AMS380 - Homework 02

Tuan Bui, SBU ID: 113141951

9/9/2021

Question 1:

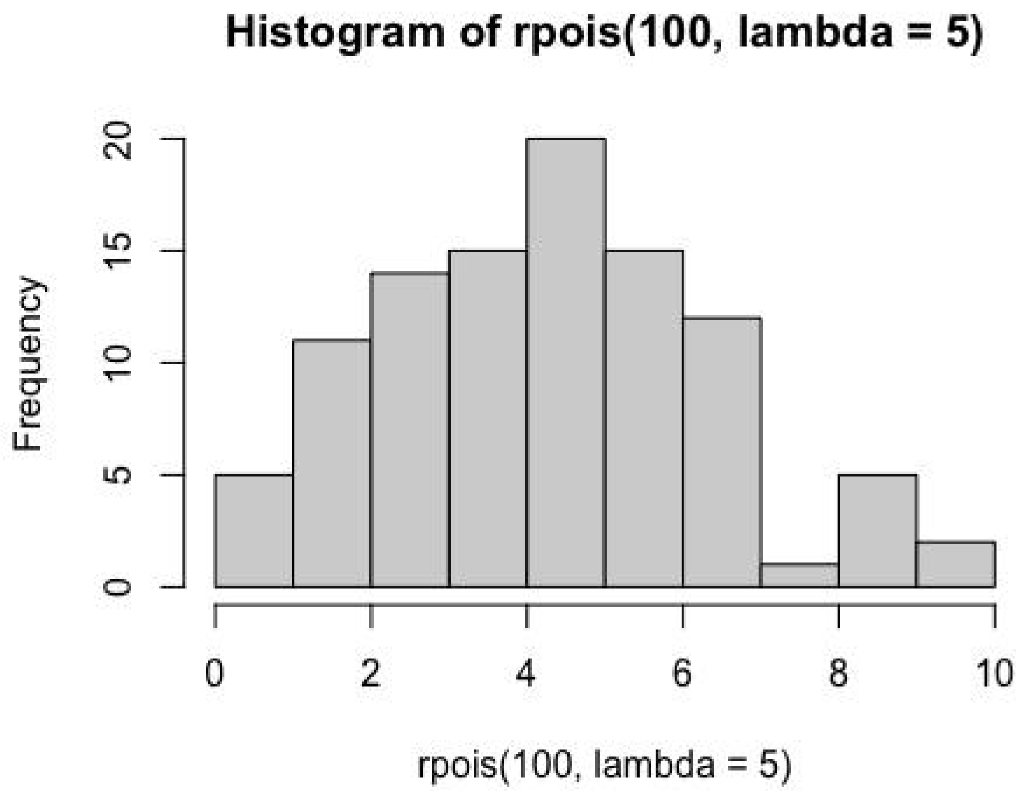
Generate 100 random samples following Poission distribution with lambda = 5.

```
rpois(100, lambda = 5)
```

```
##   [1]  7  6  2  9  3  1  3  5  7  7  2  2  8  6  5  3  5 12  6  6  4  9  5  4  3
##  [26]  2  7  7  3  5  6  2  2  7  5  6  2  7  5  6  5  3  5  3  3  6  3  3  7  1
##  [51]  4  2  5  3  5  6  6  7  8  4  6  6  3  5  6  7  4  3  5  2  5  5  1  4  4
##  [76]  6  8  5  4  7  7  2  2  3  3  8  5  4  7  4  3  6  4  9  6  7  6  7  7  4
```

Histogram of the sample

```
hist(rpois(100, lambda = 5))
```



Question 2

a. 6 < X < 8 if X follows an exponential distribution with rate is 3

```
pexp(8, rate = 3) - pexp(6, rate = 3)

## [1] 1.519223e-08
```

Answer: The probability of 6 < X < 8 if X follows an exponential distribution with rate is 3: 1.519223e-08

b. X > 10 if X follows a normal distribution with mean is 8 and standard is 3

```
pnorm(10, mean = 8, sd = 3, lower.tail = F)

## [1] 0.2524925
```

