IT2120 - Probability and Statistics

Lab Sheet 08

Dulash A.K - IT24102769

```r
##Setting the directory
setwd("D:\\Probability and Statistics Y2S1\\IT24102769 Lab 08 PS")

##Importing the data set
data<-read.table ("Data - Lab 8.txt", header=TRUE)
fix(data)
attach (data)

#Question 01
popmn<-mean (Nicotine)
popvar<-var (Nicotine)

#Question 02
#1st create null vectors to store sample data sets.
samples<-c()
n<-c()

for (i in 1:30) {
  s<-sample (Nicotine, 5, replace=TRUE)
  samples<-cbind(samples,s)
  #correction: This ensures column names are 's1', 's2', ..., 's30' instead of all 's1'.
  n<-c(n,paste('s',i,sep=""))
}

#assign column names for each sample created. Names have stored earlier under "n" variable.
colnames (samples)=n

s.means<-apply(samples,2,mean)
s.vars<-apply(samples,2,var)

#Question 03

samplemean<-mean (s.means)
samplevars<-var(s.means)

#Question 04
#Compare the population mean and mean of sample means.
popmn
samplemean

#Question 05
#compare Population varience and vae of sample means
truevar=popvar/5
samplevars
```

Data	
data	40 obs. of 1 variable
samples	num [1:5, 1:30] 1.82 1.69 2.09 2.37 0.85 2.09 1.97 2.31 2.17 1.97 ...
Values	
i	30L
n	chr [1:30] "s1" "s2" "s3" "s4" "s5" "s6" "s7" "s8" "s9" "s10" "s11" "s12" "s13…
popmn	1.77425
popvar	0.152455833333333
s	num [1:5] 2.17 1.09 0.85 1.75 2.55
s.means	Named num [1:30] 1.76 2.1 1.58 1.74 1.75 ...
s.vars	Named num [1:30] 0.3294 0.0207 0.2445 0.2747 0.056 ...
samplemean	1.78486666666667
samplevars	0.029177016091954
truevar	0.0304911666666667

Exercise

1. Calculate the population mean and population standard deviation of the laptop bag weights.

```r
#Question 1

# Set the working directory
setwd("D:\\Probability and Statistics Y2S1\\IT24102769 Lab 08 PS")

#import the dataset
data <- read.table("Exercise - LaptopsWeights.txt", header = TRUE)
fix(data)
attach(data)

#calculate the population mean of laptop bag weights
popmn <- mean(weight.kg.)

#calculate the population variance
popvar <- var(weight.kg.)

popmn
popvar
```

```
> popmn
[1] 2.468
> popvar
[1] 0.06559077
```

2. Draw 25 random samples of size 6 (with replacement) and calculate the sample mean and sample standard deviation for each sample.

```r
#Question 2

#initialize empty containers for samples and sample names
samples <- c()
n <- c()

#loop to draw 25 samples of size 6 with replacement
for(i in 1:25){
  s <- sample(weight.kg., 6, replace = TRUE)   # Sample of size 6
  samples <- cbind(samples, s)
  n <- c(n, paste("s", i))
}

#assign column names to the sample matrix
colnames(samples) <- n

#calculate sample means and variances for each sample
s.means <- apply(samples, 2, mean)
s.vars <- apply(samples, 2, var)

s.means
s.vars
```

```
> s.means
     s 1      s 2      s 3      s 4      s 5      s 6      s 7      s 8      s 9
2.561667 2.498333 2.586667 2.531667 2.563333 2.391667 2.490000 2.363333 2.598333
    s 10     s 11     s 12     s 13     s 14     s 15     s 16     s 17     s 18
2.221667 2.451667 2.505000 2.290000 2.480000 2.596667 2.463333 2.360000 2.363333
    s 19     s 20     s 21     s 22     s 23     s 24     s 25
2.573333 2.640000 2.585000 2.515000 2.361667 2.315000 2.341667
> s.vars
       s 1        s 2        s 3        s 4        s 5        s 6        s 7
0.03337667 0.05505667 0.03198667 0.07225667 0.01370667 0.18793667 0.06656000
       s 8        s 9       s 10       s 11       s 12       s 13       s 14
0.06470667 0.03501667 0.11825667 0.14029667 0.08063000 0.11964000 0.05160000
      s 15       s 16       s 17       s 18       s 19       s 20       s 21
0.05774667 0.07454667 0.17928000 0.04710667 0.02250667 0.03568000 0.01867000
      s 22       s 23       s 24       s 25
0.06055000 0.07497667 0.12683000 0.04969667
```

3. Calculate the mean and standard deviation of the 25 sample means and state the relationship of them with true mean and true standard deviation.

```
# Question 3:
#mean of the 25 sample means
samplesmean <- mean(s.means)

#variance of the 25 sample means
samplevars <- var(s.means)

#display population mean and mean of sample means
popmn
samplesmean

# n = 6 (sample size)
truevar <- popvar / 6

samplevars
truevar
```

```
> popmn
[1] 2.468
> samplesmean
[1] 2.465933
> # n = 6 (sample size)
> truevar <- popvar / 6
> samplevars
[1] 0.01298567
> truevar
[1] 0.01093179
```

data	40 obs. of 1 variable
samples	num [1:6, 1:25] 2.28 2.75 2.43 2.71 2.53 2.67 2.57 2.89 2.41 2.32 ...
Values	
i	25L
n	chr [1:25] "s 1" "s 2" "s 3" "s 4" "s 5" "s 6" "s 7" "s 8" "s 9" "s 10" "s 11" "s 12" "s 13" "s 14" "s 15" "s 16" "s 1…
popmn	2.468
popvar	0.0655907692307692
s	num [1:6] 2.2 2.7 2.05 2.32 2.32 2.46
s.means	Named num [1:25] 2.56 2.5 2.59 2.53 2.56 ...
s.vars	Named num [1:25] 0.0334 0.0551 0.032 0.0723 0.0137 ...
samplemean	1.78486666666667
samplesmean	2.46593333333333
samplevars	0.0129856666666667
truevar	0.0109317948717949