Q2

Huiyan Li

3/29/2020

Using for loop to find all combination from this population with replacement and sample size of n = 3

```
# Calculate the 125 combination of the population possible
combination_list = vector()
for (valA in c(21:25)) {
  for (valB in c(21:25)) {
    for (valC in c(21:25)) {
      combination_list <- c(combination_list, c(valA, valB, valC))
    }
  }
}
combination_list <- matrix(combination_list, ncol = 3, byrow=TRUE)
colnames(combination_list) <-  c("Col1", "Col2", "Col3")
```

Given that in the assignment hand out, the equation is using $X_i - \overline{X}$
Therefore, I am assuming $\overline{X}$ is referring to the sample mean of the combination list
As we don't know the actual global mean μ

```
sample_mean = mean(combination_list)
sample_variance_list = vector()
sigma_variance_list = vector()
for (row in 1:nrow(combination_list)) {
  sample = c(combination_list[row, 0:3])

  # Compute the sample variances using the first equation
  sample_variance = sum((sample - sample_mean)**2)/2
  sample_variance_list <- c(sample_variance_list, sample_variance)

  # Compute the sample variances using the second equation
  sigma_variance = sum((sample - sample_mean)**2)/3
  sigma_variance_list <- c(sigma_variance_list, sigma_variance)
}
```

```
s = sample_variance_list
s
```

```
##   [1] 6.0 4.5 4.0 4.5 6.0 4.5 3.0 2.5 3.0 4.5 4.0 2.5 2.0 2.5 4.0 4.5 3.0 2.5
##  [19] 3.0 4.5 6.0 4.5 4.0 4.5 6.0 4.5 3.0 2.5 3.0 4.5 3.0 1.5 1.0 1.5 3.0 2.5
##  [37] 1.0 0.5 1.0 2.5 3.0 1.5 1.0 1.5 3.0 4.5 3.0 2.5 3.0 4.5 4.0 2.5 2.0 2.5
##  [55] 4.0 2.5 1.0 0.5 1.0 2.5 2.0 0.5 0.0 0.5 2.0 2.5 1.0 0.5 1.0 2.5 4.0 2.5
##  [73] 2.0 2.5 4.0 4.5 3.0 2.5 3.0 4.5 3.0 1.5 1.0 1.5 3.0 2.5 1.0 0.5 1.0 2.5
##  [91] 3.0 1.5 1.0 1.5 3.0 4.5 3.0 2.5 3.0 4.5 6.0 4.5 4.0 4.5 6.0 4.5 3.0 2.5
## [109] 3.0 4.5 4.0 2.5 2.0 2.5 4.0 4.5 3.0 2.5 3.0 4.5 6.0 4.5 4.0 4.5 6.0
```

```
sigma = sigma_variance_list
sigma
```

```
##   [1] 4.0000000 3.0000000 2.6666667 3.0000000 4.0000000 3.0000000 2.0000000
##   [8] 1.6666667 2.0000000 3.0000000 2.6666667 1.6666667 1.3333333 1.6666667
##  [15] 2.6666667 3.0000000 2.0000000 1.6666667 2.0000000 3.0000000 4.0000000
##  [22] 3.0000000 2.6666667 3.0000000 4.0000000 3.0000000 2.0000000 1.6666667
##  [29] 2.0000000 3.0000000 2.0000000 1.0000000 0.6666667 1.0000000 2.0000000
##  [36] 1.6666667 0.6666667 0.3333333 0.6666667 1.6666667 2.0000000 1.0000000
##  [43] 0.6666667 1.0000000 2.0000000 3.0000000 2.0000000 1.6666667 2.0000000
##  [50] 3.0000000 2.6666667 1.6666667 1.3333333 1.6666667 2.6666667 1.6666667
##  [57] 0.6666667 0.3333333 0.6666667 1.6666667 1.3333333 0.3333333 0.0000000
##  [64] 0.3333333 1.3333333 1.6666667 0.6666667 0.3333333 0.6666667 1.6666667
##  [71] 2.6666667 1.6666667 1.3333333 1.6666667 2.6666667 3.0000000 2.0000000
##  [78] 1.6666667 2.0000000 3.0000000 2.0000000 1.0000000 0.6666667 1.0000000
##  [85] 2.0000000 1.6666667 0.6666667 0.3333333 0.6666667 1.6666667 2.0000000
##  [92] 1.0000000 0.6666667 1.0000000 2.0000000 3.0000000 2.0000000 1.6666667
##  [99] 2.0000000 3.0000000 4.0000000 3.0000000 2.6666667 3.0000000 4.0000000
## [106] 3.0000000 2.0000000 1.6666667 2.0000000 3.0000000 2.6666667 1.6666667
## [113] 1.3333333 1.6666667 2.6666667 3.0000000 2.0000000 1.6666667 2.0000000
## [120] 3.0000000 4.0000000 3.0000000 2.6666667 3.0000000 4.0000000
```

A. By calculating $Bias[S^2]$ and $Bias[\hat{\sigma}^2]$ check the unbiasedness of these two estimators

We know that the population variance is 2 given from the assigment

$Bias[S^2] = E(S^2) - S^2$
$Bias[S^2] = E(\overline{S^2}) - S^2$
$Bias[S^2] = E(mean(S^2)) - S^2$
$Bias[S^2] = mean(S^2) - S^2$
$Bias[S^2] = mean(S^2) - 2$

```
# Given from the assignment
population_variance = 2
s_bias = mean(s) - population_variance
s_bias
```

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

Simiarly for $Bias[\sigma^2]$
$Bias[\sigma^2] = E(\sigma^2) - \sigma^2$
$Bias[\sigma^2] = E(\overline{\sigma^2}) - \sigma^2$
$Bias[\sigma^2] = E(mean(\sigma^2)) - \sigma^2$
$Bias[\sigma^2] = mean(\sigma^2) - \sigma^2$
$Bias[\sigma^2] = mean(\sigma^2) - 2$

```
# Given from the assignment
population_variance = 2
sigma_bias = mean(sigma) - population_variance
sigma_bias
```

```
## [1] 0
```

From this, we can conclude that the S^2 is bias, and the σ^2 is not bias.

B. By calculating all three components separately check the following identity

Given the equation, $MSE[\hat{\sigma^2}] = var[\hat{\sigma^2}] + (Bias[\hat{\sigma^2}])^2$
$var[\hat{\sigma^2}]$

```
# Given from the assignment
population_variance = 2
sigma_variance = sum((sigma - population_variance)**2)/125
sigma_variance
```

```
## [1] 0.9333333
```

$Bias[\hat{\sigma^2}]^2$

```
sigma_bias_square = sigma_bias**2
sigma_bias_square
```

```
## [1] 0
```

$MSE[\hat{\sigma^2}]$

```
mse = sigma_variance + sigma_bias_square
mse
```

```
## [1] 0.9333333
```

Knowing that $Bias[\hat{\sigma^2}]^2 = 0$, implies $MSE[\hat{\sigma^2}] = var[\hat{\sigma^2}]$
Which decreases as N (sample size) increases, we can conclude MSE is consistent