Stats: The P-value approach to Hypothesis Testing

We've been studying hypothesis testing the last few days, and the approach we've been following is called the Classical Approach. Here's a problem we did recently. I hope most of this is review, so you can read through it quickly.

<u>Example 1</u>: Suppose I told you that 40% of my Intro to Stats students get an A. You don't believe it's that high, so you randomly sample 60 past students and get 18 A's. Test my claim?

Solution: I'll put my comments between < & > signs, so you can distinguish what you should think about from what you should write down.

<We know this is a hypothesis testing problem (because we see the word "test" in there), and we have to decide if it is an average problem or a proportion problem. I hope it is clear to you that we are testing about the proportion being 40% or 0.40.>

HT for p   <We write that down.>

<We need to identify the hypotheses. The two relevant proportion things are "40% is the proportion" vs "you think it is not that high," so that translates into "$p=0.40$" and "$p<0.40$" for our two hypotheses. We remember that the one with the equal sign is the one that becomes the null hypothesis H_o, and the other becomes H_1.>

H_o: $p = 0.4$
H_1: $p < 0.4$

<We need to choose an α-value, which is how much error we are willing to live with, also called the significance level of the test. Sometimes the problem tells you what to use, but if not, then use $\alpha = 0.05$.>

$\alpha = 0.05$

<Now we must identify our test statistic. Our parameter of interest is p, and the best statistic to discuss p is $\hat{p}$. Since $\hat{p}$ turns into a z-value (not t), our test statistic is $z_{\hat{p}}$. We also notice that our test is a 1-tail test (from H_1), and it happens to be the "less than" tail, so now we can form our decision rule.>

DR: If $z_{\hat{p}} <$ ________ then rej H_o.

<We must find a number in the z-chart, called z_C, to fill in that blank. We will do that on the next page.>

<The z_c we need must be the z-score that chops off a lower area of α, so we look in the z-chart for the area 0.0500 (that's α). The chart has 0.0495 and 0.0505, and those areas match with z-scores of -1.64 and -1.65, so you can use either of those. But most commonly, since 0.0500 is halfway between those areas, we go halfway between the z-scores and use z_c = -1.645. Fill that value into the blank, so now ... >

DR: If $z_{\hat{p}} < -1.645$ then rej H_o.

<Now we must calculate $z_{\hat{p}}$, and the formula is $z_{\hat{p}} = \dfrac{\hat{p}-p}{\sqrt{\dfrac{pq}{n}}}$. We observe that

$\hat{p}$ 18/60 = 0.30 . >

$$z_{\hat{p}} = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{0.30-0.40}{\sqrt{\frac{0.40x0.60}{60}}} = \frac{-0.10}{0.0632} = -1.582 \text{ .}$$

<Now we observe, by looking at the DR, that -1.582 is not less than -1.645, so ... >

Fail to reject H_o

<And now, back into English, because we are voting for H_o, ...)

The proportion is 40% .

Here's all you need to show, compressed:

HT for p

H_o: p = 0.4
H_1: p < 0.4

α = 0.05

DR: If $z_{\hat{p}} < -1.645$ then rej H_o.

$$z_{\hat{p}} = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{0.30-0.40}{\sqrt{\frac{0.40x0.60}{60}}} = \frac{-0.10}{0.0632} = -1.582$$

Failr to reject H_o

The proportion is 40% .

Now we're going to make a small but important change to our test. For our purposes, the change will affect only the DR and the statistic we calculate.

Remember that a decision rule divides the number line into two parts: the rejection region (RR) and the acceptance region (AR). On our example, the RR is $z_{\hat{p}} < -1.645$, and the AR is $z_{\hat{p}} \geq -1.645$. Also remember that the cutoff point between the two regions (-1.645) was determined by the α-value.

So, our classical approach is:

Choose α.
Find z_C, the z-score that chops off an α-sized area.
Calculate the z-score for our data.
Observe if the z-score falls below or above z_C, and declare rej or fail to rej H_o.

Our p-value approach will be:

Choose α.
Calculate the z-score for our data.
Find the area into the tail from our calculated z-score.
Compare that area to the α-value.

If that area we calculate is smaller than α, that means our data was farther into the tail than the z_C is, and we should reject H_o. If not, then we fail to reject H_o.

This area from our observed z-score to the tail is called the p-value. When we are doing the p-value approach, our decision rule will ALWAYS be:

If p-value $< \alpha$ then reject H_o.

Or to put it another way, we reject H_o if the p-value is small, but we fail to reject H_o if the p-value is large.

So, in the p-value approach, the DR changes (as I've just described), and the statistic changes (we calculate the area into the tail from our observed z-score, we do NOT calculate z_C).

Some comments:
- If H_1 is a < statement, then our p-value will be the area below our observed z-score, as we observed above.
- If H_1 is a > statement, then our p-value will be the area above our observed z-score, as we will observe next.
- If H_1 is a $\neq$ statement, then we find the area from our z-score to the closest tail, but then we double it because it is a 2-tail test. If it's not clear why this is, you can read about it in your text. But you might just accept it for now.

