Lecture 15: Hypothesis Testing Part II

Chapter 4.3

Previously... Statistical Hypothesis Testing

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.

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

- reject H_0
- fail to reject H_0

Previously... 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 (test result)
- Type II Error: a false negative (test result)

Goals for Today

- Tie-in p-Values with sampling distributions
- Revisit sleep example and consider two-sided hypotheses
- Go over the General Hypothesis Testing Procedure

Recall our Coin Example

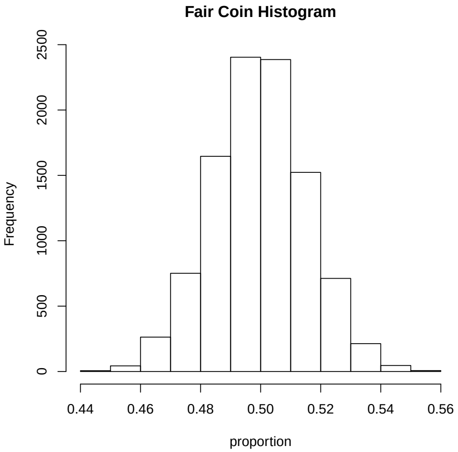
You have a coin that you want to test if it's fair via $n = 1000$ flips. Set $p_0 = 0.5$ and define a "success" as getting heads. i.e.

$$
\begin{aligned}
& H_0 : p = p_0 \text{ i.e. coin is fair} \\
vs \qquad & H_A : p \neq p_0
\end{aligned}
$$

- The point estimate $\widehat{p}$ of p is $\frac{\#\text{ of successes}}{\#\text{ of trials}}$.
- Since it is based on a sample, $\widehat{p}$ has a sampling distribution
- The standard error is $\sqrt{\frac{p(1-p)}{n}}$ (Chapter 6).
- Furthermore, the sampling distribution is Normal in this case (Central Limit Theorem)

Sampling Distribution of $\widehat{p}$

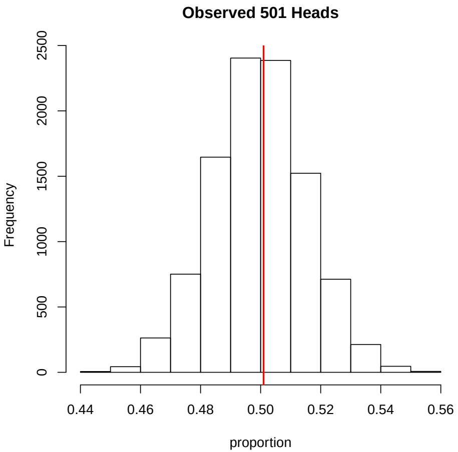
Under the null hypothesis that the coin is fair i.e. $p = p_0 = 0.5$, the sampling distribution of $\widehat{p}$ when $n = 1000$ is:



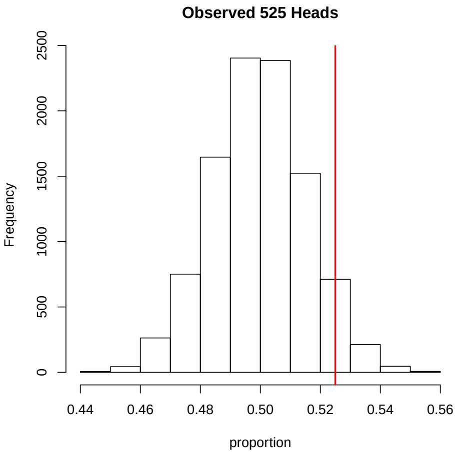
Say we observe...

