Chapter 10

Homework: #6 and #16

Popper 10 at the end

No essay

Hypothesis Testing

This is a totally prescribed way to come to a decision about competing claims about a population. It is used by the court system all the time. Also with journal articles making claims and contesting them.

The steps:

1. Setting up the two hypotheses.
2. Selecting the appropriate test statistic.
3. Determining the decision rule
4. Taking a random sample (n > 30, please)
5. Evaluating the test statistic
6. Making a decision

We will look at each step in turn and then see some examples.

Now there are two competing hypotheses. Hnought and Halternate.

H_o H_a

Hnought is the prevailing claim. (AKA the null hypothesis) The alternate hypothesis is brought by someone who disagrees with the prevailing claim.

Let's look at the common way to write these:

$H_0: \ \bar{x} = 100$

$H_a: \ \bar{x} > 100 \qquad \bar{x} < 100 \qquad \bar{x} \neq 100$

Now for us, the null hypothesis will ALWAYS be an equality statement!

a cereal mfg says there are more than 100 raisins in a box

$H_0 \quad \bar{m} = 100$ in this class H_0 is always =

The alternate hypothesis comes in 3 flavors: "I disagree"

$H_a \quad \bar{m} < 100$ $>$ or $\neq$

Example: a cereal maker claims there are more than 100 raisins in their box of cereal. A consumer groups says this is exaggerated.

$H_0 \quad \bar{x} = 100$

$H_a \quad \bar{x} < 100$

TS

$\downarrow$

Next up is the test statistic. In real life there are MANY. For us, now, only z-score because we'll only be working with sample means and sample proportions. These two descriptive statistics have a normal distribution! (Central Limit Theorem).

Let's write them out: x_0

$$z = \frac{\bar{x} = H_0\#}{\frac{s}{\sqrt{n}}} \qquad z = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}\hat{q}}{n}}}$$

Step 3: deciding on a decision rule. We need to decide exactly how big a z-score needs to be for the alternate hypothesis to prevail. Sometimes there are industry standards. Sometimes it's a negotiation.

The standards are $\alpha = 10\%$ $\alpha = 5\%$ $\alpha = 1\%$

What if we can't reject the null hypothesis

H_0 $\bar{x} = 100$ $\alpha = 10\%$

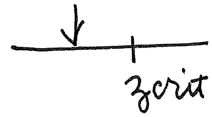
H_a $\bar{x} < 100$ zcrit 1.645

zscore 1.4

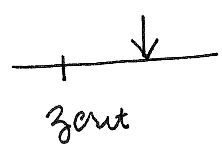
nope those people w/. a problem aren't making their case

What about the three flavors of the alternate hypothesis?

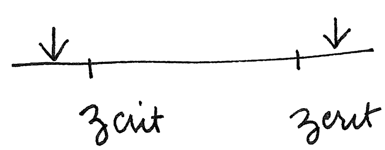
H_a $<$



H_a $>$



H_a $\neq$



$\alpha/2$

4. Collect a random sample. Well, this is the basis for everything inferential, isn't it?

5. Use the sample mean or sample proportion to evaluate the z-score

   get a number for z- score

6. Make a decision based on the previously agreed upon decision rule.

   which side of z crit is it on?

Let's look closer at #6.

Now a four part chart sums up what's happening

Across the top: the hypothesis (Hnought) is true or false (in reality)
Down the side: Reject; do not reject

Ho	T	F
reject Ho	reject Ho when its True I E	rej Ho when it's false ✓
do not reject H	do not rej Ho when it's true ✓	do not rej Ho when its F II E

Note that we are trying to minimize errors. A Type 1 error (reject Ho when it is really true) and a Type 2 error (don't reject it when it is actually false). These are linked! Minimizing one generally maximizes the other! The other two outcomes are correct decisions!

Let's look at a courtroom for a non-numerical example

Ho the defendant is innocent
Ha the defendant is guilty

The defendant is convicted or not

Set up the 2x2 table. Which error does our system want to minimize!

	innocent	guilty
convicted	convicted when innocent I	convicted when guilty ✓
not guilty	innocent & found "not guilty"	guilty & found "not" II

Now let's look at some examples.

