Homework 1

Ranjitha Bukkapattama Krishnaiah(24221339)

Introduction

The *sleep* dataset is a built-in R dataset from the `datasets` package, commonly used for analyzing paired samples. It records the *increase in hours of sleep* observed in 10 patients after taking two different drugs. Each patient received both treatments, allowing comparison of drug effects within individuals. The dataset includes three variables: `extra` (sleep increase), `group` (drug: 1 or 2), and `ID` (patient identifier). It is often used to demonstrate statistical methods such as paired t-tests and visualization of treatment effects.

Overview of Dataset

In this report we will utilize the `sleep` dataset to prepare a series of plots in this report. For more information about the dataset, we can read the help file containing the following description for the dataset.

The sleep data, part of base R's datasets package, is often employed to illustrate paired sample testing. Data is reported about gain in hours of sleep for 10 patients taking two different drugs. Both drugs were taken by each patient, which allows for direct comparison between their impacts. There are three variables in the data: extra (gain in sleep), group (drug 1 or drug 2), and ID (id for each patient). It is well suited to show paired t-tests and comparative plots for the effect of treatments.

```
'data.frame':   20 obs. of  3 variables:
 $ extra: num  0.7 -1.6 -0.2 -1.2 -0.1 3.4 3.7 0.8 0 2 ...
 $ group: Factor w/ 2 levels "1","2": 1 1 1 1 1 1 1 1 1 1 ...
 $ ID   : Factor w/ 10 levels "1","2","3","4",..: 1 2 3 4 5 6 7 8 9 10 ...
```

Data Summary

Below are descriptive statistics for the increase in sleep (extra hours) in each drug group:

```
# A tibble: 2 x 5
  group  Mean    SD   Min   Max
  <fct> <dbl> <dbl> <dbl> <dbl>
1 1      0.75  1.79  -1.6   3.7
2 2      2.33  2.00  -0.1   5.5
```

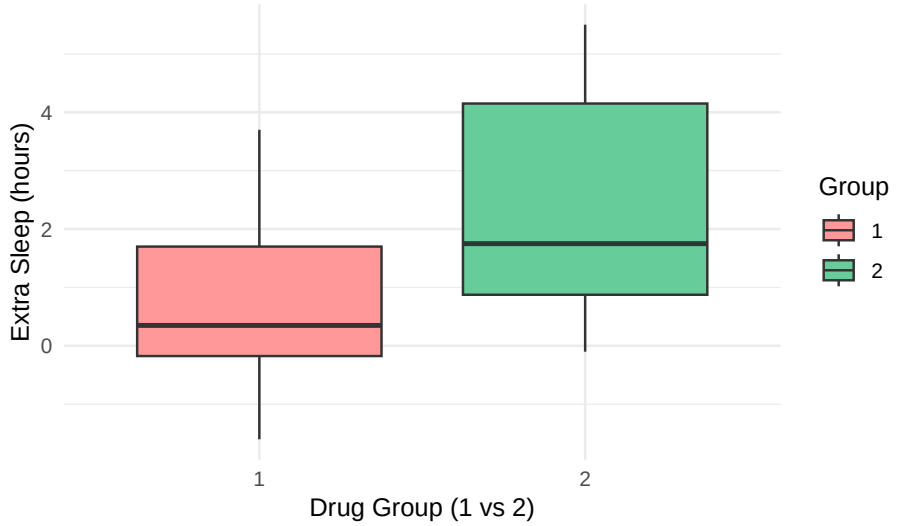
- Drug 2 shows a higher average increase in sleep (2.33 hours) compared to Drug 1 (0.75 hours), suggesting it may be more effective overall.

- Drug 1 has a wider negative range (min = -1.6),indicating some participants experienced a decrease in sleep.
- Both drugs show variability in responses, with Drug 2 having a slightly higher standard deviation (2.00 vs. 1.79).

Plot

The following boxplot (Figure 1) visualizes the sleep increase for both drugs. This helps in comparing their distributions and central tendencies.

Figure 1: Sleep Increase by Drug Group



- Drug Group 2 shows a higher median sleep increase than Group 1, indicating it may be more effective overall.
- The spread of values (IQR) and Highest recorded sleep gain are both larger for Group 2, suggesting greater variability and potential for higher effectiveness.
- Group 1 has a lower median and includes some negative values, showing that it was less consistent and less effective for some individuals.