

TBD

TBA

Daniel Johnsen, Gonde Winkelmann, Linda Knudsen
Morten Thygesen, Niklas Andersen, Salma Hirsi

Datavidenskab & Machine Learning, cs-23-dvml-2-02, March 16, 2023

P2 projekt



Copyright © Aalborg University 2015

Here you can write something about which tools and software you have used for typesetting the document, running simulations and creating figures. If you do not know what to write, either leave this page blank or have a look at the colophon in some of your books.



Elektronik og IT
Aalborg Universitet
<http://www.aau.dk>

AALBORG UNIVERSITET

STUDENTERRAPPORT

Titel:

TBD

Abstract:

Her er resuméet

Tema:

Fra data til videnskab

Projektperiode:

Forårssemester 2023

Projektgruppe:

cs-23-dvml-2-02

Deltager(e):

Daniel Elsborg Johnsen
Gonde Leon Winkelmann
Linda Knudsen
Morten Tejlmand Thygesen
Niklas Kjærgaard Andersen
Salma Ahmed Hassan Hirsi

Vejleder(e):

Søren Byg Vilsen

Oplagstal: 0

Sidetal: 10

Afleveringsdato:

16. marts 2023

Rapportens indhold er frit tilgængeligt, men offentliggørelse (med kildeangivelse) må kun ske efter aftale med forfatterne.

Contents

Preface	v
1 Introduction	1
1.1 Examples	1
1.2 How Does Sections, Subsections, and Subsections Look?	1
1.2.1 This is a Subsection	1
2 Problemanalyse	3
2.1 Teori: Statistik	3
2.1.1 Population og stikprøve	3
2.1.2 Middelværdi	4
2.1.3 Varians	4
2.1.4 Standardafvigelse	5
2.2 Teori: PRNG	5
2.2.1 LCG	6
3 Chapter 2 name	7
4 Conclusion	8
Bibliography	9
A Appendix A name	10

Todo list

<div></div> Is it possible to add a subsubparagraph?	2
<div></div> I think that a summary of this exciting chapter should be added.	2
<div></div> I think this word is misspelled	7
Figure: We need a figure right here!	7

Preface

Here is the preface. You should put your signatures at the end of the preface.

Aalborg University, March 16, 2023

Author 1

<username1@XX.aau.dk>

Author 2

<username2@XX.aau.dk>

Author 3

<username3@XX.aau.dk>

Chapter 1

Introduction

Here is the introduction. The next chapter is chapter 3.
a new paragraph

1.1 Examples

You can also have examples in your document such as in example 1.1.

Example 1.1 (An Example of an Example)

Here is an example with some math

$$0 = \exp(i\pi) + 1 . \tag{1.1}$$

You can adjust the colour and the line width in the `macros.tex` file.

1.2 How Does Sections, Subsections, and Subsections Look?

Well, like this

1.2.1 This is a Subsection

and this

This is a Subsubsection

and this.

A Paragraph You can also use paragraph titles which look like this.

A Subparagraph Moreover, you can also use subparagraph titles which look like this. They have a small indentation as opposed to the paragraph titles.

I think that a summary of this exciting chapter should be added.

Is it possible to add a subsub-paragraph?

Chapter 2

Problemanalyse

2.1 Teori: Statistik

Statistik er en videnskabelig metode, der systematisk og empirisk beskriver data ved hjælp af forskellige matematiske operationer og metoder. En grundig statistisk analyse kræver indsamling og organisering af data, efterfulgt af analyse ved hjælp af forskellige metoder og statistiske tests for at udvælge information. Typen af data varierer alt efter, hvad der undersøges, og kan derfor omfatte numeriske data såvel som kvalitative data, der skal systematiseres, før de kan analyseres statistisk. I dette kommende afsnit vil der blive introduceret forskellige statistiske begreber, såsom stikprøve og middelværdi, for at danne en overordnet forståelse af teorien. Dette er vigtigt for at kunne udføre diverse statistiske tests senere hen. Derfor vil teorien og metoden blive undersøgt i det kommende afsnit.

