IT2120- Probability and Statistics

Lab Sheet 08

Dushmanthi W. D. H.

IT24103279

Exercise

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

IT24103279.R

Source on Save

```r
setwd("C:\\Users\\Hiruni\\Desktop\\IT24103279")

data <- read.table("Exercise - LaptopsWeights.txt", header = TRUE)
# R converts "Weight(kg)" to "Weight.kg."
laptop_weights <- data$Weight.kg.

n_population <- length(laptop_weights)
population_mean <- mean(laptop_weights)
# Adjust R's sample sd to get the population sd
population_sd <- sd(laptop_weights) * sqrt((n_population - 1) / n_population)

print(paste("Population Mean:", population_mean))
print(paste("Population Standard Deviation:", population_sd))
```

```
> setwd("C:\\Users\\Hiruni\\Desktop\\IT24103279")
>
> data <- read.table("Exercise - LaptopsWeights.txt", header = TRUE)
> # R converts "Weight(kg)" to "Weight.kg."
> laptop_weights <- data$Weight.kg.
> n_population <- length(laptop_weights)
> population_mean <- mean(laptop_weights)
> # Adjust R's sample sd to get the population sd
> population_sd <- sd(laptop_weights) * sqrt((n_population - 1) / n_population)
>
> print(paste("Population Mean:", population_mean))
[1] "Population Mean: 2.468"
> print(paste("Population Standard Deviation:", population_sd))
[1] "Population Standard Deviation: 0.252885349516337"
```

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

```r
samples <- c()
number_of_samples <- 25
sample_size <- 6

for (i in 1:number_of_samples) {
  s <- sample(laptop_weights, sample_size, replace = TRUE)
  samples <- cbind(samples, s)
}
colnames(samples) <- paste("Sample", 1:number_of_samples)

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

print(sample_means)
print(sample_sds)
```

```
> samples <- c()
> number_of_samples <- 25
> sample_size <- 6
>
> for (i in 1:number_of_samples) {
+   s <- sample(laptop_weights, sample_size, replace = TRUE)
+   samples <- cbind(samples, s)
+ }
> colnames(samples) <- paste("Sample", 1:number_of_samples)
>
> sample_means <- apply(samples, 2, mean)
> sample_sds <- apply(samples, 2, sd)
>
> print(sample_means)
 Sample 1  Sample 2  Sample 3  Sample 4  Sample 5  Sample 6  Sample 7  Sample 8  Sample 9 Sample 10
 2.521667  2.456667  2.413333  2.508333  2.493333  2.431667  2.445000  2.581667  2.355000  2.435000
Sample 11 Sample 12 Sample 13 Sample 14 Sample 15 Sample 16 Sample 17 Sample 18 Sample 19 Sample 20
 2.495000  2.503333  2.466667  2.471667  2.295000  2.416667  2.468333  2.388333  2.350000  2.298333
Sample 21 Sample 22 Sample 23 Sample 24 Sample 25
 2.460000  2.613333  2.531667  2.470000  2.383333
> print(sample_sds)
 Sample 1  Sample 2  Sample 3  Sample 4  Sample 5  Sample 6  Sample 7  Sample 8  Sample 9 Sample 10
0.3048552 0.3074844 0.3366700 0.2144217 0.2067527 0.2847045 0.1827293 0.1658212 0.1211198 0.2234055
Sample 11 Sample 12 Sample 13 Sample 14 Sample 15 Sample 16 Sample 17 Sample 18 Sample 19 Sample 20
0.2847279 0.1672922 0.2924152 0.1954908 0.2285388 0.2258908 0.3343302 0.3703197 0.1938040 0.3996207
Sample 21 Sample 22 Sample 23 Sample 24 Sample 25
0.2084226 0.2406380 0.2069219 0.3234811 0.2018580
```

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.

```r
mean_of_sample_means <- mean(sample_means)
sd_of_sample_means <- sd(sample_means) # This is the Standard Error

print(paste("Mean of the 25 Sample Means:", mean_of_sample_means))
print(paste("Standard Deviation of the 25 Sample Means (Standard Error):", sd_of_sample_means))

# Compare with population parameters
theoretical_standard_error <- population_sd / sqrt(sample_size)

print(paste("Population Mean:", population_mean))
print(paste("Theoretical Standard Error (Population SD / sqrt(n)):", theoretical_standard_error))
```

```
> mean_of_sample_means <- mean(sample_means)
> sd_of_sample_means <- sd(sample_means) # This is the Standard Error
>
> print(paste("Mean of the 25 Sample Means:", mean_of_sample_means))
[1] "Mean of the 25 Sample Means: 2.45013333333333"
> print(paste("Standard Deviation of the 25 Sample Means (Standard Error):", sd_of_sample_means))
[1] "Standard Deviation of the 25 Sample Means (Standard Error): 0.0777116584036903"
>
> # Compare with population parameters
> theoretical_standard_error <- population_sd / sqrt(sample_size)
>
> print(paste("Population Mean:", population_mean))
[1] "Population Mean: 2.468"
> print(paste("Theoretical Standard Error (Population SD / sqrt(n)):", theoretical_standard_error))
[1] "Theoretical Standard Error (Population SD / sqrt(n)): 0.103240011623401"
>
```

Environment History Connections Tutorial

Import Dataset 106 MiB

R Global Environment

Data	
data	40 obs. of 1 variable
samples	num [1:6, 1:25] 2.76 2.85 2.05 2.7 2.45 2.32 2.43 2.2 2.46 2.85 ...
Values	
i	25L
laptop_weights	num [1:40] 2.46 2.45 2.47 2.71 2.46 2.05 2.6 2.42 2.43 2.53 ...
mean_of_sample_means	2.45013333333333
n_population	40L
number_of_samples	25
population_mean	2.468
population_sd	0.252885349516337
s	num [1:6] 2.05 2.43 2.23 2.51 2.51 2.57
sample_means	Named num [1:25] 2.52 2.46 2.41 2.51 2.49 ...
sample_sds	Named num [1:25] 0.305 0.307 0.337 0.214 0.207 ...
sample_size	6
sd_of_sample_means	0.0777116584036903
theoretical_standard_error	0.103240011623401