Building specifications in Houston require that residential sewer pipe have a minimum mean breaking strength of 2400 pounds per foot (ppf). A contractor has been having problems with pipe bought from a particular manufacturer and this contractor thinks the pipe does not meet the minimum standard. In an attempt to substantiate this feeling, the contractor hires a testing lab to test a random sample of 55 sections of pipe and finds the following: $n > 30$ ✓

$\bar{x} = 2340\,ppf$

$s = 200\,ppf$

Is there sufficient evidence to conclude that the contractor is correct?

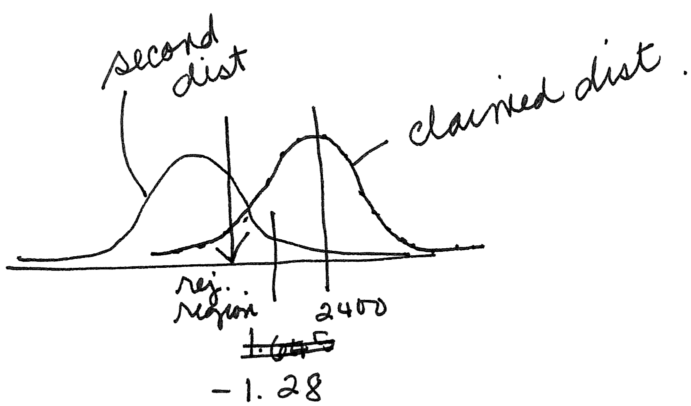
Use an alpha of 10%...ie a confidence level of 90%

Note that n = 55 is enough to use the idea that the distribution of SAMPLE means is normally distributed with an adjusted standard deviation.

Ho the mean = 2400 ppf $\mu = 2400$

Ha nope, it's less than 2400 ppf $\mu < 2400$

Picture:



rej Ho / cannot rej Ho

So our critical z-score is z = −1.28 This defines our rejection region! If our test statistic is below this critical value we'll reject Ho and go with Ha.

Calculate our test statistic. NOTE the adjusted sd:

$$z = \frac{2340 - 2400}{\frac{200}{\sqrt{55}}} = -2.22$$

WOW decision time!

What do we decide? $-2.22 < -1.28$

What do we do next?

make the mfg make it better the contractor wins!

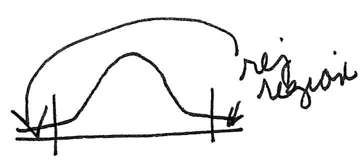
Give up in despair? Mediation? Court? Nope

Each step is totally fixed.

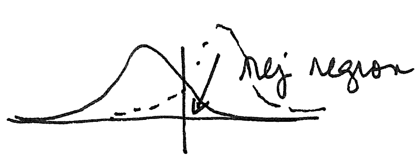
Note different pictures for different Ha's

Not equal alpha of , for example, 15% two tailed

$\neq$



Greater than alpha of 20% one tailed



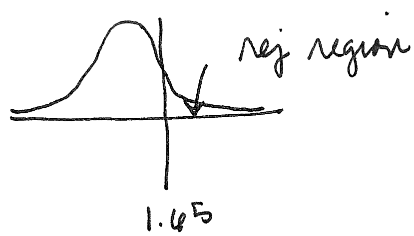
Less than alpha of 5% one tailed

Another example:

A research psychologist will administer a test designed to measure self-confidence to a random sample of 50 professional athletes. The psychologist thinks that professional athletes are more self-confident than the population at large. Since the national average on the test is known to be 72, the psychologist does his testing.

He finds a sample mean of 74.1 and a standard deviation of 13.3. Is he right with an alpha of .05?

What is the picture?



Ho: mean = 72
Ha: mean > 72

Test statistic: $z = \frac{74.1 - 72}{\frac{13.2}{\sqrt{50}}} = 1.12$ Not very unusual, but is it enough?

Well, our alpha is .05. We find that .05 of the area corresponds to a z score cut off of 1.65
Check your chart.