2.1.1 Population og stikprøve

I statistisk sammenhæng forstås "population" som en gruppe af enheder eller individer, som skal undersøges og udvindes information fra. Enhver gruppe af ting, der skal undersøges statistisk, kaldes en population. Nogle eksempler på populationer, der kan undersøges, er mænd født efter 1990, Startups i Danmark eller generelt en gruppe af enheder med lignende karakteristika. En "stikprøve" er en tilfældigt udvalgt prøve, der udtages fra en større population. Stikprøven bør have en statistisk signifikant størrelse i forhold til den samlede population, så den kan give et mere eller mindre pålideligt indtryk af hele mængdens egenskaber. Stikprøver bruges ofte, fordi det enten ikke er muligt eller er for krævende at undersøge hele populationen. Et eksempel på en stikprøve er en meningsmåling af befolkningen, hvor der undersøges en mindre del af populationen, og observationerne tilskrives populationen som helhed. Når man estimerer observationer fra en stikprøve til en samlet population, vil der altid være en vis grad af statistisk usikkerhed. Dog kan man

reducere usikkerheden ved at tage højde for faktorer som stikprøvens repræsentativitet for den samlede population, muligt bias i stikprøven, som kan påvirke og gøre resultaterne mindre generaliserbare, og størrelsen på ens stikprøve, hvor større stikprøver er mere repræsentative for den samlede population.

En vigtig del af statistik er at drage konklusioner om parametre i populationen, såsom middelværdi μ og standardafvigelse σ eller tilsvarende for stikprøver: stikprøvemiddelværdi \bar{x} og stikprøvestandardafvigelse s . Dette gøres ved at bruge formler for beregning af middelværdi, varians og standardafvigelse, som er forskellige alt efter, om man har en population eller en stikprøve.

2.1.2 Middelværdi

For populationer beregnes middelværdien ved at tage summen af alle observationerne og dividere det med antallet af observationer i populationen.

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i$$

Hvor n er antallet af observationer og x_i er den i 'te observation i populationen.

For stikprøven beregnes middelværdien på samme måde dog ud fra antallet af observationer i stikprøven:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Hvor n er antallet af observationer og x_i er den i 'te observation i stikprøven.

2.1.3 Varians

Variansen for populationer beregnes ved at tage summen af kvadratet af forskellen mellem hver observation og middelværdien og derefter dividere det med antallet af observationer i populationen:

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

Hvor x_i er den i 'te observation i populationen, n er antallet af observationer i populationen og μ er middelværdien af populationen.

Igen beregnes variansen for stikprøven på samme måde som populationen dog benyttes antallet af observationer i stikprøven:

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Hvor x_i er den i 'te observation i populationen, n er antallet af observationer i stikprøven, \bar{x} er stikprøvens middelværdi og s^2 er stikprøvens varians.

2.1.4 Standardafvigelse

Standardafvigelse er et mål for spredningen af data i en population eller stikprøve. Det er defineret som kvadratroden af variansen. Formlen for standardafvigelse i en population er:

$$\sigma = \sqrt{\sigma^2}$$

Hvor σ er standardafvigelsen og σ^2 er variansen i populationen.

I en stikprøve beregnes standardafvigelsen på samme måde, men variansen er beregnet ud fra stikprøven i stedet for hele populationen. Formlen for standardafvigelse i en stikprøve er:

$$s = \sqrt{s^2}$$

Hvor s er standardafvigelsen og s^2 er variansen i stikprøven.

2.2 Teori: PRNG

Tilfældighed refererer til egenskaberne ved et system eller en proces, der mangler forudsigelighed eller orden, og producerer resultater, der ikke kan forudsiges med sikkerhed. Tilfældighed er afgørende i mange felter, såsom statistik, kryptografi og spil, hvor det bruges til at sikre lighed, sikkerhed og forudsigelighed.

