-1994 NHANES III survey and the 2005-2006 NHANES study.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **1998-1994** | **2005-2006** | **Total** |
| **Some hearing loss** | 480 | 333 | 813 |
| **No hearing loss** | 2448 | 1438 | 3886 |
| **Total** | 2928 | 1771 | 4669 |

(a) Identify the observational units, explanatory variable, and response variable. Is this an observational study or a randomized experiment? Did the study involve random sampling, random assignment, both, or neither?

(b) Calculate the statistic: the difference in the proportion of teens with some level of hearing loss between the two years.

(c) Define the parameter of interest and state appropriate null and alternative hypotheses.

(j) As before, we will begin our inferential analysis by assuming the null hypothesis is true. What

“random process” are we simulating?/What is the source of the randomness in this study? What do we

need to assume to perform this simulation?

(k) So under the null hypothesis we really only have one value of to estimate – the common

population proportion with hearing loss for these two years. What is your best estimate for from the

sample data?

(d) Open the **Two Population Proportions** applet. Specify .173 for both process probabilities and 2928 and 1771 for the two sample sizes. Press **Draw Samples**. Briefly explain what the simulation is doing and why it is appropriate for this study. Now draw 10000 samples. Record the shape, center, and standard deviation of the distribution of the differences in sample proportions.

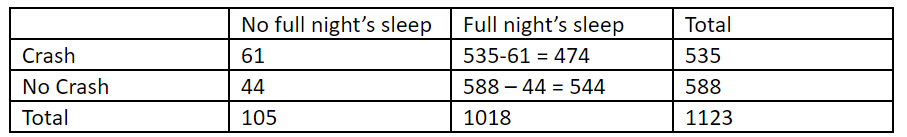
It turns out that there is no “exact” method for calculating the p-value here, because the difference in two binomial variables does not have a binomial, or any other known, probability distribution. However, there is a Central Limit Theorem for the difference in two sample proportions that predicts the distribution will be centered at 1 – 2, with standard deviation (assuming the populations are large compared to the samples) and, if there are at least 5 successes and at least 5 failures in both samples, the shape will be approximately normal.

(e) Should the CLT apply for this study? Does it seem to predict the behavior of the distribution of the difference in sample proportions well (check shape, mean, and SD)?

Note: when conducting a test of significance, we will calculate the standard error using the common estimate. But for a confidence interval, we use the separate and values.

**Quiz 22:** Carry out a two-sample *z*-test and confidence interval (p. 192, for JMP use journal version). Summarize your conclusions from this study. Be sure to address (and support) statistical significance, statistical confidence, and the populations you are willing to generalize the results to. Also, are you willing to conclude that the change in the prevalence of hearing loss is due to the increased use of ear buds among teenagers between 1994 and 2006? Explain why or why not.

**Example 2:** The study described in Investigation 3.11 identified drivers involved in car crashes in New Zealand and a comparable sample of drivers not involved in car crashes, determining whether or not the driver had received at least one full night’s sleep in the past week.



(a) Identify the observational units, explanatory variable, and response variable.

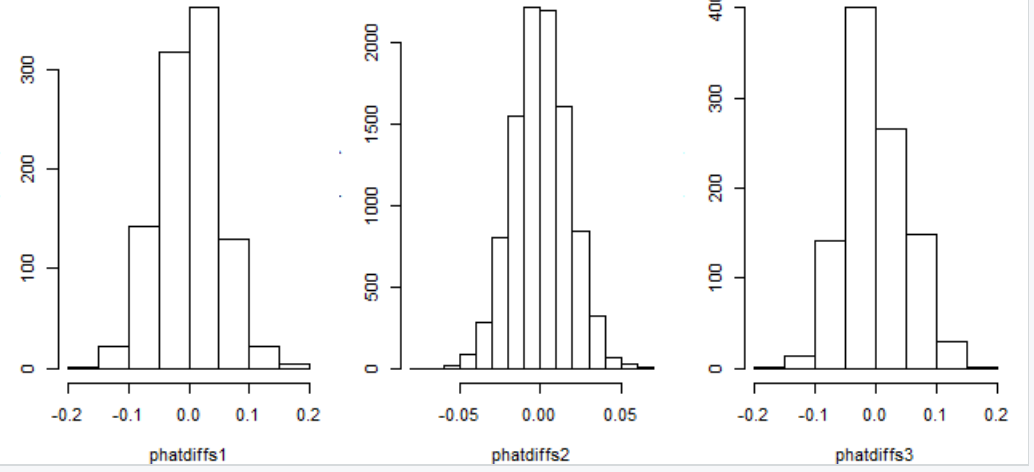
(b) Would you consider this a randomized experiment or an observational study?

(c) What is a little different about how this study was collected? What other study have we looked at that had this property? What are the consequences of this study design?

(d) How would you use the applet to carry out a simulation that mimics this sampling design? Carry out the simulation and record the standard deviation. (Verify these simulations using the R and JMP scripts in HW 6.)

(e) Now use the Two-way Tables [applet](http://www.rossmanchance.com/applets/ChiSqShuffle.html) to carry out a randomization test. How does the standard deviation of the null distribution compare? Which has a larger standard deviation and why?