Nope. The TS needed to be HIGHER than 1.65 for him to claim he's right. We do not reject Ho.

This doesn't mean he's totally wrong and too stupid to live. It does mean that the results of this test with this random sample don't support his belief. He won't get a published article out of this research.

Let's review the types of alternate hypotheses and the rejections regions all in one place

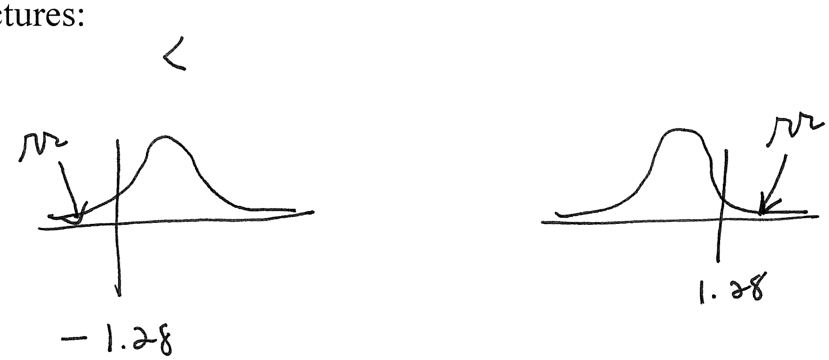
Alpha:

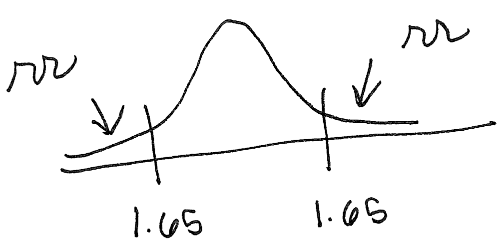
10%

Less than	$z < -1.28$	left-tailed test
Greater than	$z > 1.28$	right-tailed test
Not equal	$z > 1.65$ OR < -1.65	two-sided test

Why the change in z?

Pictures:





whoa! how come not 1.28?

05%

Less than $z < -1.65$

Greater than $z > 1.65$

Not equal $z > 1.96$ OR < -1.96

Pictures:

01%

Less than $z < -2.33$

Greater than $z > 2.33$

Not equal $z > 2.58$ OR < -2.58

Pictures:

ACTIVITIES 10 #3

Another example:

A consumer advocate group thinks that a cereal manufacturer is wrong about the bran content of their cereal. The manufacturer claims that the cereal has 1.2 oz of bran per serving; the advocacy thinks it's less than the claimed amount.

The group selects 60 boxes randomly from grocery stores all over the country and has an analysis done by an outside lab.
The mean is 1.170 and the standard deviation is .111

The group wants an alpha of .05

Who's claim is supported by the evidence?

$H_0 \quad \mu = 1.2$

$H_a \quad \mu < 1.2$

one tailed <

z crit 1.65

T.S. $\quad z = \dfrac{1.17 - 1.2}{\frac{.111}{\sqrt{50}}}$ $\quad \dfrac{1.17 - 1.2}{\frac{.111}{7.07}} = \dfrac{-.03}{.016}$

$= 1.875$

decision rej H_0

court; article in Consumer Review

Now let's look at an example for propotions.

Of 880 randomly selected drivers 56% admitted to running red lights. This information was used to write that "the majority of Americans run red lights".

Is this accurate at an alpha of .05? z crit 1.65

$H_0 \quad p = .50$

$H_a \quad p > .50$ one tailed, >

$$TS \quad Z = \frac{.56 - .5}{\sqrt{\frac{.56 \cdot .44}{880}}} = \frac{.06}{\sqrt{.28}} = \frac{.06}{.529}$$

$Z = .11$

decision – cannot rej. H_0

Popper 10

1. The null hypothesis is always an equality in this class

a T
b F

2. A not-equals test splits the rejection region to an upper tail and a lower tail.

a T
b F

3. You agree on the decision rule BEFORE doing your sample, usually.

a T
b F

4. There are two one-tailed tests; a greater than and a less than.

a T
b F

5. Hypothesis testing is descriptive statistics.

a T
b F