

PSTAT274 Lab 01

AO XU

2022-10-03

Q1

```
set.seed(100)
x = runif(1000,-1,1)
y = x^2
cor(x,y)
```

```
## [1] 0.06219732
```

```
cov(x,y)
```

```
## [1] 0.01079364
```

X and y are uncorrelated since $\text{cor}(x,y) = 0.06219732$ which is very close to zero.

X and Y are clearly dependent since $y = x^2$.

It could happen that two variables are uncorrelated but not dependent.

Q2

```
set.seed(1011)
x1 = runif(10,-1,1)
x2 = runif(100,-1,1)
x3 = runif(1000,-1,1)
mean(x1)
```

```
## [1] 0.06742002
```

```
mean(x2)
```

```
## [1] 0.03776725
```

```
mean(x3)
```

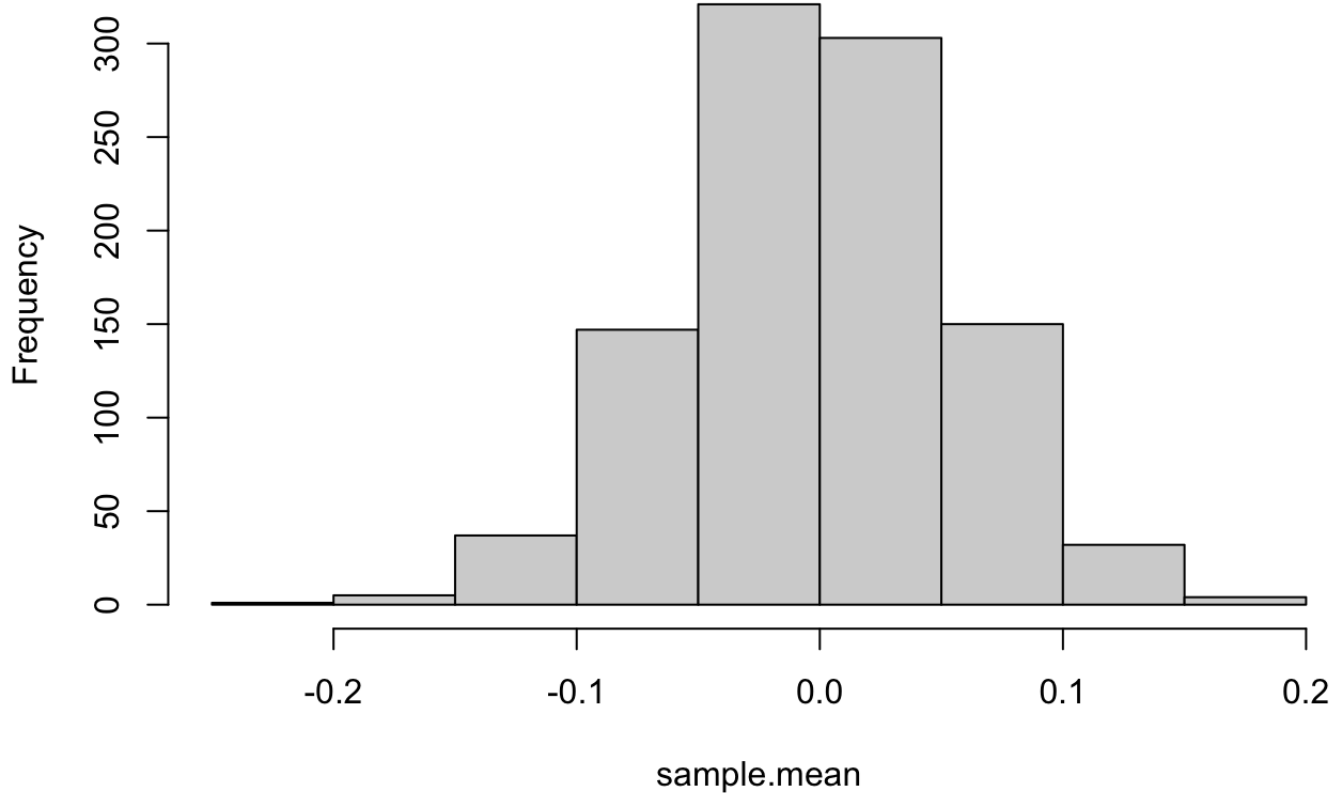
```
## [1] 0.01280656
```

As sample size increases, sample mean is approaching to the true mean.