Also note, when we DO have to double it, if we are using a t-score (instead of a z-score), remember that the t-chart has 1-tail & 2-tail lines, so we will use that. That is, the chart does the doubling for us. For the z-chart, we must double it ourselves.

Example 2. Again using parameter p, this time with an upper tail.
In a study of smokers who tried to quit smoking with nicotine patch therapy, 39 were smoking one year after the treatment, and 32 were not smoking. Use a 5% significance level to test the claim that among smokers who try to quit with nicotine patch therapy, the majority are smoking a year after the treatment. Use a p-value approach.

Solution:

HT for p

H_o: $p \leq 0.50$
H_1: $p > 0.50$

$\alpha = 0.05$

DR: If p-value $< \alpha$ then rej H_o

$\hat{p} = 39/71 = 0.549$, so $z_{\hat{p}} = \dfrac{\hat{p}-p}{\sqrt{\dfrac{pq}{n}}} = \dfrac{0.549-0.50}{\sqrt{\dfrac{0.50x0.50}{71}}} = \dfrac{0.049}{0.059} = 0.83$.

<Our p-value will be the area above 0.83 (above because H_1 points to the upper tail), so we look up 0.83 in the z-chart and find 0.7967. But we want the area above, so...>

p-value = 1 - 0.7967 = 0.2033 .

Conclusion: Fail to reject H_o <Since 0.2033 is > α>

Conclusion: The proportion is 50% or below.
<Or we could say that there is not sufficient evidence that the majority are smoking a year after treatment.>

Next page for another example.

Example 3: This time using parameter μ, and a 2-tail test.
A manufacturer claims that the thickness of the spearmint gum it produces is 7.5 (one-hundredths of an inch). A quality control specialist regularly checks this claim. On one production run, he took a random sample of 10 pieces of gum and measured their thickness, and found a mean of 7.55 (one-hundredths of an inch), with standard deviation of 0.1027 (one-hundredths of an inch). Test the manufacturer's claim, using a p-value.

Solution:

HT for μ

H_o: $\mu = 7.5$ <since we are checking if it's 7.5 or not>
H_1: $\mu \neq 7.5$

$\alpha = 0.05$

DR: If p-value $< \alpha$ then reject H_o.

$\overline{X} = 7.55$, n = 10, and s = 0.1027, so

$$t_{\overline{X}} = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{7.55 - 7.50}{0.1027/\sqrt{10}} = \frac{0.05}{0.0325} = 1.538$$

<Note that n=10 is small, so ... >
Assume pop ND.

<We need to find the 2-tail area beyond t = 1.538. We go to the t-chart, using 9 degrees of freedom. Across the 9 row we observe:

Area	0.01	0.02	0.05	0.10	0.20
df=9	3.250	2.821	2.262	1.833	1.383

We look across the 9 row to see where 1.538 would be, and we see it would be between the last two numbers, 1.833 & 1.383. Those scores match with areas of 0.10 & 020, so that means our p-value is between 0.10 & 0.20, and we note that.>

p-value is between 0.10 & 0.20.

Since p-value > 0.05, we fail to reject H_o.

The mean is 7.5 (hundredths of an inch)

Some final notes:

- We are now using p for two different things, so let's always use p-value for the p-value, and p for the parameter p.

- The p-value = the area from the observed statistic value to the appropriate tail of the bell curve.
(Doubled in the case of 2-tail on a z-chart, but t-chart takes care of that for you.)

- The p-value is always calculated under the assumption that H_o is true.

So: The p-value is the probability we get data as extreme as we actually got, assuming H_o is true

or: The p-value is the probability we make a mistake by rejecting H_o

So, if the p-value is small, since the chance of making a mistake by rejecting H_o is small, we should go ahead and reject H_o. But if the p-value is large, then the chance of making a mistake by rejecting H_o is large, so we should not reject H_o.

Example 4:
Suppose α = .05, and p-value = .02. What should our conclusion be?
(Answer on the top of the next page.)

Example 5:
Suppose α = .05, and p-value = .23. What should our conclusion be?
(Answer on the top of the next page.)

So:

Small p-value means Reject H_o
(small chance of making a mistake by doing so)

Large p-value means Fail to Reject H_o
(Too large of a chance of a mistake by doing so)

Answers from previous page

Example 4: Reject H_o; Example 5: Fail to reject H_o

Your homework for this lesson

Please do the three problems assigned on the homework page

Section 8.3 #11

Section 8.4 #13

Section 8.5 #19

This time you are to use the p-value approach.

Options for submitting your homework this time:

A) You can photograph or scan your pages and email them to me.

B) You can type your answers into an email and send them to me. You don't need to show any calculations, and just do the best you can with any notations you want to use. (Like say X-bar or z-p-hat, whatever you find convenient.)