Hypothesis Testing on South Western Sydney Hospital

Christian Themin

Problem Statement

It has been claimed that the South Western Sydney hospitals have an average length of stay (ALOS) of 4.5 days. Investigator wants to check by performing one-sample t-test on the data that was collected from Australian Institute of Health and Welfare website.

The hypothesis test is set as: H0:mu = 4.5 Ha:mu != 4.5

Load Packages

```
library(readxl)
library(magrittr)
library(dplyr)
library(lattice)
library(ggplot2)
library(car)
```

Data

```
# Read the file
Hospital <- read_xlsx('average-length-of-stay-multilevel-data.xlsx', skip =12)

# Change the datatype of ALOS
Hospital$`Average length of stay (days)` <- Hospital$`Average length of stay (days)` %>% as.numeric()
```

Descriptive Statistics

```
# Descriptive Statistics

Hospital %>%  filter(`Local Hospital Network (LHN)`=="South Western Sydney") %>%  summarise(
                Min = min(`Average length of stay (days)`, na.rm = TRUE),
                Q1 = quantile(`Average length of stay (days)`,probs = .25,na.rm = TRUE),
                Median = median(`Average length of stay (days)`, na.rm = TRUE),
                Mean = mean(`Average length of stay (days)`, na.rm = TRUE),
                Q3 = quantile(`Average length of stay (days)`,probs = .75,na.rm = TRUE),
                Max = max(`Average length of stay (days)`,na.rm = TRUE),
                SD = sd(`Average length of stay (days)`, na.rm = TRUE),
                n = n(),
                Missing = sum(is.na(`Average length of stay (days)`)))
```

Min <dbl>	Q1 <dbl>	Median <dbl>	Mean <dbl>	Q3 <dbl>	Max <dbl>	SD <dbl>	n <int>	Missing <int>
1.3	2.9	3.8	4.643396	5.5	11.9	2.434411	499	128

1 row

Boxplot

When performing boxplot to show the relationship between South Western Sydney hopitals and ALOS, there seems to be many outliers in the data. However, we do not need to worry about this as the outliers appears to be real and we need to keep it.

```
#Filtering SWS
SWS <- Hospital %>% filter(`Local Hospital Network (LHN)`=="South Western Sydney")

# Boxplot
SWS$`Average length of stay (days)` %>% boxplot
```
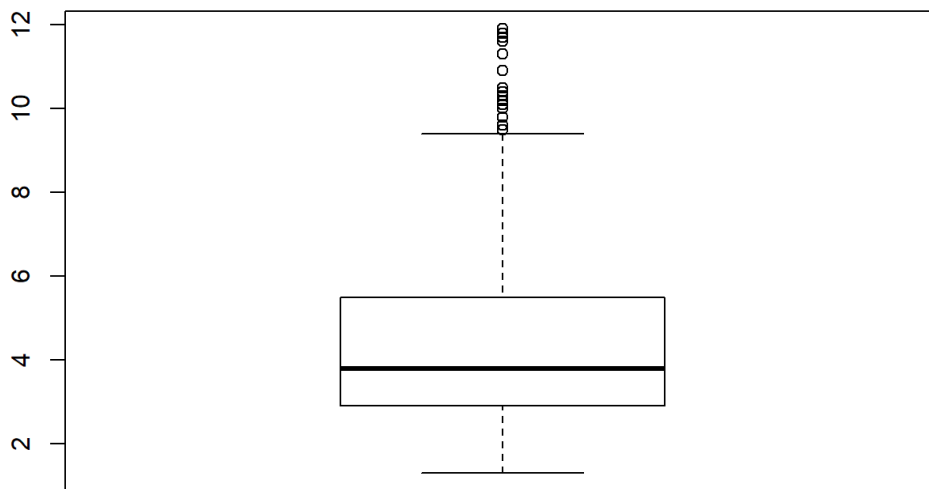
One-sample t-test

We will perform one-sample t-test to investigate whether the mean of 4.5 days is captured and check the p-value.

```
# t-test
t.test(SWS$`Average length of stay (days)`, mu = 4.5, alternative="two.sided", na.rm=TRUE)
```

```
##
##  One Sample t-test
##
## data:  SWS$`Average length of stay (days)`
## t = 1.1346, df = 370, p-value = 0.2573
## alternative hypothesis: true mean is not equal to 4.5
## 95 percent confidence interval:
##   4.394867 4.891926
## sample estimates:
## mean of x
##   4.643396
```

```
# t-critical
qt(p=0.05, df=499-1, lower.tail=TRUE)
```

```
## [1] -1.647919
```

Conclusion

A two-tailed, one-sample t-test was used to determine if the mean of South Western Sydney in Average Length Of Stay (ALOS) is significantly different from the assumed mean of 4.5 days.

The result of the hypothesis test was not statistically significant. The estimated mean is 4.643, p-value > 0.001, 95% CI[4.395, 4.892] captured the Ho mean of 4.5. The t critical value is ± 1.648.