Sri Lanka Institute of Information Technology

Lab Submission

Lab sheet No 08

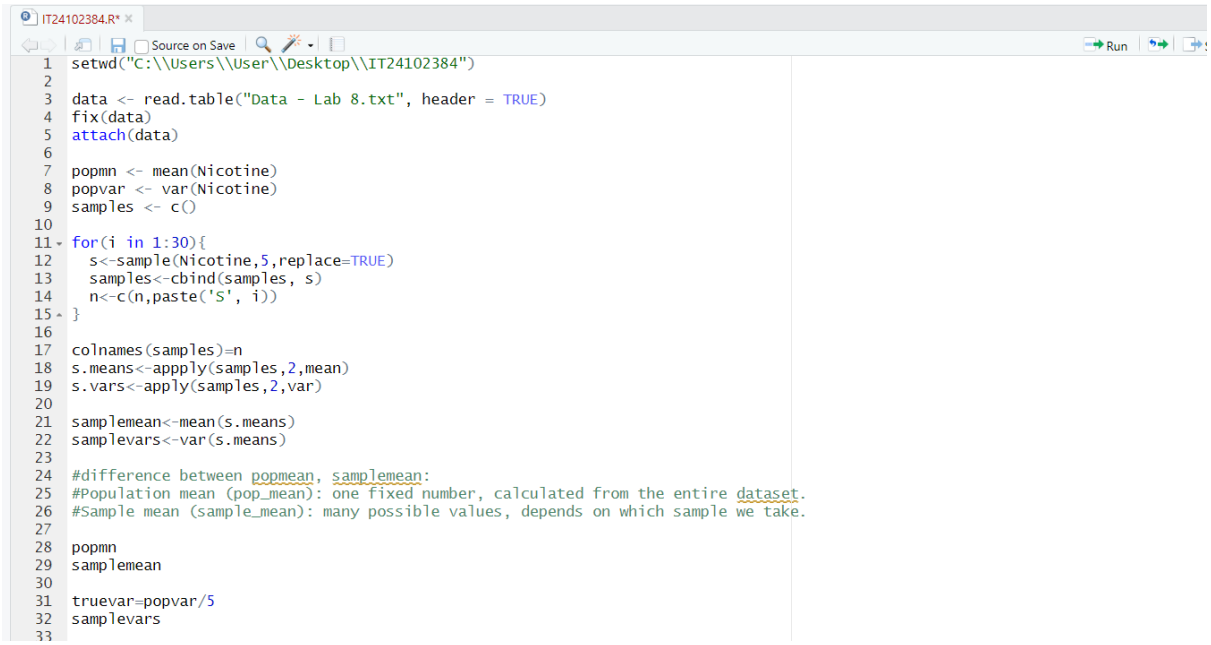
IT24102384

Asjath Haamees M.N.M

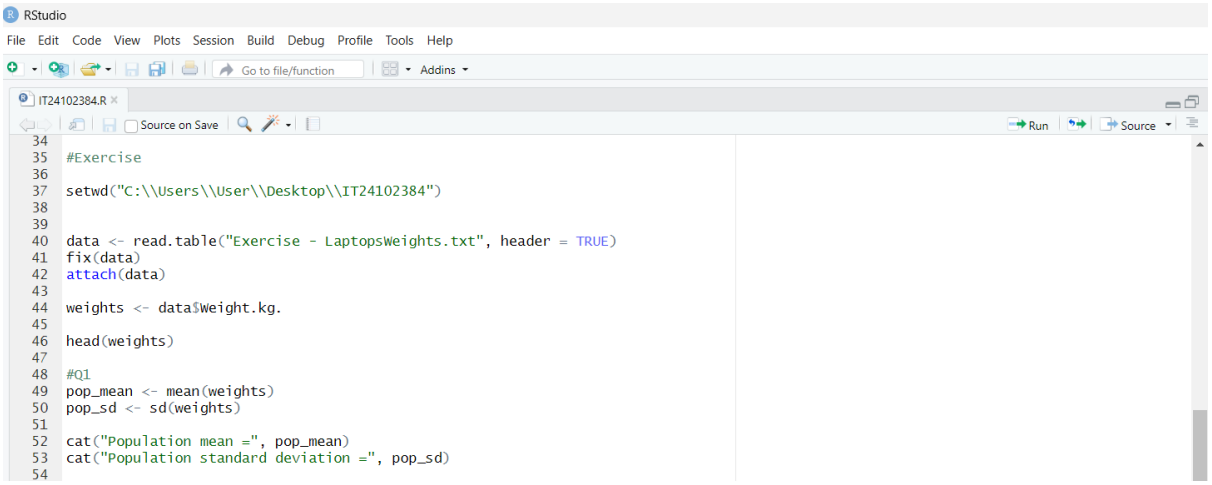
Probabilities and Statistics | IT2120

B.Sc. (Hons) in Information Technology

Exercise



```r
setwd("C:\\Users\\User\\Desktop\\IT24102384")

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

popmn <- mean(Nicotine)
popvar <- var(Nicotine)
samples <- c()

for(i in 1:30){
  s<-sample(Nicotine,5,replace=TRUE)
  samples<-cbind(samples, s)
  n<-c(n,paste('S', i))
}

colnames(samples)=n
s.means<-appply(samples,2,mean)
s.vars<-apply(samples,2,var)

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

#difference between popmean, samplemean:
#Population mean (pop_mean): one fixed number, calculated from the entire dataset.
#Sample mean (sample_mean): many possible values, depends on which sample we take.

popmn
samplemean

truevar=popvar/5
samplevars
```



```r
#Exercise

setwd("C:\\Users\\User\\Desktop\\IT24102384")


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

weights <- data$Weight.kg.

head(weights)

#Q1
pop_mean <- mean(weights)
pop_sd <- sd(weights)

cat("Population mean =", pop_mean)
cat("Population standard deviation =", pop_sd)
```

```r
#Q2
set.seed(123)
num_samples <- 25
sample_size <- 6

samples <- matrix(nrow = sample_size, ncol = num_samples)

for (i in 1:num_samples) {
  samples[, i] <- sample(weights, size = sample_size, replace = TRUE)
}

colnames(samples) <- paste("Sample", 1:num_samples, sep = "_")

sample_means <- apply(samples, 2, mean)
sample_sds <- apply(samples, 2, sd)

head(sample_means)
head(sample_sds)

#Q3
mean_of_sample_means <- mean(sample_means)
sd_of_sample_means <- sd(sample_means)

cat("Mean of sample means =", mean_of_sample_means)
cat("Standard deviation of sample means =", sd_of_sample_means)
```