Tilfældige tal er tal, der genereres uden nogen forudsigelig mønster og følger en sandsynlighedsfordeling. De anvendes ofte i computersimuleringer, kryptografi og statistisk analyse. Tilfældige tal kan genereres ved hjælp af forskellige teknikker, såsom fysiske processer som radioaktivt henfald, atmosfærisk støj eller matematiske algoritmer, der simulerer tilfældighed.

Pseudo Random Number Generator (PRNG) er metoder som bliver brugt til at skabe tilfældige tal. En computer vil altid være deterministisk og vil derfor aldrig kunne generere tilfældige tal. Dertil kan der bruges PRNG, hvor der er udarbejdet forskellige metoder, hvorpå man kan genere tal, der så vidt muligt kan give tal, som ligner de er tilfældige. Dette ændrer ikke på at tallene, som algoritmen ender med at producere, ikke er deterministiske, algoritmen får det til at ligne at de er tilfældige, derfor pseudo.

2.2.1 LCG

Der vil i projektet bruges metoden Linear Congruential Generator (LCG), som er en af mange metoder til at skabe pseudo tilfældige tal på.

$$X_{i+1} = (aX_i + c) \mod m$$

Metoden består af en startværdi, X_0 , som overholder $0 \leq X_0 \leq m$. Dernæst bliver a , også kaldet multiplikatoren, gange på X_i ledet, hvorefter der bliver adderet en konstant c . Dette bliver sat modulus m , hvor resultat af det tidligere led, bliver divideret med m , hvilket giver vores tilfældige tal, samt en ny multiplikator. Derudover, skal m overholde $0 \leq m$, a skal $0 \leq a \leq m$ og c overholde $0 \leq c \leq m$

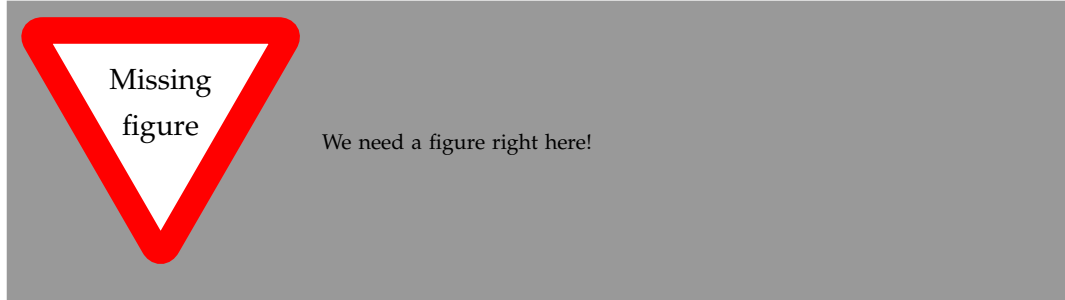
Som der blev skrevet tidligere er det stadig deterministiske tal som den genererer. Da den skaber tal på baggrund af ovenstående formel. Det vil resultere i at tallene på et tidspunkt vil gengive sig selv. Længden fra det første tal bliver gengivet og bliver kaldt perioden, da det er perioden mellem det første originale talt bliver gengivet. Måden hvorpå man kan ændre hvor stor perioden er, gør man ved at ændre på variablerne a , c og m , ved at ændre på dem ændrer man perioden.

Chapter 3

Chapter 2 name

Here is chapter 2. If you want to leearn more about $\text{\LaTeX}2_{\epsilon}$, have a look at [1], [3] and [2].

I think this word is misspelled



Chapter 4

Conclusion

There are probably still some bugs in the theme. If you should find one, then please submit it on <https://github.com/jkjaer/aauLatexTemplates>.

Bibliography

- [1] Lars Madsen. *Introduktion til LaTeX*. <http://www.imf.au.dk/system/latex/bog/>. 2010.
- [2] Frank Mittelbach. *The LATEX companion*. 2. ed. Addison-Wesley, 2005.
- [3] Tobias Oetiker. *The Not So Short A Introduction to LaTeX2e*. <http://tobi.oetiker.ch/lshort/lshort.pdf>. 2010.

Appendix A

Appendix A name

Here is the first appendix