Hypothesis Testing - Recap

Introduction

You just learned how to create an experiment and interpret the results! Let's review some of the specific things you have learned.

Key Takeaways

Some of the key takeaways from this section include:

- It's important to have a sound approach to experimental design to be able to determine the significance of your findings
- Start by examining any existing research to see if it can shed light on the problem you're studying
- Start with a clear alternative and null hypothesis for your experiment to "prove"
- It's important to have a thoughtfully selected control group from the same population for your trial to distinguish effect from variations based on population, time or other factors
- Your sample size needs to be selected carefully to ensure your results have a good chance of being statistically significant
- Your results should be reproducible by other people and using different samples from the population
- The p-value for an outcome determines how likely it is that the outcome could occur under the null hypothesis
- α is the marginal threshold at which we're comfortable rejecting the null hypothesis
- An α value of 0.05 is a common choice for many experiments
- Effect size measures just the size of the difference between two groups under observation, whereas statistical significance combines effect size with sample size
- A one sample t-test is used to determine whether a sample comes from a population with a specific mean
- A two sample t-test is used to determine if two population means are equal
- Type 1 errors (false positives) are when we accept an alternative hypothesis which is actually false
- The α that we pick is the likelihood that we will get a type 1 error due to random chance
- Type 2 errors (false negatives) are when we reject an alternative hypothesis which is actually true
- Resampling methods allow for improved precision in estimating sample statistics and validating models by using random subsets
- Common resampling techniques include bootstrapping, jackknifing and permutation tests