a. The population mean, which is the true mean, is the sum of the observations divided by the number of samples which is N = 7 in our case.

```
X = c(11, 13, 15, 17, 19, 21, 23)
mean(X)
[1] 17
```

b. The population variance is:

```
u = mean(X)
v = (X - u)^2
Var = sum(v) / 7
Var
[1] 16
```

c. Taking sample size of size 3 (i.e n = 3) from the given population (with replacement)

```
X=c(11, 13, 15, 17, 19, 21, 23)
d=expand.grid(X,X,X) #343 obs. of 3 variables
```

d. Doing row-wise operation on our data set d, calculating the mean for each row, we end up with a vector of 343 elements.

```
X_bar = apply(d, 1, mean) #mean of each row saved in variable X_bar
```

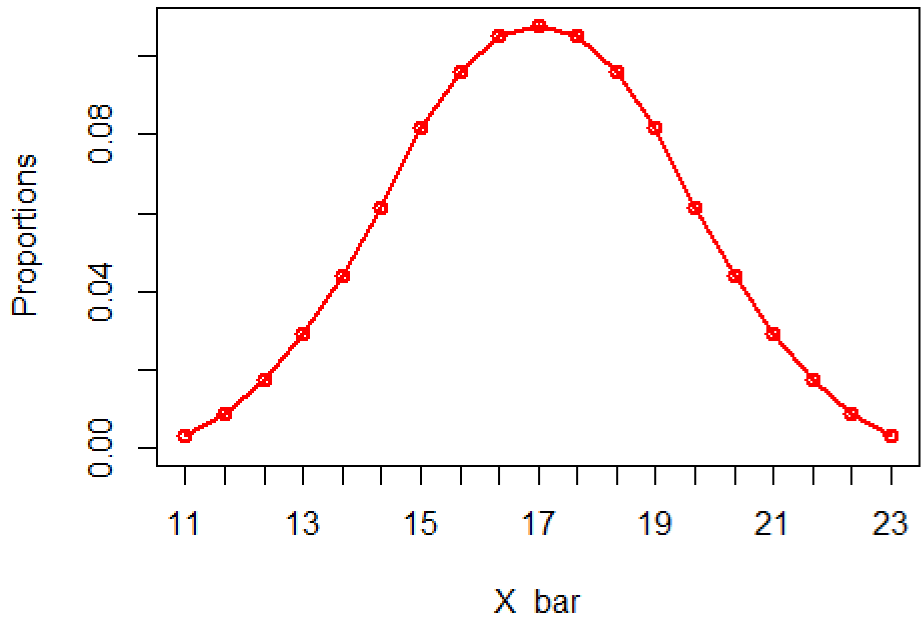
e. Constructing frequency table based on X_bar:

```
 freqTable = as.data.frame(table(X_bar))
 freqTable
         X_bar        Freq
1                11     1
2   11.6666666666667    3
3   12.3333333333333    6
4                13    10
5   13.6666666666667   15
6   14.3333333333333   21
7                15    28
8   15.6666666666667   33
9   16.3333333333333   36
10               17    37
```

```
11 17.6666666666667   36
12 18.3333333333333   33
13               19   28
14 19.6666666666667   21
15 20.3333333333333   15
16               21   10
17 21.6666666666667    6
18 22.3333333333333    3
19               23    1
```

```
#constructing proportion table
propTable <- prop.table(table(X_bar))
propTable <- as.data.frame(propTable)
colnames(propTable) <- c('X_bar', 'Prop')
propTable
              X_bar        Prop
1                11 0.002915452
2  11.6666666666667 0.008746356
3  12.3333333333333 0.017492711
4                13 0.029154519
5  13.6666666666667 0.043731778
6  14.3333333333333 0.061224490
7                15 0.081632653
8  15.6666666666667 0.096209913
9  16.3333333333333 0.104956268
10               17 0.107871720
11 17.6666666666667 0.104956268
12 18.3333333333333 0.096209913
13               19 0.081632653
14 19.6666666666667 0.061224490
15 20.3333333333333 0.043731778
16               21 0.029154519
17 21.6666666666667 0.017492711
18 22.3333333333333 0.008746356
19               23 0.002915452
```

f. Ploting these proportions against the values and connecting the points to form a curve: It takes the form the Normal Distribution.



