

A Short Note on Randomization Test

Qiao Zhou¹

All seemingly statistically significant relationships can still be due to chance and subject to sampling error. To further examine the significance of a test result, randomization test can be exploited.

Methodology

Under a statistical experiment, one group (group 1) of samples with sample size n_1 is selected from a total sample space of size n . As a result of selection, n_2 samples are left unselected in group 2, where $n_2 = n - n_1$. One can form a hypothesis that the mean of group 1 is significantly different from group 2. To test the hypothesis,

Firstly, a two-sample t-test is performed between the two independent groups. A t-stat, t_r , is obtained.

Secondly, repeat the following procedure N times,

- Shuffle the data,
- Assign the first n_1 samples to group 1 and the remaining n_2 samples to group 2²
- Calculate the test statistics (denoted t_i) for the reshuffled data

Therefore, we generate an empirical distribution using the t-statistics $t = \{t_1, t_2, \dots, t_N\}$ calculated from the randomized bootstrap data.

The t-statistic from the real data t_r is compared with the distribution of t-statistics from the bootstrap data. We calculate an empirical p-value by computing the proportion of bootstrap statistics that have a more extreme value than the t-statistic from the real data, where

$$p = \frac{\sum_i^N \mathbf{1}_{t_i > t_r}}{N}$$

Interpretation

If the real t-statistic is in the belly of the distribution, then it is indistinguishable from t-statistics generated from randomized data. If the statistic from the real data is towards the edge of the bootstrap distribution, then it is unlikely that the experimental result can have arisen by chance, and we would conclude that the difference in mean is statistically significant.

Implementation

In MATLAB, the test can be performed using the following code:

¹ The author can be contacted via email: qiao_zhou@mfe.berkeley.edu

² This step can be performed with or without replacement.

```
tot = [x1;x2];
n1 = length(x1);
n2 = length(x2);
n = n1 + n2;
[~,pobs,~,stats]=ttest2(x1,x2);
tobs = stats.tstat;
tboot=zeros(b,1);
for i = 1:b
    boot1 = tot(randi(n,n1,1));
    boot2 = tot(randi(n,n2,1));
    [~,~,~,stats]=ttest2(boot1,boot2);
    tboot(i)=stats.tstat; % record the t value
end
p=sum(abs(tboot)>=abs(tobs))./b;
```