STAT371 Discussion #7

I. Paired two samples and unpaired two samples

1. Paired two samples

   Paired two samples are also called dependent two samples. In paired samples, each data point in one sample is matched to a unique data point in the second sample.

   Example – Amount of cancer cells of a group of patient before taking treatment and after taking treatment. In this example, the amount of cancer cells before treatment is matched to the amount of cancer cells after treatment for each patient. You take the difference and use the difference to study the effect of treatment

2. Unpaired two samples

   Unpaired two samples are also called independent two samples. Two samples are independent when the data points in one sample are unrelated to the data points in the second sample.

   Example – Amount of cancer cells of group taking new developed treatment and amount of cancer cells of group taking placebo.

II. How to use R to conduct hypothesis test for two samples data

Ho: muA=muB, Ha: muA is not equal to muB

1. Unpaired two samples

```
x1<- c(43, 57, 35, 50, 38, 61)
y1<- c(51, 95, 90, 49, 101, 74, 84, 46, 75)
```

1.1 t.test: `t.test (x1, y1)`

1.2 Permutation test

1) Step 1: Load the permutation test functions into R

```
source("http://www.biostat.wisc.edu/~kbroman/teaching/stat371/permfunc.R")
```

Step 2: Use perm.test(x1, y1, n.perm=) function for permutation test

Notes: Use n.perm=NULL if you want to do the exact test with all possible permutation. Use n.perm=some number, if you want to use a certain number of permutation to do the permutation test.

2) Permutation test: use all possible permutations

Find observed T: `tobs <- t.test(x1, y1)$statistic`

Permutation test: `tperm1 <- perm.test(x1, y1, n.perm=NULL)`

Calculate the p-value: `mean(abs(tperm1) >= abs(tobs))`

3) Permutation test: 1000 permutations

Find observed T: **tobs <- t.test(x1, y1)$statistic**
Permutation test: **tperm2 <- perm.test(x1, y1, n.perm=1000)**
Calculate the p-value: **mean(abs(tperm2) >= abs(tobs))**

1.3 Rank-sum test: **wilcox.test(x1, y1)**

The p-value will be displayed in the results of rank-sum test

2. Paired two samples

x2<- c(117, 100, 94, 135, 92, 118, 144, 103, 103, 153, 163)
y2<- c(145, 95, 107, 123, 129, 143, 149, 137, 91, 162, 202)
diff<- x2-y2

1.1 t.test: **t.test (diff)**

1.2 Permutation test

1) Permutation test: use all possible permutations

Find observed T: **tobs <- t.test(diff)$statistic**
Permutation test: **tperm1 <- paired.perm.test(diff, n.perm=NULL)**
Calculate the p-value: **mean(abs(tperm1) >= abs(tobs))**

2) Permutation test: 1000 permutations

Find observed T: **tobs <- t.test(diff)$statistic**
Permutation test: **tperm2 <- paired.perm.test(diff, n.perm=1000)**
Calculate the p-value: **mean(abs(tperm2) >= abs(tobs))**

1.3 Signed-rank test: **wilcox.test(diff)**