Option for R users: Run lines 15-37. Include the final output. What is the main feature of the distribution of the difference in sample proportions that changes across the three simulations?



ill people respond less thoughtfully in text because it is often used for casual communication, or more thoughtfully because there is less time pressure to respond? Will they respond less honestly because they aren’t hearing or speaking directly to a human interviewer, or more honestly because they feel less inhibited without spoken contact? Will the lasting visual record of text messages, which others might see, make people answer more honestly because they feel accountable, or less honestly because they feel embarrassed?

be interviewed on an iPhone

and they randomly assigned subjects to either be surveyed through a phone call or through a text. One question that was asked was whether they exercise less than once per week in a typical week (an example of a question in which a yes would be considered socially undesirable). The researchers were able to determine whether or not the respondents answered honestly. Suppose the results are as follows.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Text** | **Call** | **Total** |
| **Yes** | 8 | 4 | 12 |
| **No** | 23 | 28 | 51 |
| **Total** | 31 | 32 | 63 |

Tuesday

Fisher’s Exact Test (really best for experiments) – Yawning (3.7)

Follow-up 3.8

Wednesday (lab)

Simulation – independent random samples (3.1, teen hearing loss)

Normal approximation

Thursday

Relative Risk (3.9)

Log transformation

But sometimes don’t even want response

Case control study (3.10)/Odds ratio (3.11)- smoking or crashes

Monday Review

Tuesday Exam

HW

Fisher’s Exact Test – normal approximation?

Adjustments to normal approximation

could contrast sampling methods

Old lab about this?

Also which one condition on?

Relative risk vs. odds ratio

Still need to consider

Normal approximation

Independent random samples vs. Randomized experiment

Case control study

Fisher’s Exact Test

When surveys are administered, it is hoped that the respondents give accurate and honest answers. American researchers investigated whether the mode of survey delivery affected the accuracy and honesty of responses (Schober et al., 2015). They had 634 people agree to be interviewed on an iPhone and they were randomly assigned to either be surveyed through a phone call or through a text. One question that was asked was whether they exercise less than once per week in a typical week (an example of a question in which a yes would be considered socially undesirable). Did the respondents answer honestly? The results of the respondents’ answers to this question are shown in the table.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Text** | **Call** | **Total** |
| **Yes** | 80 | 41 | 121 |
| **No** | 235 | 278 | 513 |
| **Total** | 315 | 319 | 634 |

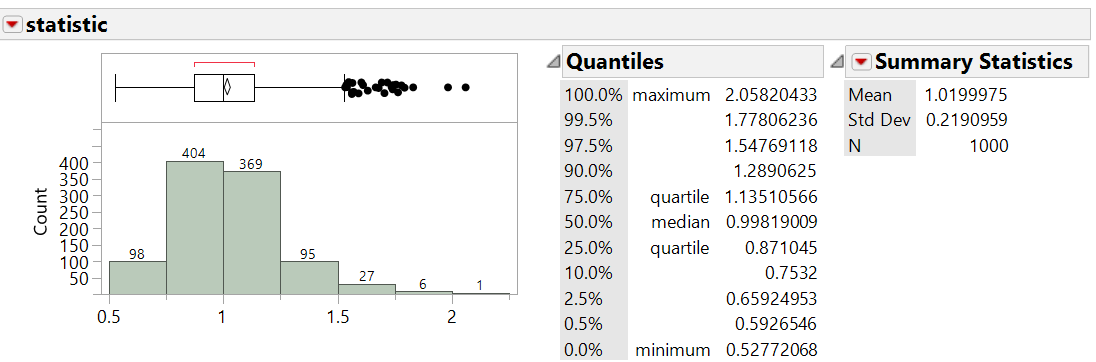
Are metal bands used for tagging harmful to penguins? Researchers Saraux et al. (2011) reported in *Nature* the results of an investigation done to answer this question. A sample of 100 penguins near Antarctica had already been tagged with RFID chips, and the researchers randomly assigned 50 of them to receive a metal band on their flippers in addition to the RFID chip. The other 50 penguins did not receive a metal band. Researchers found that 16 of the banded penguins were still living 4.5 years into the study, and 31 of the un-banded penguins were still living. We are interested in whether metal bands have an effect on whether or not the penguins are living after 4.5 years.

Does the color of a sign that asks someone to do something make a difference in a person’s obedience to that sign? This is what a student researcher investigated. There are double doors at the entrance of her college’s library. She put a sign in the door on the right as you exit the library. The sign read, USE OTHER DOOR with an arrow pointing to the door on the left. She counted students as they exited the library to see what proportion used the door on the left when the sign was printed on red paper and when it was printed on yellow paper. The results are shown in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Yellow** | **Red** | **Total** |
| **Left Door** | 14 | 24 | 38 |
| **Right Door** | 40 | 29 | 69 |
| **Total** | 54 | 53 | 107 |

A study published in the *Journal of Personality and Social Psychology* (Butler & Baumeister, 1998), investigated a conjecture that having an observer with a vested interest would decrease subjects’ performance on a skill-based task (Think about a fan watching Tom Brady through a football throw a tire to win the fan one million dollars.). Subjects were given time to practice playing a video game that required them to navigate an obstacle course as quickly as possible. They were then told to play the game one final time with an observer present. Subjects were randomly assigned to one of two groups. One group (A) was told that the participant and observer would each win $3 if the participant beat a certain threshold time, and the other group (B) was told only that the participant would win the prize if the threshold were beaten. The threshold was chosen to be a time that they beat in 30% of their practice turns. The following results are very similar to those found in the experiment: 3 of the 12 subjects in group A beat the threshold, and 8 of 12 subjects in group B achieved success.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Group A** | **Group B** | **Total** |
| Beat threshold | 3 | 8 | 11 |
| **Didn’t beat threshold** | 9 | 4 | 13 |
| **Total** | 12 | 12 | 24 |



> #Simulation 2

> phatC = rbinom(10000, 535, .0935)/535

> phatNC = rbinom(10000, 588, .0935)/588

> phatdiffs2 = phatC - phatNC

> oddsratio = phatC\*535/(535-phatC\*535)/(phatNC\*588)\*(588-phatNC\*588)

> mean(oddsratio); sd(oddsratio)

[1] 1.020025

[1] 0.2153523