

Worksheet 8

Author:

Discussants:

=====

The following exercises will give you practice running hypothesis tests for a single proportion. Please answers the following questions and submit a compiled pdf with your answers to Moodle by 11:59pm on Sunday November 4th. Function you will find useful include `rbinom()`, `barplot()` and `sample()`. Some useful symbols are: H_0 , H_a , μ , π , \neq

Part 1: Lock5 questions to understand hypothesis tests

Lock5 exercise 4.18 (first edition): Guilty Verdicts in Court Cases A reporter on cnn.com stated in July 2010 that 95% of all court cases that go to trial result in a guilty verdict. To test the accuracy of this claim, we collected a random sample of 2000 court cases that went to trial and record the proportion that resulted in a guilty verdict.

- What is/are the relevant parameters(s)? What sample statistic(s) is/are used to conduct the test?
- State the null and alternative hypotheses in words and using symbols.
- We assess evidence by considering how likely our sample results are when H_0 is true. What does that mean in this case?

Answers:

-
- The hypotheses(in words and using symbols) are:
-

Inspired by Lock5 Questions: Suppose we wanted to test whether the average body temperature of a beaver was greater than 37.6 degrees Celsius. Please state what the population parameter of interest would is, and state the null and alternative hypotheses using both words and symbols.

Answers:

Lock5 exercise 4.46 (first edition): Figure 4.15 shows a null distribution for testing $H_0 : \pi = 0.3$ vs $H_a : \pi < 0.3$. In each case, use the distribution to decide which value is closer to the p-value for the observed sample proportion.

- The p-value for $\hat{p} = 0.25$ is closest to: 0.001 or 0.30?
- The p-value for $\hat{p} = 0.15$ is closest to: 0.04 or 0.40?
- The p-value for $\hat{p} = 0.35$ is closest to: 0.30 or 0.70?

Answers:

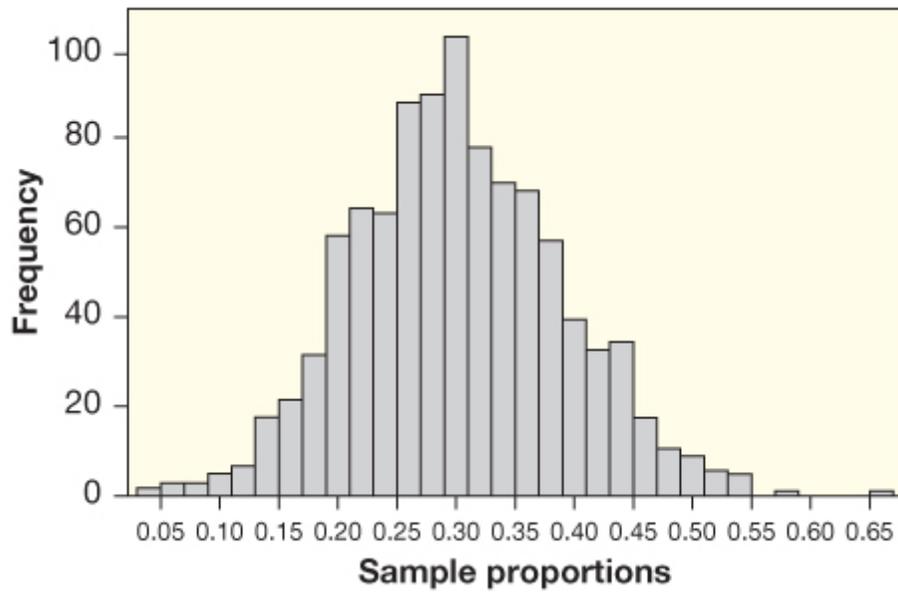


Figure 1: Figure 4.15

- a)
- b)
- c)

Lock5 exercise 4.54 (first edition): Dogs and Owners To investigate whether dogs resemble their owners, Roy and Christenfeld conducted a study testing people's ability to pair a dog with its owner. Pictures were taken of 25 owners and their purebred dogs, selected at random from dog parks. Study participants were shown a picture of an owner together with pictures of two dogs (the owner's dog and another random dog from the study) and asked to choose which dog most resembled the owner. Of the 25 owners, 16 were paired with the correct dog.

To examine whether it is likely that one would get 16 out of 25 owner-dog pairs matched correctly by chance, a null distribution was created by simulating 10,000 times the number of correct dog-owner matches that would occur by chance out of 25 guesses. The results are shown in the table below.

correct matches	4	5	6	7	8	9	10	11	12
frequency	1	17	54	148	341	599	972	1302	1549
correct matches	13	14	15	16	17	18	19	20	21
frequency	1551	1334	997	612	322	142	51	14	4

We are testing $H_0 : \pi = 0.5$ vs. $H_a : \pi > 0.5$ (evidence of dog-owner resemblance). Please answer the following questions:

- Use the data in the table to verify that the p-value for the observed statistic of 16 correct matches is 0.1145
- Use the data to calculate a p-value for an observed statistic of 20 correct matches.
- Use the data to calculate a p-value for an observed statistic of 14 correct matches.
- Which of the three p-values in parts a) to c) gives the strongest evidence against H_0
- If any of the p-values in parts a) to c) indicate statistical significance, which ones would it be?

