Hw4

Haichuan Du

6/23/2020

```r
library(clinfun)
library(jmuOutlier)
#library(NSM3) #Associated with the textbook.
#library(lsmeans) #For some of the ANOVA functions.
```

#3

```r
x=c(3500, 3500, 3500, 4000, 4000, 4000, 4300, 4500, 4500, 4900, 5200, 6000, 6750, 8000)
y=c(5710, 6110, 8060, 8080, 11400)
z=c(2930, 3330, 3580, 3880, 4280, 5120)
a=c(6320, 6860, 11400, 14000)
b=c(3230, 3880, 7640, 7890, 8280, 16200, 18250, 29900)

kruskal.test(list(x,y,z,a,b)) #p-value = 0.002242
```
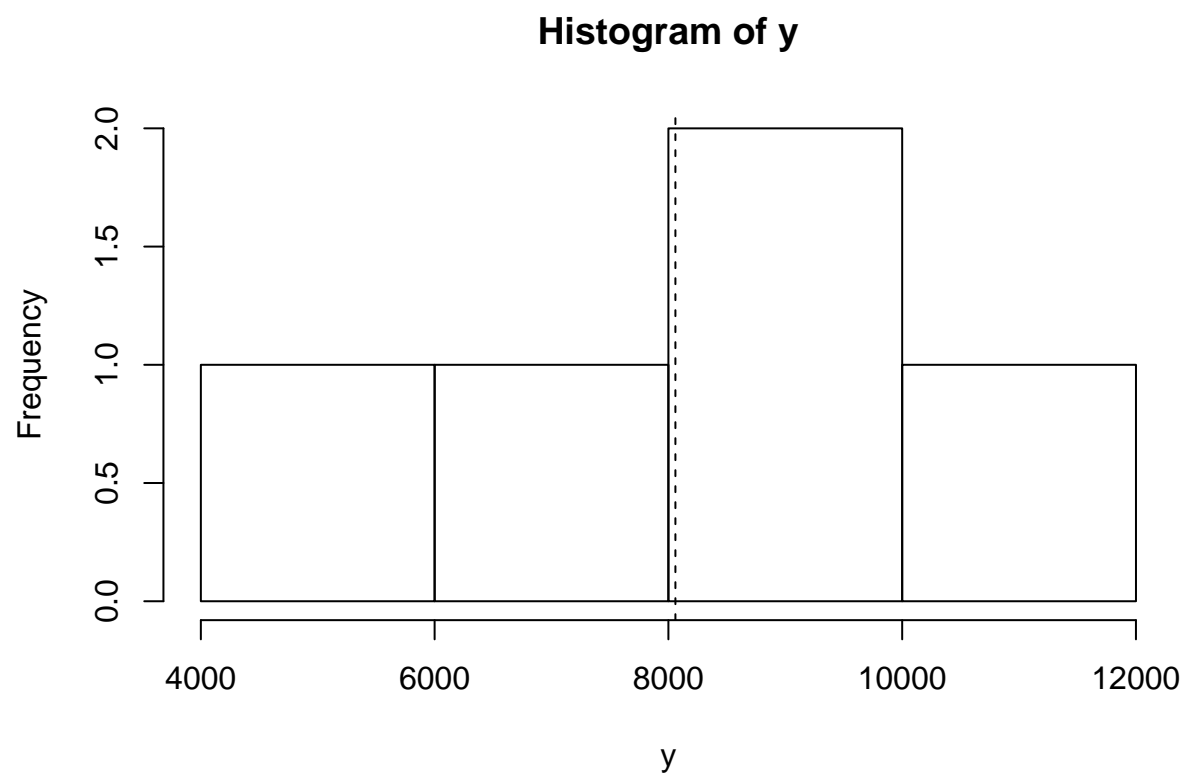
```
##
##  Kruskal-Wallis rank sum test
##
## data:  list(x, y, z, a, b)
## Kruskal-Wallis chi-squared = 16.668, df = 4, p-value = 0.002242
```
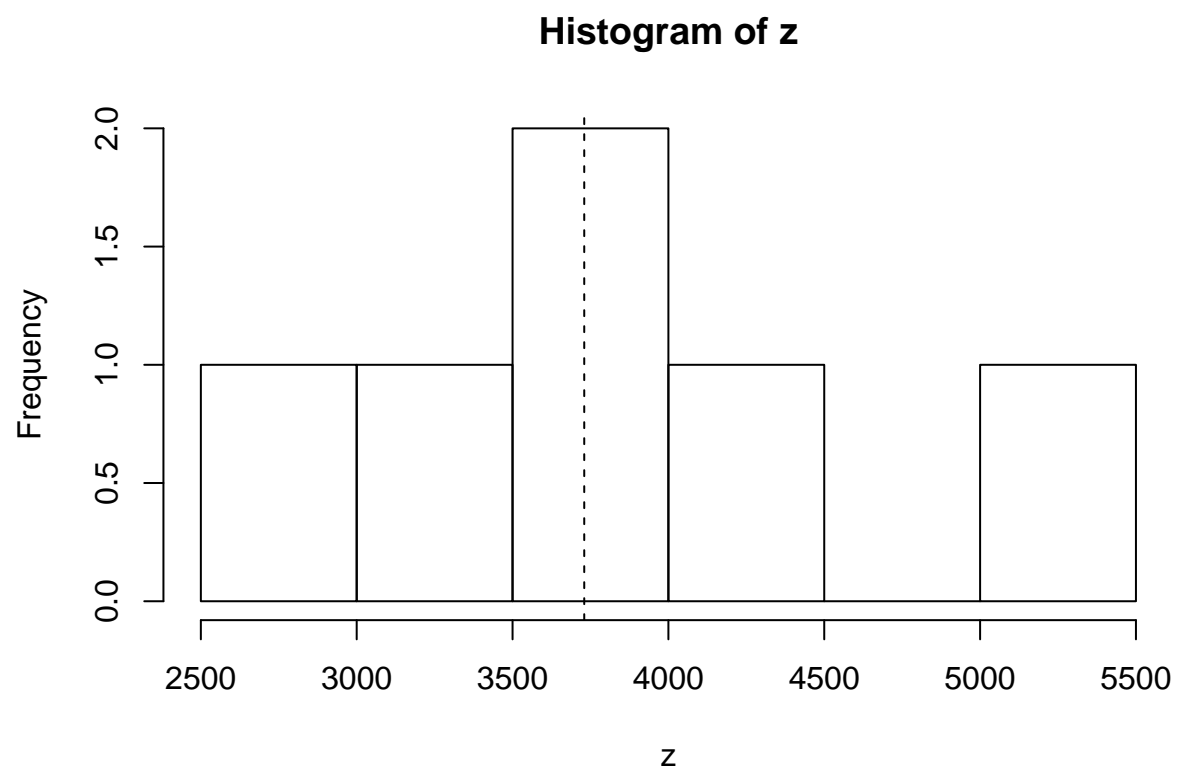
```r
hist(x)
abline(v=median(x), lty=2)
```

