

# IT2120 - Probability and Statistics

## Lab Sheet 08

### IT24103526 - Senaratne P.A.R.T.

#### Exercise

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

The first screenshot shows the R Studio interface with the following code in the script editor:

```
1 setwd("C:\\Users\\HP\\Desktop\\IT24103526")
2
3 ##Importing dataset
4 data <- read.table("Exercise - Laptopsweights.txt", header = TRUE)
5 fix(data)
6 attach(data)
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31 (Top Level) z
```

The Environment pane on the right shows the following variables:

Variable	Value
iq_95th	124.672804404272
p_above_130	0.0227501319481792
p_at_most_2	0.486582880967408
p_between_10_25	0.375

The Data Editor pane shows a table with 40 rows and 6 columns:

Weight.kg	var2	var3	var4	var5	var6
22	2.06				
23	2.41				
24	2.45				
25	2.76				
26	2.43				
27	2.61				
28	2.57				
29	2.73				
30	2.17				
31	2.67				
32	2.05				
33	1.71				
34	2.32				
35	2.23				
36	2.76				
37	2.7				
38	2.13				
39	2.75				
40	2.2				

The second screenshot shows the R Studio interface with the following code in the script editor:

```
1 setwd("C:\\Users\\HP\\Desktop\\IT24103526")
2
3 ##Importing dataset
4 data <- read.table("Exercise - Laptopsweights.txt", header = TRUE)
5 fix(data)
6 attach(data)
7
8 ##Q1
9 popmn <- mean(weight.kg.)
10 popsd <- sd(weight.kg.)
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31 (Top Level) z
```

The Environment pane on the right shows the following variables:

Variable	Value
iq_95th	124.672804404272
p_above_130	0.0227501319481792
p_at_most_2	0.486582880967408
p_between_10_25	0.375
popmn	2.468
popsd	0.256106948813907

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

The screenshot shows an RStudio session with the following code in the script editor:

```

12
13
14 #RQ2
15 #First null vector to store sample data sets.
16 samples <- c()
17 n <- c()
18
19
20 for(i in 1:25){
21   s <- sample(weight.kg., 6, replace = TRUE)
22   samples <- cbind(samples, s)
23   n <- c(n, paste('S', i))
24 }
25
26 #Assign column names for each sample created
27 colnames(samples) = n
28
29 # considered the second argument as "2" we can calculate either mean/variance column wise
30 s.means <- apply(samples, 2, mean)
31 s.stds <- apply(samples, 2, sd)
32
33 |
34
35
36
37
38 |
39 |
40 |
41 |
42 |
43 |
44 |
45 |
46 |
47 |
48 |
49 |
50 |
51 |
52 |
53 |
54 |
55 |
56 |
57 |
58 |
59 |
60 |
61 |
62 |
63 |
64 |
65 |
66 |
67 |
68 |
69 |
70 |
71 |
72 |
73 |
74 |
75 |
76 |
77 |
78 |
79 |
80 |
81 |
82 |
83 |
84 |
85 |
86 |
87 |
88 |
89 |
90 |
91 |
92 |
93 |
94 |
95 |
96 |
97 |
98 |
99 |
100 |

```

The Environment pane shows the following objects:

Object	Class	Attributes
data	data.frame	40 obs. of 1 variable
samples	matrix	num [1:6, 1:25] 2.47 2.46 2.6 2.32 2.53 2.85 2.57 2.7 2.85 2.05...
n	character	1:25
weight.kg.	numeric	1:6
s.means	numeric	1:25
s.stds	numeric	1:25

The Console shows the execution of the code, with the following output:

```

> #RQ2
> #First null vector to store sample data sets.
> samples <- c()
> n <- c()
> for(i in 1:25){
+   s <- sample(weight.kg., 6, replace = TRUE)
+   samples <- cbind(samples, s)
+   n <- c(n, paste('S', i))
+ }
> #Assign column names for each sample created
> colnames(samples) = n
> # considered the second argument as "2" we can calculate either mean/variance column wise
> s.means <- apply(samples, 2, mean)
> s.stds <- apply(samples, 2, sd)
>

```

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.

The screenshot shows an RStudio session with the following code in the script editor:

```

34
35 #RQ3
36 samplemean <- mean(s.means)
37 samplestd <- sd(s.means)
38
39 #compare the population mean and mean of sample means.
40 popmean
41
42
43 #compare the population std and std of sample means.
44 truesd = popsd/6
45 samplestd
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```

The Environment pane shows the following objects:

Object	Class	Attributes
data	data.frame	40 obs. of 1 variable
samples	matrix	num [1:6, 1:25] 2.47 2.46 2.6 2.32 2.53 2.85 2.57 2.7 2.85 2.05...
n	character	1:25
weight.kg.	numeric	1:6
s.means	numeric	1:25
s.stds	numeric	1:25
samplemean	numeric	1
samplestd	numeric	1
truesd	numeric	1

The Console shows the execution of the code, with the following output:

```

> #RQ3
> samplemean <- mean(s.means)
> samplestd <- sd(s.means)
> #compare the population mean and mean of sample means.
> popmean
[1] 2.468
> samplemean
[1] 2.447133
> #compare the population std and std of sample means.
> truesd = popsd/6
> samplestd
[1] 0.09953424
>

```