g. The mean of **X_bar** is:

```
mean(X_bar)
[1] 17
```

Compared with the mean in part a, they are the same.

h. The Variance of **X_bar** is:

```
(342/343) * var(X_bar)
[1] 5.333333

var(X) * (6/7) # population var
[1] 16
16 / 3 # where n = 3 is the sample size of each obs.
[1] 5.333333
```

It appears the variance of **X_bar** is different from the variance of the population mean calculated in part b, upon further investigation we discover the relationship between the two answers is a constant, that constant is the sample size (n=3)

i. We demonstrated **Central Limit Theorem (CLT),** which states that any random variable, as n goes to infinity, converges in distribution to Normal.

2a. Calculating S^2 and **sigma_hat**

```
X_M = data.matrix(d)
X_M = X_M - X_bar
X_M = X_M^2
S_sq = apply(X_M, 1, sum) * (1/2)
sig_hat = apply(X_M, 1, sum) * (1/3)
```

<u>TO SHOW IF BOTH ESTIMATOR IS UNBIASED THEY MUST OUTPUT THE FOLLOWING:</u>

$E[S^2]$ = $sigma^2$ where **$sigma^2$ = 16**
E[sigma_hat] = $sigma^2$

```
mean(S_sq)
[1] 16  # population mean is 16, Unbiased

mean(sig_hat)
[1] 10.66667 # population mean is 16, biased
```

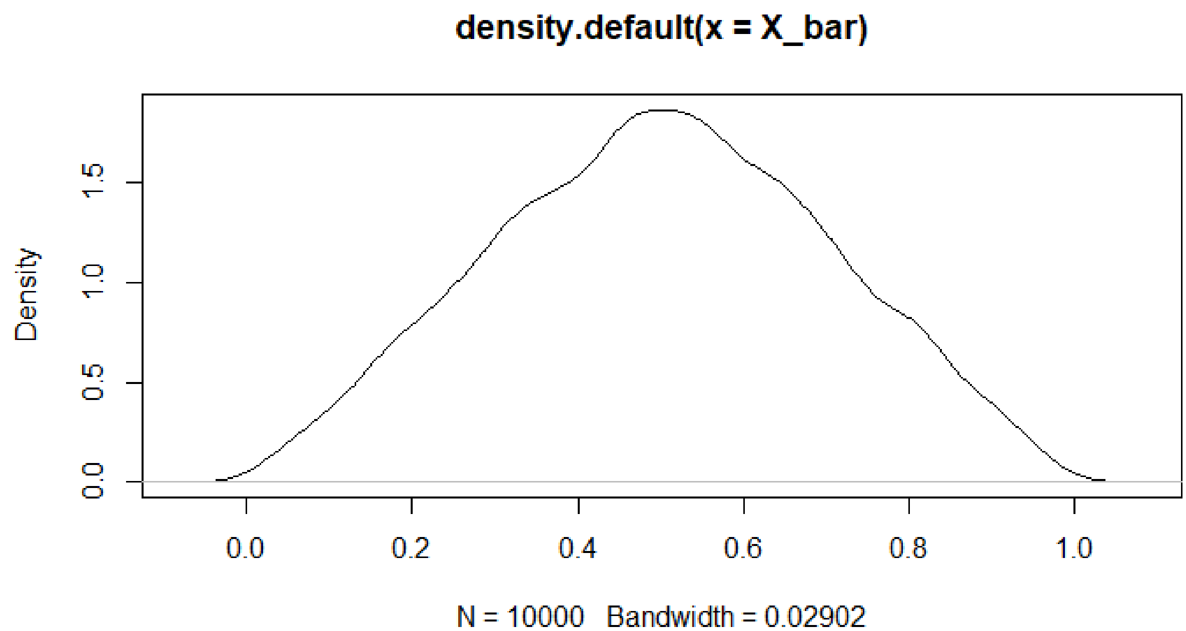
Since the population variance is **16** and **$E[S^2]$ = 16**, it shows that **S^2** is an unbiased estimator of **$sigma^2$**.

