Hypothesis Testing and Resampling Methods

Katie Pollard

BMI 206

In this unit we will learn …

- What to do if test statistics do not have an obvious parametric distribution under the null hypothesis
- How to simulate null distributions using permutations and bootstrap methods
- Common challenges with resampling based null distributions
- How to explore genomic distributions of discrete regions such as ChIP-seq peaks or GWAS hits
- How to test for enrichment of genomic features using resampling methods

Components of a Hypothesis Test

1. **Parameter:** quantity of interest

2. **Null and alternative hypotheses:** statement about parameter value

3. **Test statistic:** quantify evidence

4. **Error rate:** control mistakes

5. **Null distribution:** assess significance

6. **Procedure:** decision rule

Null Distribution & P-Value

The probability of obtaining a test statistic as large (or larger) than the one observed under a null distribution (i.e., assuming H_0 is <u>true</u>) is called a p-value.

- The p-value is small if the observed statistic would be very unusual under the null.

- The p-value is a single number that summarizes the evidence for/against H_0 in the data.

- If the sample data is inconsistent with H_0, then the test statistic will be large in magnitude (i.e., in the tail of the null distribution) and the p-value will be small.

Null Distributions

- Parametric (a.k.a. tabled distributions)

  Normal distributions   z-statistics

  Student's t-distribution   t-statistics

  F distribution   F-statistics

  Wilcoxon/Mann-Whitney U   U-statistics

- Non-parametric (i.e., resampling based)

  Permutation (2+ groups or continuous)

  Bootstrap (various types)

Permutations

- Randomize group labels, positions, locations, ...

    - Estimates a distribution that is the pool of the groups (e.g., same mean, same variance, etc)

    - Usually easy to implement

- Some issues to consider

    - What to permute is not always obvious

    - Permuting into regions that cannot be observed

    - Strict null distribution because all parameters are different from the observed data, potentially including parameters other than in null hypothesis

Bootstrap

- Resampling observed data with replacement estimates the variability in the empirical distribution

- Statistics over bootstrap iterations will have a range of values, providing an empirical test statistics distribution

- If this can be adjusted so the null hypothesis holds, it provides a suitable test statistics null distribution

  - Can be easy, e.g., make means the same in each group by computing sample means and subtracting

  - Need to think explicitly about the null hypothesis to make this adjustment to the bootstrap

  - Does not involve changing the labels, positions, etc.

Implementing a permutation test

- Simulate two vectors of numbers (n=10 random normal variables per group).

- Perform a parametric t-test.

- Generate b=100 permutations.

- Compute a t-statistic for each permutation.

- Calculate a permutation p-value.

- Compare parametric and permutation results.

- Repeat for different values of n (possibly unbalanced) and b. Also try different means in the two groups.

Implementing a bootstrap test

- Simulate two vectors of numbers (n=10 random normal variables per group).

- Standardize to have mean zero in each group. Generate b=100 bootstrap samples.

- Compute a t-statistic for each bootstrap.

- Calculate a bootstrap p-value.

- Compare parametric, permutation, and bootstrap results.

- Repeat for different values of n (possibly unbalanced) and b. Also try different means in the two groups.