

Homework 9: Due on Apr. 21

Guideline

- Homework should be submitted via Gradescope by Friday midnight (11:59 pm. CDT).
- Homework answers to Simulations and data analysis should be written in R Markdown. Please use `\newpage` in R Markdown to separate answers to different questions. For simulations, please use `set.seed(4211)` to make your result reproducible.
- Please also upload your executable “xxx.Rmd” file to the corresponding homework assignment on Canvas.
- The total score for the homework is 50 points.

1. Recall the following question from previous homework. Please write a permutation test in R to test the hypothesis that Native American MSCE was larger on average. You should use the difference in group means as the test statistic.

In a genetic inheritance study discussed by Margolin [1988], samples of individuals from several ethnic groups were taken. Blood samples were collected from each individual and several variables were measured. We shall compare the groups labeled “Native American” and “Caucasian” with respect to the variable mean sister chromatid exchange (MSCE). The data is as follows:

Native American: 8.50 9.48 8.65 8.16 8.83 7.76 8.63

Caucasian: 8.27 8.20 8.25 8.14 9.00 8.10 7.20 8.32 7.70

- (a) Draw the histogram of the test statistics in the permuted samples.
 - (b) Decide the p-value of the permutation test.
2. Permutation test can be also used to test the correlation between two continuous variables. In R, “cor(x,y)” function calculates Pearson’s correlation coefficient between x and y.
 - (a). Please use the following program to generate data from H_0 and H_1 .

```
set.seed(4211)
x=rnorm(100,0,4)
y.null=rnorm(100,0,5) # null hypothesis
y.alt=x+rnorm(100,0,3) # alternative hypothesis
## from a location model with normal errors
```

Draw scatter plots of y against x, and report their correlation for both `y.null` and `y.alt`. Please use “cor.test” function to test these correlation.

- (b). Please write a permutation test to test whether the correlation between x and y is 0 or not. Draw the histogram of the correlations in the permuted samples and decide the p-value of the permutation test for both `y.null` and `y.alt`.

3. Nonparametric Bootstrap for the variance of sample correlation. Use the same setting as in Q2.

```
set.seed(4211)
x=rnorm(100,0,4)
y.null=rnorm(100,0,5) # null hypothesis
y.alt=x+rnorm(100,0,3) # alternative hypothesis
## from a location model with normal errors
```

- (a). For each case (`y.null` and `y.alt`), please report the bootstrap variance of sample correlation between `x` and `y`.
- (b). For each case (`y.null` and `y.alt`), please draw the histogram of sample correlations in 1000 Bootstrap samples. Put the two histograms side-by-side and compare them.