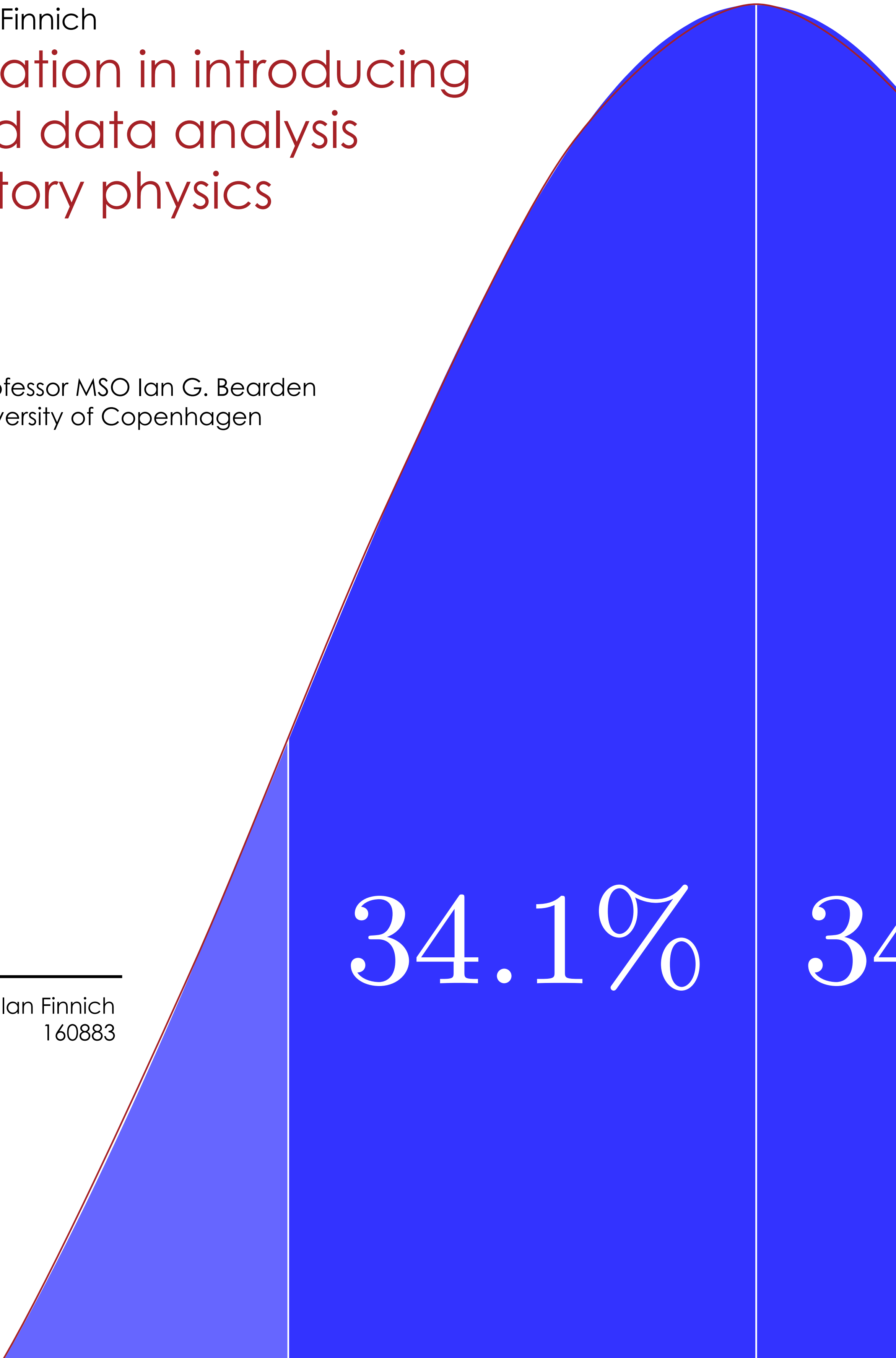
UNIVERSITY OF COPENHAGEN

Master thesis: Allan Finnich

An investigation in introducing Matlab and data analysis in introductory physics

Academic advisor: Professor MSO Ian G. Bearden
Niels Bohr Institute, University of Copenhagen



Allan Finnich
160883

I Abstract/Resumé and acknowledgements

Abstract

In 2009 I began my physics education at the University of Copenhagen. In connection to the laboratory part, of the very first attended course Mek1 on introductory physics, we were to use the CAS-software Matlab and apply error analysis with little instruction herein. Personally, this gave rise to much frustration. This experience inspired the idea to author teachings materials on the basic use of Matlab and data analysis to help future students avoid the frustration I felt. The teaching materials authored were a book titled *"Grundlæggende Matlab og dataanalyse"* (*the book*) and in total 11 screencasts: Both conveyed the same in terms of basic Matlab but the screencasts did not present any statistical background.

The main question we investigated in this thesis was whether students learned more using either the book or the screencasts; but also if preference of a certain media in general was more efficient at schooling purposes relative to others; here we looked at books, screencasts/videos, and notes on the internet in general. In addition we also asked them whether they found either the book, screencasts, or the internet most educational.

To measure learning outcomes three tests were given: A self authored test on basic Matlab and data analysis (Lab Test), the Data Handling Diagnostics test (DHD), and lastly the Force Concept Inventory test (FCI). The DHD was solely on both basic and advanced data analysis and the FCI tested for skills in Newtonian mechanics; mostly misconceptions. Overall we found that the students did quite well on the three tests as $\mu_{\text{Lab Test}} = 14.92 \pm 0.17, \mu_{\text{DHD}} = 10.10 \pm 0.23$, and a gain in the FCI of 3.82 ± 1.01. The gain of the FCI was found to be significantly greater than zero; as such we found statistical proof that the students did learn something about Newtonian mechanics by attending the course Mek1. In terms of statistical concepts we found that future courses of Mek1 should give more attention to SDOM whereas the spread and arithmetic mean warrants no further attention. In addition we found that learning two to different CAS-softwares, here Matlab and Maple, have given rise to confusion for the students.

The first four questions of Lab Test asked the students about their media preference. From these we found that the book was best received as 48.89% of 136 students used it, and of the those 65.15% found it the most educational. The screencasts where not very popular: Here only 28.68% used them and of those 25.64% also found them most educational. We argued that the durations of the screencasts were too long and need trimming and revision to become a viable future resource. Additionally we found evidence by coupling the students' media preference to the other tests that the book along the teachings of Mek1 significantly improved the students' ability to use the error propagation formula and increased their knowledge of precision of a physical quantity.

We on the other hand did not find evidence that the preference of a general media significantly increased the mean test scores of either tests; also that the number of total correct and wrong answers (NTC) of either tests was independent of general media preference. The book and the screencasts was also not found to improve the overall mean test score and they were also independent of the NTC.

Lastly, we investigated what the term "data analysis" entails in the Danish secondary: Using the written exams problems from STX and HTX we found that transformation, plotting and fitting of data were common in both educations. In addition we gave a rough sketch as a suggested solution to introduce error analysis in STX and HTX which we dubbed “Error analysis bootcamp”.

Resumé

I 2009 påbegyndte jeg min uddannelse i fysik ved Københavns Universitet. I forbindelse med det første fulgte kursus Mek1, hvilket omhandlede indledende klassisk mekanik, skulle vi benytte CAS-værktøjet Matlab og anvende usikkerhedsberegning uden den store instruktion heri. Dette gav, for mig, ophav til megen frustration, hvilket inspirerede idéen at forfatte undervisningsmaterialer omhandlende grundlæggende anvendelse af Matlab og data analyse. Disse havde til formål at afhjælpe fremtidige studerendes eventuelle frustrationer. Undervisningsmaterialerne omfatter en bog med titlen “Grundlæggende Matlab og dataanalyse” (bogen) og i alt 11 screencasts.

Hovedspørgsmålet undersøgt i denne afhandling er om studerende lærte mere ved brug af enten bogen eller screencastsne. Men også om et bestemt generelt medie bedre formidlede fagligt stof relative the andre. Af generelle medier kiggede vi på bøger, screencasts/videoer og noter på internettet. Endvidere blev de studerende også spurgt om fandt bogen, screencastsne eller internettet mest lærerig.

For at måle udbyttet af undervisningen blev tre tests givet: En forfattet test omhandlende basal anvendelse af Matlab og data analyse (Lab Test), Data Handling Diagnostics test (DHD) og Force Concept Inventory test (FCI). DHD’en omhandlede kun basal og avancerede data analyse og FCI’en testede for evener i klassisk mekanik. Generelt set fandt vi, at de studerende klarede ganske udemærket i og med $\mu_{\text{Lab Test}} = 14.92 \pm 0.17, \mu_{\text{DHD}} = 10.10 \pm 0.23$ og et gain i FCI’en på 3.82 ± 1.01. Gain’et i FCI’en fandt vi var signifikant større end nul - altså har vi statistisk evidence at de studerende har lært noget omkring klassisk mekanik ved at fulge kurset Mek1. Af statistiske begreber fandt vi, at fremtidige Mek1 kurser skal fokusere mere på SDOM i modsætning til spredning og middelværdi. Endvidere fandt vi, at tilegnelsen af færdigheder i to CAS-værktøjer, her Matlab og Maple, samtidig har medført forvirring af de studerende.

De første fire spørgsmål i Lab Test’en adspurgte de studerende omkring deres medie præference. Via disse fandt vi, at bogen blev modtaget mest positivt da 48.89% af 135 studerende benyttede den, og af disse 65.15% fandt den mest lærerig. Derimod var screencasts’ne ikke særlig populære da kun 28.68% af 136 studerende benyttede dem, og af disse 25.64% fandt dem mest lærerige. Vi argumenterede, at deres længder var for store og skal skæres ned hvis de i fremtiden skal kunne gå for at være en virkbar ressource. Ved at sammenkoble de studerendes media præference med de øvrige tests fandt vi også, at bogen sammen med undervisningen af Mek1 øgede deres evne til at benytte ophobningsloven og viden omkring præcision af en fysisk størrelse signifikant.

Vi fandt dog ikke evidens at præferencen af et af de førnævnte generelle medier øgede den gennemsnitlige test score signifikant. Endvidere også at antallet af total korrekte og forkerte svar (NTC) i alle tests er uafhængig af generel medie præference. Bogen og screencastsne fandt vi heller ikke øgede den gennemsnitlige test score ej heller, at deres brug afhæng af testsnes NTC.

Til sidst undersøgte vi hvad der ligger bag ordet “dataanalyse” i de danske gymnasier. Ved at benytte de skriftlige eksamner fra STX og HTX fandt vi, at begge uddannelser havde transformation, plotning, og fitning af data til fælles. Endvidere gav vi en grov skitse som en foreslået måde til at introducere usikkerhedsberegninger på STX og HTX; denne navngav vi “Error analysis bootcamp”.

Acknowledgements

I would like express my deepest appreciation to my academical advisor professor MSO Ian G. Bearden for useful ongoing sparring, feedback, and support. In addition I wish to thank postdoc Kristjan H. Gulbrandsen, lector Troels C. Petersen, and postdoc Christian H. Christensen; also my fiancé Ann Katrine Olsen and father Mogens Finnich for love and support all the way through my endeavours.

Table of contents

Page

Chapter 1 Introduction

In 1642 a brilliant young French mathematician, Blaise Pascal, is credited with starting a new era; the era of mechanical computation. Étienne Pascal, Pascal's father, worked as a tax official in Rouen, and as such his profession required a lot of arithmetic with large numbers, which Pascal himself learned as he was assisting him. In order to reduce some of his father's workload, Pascal began constructing the worlds first mechanical calculator in 1642.

Ten years and fifty prototypes later, his calculator was completed which was dubbed the Pascaline [Burton, 1997, page 449-450]. Pascal's great feat of engineering is shown figure 1.1. The only arithmetic operation the Pascaline could perform was addition of positive numbers. However, by using the mathematical principle of *9's complement* subtraction of numbers could also be achieved. One problem that Pascal never did solve was the unwanted carrying for which he received a lot of criticism. The unwanted carrying mostly occurred when the machine was operated on a slanted surface or when accidentally disturbed.

Later in 1694 - 52 years after Blaise Pascal's death - the German mathematician Gottfried Wilhem von Leibniz improved the Pascaline by building a machine that also could multiply numbers. His invention, however, was partly based on Pascal's originally notes and drawings of the Pascaline.

The next milestone was in the beginning of the 19^{th} century, when the English mathematician Charles Babbage sat the course for the evolution which has led to the computer as we know them today. To prevent errors in various numerical tables written by hand he came up with the idea of a machine that could perform the calculations running on steam. This machine would be programmable by the user. The last iteration of such a machine was in 1822, which he dubbed the *Analytical Engine*, and is considered the first general-purpose computer. Work on the Analytical Engine commenced in the same year but it was never completed, as the British government over time lost interest in the project [Burton, 1997, page 626-27].

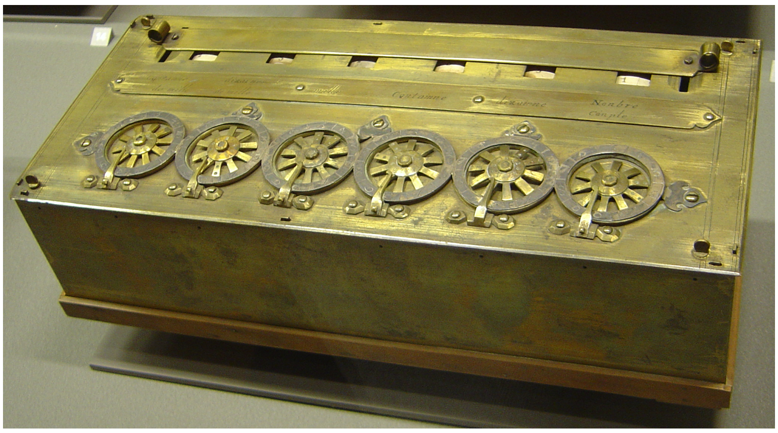


Figure 1.1 • The figure shows the Pascaline. It is the first mechanical calculator built and invented by the French mathematician Blaise Pascal, as an aid to his father's duties as a tax official in Rouen. Picture from [Wikipedia, 2013].

Today most do not question the importance and wide spread use of computers, especially in the natural sciences. When I began studying physics in 2009 the first course attended was Mek1 which was on introductory physics. In connection to the laboratory part of Mek1 the recommended CAS-software was Matlab. It was not mandatory that Matlab was to be used but it was strongly suggested. However, for many this program was uncharted territory and there was not planned any specific training as to its use. In addition, we were also supposed to apply error analysis on the basis of little instruction. Personally, this was a daunting task and it gave rise to a lot of frustration. It is this frustration which inspired the idea for this thesis: I wished to create teachings materials in order to help future students avoid the frustration I felt as a beginning student.

Every student, however, absorbs teaching differently: Therefore I decided to not only restrict the media of the teachings materials to just one, but two: a book and screencasts. On this basis we wish to investigate if either media is better relative to the other. In addition we also wish to clarify if the same is true in general. The three media investigated are books, screencasts/videos, and the internet. As a measure of what the students have learned we use their resulting test scores of three different tests the students are given during Mek1: A self authored test on basic Matlab and data analysis; the Data Handling Diagnostics test; and, lastly the Force Concept Inventory test. The students which will be part of this investigation are those attending Mek1 in 2013.

We also investigate the possibility of introducing the field of error analysis in the secondary schools based on the findings of the given tests.

Chapter 2 The teaching materials

The first step was to create the teaching materials, for the freshman physics students of 2013. In total the teaching materials comprise of a book titled *"Grundlæggende Matlab og dataanalyse"* (translated: "Basic Matlab and data analysis"), which throughout this thesis will be referred to as simply *the book*, and eleven screencasts. As the title of the book suggests, it treats the basic use of Matlab and data analysis in the context of physics experiments. The curriculum of both the book and screencasts are the same in terms of the basic use of Matlab. However, they deviate in terms of the statistical background of data analysis as the aim of the screencasts is not to provide lectures on theory, but rather to be how-to-guides on specific topics.

In this chapter, the layout and the curriculum of the book will be presented and discussed. Also, a short presentation of the technical elements and discrepancies between the book and the screencasts will be discussed.

2.1 The book

This section describes the book *"Grundlæggende Matlab og dataanalyse"*, written for the freshman physics students in 2013, which is on the basic use of Matlab and how to apply basic data analysis. When translated the title of the book is *Basic Matlab and data analysis.*

First, we discuss the layout and the reasons the layout is an important factor, and then we present and discuss its curriculum. The book can be found in appendix A.

2.1.1 The layout

In the realm of marketing and advertising, the acronym **AIDA** is often encountered; it is a list of common events that a consumer most likely goes through in order to purchase a product; it is short for **A**ttention, **I**nterest, **D**esire, and **A**ction. The AIDA-model says first to get the consumers *attention*; next get them interested in the product by arguing the benefits it gives rise to; thereafter get them to *desire* the product, i.e convince them that they "must own it", and finally leading to the *action* of buying the product.

When writing a book a pleasing layout is very important as it is what the potential reader's eyes first meet; a good layout should capture the potential reader's attention interest before he has read a single line. An unclear layout can, in worst case hinder communication rather than promote it. The worst case scenario would be that the reader simply does not want to invest more time in book, and, as a consequence, finds another source of information. As an author you want to *sell* your book as best as possible in order to attract the potential readers' attention, and a good and clear layout helps in doing exactly that. Therefore we will now present and discuss the layout of the cover and within the book.

The cover

At the top, the education institutions logo (University of Copenhagen (KU)) is placed in the left side. See figure 2.1. According to the KU Design Guide, certain specifications regarding the logo have to be met [KU, 2013a]. The specifications are here all met. In the right-hand corner, the name of the author is typeset in the same font as name of the educational institution.

To visually convey that the book is about data analysis, a histogram, containing a simulated normal distributed quantity, and a gaussian fit (blue line), is shown. The histogram is filled with a pleasing bordeaux type of red, which we henceforth will refer to as simply bordeaux. The bordeaux is the same as used in the headings for the chapters, sections and subsections of this thesis.

In the background a dotted grid is shown to further enhance the fact that the contents of the book are of an academic nature. The histogram with fit and grid are generated in Matlab along with the simulated numbers. To further emphasize that the book is also about Matlab, Mathworks' logo is placed to the right of the book title.

The initial line of thought was that the book cover consisted of a collage of some Matlab code and mathematical expressions of relevant statistical distributions. This idea, however, was dropped as it would perhaps discourage the students from ever reading the book. However, such a book cover may be intriguing, but if it ultimately discourages them it would result in no desire to read it, and thereby no action, as it may give the impression, that the complexity of the contents is high. Since both Matlab and data analysis are most likely completely new to the students a less discouraging design was chosen, which still visually conveyed the book's contents in a nice and professional manner.

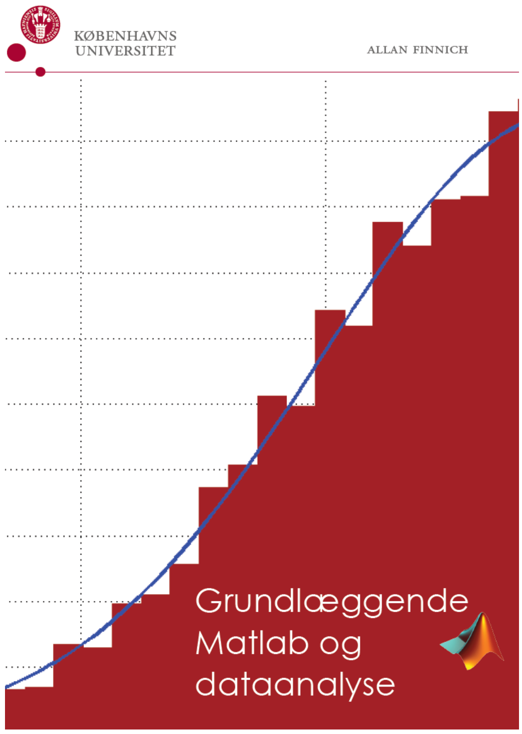


FIGURE 2.1 • THE COVER OF THE AUTHORED BOOK ABOUT BASIC MATLAB AND DATA ANALYSIS.

The structure of the layout

On every chapter page a picture, relevant to the contents of that chapter, is placed in the top right corner. To illustrate an example hereof the chapter page of chapter 3 of the book is shown on the left hand side in figure 2.2. Using bordeaux as the color, a big filled rectangular box is placed vertically midway and left justified. In this box the chapter number and title are written in white. Lastly, a written overview of the chapter's contents is placed in the bottom right corner of the page. This layout of the chapter page is the first version, and has a pleasing, clear and professional look. However, it has been considered to place a short text describing what the chapter figure shows. This was, however, not implemented as it might result in the chapter page seeming disorganized, and thereby generating unwanted noise for the reader.

The section heading consists of a similar, but smaller, bordeaux filled box in which the section number is also written in white. To the right of the box the section title is written in black. The subsection heading consists of only its title and is written in bordeaux; however, it is typeset in a smaller font size. An illustration of the section and subsection headings is found on the right hand side of figure 2.2.

The font used in both the book cover, chapter page, section heading, and the subsection heading is the same. Thus, a direct coherence between these four is the font and bordeaux. The font is called *century gothic*.

To draw special attention the main results these are summarized in a titled box, as shown in the right hand side of figure 2.2. The title of the box is written in white, while the background is a bordeaux filled ellipse. The same color is used as the frame color of the box. To make the box further stand out, its background color is yellow. This color was chosen as it provides a good contrast to the bordeux.

Lastly, when Matlab code is presented, it is done so on a gray background with a light green left frame. To make the code more organized line numbers are provided on the left side. The first version of the code blocks had alternating background color between gray and white thereby creating a zebra effect. However, this was found not desirable as it resulted in difficulty reading the contents of the code block.

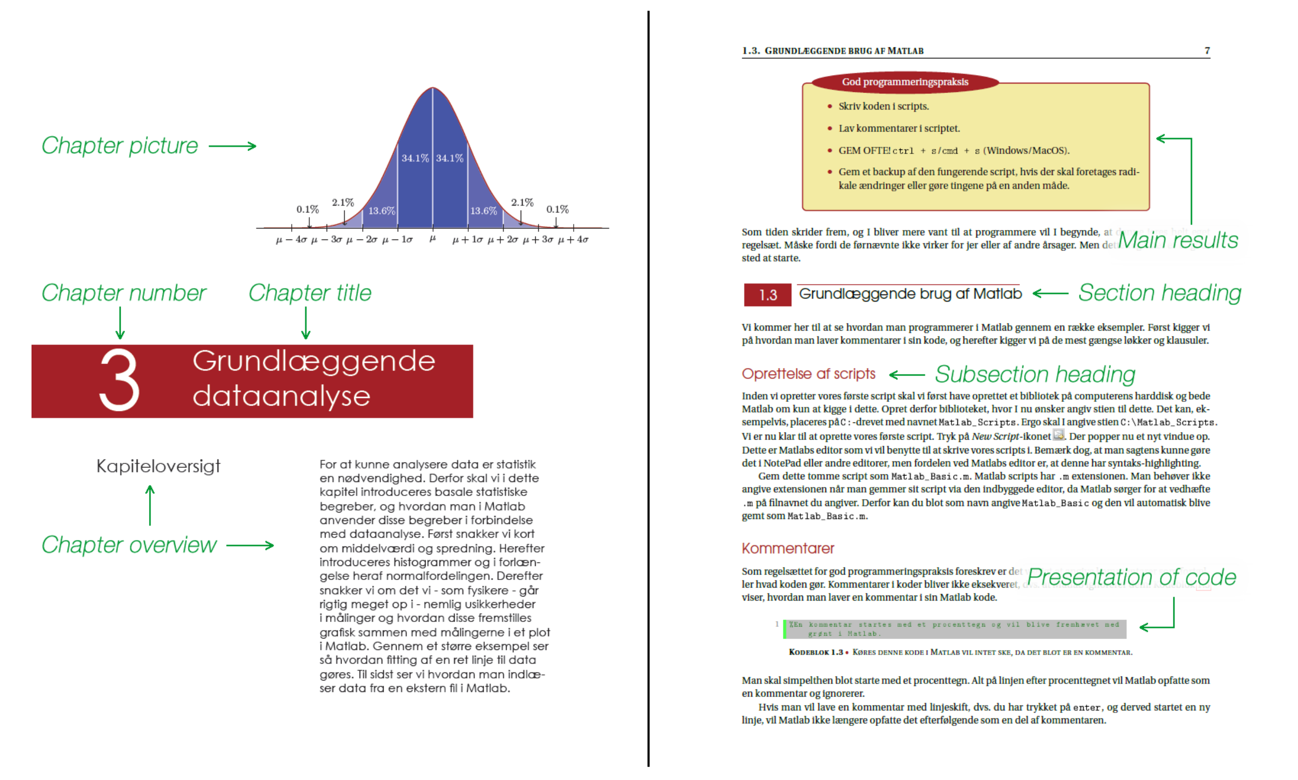


FIGURE 2.2 • THIS FIGURE ILLUSTRATES THE OVERALL LAYOUT WITHIN THE BOOK. TO THE LEFT THE LAYOUT OF THE CHAPTER PAGES IS SHOWN: CHAPTER 3 IS SHOWN AS AN EXAMPLE. TO THE RIGHT THE LAYOUT OF SECTION AND SUBSECTION HEADINGS ARE SHOWN, AND ALSO HOW MAIN RESULTS AND BLOCKS OF CODE ARE PRESENTED. NOTICE THE COLOR CONNECTION BETWEEN THE BOOK COVER (SEE FIGURE 2.1), CHAPTER PAGE, SECTION, AND SUBSECTION HEADINGS IN FORM OF THE BORDEAUX TYPE OF RED AND THE FONT TYPE. THE FONT TYPE IS CALLED CENTURY GOTHIC.

2.1.2 The curriculum

The curriculum for the term basic Matlab has been determined by an education committee at KU; however, its implementation is up the person in charge of the laboratory part of the course *Mek1*. In this case this person also happens to be the academic advisor of this thesis, professor MSo Ian Bearden. By his statement the minimum curriculum is (1) handling of one dimensional arrays; (2) plotting arrays; (3) scatter plots; (4) histograms; and (5) loops [Bearden, 2013b].

We note that the term “data analysis” as used in the course descriptions is somewhat vague. In fact, the description of Mek1 does not mention data analysis explicitly, but does mention that after the course the students are supposed to be able to *"[...] describe experimental investigations of simple mechanics phenomenon [...]"* and *"[...] use computer software to [...] report scientific results [...]"* [KU, 2014a]. These are the only references which together could be interpreted as data analysis, but there is no further specification; on the other hand "data analysis" can be found in the description of Mek2 but with no explicit detail of what is meant [KU, 2014b]. The decision was taken, therefore, to construct a minimal curriculum based on the experiments the students have been expected to conduct in previous years. To satisfactory understand and report the results of these experiments, they need at least to be familiar with (1) mean and spread; (2) histograms; and (3) error propergation formula [Bearden, 2013b].

Based on this minimal curriculum for both basic Matlab and data analysis we now present how these have been conveyed in the book starting with the foreword.

Foreword

On page one is the foreword, which describes the motivation for writing the book: The motivation was to minimize the struggle students have had with learning Matlab and basic data analysis from a personal perspective.

To further argue that the book is of interest to the readers/students it is emphasized that the applicability of the book is not restricted only to the course Mek1, wherein it is introduced; it is also very applicable in the course Mek2 as Matlab and data analysis here is also used. This is explicitly stated in the course description for Mek2:

> *"[...] Ved laboratoriearbejdet opnås forståelse for fysiske eksperimenter, måleteknikker, databehandling og statistik. Der udvikles udvidede kompetencer i anvendelse af MatLab. [...][KU, 2014b]".*

The applicability also extends to the course on introductory quantum mechanics: A mandatory part of this course is that the students are to compute numerical solutions of the one dimensional time independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}+V(x)\psi(x) \tag{2.1}$$

for different types of potentials $V(x)$ [KU, 2013b]. It is, however, not explicitly stated in the course description that the numerical solutions need to be done in Matlab. Professor Anders Søndberg Sørensen, who is responsible for the course, does confirm that Matlab is the preferred choice of program. The reason hereto is that the students cannot expect to be aided by the instructors if another software is used. Therefore, according to professor Anders Søndberg Sørensen, almost all the students chooses Matlab; no statistics exist which can support his claim [Sørensen, 2013] but it is, however, well supported anecdotally.

In addition it is in the foreword made clear that simply reading the book is not enough to lead to mastery of the subject matter; the students must actively try to use Matlab and analyze and reflect upon experimental results in order to truly master the knowledge and skills they entail. Thus, the book should be considered more a resource to fall back on rather than an exhaustive text. Thereafter a short overview of the contents of the chapters and appendices is given.

The foreword concludes with a review of relevant courses I have taken during my own physics education at KU and a good advice. This is done if the students have the ambition of becoming even better equipped experimental physicists. The courses mentioned are *EF*, which is a course on experimental physics and *Applied Statistics: From Data till Results*, which is on more advanced uses of statistics also outside the field of physics.

Given my own time as a Danish secondary school student, and that spent as a teacher at one, I have hypothesized that Microsoft's Word or, the almost identical, OpenOffice Writer are the main programs used to type in assignments and reports. When attending KU one realizes that the standard here is quite different. Therefore a good advice is given: it is informed that when studying physics at KU Microsoft's Word is passé; the standard is now LaTeX ; as such it is implied that it would be in their best interest to get acquainted with LaTeX at an early stage. If that seems to daunting a task a course in LaTeX is available; however, if they feel up to the task of learning LaTeX through self study a reference to Tobi Oetiker's book *The Not So Short Introduction to LaTeX* is given.

Chapter 1 - Basic Matlab

First step is to get the students acquainted with Matlab's interface as it is the first they will see; most likely also for the first time. To make this first encounter seem less daunting the function of the windows *Current folder*, *Command window*, *Workspace*, and *Command history* are explained.

As the target group is freshman physics students a large fraction of them will most likely not have an abundance of programming experience. To provide them a good foundation on which they can start programming a set of guidelines are presented; in total four guidelines are presented. They are as follows;

1. write comments in the code explaining what it does,
2. write the code in scripts,
3. save often, and
4. save different versions of the scripts.

No theories or statements from enlightened persons are quoted for these guidelines; they are simply based on my *"programming commandments"* to which I adhere. They are the product of personal experience and are applicable regardless of the programming language. The arguments for the guidelines given in the book are as follows: by commenting the code it saves time if one has to reacquaint himself with what the code does after some time away; writing all code in scripts is simply a necessity in order to maintain the overall view of the project and to preserve time; saving often also saves time in case of an unexpected system error or restart as the amount of work lost is hereby kept to a minimum; lastly, by saving different versions of a script one always have a working script making the frequency, for situations where one ends with a non-working script due to, say, experimental changes or an attempt to optimize the code, small. How to create scripts and write comments in them are of course presented.

The handling of one dimensional arrays is a part of the required curriculum. The definition of *handling* has been decided to entail how arrays are defined, destroyed, and manipulated. To accomplish this best the students first need to learn about variables. This is presented by showing (1) what a variable is; (2) how to define and destroy them; and (3) how they can be used in computations. This is of course presented prior to any presentation of arrays.

In connection to the manipulation of one dimensional arrays it is very relevant that the students are made aware of how Matlab uses them in computations other than addition and subtraction, i.e show the math behind element-by-element operations (EBE). To show the math behind EBE only arrays of dimension 5 are used in order not to use complicated mathematical symbols confusing the students in the process.

However, when conducting physics experiments arrays of such a low dimension are rarely the case; so to handle situations where an operation, not necessarily standard in Matlab, has to be performed on each element of arrays the use of for-loops is also needed to be explained; their importance is emphasized as they are refered to as an aspect which *"[...] almost always have to be used in connection to data analysis [...]"* [Finnich, 2013]. As some may be more experienced, and would therefore perhaps like to be further challenged, it is also demonstrated how to create simple and advanced functions which can serve as an alternative solution to the before mentioned situation. For completeness the handling of if-clauses is also presented but may be considered not as relevant as for-loops.

Chapter 2 - Grafisk fremstilling af data

The second chapter of the book is mainly on how to plot data in Matlab. Translated the chapter title is *Graphical representation of data*. This chapter covers at minimum how to plot arrays, i.e plot data which have been put into arrays, and the concept of scatter plots.

To introduce plotting in Matlab they first need to be presented to Matlab's command *plot*; it is the basic command for plotting two dimensional coordinate systems in Matlab. However, as visually demonstrated,

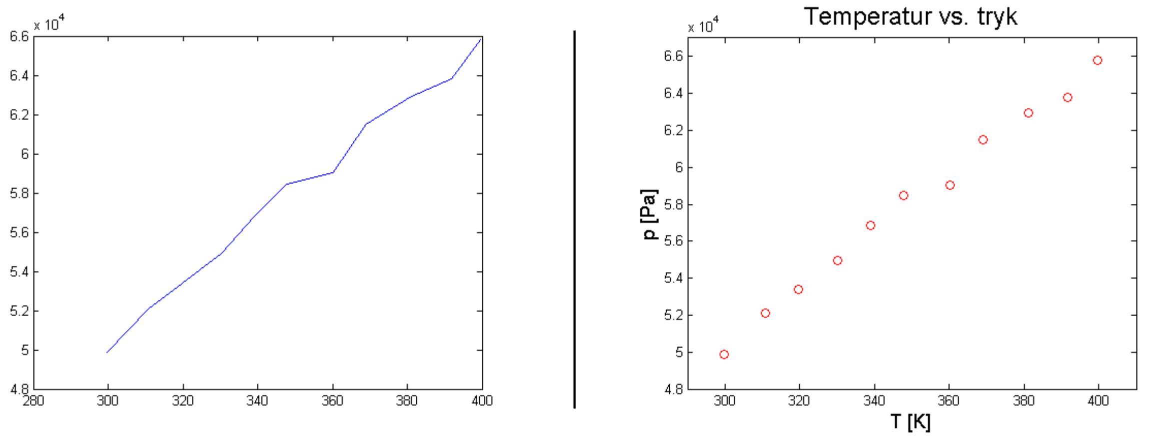


FIGURE 2.3 • TO EMPHAZISE THE INADEQUACY OF THE STANDARD STYLING OPTIONS OF PLOTS IN MATLAB A PLOT WITHOUT STYLING IS SHOWN (LEFT), AND LATER ONE WITH ONLY A LITTLE STYLING (RIGHT). THE POINT IS THAT A LITTLE STYLING MAKES A BIG DIFFERENCE.

the resulting basic plot that Matlab outputs without any form of styling is simply not satisfactory: (1) the data is not represented as points, (2) no axis labels, and (3) no title of the plot. See left side in figure 2.3. These three shortcomings are mentioned as the absolute minimum requirements for a plot. In connection hereof some basic options which styles the resulting plot are explained and examples of their use are given. To convince the students, that even though styling of plots may be boring and seem as a nuisance, they are invited to compare the unstyled plot (left side in figure 2.3) with the styled plot (right side in figure 2.3) meeting the absolute minimum requirements. In doing so the thought is that they by themselves realize that a little amount of styling makes a whole lot of difference.

Plots shown in reports or articles, etc., usually contains representations of more than one set of data in order to either compare different data sets or simply to save space. For this exact reason they are instructed in how this is done; in connection hereof legends in plots are of course introduced and treated. However, in almost all experiments more than one plot are generated when running a script; therefore they are also shown how to plot different data sets in separate plots.

In the start of chapter 2 the students are presented to a fictitious experiment:

> *"A relatively air tight container has a volume* $V = 0.5\text{m}^3$ *and contains atmospheric air. Connected to the container are an electronic thermometer and pressure gauge. When the container is heated the values of the temperature and pressure are recorded.[Finnich, 2013]".*

The resulting data from this experiment are, of course, simulated; the equation used is

$$p(T) = \frac{nR}{V} \cdot T + Q_1 = \frac{10\,\text{mol} \cdot 8.31\,\frac{\text{Pa}\cdot\text{m}^3}{\text{mol}\cdot\text{K}}}{0.5\,\text{m}^3} \cdot T + Q_2 \cdot 1000, \tag{2.2}$$

where T ranges from 300 K to 400 K in steps of 10 K and Q_2 is normally distributed real numbers with mean 0 and spread 1. The result is simply points from a straight line with gaussian noise. This fictitious experiment is used during the course of chapter 2, and is used to (1) "have" data in order to present how to generate plots; (2) enhance the explanation of what is meant by data sets being correlated or uncorrelated. A plot very similar to the one shown to the right in figure 2.3 is referred to as a scatter plot where there is a clear dependence between data sets (pressure depends on temperature). To illustrate the converse case, i.e no dependence, two sets of 100 randomly generated integers are plotted against each other.

Chapter 3 - Grundlæggende dataanalyse

To form the absolute basic statistical foundation for data analysis the definitions of mean and spread are presented, which are denoted by $\overline{x}$ and σ_x, respectively. The students are most likely familiar with the formula to compute the mean of some numbers; however, the spread is probably unfamiliar to them; as such the interpretation of this concept is a measure of is given: It is the average deviance from the mean.

To further clarify why the mean and spread are important the theoretical construct of true mean μ_s and spread σ_s are introduced. These two constructs are explained as being the mean and spread of a data set in which we recorded an infinite amount of data points; however, as pointed out, this is not possible. Therefore when recording N data points in experiments we regard the resulting data set as a sample of the infinitely large data set, and from the sample we want to give our best estimate of μ_s and σ_s which exactly are $\overline{x}$ and σ_x, respectively. Of course the students are shown how to compute the values $\overline{x}$ and σ_x in Matlab.

The only statistical distribution treated is the gaussian as it is the only relevant distribution to mention at a basic level. However, it is noted that many other statistical distributions exist but in the context of the subjects treated in the book, the normal distribution is the only one relevant. Interested readers are invited to refer to both [Taylor, 1997] and [Barlow, 1999] as they treat some of the other mentioned distributions.

The most logical way to introduce the gaussian, and in the process couple the concepts of mean and

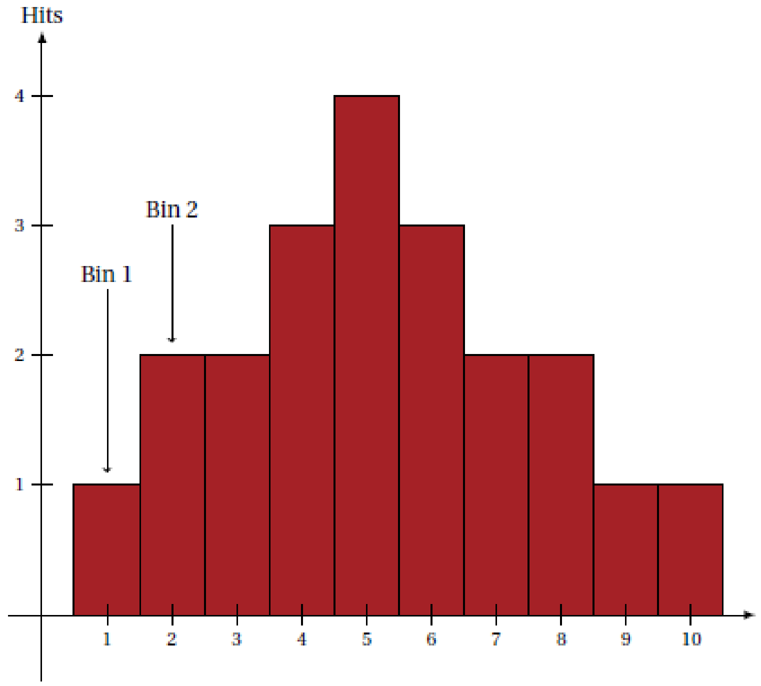


FIGURE 2.4 • TO GIVE THE STUDENTS AN EXAMPLE OF HOW A HISTOGRAM LOOKS LIKE AND ALSO EXPLAIN WHAT IS MEANT BY BINS AND ITS HITS. THIS HISTOGRAM HAS BEEN GENERATED USING SOME SEQUENCE OF 21 NATURAL NUMBERS.

spread, is via first introducing the histogram, as this is a visual statistical tool which investigates how a measured quantity is distributed. For the students to get a basic feel for histograms a step-by-step instruction on their construction is given instead of only show how to construct them in Matlab. The data used in this instruction is a sequence of 21 natural numbers; it also introduces the concepts of a histogram's bin and its hits; the same is conveyed visually. See figure 2.4.

In the presentation of the gaussian its pdf is presented both mathematically and visually. The mathematical part of the presentation is simply the formula

$$f(x;\sigma,\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-(x-\mu)^2/(2\sigma^2)}, \quad x \in \mathbb{R}, \tag{2.3}$$

and the visual is shown to the left in figure 2.5: It shows the form of the pdf and that it assumes its maximum at $x = \mu$. To ensure that the role of σ in the gaussian pdf is conveyed a plot of it is shown (see righthand side in figure 2.5); it contains five gaussians all with same μ but increasing σ in steps of 1 from 1 to 5. From this the role of σ is deduced; for increasing values of σ the curve becomes more broad, i.e greater spread.

To couple the gaussian to the experiments, the students will conduct during both Mek1 and Mek2, it is explained that the data will most likely be normally distributed implying that $\sigma \approx \sigma_x$ and $\mu \approx \overline{x}$. How the gaussian is relevant in the context of histograms is conveyed by showing the reader how to fit a gaussian to the histogram constructed earlier using the sequence of 21 whole numbers. The thought is that the reader can see a direct link between the form of both the histogram and the gaussian. On a technical note the function in Matlab which does this is *histfit*, but it has one big omission: It does not return the values for the

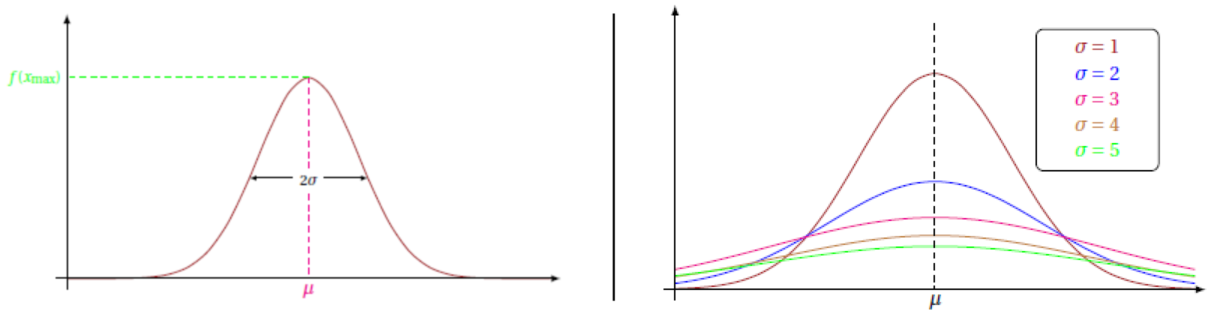


FIGURE 2.5 • THIS FIGURE SHOWS HOW THE READER IS VISUALLY SHOWN THAT THE NORMAL DISTRIBUTION IS BOTH CENTERED AND SYMMETRICAL ABOUT $x = \mu$ (LEFT) AND HOW CURVE CHANGES WHEN VARYING σ (RIGHT).

fitted mean and spread; as such it shown how to access these values via the command *fitdist* which *histfit* uses.

Uncertainties, and how to correctly report them, are also big part of data analysis. These topics are therefore to be included. This is best accomplished by setting fourth a gedanken experiment which form the basis:

> *"Imagine that all N first year physics students have made a measurement x_i of an objects length, for example the teacher's table in the auditorium, using the same experimental instrument. The individual measurements are unknown to the rest of the students except, of course, their own.[Finnich, 2013]".*

This gives rise to explain the students what an independent data set is, and also how the theoretical constructs of the true mean and spread, and sample mean and spread are used in a specific setting, i.e $\overline{x}$ and σ are the best estimates for μ_s and σ_s, respectively. By reminding them that the spread is a measure for the average deviation from the mean then must be a measure for uncertainty. In the same breath the common

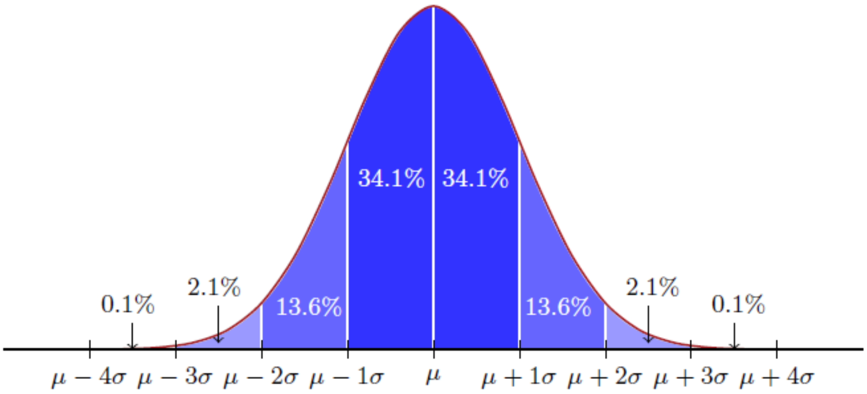


FIGURE 2.6 • THIS FIGURE IS SHOWN TO VISUALLY DEMONSTRATE THAT WHEN A QUANTITY IS NORMALLY DISTRIBUTED THEN APPROXIMATELY 68% OF THE MEASUREMENTS WILL LIE IN THE INTERVAL $\overline{x} \pm \delta x$.

notation for uncertainties is shown; also how an experimental result is reported correctly which is very relevant when writing reports during courses.

In order provide a coupling between uncertainties and the gaussian it is pointed out that, when some quantity is normally distributed, approximately 68% of the measurements will lie in the interval $\overline{x} \pm \delta x$; this is also shown visually (see figure 2.6). Why it is
68% is not demonstrated by direct calculation as it would give to confusion rather than clarity.

