9.1 Significance tests

A significance test is a formal procedure for using observed data to describe 2 competing claims (hypotheses).

The claim we seek evidence against is the null hypothesis, H_0. This is usually the case of "no difference".

Ex: $H_0: p = .8$

The claim we suspect to be true instead of the null hypothesis is called the alternative hypothesis, H_a.

Ex: $H_a: p < .8$

The alternative hypothesis can be <u>one sided</u> if it states the parameter is smaller <u>or</u> larger than the null hypothesis value.

Ex: $H_a: p < .8$ <u>OR</u>

$H_a: p > .8$

The alternative hypothesis is 2 sided if it states the parameter is different from the null hypothesis value.

Ex: $H_a: p \neq .8$

Writing hypotheses

The null hypothesis is always H_0: parameter = value

The alternative hypothesis is H_a: paramater $<$ / $>$ / $\neq$ value

* The parameter is always used, Never statistics.

Ex: Check your understanding, p. 541

1. parameter: actual proportion students who get less 8 hrs of sleep. get

$H_0: p = .85$

$H_a: p \neq .85$

2. parameter: true mean time to complete the survey.

$H_0: \mu = 10$

$H_a: \mu > 10$

Significance tests answer the question:

How likely is it to get a result like this just by chance when the null hypothesis is true.

Activity on p. 542

Without doing a simulation, we could use the binomial distribution

Binary - make or miss
Independent - each shot is independent from each other
Number - exactly 50 trials
Success - each shot has the same chance of a make

binomcdf (n = 50 shots, p = .8, k = 32 successes)

.0062 or .62%

How do we interpret this probability

1. If $H_0: p = .80$ is correct, by chance, we got a very unlikely outcome
2. If $H_a: p < .80$ is correct, and the population proportion is less than .8, so this outcome isn't all that unlikely.

An outcome that would rarely happen if the null hypothesis is true is good evidence the null hypothesis is not true.

The H_0 states a claim that we are seeking evidence against. The probability that measures the strength of this evidence against H_0 is the <u>P-Value</u>.

<u>P-Value</u>: The probability computed, assuming H_0 is true, that the statistic would take a value extreme as (or more extreme as) the one actually observed in the direction of H_a.

- small p-values are good evidence against the null hypothesis because it says the observed value is very unlikely to occur if H_0 is true.

- large p-value is good evidence to not reject the null hypothesis because it says the observed value is likely if H_0 is true.

Statistical Signficance

small Pvalue → Pvalue < α → reject H_o → convincing evidence for H_a

large Pvalue → Pvalue ≥ α → fail to reject H_o → not convincing evidence for H_a

When determing if a Pvalue is small enough to reject H_0, we use a <u>significance level</u> α.
For example if we choose $\alpha = .05$, we require evidence (Pvalue) so strong that it happens less than 5% of the time by chance when H_0 is true.

If the Pvalue is smaller than α we say the results of the study are <u>statistically significant</u> at level α. If so we reject H_0 and accept H_a because we have convincing.

* We express significance by Pvalue. For example, statistical signficance at the level .05 is expressed as: " The results were signficant, ($p < .05$)

How do we choose α?

1. must be done before collecting data.
2. If H_0 is well accepted, we need stronger evidence (small pvalue) so α should be small as well.
3. Practicality - if its time consuming or expensive to make a change, we want stronger evidence

Type I and Type II errors

If we reject H_0 when H_0 its true, we have committed a Type I error.

If we fail to reject H_0 when H_a is true, we have committed Type II error.

		Truth about population	
		H_0 true	H_a true
Conclusion on a sample	reject H_0	Type I	correct conclusion
	fail to reject H_0	correct conclusion	Type II

Ex: A fast food restaurant wants to reduce the proportion of drive thru customers who have to wait longer than 2 min. Based on store records the proportion of customers who had to wait more than 2 min. was $p = .63$. To reduce the proportion, the manager decides to assign an additional employee to assist with orders. During the next month, the manager will collect an SRS of drive thru times.

1) Form hypothesis for this situation

$H_0: p = .63$

$H_a: p < .63$

2) Describe a Type I & type II error and describe the consequences.

Type I: We think the proportion of people who wait longer than 2 minutes has decreased, however the proportion of people waiting 2 min. or more has not decreased.

A consequence is we are paying him even though he is not helping the wait.

Type II: We think the proportion of people waiting longer tha 2 min has stayed the same, when it actually has decreased.

Consequence: We let the employee helping out go which will increase the wait times again.

Significance Test + Type I error

The probability of a Type I error is equal to the significance level.

$$P(\text{type I error}) = \alpha$$