

# **Stochastics Lab Course II**

Khvam Tabougua Trevor

March 2019

# Introduction

The "Stochastics Lab course II" is an Introductory Course for statistics and stochastics applications with R programming language. The course lasted for two weeks in March 2019. The report contains results, interpretations and figures from the ten exercises that had to be solved. Along with this report, there is also the R codes, which are recommended to understand the result.

# Contents

<b>1</b>	<b>Tidyverse</b>	<b>4</b>
1.1	Problem's description . . . . .	4
1.2	Methods' description . . . . .	4
1.3	Results' discussion . . . . .	4
<b>2</b>	<b>Random number generation</b>	<b>5</b>
2.1	Problem's description . . . . .	5
2.2	Methods' description . . . . .	5
2.3	Results' discussion . . . . .	5
<b>3</b>	<b>Bootstrap</b>	<b>6</b>
3.1	Problem's description . . . . .	6
3.2	Methods' description . . . . .	6
3.3	Results' discussion . . . . .	6
<b>4</b>	<b>Generalised linear models</b>	<b>7</b>
4.1	Problem's description . . . . .	7
4.2	Methods' description . . . . .	7
4.3	Results' discussion . . . . .	7
<b>5</b>	<b>Survival analysis</b>	<b>8</b>
5.1	Problem's description . . . . .	8
5.2	Methods' description . . . . .	8
5.3	Results' discussion . . . . .	8
<b>6</b>	<b>Kernel density estimation</b>	<b>9</b>
6.1	Problem's description . . . . .	9
6.2	Methods' description . . . . .	9
6.3	Results' discussion . . . . .	9

<b>7</b>	<b>Nonparametric regression: local polynomials</b>	<b>10</b>
7.1	Problem's description . . . . .	10
7.2	Methods' description . . . . .	10
7.3	Results' discussion . . . . .	10
<b>8</b>	<b>Nonparametric regression: splines</b>	<b>11</b>
8.1	Problem's description . . . . .	11
8.2	Methods' description . . . . .	11
8.3	Results' discussion . . . . .	11
<b>9</b>	<b>Mixed models</b>	<b>12</b>
9.1	Problem's description . . . . .	12
9.2	Methods' description . . . . .	12
9.3	Results' discussion . . . . .	12
<b>10</b>	<b>Partial least squares</b>	<b>13</b>
10.1	Problem's description . . . . .	13
10.2	Methods' description . . . . .	13
10.3	Results' discussion . . . . .	13

# Chapter 1

# Tidyverse

1.1 Problem's description

1.2 Methods' description

1.3 Results' discussion

## Chapter 2

# Random number generation

### 2.1 Problem's description

### 2.2 Methods' description

Linear congruent generators: Give the algo/pseudo code. Give an exemple (with a fuul period), the drawbacks of the method. Talk a little bit about multiplicative congruent generator, then Mersenne twister. Inverse method: rejection method (Accept-Reject)

### 2.3 Results' discussion

## Chapter 3

# Bootstrap

### 3.1 Problem's description

### 3.2 Methods' description

Bootstrap: algorithm: Bootstrap confidence intervals:

### 3.3 Results' discussion

## Chapter 4

# Generalised linear models

4.1 Problem's description

4.2 Methods' description

4.3 Results' discussion



## Chapter 5

# Survival analysis

### 5.1 Problem's description

We want to analyze data where the outcome variable is the time until the occurrence of an event of interest. The event can be death, occurrence of a disease, marriage, divorce, etc. The time to event or survival time can be measured in days, weeks, years, etc. For example, if the event of interest is heart attack, then the survival time can be the time in years until a person develops a heart attack. subjects are usually followed over a specified time period and the focus is on the time at which the event of interest occurs. Why not use linear regression to model the survival time as a function of a set of predictor variables? First, survival times are typically positive numbers; ordinary linear regression may not be the best choice unless these times are first transformed in a way that removes this restriction. Second, and more importantly, ordinary linear regression cannot effectively handle the censoring of observations. Why not compare proportion of events in your groups using risk/odds ratios or logistic regression? Simply because it ignores time.

To tackle these issues, we'll use some survival analysis methods.

### 5.2 Methods' description

### 5.3 Results' discussion

## Chapter 6

# Kernel density estimation

6.1 Problem's description

6.2 Methods' description

6.3 Results' discussion

## Chapter 7

# Nonparametric regression: local polynomials

7.1 Problem's description

7.2 Methods' description

7.3 Results' discussion

## Chapter 8

# Nonparametric regression: splines

8.1 Problem's description

8.2 Methods' description

8.3 Results' discussion

## Chapter 9

# Mixed models

9.1 Problem's description

9.2 Methods' description

9.3 Results' discussion

## Chapter 10

# Partial least squares

10.1 Problem's description

10.2 Methods' description

10.3 Results' discussion