To connect the theoretical construct μ_s with uncertainty, hereby making it slightly more concrete, the introduction to the standard deviation of the mean (SDOM) is crucial as this tells that μ_s is approximately 68% likely to lie in the interval $\overline{x} \pm \delta\overline{x}$. Furthermore, SDOM gives rise to clarify that the uncertainty of $\overline{x}$ directly depends on the number of measurements. This is shown by calculation as it is considered to be simple.

When reporting experimental results the subjects precision and significant digits are very relevant and is therefore also included. The best way to introduce these was decided to be by examples: Three specific experimental results. During the analysis of these it is explained which is correctly and wrongly reported.

One of the essentials of this chapter is to show how errors propagate as the students are surely to apply it to their analysis of data from physics experiments at some point. To introduce this no theoretical derivations are shown as this would be a too daunting task for the reader to decode any meaning from. Therefore the use of the formula for error propagation, and when it is applicable is treated. Its use is shown through a numerical problem which revolves around a simple electrical circuit.

Technically they are also shown how to represent data with uncertainties in Matlab, i.e expanding the elements from chatper 2 to include errorbar plots. For the interested, and perhaps more skilled students, the procedure for fitting a straight to a data set is also treated. Note, however, that it is not just Matlab's command *fit* which is used but the formula for the slope, intersection, and their uncertainties are presented and put to use using Matlab. Again the theoretical foundation of these formulas is not treated as it is most likely too complex. Their use is therefore only demonstrated. The data used are those presented in chapter 2 from the fictitious experiment (see page 8) with a closed container containing atmospheric air; this gives a sense of connectivity between chapter 2 and 3.

It should here be noted that that the decision to only introduce the gaussian restricts presenting errors in the bins, i.e errors in connection to counting statistics. In hindsight this could also have been introduced, and automatically later give rise to apply it to the field of physics concerning radioactivity; hereby students who would like a further challenge, or are simply just curious, would have the opportunity. In short: Presenting and treating the poissonian also would serve the same purpose as the introduction to if-clauses in chapter

1 of the book.

Appendix A - Matematik appendiks

The target group of the book is, as before mentioned, freshman physics students; as such the only mathematical training they have is that from a secondary school education. In the departmental orders for these educations makes no mention as to the mathematical symbol for summation and partial differentiation being part of the mandatory curriculum. Although the course MatIntro treats these concepts but we cannot recklessly assume that they are presented there in a timely fashion in connection to Mek1. Therefore these concepts are treated in a mathematical appendix in the book. Both concepts are not treated extensively theoretical but the main point is their use and meaning.

2.2 The screencasts

The second media of the authored teaching material is the screencasts. They have been recorded with Adobe's software Adobe Captivate. Since Absalon has a restriction on the amount of space a course uses the format, in which they have been exported, is swf, i.e flash video.

In total there are eleven videos and they represent the chapters of the book in the following way: Screencasts 1 through 6 represents chapter 1 of the book; 7 and 8 represents chapter 2; and the remaining represents chapter 3. When uploaded to Absalon they have been given the following translated titles

- Screencast 1 - Matlab's interface
- Screencast 2 - Scripts and Matlab editor
- Screencast 3 - Variables and comments
- Screencast 4 - Arrays and built-in mathematical functions
- Screencast 5 - The colon-operator and for-loops
- Screencast 6 - Good programming style
- Screencast 7 - Plots and style of plots
- Screencast 8 - Plotting more data sets in one plot and legends
- Screencast 9 - Calculation of mean and spread of a data set
- Screencast 10 - Histograms and fitting of a normal distribution
- Screencast 11 - Plot of a data set with its uncertainties

They can be reviewed on the enclosed disc or at `http://www.alfin.dk/Matlab/`.

The screencasts convey roughly the same content as the book. However, they are not supposed to be lectures which is emphasized in screencast 9; this implies that some of the theory from the book are not presented: For instance the correct way to report an experimental result including significant digits and precision are not presented; also it implies that none of the calculations given in chapter 3 are not treated. Screencast 6 sadly turned out to be a lecture as this only conveys the four guidelines of good programming style; as such there is not shown any programming in Matlab but they are of great importance and are therefore chosen to be part of the screencasts. Of course the guidelines are upheld in every video, i.e in all screencasts the code is written in scripts and comments are provided where relevant.

In regards to the order of the subject presented in the screen casts the guidelines for good programming style are presented after for-loops instead of before Matlab's interface. The reason for this is that it was deemed more logical as at this point the creation of a script and writing comments have been introduced. Furthermore, the guidelines have already been put into action in the three preceding screen casts. It was the thinking that the students would notice that writing comments and using scripts are both powerful tools to keep the overall view.

2.3 Feedback

In order to collect feedback from the students, in regards to their opinion of both the book and screencasts, it was published on Absalon that I sought ten students to participate in a focus group. The first planned focus group was in the week between courses (week 46 2013) in which there was no mandatory teaching; as such participation would not have disrupted any teaching, however, no one signed up even though refreshments were offered. The reason for no willing participants is not known but we believe that (1) the students felt they required some time off after the exams in MatIntro and Mek1 or (2) they had no interest in the focus group. A second attempt was planned in the following week (week 47) which was the first teaching week of Mek2. It was announced on Absalon in week 46; again there were no takers. After the second attempt this endeavour was dropped.

It should be noted that the focus group was not the only opportunity to provide feedback; as the teaching materials was uploaded to Absalon a description was also published. Herein all the students who had constructive comments, with emphasis on constructive, to either the book or screencasts were invited to send them to me. Therefore they have had ample opportunity to provide feedback, however, they simply chose not to. The reason hereto is probably that the teaching materials were not classified as a mandatory part of the curriculum in Mek1. This will, however, be discussed further in 5.1.

The point of the focus group was to collect qualitative data to clarify (1) whether the layout of the book has had the desired effect; (2) the literacy; and (3) the level of complexity of the book. Also what their overall opinions of the book and screencasts were; and lastly how they could have been improved.

Chapter 3 The DHD and FCI tests and the authoring of the Lab Test

During the course of Mek1 the students took three tests. These are a self authored test called *Lab Test*; the Scottish Data Handling Diagnostic test (DHD); and the American Force Concept Inventory test (FCI) which was given twice.

The order in which the test were given were (1) before any teaching in connection to Mek1 was given the students took the DHD and the FCI; and (2) after the curriculum of Mek1 was taught they were given the Lab Test and the FCI again.

In this section the three tests will be presented with emphasis on the self authored Lab Test. The presentations of the remaining tests, i.e the DHD and FCI, will not be in depth. This is due to the fact that the authors of these tests wish to keep them from the general public; otherwise they will lose they effectiveness as students then simply could study the tests beforehand and thereby taint the results.

3.1 Lab Test

The Lab Test is a self authored test and it was given at the end of Mek1; it is a multiple choice test written in Danish and its purpose is to test the students on basic use of Matlab and data analysis. In total the test consists of 23 questions in which there are three possible answers of which only one, in some questions, is correct. The questions are segmented into three groups which is illustrated in table 3.1.

In case of a question the student does not understand it is noted that he can convey this by submitting an ***F*** to the left of the question number. The time allotted for taking this test was 30 minutes and the only permitted tools was pen, paper, and a calculator.

About media		Matlab	Data analysis
1,2,3,4	**Novice**	5,6,7,8,9	14,15,16,18,19
	Intermediate	10,11,12,13,20,23	17,21,22

Table 3.1 • The table shows how the questions of the Lab Test can be segmented into three different groups. The questions in group about media are not right/wrong questions they instead asked the student about media preference. The questions in the remaining groups are multiple choice of which only one possibility is correct. These can furthermore be divided into two levels of difficulty: Novice and intermediate.

3.1.1 The questions

The first group of questions is questions 1 through 4; these comprises the group about media and are not of a right or wrong nature. They are the basis for analysis into media preference. Translated the questions are as follows:

Q. 1 *Have you at any point used the book "Grundlæggende Matlab og dataanalyse", which can be found on Absalon*[1] *?* Possible answers were *yes, no* and *never heard of it.*

Q. 2 *Have you seen any of the screencasts, which also can be found on Absalon?* Possible answers were *yes, no* and *never heard of them.*

Q. 3 *Which of the following three media did you find most educational?* Possible answers were *the book from Q. 1, the screencasts from Q. 2,* and *internet (youtube.com, computerfysik.dk,.....).*

Q. 4 *Which of the following three media do you feel you benefit most from when learning something new?* Possible answers were *books, screencasts/videos,* and *Notes on the internet.*

Questions within the group Matlab are those on the basic use of Matlab. They test the students on selected commands, used in Matlab. Questions 5,6,7,8, and 9 are categorized as novice questions. These test for the very basic syntax of Matlab: How a variable and an array is defined, what the meaning by placing a semi-colon at the end of a command line is, and how to clear Matlab's Workspace and output in the Command Window.

The remaining questions in the Matlab group are of intermediate difficulty. These tests the student in how to change an element of an already existing array; create an array containing all the even numbers from 1 to 20 with the colon-operator; how to use the plot command to assign arrays to certain axis and to get Matlab to represent the data as points with a specific shape in a plot; put a title in plots, histograms etc.; which command is needed when creating a plot with errorbars; and, lastly, the syntax of a for-loop in a specific setting. Recall that the handling and use of if-clauses were only presented in the book for completeness; as such questions on this topic have not been included.

The last group of questions are on concepts and use of basic data analysis; those categorized as novice are questions 14, 15, 16, 18 and 19; and intermediate 17, 21, and 22.

Questions 14, and 15 tests the students in what the best estimators are for a quantity's true value based on N measurements and the average uncertainty in each of said measurements, i.e whether they know the concepts mean and spread of a data set is the said estimators. In question 16 they are tested for the meaning of a histogram, i.e that a histogram is a visual tool for inspecting how a set of measurements are distributed.

Questions 18, and 19 are simply the computation of the mean and standard deviation of the mean (SDOM) of a data set consisting of five fictitious measurements of the acceleration due to gravity. These questions are linked to question 14 and 15; if the student does not know the best estimator of a quantity's true value to be the mean of the measurements of said quantity, then it is likely he will not know that the mean of the five measurements in question 18 is to be computed. The linkage of questions 15, and 19 are not as straight forward as with 14 and 18, as 14 tests for the concept of spread, whereas in 19 SDOM is to be computed using the provided five measurements. One could suggest that the students, when answering question 19, are somewhat biased as 14 asked about the spread; thus they are perhaps likely to simply compute the spread and not SDOM in question 19.

Theoretical statistical questions are in the Lab Test kept to a minimum as the application of data analysis is in focus; however, since the normal distribution is perhaps *the* most applied distribution in this field, depending on the situation of course, one theoretical question regarding the distribution seemed justified. Question 17 links the concept of spread to the normal distribution. It asks: *Assume that a set of measurements of a certain quantity is normally distributed. Within how many standard deviations will* 68.2% *of the measurements lie?* The answer to this question, 1σ, is actually important in connection to identify possible outliers when inspecting errorbar plots, e.g if the threshold of the uncertainty is 1σ and they are normally distributed.

The last two questions in the data analysis group, 21 and 22, are the most difficult questions. Question

[1] Absalon is where all information and texts can be found online for the courses attended by a student.

21 is directly taken from the DHD question 17. The point of the question is to test whether the student knows if (1) the error propagation formula for independent measurements is to be applied, and (2) how to use it. Question 22 simply tests if the students have grasped the concept of precision of a physical quantity; it bears much resemblance to question 5 of the DHD.

3.1.2 Response errors

When authoring questionnaires *response errors* can arise due to the wording of the questions; more precisely it occurs when the wording is biased. This bias may lead the respondents to submit an answer which they find acceptable rather than their own opinion. An example of this can be found in [Weiers, 1999]:

> *"Shoplifting is not only illegal, but it makes prices higher for everyone. Have you ever shoplifted?".*

By stating the fact that shoplifting is illegal and results in higher prices for everyone, before asking if they ever have shoplifted, will most likely increase the chance that the respondents will answer *no* to ever having shoplifted even when some may actually have.

Such errors could also occur in the Lab Test: If, for example, Q. 1 was formulated as: *Have you at any point, which you should, used the book "Grundlæggende Matlab og dataanalyse", which can be found on Absalon?*. By changing the formulation of Q. 1 to include the underlined part would most likely lead some students to answer *yes* when they in fact may not have used the book or never have heard of it; the wording gives the impression that it is "illegal" not to have used the book.

In light of this the wording of the questions in the Lab Test have been carefully chosen as to not imply bias towards any of the possible choices.

3.2 Data Handling Diagnostic test

The Data Handling Diagnostic test (DHD) is developed by Simon Bates and Ross Galloway from the Edinburgh Physics Education Research Group which is a multiple choice test: As its name suggests, it tests the students' knowledge in the handling of data in the context of physics experiments. Examples on the type of questions are the following: identify the mean, mode, and median for a listed set of data, which fit is the best for a specific set of data, and correctly stating an experimental result among others. In total there are 23 questions and in each there are four possible answers of which only one is correct.

This test has been implemented by NBI to not only shed on light on the knowledge the students have when they start their physics education at KU in regards to data analysis, but also to determine if one full year of teaching have made a positive impact. Thus the students are given the DHD in the start of their first physics course (MEK1), and then again in the course on electrodynamics and waves (EM2), which is directly after their first summer holiday.

As mentioned in the chapter introduction, the authors of this test wish not to make it available to the general public. If this where the case the students, who perhaps knew that it will be given later on, could simply study the test beforehand thereby making the conclusions inferred on the basis of the responses extremely biased. Therefore the DHD is not submitted in this thesis as an appendix, however, see [EdPER, 2013] for contact information to get an electronic copy.

3.3 Force Concept Inventory test

The Force Concept Inventory test (FCI) has, like the DHD, an aptly named title because it directly conveys the subject matter it tests for, which is the concept of force in the field of classical mechanics, i.e Newtonian mechanics. It is authored by Ibrahim Halloun and David Hestenes et al [Halloun, 2013].

In total there are 30 questions; each multiple choice where there are five possible answers of which only one is correct. Every question, and often more than one, is set fourth in a specific context, e.g question 1 reads:

> *"A stone which is dropped from the roof of a one storey building and drops toward the surface [Bearden, 2013a]".*

, and the student is then asked to finish the sentence choosing between the provided five possibilities. Furthermore, the five possible answers in each question are constructed based on common misconceptions in Newtonian mechanics. An example of a common misconception in Newtonian mechanics is: If an object is in motion a force must be present pointing in the direction of motion, and it stops when said force has been "used up" [Angell, 2011].

Again the authors of the FCI do not wish to make it accessible to the general public due to the same arguments given in section 3.2. Therefore the FCI will also not be submitted as an appendix. See [FCI, 2014] for contact to get an electronic copy.

Chapter 4 Theory of statistical tests and concepts

In this chapter we will present the theory for the chosen statistical tests; also we present some needed concepts when dealing with statistical tests. Of statistical concepts we will describe the importance of hypothesis and which types of errors in connection to performing statistical tests; in addition the concept of statistical power is also described.

The tests chosen are the independence test, Analysis of Variance (ANOVA), and Student's paired t-test. Two different methods for testing for independence are here included: The G- and χ^2-independence test. However, when testing for independence using paired data McNemar's test is needed which we therefore also present. In order to perform an ANOVA certain conditions have to be met; to determine if these are met we need to introduce the χ^2-goodness of fit and Levene's test. It should be noted, however, that none of the distributions used in presentation of the statistical tests will be described here. For details of the relevant distributions we refer to appendix C. In addition, we will describe how the power of each statistical test - not including that of the χ^2-goodness of fit and Levene's test - is computed using non-central distributions which are also briefly described in appendix C. Lastly, a discussion as to why we do not use Yates continuity correction in the tests is given.

4.1 Statistical concepts

In order to present the statistical tests we first need to present some concepts used in the field of hypothesis testing. First the basis in hypothesis testing is to know what a null and alternative hypothesis is. Next we see what is meant by type I (α) and type II errors (β); and also how they are related to the the level of significance and power of a test. When the data is categorical we need to make use of contingency tables; lastly we therefore present what a two way contingency table is and what it contains.

4.1.1 The null and alternative hypothesis

In hypothesis testing the experimenter first has to set fourth two hypothesis; the *null* and *alternative hypothesis*. These are mathematical formulations of what he seeks to investigate. The symbols often used in the statement of the null and alternative hypothesis are $>, <, =$ and $\neq$.

Most professional journals always use = to state the null hypothesis as it usually is a statement which equates a quantity to a specific value [Blumann, 2012, page 402]. The remaining symbols are used in the mathematical formulation of H_A which usually is the research question sought to be investigated; as such it is also often referred to as the research hypothesis [Blumann, 2012, page 402].

The null hypothesis, H_0, is assumed to be valid from the start. By performing the relevant statistical test it is determined whether or not the data shows significant evidence against H_0 or not, i.e if we can infer that the observations are not due to random chance. If so H_0 is rejected and the alternative hypothesis, H_A, is accepted. Conversely, if there is no statistical significant evidence against H_0 one states "we fail to reject

H_0". Note the formulation "fail to reject", which is due to the fact that even though there is no significant evidence does not prove that H_0 is true! In order to prove H_0 would require data from the entire population, which is exceedingly rare, rather than a sample from it.

A quote from Ronald Fisher, who is regarded as the chief architect of the foundations for modern statistical science [Hald, 2004, page 147], and also coined the term null hypothesis, describes the concept of the null hypothesis quite well:

> *"In relation to any experiment we may speak of this hypothesis as the 'null hypothesis', and it should be noted that the null hypothesis is never proved or established, but is possibly disproved, in the course of experimentation. Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis [WikiQuote, 2013].".*

4.1.2 Types of errors

When performing hypothesis tests there exists certain types of errors which can occur; these are commonly called type I, and type II errors, respectively. A type I error is rejecting H_0 when it is actually true, i.e a false negative, and a type II error is not rejecting H_0 when it is false, i.e a false positive. Both [Barlow, 1999, page 142] and [Trosset, 2008, page 211-212] explains this very well by using a court of law as an analogy:

> *"In the law courts, the accused proclaims the hypothesis that he is innocent. If the jury rejects this and wrongly convict him when he is really innocent, that is a type I error. If they accept his hypothesis and let him off when he is really guilty, that is a type II error. [Barlow, 1999, page 142]".*

When investigating a quantity it will under H_0 be distributed according to a some probability density function (pdf) $p_0(x)$ and under H_A according to $p_A(x)$. This is shown in figure 4.1. Beforehand a limit has to be set in regards to the probability for committing a type I error which we are willing to accept. This

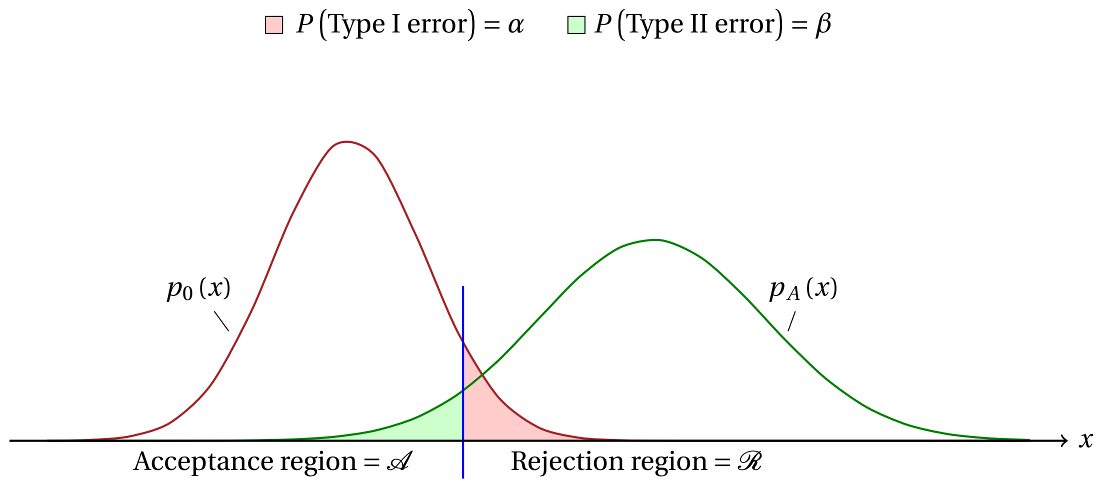


FIGURE 4.1 • THIS FIGURE SHOWS VISUALLY HOW TO PERCEED THE CONCEPTS OF TYPE I AND II ERRORS. UNDER H_0 THE PROBABILITY OF COMMITING A TYPE I ERROR, I.E THE PROBABILITY OF REJECTING H_0 WHEN IT IS IN FACT TRUE, IS OBTAINED BY INTEGRATING THE PDF FOR H_0 (p_0) OVER THE REJECTION REGION. CONVERSELY, THE PROBABILITY OF NOT REJECTING H_0 WHEN IT IS IN FACT FALSE, I.E TYPE II ERROR, IS OBTAINED BY INTEGRATING THE PDF FOR H_A (p_A) OVER THE REJECTION REGION.

probability is called the level of significance denoted by α. The popular choices of α are 1%, 5%, and 10%; however, 5% is, of the three, the most common. Once an α has been chosen we then know our acceptance and rejection region. Now, say, we actually knew H_0 to be true and the measurements results in a value which falls in the acceptance region, then we would correctly fail to reject H_0; on the other hand if the value falls in the rejection region, we wrongly reject H_0, i.e we commit a type I error. If instead we knew H_A to be true but our measurements resulted in a value which falls in the acceptance region then we would wrongly fail to reject H_0; as such we have committed a type II error which is denoted by β. Again we refer to figure 4.1.

Mathematically, the probability of committing a type I and type II error is given by

$$P\left(\text{Type I error}\right) = \alpha = \int_{\mathscr{R}} p_0(x)\,dx \quad \text{and} \quad P\left(\text{Type II error}\right) = \beta = \int_{\mathscr{A}} p_A(x)\,dx \tag{4.1}$$

where $\mathscr{R}$ and $\mathscr{A}$ is the rejection and acceptance region, respectively.

Statistical power of a test is defined as the probability of *not* committing a type II error. Due to the property that a pdf integrates to unity over its defined region and the properties of integration then statistical power is simply given by

$$1 = \int_{\mathscr{A}+\mathscr{R}} p_A(x)\,dx = \underbrace{\int_{\mathscr{A}} p_A(x)\,dx}_{\text{Power}} + \underbrace{\int_{\mathscr{R}} p_A(x)\,dx}_{\beta} \Rightarrow \text{Power} = 1-\beta. \tag{4.2}$$

Therefore statistical power is the probability of correctly rejecting H_0. See figure 4.2.

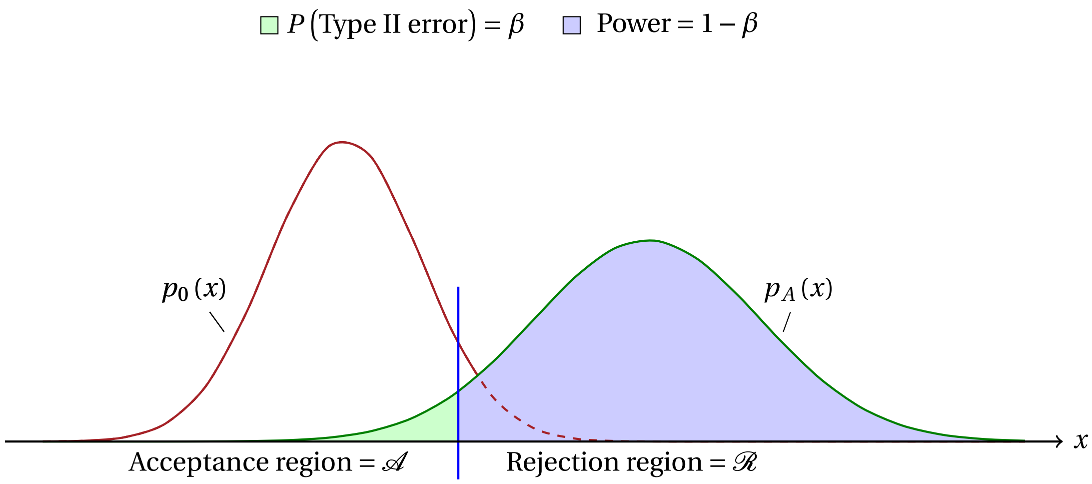


FIGURE 4.2 • THE POWER OF A STATISTICAL TEST IS FOUND BY INTEGRATING THE PDF FOR H_A OVER THE REJECTION REGION $\mathscr{R}$, WHICH IS EQUAVIVALENT TO $1-\beta$. AS β IS THE PROBABILITY OF COMITTING A TYPE II ERROR THE POWER IS THUS THE PROBABILITY OF NOT COMMITING SAID ERROR, I.E CORRECTLY REJECTING H_0.

4.1.3 Contingency tables

The, in the next section presented, independence tests make use of two way contingency tables which form the basis of the calculation of the test statistics. A contingency table groups the data into categorical variables and their respective levels. The respective categorical levels are equivalent to the number of rows and columns, which we shall denote r and c, respectively. The ij'th cell of the table contains observed frequencies O_{ij} which falls within the i'th level of category A and the j'th level of category B, i.e the i'th row and j'th column; this is illustrated in table 4.1.

Cat. A \ Cat. B	**Level** 1	**Level** 2	$\cdots$	**Level** j	$\cdots$	**Level** c	
Level 1	O_{11}	O_{12}	$\cdots$	O_{1j}	$\cdots$	O_{1c}	R_1
Level 2	O_{21}	O_{22}	$\cdots$	O_{2j}	$\cdots$	O_{2c}	R_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Level i	O_{i1}	O_{i2}	$\cdots$	O_{ij}	$\cdots$	O_{ic}	R_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Level r	O_{r1}	O_{r2}	$\cdots$	O_{rj}	$\cdots$	O_{rc}	R_r
	C_1	C_2	$\cdots$	C_j	$\cdots$	C_c	$N = \sum_{i,j=1}^{r,c} O_{ij}$

Row sums

Column sums

TABLE 4.1 • THIS TABLE ILLUSTRATES HOW A TWO WAY $r \times c$ CONTINGENCY TABLE LOOKS LIKE. THE O_{ij}'S ARE THE OBSERVED FREQUENCIES, R_i AND C_j ARE THE i'TH AND j'TH ROW AND COLUMN SUM, RESPECTIVELY.

4.2 Independence tests

In order to determine whether two categorical variables, say A and B, are dependent on each other, or not, an independence test can be performed. Depending on the situation the null hypothesis is usually that A and B are independent and the alternative is the converse, i.e dependent. Mathematically this is represented as

$$H_0 : P(AB) = P(A)P(B)$$
$$H_A : P(AB) \neq P(A)P(B).$$

In this section we will present the theory behind an independence test in which H_0 and H_A are given as above: First we see how it is done by the method of maximum likelihood, and thereafter the widely known Pearson χ^2 method. Lastly, we also present McNemar's χ^2-test which is used when testing for dependence when the data is paired. The notation constructed in section 4.1.3 will be used throughout this section.

4.2.1 G-test of independence

The method of maximum likelihood is not some mathematical justified method. It is, however, just a sensible way of producing an estimator which is a mathematical expression for the greatest probability of observing the recorded data [Barlow, 1999, page 89].

In this case we wish to test for independence between two categorical variables A and B using this method; such a test is usually called a G-test of independence, or more commonly a likelihood ratio test of independence. First we need to determine the likelihood function; from this function we then determine an estimator under (1) H_0 of independence, and (2) H_A of dependence; and lastly we then compute the ratio between the likelihood function under H_0 and that the H_A which is the test statistic.

In the following we denote the probability of being in the ij'th cell of the contingency table (see table

4.1) by θ_{ij}, and that of being in the i'th row and j'th column by ϕ_i and ρ_j, respectively. Thus our hypothesis are $H_0 : \theta_{ij} = \phi_i \rho_j$, and $H_A : \theta_{ij} \neq \phi_i \rho_j$.

DETERMINING THE LIKELIHOOD FUNCTION: The probability of obtaining a contingency table shown in table 4.1, i.e the observed data, is given by

$$\mathscr{L}\left(p_{11}, \ldots, p_{ij}, \ldots, p_{rc}\right) = \prod_{i,j}^{r,c} p_{ij}^{O_{ij}}, \tag{4.3}$$

where p_{ij} is the probability of an event resulting in O_{ij}, and the real function $\mathscr{L}$ is called the *likelihood function*. The observed frequencies O_{ij} will be distributed according to a multinominal distribution with parameters $\left(N, \theta_{11}, \ldots, \theta_{ij}, \ldots, \theta_{rc}\right)$ [Skovgaard, 1999, page 170], implying that the likelihood function is specifically given by

$$\mathscr{L}(\theta) = N! \prod_{i,j=1}^{r,c} \frac{\left(\theta_{ij}\right)^{O_{ij}}}{O_{ij}!}. \tag{4.4}$$

Taking the natural logarithm on both sides of eq. (4.4) we obtain the log-likelihood function:

$$\ell\left(\theta_{ij}\right) = \ln\left(\mathscr{L}\left(\theta_{ij}\right)\right) = \ln\left(N! \prod_{i,j}^{r,c} \frac{\left(\theta_{ij}\right)^{O_{ij}}}{O_{ij}!}\right) = \sum_{i,j=1}^{r,c} \ln(N!) + O_{ij} \ln\left(\theta_{ij}\right) - \sum_{i,j=1}^{r,c} \ln\left(O_{ij}!\right), \tag{4.5}$$

which is much easier mathematically to handle. We note that the log-likelihood function can be used instead of the original: The natural logarithm is a monotonically increasing function which implies that $\mathscr{L}\left(\theta_{ij}\right)$ and $\ell\left(\theta_{ij}\right)$ peaks at the same θ_{ij} [Skovgaard, 1999, page 110].

ESTIMATOR UNDER H_0: Under H_0, i.e if H_0 is true, we set $\theta_{ij} = \phi_i \rho_j$ and maximize $\ell\left(\theta_{ij}\right)$ w.r.t ϕ_i and ρ_j under the conditions that

$$\sum_{i=1}^{r} \phi_i = \sum_{j=1}^{c} \rho_j = 1, \quad \text{and} \quad \sum_{i=1}^{r}\sum_{j=1}^{c} O_{ij} = N = \sum_{i=1}^{r} R_i = \sum_{j=1}^{c} C_j. \tag{4.6}$$

Determining the partial derivatives w.r.t ϕ_i and ρ_j, respectively, and equating to zero yields

$$\frac{\partial \ell\left(\phi_i, \rho_j\right)}{\partial \phi_i} = \sum_{j=1}^{c} \frac{O_{ij} \rho_j}{\phi_i \rho_j} = \sum_{j=1}^{c} \frac{O_{ij}}{\phi_i} = 0 \tag{4.7}$$

$$\frac{\partial \ell\left(\phi_i, \rho_j\right)}{\partial \rho} = \sum_{i=1}^{r} \frac{O_{ij} \phi_i}{\phi_i \rho_j} = \sum_{i=1}^{r} \frac{O_{ij}}{\rho_j} = 0. \tag{4.8}$$

$\sum_{i=1}^{r} O_{ij}$ and $\sum_{j=1}^{c} O_{ij}$ equals the j'th column sum C_j and i'th row sum R_i, respectively, of the contingency table; as such we have $R_i \rho_j = C_j \phi_i$. If we sum over the columns, i.e index j, we obtain the maximum likelihood estimator $\hat{\phi}_i$; if we instead sum over the rows, i.e index i, we obtain $\hat{\rho}_j$:

$$\sum_{i=1}^{r} R_i \rho_j = \sum_{i=1}^{r} C_j \phi_i \Rightarrow \hat{\rho}_j = \frac{C_j}{N} \tag{4.9}$$

$$\sum_{j=1}^{c} R_i \rho_j = \sum_{i=j}^{c} C_j \phi_i \Rightarrow \hat{\phi}_i = \frac{R_i}{N}, \tag{4.10}$$

where we have used the conditions from eq. (4.6).

To satisfy that $\hat{\phi}_i$ and $\hat{\rho}_j$ are indeed maximums we compute the second derivative of the log-likelihood function and evaluate:

$$\left.\frac{\partial^2 \ell\left(\phi_i,\rho_j\right)}{\partial {\phi_i}^2}\right|_{\phi_i=\hat{\phi}_i} = -\sum_{j=1}^{c}\frac{O_{ij}}{\left(\hat{\phi}_i\right)^2} \quad \text{and} \quad \left.\frac{\partial^2 \ell\left(\phi_i,\rho_j\right)}{\partial {\rho_j}^2}\right|_{\rho_j=\hat{\rho}_j} = -\sum_{i=1}^{r}\frac{O_{ij}}{\left(\hat{\rho}_j\right)^2}. \tag{4.11}$$

Both second derivatives evaluated at $\hat{\phi}_i > 0$ and $\hat{\rho}_j > 0$, respectively, are less than zero as O_{ij} is always equal to or greater than zero; this implies that they are indeed maximums. The argument that $O_{ij} \geq 0$ is simply that negative observed frequencies is not possible and it would furthermore not make sense.

ESTIMATOR UNDER H_A: Under H_A we now maximize $\ell\left(\theta_{ij}\right)$ in eq. (4.5) w.r.t to θ_{ij} under the conditions

$$\sum_{i=1}^{r}\sum_{j=1}^{c}\theta_{ij} = 1, \quad \text{and} \quad \sum_{i=1}^{r}\sum_{j=1}^{c}O_{ij} = N. \tag{4.12}$$

Hereto we use the method of Lagrange multipliers; thus

$$0 = \frac{\partial}{\partial\theta_{ij}}\left(\ell\left(\theta_{ij}\right) - \lambda\left(1 - \sum_{i,j=1}^{r,c}\theta_{ij}\right)\right) = \frac{O_{ij}}{\theta_{ij}} - \lambda \Rightarrow \lambda = \frac{O_{ij}}{\theta_{ij}} \tag{4.13}$$

$$\tag{4.14}$$

Now summing over both the i'th and j'th index using the conditions in eq. (4.12) we find

$$\sum_{i,j=1}^{r,c}\lambda\theta_{ij} = \sum_{i,j=1}^{r,c}O_{ij} \Rightarrow \lambda = N, \tag{4.15}$$

which implies that the maximum likelihood estimator for θ_{ij} under H_A is $\hat{\theta}_{ij} = O_{ij}/N$. The estimator is indeed a maximum as

$$\left.\frac{\partial^2 \ell\left(\theta_{ij}\right)}{\partial {\theta_{ij}}^2}\right|_{\theta_{ij}=\hat{\theta}_{ij}} = -\frac{O_{ij}}{\hat{\theta}_{ij}} \tag{4.16}$$

which clearly is a maximum.

THE TEST STATISTIC: Here the test statistic is defined to be the ratio between the likelihood function evaluated under H_0 and that of H_A, i.e

$$\Lambda = \frac{L\left(\hat{\phi}_i,\hat{\rho}_j\right)}{L\left(\hat{\theta}_{ij}\right)} = \prod_{i,j=1}^{r,c}\frac{\left(\hat{\phi}_i\hat{\rho}_j\right)^{O_{ij}}}{\left(\hat{\theta}_{ij}\right)^{O_{ij}}} = \prod_{i,j=1}^{r,c}\left(\frac{NR_iC_j}{N^2O_{ij}}\right)^{O_{ij}} = \prod_{i,j=1}^{r,c}\left(\frac{R_iC_j}{NO_{ij}}\right)^{O_{ij}} \tag{4.17}$$

It is, however, difficult to determine how Λ will be distributed. Luckily we can use an approximation: The American statistician Samuel S. Wilks showed that $G = -2\ln\Lambda$ will be distributed according to a χ^2-distribution with $\mathrm{df}(H_A) - \mathrm{df}(H_0)$ degrees of freedom for $N \to \infty$ [Shalabh, 2009, page 339]. Note that $\mathrm{df}(Q)$ is a short hand notation meaning "number of degrees of freedom under Q". The approximate test statistic G is therefore:

$$G = 2\sum_{i,j=1}^{r,c}O_{ij}\ln\left(\frac{NO_{ij}}{R_iC_j}\right). \tag{4.18}$$

Lastly, we need to determine the number of degrees of freedom represented by the difference $\mathrm{df}(H_A) - \mathrm{df}(H_0)$. Under H_A there are rc cell probabilities θ_{ij} which must sum up to 1. Due to this constraint we lose a degree of freedom, i.e $\mathrm{df}(H_A) = rc - 1$. Under H_0 we have r row and c column probabilities which both must sum up to 1 hereby losing two degrees of freedom, i.e $\mathrm{df}(H_0) = r + c - 2$. Therefore we have

$$\mathrm{df}(H_A) - \mathrm{df}(H_0) = rc - 1 - (r + c + 2) = rc - r - c + 1 = (r-1)(c-1). \tag{4.19}$$

The probability of independence between categorical variable A and B is therefore given by

$$P(G \leq \infty) = \int_G^\infty \chi^2 (x; \mathrm{df} = (r-1)(c-1))\, dx, \tag{4.20}$$

and if $P(G \leq \infty)$ is less than the level of significance α we reject H_0.

4.2.2 Pearson's χ^2 - test

Karl Pearson, an English mathematician, greatly contributed to the field of statistics: His perhaps greatest legacy is his χ^2 goodness of fit test of which we in fact will use to test for normality in section 4.3. His goodness of fit test can also be used as an independence test, and was greatly applied as it involved far less computation in comparison to the method of maximum likelihood. However, the ease of computation came at the expense of accuracy: The test statistic used is an approximation of that of the G-test which makes it an approximation of an approximation.

In section 4.2.1 the test statistic G, was found:

$$G = 2 \sum_{i,j=1}^{r,c} O_{ij} \ln\left(\frac{O_{ij}}{E_{ij}}\right), \tag{4.21}$$

which is approximately χ^2-distributed with $(r-1)(c-1)$ degrees of freedom. We note here that we have defined $E_{ij} = R_i C_j / N$ which is the expected frequency of the ij'th cell of the contingency table: Recall, that under the hypothesis of independence the best probability of being in the ij'th is $\hat{\theta}_{ij} = \hat{\phi}_i \hat{\rho}_j$ which implies that the expected number of observed frequencies is simply $N\hat{\phi}_i\hat{\rho}_j = N(R_i/N)\left(C_j/N\right) = R_i C_j/N$.

By Taylor expanding G w.r.t to O_{ij} to the second order around the point E_{ij} and keeping quadratic terms yields

$$2 \sum_{i,j=1}^{r,c} O_{ij} \ln\left(\frac{O_{ij}}{O_{ij}}\right) + 2\left(\sum_{i,j=1}^{r,c} \ln\left(\frac{O_{ij}}{O_{ij}}\right) + 1\right)\left(O_{ij} - E_{ij}\right) + 2\sum_{i,j=1}^{r,c} \frac{\left(O_{ij} - E_{ij}\right)^2}{2E_{ij}} \approx \sum_{i,j=1}^{r,c} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}. \tag{4.22}$$

This is the Pearson χ^2-test statistic, denoted by X^2, i.e

$$X^2 = \sum_{i,j=1}^{r,c} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \approx G. \tag{4.23}$$

Since X^2 is an approximation of G Wilks' result must also apply for large sample sizes, i.e X^2 will also approximately be distributed according to a χ^2-distribution with $(r-1)(c-1)$ degrees of freedom. This implies that the probability of independence between categorical variables A and B are given as the same expression as in eq. (4.20).

In cases of low statistics the approximation, that both X^2 and G being χ^2-distributed, may be very inaccurate. However, a rule of thumb exists, in form of a lower bound in regard to, perhaps somewhat oddly, the E_{ij}'s of the contingency table and not the O_{ij}'s: The rule of thumb is when (1) $E_{ij} \geq 1$ for all i and j's; and (2) at least 80% of the E_{ij}'s are equal to or larger than 5, the approximation is valid [Skovgaard, 1999, page 171].

4.2.3 McNemar's χ^2-test

The above presented G- and Pearson's χ^2-test of independence is valid when the data is categorical and the samples random. To deal with data that are paired, e.g testing the same patient before and again after a treatment, we need to use McNemar's χ^2-test of independence. Although treatment may be an ample word for the action planned by the experimenters between the measurements, in a medicinal setting, we will instead refer to it as an *event* to make it completely general.

This test uses, like the two previous independence tests, contingency tables. However, each category has, in the setting of this thesis, only two levels since we only record success and failure before and after the event; this implies a 2 by 2 contingency table shown in table 4.2. Illustratively, the cell 11 is the observed frequency which had success before *and* after the event.

The null hypothesis here is that the event has made no impact, i.e the probability of success before,

Before \ After	**Success**	**Failure**	
Success	O_{11}	O_{12}	R_1
Failure	O_{21}	O_{22}	R_2
	C_1	C_2	$N = O_{11} + O_{12} + O_{22} + O_{21}$

TABLE 4.2 • CONTINGENCY TABLE USED IN MCNEMAR'S INDEPENDENCE TEST.

$P(S|B) = R_1/N$, and after, $P(S|A) = C_1/N$, are equal; the alternative is the converse, i.e $P(S|B) \neq P(S|A)$. Formally,

$$H_0 : P(S|B) = P(S|A) = 1/2$$
$$H_A : P(S|B) \neq P(S|A).$$

Logically, those who we have recorded success or failure for, both before and after the event, does not contribute any information as to the impact, positively or negatively, of the event; as such the only cells which carry this information are the discordant pairs[1] O_{12} and O_{21}. The total number of discordant pairs D will be distributed according to a binomial distribution with $O_{12} + O_{21}$ trials and parameter of success $p = 1/2$ [Shalabh, 2009, page 38], as the only outcomes are binary, i.e success or failure, and the chance of observing two out of four possible scenarios, randomly, is 50%. Therefore the expected value of observations, with success before and failure after, is $pn_{\text{trials}} = D/2$ with variance $n_{\text{trials}}p(1-p) = D \cdot 1/2 \cdot 1/2$ [Shalabh, 2009, page 38]. Symmetrically, this also holds for the observations with failure before and success after. This implies that the ratio

$$M = \frac{O_{12} - D/2}{\sqrt{D \cdot 1/2 \cdot 1/2}} = \frac{O_{12} - O_{21}}{\sqrt{O_{12} + O_{21}}} \tag{4.24}$$

will, under H_0, be approximately distributed according to a standard normal distribution, i.e a gaussian with zero mean and unit variance, when D is sufficiently large. It should be noted that some disagreement in literatures exist as to when D is sufficiently large; [Shalabh, 2009, page 38] claims $D \geq 20$ and [Agresti, 2007, page 246] $D > 10$. We choose, here, to accept the approximation when $D \geq 10$.

By squaring both sides in eq. (4.24) we get the final test statistic

$$M^2 = \frac{(O_{12} - O_{21})^2}{O_{12} + O_{21}} \tag{4.25}$$

[1] A pair (X, Y) is discordant if X from category, say A, is higher or lower in rank than Y from category, say B, assuming both categories have same ranking. Therefore if we record success/failure before and failure/success after it is a discordant pair.

which is computationally more simple; however, since we have squared a standard normally distributed variable it will now, as explained in section C.1 in appendix C, be χ^2-distributed with df = (r-1)(c-1) = (2-1)(2-1) = 1.

Knowing that M^2 will approximately be χ^2-distributed with df = 1 the probability of the event to have made no impact, i.e independence, is given by

$$P\left(M^2 \leq \infty\right) = \int_{M^2}^{\infty} \chi^2\left(x; \mathrm{df} = 1\right)\, dx, \tag{4.26}$$

and if $P\left(M^2 \leq \infty\right)$ is less than the level of significance α we reject H_0.

4.3 Analysis of Variance

In the field of social sciences and medicine this statistical test is often applied. The name *Analysis of Variance*, abbreviated ANOVA, implies that it is a statical test of population variances. This is, however, not the case: It instead uses variances to infer if populations can be considered to have equal means. The null and alternative hypothesis thus are

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \tag{4.27}$$
$$H_A : \mu_i \neq \mu_j, \quad \text{for at least one pair } (i, j) \in \mathbb{N}. \tag{4.28}$$

Several types of ANOVA exists; specifically the one described and relevant here is the *one way ANOVA with fixed effects*. The term *one way* means that the populations differ with respect to only one factor, e.g health condition, and *fixed effects* implies that conclusions can only be drawn for the populations that were sampled [Mickey, 2004, page 38].

The one way ANOVA with fixed effects assumes that [Mickey, 2004, page 39]:

1. the $k \in \mathbb{N} \setminus \{1\}$ samples, each consisting of n_i measurements x_{ij}, are independent random samples from k populations ($i \in [1, k]$ and $j \in [1, n_i]$),

2. each of the k populations is normally distributed, and

3. the variances of the k populations σ_i are equal.

Therefore it is necessary to check whether these assumptions can be considered to be met before an ANOVA can be performed.

Note that the samples in this thesis are de facto randomly independent drawn due the nature by which the data has been collected. The argument for this statement is given in section 5.1 in chapter 5.

4.3.1 Testing for normality

In order to investigate whether an underlying population can be considered normally distributed or not - on the basis of a sample - we turn to the χ^2 - goodness of fit test for binned data. The null and alternative hypothesis are

$$H_0 : \text{The measurements are normally distributed.}$$
$$H_A : \text{The measurements are } \textit{not} \text{ normally distributed.}$$

Let x_r be the r'th element of N measurements, and b_r be the r'th bin - of which there in total are B - with bin width A and centred around the point x_r. Here we only consider a uniform bin width. If the measurements are normally distributed, i.e distributed according to the pdf

$$p(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \tag{4.29}$$

where $x \in \mathbb{R}$, μ is the mean, and σ the spread, then the expected number of events in b_r is $f_r(x_r) = NAp(x;\mu,\sigma)$. Given that the measurements are binned we deal with counting statistics which implies that the error of each bin is described by Poisson statistics; as such the error in bin b_r is $\varepsilon_r = \sqrt{n_r}$ [Barlow, 1999, page 105-106].

