Pstat 174-Lab 1

Hongxu Ma

2022-10-03

- Question 1

```
set.seed(1)
X=runif(1000,-1,1)
Y=X^2
cor(X,Y)
```

```
## [1] 0.04042929
```

```
knitr::opts_chunk$set(echo = TRUE)
```

They are uncorrelated since correlation of X and Y is 0.04042929. They are not independent because there is quadratic relationship between X and Y. They are uncorrelated if X and Y are independent We can't conclude that they are independent or not if X and Y are uncorrelated.

- Question 2

```
set.seed(1)
x1 = runif(10, min = -1, max = 1)
mean(x1)
```

```
## [1] 0.1030277
```

```
x2 = runif(100, min = -1, max = 1)
mean(x2)
```

```
## [1] 0.03034921
```

```
x3 = runif(1000, min = -1, max = 1)
mean(x3)
```
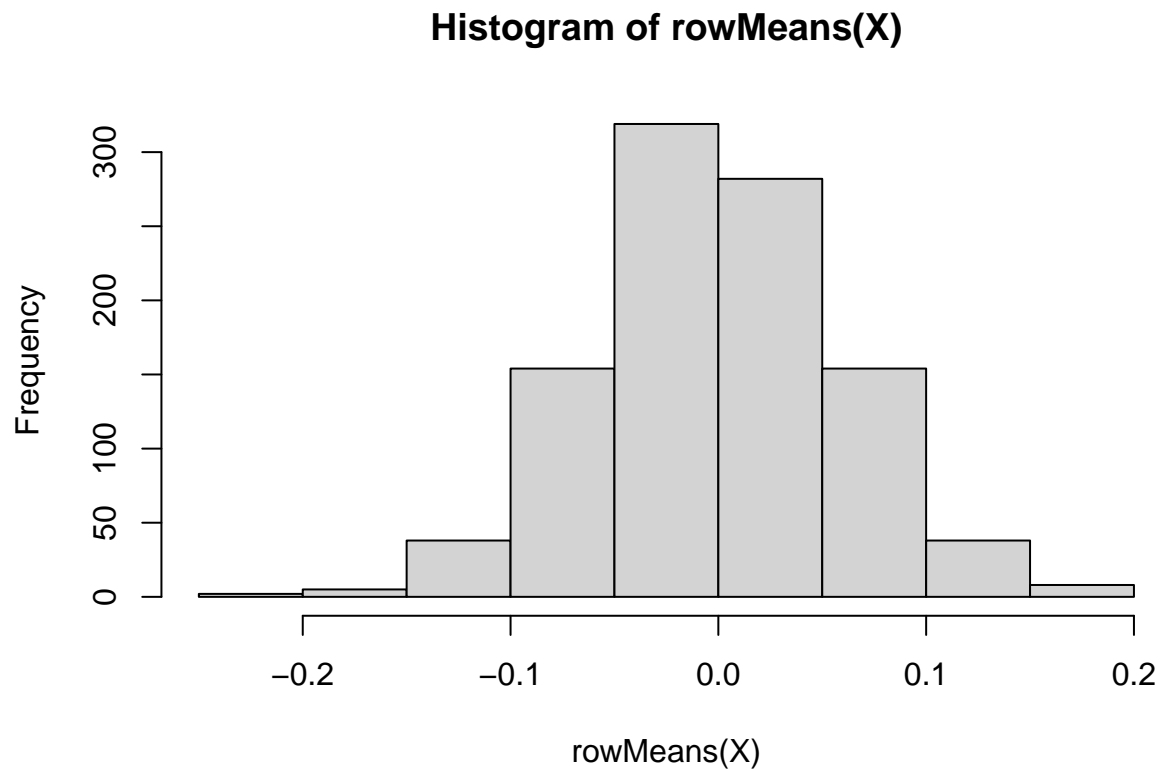
```
## [1] -0.007201965
```

We can conclude that the sample mean gets closer to the true mean as the sample size increases.

- Question 3

```
set.seed(1)
n = 100
rows = 1000
X <- matrix(runif(n*rows, -1, 1), rows)
hist(rowMeans(X))
```

Histogram of rowMeans(X)