Since the population variance is **16** and **E[sigma_hat] = 10.66667 (⅔ * *$sigma^2$*)**, it shows that **$sigma\_hat^2$** is an biased estimator of **$sigma^2$**.

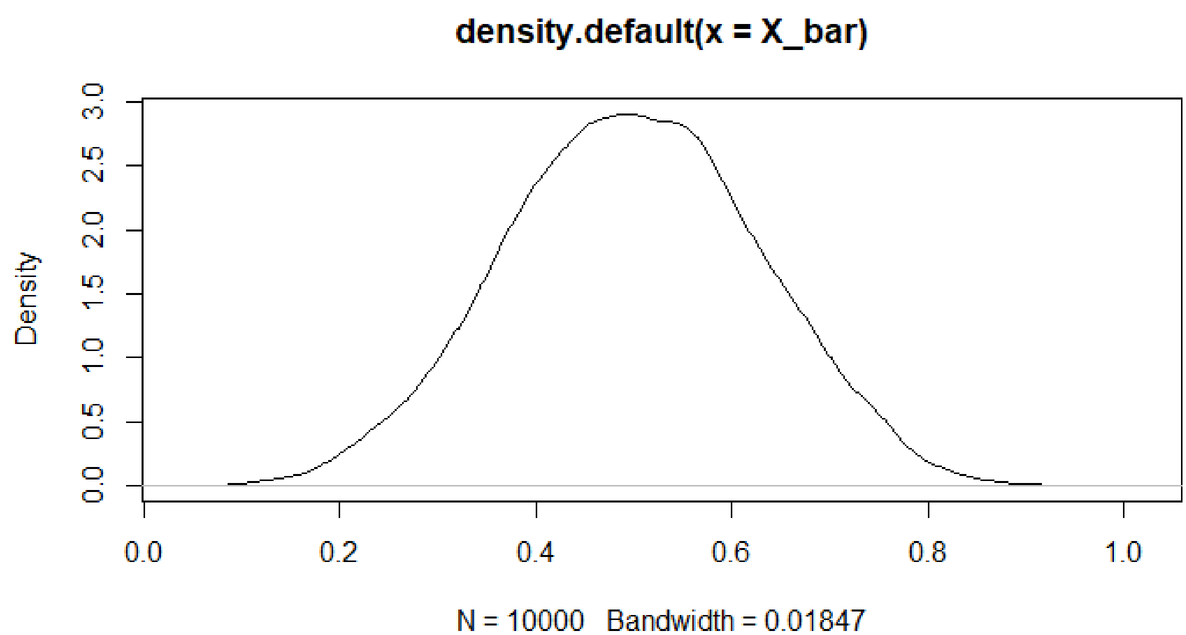
2b.

```
mse = mean((sig_hat-16)^2)
v = var(sig_hat)*(342/343)
bias = mean(sig_hat) - 16
> (v + bias^2) == mse
[1] TRUE #shown
> mse
[1] 94.81481
> v + bias^2
[1] 94.81481
```