On this basis the test statistic, which we will denote $(X_{\text{GOF}})^2$, is

$$(X_{\text{GOF}})^2 = \sum_{r=1}^{B} \left(\frac{n_r - f_r(x_r)}{\varepsilon_r}\right)^2. \tag{4.30}$$

We fit for μ, σ, and normalization constant which implies that the degrees of freedom are $B-3$. Note that the term normalization constant is here used differently than the usual convention: Given some data that are normally distributed the usual normalization constant K is such that

$$\int_{-\infty}^{\infty} \frac{K}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} = 1. \tag{4.31}$$

Here we define $K/\left(\sqrt{2\pi}\sigma\right)$ as the normalization constant.

The test statistic $(X_{\text{GOF}})^2$ will be approximately $\chi^2(x; \text{df} = B-3)$-distributed; as such the probability of the binned measurements being normally distributed is given by

$$P\left((X_{\text{GOF}})^2 \leq \infty\right) = \int_{(X_{\text{GOF}})^2}^{\infty} \chi^2(x; \text{df} = B-3)\, dx. \tag{4.32}$$

As usual if $P\left((X_{\text{GOF}})^2 \leq \infty\right) < \alpha$ we reject H_0.

4.3.2 Testing for homoscedasticity

To investigate whether the underlying k populations can be considered, statistically, to have equal variances, i.e homoscedasticity, we perform the Levene's test using the k samples. The null and alternative hypothesis are

H_0 : Homoscedasticity of the k samples.

H_A : At least one sample has a different variance than the remaining.

The test statistic for the Levene's test requires that we transform the measurements of the k samples: We define

$$Z_{ij} = |r_i - \mu_i|, \tag{4.33}$$

where μ_i is the mean of the i'th sample. Furthermore we denote N as the total number of measurements, i.e $N = n_1 + n_2 + \cdots + n_k$; the overall mean of the Z_{ij}'s by $\overline{Z}$ given by

$$\overline{Z} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} Z_{ij}; \tag{4.34}$$

and the i'th transformed sample mean by $\overline{Z}_i$, i.e

$$\overline{Z}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} Z_{ij}. \tag{4.35}$$

Using this notation the test statistic W is then given by

$$W = \frac{N-k}{k-1} \cdot \frac{\sum_{i=1}^{k} n_i \left(\overline{Z}_i - \overline{Z}\right)^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(Z_{ij} - \overline{Z}_i\right)^2}. \tag{4.36}$$

W will be distributed according to Fisher's F-distribution with $k-1$ and $N-k$ degrees of freedom [US, 2013], i.e the probability of homoscedasticity is given by

$$P(W \leq \infty) = \int_W^{\infty} F(x; \mathrm{df}_1 = k-1, \mathrm{df}_2 = N-k)\; dx. \tag{4.37}$$

If $P(W \leq \infty)$ is lower than the level of significance α we reject H_0.

4.3.3 The ANOVA test statistic

When normality and homescedasticity statistically both have been established we can perform the ANOVA; hereto we need a test statistic which we now will present along with relevant expressions.

Let

$$\overline{\mu} = \frac{1}{N}\sum_{i=1}^{k} n_i \mu_i, \quad \text{where} \quad N = \sum_{i=1}^{k} n_i, \tag{4.38}$$

denote the overall mean, which is a weighted average of the individual k population means. By introducing the quantity

$$\gamma = \sum_{i=1}^{k} n_i \left(\mu_i - \overline{\mu}\right)^2 = \sum_{i=1}^{k} n_i \left(\tau_i\right)^2 \geq 0, \tag{4.39}$$

directly implies that H_0 and H_A in eq. (4.27) and (4.28), respectively, are equivalent to

$$H_0 : \gamma = 0 \tag{4.40}$$

$$H_A : \gamma > 0. \tag{4.41}$$

τ_i is generally referred to as the effect of the i'th sample [Mickey, 2004, page 37].

Since we know nothing about the underlying populations we need to estimate the relevant measures using the samples. Estimating γ we use [Trosset, 2008, page 315]

$$ss_B = \sum_{i=1}^{k} n_i \left(\overline{x}_i\right)^2 - \frac{1}{N}\left(\sum_{i=1}^{k} n_i \overline{x}_i\right)^2, \tag{4.42}$$

where $\overline{x}_i$ is the i'th sample mean. Also we have to provide an estimate of the variance of the populations. Hereto we use the pooled sample variance given by

$$(s_P)^2 = \frac{1}{N-k} \sum_{i,j=1}^{k,n_i} \left(x_{ij} - \overline{x}_i\right)^2 \tag{4.43}$$

which then is used to construct the quantity

$$ss_W = (N-k)\,(s_P)^2 = \sum_{i,j=1}^{k,n_i} \left(x_{ij} - \overline{x}_i\right)^2. \tag{4.44}$$

According to corollary 12.1 in [Trosset, 2008, page 319] the fraction

$$F = \frac{ss_B\,/\,(k-1)}{ss_W\,/\,(N-k)} = \frac{1}{k-1}\frac{ss_B}{(s_P)^2} \tag{4.45}$$

will under H_0 be distributed according to Fisher's F-distribution with $\mathrm{df}_1 = k-1$ and $\mathrm{df}_2 = N-k$ degrees of freedom. See section C.2 in appendix C for further details of the F-distribution.

Thus if the probability of observing an value equal to or greater than F is less than the level of significance α, i.e

$$P\,(F \leq \infty) = \int_F^\infty F_{k-1,N-k}\,(x)\;dx < \alpha, \tag{4.46}$$

we reject H_0.

4.4 Student's paried t-test

Student's paired t-test is used, as the name suggests, when samples can be paired, i.e they are dependent on each other in some way; specifically, when two samples can be combined into one consisting of matched pairs. This matched design is, among others, applied in psychology, medicine, and agriculture to determine whether, for instance, a specific treatment is more effective than another, or simply to test the effects of a new drug vs. a placebo by determining if the sample means can be considered unequal. As in the presentation of the McNemar's χ^2-test we will, in the following, refer the action taken by the experimenter between measurements of the subjects as event to make such action, whatever it may have been, general. In addition we here only consider the case where the underlying variances are not known.

We denote the two samples as k_1 and k_2, respectively, each with sizes n. The two samples are combined to one consisting of n differences, δ_i, between k_2 and k_1, i.e

$$\delta_i = k_{2i} - k_{1i}. \tag{4.47}$$

If the event has had no effect we would expect the mean of the δ_i's, $\overline{\delta}$, to be equal to zero; if it have had an positive effect, which we would like to investigate, $\overline{\delta}$ should be significantly greater than zero. Therefore, the mathematical formulation of H_0 and H_A is as follows:

$$H_0 : \overline{\delta} = 0 \quad \text{and} \quad H_A : \overline{\delta} > 0. \tag{4.48}$$

As we know nothing about the underlying variance of the sample of δ_i's, $(\sigma_\delta)^2$, it is necessary to estimate it; an unbiased estimator for this is given by

$$(s_\delta)^2 = \sum_{i=1}^{n} \frac{\left(\delta_i - \overline{\delta}\right)^2}{n-1}. \tag{4.49}$$

Assuming that the d_i's, are normally distributed, under H_0, with mean 0 and variance $(\sigma_\delta)^2$, implies that the relation

$$T = \frac{\overline{\delta} \cdot \sqrt{n}}{s_\delta} \tag{4.50}$$

will be distributed according to a t-distribution with $n-1$ degrees of freedom [Shalabh, 2009, page 22]. See section C.3 in appendix C for further details of Student's t-distribution. Therefore, the probability of observing a value greater than or equal to T is given by

$$P(T \leq \infty) = \int_T^\infty t(x; \mathrm{df} = n-1)\, dx. \tag{4.51}$$

As usual if $P(T \leq \infty) < \alpha$ we reject H_0.

4.5 Estimating power of a statistical test

Power is most often very complicated to compute, and there are entire books devoted to this very concept, i.e power analysis. One of the great names in this field is Jacob Cohen; it is his procedures that we will implement when estimating the power of the performed statistical tests.

All of the test statistics yield a value which, under H_0, will be distributed according to a certain statistical distribution; in the independence tests they all followed a χ^2-distribution, in ANOVA the F-distribution, and in Student's paired t-test the Student's t-distribution. Recall, that power is defined as $1-\beta$ where β is the probability of committing a type II error; as such when computing the power of a test we need to use the distribution of the test statistic when H_A is true. Under H_A the test statistic will be distributed according to its corresponding non-central distribution. Section C.4 in appendix C briefly describes the relevant non-central distributions. The main difference is that non-central distributions takes an additional parameter which is commonly referred to as the non-centrality parameter. To distinguish the non-centrality parameters we will hence forth denote them as follows:

- λ_{χ^2} for the non-centrality parameter of the non-central χ^2-distribution,
- λ_F for the non-centrality parameter of the non-central F-distribution, and
- λ_t for the non-centrality parameter of the non-central Student's t-distribution.

When $\lambda_{\chi^2} = \lambda_F = \lambda_t = 0$ the non-central distributions becomes central, i.e the ordinary χ^2-, F-, and Student's t-distribution.

To compute the power of each test the non-centrality parameter has to be estimated; this is done by estimating the effect size - a standardized measure of the magnitude of observed effect - and simply knowing the sample size. The effect size is computed differently depending of the performed test; using Cohen's procedures, the estimated effect size of a test is in a χ^2-independence test $w = \sqrt{U/N}$ where U can be X^2, G or M^2; in ANOVA $f = \sigma_m / (s_P)^2$ where σ_m is given by

$$\sigma_m = \sqrt{\frac{1}{N}\sum_{i=1}^{k} n_i (\tau_i)^2}; \tag{4.52}$$

and in Student's paired t-test $d = \overline{\delta} / s_\delta$ [Cohen, 1988, page 48,216-217,360].

Using the estimated effect sizes and the sample sizes the non-centrality parameters is given by [Cohen, 1988, page 216-217,544-551]

$$\lambda_{\chi^2} = Nw^2 = X^2 = G, \quad \lambda_F = Nf^2, \quad \text{and} \quad \lambda_t = d\sqrt{n}, \tag{4.53}$$

where n are the sample size of pairs, i.e the number of δ_i's. Furthermore, Cohen divided the possible values of effect sizes in to three classifications; small, medium, and large. Depending on which test is performed table 4.3 shows how the value of effect size is classified [Cohen, 1988, page 40,226,355].

The estimates of the non-centrality parameters eq. (4.53) form the basis in some studies to a priori

	Independence test	**ANOVA**	**Student's paired t-test**
Small	≥ 0.1	≥ 0.1	≥ 0.2
Medium	≥ 0.3	≥ 0.25	≥ 0.5
Large	≥ 0.5	≥ 0.4	≥ 0.8

TABLE 4.3 • THIS TABLE SHOWS HOW COHEN DEFINES THE CLASSIFICATIONS SMALL, MEDIUM, AND FOR THE EFFECT SIZE DEPENDING THE WHICH TEST IS PERFORMED.

compute the sample size needed can also be used to determine the required sample size in a study to obtain a certain amount of power. Usually a power above 80% is desired for a statistical test [Mickey, 2004, page 56].

With the non-centrality estimated we can now compute the power:

$$\text{Power}_{\chi^2} = 1 - \int_0^{Q_{1-\alpha}} \chi^2_{\text{nc}}\left(x; \lambda_{\chi^2}; \text{df} = (r-1)(c-1)\right) dx \tag{4.54}$$

$$\text{Power}_{\text{ANOVA}} = 1 - \int_0^{Q_{1-\alpha}} F_{\text{nc}}\left(x; \lambda_F; \text{df}_1 = (k-1); \text{df}_2 = (N-k)\right) dx \tag{4.55}$$

$$\text{Power}_{t\text{-test}} = 1 - \int_0^{Q_{1-\alpha}} t_{\text{nc}}\left(x; \lambda_t; \text{df} = \nu\right) dx, \tag{4.56}$$

where $Q_{1-\alpha}$ is the $1-\alpha$ quantile of the corresponding central distribution. In words: The power is equal to the probability mass of the non-central distribution to the right of the $1-\alpha$ quantile of the corresponding central distribution; as such the power is also greatly affected by the level of significance α set for the test.

4.6 Continuity correction

In the presented independence tests we, in every case, approximated a discrete distribution with a continuous. In both the G- and Pearson's χ^2-test the test statistics were originally distributed according to the multinomial distribution, and in McNemar's test according to a binomial. When the cells in the relevant contingency table contain sufficiently large counts we approximate by using the χ^2-distribution. However, the English statistician Frank Yates presented in 1934 an article in which he noted that in doing so would result in a misleading probability for the null hypothesis. To compensate for this he proposed a correction term to the test statistics [Hitchcock, 2009, page 1]. Yates, however, only proposed this for the Pearson χ^2-tests, i.e both independence and goodness of fit test. The important question remains whether to use his correction or not.

[Hitchcock, 2009, page 10] notes that many introductory and advanced textbooks presents Pearson's χ^2-test without the correction which may serve as an indicator that the convention today is not to use it; however, he also notes that this trend simply may be to ease the introduction of the test. This does, however, not explain why many advanced textbooks also omit the correction. In addition [Blumann, 2012, page 613] directly states: *"Since the chi-square test is already conservative[2], most statisticians agree that the Yates correction is not necessary"*.

In regards to the use of Yates continuity correction in McNemar's test a research article [Fagerland, 2013]

[2] By conservative he means that it does not yield an overestimated probability for the null hypothesis.

concludes that the use of the correction term is not advised; in addition it also found that it is the most powerful [Fagerland, 2013, page 8].

Since no clear cut guideline exists in regards to use the continuity correction, or not, it has simply been decided not to be implemented it in the test statistics even though many other statisticians advocates its use, e.g [Shalabh, 2009, page 39] and [Fleiss, 2003, page 58]. However, as noted, many advocates no use of the term.

Chapter 5
Data analysis and interpretation

The theory presented in the previous chapter is now to be put to use as we in this chapter analyze the data obtained. The raw data which we have are the responses from the three tests the students of Mek1 have been given during the course: They are (1) the Lab Test, (2) the DHD, and lastly (3) the FCI.

The first section presents the design of the experiment and how it has changed as time progressed. Following we present the analysis of the data in which we provide preliminary interpretation of the respective results.

5.1 Design of experiment

A month prior to the school summer holiday in 2013 (July to September) the outline of this thesis was formulated; firstly it was to create teaching materials, in form of the book *"Grundlæggende Matlab og dataanalyse"*, shown in appendix A, and screencasts, to aid the first year students in learning how to use Matlab and basic data analysis. Both the book and screencasts were, after revision by the responsible lecturers for Mek1 of course, then supposed to be suggested as the main source on how the students could learn the curriculum in the use of Matlab and basic data analysis which both are required subjects in Mek1. On this basis it was to be investigated whether

1. the students' use of either the book or screencasts is correlated with the number of total correct and wrong answers (NTC) in relevant tests taken during Mek1;

2. a preference of a particular media is correlated with the NTC in the same tests;

3. a statistical significant difference, between the means of the students' correct answers of students preferring a particular media can be detected.

The media, on which we focus, are the book, screencasts, and the internet; but also books, screencasts/videos, and notes on the internet in general.

This outline was based on the assumption that no specific instruction in either the use of Matlab or basic data analysis was to be provided. However, in the 11'th hour (late August) KU decided to hire a Matlab instructor to provide lectures on the use of Matlab in Mek1. The decision to actually hire a Matlab instructor supports the idea behind this thesis; that action is needed in regards to the students learning Matlab. In addition Børge Svane Nielsen would, give lectures on basic data analysis. As a consequence both of these new initiatives, which would be part of the mandatory curriculum of Mek1, have obscured the original design: Both the book and screencasts would not be suggested as main sources but as secondary literature, thereby not including them actively in the teaching of Mek1. Therefore, the investigation this thesis endeavours to investigate would most likely be more difficult. Ample time in order to adjust the experiment to accommodate these new initiatives was, however, not present.

5.2 Data analysis

In this section the responses of the tests are analyzed and the results presented. However, we first need to present some preliminary considerations and definitions in order to form a meaningful basis.

As mentioned in chapter 3 the students, were in connection to Mek1, given the Lab Test, DHD, and FCI. In every of these tests the students were to supply their date of birth and full name. Hereby we can identify the individual student's preference of media, which they supplied in the Lab Test, between the different tests. The responses of the tests are the raw data. In the course of processing the raw data various choices had to be made; as such a set of definitions has been set fourth to which we turn if an answer, to one or more questions, is (1) not readable; (2) missing, i.e the student did not provide an answer; and (3) ambiguous, i.e more than one answer provided. To accommodate for these cases we define an unreadable, missing, or ambiguous answer as if no answer has been provided. For questions which are not of a right and wrong nature, i.e questions 1 through 4 of the Lab Test, which are on media preference, we define a response meeting definitions (1) to (3) as an "informed answer". When questions *are* of a right or wrong nature we interpret "no answer" or "not understood" both to constitute a wrong answer; as such this type of question can only be considered correctly answered if an answer has been provided *and* it is correct.

The raw data comprises of sheets of paper: This means that the students have written their responses to the individual questions of each test on separate sheets. These have then been manually entered into spreadsheets such they could be exported to a csv-file which then is readable by the analysis software; we have here used ROOT. The point is that the manual entering of the data into spreadsheets has introduced human error which is not quantifiable. Therefore we have no way of knowing the error when we claim to have counted, say, 30 students of 100. It has therefore been decided that if/when we state 30 of 100 students in percentage that we state it to two decimal places, i.e 33.33%.

In section 4.3 of the previous chapter it was stated that the raw data are de facto independent and random samples. The independence of the raw data is due to the simple fact that when the students were given the various tests they were not allowed to help or speak to each other. Therefore the students took the tests alone making the individual student's responses independent of the remaining. The de facto randomness is due to the following argument: The students, who fulfilled the criteria for admission into the physics education at KU from the various Danish secondary schools, are assumed not to be selected specifically from a certain region of Copenhagen or surrounding cities, or the remaining part of Denmark. Logic would therefore suggest that the student body of first year physics students at KU must hail from a random mix of both regions and schools making the raw data random. Note that this is an argument based on common sense and logic and not on any hard evidence; nevertheless the argument is considered sound and no further investigation into the randomness of the data will be conducted.

When representing results of performed - unpaired - independence tests we will state the results of both the χ^2- and G-test: The G-test, however, takes precedence, especially in cases where disagreement between the two exists; as such the χ^2-independence test is presented merely as support and for thoroughness. In addition we will, for every statistical test performed, state its power with the exception of χ^2-goodness of fit - and Levene's test. However, the power will not be included in any discussion or interpretation until chapter 6.

In order to perform an ANOVA and Student's paired t-test one condition is that the samples have to be normally distributed, as noted in chapter 4. To ascertain whether this condition is met is to perform a χ^2-goodness of fit test for binned data. This test, however, has some caveats; (1) it approximates counting statistics as normally distributed and not Poissonian, which implies that the errors are normally distributed, hereby making it only applicable when 80% of the bins contains at least five hits [Taylor, 1997, page 266-267]; (2) when data is binned information is lost as the test does not take bins containing zero hits into account. Despite these caveats this test is preferred as we in the end can determine a probability for the sample distribution to be gaussian. However, due to the loss of information through binning we can in some cases get a very poor fit even though the sample distribution may very well be normally distributed. To compensate for this it is possible to conduct a fit using the principles of maximum likelihood, however, this does not yield a goodness of fit measure; this implies that even though the fit is visually consistent it may very well be poor. In the end it has been decided that when a χ^2-goodness of fit is not possible due to lack of hits in the bins, or when the fit is poor due to loss of information through binning, we rely on the error of the sample mean, i.e standard deviation of the mean (SDOM), given by the standard deviation of the sample divided by the square root of the number of observations.

The level of significance α - the cutoff value - for every performed statistical test has been chosen to be $\alpha = 0.05 = 5.00\%$ which is the most conventional. We have no arguments or evidence that α should set to a specific value; as such we simply set it to the conventional value.

5.2.1 Lab Test

The Lab Test was taken by 137 students in total. In table 5.1 the response rate and the rate of students who understood the respective questions, which hence fourth will be denoted by *understood rate*, for the Lab Test. The lowest observed response rate is 92.7%, which corresponds to 127 students, i.e a maximum of 10 students have not provided an answer, which is considered to be extremely good. Furthermore the lowest understood rate is 96.4% implying that we can be fairly confident in that the students knew what was being asked in the respective questions.

Question	1	2	3	4	5	6	7	8	9	10
Response rate [%]	98.54	99.27	92.70	92.70	99.27	99.27	97.81	99.27	99.27	97.08
Understood [%]	100.00	100.00	99.27	99.27	100.00	99.27	98.54	99.27	99.27	97.08
Question	11	12	13	14	15	16	17	18	19	20
Response rate [%]	98.54	98.54	98.54	97.81	98.54	98.54	97.08	99.27	97.81	97.81
Understood [%]	97.81	100.00	100.00	100.00	100.00	100.00	96.35	100.00	99.27	99.27
Question	21	22	23							
Response rate [%]	97.81	98.54	97.81							
Understood [%]	97.81	100.00	97.81							

TABLE 5.1 • A TABLE OF THE RESPONSE RATE AND THE STUDENTS WHO DID NOT UNDERSTAND THE RESPECTIVE QUESTIONS IN PERCENT IN THE LAB TEST. THE LOWEST RESPONSE RATE AND THE PERCENTAGE OF STUDENTS NOT UNDERSTANDING A QUESTION ARE 92.70% AND 96.35% RESPECTIVELY WHICH IS CONSIDERED EXTREMELY GOOD. THE TOTAL NUMBER OF STUDENTS WHO WERE GIVEN THE TEST IS $N = 137$.

We initially noted that of the 137 students who were given the Lab Test one has zero correct answers in total; it is not that he answered every question wrongly but simply chose not to provide any answers to any questions! Therefore, the data for this student will hence fourth be omitted from any further calculations to draw more precise conclusions, as the only explanation must be that not answering any questions was his intention from the beginning of taking the test. The reason for this intention is, however, not known.

With this in mind we now turn our attention to how the remaining 136 students did on the Lab Test; for each student we count the number of correct answers for questions 5 to 23. The distribution of correctly answered questions is shown in figure 5.1. The mean is 14.92 ± 0.17, which is in the high end; as such the students did overall quite well on the Lab Test. However, this does not reveal anything about the individual questions being overall difficult. Therefore, we inspect how the students did overall questionwise; in figure 5.2 the percentage of correct answers questionwise is illustrated. We quickly note that the questions within the Matlab group (see table 3.1 on page 15) are answered correctly by over 70% of the students, with the

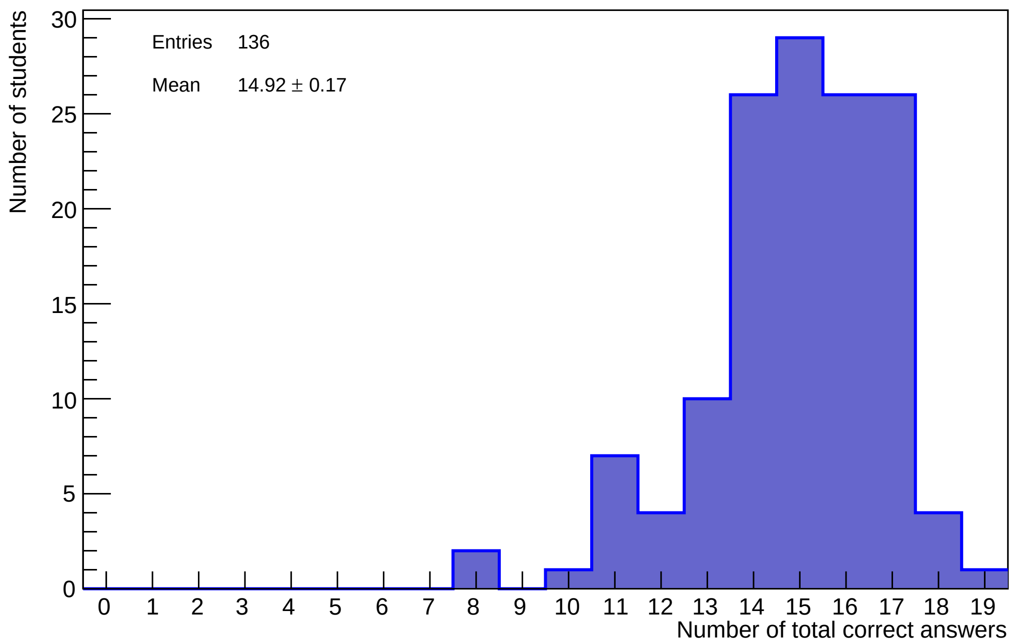


FIGURE 5.1 • THE DISTRIBUTION OF TOTAL CORRECT ANSWERS IN THE LAB TEST. THIS SHOWS THAT THE STUDENTS DID QUITE WELL OVERALL ON THE LAB TEST.

	MATLAB GROUP			DATA ANALYSIS GROUP		
	Question 5	**Question 8**	**Question 10**	**Question 17**	**Question 19**	**Question 21**
Possibility 1	**93**	**91**	27	**72**	**21**	34
Possibility 2	39	9	**92**	46	33	**85**
Possibility 3	3	35	10	10	79	12
Not responded	1	0	3	3	2	2
Not understood	0	1	4	5	1	3
Total	136	136	136	136	136	136

TABLE 5.2 • THE NUMERICAL DISTRIBUTION OF QUESTIONS 5,8, AND 10, WHICH FALL WITHIN THE MATLAB GROUP, AND QUESTIONS 17,19, AND 21, WHICH FALL WITHIN THE DATA ANALYSIS GROUP (SEE TABLE 3.1 ON PAGE 15). THE BOLD NUMBERS INDICATES THE CORRECT POSSIBILITY.

exception of questions 5,8, and 10. Furthermore, questions 17,19, and 21, which fall in the group data analysis, are of interest; especially question 19. How the answers are distributed according to the three possible answers are shown in table 5.2. For the formulation of the respective questions see appendix B.

In question five we see that 39 of the students chose the second possibility: One likely explanation for this may lie in the fact that the students, parallel to Mek1, are following the course MatIntro; it is an introductory mathematics course wherein the used CAS-software is Maple by Maplesoft, and the format for defining a variable called MyFirstVar is exactly possibility 2. Therefore as the students had to learn two different CAS-softwares at the same time, may have given rise to some confusion resulting in roughly a third answering possibility 2. Another explanation may be that those who actually did not know how to define variables in Matlab, but did in Maple, simply scanned the possible answers, and submitted possibility 2 as it is the only possibility which then would look familiar to them.

We observe in question 8 that 35 of the students have submitted possibility 3 as the correct answer.

Here there can be no confusion due to Maple, as it does not have a similar command. In this case a possible interpretation is that if the students did not know the correct answer for clearing Matlab's Command Window, which is *clc*, they may simply have submitted the possibility which seemed most logical. Here the most logical answer, when one does not know the correct answer, would be possibility 3 which is the command *clear*.

In regards to question 10 a more refined interpretation is difficult to provide; roughly 20% and 7% have submitted possibility 1 and 2, respectively, as the correct answer which both are incorrect. Possibility 1 and 2 do not resemble the syntax of Maple meaning that the confusion Maple could give rise to is most likely not in play here. However, the syntax in both possibilities does, but not 100%, look like C++ or PHP, but these programming languages have not been introduced at the time the Lab Test was given. C++ will most likely be introduced in their third year of the bachelor or the first year of the master in connection to using ROOT. PHP on the other hand will not at any point be introduced in either the bachelor or master.

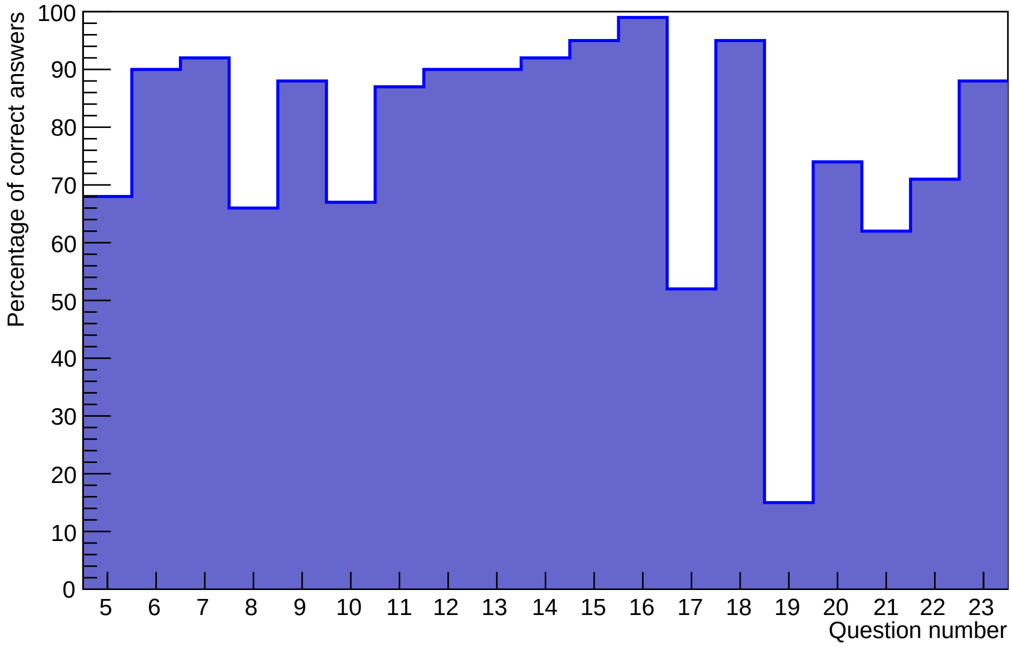


FIGURE 5.2 • PERCENTAGE OF CORRECT ANSWERS QUESTIONWISE OF THE LAB TEST. QUESTIONS 5, 8, 10, 17, 19, AND 21 ARE ALL ANSWERED CORRECTLY BY UNDER 70% OF THE STUDENTS.

Now we look at the relevant questions, which fall in the data analysis group; we again refer to table 5.2 for the numerical distribution of the three possibilities the students submitted as the correct answer. In question 17 roughly half of the students answered correctly. This question is particularly interesting as its answer is directly provided in chapter 3 of the book and also in lectures on data analysis held by Børge Svane Nielsen during Mek1 judging from the slides he uploaded to Absalon. It asked the student: *"Assume that a set of measurements of a certain quantity are normally distributed. Within how many standard deviations will* 68.2% *of the measurements lie?"*. This could be construed as a theoretical statistics question, however, still very applicable in data analysis; as such it is positive that roughly half of the students answered such a question correctly when considering that they, from their secondary education, only have limited knowledge of the subject. Especially since the formulation of the question does not specify relative to what the 68.2% of the measurements will lie within: Implicitly it is relative to the mean, which half of the students seemed to deduce. A further discussion on how limited the Danish secondary school students' knowledge is provided in chapter 7. In which we look at the possibilities of introducing data analysis in said educations.

In question 19 only 21 students submitted the correct answer. It asked them what the uncertainty of the mean of the set of measurements $9.7\,\mathrm{m/s^2}, 9.9\,\mathrm{m/s^2}, 10.1\,\mathrm{m/s^2}, 10.1\,\mathrm{m/s^2}$, and $10.3\,\mathrm{m/s^2}$, given in question 18, was; as such they were supposed to apply the formula for the standard deviation of the mean (SDOM), i.e $\delta_{\overline{x}} = \sigma_x/\sqrt{N}$ which in this case numerical is equal to $0.1\,\mathrm{m/s^2}$ when stated with one decimal. Since question 15 asked them what the best estimate of the uncertainty, in each of N measurements of a quantity, is it would be interesting to investigate if we can observe a correlation between the NTC of question 15 and 19. Logically, we would expect some correlation as the computation of SDOM requires the knowledge of the spread of a data set. Preliminary to performing an independence test we find that of the constructed contingency table, shown in table 5.3, $E_{21} = 0.93$; as a consequence one of the requirements for the test is not met. However, as $E_{21} \approx 1$ we still perform the test but keep in mind that its validity may be in question. Performing the test yields a probability of independence of 93.30%. Therefore if we accept the approximation $E_{21} \approx 1$ then we observe no correlation between the responses of question 15 and 19. Surprisingly, this implies in this case that answering question 19 correct or wrong is not dependent on the answer provided in question 15 and vice versa. From a standpoint of common sense would, however, dismiss this notion as utter nonsense as the formula of the spread of a data set is directly a part of that of the SDOM: This is, however, not what we have observed.

In the search for an explanation of the paradoxical result we look at the remaining two possibilities: Possibility two was $0.2\,\mathrm{m/s^2}$ and the last $0.3\,\mathrm{m/s^2}$. If the students were under the impression that the spread was to be calculated, in order to answer correctly, they should have responded $0.2\,\mathrm{m/s^2}$, i.e possibility two; however, only 33 students did this. Most peculiar is that the third possibility dominates in terms of number of responses as 79 chose to submit it that the correct answer. We can simply offer no explanation to this response pattern: If the students forgot to take the square root, i.e they computed $\delta_{\overline{x}}$ as σ_x/N, would yield $0.05\,\mathrm{m/s^2}$. In short: The reason as to the high number of students submitted possibility three as the correct answer is an enigma. But given the data we simply observe that the students did learn that the spread is the best estimate of the uncertainty in a single measurement, however, they have not learned how to compute SDOM.

We investigate accordingly in regards to the concept of the mean, i.e have the students actually learned

CONTINGENCY TABLE

	Q. 19 correct	Q. 19 wrong	
Q. 15 correct	20	110	130
Q. 15 wrong	1	5	6
	21	115	136

RESULTS

	G-test	χ^2-test
Test statistic	0.007078	0.007220
Effect size	0.007214	0.007286
λ_{χ^2}	0.007078	0.007220
Probability	93.30%	93.29%
Power	5.08%	5.08%

TABLE 5.3 • CONSTRUCTED CONTINGENCY TABLE AND TEST RESULTS WHEN TESTING FOR INDEPENDENCE BETWEEN THE RESPONSES OF QUESTION 15 AND 19. NOTE THAT SINCE $E_{21} = 0.93 < 1$ WE VIOLATE THE VALIDITY OF THE APPROXIMATION. HOWEVER, BOTH TESTS YIELDS ROUGHLY THE SAME PROBABILITY OF INDEPENDENCE WHEN ACCEPTING THE SPPROXIMATION $E_{21} \approx 1$.

both its meaning and how to compute it. This is done investigating for correlation between the NTC of question 14 and 18. However, we also here run into the problem of one of the expected frequencies of the constructed contingency table (see table 5.4) not being greater than or equal to 1 as we find $E_{22} = 0.44$. Here E_{22} is too far from 1 that we would perform an independence test, but from the constructed contingency table we see that 122 students answered correctly in both question 14 and 18; as such 89.71% of the students knew both that the mean is the best estimate of the true mean, and how to compute it. Only 1.47% did not. Alone from the very high/low percentage of students who have/have not learned both the meaning and computation of the mean we, in case of these students, observe no notable problems with the concept mean.

Question 21 was essentially on the use of the error propagation formula as it was to be applied in order

CONTINGENCY TABLE	Q. 18 correct	Q. 18 wrong	
Q. 14 correct	122	4	160
Q. 14 wrong	8	2	10
	130	6	136

TABLE 5.4 • CONSTRUCTED CONTINGENCY TABLE IN THE PURSUIT OF TESTING FOR INDEPENDENCE BETWEEN THE RESPONSES OF QUESTION 14 AND 18. HOWEVER, WE FIND THAT $E_{22} = 0.441176 < 1$ WE VIOLATE THE VALIDITY OF THE APPROXIMATION; AS SUCH WE DO NOT PERFORM AN INDEPENDENCE TEST AS WE WOULD NOT BE ABLE TRUST ITS RESULTS.

to determine the correct answer. Recall that the error propagation formula for independent measurements is given by

$$\delta f(z_1, z_2, \ldots, z_M) = \sqrt{\sum_{i=1}^{M} \left(\frac{\partial f}{\partial z_i} \cdot \delta z_i \right)^2}, \tag{5.1}$$

where δz_i denotes the uncertainty in the z_i'th variable. In terms of difficulty the use of the error propagation formula contra the computation of SDOM of a small data set, i.e question 19, is greater. Therefore it is odd that a larger percentage of students answered correctly here than in 19. Thus we investigate whether there exists a dependence between question 19 and 21, i.e if we statistically can infer a tendency to answer question 21 correctly given that 19 was answered correctly. Performing an independence test (see table 5.5) reveals that this is not the case. One explanation could therefore be that the students simply guessed more favourable in question 21 than in 19 or actually have learned to apply the formula.

Validity and interpretation of media responses

Recall that questions 1 through 4, the about media group in table 3.1, in the Lab Test were on media preference. Before we investigate the various correlations between test scores and media preferences we need to validate the responses. Validation in this context means if the responses make sense. For instance, say, x students claimed to never have heard of the book and y that they found it the most educational, if it turns out that y was greater than x it would simply not make sense; from a logical standpoint there could never be more students who found it most educational, which would require that they have used it to make such a statement, than those who claimed to never have used it! Actually, even if y was in the neighbourhood of x would also be considered as suspicious.

This section aims to investigate such possible illogical response patterns. If the responses were not investigated the conclusions inferred from them could be utter useless. The tool here applied are common sense and elements of the mathematical discipline set theory. The notation used is that which is standard in regards to this field: If A and B are sets then $A \cap B$ denotes the intersection between A and B; and $A \cup B$ the union. Figure 5.3 shows Venn diagrams illustrating the notation and concepts. In addition the symbol

CONTINGENCY TABLE

	Q. 21 correct	Q. 21 wrong	
Q. 19 correct	11	10	21
Q. 19 wrong	74	41	115
	85	51	136

RESULTS

	G-test	χ^2-test
Test statistic	1.060810	1.084992
Effect size	0.088318	0.089319
λ_{χ^2}	1.060810	1.084992
Probability	30.30%	29.76%
Power	17.78%	18.06%

TABLE 5.5 • CONTINGENCY TABLE AND TEST RESULTS FOR TESTING DEPENDENCE BETWEEN QUESTION 19 AND 21 IN THE LAB TEST AND THE TEST RESULTS. BOTH TEST STATISTICS, I.E G AND X^2 IMPLIES THAT WE FAIL TO REJECT THE NULL HYPOTHESIS OF INDEPENDENCE.

$\emptyset$ denotes the empty set.

The questions in the about media group were as follows:

Q. 1 *Have you at any point used the book "Grundlæggende Matlab og dataanalyse", which can be found on Absalon?* Possible answers were *yes*, *no* and *never heard of it*.

Q. 2 *Have you seen any of the screencasts, which also can be found on Absalon?* Possible answers were *yes*, *no* and *never heard of them*.

Q. 3 *Which of the following three media did you find most educational?* Possible answers were *the book from Q. 1*, *the screencasts from Q. 2*, and *internet (youtube.com, computerfysik.dk,.....)*.

Q. 4 *Which of the following three media do you feel you benefit most from when learning something new?* Possible answers were *books*, *screencasts/videos*, and *Notes on the internet*.

How the students answered these questions is shown in table 5.6. It is clear that illogical response patterns are only to be found in the responses of question 1,2, and 3 since (1) they are highly mutually dependent, and (2) it is quite reasonable that, say, a student answered books in general are his preferred media when

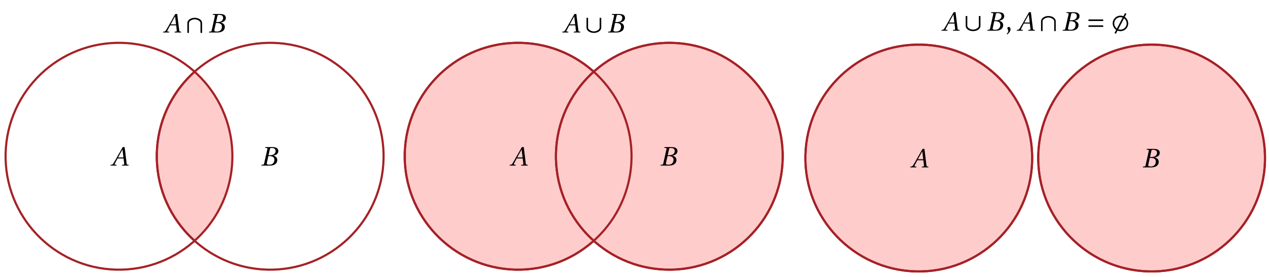


FIGURE 5.3 • VENN DIAGRAMS ILLUSTRATING THE STANDARD NOTATION USED FOR SETS. AS SHOWN THE SYMBOL $\cap$ IS THE INTERSECTION, $\cup$ THE UNION, AND $\emptyset$ REPRESENTS THE EMPTY SET. THE DIAGRAM TO THE FAR RIGHT ILLUSTRATES THE CONCEPT OF DISJOINT SETS, I.E $A \cap B = \emptyset$.

	Question 1	Question 2	Question 3	Question 4
Possibility 1	66(B)	39(S)	45(BO)	66
Possibility 2	39($\overline{B}$)	74($\overline{S}$)	14(SC)	34
Possibility 3	30($\dot{B}$)	23($\dot{S}$)	67(IN)	26
Informed answers	135	136	126	126

TABLE 5.6 • THE DISTRIBUTION OF THE STUDENTS' ANSWERS OF THE THREE POSSIBILITIES TO QUESTIONS 1 - 4 IN THE LAB TEST, WHICH ARE THOSE IN THE ABOUT MEDIA GROUP. FROM THIS DATA RELEVANT SETS ARE CONSTRUCTED TO CHECK FOR INCONSISTENCIES. THE ENTRIES ARE NUMBER OF STUDENTS AND THAT IN THE PARENTHESES ARE NOTATION FOR USE IN TABLE 5.7.

learning a new topic even though he has either used, not used, or never heard of the book.

Following this line of thought we in total construct ten sets from the media responses which are then divided into two groups; *consistency check* and *illogical responses*, respectively. This is shown in table 5.7. Sets within the group consistency check are used to check if they are subsets of the relevant sets in table 5.6; illogical responses consists of, as the name suggests, sets used to check for or contain illogical responses.

Cross referencing the sets, within the consistency check group, with those in table 5.6 we clearly observe no inconsistencies for the sets $B \cap BO$ and $S \cap SC$ as they are subsets of B and S, respectively. Hereby implying that the number of students who used the book and screencasts, respectively, and also found it the most educational, are not greater than each of the totals who used the two media. Also we observe that the number of students, who either have not used or never heard of the book and answered internet as the most educational media, is smaller than the total of 67, i.e $\left(\overline{B} \cup \dot{B}\right) \cap IN \subset IN$ which is consistent. The same is true in the case of the screencasts, i.e $\left(\overline{S} \cup \dot{S}\right) \cap IN \subset IN$. We observe, however, one inconsistency; due to the mathematical properties of set operations we have that

$$\left(\overline{B} \cup \dot{B}\right) \cap IN \cup (B \cap IN) = \left(B \cup \overline{B} \cup \dot{B}\right) \cap IN \tag{5.2}$$

which simply is the set of students who found the internet most educational which contains 67 elements; however, summing the respective number of elements does not yield 67 as it should but $13 + 53 = 66$. The reason for this particular inconsistency is because the one student who did not provide an answer to question 1 have answered the internet as most educational.

Logically, there are only four cases which could constitute an illogical response: If a student has not used/never heard of the book/screencasts, and he in question 3 answered the book/screencasts. Therefore the sets of interest are $\overline{B} \cap BO, \overline{S} \cap SC, \dot{B} \cap BO$, and $\dot{S} \cap SC$. It turns out that $\dot{B} \cap BO = \dot{S} \cap SC = \emptyset$ which implies that those who never have heard of the book or screencasts have not voted either of these media as the most educational which is consistent. Following the same logic the relation $\overline{B} \cap BO = \overline{S} \cap SC = \emptyset$ should also be true; this is, however, only true for $\overline{S} \cap SC$. The set $\overline{B} \cap BO$ contains two elements which implies that two students found the book the most educational media even though they have never used it. This is illogical because if the book has never been used they have no foundation to deem it most educational.

Somewhat illogical is also those who claimed to have used the book and at the same time never have heard of the screencasts, i.e the $B \cap \dot{S}$, and vice versa, i.e $S \cap \dot{B}$. This is due to the fact that both the book and screencasts were published on Absalon which in turn suggests that if a student has used the book he in all likelihood would have been aware of the existence of the screencasts. However, a simple count reveals that there are 7 elements in the set $B \cap \dot{S}$ and 8 in $S \cap \dot{B}$. These sets are not included in the group illogical responses as there do exist a minute chance that the students used the book but not knew of the screencasts, and vice versa. Therefore we only note that this pattern seems quite odd but do not consider it further.

We have found that there are some inconsistencies due to missing data in form of students not answering questions and illogical responses. To properly compensate for missing data would require to apply theory from the field of statistics known as *missing data*; this field is, however, out of the scope of this thesis. A second solution could simply be to let logic dictate the value of the missing responses where possible; however, in doing so we would contaminate the data pool which is not desirable. The solution we will apply, when testing for correlation between test scores and media preference, is simply to disregard the students who failed to provide an answer to their media preference. This will in turn result in loss of valuable data; but as previously noted the response and understood rate are both close to 100% for the questions in the about media group which was shown in table 5.1; as such the exclusion of students not providing a media preference will most likely not impact the inferred conclusions significantly; as such this is the main argument for the implementation of this strategy.

Therefore, despite the observed inconsistencies, which actually are small relative to the total data set, we overall conclude that the responses of media preferences are consistent and find no further alteration of the data warranted.

Estimating the reception of the book and screencasts

We mentioned in section 2.3 that it was not possible to rally any students willing to be part of a focus group; the purpose was to shed some light as to how the students received the book and screencasts and how they could be improved for future students. Therefore the only alternative we have to provide some insight into this is to look at the quantitative data, i.e their responses shown in shown in table 5.6.