sampling distribution of the sampling means approaches a normal distribution
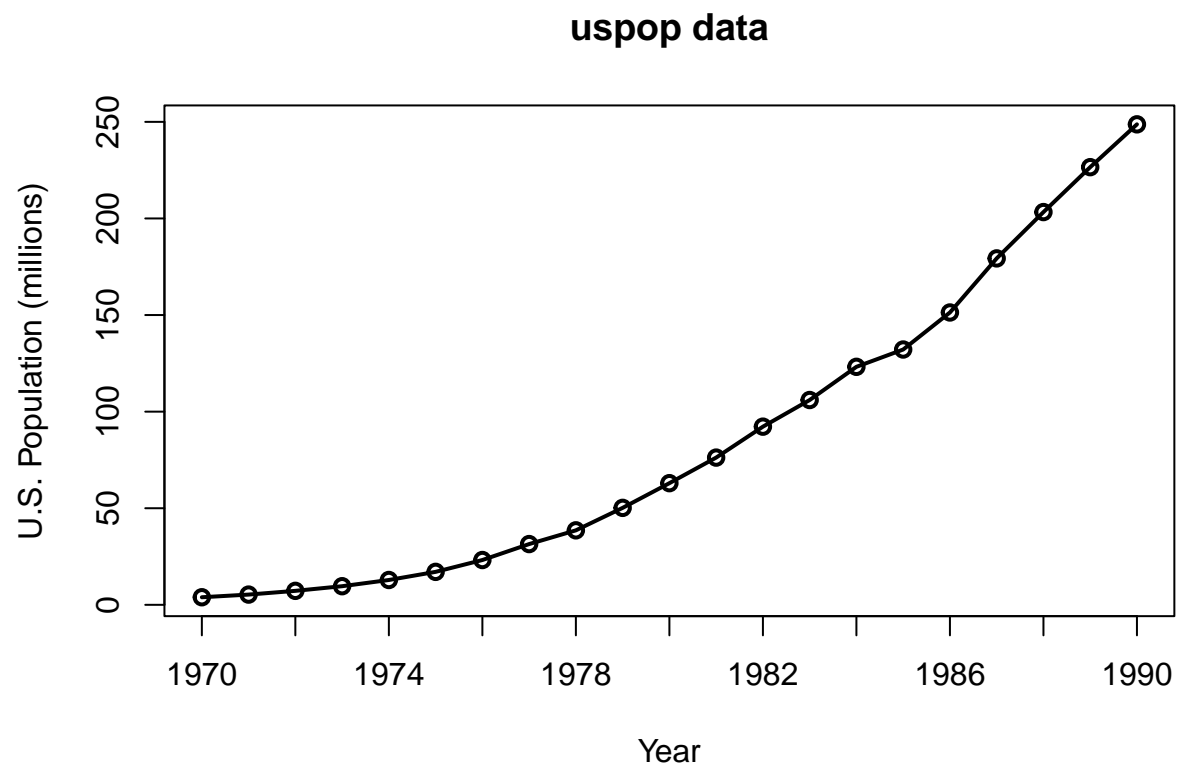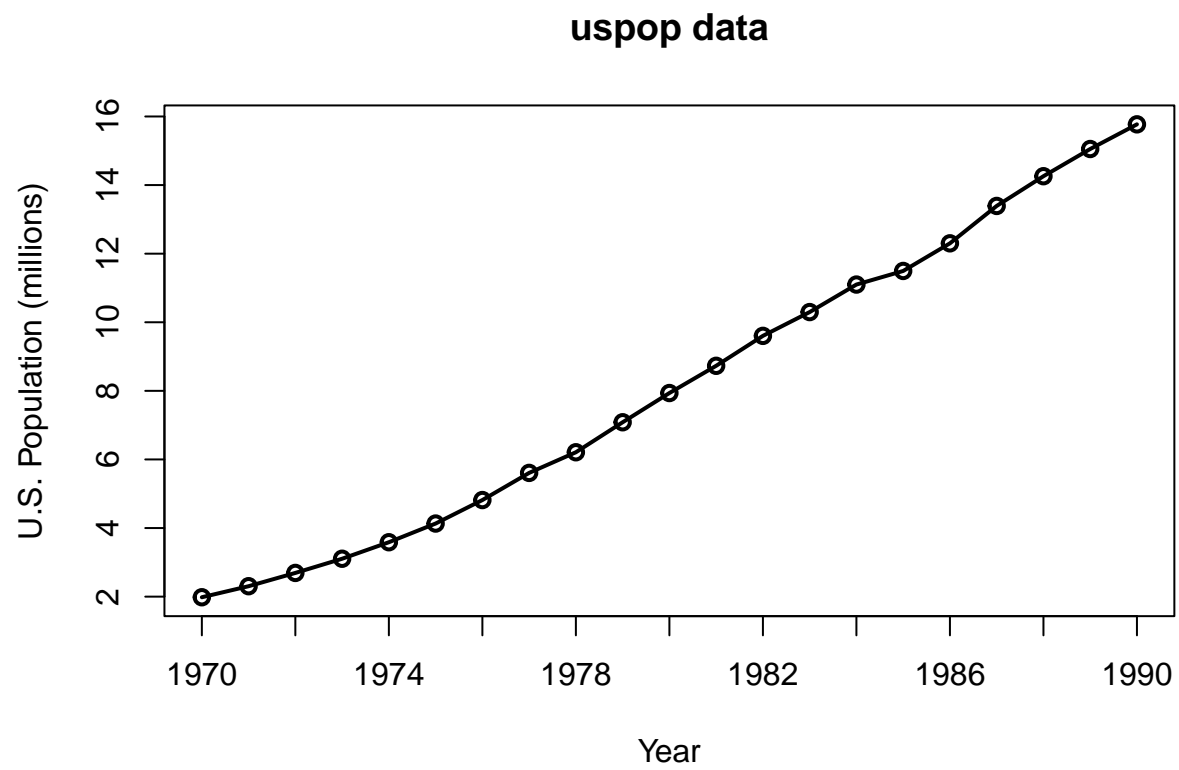
- Question 4

```
uspop <- scan('uspop.txt')

plot(uspop/1000000, xaxt = 'n', type = 'o', main = "uspop data", lwd = 2, xlab = "Year", ylab = "U.S. Po
axis(1, seq(1,21,2), seq(1970,1990,2))
```

uspop data


```
uspop_sqrt <- sqrt(uspop/1000000)

plot(uspop_sqrt, xaxt = 'n', type = 'o', main = "uspop data",
     lwd = 2, xlab = "Year", ylab = "U.S. Population (millions)")
axis(1, seq(1,21,2), seq(1970,1990,2))
```

uspop data

The image of the second graph tends to be more linear, which means that the slope grows more smoothly.