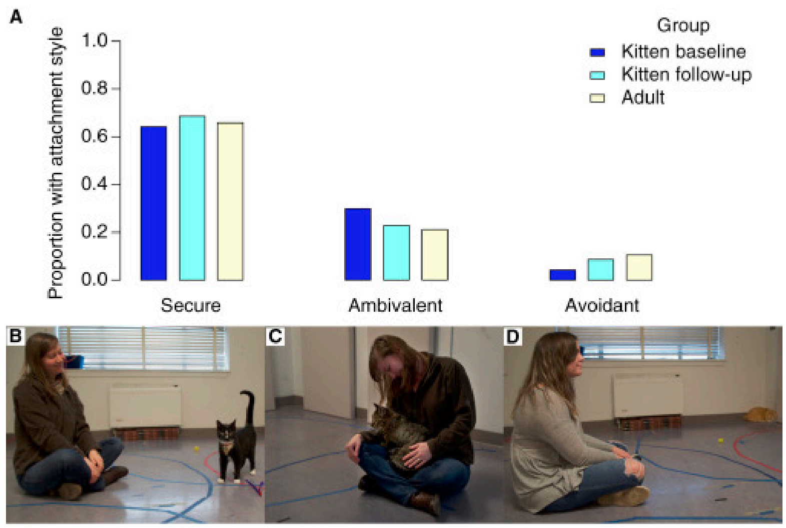
Guide: Hypothesis Testing for population proportion

Based off of p1 & 2 of this worksheet: Week 7 Worksheet



This study from Current Biology studies the types of human attachment that different cats exhibit. For the sake of simplicity, we assume that there are two types of attachment:

- Secure: cats with secure attachment display a reduced stress response and contact-exploration balance with the caretaker
- Non-secure: do not exhibit secure attachment

Let's say that you're interested in doing some hypothesis tests on p, the % of cats that are securely attached. 🐱

A. Your cat loving friend thinks that p is right around 80%. Construct a test to test their null hypothesis against the alternate hypothesis that it is not around 80%.

$H_0 : p = 0.8$ $\quad$ $H_1 : p \neq 0.80$ This is called a two-sided test (see part D).

B. Your other friend thinks the percentage of people who like cats is less than 80%. Construct a test to test using their hypothesis as the alternative hypothesis.

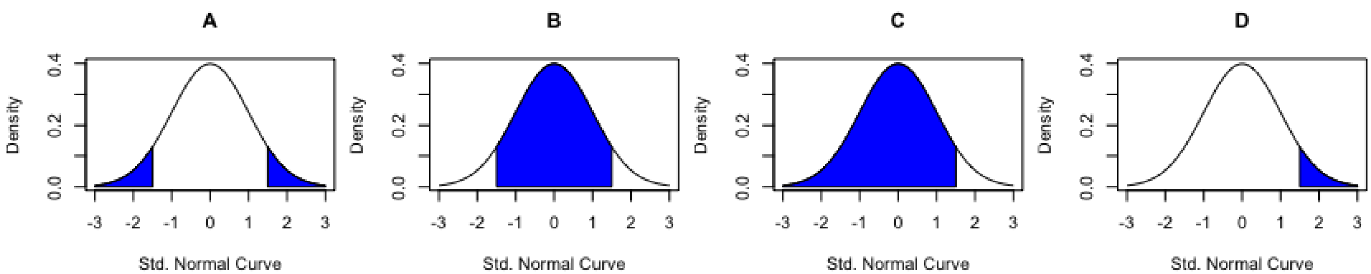
$H_0 : p = 0.8$ $\quad$ $H_1 : p < 0.80$ This is called a one-sided test (see part D).

C. Let's say you observed a sample proportion of Z-value of 1.5. What does this mean? What was the observed sample proportion (assuming n=100)?

This means that you observed a value 1.5 SE's above what you expected above the null hypothesis.   Since the SE here is $\sqrt{\frac{0.2\times 0.8}{100}}=0.04$ (which is calculated assuming the null is true), the observed sample proportion was 0.86.

D. Let's say you observed a sample proportion with a Z-value of 1.5.

Select the areas of the curve that correspond to the p-value of each of the tests in part A & B.