It is positive that 48.89% of the students have at some point during Mek1 used the book when taking into account that the book was classified as secondary literature. In addition, of those who used the book, 65.15% claimed it to be the most educational media; this may be interpreted as these found its contents, or elements hereof, usable in connection to Mek1. On the other hand it is a bit mysterious that 22.22% claimed to never have heard of it as it has been advertised on Absalon on, at least, three separate occasions; logic would, therefore, dictate that they would have heard of the book as they, at the very least, once a week have to check Absalon for homework and which experiments to conduct in the laboratory exercises. The same argument can be made as to the awareness of the screencasts in question 2 as 16.91% here claimed not to have heard of them. In addition to both the book and screencast have been advertised on Absalon, Ian Bearden also held two lectures wherein he asked the attending students if they have heard of the book or the screencasts; the results where consistent with what we here have observed [Bearden, 2013b]. Therefore the effort in advertising the existence of the created teaching materials was not inadequate. The exact reason why some students still not knew of the teaching materials existence at the end of Mek1 is simply not known.

Somewhat surprising is the fact that 54.41% have not used the screencasts as this is quite a high rate.

Consistency check		Illogical responses	
Set	**#elements**	**Set**	**#elements**
$B \cap BO$	43	$\overline{B} \cap BO$	2
$S \cap SC$	10	$\dot{B} \cap BO$	0
$B \cap IN$	13	$\overline{S} \cap SC$	0
$S \cap IN$	13	$\dot{S} \cap SC$	0
$\left(\overline{B} \cup \dot{B}\right) \cap IN$	53		
$\left(\overline{S} \cup \dot{S}\right) \cap IN$	54		

TABLE 5.7 • THIS TABLE SHOWS THE RELEVANT SETS TO CHECK FOR INCONSISTENCIES IN THE MEDIA RESPONSES OF THE STUDENTS. ONE SMALL INCONSISTENCY ARE FOUND IN THE GROUP *consistency check*; SUMMING THE ELEMENTS OF THE SETS $B \cap BO$ AND $\left(\overline{B} \cup \dot{B}\right)$ DOES NOT YIELD 67 AS IT SHOULD (SEE TABLE 5.6). IN THE *illogical responses* WE DO, HOWEVER, OBSERVE TWO SMALL INCONSISTENCIES AS TWO STUDENTS HAVE ANSWERED THAT THEY LIKED THE BOOK BUT HAVE NOT USED IT.

One explanation for this may be connected to the choice of software in which the screencasts have been made: Recall that the software used was Adobe Captivate and the format of the screencasts was Flash, i.e the file extension "swf". When a student wished to see a screencast he had to login to Absalon and click on the desired; when clicked the screencast had to be fully loaded in order to begin playback. It is this loading time, however short, that may be one of the main reasons for the students' lack of interest in them. In addition, their duration may also contribute to this; only three out of eleven screencasts have a duration under nine minutes and the remaining ranges from 17 to 29.5 minutes. Such long durations may discourage the students from using them as up to 30 minutes may not be a period of time they could or wished to spend on a video; they perhaps therefore regarded them as lectures which was not intended. The solution could be to instead create screencasts of maximum three to five minutes in length. We note that it may also simply be that this particular class of students is not fond of screencasts as a media, or they did not find the content satisfactory. This is somewhat supported by the fact that only 28.68% claimed to have used them, and only 26.98% found the media in general most educational when learning a new subject, which is quite low relative to the book.

If we construct the set $B \cap S$ we find that it contains 24 elements, i.e 24 students of the 136 have used both the book and screencasts. Of these only 5 voted the internet as the most educational media, i.e the set $(B \cap S) \cap IN$ only contains 5 elements. This is quite positive considering the teaching materials only were suggested as secondary literature, and also that after their respective use they were by many also voted as most educational. From this observation the teaching materials can be interpreted as well received by the students: The book was, however, the most well received teaching material as 65.15% of those who used it found it most education. The questions concerning which aspects of the teaching materials could be improved can not on the basis of the data be answered.

Judging from the responses in media preference the book is without question the teaching material the students have used the most, and at the same time found most educational, when comparing to the screencasts. This high rate may be due to 52.38% of this particular class of students in general preferred books when learning a new topic and only 26.98% claimed screencasts/videos. Note, however, that 53.17% found the internet most educational and of those 79.10% have not used or never heard of the book and 80.60% of the screencasts; as such the internet was overall the most popular media. This trend would probably have been somewhat different if the teaching materials were classified as main literature.

Testing for best media

We now turn our attention to use the media responses to determine whether we can observe if there are some correlation between media preference and test scores. See table 5.6 for the numerical distribution.

Q. 1 *Have you at any point used the book "Grundlæggende Matlab og dataanalyse", which can be found on Absalon?* Possible answers were *yes* ($N = 66$), *no* ($N = 39$) and *never heard of it* ($N = 30$).

First we investigate possible dependence between the NTC and whether the students used the book or not. In total 135 students provided an informed answer to question 1. For these 135 we count the NTC for every student, who has used the book; the same procedure is applied for those who did not and never heard of it. Performing an independence test will investigate whether there exists some correlation between having used, not used, or never heard of the book, and the NTC of the Lab Test. Referring to table 5.8, which contains the resulting contingency table and test results, it is seen that we fail to reject the null hypothesis of independence; thus there is no evidence that a correlation exists.

During Mek1 the students have had lectures in both the use of Matlab and basic data analysis. With this in mind some correlation between having used and not used the book and the NTC may exist when only considering the questions within the Matlab and data analysis group of the Lab Test; however, when performing an independence test considering only these questions this turns out, overwhelmingly, not to be the case as an independence test results in 94.90% chance of independence when only considering the questions within the Matlab group; same conclusion is drawn when only considering the questions within the data analysis group as it here results in 62.31%. See tables D.1 and D.2, respectively, in appendix D for further details.

Q. 2 *Have you seen any of the screencasts, which also can be found on Absalon?* Possible answers were *yes* ($N = 39$), *no* ($N = 74$) and *never heard of it* ($N = 23$).

All of the 136 students provided an informed answer to question 2. When testing for independence between the NTC, and whether the students have seen any of the screencasts, or not, we again fail to reject the null hypothesis of independence as the test results in 50.52%. The test results are shown in table 5.9. This is somewhat expected as 54.41% of the students claimed to not have seen any of the screencasts and 16.91% to never have heard of them. If the test is performed considering only the questions within the Matlab and data analysis group, respectively, it amounts to the same conclusion. The results of these tests and contingency tables can be found in appendix D in tables D.3 and D.4, respectively.

At this point we can conclude that no significant correlation is observed, between the students ever having used, not used or never heard of either the book or screencasts during Mek1 and the NTC in the Lab Test. Considering that the original design of this experiment was that the book and the screencasts was to be the main suggested literature, for the students to learn about the use of Matlab and basic data analysis, we are somewhat not surprised of failure to observe correlation: The new initiatives implemented by KU in the school year 2013/14 for the new physics students, however, changed the design, as remarked in section 5.1. This is believed to have played a central role since we do not detect any noticeable correlation between the students having used the book or screencasts and their NTC of the Lab Test. It should, however, be noted that during Mek1 the Matlab instructor kindly reminded the students of the existence of both the book and screencasts but stated clearly that it was not part of the mandatory curriculum.

Q. 3 *Which of the following three media did you find most educational?* Possible answers were *the book*

CONTINGENCY TABLE

	Correct	**Wrong**	
Used the book ($N = 66$)	987	267	1254
Not used the book ($N = 39$)	579	162	741
Never heard of it ($N = 30$)	449	121	570
	2015	550	2565

RESULTS

	G**-test**	χ^2**-test**
Test statistic	0.109696	0.109993
Effect size	0.006540	0.006548
λ_{χ^2}	0.109696	0.109993
Probability	94.66%	94.65%
Power	5.83%	5.83%

TABLE 5.8 • CONTINGENCY TABLE AND TEST RESULTS WHEN TESTING FOR DEPENDENCE BETWEEN THE STUDENTS' NTC IN THE LAB TEST AND WHETHER THEY EVER HAVE USED, NOT USED OR NEVER HEARD OF THE BOOK. BOTH TEST STATISTICS, I.E G AND X^2 IMPLIES THAT WE FAIL TO REJECT THE NULL HYPOTHESIS OF INDEPENDENCE. THE CONTINGENCY TABLE IS BASED ONLY ON THE STUDENTS WHO PROVIDED AN INFORMED ANSWER TO ALL THREE POSSIBILITIES IN QUESTION 1 OF THE LAB TEST OF WHICH THERE ARE 135.

from Q. 1 ($N = 45$), *the screencasts from Q. 2* ($N = 14$), and *internet (youtube.com, computerfysik.dk,.....)* ($N = 67$).

In total 126 of the 136 students provided an informed answer to question 3. We now for each of the three media, the 126 students found to be most educational, count the NTC and perform an independence test which results in 57.70% chance of independence; as such there is not enough evidence to reject the null hypothesis of independence between the three media and the NTC of the Lab Test. The contingency table and numerical details are both shown in table 5.10.

No correlation does, however, not automatically imply no significant difference in the means of total correct answers for each media; by simple visual comparison of the distributions of total correct answers for each of the three media (see figure 5.4) we see they roughly have the same mean. Initially, we would

CONTINGENCY TABLE

	Correct	**Wrong**	
Seen some of the screencasts ($N = 39$)	581	160	741
Not seen some of the screencasts ($N = 74$)	1096	310	1406
Never heard of them ($N = 23$)	352	85	437
	2029	555	2584

RESULTS

	G**-test**	χ^2**-test**
Test statistic	1.365769	1.341800
Effect size	0.022990	0.022788
λ_{χ^2}	1.365769	1.341800
Probability	50.52%	51.12%
Power	16.59%	16.37%

TABLE 5.9 • CONTINGENCY TABLE AND TEST RESULTS WHEN TESTING FOR DEPENDENCE BETWEEN THE STUDENTS' NTC IN THE LAB TEST AND WHETHER THEY HAVE SEEN SOME OF THE SCREENCASTS OR NOT. BOTH TEST STATISTICS IMPLIES THAT WE FAIL TO REJECT THE NULL HYPOTHESIS OF INDEPENDENCE.

CONTINGENCY TABLE

	Correct	**Wrong**	
The book ($N = 45$)	660	195	855
The screencasts ($N = 14$)	208	58	266
The internet ($N = 67$)	1007	266	1273
	1875	519	2394

RESULTS

	G**-test**	χ^2**-test**
Test statistic	1.099710	1.074010
Effect size	0.093423	0.093580
λ_{χ^2}	1.099710	1.074010
Probability	57.70%	57.60%
Power	14.16%	14.20%

TABLE 5.10 • CONTINGENCY TABLE TO TEST FOR DEPENDENCE BETWEEN THE STUDENTS' TOTAL CORRECT AND WRONG ANSWERS IN THE LAB TEST AND WHICH MEDIA THEY FOUND TO BE MOST EDUCATIONAL. THIS TABLE IS CONSTRUCTED USING THE 126 INFORMED ANSWERS PROVIDED.

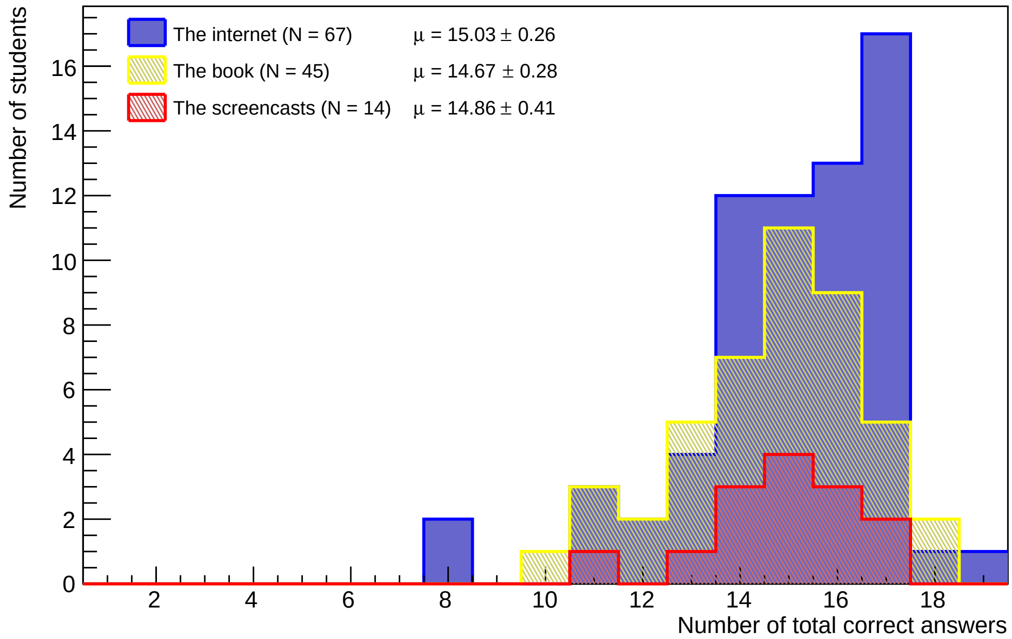


FIGURE 5.4 • THIS FIGURE SHOWS THE DISTRIBUTION OF TOTAL CORRECT ANSWERS FOR ALL THREE MEDIA IN QUESTION 3, I.E THE BOOK, SCREENCAST, AND INTERNET. THE DISTRIBUTIONS ROUGHLY HAVE THE SAME MEAN, WHICH SUPPORTS THAT NONE OF THE THREE MEDIA HAVE HAD AN IMPACT ON THE STUDENTS' ANSWERS OF THE LAB TEST.

therefore expect that they also are statistically equal; to test whether this claim is supported we perform an ANOVA test. However, we first have to establish if we safely can assume that the respective distributions of the three media are gaussian and, if so, whether they are homoscedastic. This is done by performing a χ^2-goodness of fit and Levene's test, respectively. Part of table 5.11 shows the results of the χ^2-goodness of fit and Levene's test; we find that each of the three distributions meets the requirements for normality and homoscedasticity. Note, however, that the distribution of those students who answered internet (blue distribution in figure 5.4) is only 5.69% likely to be drawn from a gaussian. This low probability might suggest that it is in fact not gaussian. To investigate the normality further would require more data which we do not have; as such we can only note that there may be reasons for concern, in regards to the normality of this particular distribution, and continue under the assumption that it meets the requirement.

Under the assumption that all three distributions are gaussian we can perform an ANOVA; the test yields a probability of 53.38% implying that we cannot reject the hypothesis that the means of total correct answers, with respect to the three media (see figure 5.4), are statistically equal. See table 5.11 for further numerical details.

Further support is found by looking at the SDOM's of the respective distributions in figure 5.4: We note that the distribution of those who found the screencasts most educational (red distribution in figure 5.4) has very low hits in the bins, i.e low statistics. Despite the χ^2-goodness of fit test resulted in 83.27% chance of the distribution being gaussian it could very well not be. If so the SDOM of 0.41 calculated is not valid as it assumes the distribution is gaussian; instead we compute the SDOM under the assumption that it is poissonian, i.e SDOM = $\sqrt{\mu/N}$. This results in $\mu_{\text{SC}} = 14.86 \pm 1.03$ and we then see that both μ_{Book} and μ_{Internet} lies within this interval and μ_{SC} in theirs.

Q. 4 *Which of the following three media do you feel you benefit mostly from when learning something new?* Possible answers were *books* ($N = 66$), *screencasts/videos* ($N = 34$), and *notes on the internet* ($N = 26$).

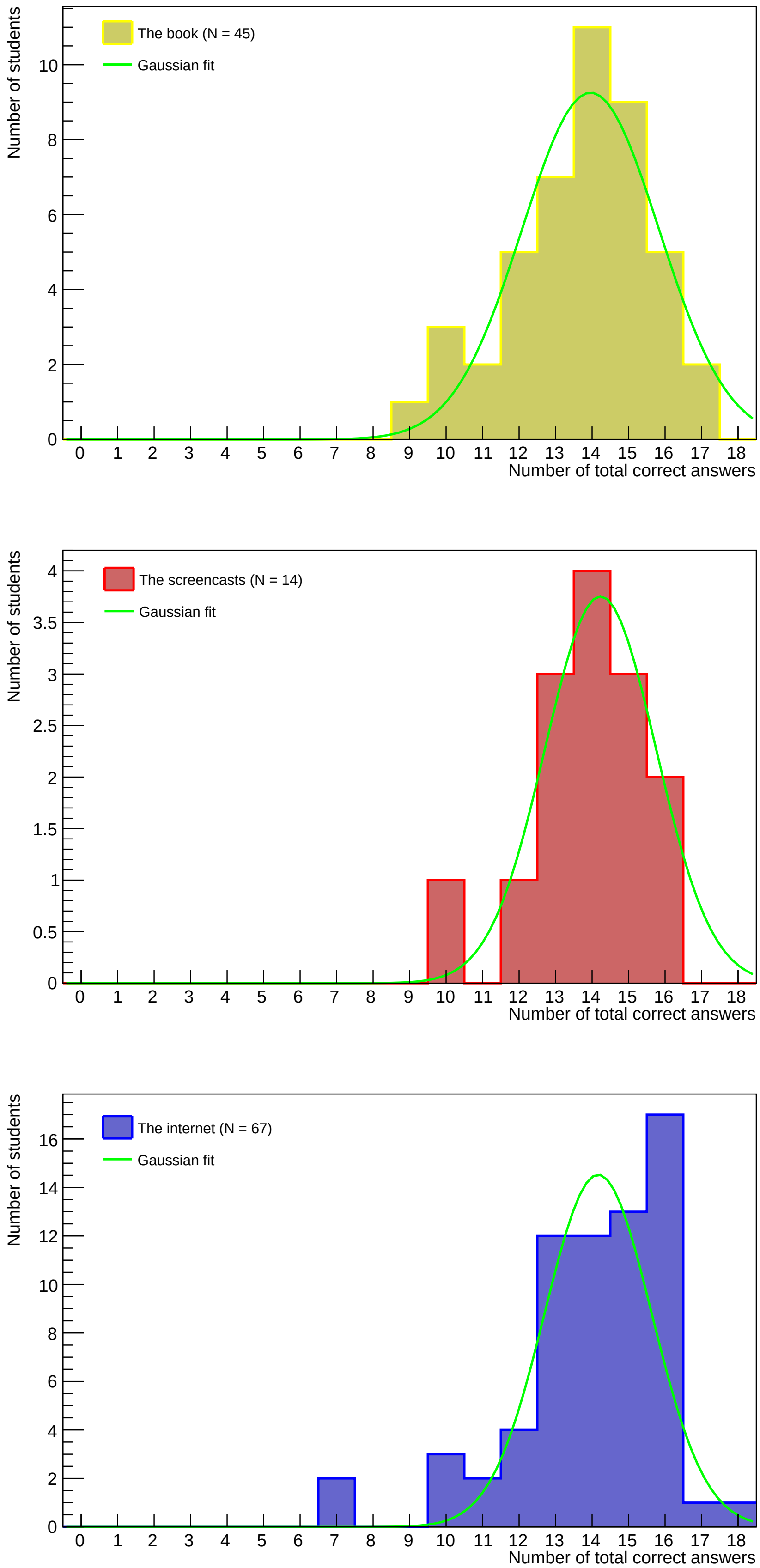


FIGURE 5.5 • THIS FIGURE SHOWS THE RESPECTIVE DISTRIBUTIONS FOR THE THREE MEDIA FROM FIGURE 5.4: THE BOOK (TOP), THE SCREENCASTS (MIDDLE); AND THE INTERNET (BOTTOM). FOR THE NUMERICAL RESULTS OF THE χ^2-GOODNESS OF FIT TEST WE REFER TO TABLE 5.11.

χ^2-GOODNESS OF FIT TEST	Book (yellow)	Screencasts (red)	Internet (blue)
Test statistic $(X_{\mathbf{GOF}})^2$	2.806849	1.012600	13.695942
Degrees of freedom	6	3	7
Normality constant	9.25 ± 2.15	3.75 ± 1.69	14.53 ± 2.58
Mean	14.96 ± 0.32	15.23 ± 0.56	15.17 ± 0.22
Spread	1.87 ± 0.38	1.52 ± 0.83	1.47 ± 0.17
Test probability	83.27%	79.82%	5.69%

LEVENE'S TEST	
Test statistic W	0.700982
Degrees of freedom $(\mathrm{df}_1,\mathrm{df}_2)$	2, 123
Probability	49.81%

ANOVA	ss_B	ss_W	$(s_P)^2$
Value	3.559700	481.244000	3.91255
Degrees of freedom	2	123	
Test statistic F	0.630903		
Effect size f	0.084975		
λ_F	0.909816		
Overall mean $\overline{\mu}$	14.88		
Probability	53.38%		
Power	12.28%		

TABLE 5.11 • THIS TABLE SHOWS THE SUMMERY OF A χ^2-GOODNESS OF FIT -, LEVENE'S -, AND AN ANOVA TEST. THE COLORS REFER TO THE DISTRIBUTIONS SHOWN IN FIGURE 5.4. NOTE THAT THE DISTRIBUTION OF TOTAL CORRECT ANSWERS, FOR THOSE STUDENTS WHO CLAIMED THAT THE INTERNET WAS MOST EDUCATIONAL, IS ONLY 5.69% LIKELY TO BE GAUSSIAN. THIS COULD IMPLY THAT IT IS IN FACT NOT GAUSSIAN. WITH A PROBABILITY OF 53.38% OF THE ANOVA IMPLIES THAT WE CANNOT REJECT THE HYPOTHESIS OF THE MEANS BEING EQUAL. SEE FIGURE 5.5 FOR THE FITTED GAUSSIANS WITH THE RESPECTIVE DISTRIBUTIONS.

Again 126 of the 136 students provided an informed answer to question 4. Here the same tendency as in question 3 is observed: No dependence between which of the three media the students felt they benefited most from when learning a new subject, and the NTC of the Lab Test; the independence test results in 67.68% probability of independence. See table 5.12 for further details.

The respective means of total correct answers with respect to the three media are again roughly equal which initially suggests that general media preference has no impact on the mean of total correct answers. See figure 5.6. By performing χ^2-goodness of fit test reveals that all three distributions can, statistically, be regarded as gaussian. However, there is a relatively high uncertainty in the three fit parameters, i.e normality constant, mean, and spread, for the internet distribution (blue distribution in figure 5.6). This high level of uncertainty in the parameters is most likely the reason for the high χ^2-goodness of fit probability of 86.31% of which we therefore remain critical. In addition, the difference between the mean of the fitted gaussian and the mean of the data in the internet-case is 3.48 which is relatively much greater than that in the books- and screencasts/videos-case. Thus we can at best safely assume normality for these distributions. A Levene's test results in a 6.10% chance of homescedasticity between the books and screencasts/videos distributions. This is very close to the cut off percentage of 5.00%. However, since we do not have more data to further investigate we can again only note a reason for concern regarding the homoscedasticity between these distributions; bottom line is that we still fail to reject the hypothesis of homoscedasticity.

An ANOVA then reveals that the means of the distributions of correct answers of those who found that they benefit most from books and screencasts/videos, respectively, when learning a new topic are 14.19% likely to be equal. We note that this is a somewhat low probability relative to the ANOVA performed in question 3 (see table 5.11) which may be due to a larger deviation from homoscedasticity. See table 5.13 for

CONTINGENCY TABLE

	Correct	Wrong	
Books ($N = 66$)	993	261	1254
Screencasts/videos ($N = 34$)	514	132	646
Notes on the internet ($N = 26$)	383	111	494
	1890	504	2394

RESULTS

	G-test	χ^2-test
Test statistic	0.780736	0.788996
Effect size	0.078717	0.079132
λ_{χ^2}	0.780736	0.788996
Probability	67.68%	67.40%
Power	11.34%	11.41%

TABLE 5.12 • CONTINGENCY TABLE TO TEST FOR DEPENDENCE BETWEEN THE NTC ON THE LAB TEST AND WHICH MEDIA THE STUDENTS FOUND MOST BENEFICIAL WHEN LEARNING A NEW SUBJECT. THIS TABLE IS CONSTRUCTED USING THE 126 INFORMED ANSWERS FROM QUESTION 4 OF THE LAB TEST.

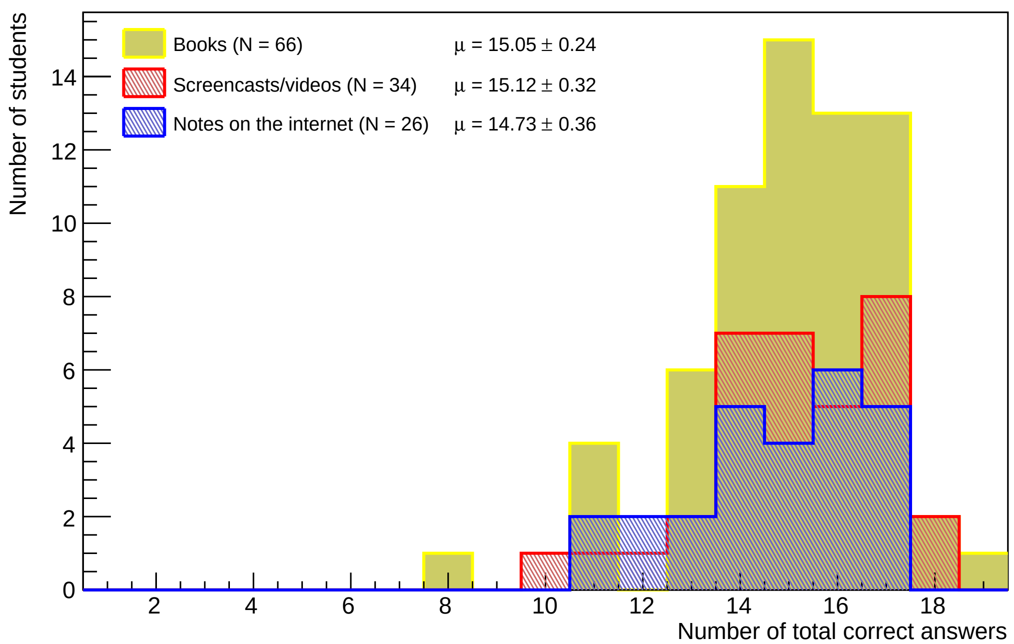


FIGURE 5.6 • THIS FIGURE SHOWS THE DISTRIBUTION OF TOTAL CORRECT ANSWERS FOR ALL THREE MEDIA IN QUESTION 4. THE DISTRIBUTIONS ROUGHLY HAVE THE SAME MEAN, WHICH SUPPORTS THAT NONE OF THE THREE MEDIA HAVE HAD AN IMPACT ON THE STUDENTS' ANSWERS OF THE LAB TEST.

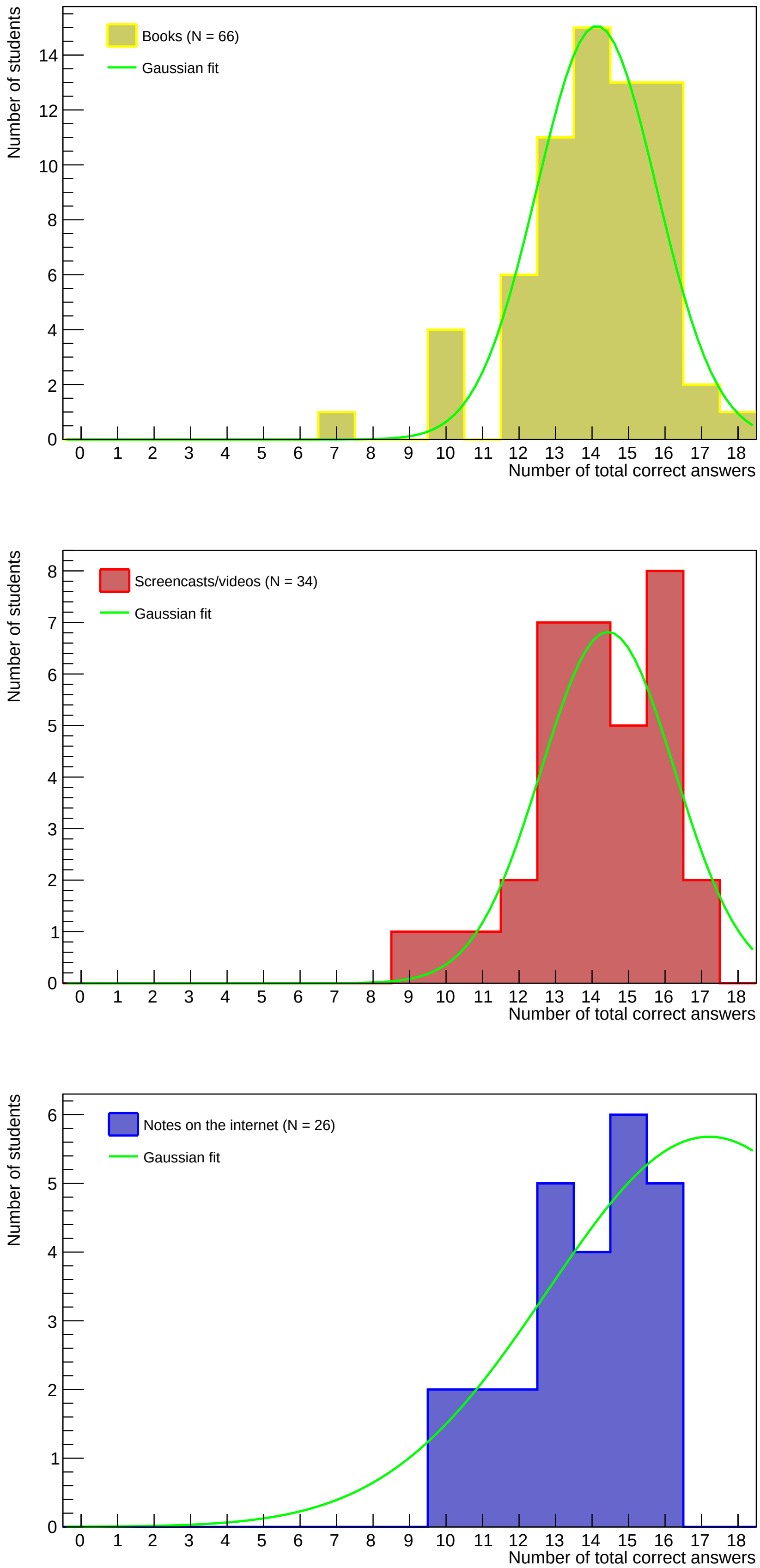


FIGURE 5.7 • THIS FIGURE SHOWS THE RESPECTIVE DISTRIBUTIONS FOR THE THREE MEDIA FROM FIGURE 5.6: BOOKS (TOP), SCREENCASTS/VIDEOS (MIDDLE); AND NOTES ON THE INTERNET (BOTTOM). FOR THE NUMERICAL RESULTS OF THE χ^2-GOODNESS OF FIT TEST WE REFER TO TABLE 5.13. NOTE THE POOR FIT FOR THE NOTES ON THE INTERNET-DISTRIBUTION.

χ^2-GOODNESS OF FIT TEST	Book (yellow)	Screencasts (red)	Internet (blue)
Test statistic $(X_{\mathbf{GOF}})^2$	6.768023	4.132125	1.107186
Degrees of freedom	6	6	4
Normality constant	15.06 ± 2.79	6.82 ± 1.71	5.68 ± 4.64
Mean	15.13 ± 0.31	15.43 ± 0.39	18.21 ± 10.09
Spread	1.64 ± 0.28	1.83 ± 0.39	4.42 ± 5.36
Probability	34.28%	65.88%	89.31%

LEVENE'S TEST (WITHOUT INTERNET)	
Test statistic W	3.592588
Degrees of freedom (df_1,df_2)	1,98
Probability	6.10%

ANOVA (WITHOUT INTERNET)	ss_B	ss_W	$(s_P)^2$
Value	0.116952	359.137119	3.664664
Degrees of freedom	1	98	
Test statistic F	2.191877		
Effect size f	0.017864		
λ_F	0.031913		
Overall mean $\overline{\mu}$	15.07		
Probability	14.19%		
Power	5.36%		

TABLE 5.13 • THIS TABLE SHOWS THE SUMMERY OF A χ^2-GOODNESS OF FIT TEST. THE COLORS REFERS TO THE DISTRIBUTIONS SHOWN IN FIGURE 5.6. NOTE THAT THE RESULTS BY FITTING A GAUSSIAN TO THE DISTRIBUTION OF TOTAL CORRECT ANSWERS, OF THOSE STUDENTS WHO IS MOST COMFORTABLE WITH USING THE INTERNET AS A SOURCE WHEN LEARNING A NEW SUBJECT, SHOWS VERY HIGH UNCERTAINTIES IN ALL THREE FIT PARAMETERS! THESE HIGH UNCERTAINTIES MAY IMPLY THE ACCORDINGLY HIGH PROBABILITY.

further details of the χ^2-goodness of fit-, Levene's test, and ANOVA.

To provide support for this claim we look at the errors in the distributions means in figure 5.6; we found that $\mu_{\text{Books}} = 15.05 \pm 0.24, \mu_{\text{Videos}} = 15.12 \pm 0.32$, and $\mu_{\text{Notes}} = 14.73 \pm 0.36$. We at once see that μ_{Books} and μ_{Videos} lies within their respective error intervals and μ_{Notes} does not. The SDOM, however, calculated here is under the assumption that the distributions are gaussian but in the case of notes of the internet we argued, on the basis of the χ^2-goodness of fit test, that this distribution is most likely not gaussian. Therefore we instead assume that it is poissonian as we *count* the number of correct answers for each student. The SDOM for a poisson distribution is $\sqrt{\mu/N}$; this is due to the fact that the standard deviation for a poisson distribution is equal to the square root of the mean, i.e $\sigma = \sqrt{\mu}$. Under this assumption we find that $\mu_{\text{Notes}} = 14.73 \pm 0.75$ which implies that μ_{Books} and μ_{Videos} lies in the error interval and μ_{Notes} in theirs; as such this points to the conclusion that general media preference does not affect the mean of total correct answers of the Lab Test.

In conclusion we cannot detect any correlation between the three general media, i.e books, screencasts/videos, and notes on the internet, and the NTC of the Lab Test. As the normality of the distribution, of total correct answers for those students who found notes on the internet to be most beneficial when learning a new topic in question 4, is seriously questionable we cannot infer anything, about whether preferring

CONTINGENCY TABLE

	Correct	Wrong	
The book ($N = 45$)	660	195	855
The screencasts ($N = 14$)	208	58	266
	868	253	1121

RESULTS

	G-test	χ^2-test
Test statistic	0.117335	0.116678
Effect size	0.044595	0.044470
λ_{χ^2}	0.117335	0.116678
Probability	73.19%	73.27%
Power	6.35%	6.35%

TABLE 5.14 • CONTINGENCY TABLE TO TEST FOR DEPENDENCE BETWEEN THE STUDENTS' TOTAL CORRECT AND WRONG ANSWERS OF THE LAB TEST AND WHETHER THEY FOUND THE BOOK OR THE SCREENCASTS THE MOST EDUCATIONAL DURING MEK1. THIS TABLE IS CONSTRUCTED USING IN TOTAL 59 INFORMED ANSWERS.

this media made a difference in regards to the mean of total correct answers or not relative to the others, using ANOVA: However, it does results in failure to reject that the means in total correct answers for those who answered books and screencasts/videos, respectively, are equal. Therefore whether a student finds that he prefers either books or screencasts/videos in general when learning a new topic does not impact the mean number of correct answers of the Lab Test according to the ANOVA. Considering the SDOM's of the respective distributions does support that preference of either general media, i.e including notes on the internet, does not impact the mean number of correct answers of the Lab Test.

In regards to the three possible answers in question 3 one could hypothesize that it may obscure any evidence of correlation. Recall that the possible answers in question 3 were *the book*, *the screencasts*, and *the internet*. An argument could be made that there is a greater probability of the students answering *the internet* simply due to a lack of more refined options, i.e *the internet* represents a too broad spectre in comparison. Therefore we could try to investigate whether there exists a correlation between only those who found the book and screencasts, respectively, most educational and the NTC. Performing an independence test for this scenario does, however, not yield a different result; there is still no detectable correlation. Table 5.14 shows the contingency table and independence test results. One solution could be to not supply *the internet* as a possible answer but instead give the option *neither*, i.e that nor the book or screencasts was the most educational in the setting of Mek1. Another could be to only have provided the options *the book* and *the screencasts*.

5.2.2 DHD

Recall that the DHD was given at the very beginning of Mek1 and the Lab Test in the very end. It may therefore seem strange that we investigate media preference since this was provided after the duration of Mek1; as such the students did not even know of the the book and the screencasts at the time the test was given. However, it would be quite interesting to investigate whether the students' future use of a certain media had an impact of test scores in the past, i.e DHD test scores.

In total 175 students took the DHD, and the response and understood rates are shown in table 5.15. We note that the lowest response rate is 94.3% corresponding to 165 students, which is considered a good response rate. Note that this is not a measure for the students actually submitting an answer according to their own beliefs; only that a large percentage of them have submitted an informed answer. Of the understood rates we observe an outlier in question 3 with 57.10%, which is approximately four standard deviations away from the mean! The most likely explanation for this large deviation lies in the topic of the question: Question 3 asked the students what the mode of a set of whole numbers was; the only secondary educations where they might have encountered the concept of mode, prior to beginning their bachelor in physics, are HF and HHX as descriptive statistics is part of the core subjects of mathematics level C and B, respectively [Retsinformation, 2013a][Retsinformation, 2013b]. It should be noted that the subject matter of descriptive statistics in the HF and HHX educations is not explicitly described; therefore the concept of mode may or may not have been taught. In addition, physics is not part of the HHX curriculum; as such students who have attended this particular secondary education are not part of the student body of Mek1 unless they explicitly have received supplementary schooling in physics afterwards which of course is a possibility. These considerations implies that a large portion of students most likely have never heard the term before, which as a consequence have resulted in the observed low understood rate.

Question	1	2	3	4	5	6	7	8	9	10
Response rate [%]	98.86	98.29	96.57	99.43	99.43	94.29	99.43	99.43	98.29	99.43
Understood [%]	90.29	94.29	57.14	100.00	99.43	98.29	94.86	92.57	97.71	93.71
Question	11	12	13	14	15	16	17	18	19	20
Response rate [%]	98.29	99.43	99.43	100.00	100.00	98.86	99.43	98.86	99.43	98.86
Understood [%]	92.57	100.00	99.43	100.00	96.00	94.86	96.57	95.43	98.29	92.57
Question	21	22	23							
Response rate [%]	98.29	98.86	98.86							
Understood [%]	84.57	96.57	94.86							

TABLE 5.15 • A TABLE SHOWING THE RESPONSE RATE AND THE STUDENTS WHO DID NOT UNDERSTAND THE RESPECTIVE QUESTIONS IN PERCENT IN THE DHD. THE STUDENTS ARE THOSE IN THEIR FIRST YEAR OF 2013. THE LOWEST RESPONSE RATE AND THE PERCENTAGE OF STUDENTS NOT UNDERSTANDING A QUESTION ARE 94.29% AND 57.14% RESPECTIVELY. THE TOTAL NUMBER OF STUDENTS WHO TOOK THE TEST IS $N = 175$.

Overall the students did relatively well with a mean score of 10.10 ± 0.23 when considering that this is a much more difficult test than the Lab Test, wherein the mean score was 14.92 ± 0.17 (see figure 5.1). The distribution of total correct answers of the DHD is shown in figure 5.8. We note that this distribution have a greater spread than that of the Lab Test as $\sigma_{\text{DHD}} = 0.23 \cdot \sqrt{175} = 3.04 > \sigma_{\text{Lab Test}} = 0.17 \cdot \sqrt{135} = 1.98$. This relatively greater difference in the spread of the distributions may be explained by the level of difficulty of both the Lab Test and the DHD as (1) the DHD was written in English, and not in Danish like the Lab Test, meaning that the wording of the individual questions may have given rise to confusion as the students only relied on their secondary education English at the time the test was given, in which some of the technical terms in connection to data analysis may not be part of; and (2) the DHD was taken at the very beginning of Mek1 before any teaching implying the students answered the DHD using only the overall knowledge

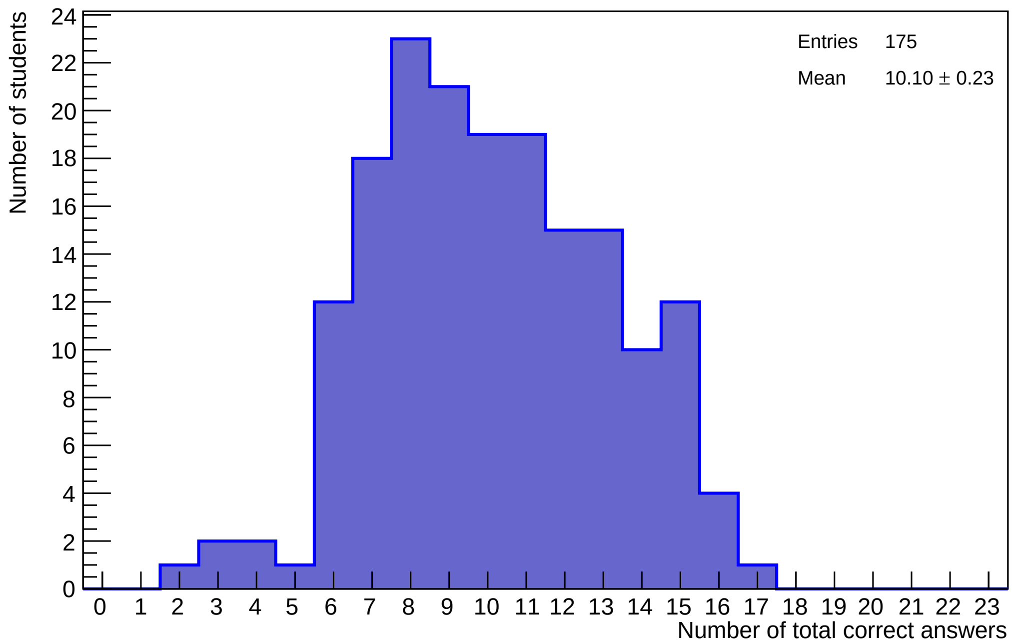


FIGURE 5.8 • SHOWS THE DISTRIBUTION OF TOTAL CORRECT ANSWERS PER STUDENT OF THE DHD. COMPAIRED TO THE DISTRIBUTION OF THE LAB TEST (SEE FIGURE 5.1) THE SPREAD IS LARGER AND THE MEAN IS LOWER. HOWEVER, WHEN WE TAKE INTO CONSIDERATION THAT THE DHD WAS GIVEN AT THE VERY BEGINNING OF MEK1, AND IT IS IN ENGLISH, THE STUDENTS DID OVERALL QUITE WELL.

obtained in their secondary education. In, comparison the Lab Test was given at the end of Mek1; as such the students have been taught the entire curriculum of the course, but not necessarily absorbed all of it, thereby having expanded their knowledge. In addition, the DHD contains questions addressing topics the students never have heard of; one example is the before mentioned mode of a data set.

Looking at the percentages of correct answers questionwise, shown in figure 5.9, only 11 questions have a percentage above 40: Questions 1,2,4,5,6,7,9,10,14,19,22. It is with the DHD, however, not possible to give a detailed description of what the remaining questions are as its authors wish to keep the questions from the general public, but the topics of the remaining questions which have a percentage below 40 are systematic errors, determine best trend lines, and determining the best mathematical model to shown data. We will, however, analyze questions 5 and 17 in greater detail as they are almost identical to 22 and 21, respectively, of the Lab Test. This is treated in section 5.2.3.

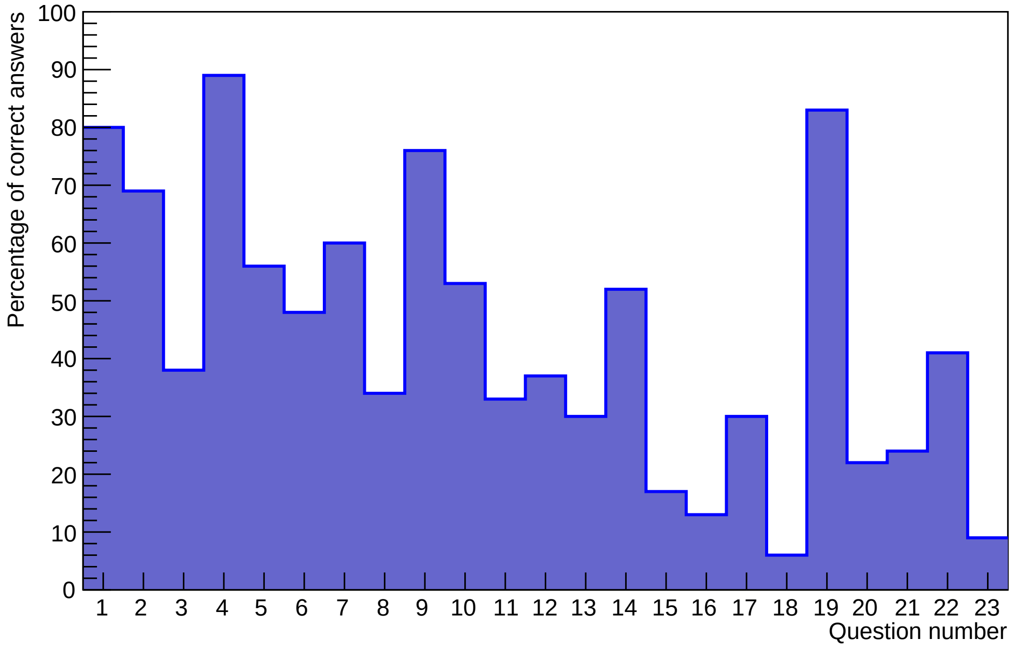


FIGURE 5.9 • THIS FIGURE SHOWS HOW THE TOTAL CORRECT ANSWERS QUESTIONWISE. ONLY 11 QUESTIONS HAVE A PERCENTAGE ABOVE 40 WHICH ARE 1,2,4,5,6,7,9,10,14,19,22. COMPARED TO THOSE OF THE LAB TEST (FIGURE 5.2) THE PERCENTAGES ARE LOW.