Q3

```
n = 100
rows = 1000
sim = matrix(runif(n*rows,-1,1),rows)
sample.mean = rowMeans(sim)
hist(sample.mean)
```

Histogram of sample.mean



mean = 0, variance = $1/300$ by using central limit theorem, it's normal distribution

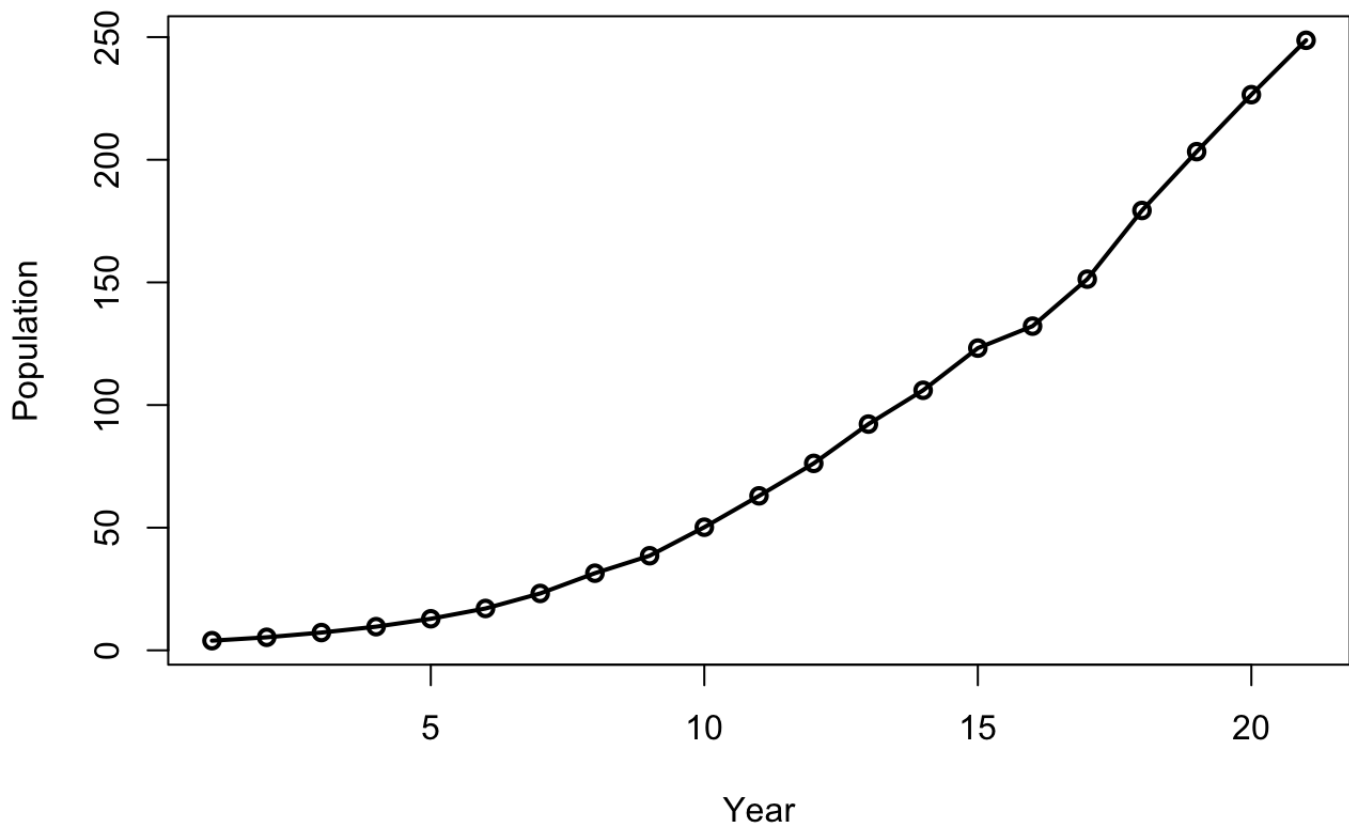
Q4

```
uspop = scan("uspop.txt")  
uspop
```

```
## [1] 3929214 5308483 7239881 9638453 12860702 17063353 23191876  
## [8] 31443321 38558371 50189209 62979766 76212168 92228496 106021537  
## [15] 123202624 132164569 151325798 179323175 203302031 226542203 248709873
```