Remember that in hypothesis testing, we are looking for the probability that the difference between what we observe (86%) and what we expect (80% under null) is purely due to chance.

Note that in test A, we don't care if the difference is positive or negative. So the p-value is the probability that we observe an outcome as or more different than what we observed (the difference was 6%).

Thus the p-value is the probability that:

- we see an observed value equal to or greater than 86% (80% + 6%)
- OR we see an observed value equal to or less than 74% (80% - 6%)

So the probability we want is covered by A.

In test B, the alternative only is supported by evidence that p is as or more smaller than what we observed (86%). Thus the p-value is the probability that we see an observed value 86% (Z=1.5) or below. This is represented by C.

Note: histogram D is the p value for H_1: $p > 0.8$.

E. Suppose you have the following data:

Sample size: ***n = 100*** *Sample proportion securely attached 🐱 s:* ***0.73***

Check the normality assumption and calculate p-values for both tests.

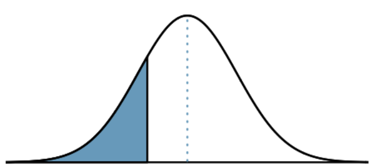
Check: $np(1 - p) = 19.71 \geq 10.$
We can assume that the sample proportion is normally distributed.

Remember that in hypothesis testing, we expect the null to be true. So we use the value assumed under the null hypothesis (p=0.8) to calculate the SE

So.... $SE = \sqrt{\frac{0.8\,(0.2)}{100}} = 0.04.$

Therefore...

$$Z = \frac{Observed\ (sample) - Expected\ Under\ Null}{SE} = \frac{0.73-0.8}{0.04} = \frac{-0.07}{0.04} = -1.75$$



Looking up in the Negative Z-table you get.... **0.0401**

So you need to adjust each of the areas... to get what you want.

The above gives the P-value for test B (one sided):
$p = 0.0401.$

Adjusting for a two-sided case... P-value for test A (two sided): $p \approx 0.08$

E. State your conclusion given significance levels 0.05 and 0.10

Test A: $H_0: p = 0.8$ $H_1: p \neq 0.8$ P-value: 0.08

At $\alpha = 0.05$: fail to reject null hypothesis
At $\alpha = 0.10$: reject null hypothesis in favor of the alternative

Test B: $H_0: p = 0.8$ $H_1: p < 0.8$ P-value: 0.04

At $\alpha = 0.05$: reject null hypothesis in favor of the alternative
At $\alpha = 0.10$: reject null hypothesis in favor of the alternative

Summary: How to calculate p-values for population proportion

Step 1: Calculate the Z value as instructed above.

Follow the approach depending on positive/negative.

Positive Z-value approach:

Step 2: Look up the Z value in the positive Z-table.

Example: Z = 1.75 → 0.9599

Step 3: Determine what you are testing.

A. One sided test with alternate hypothesis, less than some % (Ex: $H_1 : p < 0.8$)
- Then your p-value is 0.9599

B. One sided test with alternate hypothesis, greater than some % (Ex: $H_1 : p > 0.8$)
- Subtract 0.9599 from 1. Ex: 1-0.9599 = 0.0401

What you calculated in B. 

C. Two-sided test alternative (not equal to) $H_1 : p \neq 0.8$
- Subtract 0.9599 from 1. Ex: 1-0.9599 = 0.0401 and multiply by 2

Negative Z-value approach:

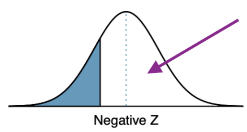
Step 2: Look up the Z value in the negative Z-table.

Example: Z = -1.75 → 0.0401.

Step 3: Determine what you are testing.

D. One sided test with alternate hypothesis, less than some % (Ex: $H_1 : p < 0.8$)
- Then your p-value is 0.0401

E. One sided test with alternate hypothesis, greater than some % (Ex: $H_1 : p > 0.8$)
- Subtract 0.0401 from 1. Ex: 1-0.0401 = 0.9599

What you calculated in E 

.

F. Two-sided test alternative (not equal to) $H_1 : p \neq 0.8$

- Multiply 0.0401 by 2