Answer: The probability of X > 10 if X follows a normal distribution with mean is 8 and standard is 3: 0.2524925

c. X^2 < 5 if X follows a binomial distribution with size 20 and prob is 0.4

```
pbinom(sqrt(5), size = 20, prob = 0.4)

## [1] 0.003611472
```

Answer: The probability of X^2 < 5 if X follows a binomial distribution with size 20 and prob is 0.4

Question 3

```
combination_base <- choose(300, 110)^2
prop <- 0
sq <- 55
dq <- 55


#probability of 2 students getting at least 50% of the test
repeat{
  prop <- ((choose(300,110)*choose(110,sq)*choose(190,dq))/combination_base) + prop
  sq <- sq + 1
  dq <- dq -1
  if(sq == 110) break }
prop
```

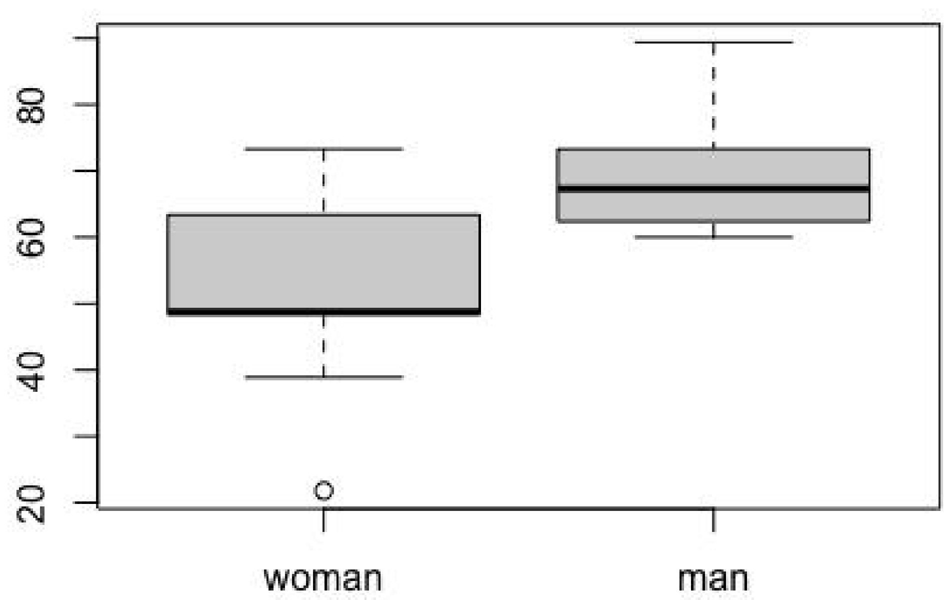
```
## [1] 0.0002275255
```

The probability that 2 students get at least the same 50% of the test is 0.0002275255

Question 4

```
# H0: mu_man = mu_woman Ha: mu_man != mu_woman
# load data
woman <- c(38.9, 61.2, 73.3, 21.8, 63.4, 64.6, 48.4, 48.8, 48.5)
man <- c(67.8, 60.0, 63.4, 76.0, 89.4, 73.3, 67.3, 61.3, 62.4)

# visualize data by box plot
boxplot(woman, man, names = c('woman', 'man'))
```




```
shapiro.test(woman)

##
##  Shapiro-Wilk normality test
##
```

```
## data:  woman
## W = 0.94266, p-value = 0.6101
```

```
shapiro.test(man)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  man
## W = 0.86425, p-value = 0.1066
```

```
# p-value of the shapiro-test of both samples are 0.6101 and 0.1066, which are greater than the significance level(0.05), not reject H0, both samples are normal.

var.test(woman,man)
```

```
##
##  F test to compare two variances
##
## data:  woman and man
## F = 2.7675, num df = 8, denom df = 8, p-value = 0.1714
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##   0.6242536 12.2689506
## sample estimates:
## ratio of variances
##           2.767478
```

```
var.test(man,woman)
```

```
##
##  F test to compare two variances
##
## data:  man and woman
## F = 0.36134, num df = 8, denom df = 8, p-value = 0.1714
```

```
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.08150656 1.60191315
## sample estimates:
## ratio of variances
##          0.3613398
```

```
# p-value of the variance test is 0.1714, which is greater than the significance level
(0.05), not reject H0, the variance of both samples are assumed to be the same.

t.test(woman, man, mu = 0, var.equal = T)
```

