Lecture 14: Hypothesis Testing Part I

Chapter 4.3

Goals for Today

- Introduce Hypothesis Testing Framework
- Testing Hypotheses Using Confidence Intervals
- Types of Errors
- Testing Hypotheses Using p-Values

Statistical Hypothesis Testing

(For now) A hypothesis is a claim about a population parameter.

A hypothesis test is a method for using sample data to decide between two competing hypotheses about the population parameter:

- A null hypothesis H_0.
  i.e. the status quo that is initially assumed to be true, but will be tested.
- An alternative hypothesis H_A.
  i.e. the challenger.

Example

We flip a coin many times and start to suspect that it is biased:

- H_0: the coin is fair. i.e. the probability of heads is $p = 0.5$
- H_A: the coin is not fair. i.e. $p \neq 0.5$

Crucial Concept: Conclusions of Hypothesis Tests

There are two potential outcomes of a hypothesis test. Either we

- reject H_0 in favor of H_A
- fail to reject H_0

Note the difference between accepting H_0 & failing to reject H_0

- "accepting H_0" is saying we are sure H_0 is true
- "failing to reject H_0" is saying something not as strong: we do not have enough evidence to reject H_0.

Analogy: US Criminal Justice System

In the criminal justice system, the jury's verdict does NOT make any statement about the defendant being innocent, rather that there was not enough evidence to prove beyond a reasonable doubt that they were guilty.

Analogy: US Criminal Justice System

Let's compare criminal trials to hypothesis tests:

Truth:

- Truth about the defendant: innocent vs guilty
- Truth about the hypothesis: H_0 or H_A

Decision:

- Verdict: not guilty vs guilty
- Test outcome: "Do not reject H_0" vs "Reject H_0"

Testing Hypotheses Using Confidence Intervals

Example on page 173: The average 10 mile run time for the Cherry Blossom Run in 2006 μ_{2006} was 93.29 min. Researchers suspect μ_{2012} was different:

- H_0: average time was the same. i.e. $\mu_{2012} = 93.29$
- H_A: average time was different. i.e. $\mu_{2012} \neq 93.29$

Testing Hypotheses Using Confidence Intervals

Decision Errors

Hypothesis tests will get things right sometimes and wrong sometimes:

		Test conclusion	
		do not reject H_0	reject H_0 in favor of H_A
Truth	H_0 true	OK	Type I Error
	H_A true	Type II Error	OK

Two kinds of errors:

- Type I Error: a false positive
- Type II Error: a false negative

Decision Errors

- Trade-off between these two error rates
  - procedures with lower type I error rates typically have higher type II error rates
  - vice-versa
- In other words, there is almost never a procedure that makes no type I errors and no type II errors. Some sort of balance between the two is required

Next Time

- More Hypothesis Testing