Johns Hopkins Engineering

625.464 Computational Statistics

Randomization Methods

Module 8 Lecture 8A

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Randomization Methods

Basic idea: Compare an observed configuration of the outcomes with all possible configurations.

Process 1: $x_1, x_2, \ldots, x_{n_1}$

Process 2: $y_1, y_2, \ldots, y_{n_2}$

null Hypothesis: H_0 = the means are equal

$$t_0 = \bar{x} - \bar{y}$$

Randomization Test

$$t_0 = \bar{x} - \bar{y}$$

$\bar{x}_{new}$ is mean of

$$y_1, x_2, x_3, \cdots, x_{n_1}$$

$\bar{y}_{new}$ is mean of

$$x_1, y_2, y_3, \cdots, y_{n_2}$$

$$t_1 = \bar{x}_{new} - \bar{y}_{new}$$

$\binom{n_1+n_2}{n_2}$ different configurations $\Rightarrow$ $\binom{n_1+n_2}{n_2}$ test statistics

if $\binom{n_1+n_2}{n_2} = m$ compare to other m. It will be kth largest. And we will reject H_0 with sig level k/m

Randomization Example

For each rat i the outcome X_i is measured.

- null hypothesis: outcome does not depend on if the rat was treatment or control
- alternate hypothesis: outcomes for rats labeled treatment will be larger

$$T = \text{mean}(\text{treatment}) - \text{mean}(\text{control})$$

and has value $T = t_0$ for the observed data

calculate $t_2, t_3, \ldots, t_M$ for all M possible permutations and compare to t_0