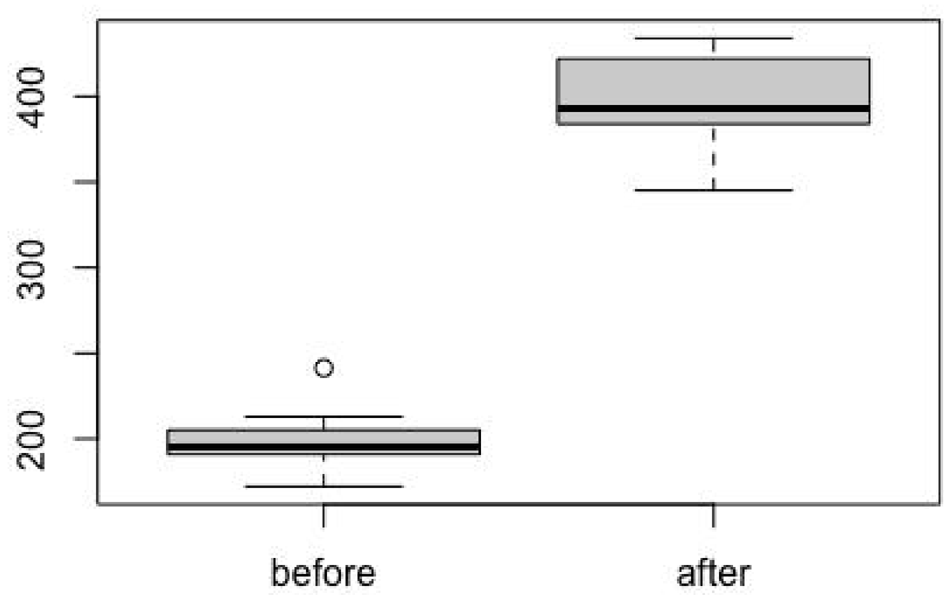
```
##
##  Two Sample t-test
##
## data:  woman and man
## t = -2.7842, df = 16, p-value = 0.01327
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -29.748019  -4.029759
## sample estimates:
## mean of x mean of y
##  52.10000  68.98889
```

```
# p-value of the t-test is 0.01327, which is less than the significance level 0.05, rej
ect H0, the mean of woman'weight and man's weight are significantly different.
```

Question 5

```
#H0: mu_before = mu_after, Ha: mu_before <= mu_after
# load data
before <- c(200.1, 190.9, 192.7, 213.0, 241.4, 196.9, 172.2, 185.5, 205.2, 193.7)
after <- c(392.9, 393.2, 345.1, 393.0, 434.0, 427.9, 422.0, 383.9, 392.3, 352.2)
diff <- after - before

# visualize data by box plot
boxplot(before, after, names = c('before', 'after'))
```




```
shapiro.test(diff)

##
##  Shapiro-Wilk normality test
##
```

```
## data:  diff
## W = 0.94536, p-value = 0.6141
```

 # p-value of shapiro test of the different is 0.6141, greater than the significance level 0.05, not reject H0, the different can be assumed normal.

```
t.test(diff, mu = 0, alternative = 'greater', conf.level = 0.9)
```

```
##
##  One Sample t-test
##
## data:  diff
## t = 20.883, df = 9, p-value = 3.1e-09
## alternative hypothesis: true mean is greater than 0
## 90 percent confidence interval:
##  181.6095      Inf
## sample estimates:
## mean of x
##    194.49
```

 # p-value of the paired sample t-test is less than the significance level 0.10, reject H0, the mean of the weight before is significantly greater than one before.