Lecture 23

Today we address the topic of Hypothesis Testing.

A Hypothesis Test is a formal method of testing a claim about the population parameter (or model) based upon a sample of data.

Components of a test

1. Null hypothesis denoted by H_0
2. Alternative hypothesis denoted by H_A
3. Test statistic
4. Pvalue or Level of significance
5. Conclusion.

Null/Alternative hypothesis

An Hypothesis test is designed to assess the strength of evidence against the null hypothesis.

The null hypothesis is usually a statement of "no difference" or "no effect"

The alternative hypothesis is what we hope to prove is true. It should be the opposite to the null hypothesis.

<u>Note</u> H_0 and H_A should always be stated in terms of the population parameter (not the sample statistic)

<u>Examples</u>

	Null	Alternative	
1)	$H_0: \mu = 123$	$H_A: \mu \neq 123$	Sometimes called two-sided alternative
2)	$H_0: \mu \leq 123$	$H_A: \mu > 123$	Sometimes called one sided alternative
3)	$H_0: \mu \geq 123$	$H_A: \mu < 123$	

<u>Test statistics</u>

The test statistic is a quantity we compute based upon our sample data. We use it to judge whether ~~the~~ null or alternative hypothesis is more likely.

P-values

The p-value is the probability, if we assume that H_0 is true, that the test statistic would take the value observed or something more extreme.

In practise this means the smaller the p-value the more evidence we have against H_0.

Some example P-values

P-value	Evidence against H_0
$< .1$	very weak evidence
$< .05$	weak evidence
$< .01$	evidence
$< .001$	strong evidence
$< .0005$	Very strong evidence.

You could think of it this way

H_0: not guilty

H_A: guilty

Then the smaller you make α the less likely it is that you convict an innocent person

Drawing a Conclusion

If the p-value is smaller than same value α we say that the data are statistically significant at level α

For this class unless otherwise stated assume that you are using a 5% level of significance (ie $\alpha = .05$).

So if p-value $\leq .05$ say "I reject H_0 in favor of H_A"

if p-value $> .05$ say "I can not reject H_0"

<u>Testing for population mean</u> (assuming σ is known)

Typical null hypothesis

$H_0: \mu = \mu_0$ (μ_0 is a constant chosen before test)

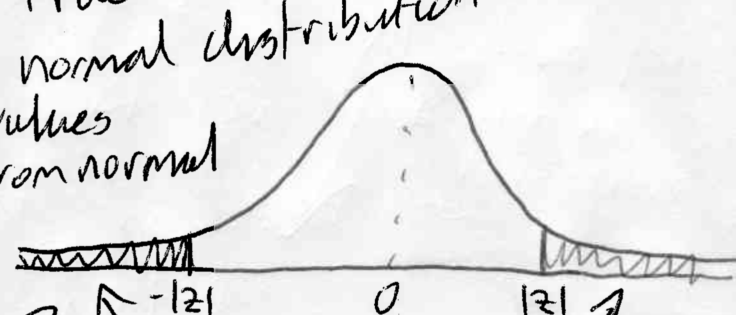
$H_A: \mu \neq \mu_0$

Test statistic

Assuming that H_0 is true this statistic has normal distribution and

$$z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

so p-values come from normal



p-value is $2P(Z \geq |z|)$ ie this area

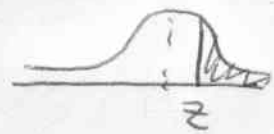
If the alternative changes then so does the ~~...~~ region under the curve used to compute the p-value. The test statistic remains the same. eg

if $H_0: \mu \geq \mu_0 \quad H_A: \mu < \mu_0$

p-value is $P(Z < z)$

if $H_0: \mu \leq \mu_0 \quad H_A: \mu > \mu_0$

p-value is $P(Z > z)$



Example

Suppose that a researcher is interested in discovering whether students a SFSU are smarter than the general populace. The researcher uses IQ scores to assess whether people are smarter or not. In the general population the mean IQ score is 100 and the standard deviation is 15.

The researcher takes a SRS of size 25 from all SFSU students and finds a sample mean IQ score of 107? Is this evidence to conclude that SFSU students are smarter than average?

To answer this problem use a hypothesis test

<u>Step 1</u> State the null and alternative hypothesis

$\mu \equiv$ "mean IQ score for all SFSU students"

$H_0: \mu \leq 100$ $\qquad$ $H_A: \mu > 100$

this means "not smarter than general population

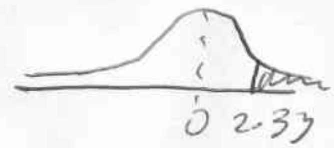
"Smarter than general population"

<u>Step 2</u> Compute value of test statistic

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{107 - 100}{15/\sqrt{25}} = \frac{7}{3} = 2.33$$

Step 3 Compute the P-value

$$\text{P-value} = P(Z > z) = P(Z > 2.33)$$



$$= 1 - P(Z < 2.33)$$
$$= 1 - .9901$$
$$= .0099$$

Step 4 Make a conclusion

Since P-value $= .0099 < .05$ we may reject H_0 and accept H_A. In other words we have statistically significant evidence that SFSU students are more intelligent than the general population

(note that all figures in this example have been invented, it does not represent real life data)