STAT 403 Tutorial · Week 4

1. Review of Randomization Test

- Suppose there are 6 units (plots) : labelled as 1,2,3,4,5,6;
  2 treatments (fertilizers) : labelled as A,B.

We hope to conduct an experiment to compare two fertilizers' effects over yield and assess whether B is better than A, i.e.

H_0: A & B are the same v.s. H_a: B is better than A.

- Randomization : treatments should be assigned to units randomly.
  (3 A's and 3 B's)

(i). 

: randomly take 3 balls to assign A; for the rest, assign B.

(ii) Using R : " sample (6, 3)" $\curvearrowleft$ number of items to choose.
$\nearrow$ number of items to choose from

- Testing statistic : $D = \bar{Y}_B - \bar{Y}_A$

units:	1	2	3	4	5	6
trts.	A	B	B	A	A	B
yields	11	12	16	13	9	14

$D_{obs} = \frac{1}{3}(12+16+14) - \frac{1}{3}(11+13+9) = 3$

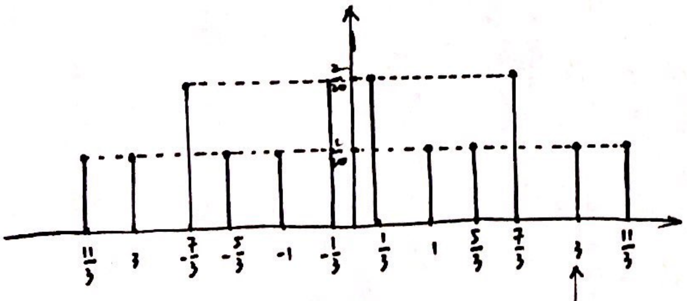
- Since there are $\binom{6}{3} = 20$ (in R: "choose (6,3)") possible treatment assignments, we can calculate 20 D values, one for each assignment.

(See last tutorial notes for reference)

In addition, each of the 20 D values occurs with probability $\frac{1}{20}$.

This leads to a discrete distribution of D under H_0.

D	-3.67	-3	-2.33	-1.67	-1	-0.33	0.33	1	1.67	2.33	3	3.67
Prob.	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{1}{20}$ ↑ obs.	$\frac{1}{20}$



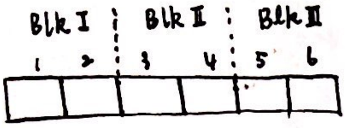
prob. mass function of D

$$\text{p-value} = P(D > D_{obs}) + \frac{1}{2} P(D = D_{obs})$$
$$= \frac{1}{20} + \frac{1}{2} \times \frac{1}{20}$$
$$= 7.5\%$$

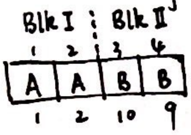
Conclusion: cannot reject null hypothesis at significance level of 5%.

2. paired t test and paired randomization test.

Now suppose we have the same treatments and research goal (assess whether B is batter than A). But the 6 plots (units) are divided into 3 pairs (blocks), each consisting of two plots, because the plots within the same block are considered to be homogeneous (for example, they may have similar soil quality).



- We would like to assign one A and one B to every two units in each block. Why?



~> Because we want to distinguish whether the treatment (fertilizer) or the block factor (soil quality) caused the difference.

- Randomization : (i) Toss a coin for each block, assign A to the left unit if observing heads ;
  (ii) In R, "sample(0:1, 3, replace=True)".

- Suppose below is what observed.

blks:	Blk I		Blk II		Blk III	
units:	1	2	3	4	5	6
trts:	A	B	B	A	A	B.
Y:	4	6	8	7	15	14

Ho: A & B the same

Ha: B is better than A

- paired t test.

Let $D_i = Y_B - Y_A$ in i-th block, then $D_1 = 2$, $D_2 = 1$, $D_3 = -1$.

The testing statistic $T = \dfrac{\bar{D}}{\sqrt{s^2/m}} \overset{H_0}{\sim} t_{m-1}$

$\bar{D} = \frac{1}{m}(D_1 + D_2 + \dots + D_m)$

$s^2 = \frac{1}{m-1}\sum_{i=1}^{m}(D_i - \bar{D})^2$.

Therefore, $T_{obs} = \dfrac{2/3}{\sqrt{7/3 \cdot 1/3}} = 0.53$

p-value = $P(T \geq T_{obs}) = 0.32$ ("1- pt(0.53, 2)")

So basically it's a one-sample t test once Di's have been calculated.

- For each block there are two possible treatment assignments. Since there are 3 blocks, in total there will be $2^3 = 8$ treatment assignments. Each treatment assignment corresponds to a D value, where the testing statistic $D = \bar{D} = \bar{Y}_B - \bar{Y}_A$.

BlkI		BlkII		BlkIII	
1	2	3	4	5	6
A	B	B	A	A	B
4	6	8	7	15	14

$\rightsquigarrow \quad D_{obs} = \frac{1}{3}\left[(6-4) + (8-7) + (14-15)\right] = \frac{2}{3}$

Treatment assignment	$\bar{Y}_A$	$\bar{Y}_B$	D	
A B A B A B	9	$28/3$	$1/3$	$< D_{obs}$
A B A B B A	$26/3$	$28/3$	$2/3$	$= D_{obs}$
A B B A A B	$26/3$	$28/3$	$2/3$	→ observed.
A B B A B A	$25/3$	$29/3$	$4/3$	$> D_{obs}$
B A A B A B	$29/3$	$25/3$	$-4/3$	$< D_{obs}$
B A A B B A	$28/3$	$26/3$	$-2/3$	$< D_{obs}$
B A B A A B	$28/3$	$26/3$	$-2/3$	$< D_{obs}$
B A B A B A	$28/3$	9	$-1/3$	$< D_{obs}$

This leads to a discrete distribution of D under H_0:

D	$-\frac{4}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$4/3$
Prob.	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

Therefore,

$$\begin{aligned} \text{p-value} &= P(D > D_{obs}) + \frac{1}{2}P(D = D_{obs}) \\ &= P(D > \tfrac{2}{3}) + \frac{1}{2}P(D = \tfrac{2}{3}) \\ &= \frac{1}{8} + \frac{1}{2} \times \frac{2}{8} \\ &= 25\%. \end{aligned}$$

Conclusion: cannot reject null hypothesis at significance level of 5%.