Equality of DHD means between years

The student body of Mek1 in 2013 is not unique in the sense that they are the only year given the DHD at the very beginning of Mek1. Of available data we have the students' responses from the years 2011, 2012, and 2013 in which they attended Mek1. Given we have these responses it is possible to investigate whether the students of Mek1 2013 are either better or worse equipped than the preceding years in regards to their knowledge of data analysis. Simply put: Investigate whether the students of Mek1 2013 have an advantage or disadvantage in data analysis when beginning Mek1.

From the responses of the respective years we find that in 2011 the student body comprised of 172 students; in 2012 it was 187; and, lastly, in 2013 it was 175 as was presented in the previous section. How the number of total correct answers are distributed with respect to the different years is shown in figure 5.10. We find that the means and their errors are $\mu_{2011} = 9.77 \pm 0.25$, $\mu_{2012} = 9.94 \pm 0.23$, and $\mu_{2013} = 10.10 \pm 0.23$. Initially, we see that the error intervals of μ_{2011} and μ_{2013} does not overlap. However, the respective means all fall within each others error intervals when expanding them to two SDOM's, i.e $\mu_{2011} = 9.77 \pm 0.50$, $\mu_{2012} = 9.94 \pm 0.46$, and $\mu_{2013} = 10.10 \pm 0.46$. On this basis the means of the distributions are most likely statistically equal under the assumption they can be considered gaussian.

To support this initial claim we perform an ANOVA; however, we find that there may be an issue in regards to the normality of the 2011 distribution (red distribution in figure 5.10) as a χ^2-goodness of fit test results in a probability of only 3.89%; strictly speaking this implies we reject the hypothesis of it being gaussian. In addition the test results the 2013 distribution (blue distribution in figure 5.10) only being 6.88% likely to be gaussian which is close to the cut off 5.00%; we do, however, strictly speaking not reject the notion that it may be gaussian. For further details see table 5.16; also the gaussian fits of the respective distributions are shown in figure 5.11.

From a pure visual standpoint all the distributions seem to have a gaussian tendency in terms of their shape. The reasons then for the relative poor fits may lie in the bin hits: Looking at the 2011 distribution in

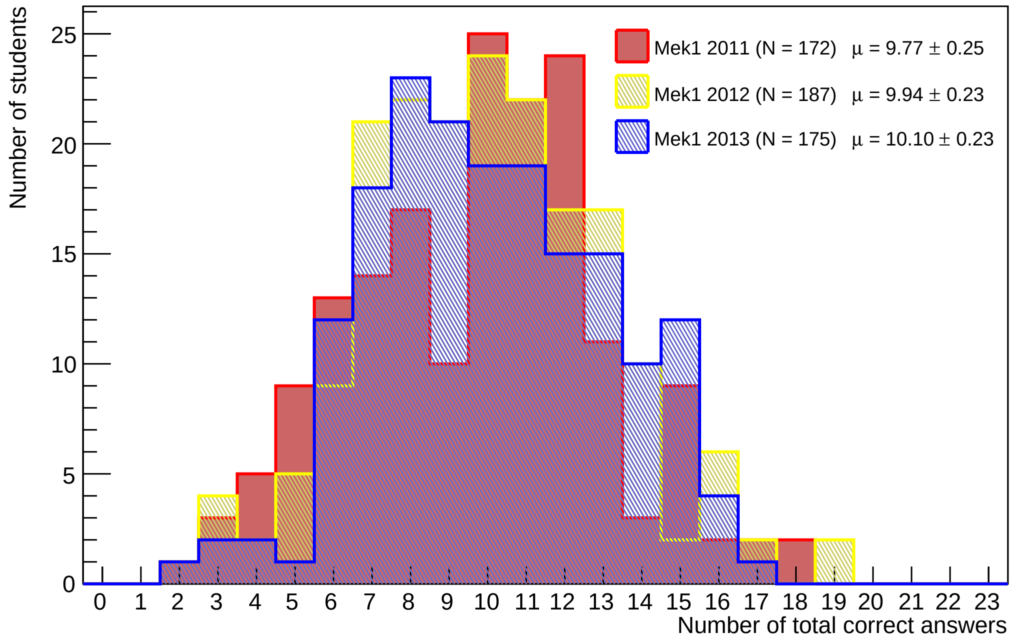


FIGURE 5.10 • DISTRIBUTIONS OF TOTAL CORRECT ANSWERS OF THE DHD OF THE STUDENT BODIES ATTENDING MEK1 IN 2011, 2012, AND 2013. NOTE THAT THE REPORTED UNCERTINTY ARE SDOM. JUDGING FROM THE MEANS AND THE ERRORS WE ARE 95.40% CONFIDENT THAT THE RESPECTIVE MEANS ARE CONTAINED IN THEIR RESPECTIVE ERROR INTERVALS, $\mu_{2011} = 9.77 \pm 0.50, \mu_{2012} = 9.94 \pm 0.46$, AND $\mu_{2013} = 10.10 \pm 0.46$.

figure 5.11 we notice that the eighth bin only contains ten hits, whereas the seventh and ninth bin contain 17 and 25 hits, respectively. The same pattern is also observed around the 13'th bin. For the 2013 distribution we observe drastic decrease in bin hits from bin four to bin five, i.e bin four has only 1 hit whereas bin five has 12. These relatively large drops in bin hits may cause the χ^2-goodness of fit test to yield a low probability of the distributions being gaussian.

Based on these arguments we proceed under the assumption that the distributions are gaussian; as such we find that they, according to a Levene's test, are homoscedastic. This enables us to perform an ANOVA which results in 54.01% chance of equal means of the distributions. Further details of both the Levene's test and ANOVA are provided in table 5.16. The ANOVA therefore support the initial claim of equality of the 2011, 2012, and 2013 distributions. This implies that the student body of Mek1 2013 was equally equipped in terms of their knowledge of data analysis when taking the DHD when comparing to that of Mek1 2011 and 2012.

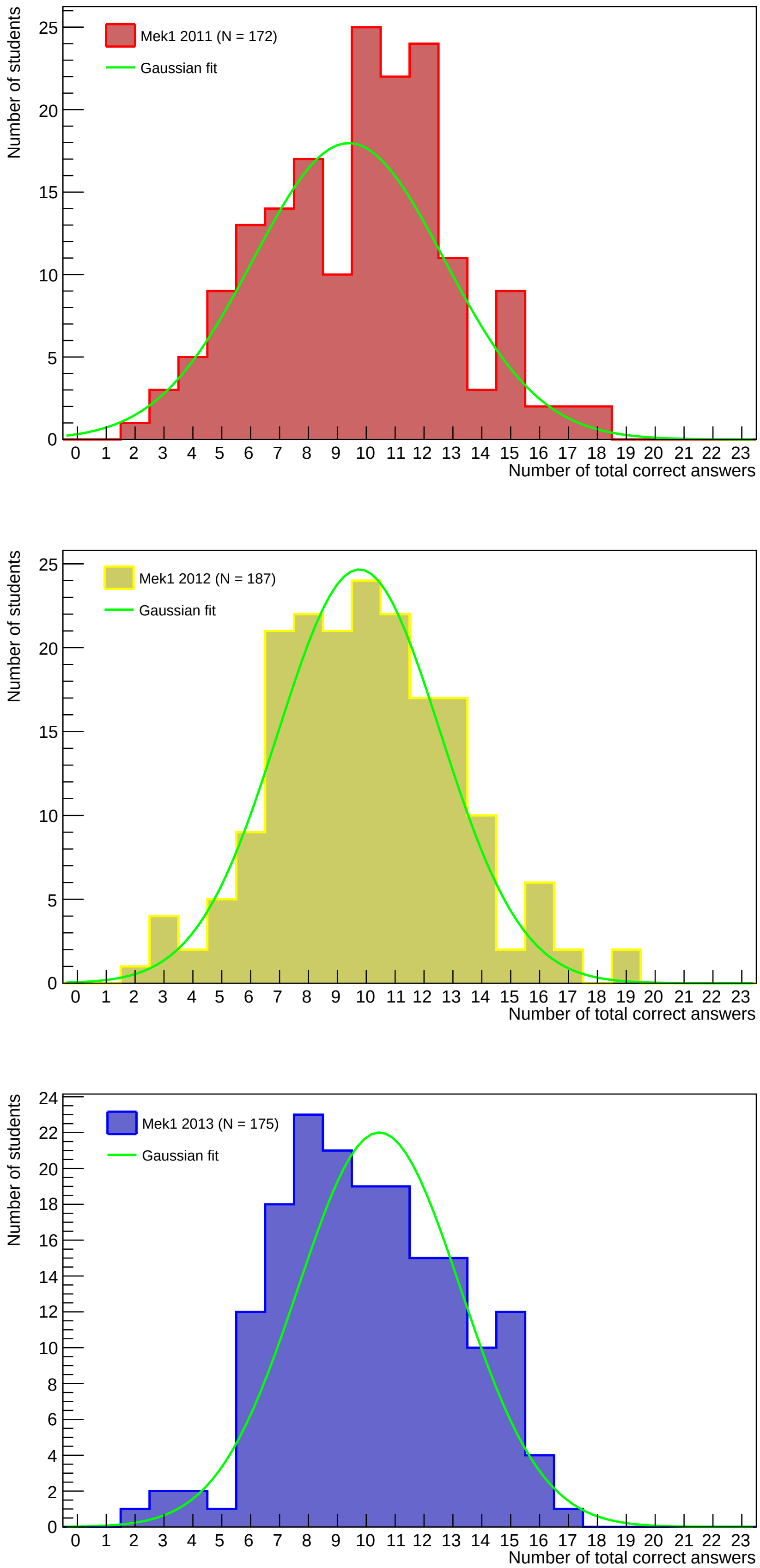


FIGURE 5.11 • SHOWS THE DISTRIBUTIONS OF 2011 (TOP), 2012 (MIDDLE), AND 2013 (BOTTOM) INCLUDING THE GAUSSIAN FIT. FITTING A GAUSSIAN TO THE 2011 DISTRIBUTION THE χ^2-GOODNESS OF FIT TEST RESULTS IN THE REJECTION OF THE NULL HYPOTHESIS OF IT BEING GAUSSIAN (SEE TABLE 5.16). MOST LIKELY IT IS THE DRASTIC DROPS, IN COMPARISON TO THE SURROUNDING BINS, IN HITS OF THE EIGHTH AND 13'TH BIN THAT RESULTS IN THE LOW PROBABILITY OF THE FIT.

χ^2-GOODNESS OF FIT TEST			
	Mek1 2011 (red)	**Mek1 2012 (yellow)**	**Mek1 2013 (blue)**
Test statistic $(X_{\mathbf{GOF}})^2$	24.583979	14.101007	21.213054
Degrees of freedom	14	14	13
Normality constant	17.97 ± 1.89	24.67 ± 2.44	22.01 ± 2.19
Mean	9.39 ± 0.29	9.76 ± 0.22	10.45 ± 0.25
Spread	3.31 ± 0.24	2.81 ± 0.18	2.81 ± 0.17
Probability	3.89%	44.22%	6.88%

LEVENE'S TEST			
Test statistic W	0.719991		
Degrees of freedom $(\mathrm{df}_1,\mathrm{df}_2)$	2,531		
Probability	48.72%		

ANOVA			
	ss_B	ss_W	$(s_P)^2$
Value	9.429800	5195.907433	9.785136
Degrees of freedom	2	534	
Test statistic F	0.616794		
Effect size f	0.042481		
λ_F	0.963686		
Overall mean $\overline{\mu}$	9.934457		
Probability	54.01%		
Power	12.90%		

TABLE 5.16 • NUMERICAL RESULTS OF A PERFORMED χ^2-GOODNESS OF FIT -, LEVENE'S TEST, AND ANOVA. THE COLORS REFER TO THE THAT OF THE DISTRIBUTIONS SHOWN IN FIGURE 5.10. WE GET LOW PROBABILITIES OF THE 2011 - AND 2013 DISTRIBUTION FROM THE χ^2-GOODNESS OF FIT TEST WHICH IS MOST LIKELY DUE TO RELATIVELY DRASTIC SHIFTS IN BINS HITS. IF DESPITE OF THIS ASSUME NORMALITY OF ALL DISTRIBUTIONS WE FIND THAT THEY ARE HOMOSCEDASTIC AND THE PERFORMED ANOVA LEADS TO THE CONCLUSION THAT THE MEANS ARE STATISTICALLY EQUAL.

Testing for best media

In order to investigate how media preference affects the outcome of the DHD we first had to identify the students who were given both the DHD and the Lab Test. Recall that the students submitted their birth date and full name in every test; this can be used to identify students who were given both tests. In the case of identical birth dates their names were additionally used to correctly pair questions 1-4 of the Lab Test and the answers of the DHD. Note that some chose not to supply their birth date; as such they cannot be included in this part of the analysis. However, it was possible in a few cases to determine the birth date, if it was not submitted, in the DHD, as it was submitted in the Lab Test and vice versa. In total we identify 120 students who took both the Lab Test and the DHD. The numerical distribution of their media responses is shown in table 5.17.

Q. 1 *Have you at any point used the book "Grundlæggende Matlab og dataanalyse", which can be found on Absalon?* Possible answers were *yes* ($N = 59$), *no* ($N = 33$) and *never heard of it* ($N = 26$).

Here 118 students in total provided an informed answer to question 1 in the Lab Test. We now test for dependence between the NTC of the DHD and whether the students during Mek1 had used, not used or never heard of the book. The contingency table and test results are shown in table 5.18. From the independence

	Question 1	Question 2	Question 3	Question 4
Possibility 1	59 (B_{DHD})	31 (S_{DHD})	40 (BO_{DHD})	55
Possibility 2	33 $\left(\overline{B}_{\mathrm{DHD}}\right)$	67 $\left(\overline{S}_{\mathrm{DHD}}\right)$	11 (SC_{DHD})	32
Possibility 3	26 $\left(\dot{B}_{\mathrm{DHD}}\right)$	21 $\left(\dot{S}_{\mathrm{DHD}}\right)$	59 (IN_{DHD})	24
Informed answers	118	119	110	111

TABLE 5.17 • THIS TABLE SHOWS THE MEDIA RESPONSES OF THE STUDENTS WHO TOOK BOTH THE DHD AND LAB TEST. THESE STUDENTS HAVE BEEN IDENTIFIED BY USING THEIR SUBMITTED BIRTH DATA AND NAMES IN BOTH TESTS.

test we do not observe a correlation between the NTC and whether the students have used the book or not as the probability of independence results in 22.39%.

Q. 2 *Have you seen any of the screencasts, which also can be found on Absalon?* Possible answers were *yes* ($N = 31$), *no* ($N = 67$) and *never heard of it* ($N = 21$).

In question 2 119 students provided an informed answer. By performing an independence test, to investigate whether the students' use of the screencasts are correlated with the NTC of the DHD, we strictly speaking observe no correlation between the two as the independence test results in 9.74%. However, the low probability may suggest that there is a correlation. The contingency table and test results are displayed in table 5.19.

One explanation for the low observed probability may be that we find a correlation between the NTC of the DHD and whether the students have not used or never heard of the screencasts; an independence test here results in 3.14% implying that we reject the null hypothesis of independence between these two categories. See table 5.20 for further numerical details. One likely interpretation of this may be that those students, who claimed they did not use the screencasts later may, have more actively sought information else where, i.e the book, the internet, or other literature whereas those who never had heard of them may have been more passive. This interpretation is somewhat supported by simple counting; we find that 33 out of the 67 students (49.25%), who claimed to not have used the screencasts, used the book later, i.e 33 ele-

CONTINGENCY TABLE

	Correct	Wrong	
Used the book ($N = 59$)	597	760	1357
Not used the book ($N = 33$)	331	428	759
Never heard of the book ($N = 26$)	286	312	598
	1214	1500	2714

RESULTS

	G-test	χ^2-test
Test statistic	2.993338	3.001013
Effect size	0.033210	0.033253
λ_{χ^2}	0.130145	0.130479
Probability	22.39%	22.30%
Power	32.09%	32.16%

TABLE 5.18 • CONTINGENCY TABLE AND RESULTS OF AN INDEPENDENCE TEST WHEN TESTING FOR CORRELATION BETWEEN THE NTC OF THE DHD AND WHETHER THE STUDENTS HAVE USED, NOT USED, OR NEVER HEARD OF THE BOOK. THE TEST PROBABILITY IMPLIES NO CORRELATION BETWEEN THE TWO.

CONTINGENCY TABLE

	Correct	Wrong	
Used the screencasts ($N = 31$)	317	396	713
Not used the screencasts ($N = 67$)	670	871	1541
Never heard of the screencasts ($N = 21$)	237	246	483
	1224	1513	2737

RESULTS

	G-test	χ^2-test
Test statistic	4.658621	4.674842
Effect size	0.041256	0.041328
λ_{χ^2}	0.202549	0.203254
Probability	9.74%	9.66%
Power	47.43%	47.58%

TABLE 5.19 • CONTINGENCY TABLE AND RESULTS OF AN INDEPENDENCE TEST WHEN TESTING FOR CORRELATION BETWEEN THE NTC OF THE DHD AND WHETHER THE STUDENTS HAVE USED ANY OF THE SCREENCASTS OR NOT. THE TEST PROBABILITY, STRICTLY SPEAKING, IMPLIES NO CORRELATION BETWEEN THE TWO WHEN USING A LEVEL OF SIGNIFICANCE OF 5.00%.

CONTINGENCY TABLE

	Correct	Wrong	
Not used the screencasts ($N = 67$)	670	871	1541
Never heard of the screencasts ($N = 21$)	237	246	483
	907	1117	2024

RESULTS

	G-test	χ^2-test
Test statistic	4.632154	4.646576
Effect size	0.047839	0.047914
λ_{χ^2}	0.201398	0.202025
Probability	3.14%	3.11%
Power	57.63%	57.76%

TABLE 5.20 • CONTINGENCY TABLE AND RESULTS OF THE PERFORMED INDEPENDENCE TEST FOR THOSE STUDENTS WHO CLAIMED TO NOT HAVE USED OR NEVER HAVE HEARD OF THE SCREENCASTS LATER AND THE NTC OF THE DHD. WE FIND THAT WE REJECT THE NULL HYPOTHESIS OF INDEPENDENCE BETWEEN THE TWO. THE INTERPRETATION OF THIS OBSERVATION IS THAT THOSE WHO CLAIMED NOT TO HAVE USED THE SCREENCASTS LATER KNEW OF THEIR EXISTENCE BUT MORE ACTIVELY SOUGHT INFORMATION IN THE BOOK, ON THE INTERNET, OR IN OTHER LITERATURE THAN THOSE WHO LATER NEVER HAD HEARD OF THEM.

ments in the set $B_{\mathrm{DHD}} \cap \overline{S}_{\mathrm{DHD}}$. Counting the elements in the set $\dot{S}_{\mathrm{DHD}} \cap \dot{B}_{\mathrm{DHD}}$ we find 12 which corresponds to 57.14% of the 21 students who claimed to never have heard of the screencasts also claimed to never have heard of the book later.

Q. 3 *Which of the following three media did you find most educational?* Possible answers were *the book from Q. 1* ($N = 40$), *the screencasts from Q. 2* ($N = 11$), and *internet (youtube.com, computerfysik.dk,.....)* ($N = 59$).

Of the 120 students who were given both the DHD and the Lab Test 110 provided an informed answer to question 3. When performing an independence test to shed light on a possible correlation between the students' choice of media, they found was most educational, it results in 8.80% which is quite low. This is most interesting as this result implies that there actually may be a correlation between the NTC of the DHD and which media they later, during Mek1, found most educational. However, strictly speaking there is not enough evidence to reject the null hypothesis of independence but the low probability warrants further investigation.

Investigating further we find that there is some correlation between the students who found either the book or screencasts most educational and the NTC of the DHD; an independence test here results in 3.75%. See table 5.22 for further details. Counting how many students who found the screencasts ($N = 11$) and the book ($N = 40$), respectively, most educational we find that all 11 students have also used the screencasts, and 39 of 40 have also used the book. In addition, if we count the number of elements in the sets $S_{\text{DHD}} \cap BO_{\text{DHD}}$ and $B_{\text{DHD}} \cap SC_{\text{DHD}}$ we find that they contain 6 and 8 elements, respectively. Therefore 54.55% of the 11 students who found the screencasts most educational have also used the book later; but only 20.00% of those who voted the book most educational also used the screencasts later. These mixed effects are believed to be the main cause for the low probability of the performed independence test.

For the three media their mean of total correct answers is computed to $\mu_{\text{Book}} = 9.68 \pm 0.45, \mu_{\text{SC}} = 11.36 \pm 0.75$, and $\mu_{\text{Internet}} = 10.37 \pm 0.42$. The distributions are shown in figure 5.12. Quickly we see that the error intervals of μ_{SC} and μ_{Internet} overlap. However, expanding the error intervals to two SDOM's we then see that they all overlap meaning that the means may very well be statistically equal.

In the pursuit of performing an ANOVA, in order to investigate this claim, we find that the assumption of normality for the distributions for those who answered the internet and the screencasts is violated; a χ^2-goodness of fit test results for the internet distribution in a probability 4.58% of being gaussian; and for the screencasts distribution it results in 95.14%, however, the fitted mean and spread are -8.81 ± 112.23 and 24.96 ± 12.75, respectively. The large uncertainties of the fitted mean and spread of the screencasts distribution (red distribution in figure 5.12) are without much doubt a result of a lack of data; here only 11 entries comprise the eight bins. Since the χ^2-goodness of fit test does not take empty bins into account a lot of information is lost giving rise to the poor fit. The low probability of the internet distribution (blue distribution

CONTINGENCY TABLE

	Correct	Wrong	
The book ($N = 40$)	387	533	920
The screencasts ($N = 11$)	125	128	253
The Internet ($N = 59$)	612	745	1357
	1124	1406	2530

RESULTS

	G-test	χ^2-test
Test statistic	4.861884	4.868586
Effect size	0.043837	0.043867
λ_{χ^2}	4.861884	4.868586
Probability	8.80%	8.77%
Power	49.19%	49.25%

TABLE 5.21 • THIS TABLE SHOWS THE PERFORMED INDEPENDENCE TEST WHEN TESTING FOR CORRELATION BETWEEN THE NTC OF THE DHD AND WHICH MEDIA THE STUDENTS FOUND MOST EDUCATIONAL. INTERESTINGLY, THE TEST RESULTS IN A LOW PROBABILITY WHICH CAN BE INTERPRETED AS SOME CORRELATION MAY EXITS. THIS MEANS THAT THE FUTURE CHOICE OF MEDIA MAY HAVE HAD AN IMPACT ON THE DHD TAKEN IN THE PAST.

CONTINGENCY TABLE

	Correct	Wrong	
The book ($N = 40$)	387	533	920
The screencasts ($N = 11$)	125	128	253
	512	661	1173

RESULTS

	G-test	χ^2-test
Test statistic	4.325781	4.348621
Effect size	0.060725	0.060887
λ_{χ^2}	4.325781	4.348621
Probability	3.75%	3.70%
Power	54.77%	54.99%

TABLE 5.22 • CONTINGENCY TABLE AND TEST RESULTS OF THE PERFORMED INDEPENDENCE TEST TO INVESTIGATE POSSIBLE CORRELATION BETWEEN STUDENTS WHO FOUND THE BOOK OR THE SCREENCASTS MOST EDUCATIONAL AND THE NTC OF THE DHD. THE TEST PROBABILITY IMPLIES THAT WE REJECT THE NULL HYPOTHESIS OF INDEPENDENCE. THE CAUSE OF THIS REJECTION IS BELIVED TO BE DUE MIXED EFFECTS AS 20.00% WHO VOTED THE BOOK MOST EDUCATIONAL HAVE ALSO USED THE SCREENCASTS AND 54.55% IN THE CONVERSE CASE.

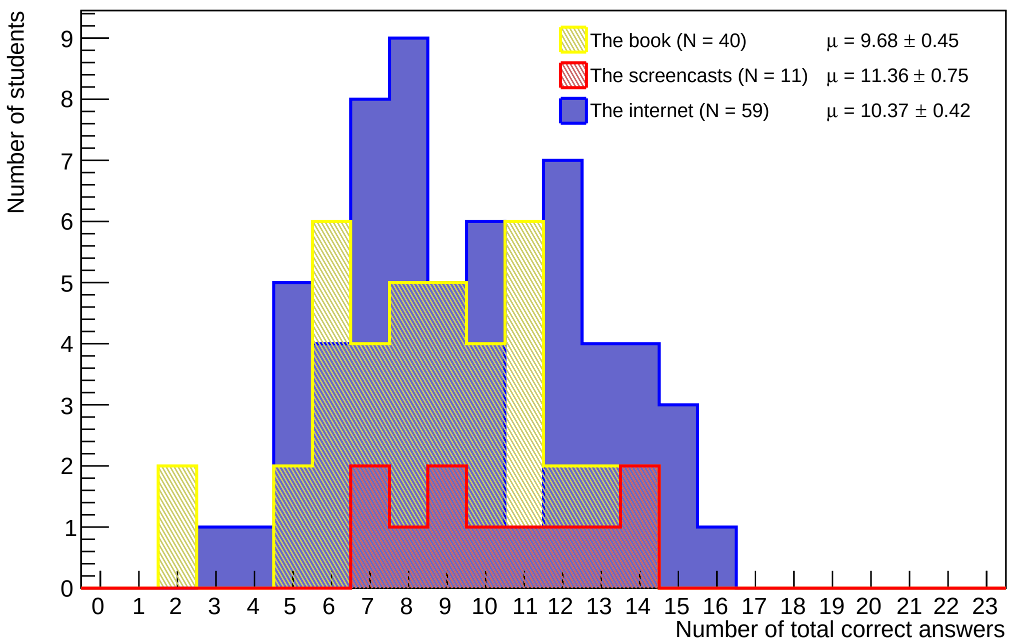


FIGURE 5.12 • DISTRIBUTIONS OF THE THOSE STUDENTS WHO FOUND THE BOOK, SCREENCASTS, AND INTERNET MOST EDUCATIONAL, I.E QUESTION 3 OF THE LAB TEST. IN TOTAL 110 PROVIDED AN INFORMED ANSWER. EXPANDSION OF THE ERRORS TO 95.00% CONFIDENCE INTERVALS SUGGESTS THAT THE MEANS ARE STATISTICALLY EQUAL AS THEY ALL OVERLAP.

in figure 5.12) is most likely a result of relatively alternating bin hits around the mean; especially around the ninth bin which only contains 1 hits whereas the tenth and 13'th contain 6 and 7 hits, respectively. Further details of the fit results are provided in table 5.23 and the respective distributions with their fitted gaussian

χ^2-GOODNESS OF FIT TEST	The book (yellow)	The screencasts (red)	The internet (blue)
Test statistic $(X_{\mathbf{GOF}})^2$	3.388719	1.129943	19.966107
Degrees of freedom	8	5	11
Normality constant	4.66 ± 1.10	1.72 ± 6.27	4.06 ± 0.93
Mean	9.32 ± 0.94	-8.81 ± 112.23	10.10 ± 1.01
Spread	3.85 ± 1.25	24.96 ± 12.75	4.28 ± 1.10
Probability	90.77%	95.14%	4.58%

LEVENE'S TEST (WITHOUT SCREENCASTS)	
Test statistic W	0.185650
Degrees of freedom (df_1,df_2)	1,97
Probability	66.75%

ANOVA (WITHOUT SCREENCASTS)	ss_B	ss_W	$(s_P)^2$
Value	11.610208	914.118394	9.423901
Degrees of freedom	1	97	
Test statistic F	0.193155		
Effect size f	0.111554		
λ_F	1.231996		
Overall mean $\overline{\mu}$	10.090909		
Probability	66.13%		
Power	19.57%		

TABLE 5.23 • THE GOODNESS OF FIT TEST FOR THE INTERNET DISTRIBUTION RESULTS IN REJECTION OF IT BEING GAUSSIAN (SEE BOTTOM IN FIGURE 5.13). THE REASON FOR THIS IS MOST LIKELY DUE TO RELATIVELY GREAT ALTERNATION IN BIN HITS AROUND THE MEAN. THE LARGE ERRORS, AND THEREBY THE HIGH PROBABILITY, OF THE FIT OF THE SCREENCASTS DISTRIBUTION IS SIMPLY DUE TO A LACK OF HITS IN THE BINS (SEE MIDDLE IN FIGURE 5.13). ASSUMING NORMALITY OF THE INTERNET DISTRIBUTION AN ANOVA YIELDS THAT ITS MEAN AND THAT OF THE BOOK DISTRIBUTION ARE STATISTICALLY EQUAL.

are shown in figure 5.13.

If we assume that the internet distribution is gaussian we find that it, and the book distribution to have the same spread as a Levene's test results in 66.75%. In addition an ANOVA reveal that their means are statistically equal. See table 5.23 for further numerical details.

Whether this also is the case with the screencasts distribution is not known due to the lack of data. However, the mean and its uncertainty are relatively far from the remaining but this is under the assumption that the distribution is gaussian. Here it would be more prudent to assume a poissonian; in that case the SDOM is $\sqrt{11.36/11} = 1.02$. But this still results in we have to expand the error intervals to two SDOM's in order to observe them overlap.

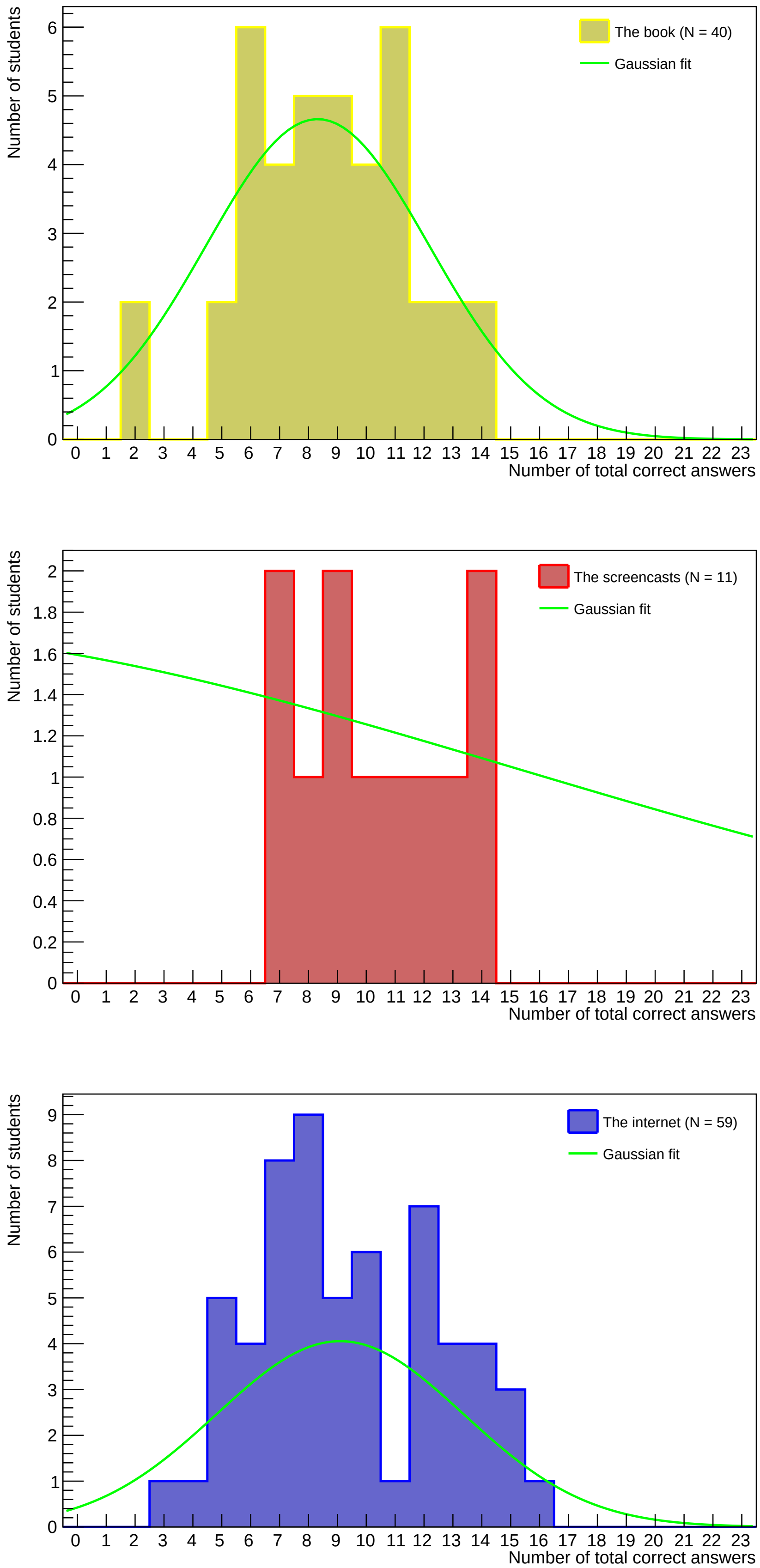


FIGURE 5.13 • DISTRIBUTIONS FOR THE STUDENTS WHO VOTED THE BOOK - (TOP), THE SCREENCASTS (MIDDLE), AND THE INTERNET (BOTTOM), WITH THEIR GAUSSIAN FIT, AS THE MEDIA THEY FOUND MOST EDUCATIONAL. FOR NUMERICAL RESULTS OF THE RESPECTIVE FITTED NORMALITY CONSTANTS, MEANS, AND SPREADS SEE TABLE 5.23.

CONTINGENCY TABLE

	Correct	Wrong	
Books ($N = 55$)	559	706	1265
Screencasts/videos ($N = 32$)	331	405	736
Notes on the internet ($N = 24$)	242	310	552
	1132	1421	2553

RESULTS

	G-test	χ^2-test
Test statistic	0.186725	0.186791
Effect size	0.008552	0.008554
λ_{χ^2}	0.186725	0.186791
Probability	91.09%	91.08%
Power	6.43%	6.43%

TABLE 5.24 • CONTINGENCY TABLE AND RESULTS WHEN TESTING OF DEPENDENCE BETWEEN THE NTC OF THE DHD, AND WHICH MEDIA THE STUDENTS FOUND MOST BENEFICIAL WHEN LEARNING A NEW TOPIC DURING MEK1. THE TEST PROBABILITY IMPLIES THAT THERE IS OVERWHELMING NO CORRELATION BETWEEN THE TWO CATEGORIES.

Q. 4 *Which of the following three media do you feel you benefit mostly from when learning something new?* Possible answers were *books* ($N = 55$), *screencasts/videos* ($N = 32$), and *Notes on the internet* ($N = 24$).

111 students provided an informed answer to question 4 in the Lab Test. Performing an independence test results in a probability of 91.09%; as such no correlation between the three general media and the NTC of the DHD is observed. Therefore which media the students found most beneficial when learning a new topic, later, is not correlated to the NTC of the DHD. See table 5.24 for further details.

The sample means of the three groups are $\mu_{\text{Books}} = 10.16 \pm 0.40, \mu_{\text{Videos}} = 10.34 \pm 0.45$, and $\mu_{\text{Notes}} = 10.08 \pm 0.68$. We quickly see that the error intervals overlap; as such they are most likely statistically equal. The distributions are shown in figure 5.14.

In the preparation of performing an ANOVA we find that we can assume that all three distributions are gaussian and homoscedastic. See table 5.25 for numerical results of the χ^2-goodness of fit - and Levene's test: The respective distributions and their gaussian fit are shown in figure 5.15. Looking at the notes on the internet distribution, see bottom of figure 5.15, it is somewhat unexpected that this can be assumed gaussian as only 4 out of 12 bins contain hits greater than 2. This is most likely the reason for these large uncertainties in the fitted mean and spread; this then results in a high χ^2-goodness of fit probability of 83.20%. For the books and screencasts/videos distributions, see top and middle of figure 5.15 respectively, we on the other hand observe relatively low probabilities; 22.33% and 19.04%, respectively. The reason for this is most likely due to the drastic drops of hits in some bins. In the case of the books distribution we see such a drop in hits in the tenth bin in which there is only 1 hit whereas 4 and 5 hits are contained in the ninth and 11'th, respectively. For the screencasts/videos distribution sudden drops are observed in the fourth and ninth bin.

Performing the ANOVA we find that the distributions statistically have the same mean as the test results in a probability of 39.39%. Therefore we observe no statistical evidence that the future preference of either the general media had an impact on the mean of correct answers of the DHD.

Strictly speaking, we can at this point conclude that the future use of the media the students found most educational, i.e the book, the screencasts, or the internet, is not correlated to the NTC of the DHD. However, since the independence test resulted in only 8.80% chance of independence there is reason to believe that

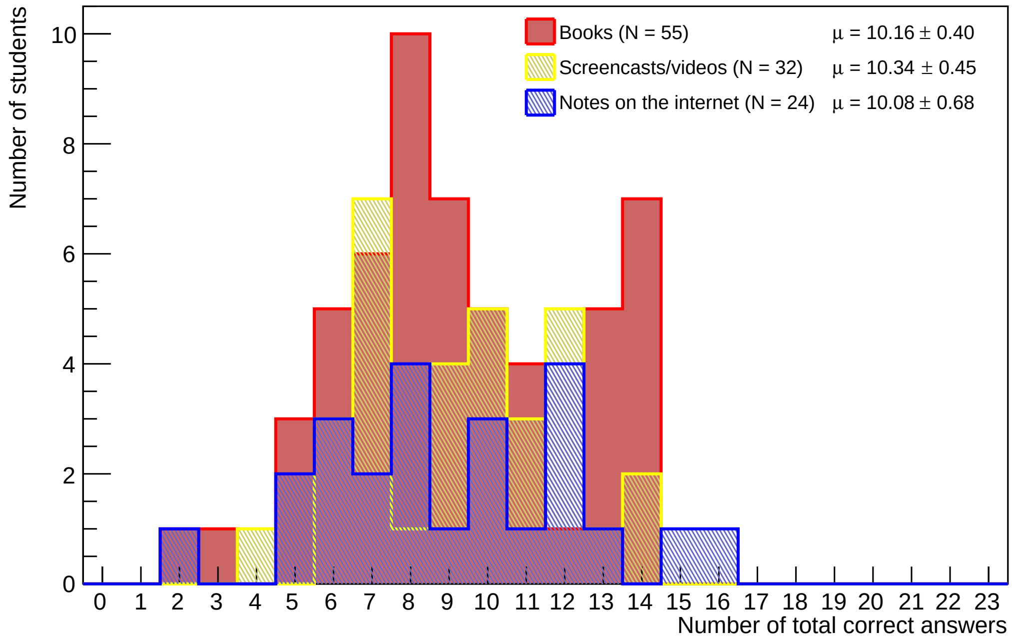


FIGURE 5.14 • SHOWS THE DISTRIBUTIONS OF CORRECT ANSWERS OF THOSE STUDENTS WHO CLAIMED THEY FOUND EITHER BOOKS, SCREENCASTS/VIDEOS, OR NOTES ON THE INTERNET MOST BENEFICIAL WHEN LEARNING A NEW TOPIC.

there may actually be some degree of association between the two. But the believed degree of association is most likely due to mixed effects from students using the screencasts or the book and found the other most educational. In addition, we found that the means of correct answers were statistically equal for those who claimed that the book and the internet, respectively, were most educational. For the screencasts distribution (red distribution in figure 5.12) a serious lack of data implied that the performed χ^2-goodness of fit test, to check for normality, was not trustworthy despite a test probability of 95.14%; this was argued to be mainly due to very large errors in the fitted parameters as the distribution only consisted of 11 hits in 8 bins! However, we found that the error intervals in the means for all three distributions overlap when expanded to two SDOM's whether the screencasts distribution were assumed gaussian or poissonian. On that basis we argued that the mean of the screencasts distribution is most likely equal to that of the book and the internet distribution.

Overwhelmingly, we also found no correlation between which media the students claimed to be most beneficial when learning a new topic and the NTC of the DHD; the independence test resulted in 91.09% chance of independence. We encountered here no problems with the normality of the three distributions. However, the χ^2-goodness of fit test for the notes on the internet distribution (blue distribution in figure 5.14) yielded a suspiciously high probability of 83.20%. The reason for this was most likely due to the relatively large errors in the fitted parameters (see table 5.25). An ANOVA revealed that the students' future preference of the three general media, i.e books, screencasts/videos, and notes on the internet, does not impact their mean in total correct answers. This was supported by the error intervals overlapping within one SDOM.

χ^2-GOODNESS OF FIT TEST			
	Books (yellow)	**Screencasts/videos (red)**	**Notes on the internet (blue)**
Test statistic $(X_{\mathbf{GOF}})^2$	11.826447	9.968557	5.026608
Degrees of freedom	9	7	9
Normality constant	7.09 ± 1.63	2.84 ± 0.81	1.98 ± 0.60
Mean	9.01 ± 0.43	10.50 ± 1.28	8.58 ± 2.51
Spread	2.57 ± 0.52	3.98 ± 1.70	6.05 ± 3.22
Probability	22.33%	19.04%	83.20%

LEVENE'S TEST	
Test statistic W	0.350464
Degrees of freedom $(\mathrm{df}_1,\mathrm{df}_2)$	2, 108
Probability	71.22%

ANOVA			
	ss_B	ss_W	$(s_P)^2$
Value	1.060284	938.163628	8.686700
Degrees of freedom	2	108	
Test statistic F	0.939766		
Effect size f	0.033161		
λ_F	0.122058		
Overall mean $\overline{\mu}$	10.198198		
Probability	39.39%		
Power	5.90%		

TABLE 5.25 • χ^2- GOODNESS OF FIT -, LEVENE'S TEST, AND ANOVA OF THOSE STUDENTS WHO FOUND BOOKS, SCREENCASTS/VIDEOS, OR NOTES ON THE INTERNET, MOST BENEFICIAL WHEN LEARNING A NEW SUBJECT. WE SEE THAT THE DISTRIBUTIONS CAN BE ASSUMED GAUSSIAN AND HOMOSCEDASTIC. IN ADDTION ANOVA REVEALS THAT THE MEANS OF THE RESPECTIVE DISTRIBUTIONS ARE STATISTICALLY EQUAL. SEE FIGURE 5.15 FOR THE DISTRIBUTIONS AND THEIR GAUSSIAN FIT.

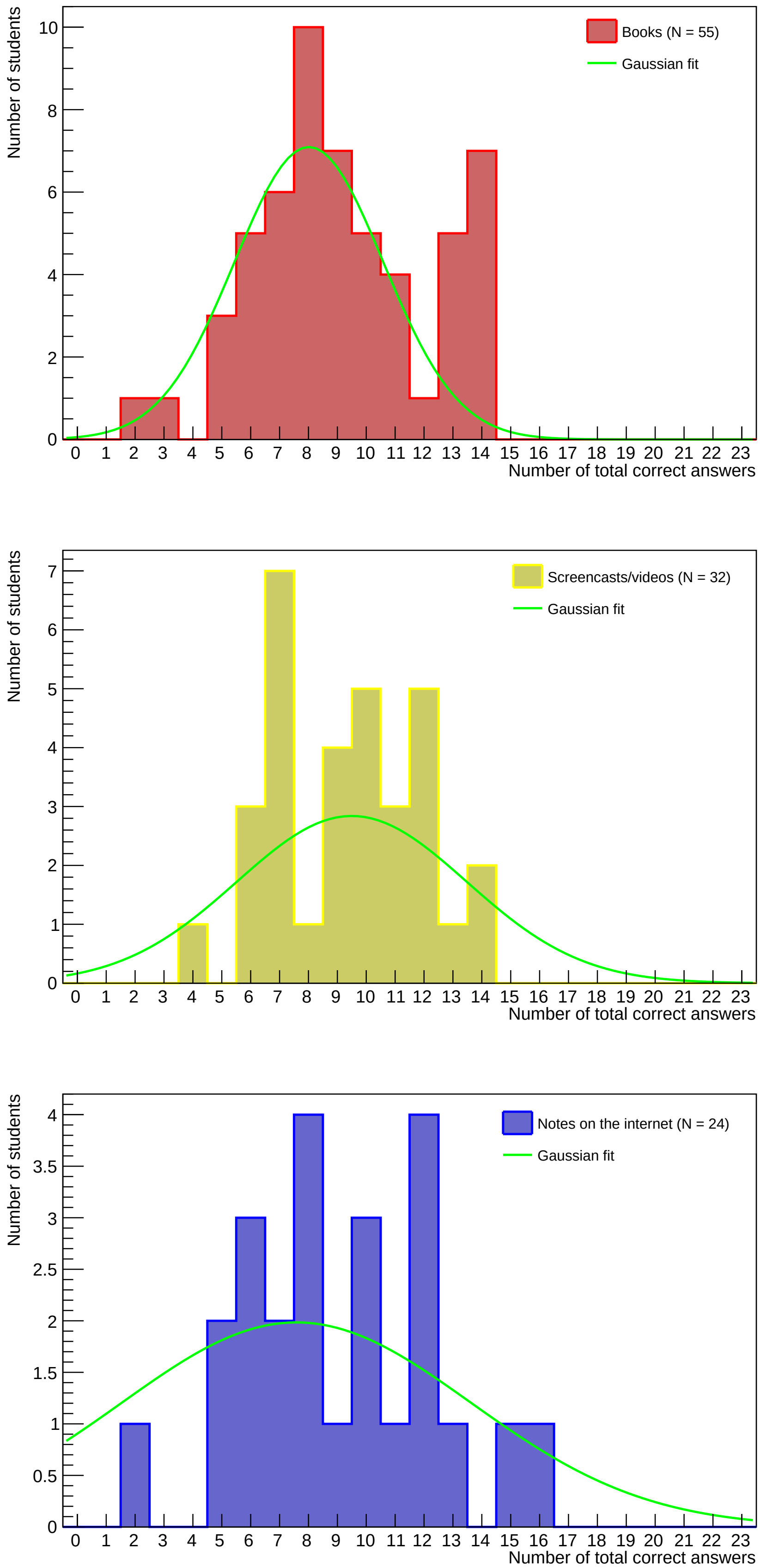


FIGURE 5.15 • DISTRIBUTIONS FOR THE STUDENTS WHO VOTED BOOKS - (TOP), SCREENCASTS/VIDEOS (MIDDLE), AND NOTES ON THE INTERNET (BOTTOM), WITH THEIR GAUSSIAN FIT, AS THE MEDIA THEY FOUND MOST BENEFICIAL WHEN LEARNING A NEW TOPIC. FOR NUMERICAL RESULTS OF THE RESPECTIVE FITTED NORMALITY CONSTANTS, MEANS, AND SPREADS SEE TABLE 5.25.