Histogram of x
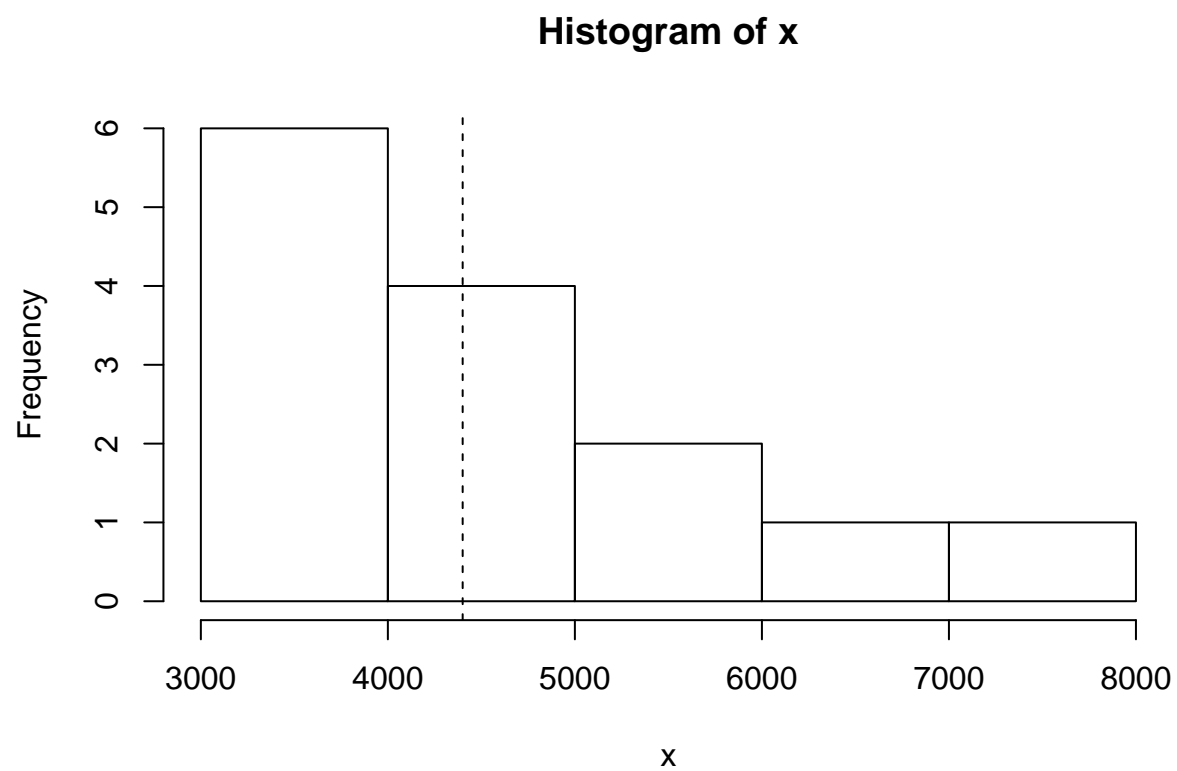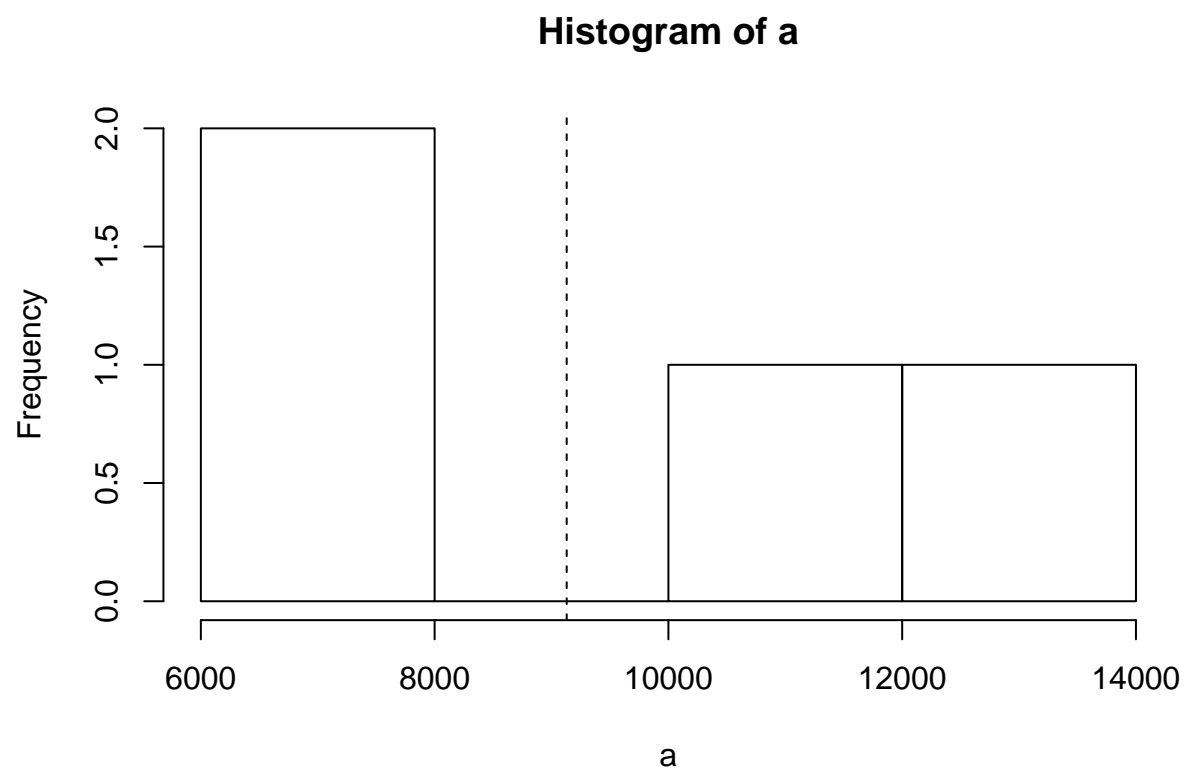