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

```r
> setwd("C:\\Users\\User\\Desktop\\IT24102384")
> data <- read.table("Exercise - LaptopsWeights.txt", header = TRUE)
> fix(data)
> attach(data)
The following object is masked from data (pos = 3):

    Weight.kg.

The following object is masked from data (pos = 4):

    Weight.kg.

The following object is masked from data (pos = 5):

    Weight.kg.

The following object is masked from data (pos = 8):

    Weight.kg.

The following object is masked from data (pos = 9):

    Weight.kg.

> weights <- data$Weight.kg.
> head(weights)
[1] 2.46 2.45 2.47 2.71 2.46 2.05
> #Q1
> pop_mean <- mean(weights)
> pop_sd <- sd(weights)
> cat("Population mean =", pop_mean)
Population mean = 2.468
> cat("Population standard deviation =", pop_sd)
Population standard deviation = 0.2561069
```

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

```
> #Q2
> set.seed(123)
> num_samples <- 25
> sample_size <- 6
> samples <- matrix(nrow = sample_size, ncol = num_samples)
> for (i in 1:num_samples) {
+   samples[, i] <- sample(weights, size = sample_size, replace = TRUE)
+ }
> colnames(samples) <- paste("Sample", 1:num_samples, sep = "_")
> sample_means <- apply(samples, 2, mean)
> sample_sds <- apply(samples, 2, sd)
> head(sample_means)
Sample_1 Sample_2 Sample_3 Sample_4 Sample_5 Sample_6
2.530000 2.573333 2.473333 2.591667 2.456667 2.401667
> head(sample_sds)
 Sample_1  Sample_2  Sample_3  Sample_4  Sample_5  Sample_6
0.1513935 0.1191078 0.1718914 0.1345239 0.2749303 0.2544340
```

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

```
> #Q3
> mean_of_sample_means <- mean(sample_means)
> sd_of_sample_means <- sd(sample_means)
> cat("Mean of sample means =", mean_of_sample_means)
Mean of sample means = 2.4668
> cat("Standard deviation of sample means =", sd_of_sample_means)
Standard deviation of sample means = 0.07624874
```