5.2.3 Paired Lab Test and DHD

As mentioned in chapter 3 two questions in the Lab Test and the DHD are almost identical; they were as follows:

- question 21 of the Lab Test is identical to question 17 in the DHD; it addresses the use of the error propagation formula,
- question 22 in the Lab Test is *similar* to question 5 in the DHD; it has been rephrased but it addresses is the exact same topic which is precision of a physical quantity.

Since the DHD was given at the very beginning of Mek1 - before any teaching - and the Lab Test in the end we have data which shows how the students answered before and after the teaching of Mek1 in regards to the same question. Note, however, that the two questions from the DHD have, in the Lab Test, been translated from English to Danish.

By performing paired independence tests, i.e McNemar's χ^2-tests, using the answers of the identical questions of the Lab Test and the DHD we can investigate if there is a correlation between the answers before and after the teachings of Mek1.

Error propagation formula - Lab Test question 21 VS. DHD question 17

First we look at question 21 in the Lab Test and 17 in the DHD. When disregarding the media preference of the students, i.e we consider all of the 120 students who took both the Lab Test and DHD, a McNemar test results in $8.60 \cdot 10^{-4}$%; as such the teachings of Mek1 certainly has made an impact! The question, however, remains to ascertain whether the impact is of a positive or negative nature; by computing the odds ratio (OR) of the McNemar test we can accomplish this. Recall from chapter 4 that the OR of a McNemar test is given by O_{12}/O_{21}; referring to the contingency table shown in table 4.2, the OR is in this setting the number of observed students who answered correct before and wrong after (O_{12}) divided by those who answered wrong before and correct after (O_{21}); as such if the taught curriculum of Mek1 has made a positive impact we would expect that O_{21} is greater than O_{12} implying that $OR < 1$. Computing the OR we find, when not considering any media preference, that $OR = 11/44 = 0.25$ which implies a positive impact. Further details of the test are shown in table 5.26.

To investigate if this significance also can be detected, when considering if the students have used the book, we perform the same McNemar test using the responses of the 59 students who claimed to have used it. Results of the test are shown in table D.6. The test again results in the rejection of the null hypothesis as the chance of no impact is 0.11%. Furthermore, $OR = 5/22 = 0.23$ implies that the use of the book, combined with the teaching, also has made a positive impact. We note that with an $OR = 0.23$ there is a small improvement relative to when not considering media preference (see table 5.26); however, since the difference is quite small it is most likely not significant.

It turns out, however, when using the responses from the 33 students who claimed to not have used the book, we again reject the null hypothesis as a McNemar test results in 2.01%; with an $OR = 0.25 < 1$ suggests that no use of the book and the teachings also have made a significant positive impact. For the remaining 26, who have never heard of the book, a McNemar test results 5.22% which is very close to the cut off of 5.00%; such a low probability, most likely, means that if more data where available we would reject the null hypothesis of independence: Again the impact is positive as $OR = 0.30 < 1$, however, not as positive when the book has been used. It is noteworthy that even though the impact is positive in the cases where the students have not used or never heard of the book, it is most positive when they have used the book. Further numerical details of the McNemar tests when considering no use and never heard of the book are shown in

CONTINGENCY TABLE
(NO MEDIA PREFERENCE)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	28	11	39
DHD Q.17 wrong	44	37	81
	72	48	120

RESULTS

	McNemar's χ^2-test
Test statistic	19.800000
Effect size	0.406202
Odds ratio	0.250000
λ_{χ^2}	19.800000
Probability	$8.60 \cdot 10^{-4}$%
Power	99.36%

TABLE 5.26 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF AN PAIRED INDEPENDENCE TEST, WHEN NOT CONSIDERING THE 120 STUDENTS' MEDIA PREFERENCE, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT. THE TEST RESULTS IN A PROBABILITY OF INDEPENDENCE MUCH LOWER THAN 5.00%, I.E WE REJECT THE NULL HYPOTHESIS OF INDEPENDENCE.

table D.7 and D.8, respectively.

When considering if the students have used, not used or never heard of the screencasts we observe the same trend as with the book: When they claimed to (1) have used the screencasts ($N = 31$) we observe correlation (0.27%) which is positive as $OR = 0.14$; (2) not have used the screencasts ($N = 67$) correlation (0.53%) is again observed which also is positive as $OR = 0.32$; and (3) never have heard of the screencasts we, strictly speaking, do not observe correlation, however, the test probability (5.78%) is very close to the cut off percentage. The detailed test results are shown in table D.9, D.10, and D.11.

So far the data shows significance both when the book or screencasts have been used; and close to significance if they never had heard of either. In addition, we in every case observe an $OR < 1$ meaning that the number of students answering correctly after the teachings of Mek1, and wrong prior, is greater than those who answered correctly prior and wrong after. These results overall suggests that the impact of the book and screencasts is most likely small compared to the teachings in regards to the use of the error propagation formula.

However, when we consider the students' responses, in regards to whether they found the book ($N = 40$), screencasts ($N = 11$) or the internet ($N = 59$) most educational, we find some results which points in the opposite direction: A McNemar test results in a chance of independence of $9.67 \cdot 10^{-2}$% and $OR = 0.13$ for the 40 students who found the book most educational. The test results are shown in table 5.27. An $OR = 0.13$ is the lowest yet observed; as such those who found the book most educational have the highest difference in answering wrong before and correct after. In addition it turns out that 39 of these students also have used the book; as such with the overwhelmingly low probability of independence this points strongly to the book having resulted in a very positive impact. The difference in wrong answers before and correct after is, for those who found the internet most educational, also significant with a chance of independence of 0.25% but $OR = 0.27$. See table D.12 for further details. When constructing the contingency table, using the responses from the 11 students, who found the screencasts most educational, the number of observed discordant pairs is unfortunately smaller than 10. This implies that the results from a performed McNemar

CONTINGENCY TABLE
(FOUND THE BOOK MOST EDUCATIONAL)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	10	2	12
DHD Q.17 wrong	16	12	28
	26	14	40

RESULTS

	McNemar's χ^2-test
Test statistic	10.888889
Effect size	0.521749
Odds ratio	0.125000
λ_{χ^2}	10.888889
Probability	$9.67 \cdot 10^{-2}$%
Power	90.99%

TABLE 5.27 • CONTINGENCY TABLE AND RESULTS OF A PERFORMED MCNEMAR TEST WHEN CONSIDERING THE RESPONSES FROM THE 40 STUDENTS WHO FOUND THE BOOK MOST EDUCATIONAL. WITH A PROBABILITY OF INDEPENDENCE OF $9.67 \cdot 10^{-2}$% AND $OR = 0.13$ WE THEREFORE OBSERVE THAT USE OF THE BOOK COMBINED WITH THE TEACHINGS OF MEK1 HAVE HAD A POSITIVE IMPACT IN REGARDS TO THE KNOWLEDGE OF PRECISION OF A PHYSICAL QUANTITY.

test, would most likely be inaccurate; as such it is not performed. The contingency table is, however, shown in table D.13.

If we instead consider which media the students, in general, found most beneficial when learning new topics McNemar's test results, in all cases, in significance and all with a chance of the independence smaller than 1.00%. However, when looking at the computed odds ratios, the data points to screencasts/videos as the best general media; here 32 students answered screencasts/videos and a McNemar test (see table 5.28) results in significance as 0.67% < 5.00% and $OR = 0.10$ which now is the lowest observed odds ratio. However, the odds ratio for those 24 who found notes on the internet most beneficial is not far from as $OR = 0.17$; for the 55 who answered books we find $OR = 0.25$ which is the same observed odds ratio when not considering media preference. The results of the McNemar tests for those who answered books and notes on the internet are shown in tables D.14 and D.15.

Overall the conclusion is that the teaching of the curriculum of Mek1 has made a positive impact in the students' ability to use the error propagation formula. This is somewhat unexpected as the its use is of an advanced nature, and it has not been been mentioned in the teachings of basic data analysis of Mek1 by Børge Svane Nielsen implying that it is not part of the mandatory curriculum of Mek1. This is based on the fact that it was not mentioned in any of his slides he used in his lectures, posted on Absalon. However, the very low change of no impact, $8.60 \cdot 10^{-4}$%, implies that the formula must have been treated at some point in connection to the teachings. It was undoubtedly mentioned and treated in chapter 3 of the book, and we found significant improvement, $9.67 \cdot 10^{-2}$% change of independence, in its use for the 40 students who voted the book most educational of which 39 also claimed to have used it. In light of this the book compared to the internet was statistically most effective. However, quite unexpected we also found that the use of the screencasts has had a positive impact; this implies that using the screencasts, of which none mentions the use of the error propagation formula, combined with the teachings of Mek1 also seems to strengthen the students' ability to use the formula. We note, however, that 19 students have used both the book and screencasts thereby resulting in mixed effects; this could obscure the findings as this is not specifically corrected for but it could also be support that the book is here the most effective media combined with the teachings.

CONTINGENCY TABLE
(FOUND SCREENCASTS/VIDEOS MOST BENEFICIAL)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	10	1	11
DHD Q.17 wrong	10	11	21
	20	12	32

RESULTS

	McNemar's χ^2-test
Test statistic	7.363636
Effect size	0.479702
Odds ratio	0.100000
λ_{χ^2}	7.363636
Probability	0.67%
Power	77.44%

TABLE 5.28 • CONSIDERING THE STUDENTS WHO FOUND THE GENERAL MEDIA SCREENCASTS/VIDEOS A MCNEMAR'S TEST YIELDS THAT PREFERENCE OF THIS MEDIA HAS A HIGH POSITIVE IMPACT WITH A CHANCE OF INDEPENDENCE OF 0.67% AND $OR = 0.10$. IN TERMS OF ODDS RATIO THIS IS THE BEST GENERAL MEDIA IN REGARDS TO THE USE OF THE ERROR PROPAGATION FORMULA AS IT IS THE LOWEST OBSERVED.

Precision of a physical quantity - Lab Test question 22 VS. DHD question 5

Following the same procedure we now investigate whether the teachings and media preference have made an impact on the students' knowledge of what precision of a physical quantity is; the questions which asked this were question 5 in the DHD and 22 in the Lab Test.

When disregarding the 120 students' media preference we again observe correlation as McNemar's test results in a chance of independence of 0.35%. See table 5.29 for numerical details. Therefore the teachings have, in regards to what precision of a physical quantity is, made an impact. In addition with an $OR = 0.36$ the impact is positive. However, the here observed odds ratio is much higher than that in the case of the use of the error propagation formula (see table 5.26) implying that not as many students answered correctly after the teaching relative to wrong prior. This is somewhat peculiar as the use of the error propagation formula can easily be construed as more difficult than precision of a physical quantity; comparison of the contingency tables in the case of no media preference, i.e tables 5.26 and 5.29, does somewhat shed light on this. We see that 23.33% answered correctly and 30.83% wrongly both before and after the teachings of Mek1 on the question about the use of the error propagation formula, and in the case of precision of a physical quantity 47.50% answered correctly before and after while 20.83% answered wrongly. In this specific case we thus observe an increase of 103.57% in students who have answered correctly both before and after the teachings in regards to what a physical quantity is relative to the use of the error propagation formula. Clearly, more students then knew what precision was from the beginning of Mek1; using this as a measure of difficulty then the use of the error propagation formula was, with this particular group of students, a more difficult topic than precision of a physical quantity which we in general previously claimed.

In addition greater frequencies are observed in the cells O_{11} and O_{21} of the contingency table which implies fewer frequencies available for cells of the discordant pairs, i.e O_{12} and O_{21}. In other words, when a larger number of students answered correctly, both before and after the teachings, it implies that the room for improvement was as a consequence smaller.

Regarding the students who have used the book a McNemar's test results in 1.05% chance for indepen-

CONTINGENCY TABLE
(NO MEDIA PREFERENCE)

	Lab Test Q.22 correct	Lab Test Q.22 wrong	
DHD Q.5 correct	57	10	67
DHD Q.5 wrong	28	25	53
	85	35	120

RESULTS

	McNemar's χ^2-test
Test statistic	8.526316
Effect size	0.266557
Odds ratio	0.357143
λ_{χ^2}	8.526316
Probability	0.35%
Power	83.15%

TABLE 5.29 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF A PERFORMED MCNEMAR TEST WHEN WE DISREGARD THE STUDENTS' MEDIA PREFERENCE. WE OBSERVE THAT THE OVERALL IMPLEMENTATION OF THE BOOK AND SCREENCASTS TOGETHER WITH THE TEACHINGS OF MEK1 HAS HAD A POSITIVE IMPACT IN REGARDS TO PRECISION OF A PHYSICAL QUANTITY.

dence and $OR = 0.29$ as shown in table D.16; as such the use of the book combined with the teachings have made a positive impact. The difference in the odds ratios between no media preference ($OR = 0.36$) and using the book is not great, but in the case of these students it would seem that using the book, combined with the teachings, has made a more positive impact which makes sense as the book directly treats the topic of precision. It is unfortunately not possible to test for independence for those who did not use and never have heard of the book, respectively, as the resulting contingency table in both cases have less than 10 discordant pairs. However, if we instead pool the responses, i.e we construct a contingency table using the responses for those who did not use and never had heard of the book, we observe no correlation as a McNemar test results in 13.36% chance of independence. Therefore the non use of the book, combined with the teachings, statistically has no impact! We still find an odds ratio less than one, $OR = 0.45$, implying that more students still have answered correct after than wrong prior. See table 5.30 for further details. Therefore it would seem that the book actually is a large contributor in teaching the students the concept of precision!

For the students, who claimed to have used the screencasts, we find that their use does not have an impact as a McNemar test results in 19.67% probability of independence as shown in table D.17; however, we find an $OR = 0.50$ which again is less than one despite the fact that we fail to reject the null hypothesis of independence. Interestingly, we do reject it when considering the students who did not use the screencasts; a test here results in 1.16% and $OR = 0.27$ which is shown in table 5.31. This is somewhat paradoxical as not one of the screencasts treated precision of a physical quantity which implies that their sole use during Mek1 should not significantly improve the chances of answering correctly after; but the teachings of Mek1 should. We found previously that use, or no use, of the screencasts had a positive impact in regards to the use of the error propagation formula even though none of them treated the topic and argued that the teachings therefore must have been the main contributor. Since precision of a physical quantity is an easier topic we would also here expect significance regardless of the students having used the screencasts, or not. One likely explanation of this paradox is that of the 74 students, who did not use the screencasts, 33 also used the book, and only 7 in the converse case. It would therefore be interesting to construct a contingency table, using the responses where we disregard those who did not use both the screencasts and the book. This, however, results in only six discordant pairs, i.e $D = 6 < 10$, and thereby violating the approximation requirements for the McNemar test. In addition it is not possible to minimize the mixed effects from the 38 students, who

CONTINGENCY TABLE
(NOT USED OR NEVER HEARD OF THE BOOK)

	Lab Test Q.22 correct	Lab Test Q.22 wrong	
DHD Q.5 correct	31	5	36
DHD Q.5 wrong	11	12	23
	42	17	59

RESULTS

	McNemar's χ^2-test
Test statistic	2.250000
Effect size	0.195283
Odds ratio	0.454545
λ_{χ^2}	2.250000
Probability	13.36%
Power	32.30%

TABLE 5.30 • RESULTS OF A PERFORMED MCNEMAR TEST AND CONTINGENCY TABLE USING THE POOLED RESPONSES OF THE STUDENTS WHO DID NOT USE AND NEVER HAVE HEARD OF THE BOOK, REFERRED TO AS NON USE OF THE BOOK. THE TEST RESULTS IN NO OBSERVED CORRELATION; AS SUCH THE NON USE OF THE BOOK COMBINED WITH THE TEACHINGS HAS NO IMPACT ON THE STUDENTS' RESPONSE PATTERN IN REGARDS TO PRECISION OF A PHYSICAL QUANTITY.

have used the book, by including the 21, who claimed to never have heard of the screencasts, as 7 of those also claimed to have used the book; as such the total number of students, who have not used or never have heard of the screencasts but have used the book, is $33+7=40$ of the $67+21=88$ students. Thus including said responses would not reduce the level of mixed effects greatly. This, however, supports that the book along with the teachings of Mek1 have made a great contribution in the students learning about precision!

As we found that use of the book results in a positive impact we would expect that the same is true for the 40 students those who found the book to be most educational since 39 of these also used it. Strictly speaking this turns out not to be the case: A McNemar test results in 8.96%, however, this probability is close to the cut off which may indicate that if a larger group of students had taken both the Lab Test and DHD it would result in significance; the odds ratio is, however, still less than one as $OR = 0.42$. The numerical details of the test are provided in table 5.32. It is unfortunately not possible to test for significance for the 11 students who found the screencasts most educational as there are only 5 discordant pairs. We do, however, find that the internet as a media made a positive impact, when regarding the responses of those who claimed it to be most educational; a McNemar test here results in 3.25% and $OR = 0.27$. See table D.18 for details.

In regards to the general media, the students found most beneficial when learning a new subject, 55 voted books, 32 screencasts/videos, and 24 notes on the internet. Clearly, books in general are, in the case of these particular students, the media which they voted most beneficial. However, the preference of this media does not significantly make an impact in answering the same question about precision of a physical quantity correctly after Mek1; a McNemar test results in 10.83% chance of independence but with $OR = 0.46$ as shown in table D.19. Even though the test does not result in observed significance we still observe an odds ratio which is less than 1 meaning that there are more students who answered wrong before but correct after than in the converse case. When regarding those who voted notes on the internet most beneficial we observe only 7 discordant pairs.

In case of screencasts/videos we observe 9 discordant pairs; this is fairly close to the required ten pairs. If we disregard that $D < 10$ a McNemar test results in 1.96% chance of independence and $OR = 0.13$; as such the preference of this media, combined with the teachings of Mek1, has actually made a positive impact.

CONTINGENCY TABLE
(NOT USED THE SCREENCASTS)

	Lab Test Q.22 correct	Lab Test Q.22 wrong	
DHD Q.5 correct	32	4	36
DHD Q.5 wrong	15	16	31
	47	20	67

RESULTS

	McNemar's χ^2-test
Test statistic	6.368421
Effect size	0.308304
Odds ratio	0.266667
λ_{χ^2}	6.368421
Probability	1.16%
Power	71.35%

TABLE 5.31 • MCNEMAR TEST RESULTS AND CONTINGENCY TABLE. WE OBSERVE THAT NO USE OF THE SCREENCASTS SIGNIFICATLY IMPROVES THE CHANCES OF ANSWERING CORRECTLY IN REGARDS TO THE QUESTION ON PRECISION AFTER THE TEACHINGS. THIS IS EXPECTED AS NONE OF THE SCREENCASTS TREATS THIS TOPIC, HOWEVER, THE TEACHINGS OF MEK1 DOES; AS SUCH WE WOULD ALSO EXPECT SIGNIFICANCE WHEN THE SCREENCASTS HAVE BEEN USED WHICH WAS THE CASE IN REGARDS TO THE USE OF THE ERROR PROPAGATION FORMULA. ONE EXPLANATION IS MIXED EFFECTS FROM 33 STUDENTS WHO DID NOT USE THE SCREENCAST BUT THE BOOK.

CONTINGENCY TABLE
(FOUND THE BOOK MOST EDUCATIONAL)

	Lab Test Q.22 correct	Lab Test Q.22 wrong	
DHD Q.5 correct	16	5	21
DHD Q.5 wrong	12	7	19
	28	12	40

RESULTS

	McNemar's χ^2-test
Test statistic	2.882353
Effect size	0.268438
Odds ratio	0.416667
λ_{χ^2}	2.882353
Probability	8.96%
Power	39.67%

TABLE 5.32 • WHEN REGARDING THE RESPONSES FROM THE STUDENTS WHO FOUND THE BOOK MOST EDUCATIONAL A MCNEMAR TEST RESULTS IN THIS MEDIA HAVING NO SIGNIFICANT IMPACT; HOWEVER, WE STILL OBSERVE AN ODDS RATIO SMALLER THAN ONE. THE TEST PROBABILITY IS CLOSE TO THE CUTOFF OF 5.00%; AS SUCH IF MORE DATA WERE AVAILBLE IT MAY THEN RESULT IN THE REJECTION OF NULL HYPOTHESIS OF INDEPENDENCE, I.E NO IMPACT.

CONTINGENCY TABLE
(FOUND SCREENCASTS/VIDEOS IN GENERAL MOST BENEFICIAL)

	Lab Test Q.22 correct	**Lab Test Q.22 wrong**	
DHD Q.5 correct	15	1	16
DHD Q.5 wrong	8	8	16
	23	9	32

RESULTS

	McNemar's χ^2-test
Test statistic	5.444444
Effect size	0.412479
Odds ratio	0.125000
λ_{χ^2}	5.444444
Probability	1.96%
Power	64.56%

TABLE 5.33 • CONTINGENCY TABLE AND RESULTS OF A MCNEMAR TEST OF THE RESPONSES FROM STUDENTS WHO FOUND SCREENCASTS/VIDEOS, IN GENERAL, MOST BENEFICIAL WHEN LEARNING A NEW SUBJECT. WHEN DISREGARDING THE FACT THAT ONLY 9 DISCORDANT PAIRS ARE OBSERVED THE TEST IS SIGNIFICANT IMPLYING THAT PREFERENCE OF THIS MEDIA HAS A POSITIVE IMPACT.

The test results are shown in table 5.33.

We found when disregarding media preference, i.e the overall coursework consisting of the teaching and availability of both the book and screencasts, had a positive impact. This was expected as the precision of a physical quantity is (1) not as difficult a topic than use of the error propagation formula which the fact that 103.57% more students here answered correctly both before and after Mek1; (2) treated directly by the teachings and the book.

The use of the book had a positive impact and not having heard of or not having used it resulted in no observed impact; as such this would suggest that use of the book, combined with the teachings of Mek1, significantly increases the students' ability to answer correctly! This was, strictly speaking, not the case when we considered those who found the book most educational; the test probability was, however, close to the cut off.

None of the screencasts treated precision and it therefore made sense that we found that their use did not have an impact. Not using them did, however, result in a positive impact: The reason hereto was argued to be due to mixed effects from the fact that 33 of 74 students who did not used the screencasts used the book; and we found that the use of the book resulted in a positive impact. Only 7 claimed to have not used the book the but the screencasts. This actually supported the claim that the book has been a significant contributor, along with the teachings, in teaching the students the concept of precision.

5.2.4 FCI

The FCI was given twice in the duration of Mek1; first time was before the beginning of the course (PRE-test) and again after the curriculum had been taught (POST-test). In total 175 students were given the PRE-test and 138 the POST-test hereby leaving 37 students unaccounted for. The reason for them not taking the POST-test may be that they either were not present the day of the POST-test or had dropped the course all together. It should be noted that some students may not even have been present the day of the PRE-test. Therefore the 175 students may not represent the full student body attending Mek1.

In table 5.34 the response and understood rates are shown for both the PRE- and POST-test. We note that both rates are very good; the lowest observed response and understood rate are both in the PRE-test which are 91.43% and 96.57%, respectively; these corresponds to 160 and 169 students, respectively. On this basis we, in both PRE- and POST-test, can be fairly certain that we have excellent response rates of each questions and also that the students knew what was asked for in the respective questions.

To identify the students who took both PRE- and POST-test, i.e we disregard if they also took the Lab Test and thereby media preference, we use their supplied name and birth date; in total we identify 129 such students. The distributions of correctly answered questions are shown in figure 5.16. We clearly see that there is an improvement of the mean of correctly answered questions; the mean of the PRE-test is $\mu_{\text{PRE}} = 20.60 \pm 0.59$ and that of the POST $\mu_{\text{POST}} = 24.42 \pm 0.44$. Therefore the increase is roughly four questions. From the uncertainties in μ_{PRE} and μ_{POST} we clearly see that each are not contained in the respective uncertainty intervals; not even within three standard deviations! Therefore the difference between μ_{PRE} and μ_{POST} is most likely

PRE-TEST										
Question	1	2	3	4	5	6	7	8	9	10
Response rate [%]	99.43	100.00	100.00	100.00	98.86	99.43	99.43	100.00	97.71	100.00
Understood [%]	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.43	98.86	99.43
Question	11	12	13	14	15	16	17	18	19	20
Response rate [%]	99.43	100.00	100.00	99.43	99.43	99.43	99.43	100.00	99.43	98.86
Understood [%]	100.00	100.00	100.00	100.00	99.43	100.00	100.00	99.43	98.29	100.00
Question	21	22	23	24	25	26	27	28	29	30
Response rate [%]	97.71	97.14	97.14	96.57	93.14	92.57	93.14	91.43	91.43	91.43
Understood [%]	97.14	98.86	99.43	98.86	99.43	99.43	100.00	96.57	100.00	97.71
POST-TEST										
Question	1	2	3	4	5	6	7	8	9	10
Response rate [%]	100.00	99.29	100.00	100.00	100.00	100.00	100.00	100.00	99.28	99.28
Understood [%]	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Question	11	12	13	14	15	16	17	18	19	20
Response rate [%]	99.28	99.28	99.28	100.00	100.00	100.00	100.00	99.28	100.00	99.28
Understood [%]	100.00	100.00	100.00	99.28	100.00	100.00	100.00	100.00	100.00	99.28
Question	21	22	23	24	25	26	27	28	29	30
Response rate [%]	99.28	100.00	100.00	100.00	99.28	99.28	100.00	98.55	97.83	99.28
Understood [%]	100.00	100.00	100.00	100.00	100.00	99.28	100.00	100.00	100.00	100.00

TABLE 5.34 • A TABLE OF SHOWING THE RESPONSE RATE AND THE STUDENTS WHO DID NOT UNDERSTAND THE RESPECTIVE QUESTIONS IN PERCENT IN THE PRE AND POST FCI. THE LOWEST RESPONSE AND UNDERSTOOD RATES ARE 91.43% AND 96.57%, RESPECTIVELY, WHICH IS CONSIDERED EXTREMELY GOOD. THE TOTAL NUMBER OF STUDENTS WHO TOOK THE PRE-TEST IS 175 AND 138 TOOK THE POST-TEST.

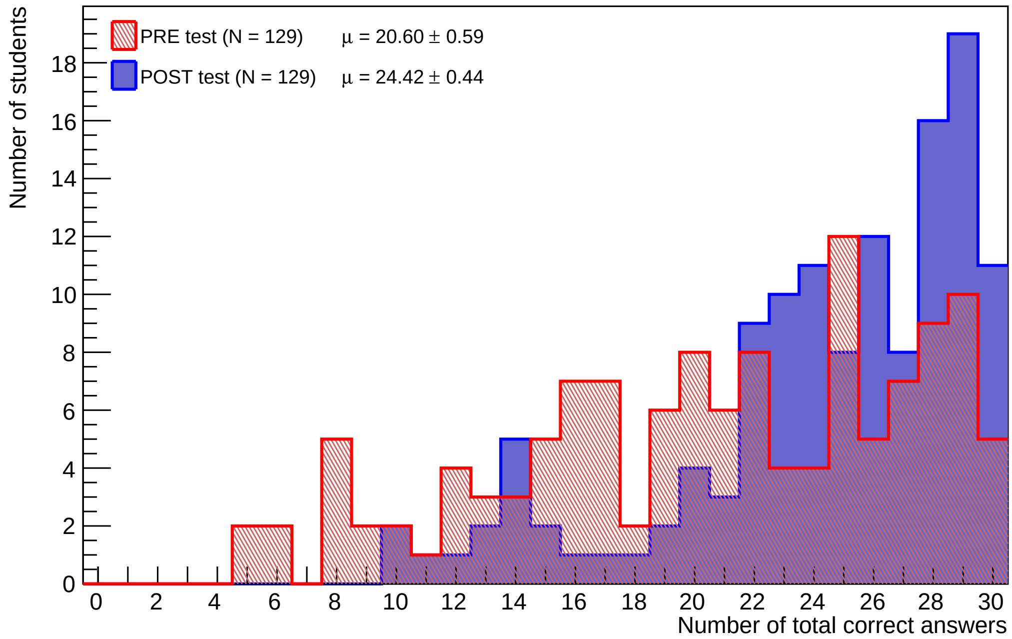


FIGURE 5.16 • DISTRIBUTION OF TOTAL CORRECT ANSWERS OF THE 129 STUDENTS WHO, IN 2013, TOOK BOTH THE FCI PRE- AND POST-TEST. WE QUICKLY NOTE THAT THE DISTRIBUTION FOR POST IS, RELATIVE TO PRE, SKEWED TO THE RIGHT. THIS IMPLIES THAT THE NUMBER OF TOTAL CORRECT ANSWERS IS GREATER THAN THAT OF THE PRE WHICH THE MEANS ALSO SHOW.

statistically greater than zero. By performing a Student's paired t-test we can ascertain whether this claim is supported. Recall, that the assumption for performing a Student's paired t-test was that the used sample of differences must be gaussian. The sample of differences in the context of the FCI is simply POST – PRE for each student. From this point on we will refer to the difference POST – PRE as the *gain* denoted by g.

A χ^2-goodness of fit test reveals that the g's are only 6.28% likely to be distributed according to a gaussian: The probability is most likely due to drastic changes in bin hits, especially around the mean 3.82 ± 0.17. However, we do not strictly speaking reject the null hypothesis of the distribution being gaussian as 6.28% > 5.00%. Therefore we assume that the normality assumption is valid. Performing a Student's paired t-test results in a $1.11 \cdot 10^{-14}$% chance of the mean of the g's, $\overline{g}$, being statistically equal to zero. We therefore reject the null hypothesis and accept the alternative that $\overline{g} > 0$. This implies that the teachings of Mek1, along with the availability of the book and screencasts, have resulted in a significant improvement in the mean test score of the FCI. See table 5.35 for further details.

χ^2-GOODNESS OF FIT TEST	
Test statistic $(X_{\mathbf{GOF}})^2$	30.455404
Degrees of freedom	20
Normality constant	9.49 ± 1.38
Mean	3.69 ± 0.52
Spread	4.25 ± 0.49
Probability	6.28%
STUDENT'S PAIRED t-TEST (NO MEDIA PREFERENCE)	
Test statistic T	9.495183
Effect size d	0.836004
λ_t	9.495183
Mean PRE $\mu_{\mathbf{PRE}}$	20.596899
Mean POST $\mu_{\mathbf{POST}}$	24.418605
Mean gain $\overline{g}$	3.821705
Degrees of freedom	128
Probability	$1.11 \cdot 10^{-14}\%$
Power	100.00%

TABLE 5.35 • 129 STUDENTS TOOK BOTH PRE- AND POST-TEST. PERFORMING A STUDENT'S PAIRED t-TEST SHOWS THAT THE MEAN GAIN, $\overline{g} = 3.82$, CONSTITUTES A SIGNIFICANT IMPROVEMENT AS THE TEST PROBABILITY IS NEAR ZERO. THE TEACHINGS OF MEK1, ALONG WITH THE AVAILABILITY OF BOTH THE BOOK AND SCREENCASTS, THEREFORE, HAVE RESULTED IN A SIGNIFICANT IMPROVEMENT OF $\overline{g}$.

5.2.5 Paired Lab Test and FCI

It may seem irrelevant to investigate whether the book or screencasts have made a significant difference in regards to the mean test score of the FCI as neither treated any topics of classical mechanics. However, an investigation may reveal something else entirely; as such it is warranted.

Following the same procedure, i.e using the supplied birth date, we determine the students who have taken both PRE- and POST-test and also the Lab Test; in total we identify 118 such students. How they answered questions 1 through 4 on the Lab Test, i.e media preference, is shown in table 5.36. Comparing tables 5.36 and 5.17 we see that the distributions of responses are almost identical and they follow the same pattern of the 136 students who took the Lab Test.

Q. 1 *Have you at any point used the book "Grundlæggende Matlab og dataanalyse", which can be found on Absalon?* Possible answers were *yes* ($N = 58$), *no* ($N = 33$) and *never heard of it* ($N = 26$).

	Question 1	**Question 2**	**Question 3**	**Question 4**
Possibility 1	58	34	41	58
Possibility 2	33	65	12	32
Possibility 3	26	19	55	20
Informed answers	117	118	108	110

TABLE 5.36 • THIS TABLE SHOWS THE MEDIA RESPONSES OF THE STUDENTS WHO TOOK THE FCI BOTH BEFORE AND AFTER MEK1 AND THE LAB TEST.

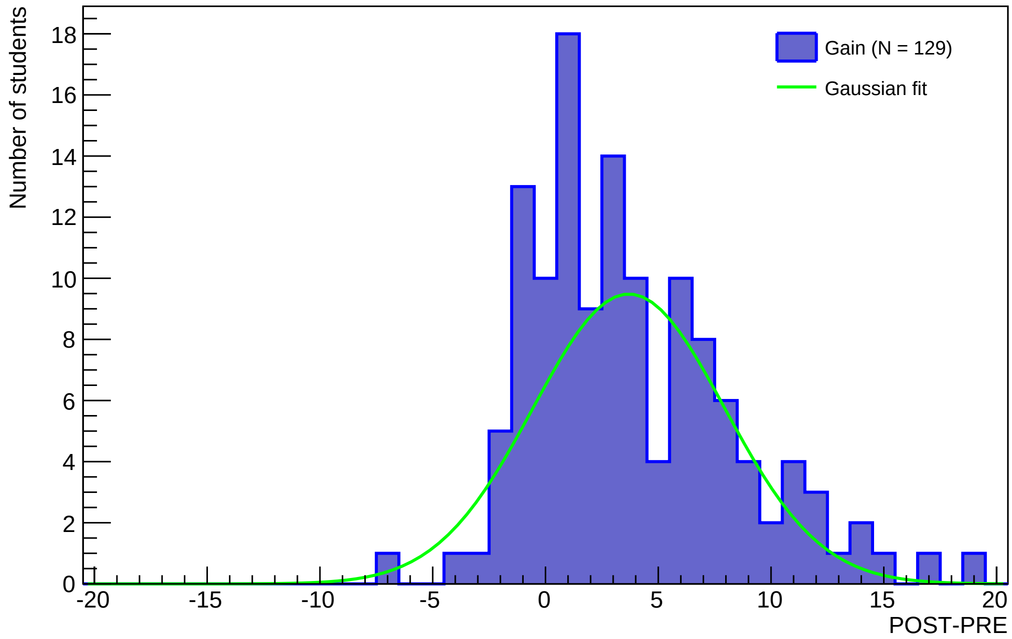


FIGURE 5.17 • DISTRIBUTION OF THE GAINS FOR THE 129 STUDENTS WHO WERE GIVEN BOTH THE FCI PRE- AND POST-TEST. A χ^2-GOODNESS OF FIT TEST RESULTS IN A PROBABILITY OF 6.28% OF THE DISTRIBUTION BEING GAUSSIAN. TABLE 5.35 SHOWS THE NUMERICAL DETAILS OF THE FIT RESULTS.

For each of the students who answered either one of the three possibilities to question 1 of the Lab Test, i.e *yes*, *no*, and *never heard of it*, we determine the gain. The distributions of the gains with respect to each of the response possibilities are shown in figure 5.18. Quickly, we see that in all three cases the mean of the gains are positive and greater than zero as $\mu_{\text{Used}} = 3.88 \pm 0.59$, $\mu_{\text{Not used}} = 3.67 \pm 0.79$, and $\mu_{\text{Never heard}} = 3.31 \pm 0.84$. We also clearly observe that a mean gain of zero is not contained in any of the error intervals implying that we observe a significant improvement of the mean gain regardless if the book has been used, not used or never heard of. This is supported by a Student's paired t-test as we find, in each case, that the teachings of Mek1 along with the availability of the book and screencasts statistically have improved the main gain of the FCI. The results of the test are shown in table 5.37.

Looking again at the mean gains in figure 5.18 we see that in this particular case the group of students who has used the book has the highest mean gain with $\mu_{\text{Used}} = 3.88 \pm 0.59$ and those who claimed to never have heard of it the lowest with $\mu_{\text{Never}} = 3.31 \pm 0.84$. Therefore the greatest difference in mean gains is $\mu_{\text{Used}} - \mu_{\text{Never}} = 0.57 \pm 1.02$ which, in all likelihood, gives rise to the notion that the mean gain of the three groups are not statistically different from each other: Performing an ANOVA provides support for this claim as we do not find sufficient evidence to reject that $\mu_{\text{Used}} = \mu_{\text{Not used}} = \mu_{\text{Never}}$ since the test results in 39.26%. See table 5.37 for further numerical details.

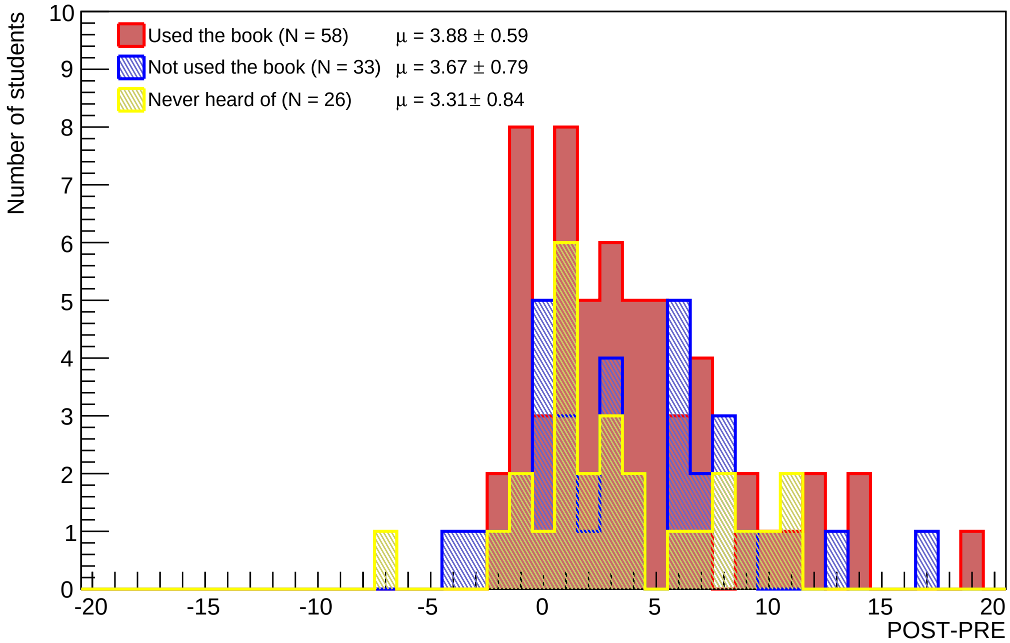


FIGURE 5.18 • THE DISTRIBUTIONS OF THE $\overline{g}$'S IN THE CASES WHERE THE STUDENTS HAVE USED, NOT USED, AND NEVER HEARD OF THE BOOK. THE GREATEST DIFFERENCE IN THE MEAN GAINS IS $\mu_{\text{USED}} - \mu_{\text{NEVER}} = 0.57 \pm 1.02$, WHICH IS QUITE LOW SUGGESTING THEY ARE STATISTICALLY EQUAL. AN ANOVA SUPPORTS THIS SUGGESTION. SEE TABLE 5.37.

χ^2-GOODNESS OF FIT TEST

	Used the book (red)	Not used the book (blue)	Never heard of the book (yellow)
Test statistic $(X_{\mathbf{GOF}})^2$	8.894607	7.673034	5.808105
Degrees of freedom	13	12	11
Normality constant	5.22 ± 1.05	2.27 ± 0.65	1.73 ± 0.57
Mean	2.80 ± 0.96	4.66 ± 2.12	2.43 ± 3.31
Spread	4.59 ± 1.15	6.81 ± 3.02	8.11 ± 5.99
Probability	78.09%	81.01%	88.59%

STUDENT'S PAIRED t-TEST

	Used the book	Not used the book	Never heard of the book
Test statistic T	6.554244	4.598692	3.850969
Effect size d	0.860614	0.800529	0.755237
λ_t	6.554244	4.598692	3.850969
Mean PRE $\mu_{\mathbf{PRE}}$	20.362069	20.060606	22.384615
Mean POST $\mu_{\mathbf{POST}}$	24.241379	23.727273	25.692308
Mean gain $\overline{g}$	3.879310	3.666667	3.307692
Degrees of freedom	57	32	25
Probability	$8.77 \cdot 10^{-7}\%$	$3.18 \cdot 10^{-3}\%$	$3.63 \cdot 10^{-2}\%$
Power	100%	99.78%	92.21%

LEVENE'S TEST

Test statistic W	0.942743
Degrees of freedom $(\mathrm{df}_1, \mathrm{df}_2)$	2, 114
Probability	39.26%

ANOVA

	ss_B	ss_W	$(s_P)^2$
Value	5.896110	2250.271553	19.739224
Degrees of freedom	2	114	
Test statistic F	0.859185		
Effect size f	0.050527		
λ_F	0.298700		
Overall mean $\overline{\mu}$	3.692308		
Probability	42.62%		
Power	7.25%		

TABLE 5.37 • TEST RESULTS OF χ^2-GOODNESS OF FIT -, STUDENT'S t-TEST, AND ANOVA. THE COLORS REFER TO THAT OF THE DISTRIBUTIONS SHOWN IN FIGURE 5.18. STUDENT'S t-TEST REVEALS THAT THE $\overline{g}$'S ARE STATISTICALLY GREATER THAN ZERO WHETHER THE STUDENTS HAVE USED, NOT USED OR NEVER HAVE HEARD OF THE BOOK. IN ADDITION AN ANOVA SHOWS THAT THE DISTRIBUTIONS OF THE $\overline{g}$'S ARE STATISTICALLY EQUAL IMPLYING THE USE, OR LACK HEREOF, DOES NOT IMPACT $\overline{g}$.

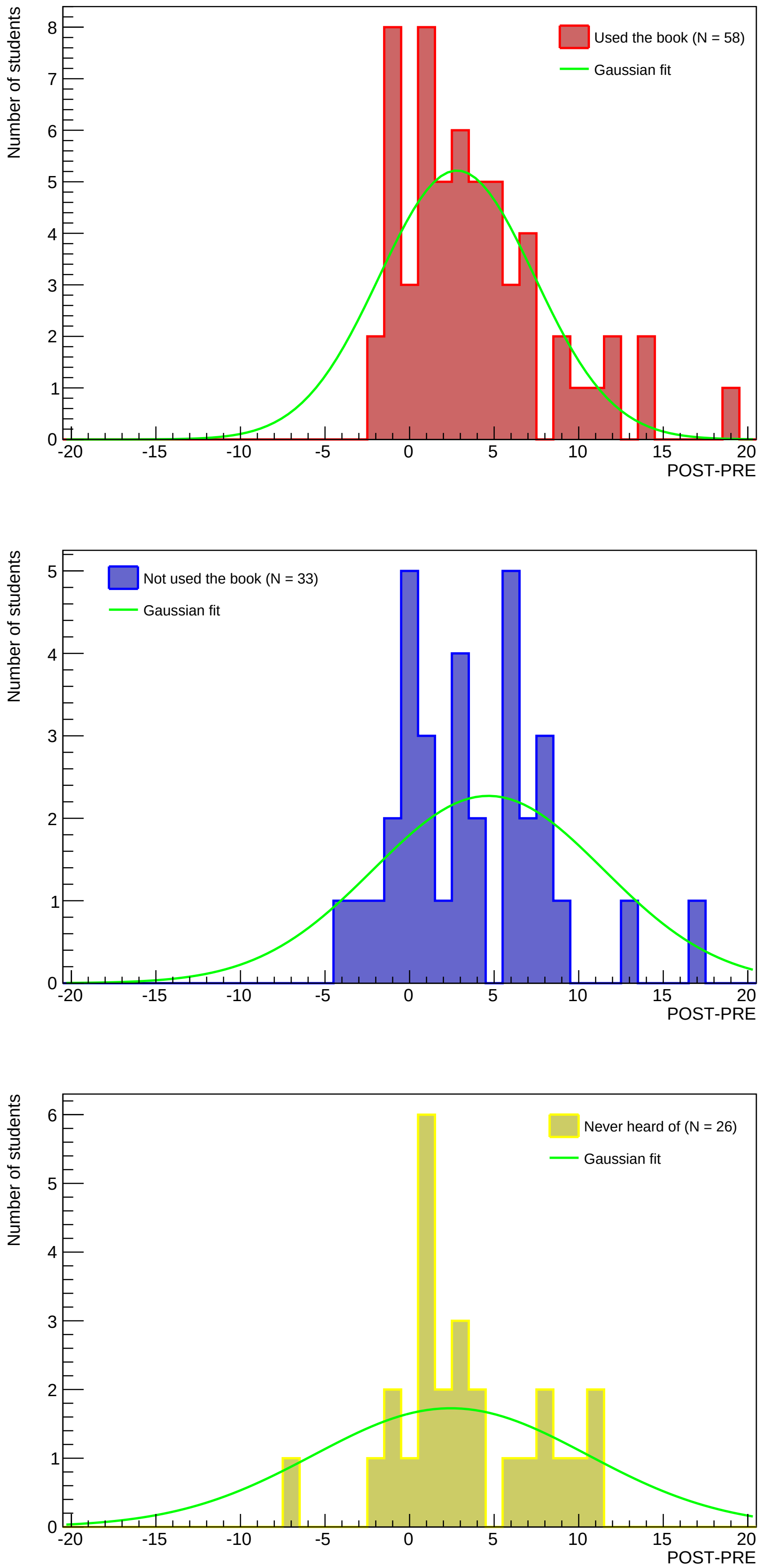


FIGURE 5.19 • DISTRIBUTIONS FOR GAINS AND THEIR GAUSSIAN FIT FOR THOSE WHO USED (TOP), AND NOT USED (MIDDLE), AND NEVER HEARD OF THE BOOK (BOTTOM). FOR NUMERICAL RESULTS OF THE RESPECTIVE FITTED NORMALITY CONSTANTS, MEANS, AND SPREADS SEE TABLE 5.37.