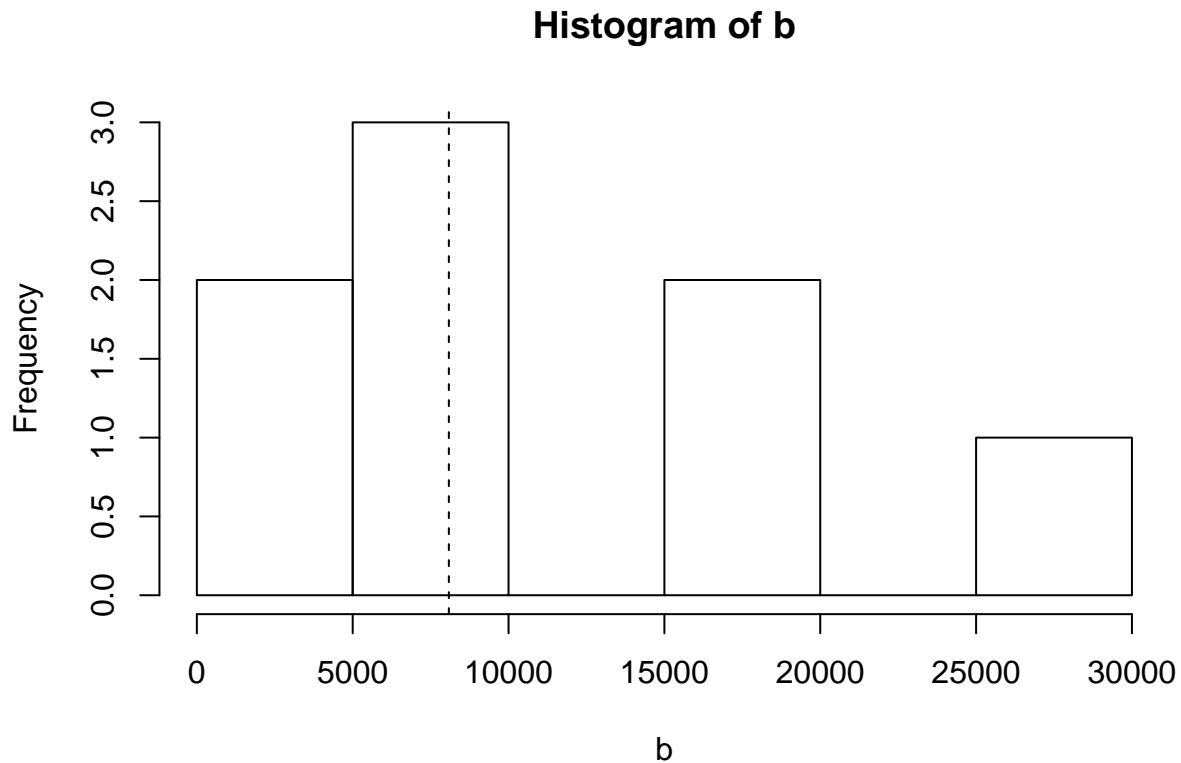


```
hist(y)
abline(v=median(y), lty=2)
```

