Module 5E: Unsupervised Learning

A smattering of options: PCA, permutations, bootstrap

Two major categories of computational methods

Null sampling distributions:

1. Simulation – hypothesis testing

2. Randomization/Permutation

Precision of estimates:

3. Bootstrapping – sampling distribution of estimate; the values for the parameter estimates that we might obtain and their probabilities.

Two major categories of computational methods

Null sampling distributions:

1. Simulation – hypothesis testing

Determine the null distribution by simulation of the sampling process

5 main steps

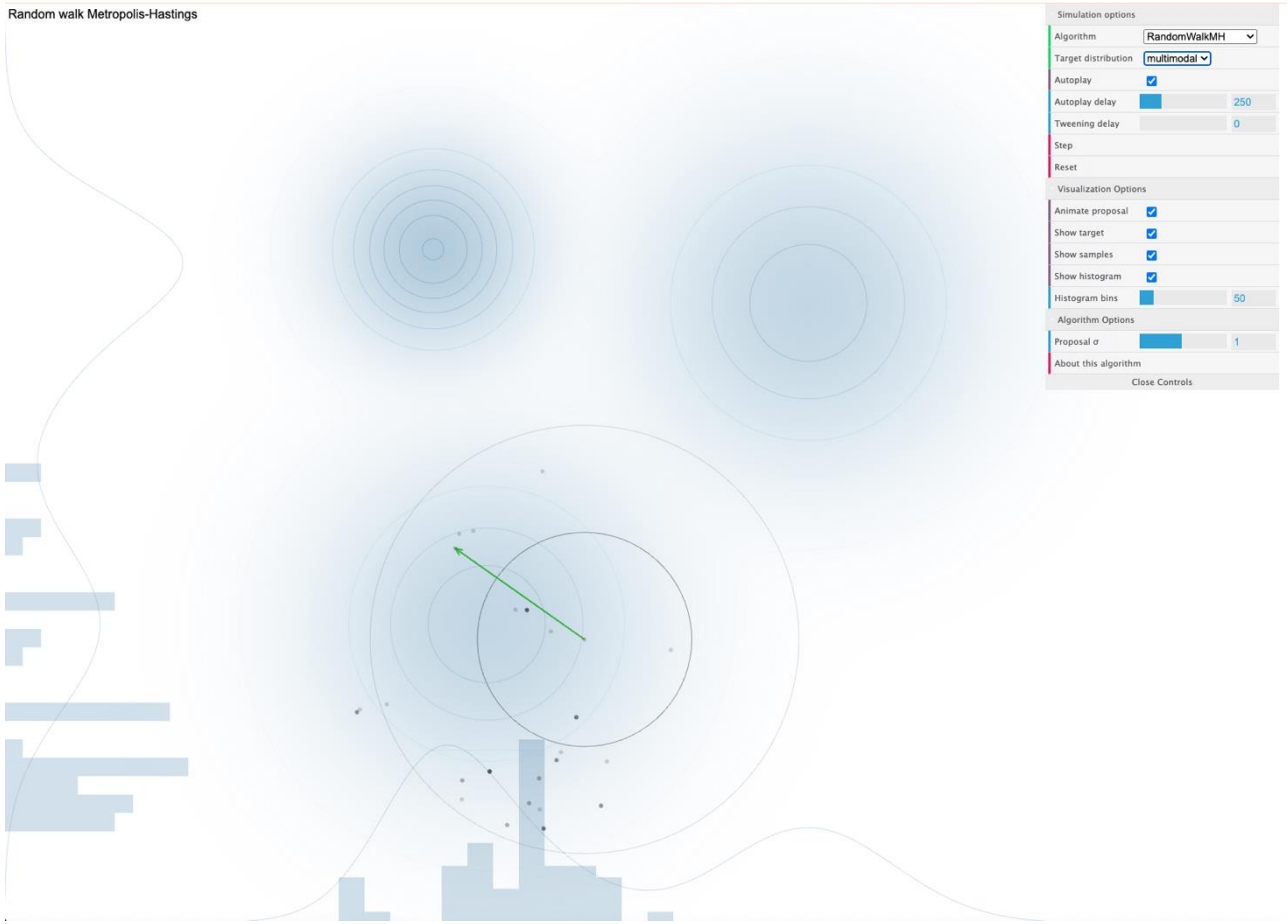
1. **Create and sample imaginary population**
   -parameters specified by null hypothesis
   -Same protocol that was used to collect real data
2. **Calculate test statistic on simulated sample**
3. **Repeat many times**
4. **Form the null distribution**
   - Gather simulated values for the test statistic
5. **Compare test statistic from the actual data to the null distribution**

This is a BROAD topic. Some of these simulations will be relevant: https://chi-feng.github.io/mcmc-demo/
This is an excellent explanation of the simulations: https://elevanth.org/blog/2017/11/28/build-a-better-markov-chain/

Two major categories of computational methods

Null sampling distributions:

1. Simulation – hypothesis testing



https://chi-feng.github.io/mcmc-demo/