Q. 2 *Have you seen any of the screencasts, which also can be found on Absalon?* Possible answers were *yes* ($N = 34$), *no* ($N = 65$) and *never heard of them* ($N = 19$).

In figure 5.20 the distributions of the gains with respect to each of the response possibilities of question 2 of the Lab Test. Again we observe that the $\overline{g}$'s in all cases are greater than zero but still fairly close to each other numerically with $\mu_{\text{Used SC}} = 3.74 \pm 0.57$, $\mu_{\text{Not used SC}} = 3.88 \pm 0.61$, and $\mu_{\text{Never SC}} = 3.11 \pm 0.98$. We also see immediately that the error intervals overlap which would suggest that the mean gains again are statistically equal. In addition a mean gain of zero is again not contained in either of the error intervals. Initially, we therefore suspect that it does not matter whether the students have used, not used, or never heard of the screencasts, and that in all cases the mean gain is significantly greater than zero. By performing a Student's paired t-test and ANOVA can shed light on whether this is supported or not; before these tests can be performed we need to check for normality of the distributions, and if they are homoscedastic.

When checking for normality the distribution of gains, for the group of students who have never heard of the screencasts, that there is most likely not sufficient data to determine an acceptable fit, as a χ^2-goodness of fit test results in the fitted gaussian to have preposterous normalization constant, mean, and spread; also their uncertainties are very great. This is due to the histogram (yellow in figure 5.20) only having nine bins of which only two contains observations greater than 2. In order to still utilize the gains for this distribution we join them with the gains for the group who did not use the screencasts which we will denote by the non-use of the screencasts. Figure 5.21 shows the distributions of gains for the groups who claimed to have used and non-used the screencasts. These distributions can be assumed to be gaussian which χ^2-goodness of fit tests shows (see table 5.38). In figure 5.22 the distributions and their gaussian fits are shown.

We, however, note that the test probability for the fit of the non-used distribution (blue distribution in figure 5.22). is only 18.17% despite the relatively high number of bin hits. The reason for this low probability

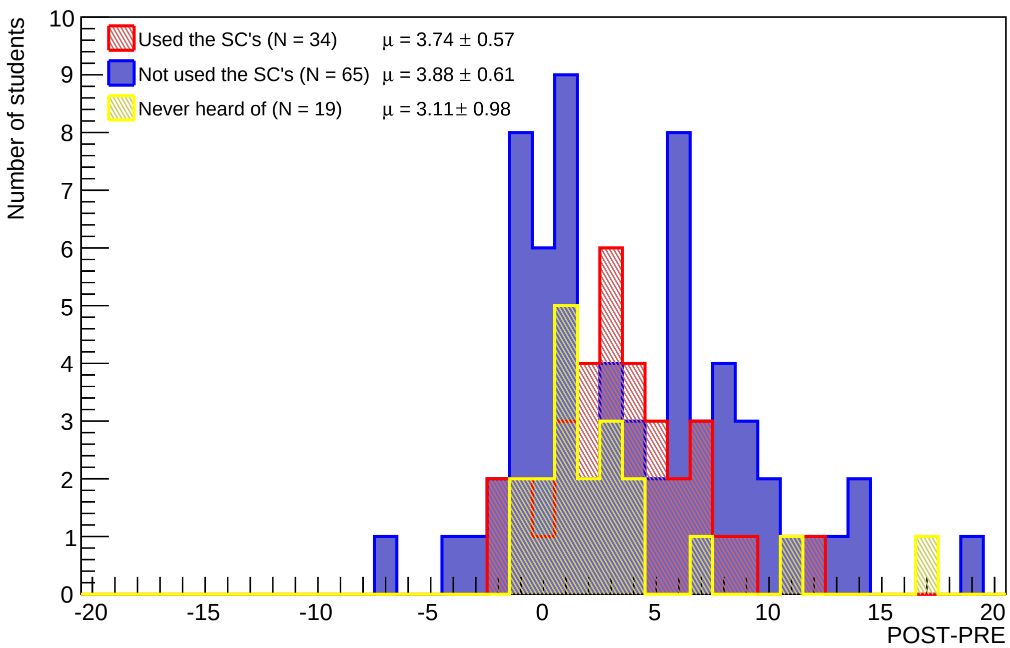


FIGURE 5.20 • SHOWS THE DISTRIBUTIONS OF GAINS WITH RESPECT TO WHETHER THE STUDENTS HAVE USED, NOT USED, OR NEVER HAD HEARD OF THE SCREENCASTS. JUDGING FROM THE SIMPLY THE MEAN GAIN OF THE DISTRIBUTIONS AND THEIR ERRORS THE MEANS ARE MOST LIKLY EQUAL AND ALSO SIGNIFICANTLY GREATER THAN ZERO, I.E $\overline{g} > 0$ FOR ALL THREE DISTRIBUTIONS.

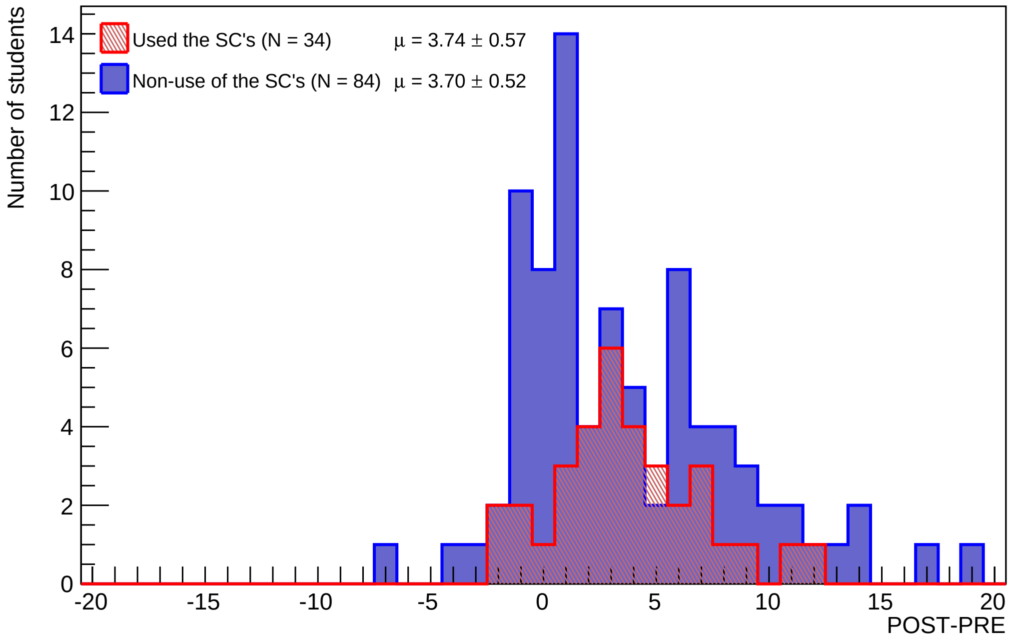


FIGURE 5.21 • DISTRIBUTIONS OF THE GAINS OF THE FCI WHEN CONSIDERING IF THE STUDENTS HAVE USED, OR NON-USED THE SCREENCASTS. NON-USED MEANS THAT THEY EITHER DID NOT USE OR NEVER HEARD OF THEM. THE MEANS AND THEIR ERROR STRONGLY SUPPORT THAT THESE DISTRIBUTIONS INDEED ARE EQUAL; ALSO THAT THE MEAN GAIN IN BOTH CASES IS SIGNIFICANTLY GREATER THAN ZERO.

is most likely drastic changes in bin hits; bin seven and nine have 14 and 7 hits, respectively, whereas bin eight only 4. Such a drastic change in bin hits is also observed around bin 11. In the end we, however, accept that the distribution is gaussian.

Having established that the used and non-used distributions can be considered are gaussian we can perform Student's paired t-tests; they reveal that the teachings of Mek1 have improved the mean gain significantly regardless of the use or non-use of the screencasts. This is not surprising as a mean gain of zero is not contained in either intervals $\mu_{\text{Used}} = 3.74 \pm 0.57$ or $\mu_{\text{non-used}} = 3.70 \pm 0.52$.

One peculiar observation reveals itself by performing an ANOVA to determine whether the distributions of the gains, in figure 5.21 are statistically equal; the test results in a probability of 0.12%. This implies that we reject the null hypothesis of the distributions being statistically equal; however, when both mean gains are contained in their respective error interval, suggests on the other hand that they should be statistically equal; as such the non-normality of the distribution of gains for the group who never have heard of the screencasts may seriously have tainted the joined data pool. If we perform a χ^2-goodness of fit test on the distribution of Not used SC-gains (blue distribution in figure 5.20) we find that this is 47.61% likely to be a gaussian; when the distributions are joined this probability drops to 18.17%, as shown in table 5.38. We therefore remain extremely sceptical of the result of the ANOVA!

χ^2-GOODNESS OF FIT TEST

	Used the SC's (red)	**Non-use of the SC's (blue)**
Test statistic $(X_{\mathbf{GOF}})^2$	5.070587	24.388503
Degrees of freedom	11	19
Normality constant	3.48 ± 1.18	4.93 ± 0.90
Mean	3.52 ± 0.78	3.73 ± 0.78
Spread	3.57 ± 1.40	5.13 ± 0.85
Probability	92.77%	18.17%

STUDENT'S PAIRED t-TEST

	Used the SC's	**Non-use of the SC's**
Test statistic T	6.447196	7.012306
Effect size d	1.105685	0.765105
λ_t	6.447196	7.012306
Mean PRE $\mu_{\mathbf{PRE}}$	20.147059	20.904762
Mean POST $\mu_{\mathbf{POST}}$	23.882353	24.607143
Mean gain $\overline{g}$	3.735294	3.702381
Degrees of freedom	33	83
Probability	$1.30 \cdot 10^{-5}\%$	$2.89 \cdot 10^{-8}\%$
Power	100.00%	100.00%

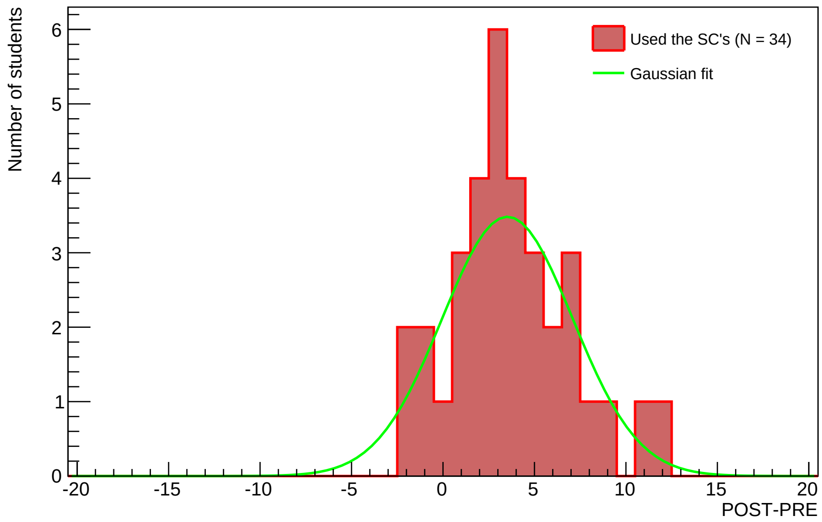
LEVENE'S TEST

Test statistic W	0.013550
Degrees of freedom (df_1,df_2)	1,116
Probability	90.75%

ANOVA

	ss_B	ss_W	$(s_P)^2$
Value	0.026219	2285.962568	19.7065174
Degrees of freedom	1	116	
Test statistic F	10.906377		
Effect size f	0.003358		
λ_F	0.001330		
Overall mean $\overline{\mu}$	3.711864		
Probability	0.12%		
Power	5.02%		

TABLE 5.38 • RESULTS OF PERFORMED χ^2-GODDNESS OF FIT -, LEVENE'S -, STUDENT'S PAIRED t-TESTS, AND ANOVA. THE COLORS REFER TO THOSE DISTRIBUTIONS SHOWN IN FIGURE 5.21. ULTIMATELY WE CAN ASSUME THAT THE DISTRIBUTIONS ARE GAUSSIAN AND HOMOSCEDASTIC. STUDENT'S PAIRED t-TEST REVEALS THAT THEIR MEAN GAIN ARE ALL SIGNIFICANTLY GREATER THAN ZERO. THE RESULT OF THE ANOVA IS, HOWEVER, VERY QUESTIONABLE. THE TWO DISTRIBUTIONS ARE SHOWED WITH THEIR GAUSSIAN FIT IN FIGURE 5.22.



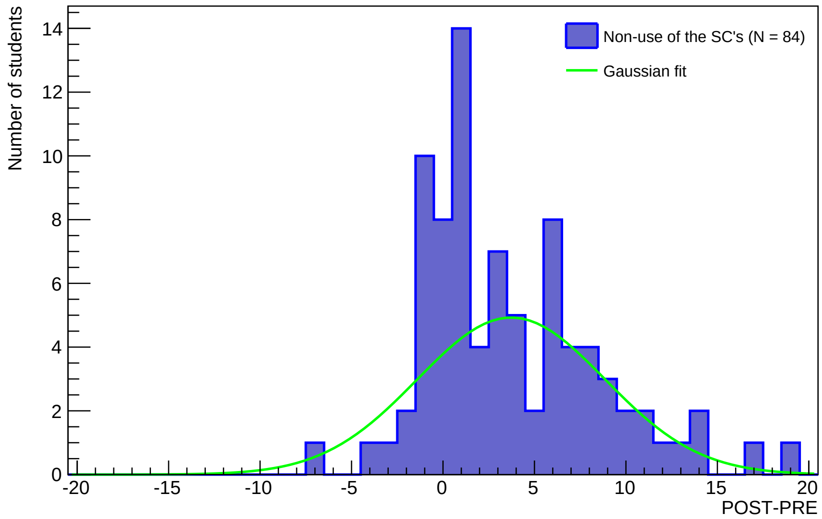


FIGURE 5.22 • DISTRIBUTIONS FOR GAINS AND THEIR GAUSSIAN FIT FOR THOSE WHO USED (TOP), AND NON-USED (BOTTOM) THE SCREENCASTS. FOR NUMERICAL RESULTS OF THE RESPECTIVE FITTED NORMALITY CONSTANTS, MEANS, AND SPREADS SEE TABLE 5.38.

Q. 3 *Which of the following three media did you find most educational?* Possible answers were *the book from Q. 1* ($N = 41$), *the screencasts from Q. 2* ($N = 12$), and *internet (youtube.com, computerfysik.dk,.....)* ($N = 55$).

With much anticipation we now turn to investigate whether we can detect if either the book, the screencasts or the internet has resulted in a statistically better $\overline{g}$. However, as mentioned in the beginning of this section neither the book or the screencasts treated any topic regarding classical mechanics. This would suggest that there should be no significant improvement in the mean gain for both of these media. If we do observe a significant result logic would dictate that it should be observed when comparing the internet against the book and screencasts as the internet contains a vast amount of information in classical mechanics and much other.

Constructing histograms of the gains with respect to each of the three media results in the distributions shown in figure 5.23. We compute the means and their respective errors to $\mu_{\text{Book}} = 4.20 \pm 0.75$, $\mu_{\text{SC}} = 4.17 \pm$

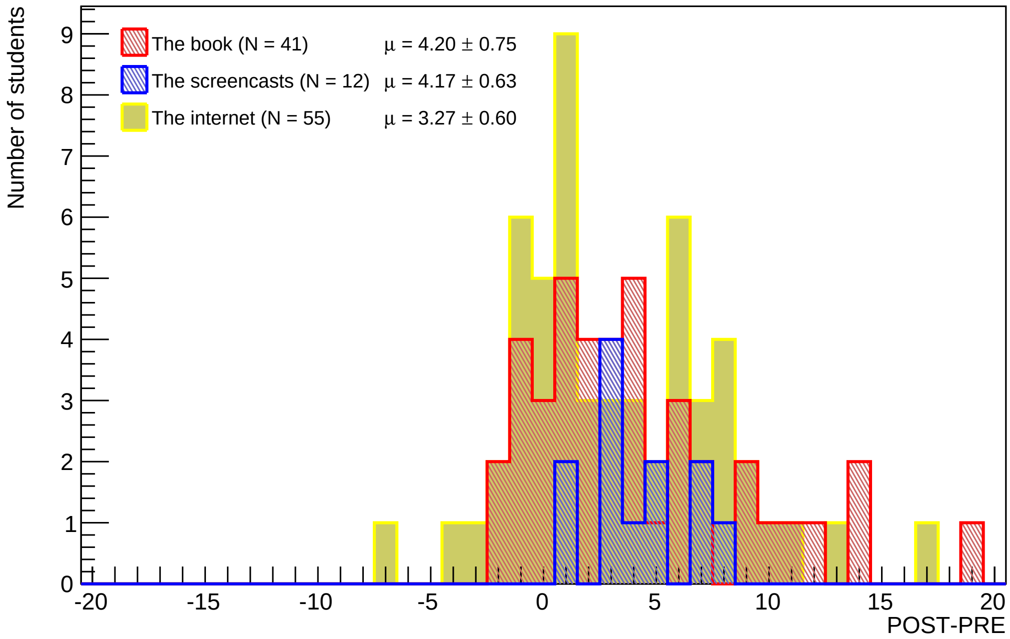


FIGURE 5.23 • DISTRIBUTIONS OF MEAN GAINS FOR THOSE WHO FOUND THE BOOK, SCREENCASTS, AND INTERNET MOST EDUCALTIONAL. ALL THREE DISTRIBUTIONS HAVE BEEN CONSTRUCTED USING THE RESPONSES OF THE 108 STUDENTS WHO PROVIDED AN INFORMED ANSWER TO QUESTION 3 OF THE LAB TEST AND ALSO TOOK BOTH THE FCI PRE- AND POST-TEST.

0.63, and $\mu_{\text{Internet}} = 3.27 \pm 0.60$.

We at once notice that the error intervals overlap which initially would suggest that the mean gains are statistically equal. Checking for normality we sadly find that in the case of the book and screencasts distributions (red and blue distributions, respectively, in figure 5.23) that we get poor χ^2-goodness of fit results. The test probabilities are both way above the cut off of 5.00% but this is most likely due the very large errors in the fitted normality constant, mean, and spread. See table 5.39 for details. Looking at the book and screencasts distributions in figure 5.23 we note that only 12 measured gains constitute the screencasts distribution and that only one of six bins have hits above 2; the book distribution comprises of 16 bins of which 7 have hits greater than 2. In addition, we note that this distribution somewhat resembles a poisson shifted to left.

Since we cannot assume that two of the three distributions are gaussian we therefore instead assume poissonian for these; as such we then have $\mu_{\text{Book}} = 4.20 \pm 0.32$, $\mu_{\text{SC}} = 4.17 \pm 0.59$, and still $\mu_{\text{Internet}} = 3.27 \pm 0.60$. We now see that the error intervals of μ_{Internet} and μ_{Book} does not overlap. However, to two SDOM's all three error intervals overlap which implies that the means of the distributions are most likely still equal.

We found that only the internet distribution (yellow in figure 5.23) could be assumed gaussian; as such we can then perform a Student's paired t-test to ascertain whether the mean gain is significantly greater than zero: We find that this is indeed the case with a probability $8.66 \cdot 10^{-5}$%. This is not surprising as a mean gain of zero is far from the error interval 3.27 ± 0.60. For a mean gain of zero to be in the error interval we see that we have to expand it with 5.45 SDOM's, i.e $0 \in 3.27 \pm 5.45 \cdot 0.60$. Therefore the the book and screencasts distributions most likely also have a mean gain significantly greater than zero.

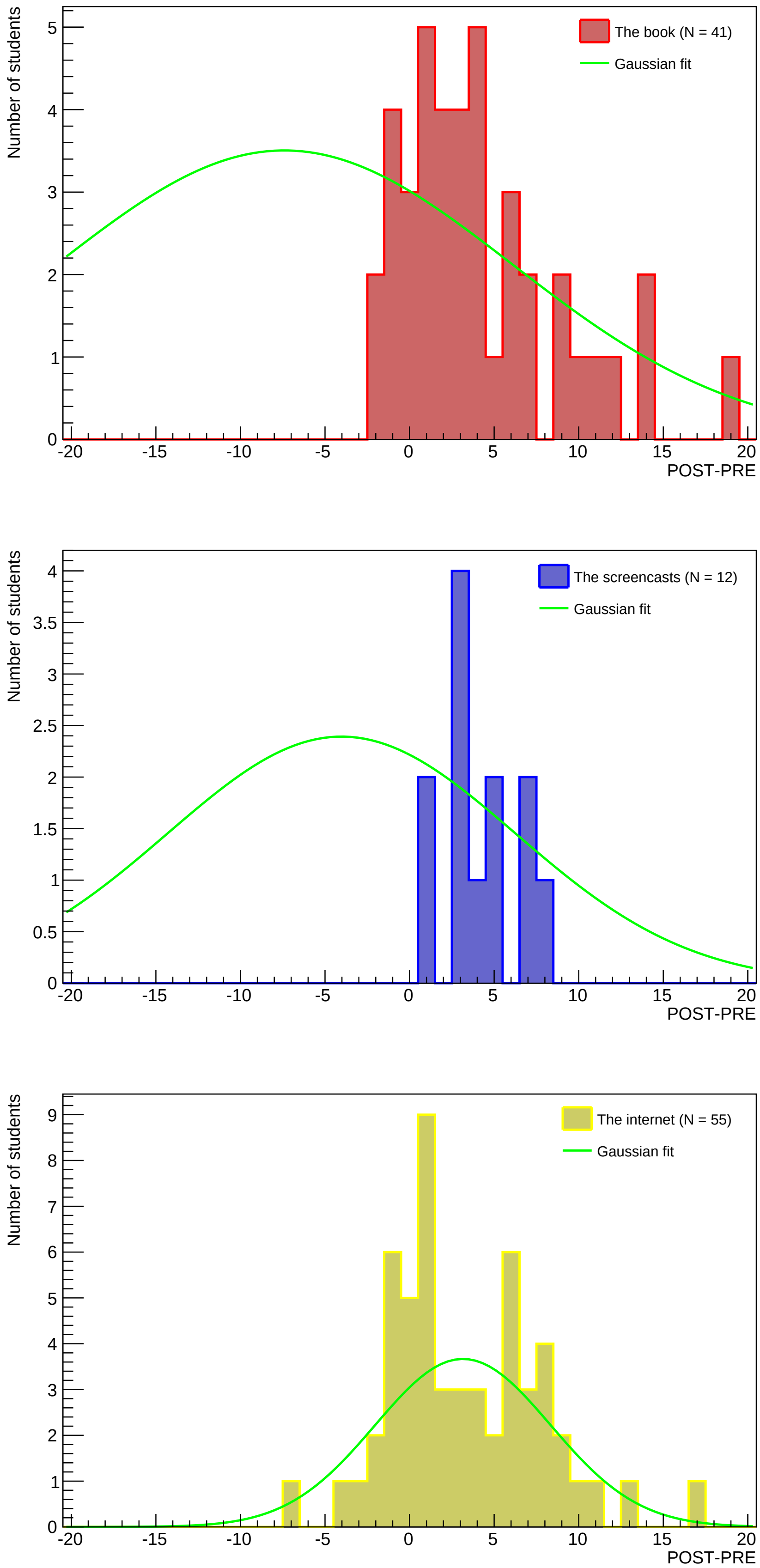


FIGURE 5.24 • DISTRIBUTIONS FOR GAINS AND THEIR GAUSSIAN FIT FOR THOSE WHO FOUND THE BOOK (TOP), THE SCREENCASTS (MIDDLE), AND THE INTERNET (BOTTOM) MOST EDUCATIONAL. FOR BOTH THE BOOK AND SCREENCASTS DISTRIBUTIONS THE RESULTING IS POOR; IN THE CASE OF THE BOOK IT MAY BE DUE TO THE DISTRIBUTION ACTUALLY BEING A POISSON SHIFTED TO THE LEFT FROM A VISUAL STANDPOINT; THE POOR FIT OF THE SCREENCASTS DISTRIBUTION IS MOST LIKELY DUE TO A LACK OF DATA. FOR NUMERICAL RESULTS OF THE RESPECTIVE FITTED NORMALITY CONSTANTS, MEANS, AND SPREADS SEE TABLE 5.39.

χ^2-GOODNESS OF FIT TEST	The book (red)	The screencasts (blue)	The internet (yellow)
Test statistic $(X_{\mathbf{GOF}})^2$	7.223734	6.303958	11.955246
Degrees of freedom	13	3	16
Normality constant	3.51 ± 3.94	2.39 ± 7.45	3.67 ± 0.77
Mean	$-7.41.80 \pm 38.35$	-4.03 ± 79.53	3.15 ± 0.97
Spread	13.47 ± 21.53	10.30 ± 47.59	5.19 ± 1.11
Probability	89.03%	56.70%	74.71%

STUDENT'S PAIRED t-TEST	The internet (yellow)
Test statistic T	5.365680
Effect size d	0.723508
λ_t	5.365680
Mean PRE $\mu_{\mathbf{PRE}}$	21.109091
Mean POST $\mu_{\mathbf{POST}}$	24.381818
Mean gain $\overline{g}$	3.272727
Degrees of freedom	54
Probability	$8.66 \cdot 10^{-5}$%
Power	99.98%

TABLE 5.39 • RESULTS OF THE PERFORMED χ^2-GOODNESS OF FIT TEST AND STUDENT'S PAIRED t-TEST WHEN CONSIDERING WHICH MEDIA THE STUDENTS FOUND MOST EDUCATIONAL. THE COLORS REFER TO THE DISTRIBUTION SHOWN IN FIGURE 5.23. ALL OF THE RESULTING TEST PROBABILITIES OF THE FITS ARE RELATIVELY HIGH. HOWEVER, THE FITTED PARAMETERS FOR THE BOOK AND SCREENCASTS DISTRIBUTIONS HAVE GREAT UNCERTAINTIES; AS SUCH WE CANNOT ASSUME THAT THESE ARE GAUSSIAN. FOR THE RESPECTIVE DISTRIBUTION WITH THEIR GAUSSIAN FITS SEE FIGURE 5.24. A STUDENT'S t-TEST SHOWS THAT $\overline{g} > 0$ FOR THOSE WHO FOUND THE INTERNET MOST EDUCATIONAL.

Q. 4 *Which of the following three media do you feel you benefit most from when learning something new?* Possible answers were *books* ($N = 58$), *screencasts/videos* ($N = 32$), and *Notes on the internet* ($N = 20$).

The last question in regards to media preference is properly the most interesting. This question asked the students which general media they found most beneficial when learning a new subject.

We construct histograms of the students' gain with respect to each of the general media, *books*, *screencasts/videos*, and *notes on the internet*. This results in the histograms shown in figure 5.25 where we find $\mu_{\text{Books}} = 3.71 \pm 0.62$, $\mu_{\text{Videos}} = 3.94 \pm 0.72$, and $\mu_{\text{Notes}} = 3.05 \pm 0.81$. Judging from a standpoint of simply means and errors we again see that their error intervals overlap; this would imply, based on what we have seen so far, that the means are statistically equal. In addition we observe that a mean gain of zero is again not contained in either of the error interval of the distributions; as such they are properly all significantly greater than zero, i.e all of the three media preferences results in a significant improvement of the FCI-test. To support these claims we check if the distributions can assumed gaussian in order to perform an ANOVA and Student's paired t-test.

Performing a χ^2-goodness of fit test they in all cases results in test probabilities greater than 20.00% with the lowest being 22.64% for the books distribution (red distribution in figure 5.25), and the highest being 91.55% for the screencasts/videos distribution (blue distribution in figure 5.25). The details of the tests are shown in table 5.40 and figure 5.26 shows the individual distributions and their gaussian fit. The somewhat low probability for the fit of the books distribution is most likely due to heavy fluctuations of the bins hits around the mean ($\mu_{\text{Books}} = 3.71 \pm 0.62$). We also note that this may be the reason for the relative

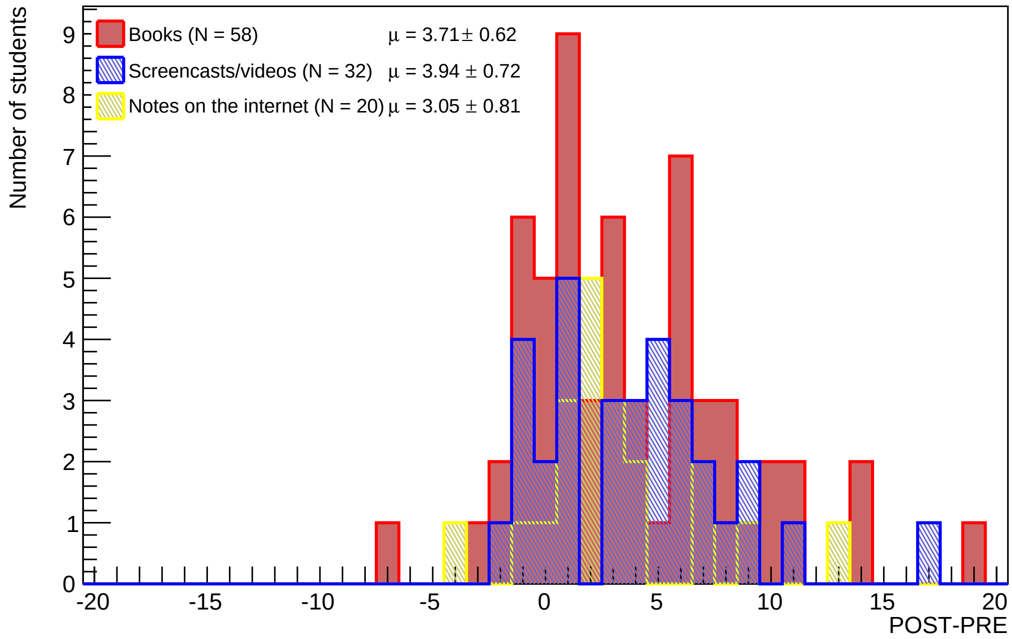


FIGURE 5.25 • DISTRIBUTIONS OF MEAN GAINS FOR THOSE WHO FOUND THE BOOKS, SCREENCASTS/VIDEOS, AND NOTES ON THE INTERNET MOST BENEFICIAL WHEN LEARNING A NEW TOPIC. ALL THREE DISTRIBUTIONS HAVE BEEN CONSTRUCTED USING THE RESPONSES OF THE 110 STUDENTS WHO PROVIDED AN INFORMED ANSWER TO QUESTION 4 OF THE LAB TEST AND ALSO TOOK BOTH THE FCI PRE- AND POST-TEST.

high uncertainty in the fitted mean and spread. The high test probabilities for the remaining distributions seems to be somewhat peculiar as also these are also subject to heavy fluctuations around the mean and in the tails. In the end we, however, accept the fit. Now we check for homoscedasticity and then perform an ANOVA to test whether the distribution means are statistically equal. A Levene's test reveals that we can assume homoscedasticity for the distribution as the test probability results in 71.57%. We then find via a performed ANOVA that the distributions are in fact statistically equal which supports our initial claim. In addition Student's paired t-test reveals that each media preference has resulted in a significant improvement, i.e $\overline{g} > 0$, of the FCI. In light of the results of the ANOVA and Student's paired t-tests supports the initial claim of equal significant increase in the mean gain regardless of preference of either general media. Numerical details of the Levene's test, ANOVA and Student's paired t-test are shown in table 5.40.

χ^2-GOODNESS OF FIT TEST

	Books (red)	Screencasts/videos (blue)	Notes on the internet (yellow)
Test statistic $(X_{\mathbf{GOF}})^2$	18.727299	4.543915	3.437547
Degrees of freedom	15	10	7
Normality constant	2.86 ± 0.78	3.45 ± 1.01	2.46 ± 1.04
Mean	3.16 ± 2.15	3.15 ± 0.97	4.00 ± 1.61
Spread	7.64 ± 3.85	4.07 ± 1.35	4.67 ± 2.85
Probability	22.64%	91.95%	84.18%

STUDENT'S PAIRED t-TEST

	Books (red)	Screencasts/videos (blue)	Notes on the internet (yellow)
Test statistic T	5.934227	5.367307	3.680552
Effect size d	0.779202	0.948815	0.822997
λ_t	5.934227	5.367307	3.680552
Mean PRE $\mu_{\mathbf{PRE}}$	21.051724	18.531250	23.950000
Mean POST $\mu_{\mathbf{POST}}$	24.758621	22.468750	27.000000
Mean gain $\overline{g}$	3.706897	3.937500	3.050000
Degrees of freedom	57	31	19
Probability	$9.21 \cdot 10^{-6}$%	$3.75 \cdot 10^{-4}$%	$7.94 \cdot 10^{-2}$%
Power	100.00%	99.98%	97.13%

LEVENE'S TEST

Test statistic W	0.334578
Degrees of freedom (df_1,df_2)	2, 107
Probability	71.57%

ANOVA

	ss_B	ss_W	$(s_P)^2$
Value	10.030486	2032.869471	18.998780
Degrees of freedom	2	107	
Test statistic F	0.764717		
Effect size f	0.069279		
λ_F	0.527954		
Overall mean $\overline{\mu}$	3.654545		
Probability	46.80%		
Power	9.07%		

TABLE 5.40 • RESULTS OF PERFORMED χ^2-GOODNESS OF FIT -, LEVENE'S -, STUDENT'S PAIRED t-TEST, AND ANOVA. THE COLORS REFER TO THOSE USED IN FIGURE 5.14. WE ARE SOMEWHAT SCEPTICAL OF THE GOODNESS OF FIT TESTS AS THERE ARE HEAVY FLUCTUATIONS IN THE DISTRIBUTIONS BIN HITS AROUND THE MEAN AND IN THE TAILS. HOWEVER, IF WE ACCEPT THE FIT RESULTS THE ANOVA REVEALS THAT THE DISTRIBUTIONS OF THE RESPECTIVE GENERAL MEDIA ARE EQUAL, AND STUDENT'S PAIRED t-TESTS RESULTS IN THEIR MEAN GAIN ALL ARE SIGNIFICANTLY GREATER THAN ZERO. SEE FIGURE 5.15 FOR THE RESPECTIVE DISTRIBUTIONS AND THEIR GAUSSIAN FIT.

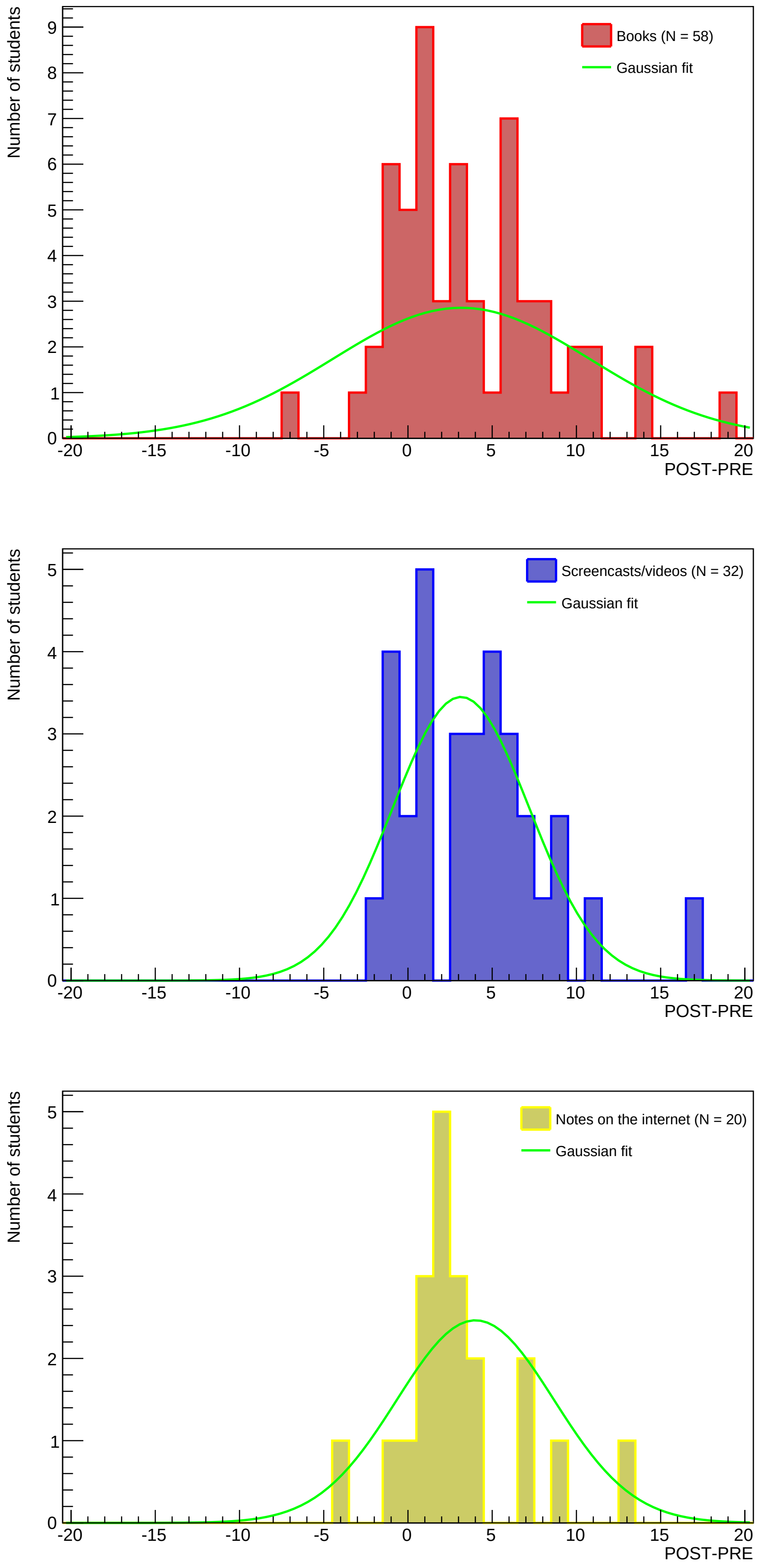


FIGURE 5.26 • DISTRIBUTIONS FOR GAINS AND THEIR GAUSSIAN FIT FOR THOSE WHO FOUND THE GENERAL MEDIA BOOKS (TOP), THE SCREENCASTS/VIDEOS (MIDDLE), AND NOTES ON THE INTERNET (BOTTOM) MOST BENEFICIAL WHEN LEARNING A NEW TOPIC. FOR NUMERICAL RESULTS OF THE RESPECTIVE FITTED NORMALITY CONSTANTS, MEANS, AND SPREADS SEE TABLE 5.40.

CHAPTER 6 Reflections

In the previous chapter we performed the data analysis. Now we will reflect on the results: We will highlight main results and discuss them. In addition we will also provide considered improvements.

LAB TEST: In the Lab Test the 136 students did quite well overall with a mean of correct answers of 14.92 ± 0.17 out of 19 questions. The least understood question by the students was question 17: 5.60% flagged it as not understood. The worst percentage of student responses to a question was found in question 3 and 4 to which 92.70% provided an answer. Therefore we were fairly certain that they knew what was being asked for in the respective questions and that a large percentage of them actually had provided an answer.

Only 6 of the 19 questions were answered correctly by less than 70.00% of the students; of these 6 questions three were on data analysis (Q.17, Q.19, and Q.21) and the remaining on the use of Matlab (Q.5, Q.8, and Q.10). Analysing the responses from these selected questions we found that learning Maple parallel to Matlab may have given rise to confusion as 28.89% provided the Maple procedure, to define a new variable, as the correct answer. In addition they did learn that the spread is a measure for the uncertainty of a single measurement, however, they did not seem to know how or when to compute SDOM. This we based on the fact that 80.88% answered correctly the use of the spread (Q.15) but wrongly on the computation of SDOM (Q.19). In light of this it was therefore found confounding that 54.41% answered wrongly on the use of SDOM, but correctly on the use of the error propagation formula, as the latter clearly was a more difficult topic. No deeper interpretation could be found other than the students may simply have guessed more favorably on the use of the error propagation formula than on the computation of SDOM or that they simply have learned to use the error propagation formula. If the latter explanation was the case it would imply that the book actually had greatly contributed in regards to the formulas use as it was only mentioned in the book; however, the computation of SDOM should, logically, have been answered correctly by a much larger percentage than we here observed (15.79%). More attention to the topic SDOM and its meaning in connection to data analysis in future Mek1 courses is therefore suggested. We found it unclear whether the use of the error propagation formula was supposed to be mandatory in Mek1; it was not specifically mentioned in the course description nor included in the slides from the teachings of basic data analysis by Børge Svane

	MATLAB GROUP			DATA ANALYSIS GROUP		
	Question 5	**Question 8**	**Question 10**	**Question 17**	**Question 19**	**Question 21**
Possibility 1	**93**	**91**	27	**72**	**21**	34
Possibility 2	39	9	**92**	46	33	**85**
Possibility 3	3	35	10	10	79	12
Not responded	1	0	3	3	2	2
Not understood	0	1	4	5	1	3
Total	136	136	136	136	136	136

TABLE 6.1 • THIS IS TABLE 5.2 SHOWN HERE FOR CONVENIENCE. THE BOLD NUMBERS INDICATE THE CORRECT POSSIBILITIES.

Nielsen. However, if it was supposed to be mandatory we suggest that this topic also needs more attention despite that 64.89% answered correctly on its use in the Lab Test as the students could simply have guessed more favourably. One concept we, on the other hand, found that the students had relatively fully grasped during Mek1 was the mean: 89.71% answered correctly both in question 14 and 18, i.e they knew that the mean was the best estimate of the true mean of a quantity and how to compute it. Only 1.47% failed to answer both questions correctly.

— RECEPTION OF TEACHING MATERIALS: From the Lab Test we obtained the individual students' media preference through questions 1 - 4. Recall, that these were:

Q. 1 *Have you at any point used the book "Grundlæggende Matlab og dataanalyse", which can be found on Absalon?* Possible answers were *yes, no* and *never heard of it.*

Q. 2 *Have you seen any of the screencasts, which also can be found on Absalon?* Possible answers were *yes, no* and *never heard of them.*

Q. 3 *Which of the following three media did you find most educational?* Possible answers were *the book from Q. 1, the screencasts from Q. 2,* and *internet (youtube.com, computerfysik.dk,.....).*

Q. 4 *Which of the following three media do you feel you benefit most from when learning something new?* Possible answers were *books, screencasts/videos,* and *Notes on the internet.*

The numerical distribution from table 5.6 are shown in table 6.2 which shows how the students responded. Using the principles set theory we investigated the validity of the responses of media preference: We observed some small inconsistencies due to missing responses and illogical responses. To compensate for the missing responses rigorously the statistical field of *missing data* could be applied. This is, however, out of the scope of this thesis. We argued that a solution would have been to let logic dictate some of the missing responses, but this would have resulted in a tainted data pool which was not desired; as such we had no other alternative than to exclude students who failed to provide an answer in regards to their media preference. We do, however, not know to which degree it may have affected the findings. We also found 2 illogical responses which was due to two students who had claimed the book most educational but had not used it during Mek1. The impact that these two illogical responses may have had was considered to negligible.

Two attempts to put together a focus group to collect qualitative data were not successful. The reason hereto was most certainly not due to a lack of trying as both attempts were advertised at least one week in advanced. Therefore the students' responses of media preference were the only source of information available to determine the reception of the book and screencasts. Hereto we found that of the screencasts and book the latter was without doubt the most used teaching material as 66 out of 135 have used it; only 39 of 139 claimed to have used the screencasts. In addition 43 of 66 have used the book and voted it to be most educational; this was considered to be quite good seen in the perspective that only 10 out of 39 users of the screencasts claimed them to be most educational. Overall, the internet was deemed as the most educational media by most as of 67 out of 126 students claimed exactly this, and it was mostly those who claimed that they had not used or never had heard of either the book or screencasts.

	Question 1	**Question 2**	**Question 3**	**Question 4**
Possibility 1	66	39	45	66
Possibility 2	39	74	14	34
Possibility 3	30	23	67	26
Informed answers	135	136	126	126

TABLE 6.2 • THIS IS TABLE 5.6 SHOWN HERE FOR CONVENIENCE.

Both the book and screencasts were classified as secondary literature; with this in mind we argued that the book had been received very positively by the students due to its before mentioned frequent use. The screencasts were on the other hand not that well received due to the fact that 74 of 136 students had not use them. The reason hereto we argued was most likely due to their lengths: Three of eleven screencasts had a duration under nine minutes, whereas the remaining ranged from 17 to 29.5 minutes. If the screencasts are to be implemented as a viable teaching material in future courses of Mek1 their duration has to be limited significantly. This could be accomplished by simply treating fewer topics in each of the screencasts. In doing so automatically gives rise to more refined and precise titling of them; as a consequence the students would perhaps be able to more specifically identify which screencasts to review in connection to solve their problem at hand.