Histogram of y


```
hist(z)
abline(v=median(z), lty=2)
```

Histogram of z


```r
hist(a)
abline(v=median(a), lty=2)
```

Histogram of a


```r
hist(b)
abline(v=median(b), lty=2)
```

Histogram of b

Ho: m1=m2=m3=m3=m5 vs Ha: The population median are not all the same

we use the Kruskal-Wallis rank sum test, which here gives a p-value of 0.002242. We reject Ho and conclude Ha, that is, we have enought evidence at 0.05 level to conclude that they are differences between the median numbers of GR sites per leukocyte cell for the population of normal subjects and the populations of patients with hairy-cell leukemia, chronic lymphatic leukemia, chronic myelocytic leukemia, or acute leukemia.

(i) The test results are consistent with the plots since the polulation medians are not all the same

(ii) The shift model seems reasonable here because we have more than two populations and some shift somewhere

#4

```
data4 <- c(1.91, 1.53, 2.08, 1.71, 2.44, 2.45, 2.04, 1.6, 2.37, 2.52, 2.36, 2.73, 2.78, 2.88, 2.10, 1.6
groups=rep(1:5,c(4,1,4,3,4)) #Specifying which values are for which samples.
jonckheere.test(data4,groups, alternative = "increasing")
```

```
##
##   Jonckheere-Terpstra test
##
## data:
## JT = 73, p-value = 0.01534
## alternative hypothesis: increasing
```

Ho: Equal distribution vs Ha: Population with bigger i tend to have bigger value

we can use the Jonckheere-Terpstra test, which here gives a p-value of 0.01534, we reject Ho and conclude at level 0.05 that the average basal area increment for a given stand of oak trees is an increasing function of the value of the stand's growing site index.

#6

```r
data2=c(1, 2, 3, 4, 5, 21, 22, 23, 24, 25, -1, -2, -3, -4, -5 )
groups=rep(1:3,c(5,5,5)) #Specifying which values are for which samples.
kruskal.test(data2,groups) #Kruskal-Wallis test.
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  data2 and groups
## Kruskal-Wallis chi-squared = 12.5, df = 2, p-value = 0.00193
```

```r
jonckheere.test(data2,groups) #Jonckheere-Terpstra test. (clinfun)
```

```
##
##  Jonckheere-Terpstra test
##
## data:
## JT = 25, p-value = 0.2098
## alternative hypothesis: two.sided
```

```r
print("The dataset for scenario (a): ")
```

```
## [1] "The dataset for scenario (a): "
```

```r
data2
```

```
##  [1]  1  2  3  4  5 21 22 23 24 25 -1 -2 -3 -4 -5
```

```r
data1=c(1, 3, -2, -4, -5, -1, 2, -3, 4, 5, -21, 22, 23, 24, 25 )
groups=rep(1:3,c(5,5,5)) #Specifying which values are for which samples.
kruskal.test(data1,groups) #Kruskal-Wallis test.
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  data1 and groups
## Kruskal-Wallis chi-squared = 4.22, df = 2, p-value = 0.1212
```

```r
jonckheere.test(data1,groups) #Jonckheere-Terpstra test. (clinfun)
```

```
##
##  Jonckheere-Terpstra test
##
## data:
## JT = 59, p-value = 0.02393
## alternative hypothesis: two.sided
```

```r
print("The dataset for scenario (b): ")
```

```
## [1] "The dataset for scenario (b): "
```

```r
data1
```

```
##  [1]   1   3  -2  -4  -5  -1   2  -3   4   5 -21  22  23  24  25
```