Hypothesis Testing: Power

Example:

You work for an insurance company and want to see if there is a difference between customers from two different advertising channels.

Specifically, you want to determine if there is a difference in the mean policy life between these two groups of customers.

Example:

You take a sample of customers from each channel and find the sample statistics on policy life in days.

	Channel A	Channel B
Mean	177.9 days	149.6 days
Standard Deviation	48.7	39.6
Sample Size	20	20

Example:

You take a sample of customers from each channel and find the sample statistics on policy life in days.

	Channel A	Channel B
Mean	177.9 days	149.6 days
Standard Deviation	48.7	39.6
Sample Size	20	20

p-value = 0.0512

Example:

You take a sample of customers from each channel and find the sample statistics on policy life in days.

	Channel A	Channel B
Mean	177.9 days	149.6 days
Standard Deviation	48.7	39.6
Sample Size	20	20

p-value = 0.0512

Conclusion: There is not enough evidence to conclude that there is a difference in average policy life between the two channels.

Example:

What might have gone wrong with this test?

Example:

What might have gone wrong with this test?

There might actually be a difference, but we concluded that there is not (**Type II Error)**.

Example:

What might have gone wrong with this test?

There might actually be a difference, but we concluded that there is not (**Type II Error)**.

We had a very small sample size (only 20 per group), so our estimates have very large margins of error.

If there actually is a difference, we might not have given ourselves enough of a chance to detect it.

Power

The **power** of a hypothesis test is the probability that we reject the null hypothesis, in the event that a specific alternative hypothesis is true.

It can also be defined as the probability of *not* committing a Type II error, in the event that a specific alternative hypothesis is true.

Power

When running a designed experiment (such as an A/B test), you should be doing power calculations prior to running the experiment to ensure that you have a good chance of avoiding an inconclusive experiment.

To see how to perform power calculations using *statsmodels*, see the corresponding notebook.