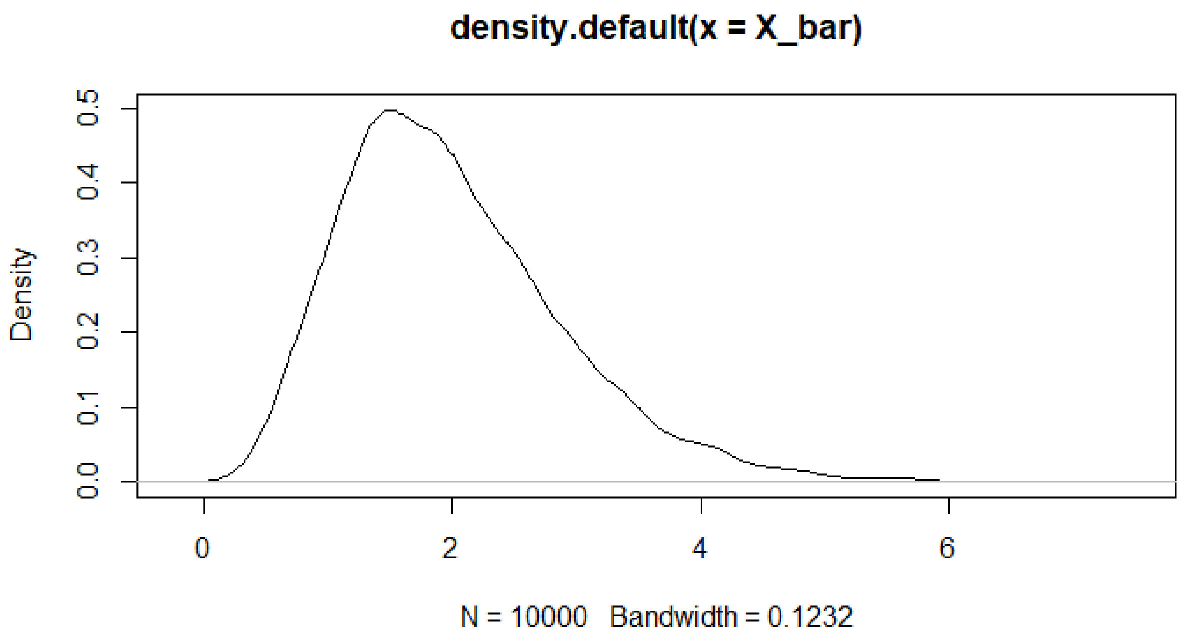
3a.


density.default(x = X_bar)
Density
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0
N = 10000 Bandwidth = 0.02902

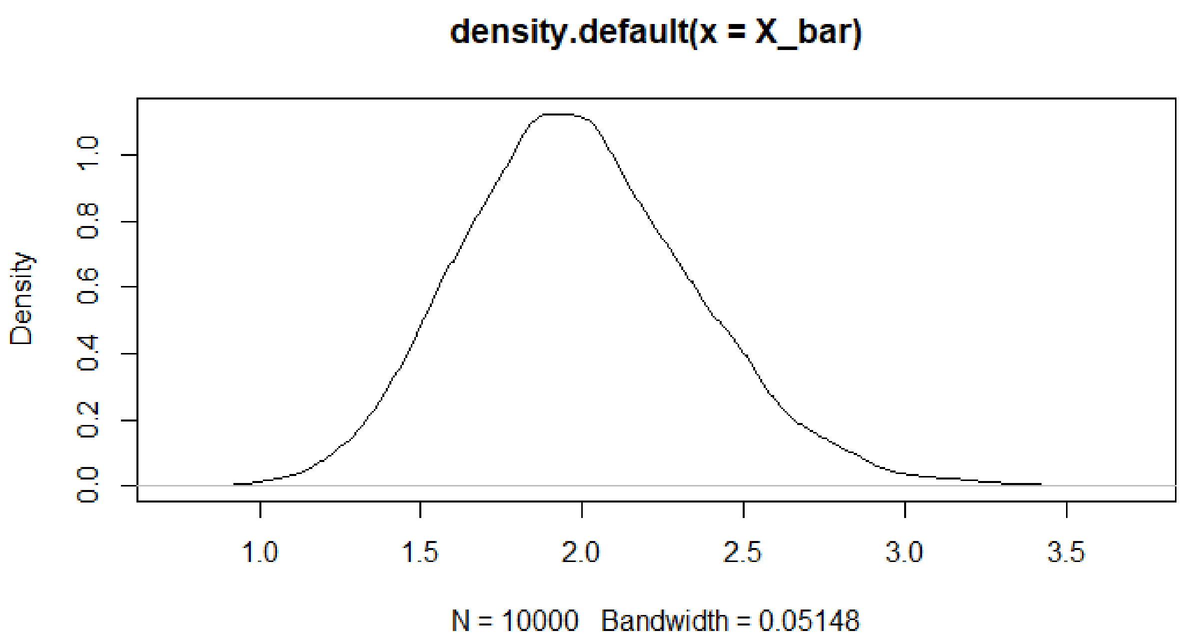
3b.


