Introduction to Randomization Tests

3/7/2011

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.

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.

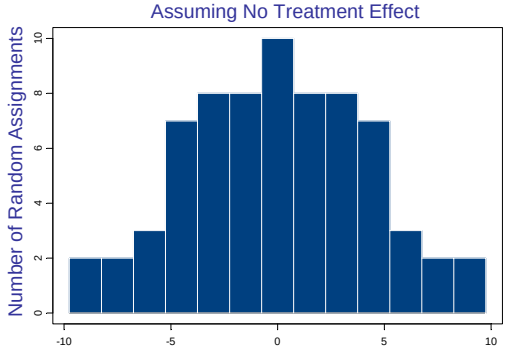
I wish to claim that this difference was caused by the drug.

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?

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

Distribution of Difference between Treatment Means Assuming No Treatment Effect



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.

7

Randomization Test

- This is an example of a randomization test.
- R.A. Fisher described such tests in the first half of the 20th century.
- Randomization tests are closely related to permutation tests (almost synonymous) which are popular for assessing statistical significance because they do not rely on specific distributional assumptions.

8