Answers:

Note, the book using p to denote the population parameter while we are using the symbol π (since all population parameters should be Greek!).

-
-
-
-
-

Lock5 exercise 4.58: Colonoscopy, Anyone? A colonoscopy is a screening test for colon cancer, recommended as a routine test for adults over age 50. A new study provides the best evidence yet that this test saves lives. The proportion of people with colon polyps expected to die from colon cancer is 0.01. A sample of 2602 people who had polyps removed during a colonoscopy were followed for 20 years, and 22 of them died from colon cancer. Does this provide evidence that the proportion of people who die from colon cancer after having polyps removed in a colonoscopy is significantly less than the expected proportion (without a colonoscopy) of 0.01?

- What are the null and alternative hypotheses?

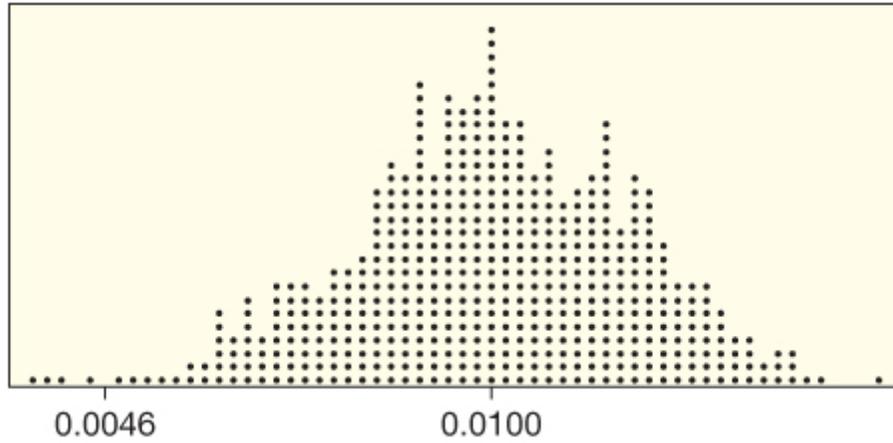


Figure 2: Figure 4.19

- b) What is the sample proportion?
- c) Below a plot of a null distribution for this test. Use the fact that there are 1000 points in this distribution to find the p-value. Explain your reasoning.

Answers:

- a)
- b)
- c)



Figure 3: Paul

Part 2: Running your own hypothesis tests in R

For the second set of exercises we will examine the soccer prediction record of Paul the Octopus to try to determine if he is psychic. To test Paul's psychic abilities, before each soccer game, two containers of food (mussels) were lowered into the octopus's tank. The containers were identical, except for country flags of the opposing teams, one on each container. Whichever container Paul opened was deemed his predicted winner. In the 2010 World Cup, Paul (in a German aquarium) became famous for correctly predicting 11 out of 13 soccer games. We will use hypothesis testing to determine the probability that Paul would get 11 out of 13 correct if he was merely guessing.

Exercise 2.1: Start by writing down the 5 steps for hypothesis testing. You can look at the lecture slides if you need to, although you should memorize these steps sometime soon.

- 1)
- 2)
- 3)
- 4)
- 5)

Exercise 2.2: Now state the null and alternative hypotheses both in words and in the appropriate symbols.

Exercise 2.3: Compute the statistic of interest and save it in the variable `paul_stat`. Do you think this statistic is likely to be obtained if Paul was guessing?

Answer:

Exercise 2.4: Now use the `rbinom()` function to generate a null distribution that would occur if Paul was guessing, and save the results in a variable called `null_distribution`. Remember that the arguments to the `rbinom(num_sims, size, prob)` are: `num_sims` is the number of simulations to run, `size` is the number of “coin flips” in each simulation, and `prob` is the probability of getting heads on each coin flip. Plot the number of times a simulation returned `k` heads as a histogram.

Exercise 2.5: Now use the variables `null_distribution` and `paul_stat` to calculate the number of simulations that had as many or more “heads” than as Paul’s correct soccer prediction answers. Convert this to a p-value by dividing by the total number of simulations. Does this p-value provide evidence that Paul is psychic?

Answer:

Exercise 2.6: Make a judgement call as to whether you believe Paul is psychic based on the p-value and any other information you think is relevant. Make sure to justify your answer to explain Paul’s prediction abilities.

Answer:

Exercise 2.7: Calculate the confidence interval for Paul’s prediction abilities using the bootstrap method

```
# a vector of Paul's answers whether they were correct or  
# incorrect (not necessarily in the order of his predictions)  
pauls_answers <- c(rep("correct", 11), "incorrect", "incorrect") # 11 correct, 2 incorrect  
  
# continue creating the bootstrap distribution from here
```

Answer:

Reflection

How did this worksheet go? If there are any concepts you are not clear about please try to review them and/or meet with the TAs on Tuesday. We will be continuing with other types of hypothesis tests through the next few weeks so the quicker you can nail down your understanding of the concepts the easier the rest of the material in the class will be.