

## Biostatistics Homework 1

將完成的作業以 pdf 檔案上傳到 E3.

考慮助教人力，只批改非論述的題目

(論述能力很重要，會在考前檢討)

1. (20 points) Two drugs, amantadine (A) and rimantadine(R), are being studied for use in combating the influenza virus. A single 100-mg dose is administered orally to healthy adults.

The variable studied is  $T_{\max}$ , the time in minutes required to reach maximum plasma concentration. These data obtained: (based on information found in Gordon Douglas, Jr., "Drug Therapy", New England Journal of Medicine, vol.322, February 1990, pp. 443-449)

$T_{\max}(\mathbf{A})$		
105	123	12.4
126	108	134
120	112	130
119	132	130
133	136	142
145	156	170
200		



Amantadine 100 mg (sold p...  
1800petmeds.com · In stock



Rimantadine 100mg Tabl...  
healthwarehouse.com · I...

$T_{\max}(\mathbf{R})$		
221	227	280
261	264	238
250	236	240
230	246	283
253	273	516
256	271	

Please use a statistical software (R or Python) to analyze the data and **also provide your explanations.**

(a) (10 points) Construct a side-by-side box plot and identify outliers.

Please add the main title to the plot and also label the boxes.

(b). (5 points)

b1) Calculate  $\bar{X}$  and  $S^2$  for the data of sets A and R. (5%)

b2) Compare the two samples based on (b1)

(c) (5%) For the set A, assume that the outlier of set A is the result of a misplaced decimal point. Correct the error by deleting the decimal and see what changes this makes in your box plot. Recompute  $\bar{X}$  and  $S^2$ , using the correct data point, and compare your results to those of part b1.

(d) Is there an outlier in set R? If so, is there an obvious legitimate reason to delete it from the data set.

**2.Coevolution. (20 points)** Different varieties of the tropical flower Heliconia are fertilized by different species of hummingbirds. Over time, the lengths of the flowers and the form of the hummingbirds' beaks have evolved to match each other.



Heliconia - Caribaea Red (COMI...



Heliconia Caribaea Phot...



Hummingbird on Heliconia in Guadeloupe ...  
dreamstime.com

Here are data on the lengths in millimeters of two variates of these flowers on the island of Dominica.

#### **H. Caribaea Red**

41.9, 42.0, 41.9, 43.1, 41.5, 41.7, 39.8, 40.6  
39.6, 42.2, 40.7, 37.9, 39.2, 37.4, 38.2, 38.1  
38.1, 38.0, 38.8, 38.2, 38.9, 37.8, 38.0

#### **H. Caribaea Yellow**

36.8, 37.0, 36.5, 36.1, 36.0, 35.5, 38.1, 37.1,  
35.2, 36.8, 36.7, 35.7, 36.0, 34.6, 34.6

(a) **(10%)** Make a back-to-back stem-leaf plot to compare the two samples.  
See the last page for the instruction of installing the related package.

(b) What are the most important differences among the two varieties of flower?

(c) **(10%)** Find Q1, Median and Q3 for each group.

3. (20 points) The following table shows the percent of men and women at least 15 years old who were literate in 2002 in the major Islamic nations. We omitted countries with populations less than 3 million. The data file named “literacy\_rates.csv” is attached.

Literacy rates (percent) in Islamic nations					
Country	Female percent	Male percent	Country	Female percent	Male percent
Algeria	60	78	Morocco	38	68
Bangladesh	31	50	Saudi Arabia	70	84
Egypt	46	68	Syria	63	89
Iran	71	85	Tajikistan	99	100
Jordan	86	96	Tunisia	63	83
Kazakhstan	99	100	Turkey	78	94
Lebanon	82	95	Uzbekistan	99	100
Libya	71	92	Yemen	29	70
Malaysia	85	92			

a. (15 points) Construct side-by-side box plots and list the 5 number summary statistics (min, Q1, median, Q3 and max) for both groups.

b. (5 points) Identify any outliers based on the boxplots. Comment on the outlier, if any.

b. Make the histograms of males and females. (5 points) Compare the shapes of the two distributions.

4. (5 points) A researcher is comparing two multiple-choice tests with different conditions.

In the first test, a typical multiple-choice test is administered. In the second test, alternative choices (i.e. incorrect answers) are randomly assigned to test takers. The results from the

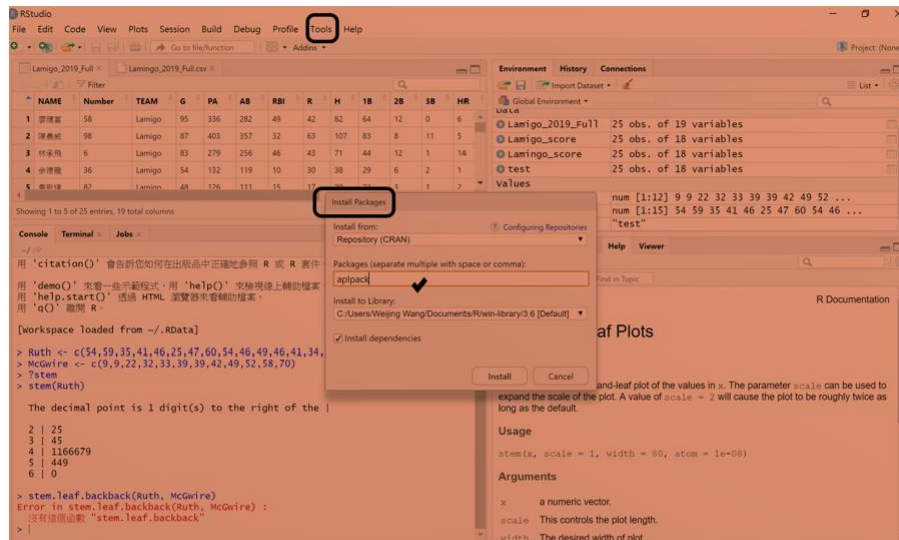
two tests are:

	Regular test	Randomized test
Mean	59.9	44.8
Standard deviation	10.2	12.7

a. (5 points) Please compute the coefficient of variation (CV) for each test.

b. Compare the two tests based on the results.

To make back-to-back stemplot, you need to install the package “aplpack”.



After installing the package “aplpack”, you still need to do the following.

```
> library(aplpack)
> stem.leaf.backback(,,|
```