Assignment 3: Design of experiments
TMA4267 Linear statistical models spring 2020

Silje Anfindsen

15 April, 2020

Introduction

We have all experienced the annoying feeling when the internet speed is slow. Maybe it takes ages to download the 2-minutes-short video on YouTube, or it does not download at all. Then you maybe try to move your computer closer to the router, or ask your sister to turn off the wifi on her phone to give your device a greater part of the internet in the room. So, what is it that makes the internet speed, more specific the download-speed fast or slow? Which factors play a big role in this experiment, and which do not? In this study I want to investigate and identify which factors contribute the most to a slow download speed on my computer.

Selection of factors and levels:

As already discussed, we all have an idea of which factors that are relvant to the download speed. I have chosen to investigate four factors, some of them which can seem obvious, while others travel around the internet as unsolved myths. I have also chosen two different levels for each factor. High level is denoted by *+1*, and low level as *-1*. I tried to find levels giving the factors enough room to possibly variate enough for me to observe it. The table below displays these factors together with the two levels.

Factor	Level: +1	Level: -1
A	Busy computer	No program running
B	Blutooth devices nearby	No bluetooth nearby
C	Distance: 5m	Negligble distance
D	Obstacle around router	No obstacles

A. I am going to run a heavy program, more explicit a numerical scheme running in Python at the same time as I measure the download speed and compare it with no program running on the computer (except the speedometer).

B. I will investigate if the effect of bluetooth devices near the router will affect the internet speed. I will use phone with bluetooth as well as headsets connected to bluetooth.

C. Here we will mesaure the internet speed for the device being in about 5 meter distance from the router vs. the device being as close as possible.

D. The last factor is meant to represent possible objects blocking the signal. Therefore I will put the router inside a pot made of iron and measure the signal compared to no pot blocking it.

I expect the increased distance from the router (factor C) will interact with the obstacle around the router (factor D). I find it quite easy to check that the factors are at the desired level as everything is visible. The most challenging factor has to be the busy computer (factor A), as it is hard to define a busy computer. It is

also hard to know when the computer is struggeling the most, and when it is not struggeling at all. In this case I will try to notice when the computer is most overheated in orer to define a busy computer.

Selection of response variable:

In order to investigate the internet speed I will measure the download speed for the network in my home. To do this, I will use my own laptop, a 1.5-year old Macbook Pro as device. The download speed is measured using the public webpage: speedtest.net. The unit is Megabits per second, Mbps. According to an article on allconnect.com, [1] a good internet speed is 25 Mbps, or above. Fast internet speeds that make your internet connection able to support multiple devices at once is in the 100+ Mbps range.

In this case I have chosen to measure the download speed as this is what we usually exploit when being on the internet; watching a video, uploading a webpage etc. The response variable could also be the upload speed if we were interested in how fast we can upload our photo album to icloud for example. In this experiment I will use the 5 GHz frequency, but we could also have measured the internet speed using 2,4 GHz, which often has better coverage but less speed. These are the two usual frequency available in most households.

The speedometer I have found is free and probably not the most advanced and precise instrument. On the other hand, it uses several seconds to find the average of the download speed.

Choice of design:

My experiment consists of 4 factors and as the costs related to the experiment are low, I have the possibility to perform a full 2^4 factorial design. As I am familiar with the fact that the network signal and speed is quite sensitive to the distance from the router, I have chosen to do my experiment in two blocks with distance from router as the block factor. The two possible levels are right beside the router and approximatly 5 meters away.

Replicates. . .

Desired resolution of the design? Infinity for full factorial. But for blocking?

Implementation of the experiment:

Randomization helps to avoid systematic error due to the order. Therefore I have made sure to randomize the runs in R before doing the experiments.

o Describe any problems with the implementation (maybe the randomization was not followed?).

o Is each experiments a genuine run replicate, that is reflects the total variability of the experiment? (Each trial should be a performed independently and constitute a full trial.)

```r
# Import SixSigma package
library(SixSigma)

# Design the experiment (2^3)
ExperimentDesign <- expand.grid(A = gl(2, 1, labels = c("-", "+")),
                                B = gl(2, 1, labels = c("-", "+")),
                                C = gl(2, 1, labels = c("-", "+")))
ExperimentDesign
```

```
##   A B C
## 1 - - -
## 2 + - -
```

```
## 3 - + -
## 4 + + -
## 5 - - +
## 6 + - +
## 7 - + +
## 8 + + +
```

```r
# Randomize the experiment
ExperimentDesign$ord <- sample(1:8, 8)
ExperimentDesign[order(ExperimentDesign$ord), ]
```

```
##   A B C ord
## 4 + + -   1
## 5 - - +   2
## 2 + - -   3
## 3 - + -   4
## 1 - - -   5
## 7 - + +   6
## 6 + - +   7
## 8 + + +   8
```

```r
# Create replicates
ss.data.doe1 <- data.frame(repl = rep(1:2, each = 8),
                           rbind(ExperimentDesign))
ss.data.doe1
```

```
##    repl A B C ord
## 1     1 - - -   5
## 2     1 + - -   3
## 3     1 - + -   4
## 4     1 + + -   1
## 5     1 - - +   2
## 6     1 + - +   7
## 7     1 - + +   6
## 8     1 + + +   8
## 9     2 - - -   5
## 10    2 + - -   3
## 11    2 - + -   4
## 12    2 + + -   1
## 13    2 - - +   2
## 14    2 + - +   7
## 15    2 - + +   6
## 16    2 + + +   8
```

```r
# Add responses
ss.data.doe1$response <- c(5.33, 6.99, 4.23, 6.61,
                           2.26, 5.75, 3.26, 6.24,
                           5.7, 7.71, 5.13, 6.76,
                           2.79, 4.57, 2.48, 6.18)
ss.data.doe1
```

```
##    repl A B C ord response
```

```
## 1     1 - - -   5     5.33
## 2     1 + - -   3     6.99
## 3     1 - + -   4     4.23
## 4     1 + + -   1     6.61
## 5     1 - - +   2     2.26
## 6     1 + - +   7     5.75
## 7     1 - + +   6     3.26
## 8     1 + + +   8     6.24
## 9     2 - - -   5     5.70
## 10    2 + - -   3     7.71
## 11    2 - + -   4     5.13
## 12    2 + + -   1     6.76
## 13    2 - - +   2     2.79
## 14    2 + - +   7     4.57
## 15    2 - + +   6     2.48
## 16    2 + + +   8     6.18
```

Analysis of data:

Conclusion and recommendations:

References

[1] Anders, David (2019), Internet speed classifications: What's fast, what's slow and what is a good internet speed?