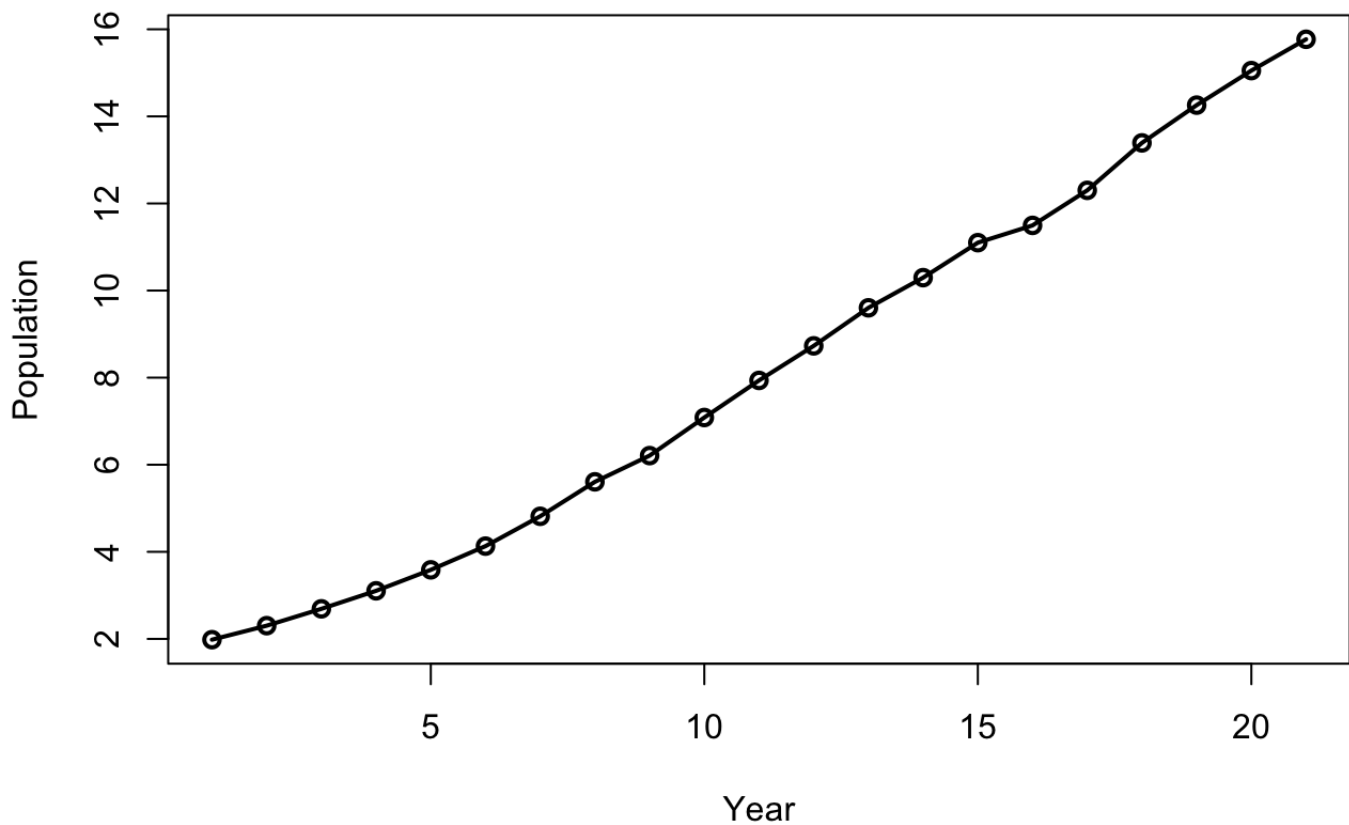
```
plot(uspop/1000000, type = "o", main = "US population", xlab = "Year", ylab = "Population", lwd = 2)
```

US population



```
plot(sqrt(uspopp/1000000), type = "o", main = "US population", xlab = "Year", ylab = "Population", lwd = 2)
```

US population



For

the sqrt, the graph is more linear ◦