HW_6_1

The 2 variables that I choose are satisfactory level (continous) ad year at company (discrete)

Code

```
> firstVar = c(0.31, 0.41, 0.38, 0.39, 0.44)


> secondVar= c( 0.4, 0.46, 0.41, 0.45, 0.44, 0.36, 0.45, 0.43, 0.38, 0.39, 0.4, 0.4, 0.41,
0.37, 0.43, 0.37, 0.4, 0.36, 0.43, 0.36)

> thirdVar = c(0.09, 0.11, 0.1, 0.11, 0.11, 0.11, 0.1, 0.1)
> fourthVar = c(0.19, 0.1, 0.11, 0.09, 0.23, 0.11, 0.24, 0.91, 0.89, 0.9)

> fifthVar = c(0.88, 0.77, 0.84, 0.83, 0.9, 0.84)

> boxplot(firstVar,secondVar,thirdVar,fourthVar,fifthVar)
```

Output

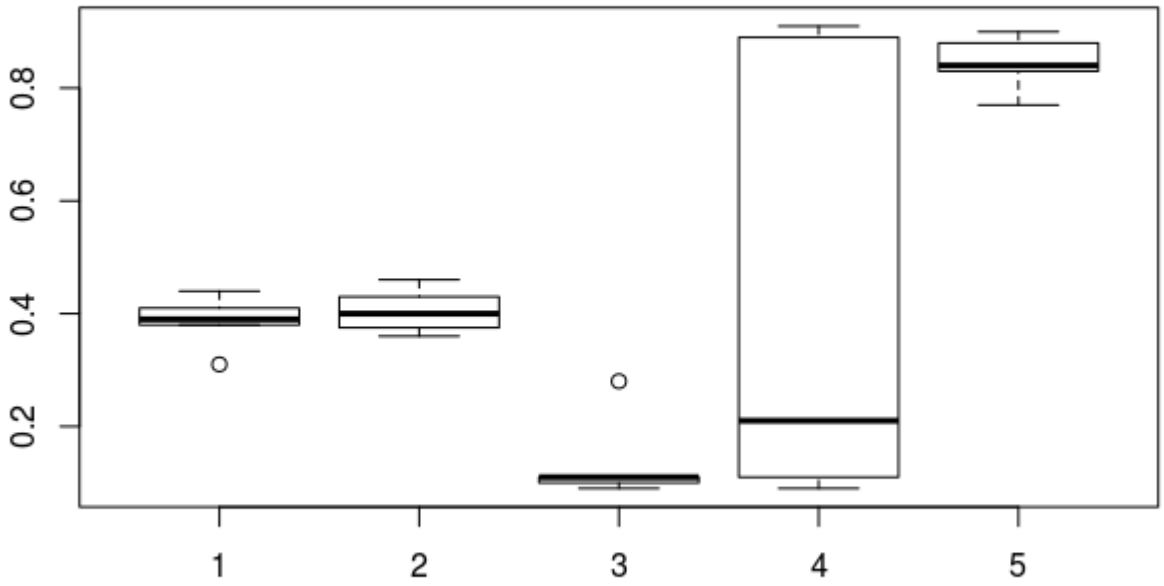

Illustration 1: Box plot of Satisfactory level of different year

Interpretation

Looking at the box plot, I can see that people that stay at the company for 1,2,3 and 5 years usually have roughly the same range of satisfactory with others with people in the same year. That is because at those years, the value does not deviate much that make the box plot longer, unlike year 4 where people have all different satisfactory level that make the box plot big. This indicate that at year 4, people do not know for sure if they like the company or hate it, making a variability in the satisfactory level. Where as for other years, people already know if they are satisfy or not.

Another interesting thing is that it seems like people who work at the company for 5 year really like the company. Where as year 1 and 2 is about average and people hate the company at year 3 and 4

HW_6_2

Looking at the box plot, I can see that the box plot is not generally affected by the sample size. However, as the sample size increase, the variability also increase as well like the min and max value

Code

```
> firstVar= rnorm(20,0,10)
> secondVar= rnorm(30,0,10)
```

```
> thirdVar= rnorm(40,0,10)
> fourthVar= rnorm(50,0,10)
> fifthVarVar= rnorm(100,0,10)
> boxplot(firstVar,secondVar,thirdVar,fourthVar,fifthVarVar)
```

Output

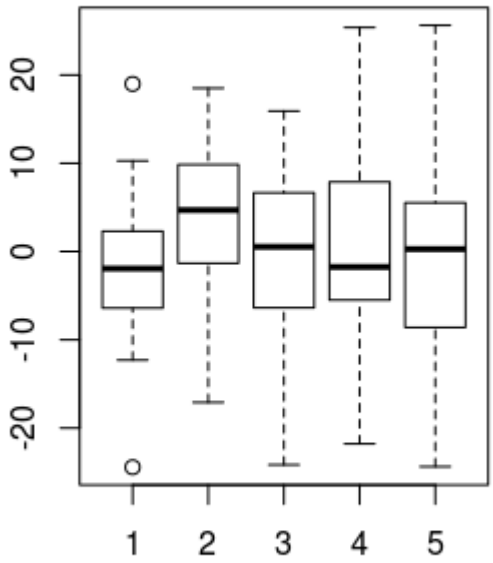

Illustration 2: Box plot of different sample size

HW8-1

c)

```
> a = dbinom(1:100,4,1/4)
> range = 1:100
> sum(a* range)
[1] 1
```

Both answer are the same using both the E(x) and n*mu