- 501 heads
- 525 heads
- 900 heads

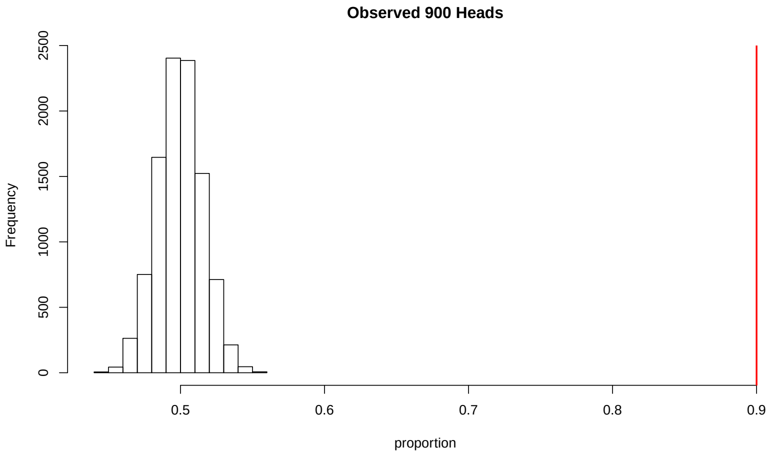
Say we observe...



Say we observe...



Say we observe...



Recall Example about Reedie Sleep Habits

A poll found that college students sleep about 7 hours a night. Researchers suspect that Reedies sleep more.

Let μ be the true population mean # of hours Reedies sleep a night. Then:

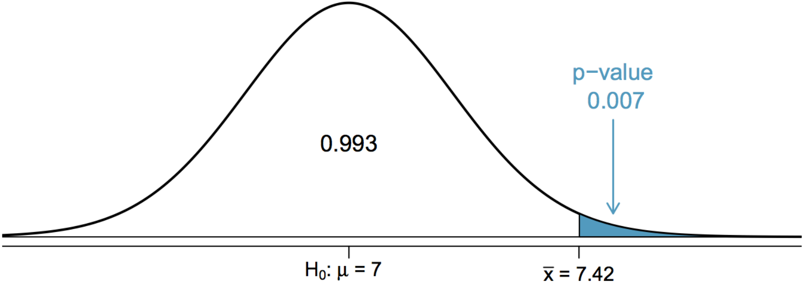
- $H_0 : \mu = \mu_0 = 7$
- $H_A : \mu > 7$

Researchers sample $n = 110$ Reedies and found that $\overline{x} = 7.42$ and $s = 1.75$, hence $z = \frac{7.42-7}{\frac{1.75}{\sqrt{110}}} = 2.47$.

Recall Example about Reedie Sleep Habits

In our case, since $H_A : \mu > 7$, more extreme means to the right of $z = 2.47$.

Hence, the p-value is 0.007:



Two-Sided Alternative Hypothesis

Say instead we had a two-sided alternative hypothesis:

- $H_0 : \mu = 7$
- $H_A : \mu \neq 7$

The the p-value would be double: $2 \times 0.007 = 0.014$

In General: Hypothesis Testing Procedure

1. Construct your hypothesis testing framework:
   - Define H_0, H_A and if applicable a null value.
   - Set your significance level α
2. Verify that the conditions hold
3. Compute your test statistic
4. Compute the p-value
   - Identify the appropriate distribution to compare the test statistic to
   - Depending on H_A, determine what constitutes being more extreme and compute the p-value using the appropriate probability table.
5. If the p-value is $< \alpha$, reject H_0. Otherwise do not.

In Reedie Sleep Example

1. Construct your hypothesis testing framework:
   - Define H_0, H_A and if applicable a null value.
   - Set your significance level α
2. Verify that the conditions hold
3. Compute your test statistic: the z-score of $\overline{x} = 7.42$
4. Compute the p-value
   - Identify the appropriate distribution to compare the test statistic to: normal distribution
   - Depending on H_A, determine what constitutes being more extreme and compute the p-value using the appropriate probability table: z-table on page 409
5. If the p-value is $< \alpha$, reject H_0. Otherwise do not.

Next Time

- How big a sample size to I need?
- Power Calculations