density.default(x = X_bar)
Density
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0.0
0.2
0.4
0.6
0.8
1.0
N = 10000 Bandwidth = 0.01847

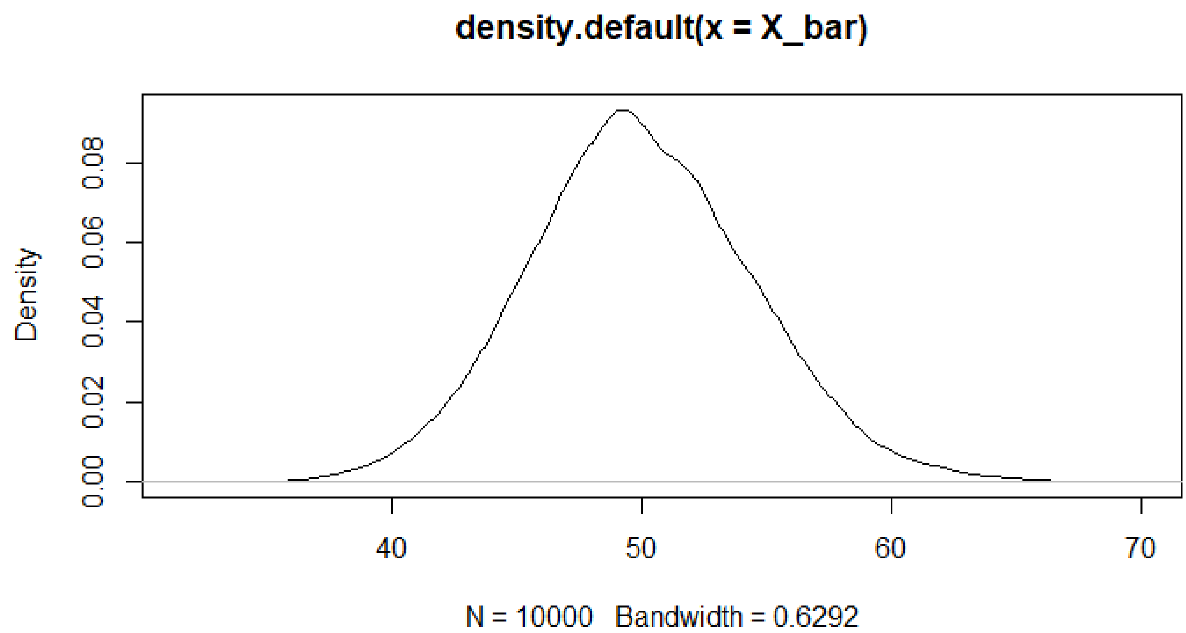
3c.


density.default(x = X_bar)
Density
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
N = 10000 Bandwidth = 0.1232

3d.


density.default(x = X_bar)
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.0
1.5
2.0
2.5
3.0
3.5
N = 10000 Bandwidth = 0.05148

3e.



3f. Central limit theorem states that all sequences of random variables converge in distribution to a normal distribution as n goes to infinity, in our example, comparing a and b, n should be greater or equal to 5 for CLT to take effect. In probability theory and statistics, skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. (source).

Comparing part a and part b, skewness does not play a role, the graph in part a is more of a mountain top, its somewhat symmetric.

For parts c, d, and e, skewness does have play a role. Here graph c is skewed to the left, increasing the number of samples n>= 30 (part d) we can see CLT being applied on the graph as it starts to take the shape of the symmetric normal curve. Interestingly enough, while not increasing the sample size for part e, but increasing the degrees of freedom, we see the curve become symmetric around its mean. This is because the Chi-squared distribution is the standard normal distribution but squared, and for Normal distribution, the CLT does not apply, because the linear combination of a Normal gives back a normal.