Introduction to
Permutation Tests

Peng Liu

3/01/12

1

Assumption for a z-test, t-test or F-test

- When conducting a z-test or a t-test, we are actually assuming that the data (or the random errors) follow a normal distribution.

- Based on this assumption, we know the distribution of the test statistic (T.S.) under the null hypothesis.

- Based on the distribution (z-distribution, t-distribution or F-distribution), we get a p-value for each observed T.S..

- This can be referred to as "parametric approaches".

2

What if the distributional assumption does not hold?

- If the normal assumption does not hold for the data and the sample size is small, the results of z-test, t- or F-test are not reliable.
- What can we do?
  - Transformation of data to make the data normal
  - Choose some tests that do not make such distributional assumptions – "nonparametric approaches"

3

Permutation test

- Permutation tests (randomization tests) can be used without the normal assumption for the distribution of data.

- Permutation test is a nonparametric approach to establish the null distribution of a test statistic.

- Permutation tests are attractive to microarray study because it makes fewer assumptions.

4

Permutation

- Permutation is the rearrangement of objects or symbols into distinguishable sequences.
- Each unique ordering is called a permutation.
- For example, for A, B, C, and D, each possible ordering of all 4 elements without repetitions is one permutation, such as B, C, D, A.

5

Calculation the number of permutations

- Definition: For a positive number n, $n!$ (read n factorial) is the product of all the positive integers less than or equal to n. That is:

  $n! = n \times (n-1) \times (n-2) \times \ldots \times 3 \times 2 \times 1$

  e.g.: $4! = 4 \times 3 \times 2 \times 1 = 24$

- The number of permutations with n objects is $n!$
- The number of permutations with r elements from n objects is $n!/(n-r)!$

6

Combination

- A combination is an un-ordered collection of unique elements.
- For example, for the letters A, B, C and D, we want to get a collection of 2 elements from the 4.
- Note the getting a collection of 2 elements from the 4 in total is equivalent to assigning them into 2 different groups with 2 of them in group 1 and the other 2 in group 2.

7

Calculation the number of combinations

- Definition: For a positive number n and r, $n \geq r$, we define

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

read n choose r.

- The number of combinations (groupings) of size r from a total set of n elements is $\binom{n}{r}$.

8

Idea of permutation test

- Under H_0 (the null hypothesis), some of the data are exchangeable.
- We permute (rearrange) the data by shuffling their labels of treatments, and then calculate our T.S. on each permutation. The collection of T.S. from the permuted data constructs the null distribution.

9

Steps of permutation test

1. Analyze the problem and identify the hypothesis.
2. Choose a T.S. and establish a rejection rule that will distinguish the null from the alternative.
3. Compute the T.S. for the original observations.
4. Rearrange the observations.
   1. Compute T.S. for the rearranged data.
   2. Compare the T.S. from original observation with the ones from re-arranged data.
5. Make conclusion.

10

An Example

- Suppose I have an instrument that measures the mRNA transcript abundance of a certain gene.

- I have developed a drug that I suspect will alter the expression of that gene when the drug is injected into a rat.

- I randomly divide a group of eight rats into two groups of four.

- Each rat in one group is injected with the drug.

- Each rat in the other group is injected with a control substance.

11

Hypothetical Data

I use my instrument to measure the expression of the gene in each rat after treatment and obtain the following results:

	Control	Drug
Expression	9 12 14 17	18 21 23 26
Average	13	22

The difference in averages is 22-13=9.

I wish to claim that this difference was caused by the drug.

12

2

Interpretation of the Results

- Clearly there is some natural variation in expression (not due to treatment) because the expression measures differ among rats within each treatment group.

- Maybe the observed difference (22-13=9) showed up simply because I happened to choose the rats with larger expression for injection with the drug.

- What is the chance of seeing such a large difference in treatment means if the drug has no effect?

13

Rearrangement of data

Random Assignment	Control				Drug				Difference in Averages
1	9	12	14	17	18	21	23	26	9.0
2	9	12	14	18	17	21	23	26	8.5
3	9	12	14	21	17	18	23	26	7.0
4	9	12	14	23	17	18	21	26	6.0
5	9	12	14	26	17	18	21	23	4.5
6	9	12	17	18	14	21	23	26	7.0
7	9	12	17	21	14	18	23	26	5.5
8	9	12	17	23	14	18	21	26	4.5
9	9	12	17	26	14	18	21	23	3.0
10	9	12	18	21	14	17	23	26	5.0
11	9	12	18	23	14	17	21	26	4.0
12	9	12	18	26	14	17	21	23	2.5
13	9	12	21	23	14	17	18	26	2.5
14	9	12	21	26	14	17	18	23	1.0
15	9	12	23	26	14	17	18	21	0.0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
69	18	21	23	26	9	12	14	17	-8.5
70	18	21	23	26	9	12	14	17	-9.0

14

Distribution of Difference between Treatment Means Assuming No Treatment Effect

Number of Random Assignments

Difference in Treatment Means (Control – Drug)
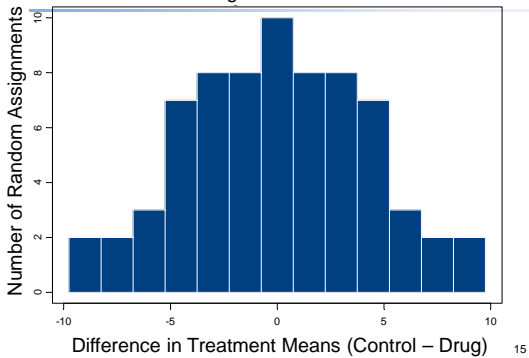
15

Conclusions

- Only 2 of the 70 possible random assignments would have led to a difference between treatment means as large as 9.

- Thus, under the assumption of no drug effect, the chance of seeing a difference as large as we observed was 2/70 = 0.0286.

- Because 0.0286 is a small probability, we have reason to attribute the observed difference to the effect of the drug rather than a coincidence due to the way we assigned our experimental units to treatment groups.

16

Permutation Test

- R.A. Fisher described such randomization (permutation) test in the first half of the 20th century.

- Permutation tests are popular for assessing statistical significance because they do not rely on specific distributional assumptions.

17

Permutation Test

- Permutation tests can be applied to more general experiments than the example here.

- This method depends on identifying units to permute, which should be the units in the experiment that are exchangeable under the null hypothesis, determined by the design of the experiment and the factor(s) being tested.

18

3

Comments about permutation tests

- When the parametric assumption (e.g. normal distribution of data) is correct, the parametric tests (z-, t-, F-tests) are more powerful.
- The permutation test is exact under relatively non-stringent conditions (less assumptions) and protect against deviations from parametric assumptions.
- Permutation test is more computationally intensive.

19

Applying permutation tests into microarray

- If sample size is small, the number of permutations is small.
- As a consequence, the p-values have a problem called "granularity".
- Sometimes, one can pool the permuted data from all genes if assuming that all genes have the same null distribution. Or, one can shuffle the residuals instead of the data.

20

Reference

- Nettleton, D. (2006) A Discussion of statistical methods for design and analysis of microarray experiments for plant scientists. *The Plant Cell*,18, 2112–2121.
- Cui, X and Churchill, G. (2003) Statistical tests for differential expression in cDNA microarray experiments. *Genome Biology*, 4:210
- Good, P. (2000) Permutation Tests, 2nd Edition, Springer
- Significance Analysis of Microarrays (SAM): check http://www-stat.stanford.edu/~tibs/SAM/ for more references and developments.

21