When investigating for correlation between media preference and the NTC of the Lab Test we in every case found no correlation. This implies that there was not enough evidence indicating that the NTC of the Lab Test was dependent on the use, or a lack hereof, of either the book, the screencasts, or the internet. We also found that the means of total correct answers of the Lab Test for each media were not significantly different from each other. This turned out also to be the case for the preference of a general media, i.e books, screencasts/videos, and notes on the internet.

When considering the claimed high usage of both the book and the internet it is believed that there may be a significant difference in the means of correctly answered questions for these media; however, since the Lab Test was only given as a POST-test, i.e given at the end of Mek1, the teachings had perhaps evened these differences out. On this basis we hypothesize that if the Lab Test was given before any teaching and again after we would be in a far better position to detect any significant difference regarding which media results in better test scores.

— IMPROVEMENT OF MEDIA PREFERENCE QUESTIONS: We argued that there was room for improvement of the questions regarding media preference of the Lab Test. In both question 1 and 2 the three possibilities were *yes*, *no*, or *never heard of it/them*. The initial intent was that the three possibilities were to be mutually independent: If a student had used the media *yes* should be checked, if he had not <u>but knew of its existence</u> *no* should be checked, and lastly if he did not know of its existence *never heard it* should be checked. However, we argued from a logical standpoint that *no* and *never heard of it* could have been slightly correlated: If a student had not heard of the media, it would logically imply that he had not used it; as such there was some bias in form of the individual student's interpretation of the three possibilities! In light of this there may have been some students who have claimed that they did not use the media even when did not know of its existence. However, we had no way of determining the number of students who answered in this fashion. Therefore *no* should have been rephrased to *no, but I have heard of it*. Another solution to this dependency issue could be to only have the possibilities *yes* and *no*.

If we accept the hypothesis that the dependency issue was not in play, i.e the students claimed to not have used either the book or screencasts but knew of their existence, it implies that 30 and 23 of 135 and 136 students, respectively, had not heard of the these. This we found to be quite confounding as both the book and screencasts were advertised on three separate occasions of which one was the Matlab instructor. In addition, they were also advertised in two lectures held by Ian Bearden in connection to clicker questions. Thus the degree of which book and screencasts have been introduced was most definitely not low.

We also hinted that the possibility *the internet* in question 3 may have been too broad compared to the remaining media, i.e *the book* and *the screencasts.The internet* as a possibility could here have been replaced with *neither* which perhaps could have given a more precise overall picture in regards to the effect both the book and screencasts may have had on the test scores. It could also be that *notes on the internet* in question 4 then should have been replaced by simply *the internet*; however, this would then present some issues as the possibilities would not be mutually exclusive: Screencasts/videos are today found mainly on the internet. This issue was exactly the reason why the third possibility was chosen to be *notes on the internet*.

DHD: The mean of the 175 students who took the DHD was found to 10.10 ± 0.23 which we argued was quite a fair result as the DHD compared to the Lab Test was a much more difficult test. The difference in the difficulty of the tests was mainly argued to the language in which the test was written; the DHD was written in English and the Lab Test in Danish. Therefore, as the DHD was given before the beginning of Mek1 the students then only relied on their secondary education in data analysis and the English language; hence technical terms, within the regime of data analysis were extremely limited and could have been lost in translation. The limited knowledge of technical terms was especially clear when the students were asked what the mode of a set of measurements was. Here only 57.10% understood what was being asked for.

Comparing the students of Mek1 who were given the Lab Test, i.e the year 2013, with those who also were given it in 2011 and 2012, we found that the mean of correctly answered questions were statistically equal. Therefore the student body of Mek1 in 2013 had no advantage or disadvantage in regards to their knowledge of data analysis when using the DHD as a measure hereof.

Some interesting results were found when analysing the responses of the 120 students who were given both the DHD and Lab Test: Independence tests revealed correlation between the NTC of the DHD and whether the students claimed later to (1) have not used, or never have heard of the screencasts; (2) have found the book, or screencasts most educational. In the case of (1) we argued that it may be due to the possibilities *no* and *never heard of them* are not fully independent as previously mentioned; on the other hand it may also be possible that the group of students who claimed later that they never had heard of the screencasts which comprised of 21 students, were more passive, i.e they did only what they were told to do and nothing more, as 57.14% of this group also claimed to never had heard of the book. In the case of (2) we suggested that the correlation may have arisen due to mixed effects as some students later have used the book and found the screencasts most educational, and vice versa. A precise interpretation of the observed correlation between the future use of the book, or screencasts and the NTC of the DHD was ultimately not found. This correlation does, however, not imply that future use of either media have resulted in a significantly different mean of correct answers of the DHD when looking at the errors of the means of the respective media.

Two questions from the DHD were also present in the Lab Test. The topic of these questions were the use of the error propagation formula and precision of a physical quantity. Using these repeated questions we could investigate further: We found that the teachings of Mek1, i.e when not considering media preference, had significantly contributed to the students' ability to use the error propagation formula; this was argued to be somewhat unexpected as the formula was not to found in Børge Svane Nielsen's slides he used in the lectures on basic data analysis. However, as we have found significant impact in the formula's use it must have been implemented in the teaching of Mek1 at some point. Most interestingly we also found that for the 40 students, who voted the book most educational, 39 claimed to also have used it, and the chance of no impact in the formula's use was only $9.67 \cdot 10^{-2}$%. In addition we also found that the use of the screencasts resulted in a significantly positive impact which was confounding as the screencasts did not mention the use of the formula; however, 19 of the 31 students, who had used the screencasts, had also used the book. We argued that these findings pointed to the book having been the most effective media in teaching the students to use the formula. In regards to precision of a physical quantity we also found that the teachings of Mek1 had significantly contributed to the their knowledge hereof. Most interestingly we also found that the students who claimed they had not used or never had heard of it did not significantly improve their ability to answer correctly before and after Mek1; however its use did result in a significant improvement.

FCI: The FCI tested the students in common misconceptions of Newtonian mechanics. We found, when using the responses from the 129 students who took both the PRE- and POST-test, that the teachings of Mek1 very significantly had improved the mean of correct answers: $\mu_{\mathrm{PRE}} = 20.60 \pm 0.59$ and $\mu_{\mathrm{POST}} = 24.42 \pm 0.44$. Not surprisingly, when using the responses from the 118 students who were given the PRE-, POST-test, and the Lab Test, we also found that the mean gain with respect to each media, i.e $\overline{g}$, was in every case significantly greater than zero but equal. This implies that media preference, general or limited to the book and screencasts, did not significantly improve the mean gain of the FCI. In the case of the use, or lack hereof, of either the book or screencasts this result was expected as neither treated topics of Newtonian mechanics.

POWER OF THE TESTS: It was mentioned in chapter 4 that when performing statistical tests the power should be at least 80%. Of all the performed statistical tests in which we computed the achieved power, i.e post hoc power analysis, the Student's paired t-tests were the only in which the desired power was reached. Only in a few cases the same was true for the McNemar-independence test. On the other hand every performed G-test of independence and ANOVA were far from the desired percentage; especially the ANOVA where the highest achieved power was 19.57%. For further details see table 6.3.

The low power of these tests was directly connected to the number of observations N and the computed effect size. Recall, the relations between the non-centrality parameters, effect sizes, and sample size in the case of the G-test of independence and ANOVA:

$$\lambda_{\chi^2} = N \cdot w^2 \quad \text{and} \quad \lambda_F = N \cdot f^2, \tag{6.1}$$

where w is the effect size of the G-test, and f that of the ANOVA.

Let us assume that all of the performed tests should result in the rejection of their respective null hypothesis, i.e the resulting test probability are all less than $\alpha = 0.05$. We can then compute the required sample size needed, using the current computed effect sizes, to achieve 80% power: The non-centrality parameter needed to achieve said power is in a G-test of independence with one degree of freedom $\lambda_{\chi^2,80\%} \approx 7.85$, and in ANOVA $\lambda_{F,80\%} \approx 12.99$. We have here only considered the cases of tests in which we observed the lowest power. Thus we have that $N = 7.85/\,(0.007078)^2 > 156{,}000$ and $N = 12.99/\,(0.003358)^2 > 1{,}150{,}000$; as such *if* the null hypothesis should be rejected at $\alpha = 0.05$ with the current computed effect sizes we would need more than $156{,}000$ samples in the G-test of independence and $1{,}150{,}000$! Such a sample would simply not be possible to obtain.

However, if we are to believe the claim, put forth by John M. Hoenig and Dennis M. Heisey in their article *The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis*, then post hoc power analysis is ultimately flawed: According to them there exists a direct relation between the test probability (p-value) and observed power, i.e the test probability directly implies a certain power:

> *"Observed power can never fulfill the goals of its advocates because the observed significance level of a test ('p value') also determines the observed power; for any test the observed power is a* $1:1$ *function of the p value. [...] Because of the one-to.one relationship between p values and observed power, nonsignificant p values always correspond to low observed power. [Hoenig, 2001]".*

On this basis the performed post hoc power analysis of each test was utterly useless. But a more in-depth analysis of the power of the tests is out of the scope of this thesis.

	G-**test**	**ANOVA**	**McNemar-test**	**Student's paired** t**-test**
Minimum power [%]	5.08	5.02	32.30	92.21
— *effect size*	0.007078	0.003358	0.195283	0.755237
Maximum power [%]	57.63	19.57	99.36	100.00
— *effect size*	0.047839	0.111554	0.406202	–

TABLE 6.3 • TABLE OF MINIMUM AND MAXIMUM ACHIEVED POWER IN PERCENT OF THE PERFORMED STATISTICAL TESTS AND THE COMPUTED EFFECT SIZES. SINCE MORE THAN ONE STUDENT'S PAIRED t-TEST RESULTED IN 100.00% POWER WE CANNOT SIMPLY STATE ONE EFFECT SIZE IN THIS CASE.

ONLINE ENVIRONMENT FOR TEST TAKING AND ANALYSIS: When the students were given the DHD and FCI each was handed a hard copy of the test and an answering sheet. It was not allowed to provide any answer to the individual questions on the hard copy; the answering sheet was to be used for this. In the case of the Lab Test the students were to supply their answers directly on the test.

The initial step to process the data, i.e the students' responses, was to manually enter them into spreadsheets, where after the spreadsheets was exported to a csv-file which then could be imported into the desired analysis software; in this thesis ROOT was used. Not only is this a very time consuming process but it also introduces the factor of human errors. Due to the extreme amount of time required, to process and analyse the test data, this would most likely not be done immediately after the duration of a course; as such the lecturers involved may not be able to interpret the results of the various statistical tests quite as effective as memory of what was done in the course would not be recent.

A suggestion to solve this issue is the construction of an online environment. This environment would consist of two sub environments; one where the students provide their personal data and responses to the given tests, and another where lecturers can perform relevant statistical tests by simple selections and one click.

— SUB ENVIRONMENT FOR THE STUDENTS: The first step in this environment is the for students to supply a test identification number. This would be a number generated when the lecturer has declared a new event. This will be described in greater details below. If the test identification number is valid they would then proceed to provide their personal information. In order to uniquely identify the students, which is of the utmost importance in paired statistical tests, the submission of their date of birth would not be ample. To accomplish unique identification they have to provide (1) their personal identification number, called a cpr-number, issued by the Danish state at their birth which is on the form *ddmmyy-xxxx*; (2) their student identification number issued at the beginning of their studies at KU; and (3) their first and last name.

It may be enough for the students to provide their cpr-number as this uniquely identifies them to the Danish state; however, this numbering system currently face critiques as it can easily be misused [Jyllands-Posten, 2014]. Therefore a newer more secure system may be implemented in the future. Due to this radical change the student identification number could serve as a backup along with their first and last name.

When the personal data has been entered the students would then be able to supply their answers to the respective questions of the test. This would simply be an electronic version of the before mentioned answering sheet. The possible response options would be grouped radio buttons; this automatically ensures that only one of the possible options can be submitted as a response. At the end they click a simple button which then submits their responses. When clicked a procedure checking if all questions have been answered should be performed: This eliminates the possibility of no responses. However, the option for marking a question as one they do not understand should still be present, but whether they still need to provide a response has not yet been determined as the implications are not clear at present. To mark a question as not understood a checkbox would be present at each question. Also whether the formulation of each questions has to be implemented electronic is at present not decided.

— SUB ENVIRONMENT FOR THE LECTURERS: In order for the students to take the tests the lecturer has to create an event, mark it as active, and of course select which test he wishes to give the students. When completed a test identification number is generated; the format of this number is not yet determined. As mentioned, the identification number would be entered by the students in order to direct them to the relevant answering sheet. In addition the event would automatically be given a systematic name, e.g *DHD_Mek1_2013/14_10092013*, to differentiate the given tests from each other. The format of the name of a created event is, however, not final.

After the students have completed the test the lecturer mark the event as inactive, thereby ensuring no further responses can be submitted. In the case where the lecturer forgets to mark the event as inactive the

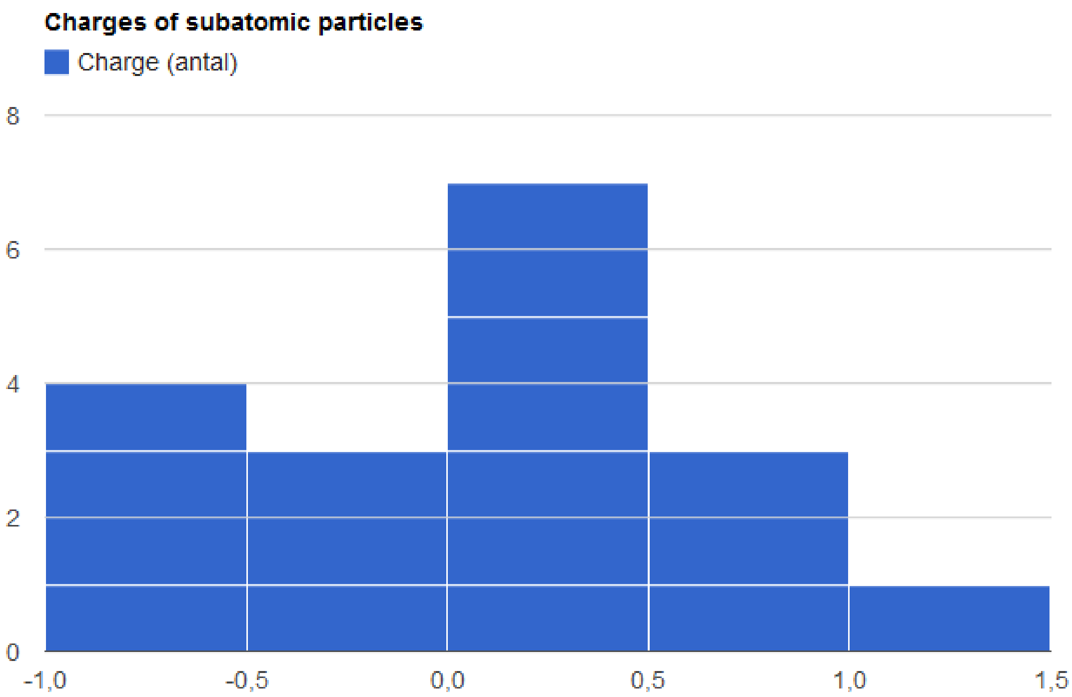


FIGURE 6.1 • ONLINE HISTOGRAM CREATED WITH GOOGLE'S FREE TO USE API *Google Charts*. THIS IS FROM THE GOOGLE CHARTS' GALLERY OF EXAMPLES[GOOGLE, 2014].

possibility of submitting a time stamp by which the responses have to be submitted as a contingency could be implemented.

As soon as the test data is available, and the event has been marked as inactive, the lecturer will be able to immediately perform the desired statistical tests, and generate relevant histograms via the environment. The possibilities for processing the data are of course limited by the functionality of the environment. At present the statistical tests relevant would be McNemar's -, G-test of independence, ANOVA, and Student's paired t-test. To generate the histograms, and display them with this environment, the use of Google's free API *Google Charts* is a good candidate. An example of a histogram created using this API is shown in figure 6.1. In addition it should also be possible for the lecturer to export the test data to a csv file if he wishes to analyze it differently.

— FURTHER CONSIDERATIONS AND APPLICABILITY: In order for this online environment to be realized a lot of additional consideration and planning lies ahead; it is at present only in the "napkin stage". For instance, say, the pdf's needed in the relevant statistical tests are by default a part of the used programming language's mathematical library; will these pdf's then be sufficiently accurate in comparison to a more professional statistical software as, say, R or SAS? Also how do we maximize security of the database where the students personal data and responses would be stored? One solution could be to encrypt the information with the Security Hash Algorithm 1 (SHA1) which is a built-in function of most popular server-side scripting languages; however, would this be sufficient or are there more secure possibilities available which should be considered?

If realized such an environment would not only be restricted to the physics department: Other faculties could use it if they too were to give some sort of multiple choice test. Actually, the test need not be a multiple choice; the environment could easily be extended to support tests where a numerical or letter based value was to be submitted as a result.

In addition it would also be quite possible for the environment to be used in investigations such the one conducted in this thesis. One example could be if a lecturer has decided to introduce a textbook as secondary literature. Then via a test, given both prior and after the course, he would be interested in determining if the use of the secondary textbook have resulted in a more significant mean gain relative to the remaining. However, such a feature is an advanced addendum which requires vast knowledge of databases and their structure but it would not be impossible.

Chapter 7
Data analysis in the Danish secondary school

As we have from the responses of the Lab Test that the students had difficulty in regards to some of the concepts regarding data analysis it seems warranted to investigate the possibility of implementing them in Danish secondary schools.

At present there are in Denmark four main secondary educations: HTX, STX, HF, and HHX.

- *"Højere Teknisk Eksamen"*, translated "Higher Technical Exam" short HTX, it is a relatively new secondary education which was established in 1982. This education, in comparison to the remaining, focuses on the interaction between the theory of the natural sciences and technological solutions [Damberg, 2011, page 101-103]. In short: HTX is mainly for those who intent to study for some degree in engineering due to its practical nature.
- *"Studenter Eksamen"*, translated "Student Exam" short STX, is the oldest as it was formed in 1850 on the basis of an older institution called *"Latin skolen"* [Damberg, 2011, page 23]. STX focuses, in comparison to HTX, mainly on theory in every course. In addition it offers the most versatile number of courses.
- *"Højere Forberedelseseksamen"* No description of the subject physics. Therefore we do not here consider HF.
- *"Højere HandelsEksamen"* has its perspective set on commerce and political sciences, and is therefore not an education of which the natural sciences, i.e physics, chemistry, biology and so fourth, are part of the curriculum [Damberg, 2011, page 97]. Therefore we do not here consider HHX.

From this we see that the relevant secondary educations are HTX and STX: First we investigate what the term data analysis entails here in order to provide a suggested solution.

7.1 Definition of data analysis in secondary schools

Each of the subjects taught in the Danish secondary school has a certain standard which is defined by the Danish government. A guideline for this standard is provided in the departmental order[1] of the subject; hereto the teachers shall refer if needed. The contents of such a departmental order are mainly (1) core and additional subjects, (2) purpose and end competencies for the students, (3) used didactical principles, (4) examination and evaluation. From this point on we will denote *departmental order* simply by DO.

From the section about core subjects in the DO's for physics level C, B, and, A we can shed some light on what data analysis, at present, is in the secondary school. In addition the written examinations in physics level A (no written exams in physics level C and B [Retsinformation, 2013d], [Retsinformation, 2013c]) are available via the website for the ministry of education (Danish: "Undervisningsministeriet") [Undervisningsministeriet, 2014].

[1] In Danish: *"læreplan"*

Of the four main secondary educations, i.e STX, HTX, HF, and HHX, we only consider STX and HTX; HHX are automatically disqualified due to physics not being taught, and HF does not have a DO for physics at any level [Retsinformation, 2013a].

When reading the DO's for physics levels C, B, and A for STX and HTX no explicit mentioning of data analysis can be found at all. Besides the description of the core subjects and the overall wording of the DO's, however, gives the impression that much is left to personal interpretation by the teacher; as such a few selected statements could easily be interpreted as some sort of data analysis has to be part of the mandatory curriculum: According to the competencies mentioned in the DO's the students should:

> HTX PHYSICS LEVEL A AND B:
> *"kunne planlægge og gennemføre enkle fysiske eksperimenter og analysere simple fysiske problemstillinger, opstille løsningsmodeller og udføre et større eksperimentelt arbejde, hvori indgår målinger, resultatbehandlinger og vurderinger."* [1] [Retsinformation, 2013c]
>
> STX PHYSICS LEVEL A AND B:
> *"kunne behandle eksperimentelle data med henblik på at diskutere matematiske sammenhænge mellem fysiske størrelser."* [2] *[Retsinformation, 2013d]*
>
> STX PHYSICS LEVEL C:
> *"kunne beskrive og udføre enkle kvalitative og kvantitative fysiske eksperimenter, herunder opstille og falsificere enkle hypoteser."* [3] *[Retsinformation, 2013d]*

However, what is meant by data analysis is neither mentioned nor clear as they only refer to the students being able to *"process experimental data"*.

No clear definition of what data analysis covers in the secondary school seems to be present. In order to get some clarity in this regard we therefore only have the published written exam problems as a measure. For HTX and STX we have looked at physics exam problems between the school years 2007/08 and 2012/13 which revolves around some given data.

We have observed that in connection to STX there is often one type of problem in which some visual analysis are to be applied, and another type which is on transformation of data and fitting. In the HTX exam problems actual measurements, stored in a downloadable file, form the basis of the problem; as such much more data is given relative to those in STX. The reason hereto is perhaps that the use of computers are more emphasized in HTX. Typically, in connection to HTX we found that one type of problem was on transformation and plotting of data; another on data interpretation.

Of the above mentioned types of problems we classify them as:

- STX TYPE 0: Transformation of data and fitting.
- STX TYPE 1: Visual analysis.

[1] English translation: *"be able to plan and conduct simple experiments and analyze physical problems; put fourth solution models and conduct larger experimental work wherein measurements, processing of results, and assessments are part of."*

[2] English translation: *"be able to process experimental data and discuss mathematical dependencies between physical quantities."*

[3] English translation: *"be able to describe and conduct simple qualitative and quantitative physics experiments including set fourth and falsifying simple hypothesis."*

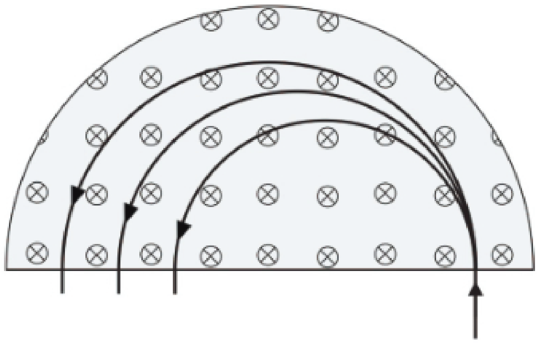


FIGURE 7.1 • THIS IS THE SHOWN ILLUSTRATION FROM PROBLEM 5 OF THE EXAM PROBLEM SET FOR STX PHYSICS A IN MAY 2013 WHICH SHOWS TRAJECTORIES OF DEFLECTED LITHIUM IONS IN A HOMOGENOUS MAGNETIC FIELD WHICH POINTS PERPENDICULARLY INTO THE PAGE RELATIVE TO THE PLANE OF MOTION OF THE IONS. THE PROBLEM SET CAN BE FOUND ON THE WEBSITE OF THE DANISH MINISTRY OF EDUCATION [UNDERVISNINGSMINISTERIET, 2013].

- HTX TYPE 0: Transformation and plotting of data.
- HTX TYPE 1: Data interpretation.

These types will be illustrated through selected problems from the written exams.

7.1.1 STX type 0 and 1

We have selected the written exam in physics A given in May 2013 to illustrate a STX type 0 and 1 problem. Type 0 is found in problem five and type 1 in both problem one and six; problem six has, however, been selected.

The title of problem five was *"Massespektrograf"* which translates to *Mass spectrometer*. The students were shown an illustration of three deflection trajectories of lithium ions due to a homogeneous magnetic field pointing perpendicularly into the page relative to the plane of the ions' motion. The illustration can be seen in figure 7.1. From this the students were first to determine whether the deflected ions were positively or negatively charged. Hereto they were to apply the Lorentz force law, which states that the magnetic force $\vec{F}_{\text{mag}}$ on a particle, with charge Q and velocity $\vec{v}$, moving in a magnetic field $\vec{B}$ is $\vec{F}_{\text{mag}} = Q\left(\vec{v}\times\vec{B}\right)$. Simply using the right hand rule for vector cross products we see that the ions have to be positively charged as $\vec{F}_{\text{mag}}$ always points toward the center of curvature.

In the second and last question of this problem the strength of the homogeneous magnetic field was to be determined on the basis of six given data pairs. The data pairs are shown in table 7.1. In addition it is reported that the ions have a speed of $v = 2.87\cdot 10^6\,\text{m/s}$ as they enter the magnetic field; also that they have singular charge, i.e $Q = -e = 1.602\cdot 10^{-19}\,\text{C}$.

As the magnetic field is homogeneous the ions will, whilst moving in it, undergo uniform circular motion; this implies

$$\left|\vec{F}_{\text{mag}}\right| = \left|\vec{F}_{\text{uniform}}\right| \Rightarrow QvB\sin(\theta) = m\frac{v^2}{R} \Rightarrow p/Q = BR, \tag{7.1}$$

as $\sin(\theta) = 1$ since $\vec{v}$ is perpendicular to $\vec{B}$. Therefore the ions' momentum per charge as a function of the radius of curvature is a straight line with slope B.

Mass [u]	5.012	6.015	7.016	8.022	9.027	10.035
Radius [m]	0.435	0.524	0.601	0.683	0.783	0.876

TABLE 7.1 • SIX PAIRS OF DATA, GIVEN IN PROBLEM FIVE OF THE STX PHYSICS A MAY 2013 EXAM PROBLEM SET, FOR THE MASS OF DEFLECTED LITHIUM IONS AND THEIR RESPECTIVE RADIUS OF CURVATURE DUE TO MOVEMENT IN A HOMOGENEOUS MAGNETIC FIELD.

Plotting p/e as a function of R we see that they lie almost perfectly on a straight line. The fit model we will use is $BR+k$ where k is the intersection; it turns out that $B = 0.341 \pm 0.007\,\mathrm{T}$ and $k = 0.002 \pm 0.005\,\mathrm{Tm}$. The data and fit are shown in figure 7.2.

A STX type 1 problem involves analysis from a purely visual standpoint. An example hereof can be found in problem six of the STX written exam in physics A from May 2013. The title of the problem is *"Roning"* which translates to *Rowing*. The topics the students here were to apply were basic kinematics and power.

The students were informed that a rower finished a $s = 2000.00\,\mathrm{m}$ race in 6 min and 57.82 s, i.e $t = 417.82\,\mathrm{s}$. From this the average speed v of the rower was first to be computed which simply is $v = s/t = 4.79\,\mathrm{m/s}$. Then they were told that a boat accelerates from rest to a speed of 19.90 km/h, i.e $v_1 = 5.44\,\mathrm{m/s}$, on a distance of $s_1 = 90.00\,\mathrm{m}$. From this the average acceleration of the boat was to be computed. Here the intent was most likely for the students to apply the formula $s_1 = \left((v_1)^2 - (v_{\mathrm{rest}})^2\right)/(2\cdot a)$, where a is the average acceleration: Using this then $a = (v_1)^2/(2s_1) = 0.16\,\mathrm{m/s^2}$.

The third, and last question, of problem six was where the use of visual analysis was to be applied. A graph was shown which showed the horizontal force on the handle of a rowing machine by a rower as a function of the horizontal position of his hands in one pull. The graph is shown in figure 7.3. From this the students were then to determine the average power of the work done on the handle by the rower in one pull.

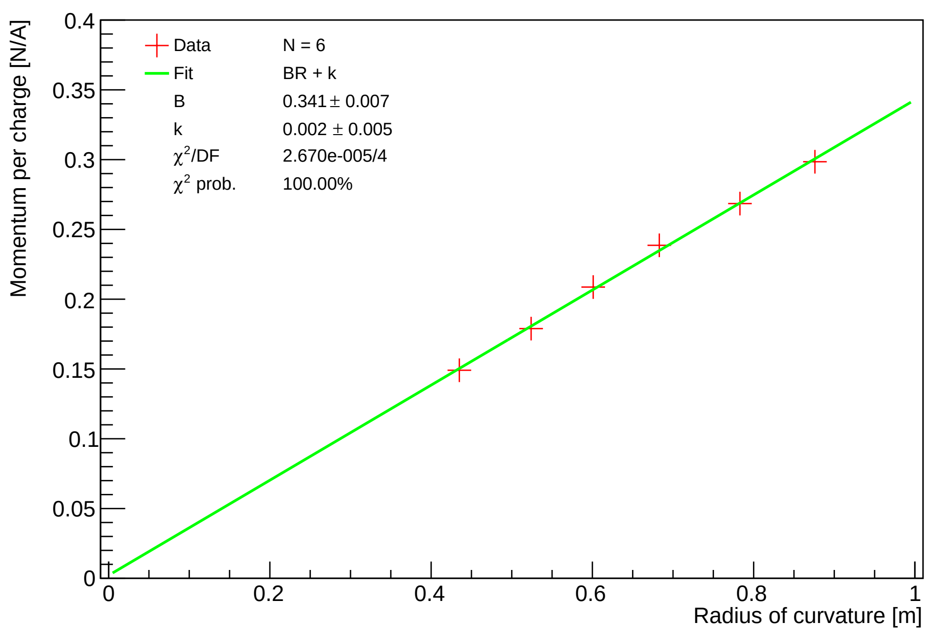


FIGURE 7.2 • PLOTTING THE IONS MOMENTUM PER CHARGE, I.E p/e, AS A FUNCTION OF THE RADIUS OF CURVATURE R RESULTS IN A STRAIGHT WHERE ITS SLOPE IS EQUAL TO THE STRENGTH OF THE HOMOGENEOUS MAGNETIC FIELD B. NOTE THAT WE ALSO FIT FOR THE INTERSECTION k WHICH IMPLIES THE FULL MODEL USED IN THIS FIT IS $BR+k$.

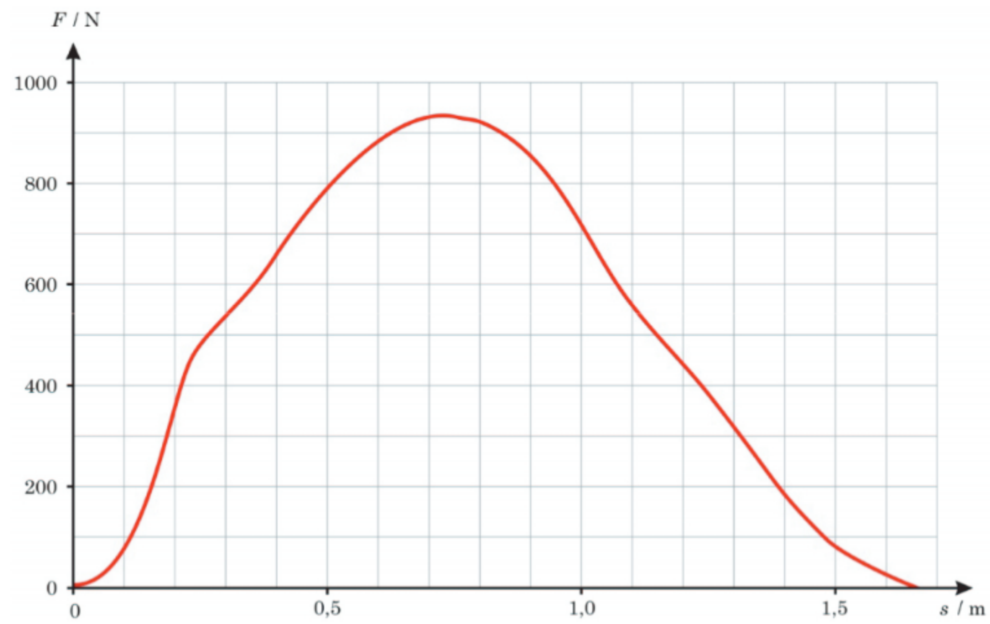


FIGURE 7.3 • THIS GRAPH WAS SHOWN IN CONNECTION TO THE THIRD AND LAST QUESTION OF PROBLEM SIX FROM THE WRITTIN EXAM IN PHYSICS A [UNDERVISNINGSMINISTERIET, 2013] GIVEN ON MAY 2013. IT SHOWED THE HORIZONTAL FORCE ON THE HANDLE OF A ROWING MACHINE BY A ROWER AS A FUNCTION OF THE HORIZONTAL POSITION OF HIS HANDS IN ONE PULL. THE INTENT WAS THAT THE STUDENTS SHOULD ESTIMATE THE AREA UNDER THE CURVE IN ORDER TO DETERMINE THE WORK DONE. WE ESTIMATE THAT THE WORK DONE IS $W = 82\,\text{SQUARES} \cdot 10\,\text{J/SQUARE} = 820\,\text{J}$.

Since the graph was shown on a grid and no mathematical expression was provided it is believed that the intent was for them to determine an estimate. For the students to determine such an estimate they were to (1) use the definition of power P, i.e $P = W/t$, where W is the work done within the time t; and (2) use the definition of work $W = F \cdot s$, i.e force multiplied by distance; in this context the area under the curve equals the work done in one pull. The grid aids the students in estimating this area. We estimate that the number of squares under the shown graph is 82 squares; as one square equals $0.1\,\text{m} \cdot 100.0\,\text{N} = 10\,\text{J}$ the work done is therefore $W = 82\,\text{squares} \cdot 10\,\text{J/square} = 820\,\text{J}$ in one pull.

When the work done by the rower in one pull has been established the students were then to apply the definition of power, i.e $P = W/t$. The time t, the average duration one pull lasted, can be determined as it was given that the rower carried out 32 pulls per minute, i.e one pull took, on average, $t = 60\,\text{s}/32$. By our estimate the average power is then $P = W/t = 437.33\,\text{W}$.

7.1.2 HTX type 0 and 1

The selected problem to illustrate both a HTX type 0 and 1 problem is problem four of the December 2011 physics A written exam; its title is *"Hurtig kompression"* which translates to *Rapid compression.* This problem revolves around actual data from an experimental setup illustrated in figure 7.4: A cylinder, containing some amount of carbon dioxide, with a piston connected to a lever. The lever is at some time pressed down quickly. From this information, and the given measurements, the students were to answer the following

a) Determine the amount of carbon dioxide before the gas is compressed.

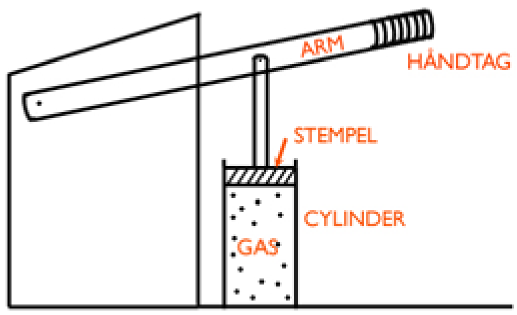


FIGURE 7.4 • THIS IS FIGURE 4.3 FROM THE EXAM PROBLEM SET FOR HTX PHYSICS A IN DECEMBER 2011 WHICH SHOWS THE EXPERIMENTAL SETUP FROM WHICH THE DATA HAS BEEN PRODUCED. THE PROBLEM SET CAN BE FOUND ON THE WEBSITE OF THE DANISH MINISTRY OF EDUCATION [UNDERVISNINGSMINISTERIET, 2011].

b) Graph the amount of substance of carbon dioxide as a function of time and assess whether the cylinder is tightly sealed during the compression.

c) Graph the pressure as a function of the volume during the compression.

d) Show, using the measurements, that the compression approximately can be considered adiabatic, and determine the adiabatic constant $\gamma = c_p/c_v$.

The measurements given were comprised of 161 values of time t, volume $V(t)$, pressure $p(t)$, and temperature $T(t)$ in the units s, cm^3, kPa, and K. Measurements not reported according to the SI standard are of course transformed to this.

Assuming that the gas is ideal, we use the following relation: $pV = nRT$, where p is the pressure, V the volume of the gas, n the amount of gas, R the ideal gas constant, and T the temperature. Therefore if we plot pV/R as a function of T we should observe a straight line with a slope of n. Such a plot is shown to the left in figure 7.5. It is quite clear from the plot that after the gas has reached about 350 K the trend breaks down.

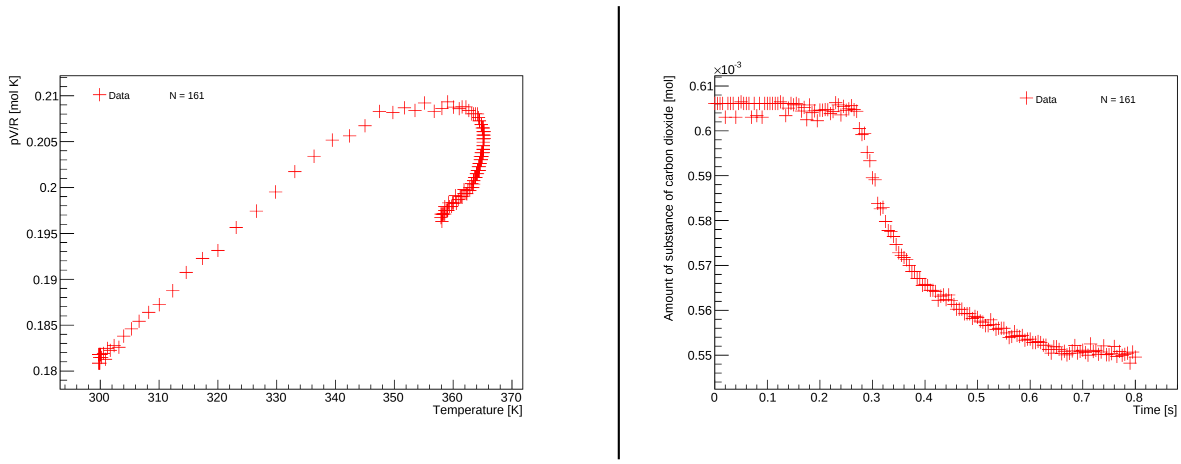


FIGURE 7.5 • LEFT: CALCULATED pV/R AS A FUNCTION OF THE TEMPERATURE T USING THE ENTIRE DATA SET ($N = 161$). HERE WE CLEARLY SEE THAT THE TREND IS A STRAIGHT LINE BUT IT BREAKS DOWN WHEN THE GAS REACHES ABOUT 350 K. THIS COULD MEAN THAT THE CYLINDER LEAKS. RIGHT: THE AMOUNT OF SUBSTANCE OF GAS $n = PV/(RT)$ AS A FUNCTION OF TIME. n SEEMINGLY DROPS EXPONENTINALLY AFTER APPROXIMATELY 0.26 SECONDS WHICH ALSO SUGGESTS THAT THE CYLINDER IS LEAKING.

Before n can be determined we need to investigate further why the trend toward a straight line breaks. Therefore we plot $n = pV/(RT)$ as a function of the time t which is shown to the right in figure 7.5. We see that in the interval $[0\,\mathrm{s}, 0.26\,\mathrm{s}]$ n remains constant whereas in the remaining interval n drops seemingly exponential; as we have no knowledge of the experimental equipment the conclusion is that cylinder is leaking gas during the compression.

With this established we can determine n, however, we can only use data from the interval $[0\,\mathrm{s}, 0.26\,\mathrm{s}]$ (53 data points) as n here is approximately the same. Fitting a straight line, i.e the model $pV/R = nT + b$, for the slope n and intersection b we find that $n = (600 \pm 6) \cdot 10^{-6}\,\mathrm{mol}$ and $b = 0.002\,\mathrm{mol \cdot K} \pm 0.001\mathrm{mol \cdot K}$. Data with the fit is shown in figure 7.6.

For adiabatic compression of an ideal gas the relation $pV^{\gamma} = k$ must hold, where γ is the dimensionless adiabatic constant, and k is some constant. Plotting the pressure as a function of the volume using the trimmed data, i.e the 53 data points from the time interval $[0\,\mathrm{s}, 0.26\,\mathrm{s}]$, results in a visually nice adiabat: Fitting for the adiabatic constant and k we find from the trimmed data that $\gamma = 1.220 \pm 0.002$ and $k = 0.127 \pm 0.003$. If we instead use the entire data set $\gamma = 1.120 \pm 0.008$ and $k = 0.42 \pm 0.04$. Thus trimming the data have indeed brought the fit result closer to the, in the problem set, given value 1.29. Data with the fit is shown in figure 7.7. It may also be that some students instead would choose to use the equivalent model $\ln(p) = -\gamma \ln(V) + \ln(k)$; this, however, does not impact the fitted value of γ by much as in this case $\gamma = 1.226 \pm 0.003$.

7.1.3 The definition

In the case of the presented problem from the HTX written exam it would seem that the primary intent was first that the students should be able to infer that the cylinder leaked gas as the straight line tendency clearly breaks after reaching a certain temperature T ($\approx 350\,\mathrm{K}$) when plotting pV/R vs. T. Secondly, when

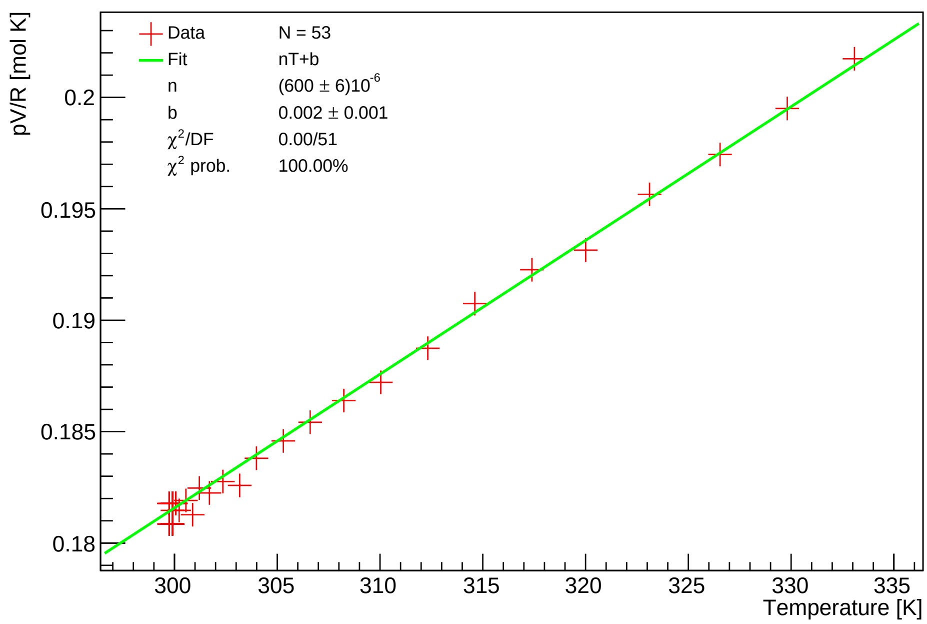


FIGURE 7.6 • DATA WITH THE PERFORMED FIT WHEN USING THE TRIMMED DATA SET, I.E ONLY DATA FROM THE INTERVAL $[0\,\mathrm{s}, 0.26\,\mathrm{s}]$ WHERE THERE APPEARS TO BE INSIGNIFICANT LEAKAGE, TO DETERMINE THE AMOUNT OF SUBSTANCE OF GAS IN THE CYLINDER.

the students were asked to plot $n = pV/(RT)$ vs. the time t was serve as support for the leakage. Thirdly, by fitting either a power curve to a plot of p vs. V or a straight line to $\ln(p)$ vs. $\ln(V)$ the adiabatic constant γ could be determined.

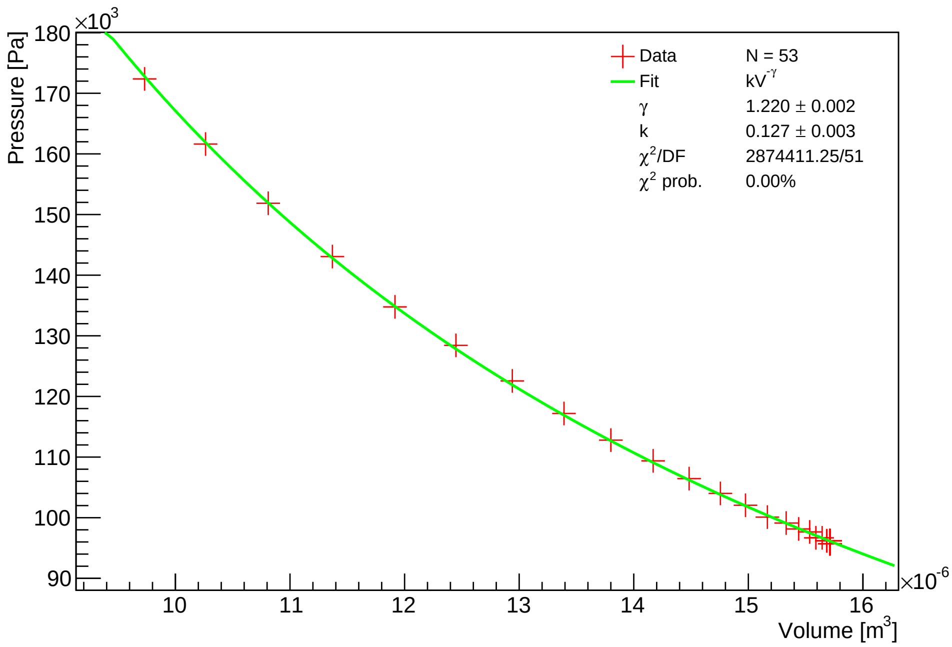


FIGURE 7.7 • DATA WITH THE PERFORMED FIT USING THE TRIMMED DATA SET. WE SEE WE GET A FAIRLY NICE ADIABAT.

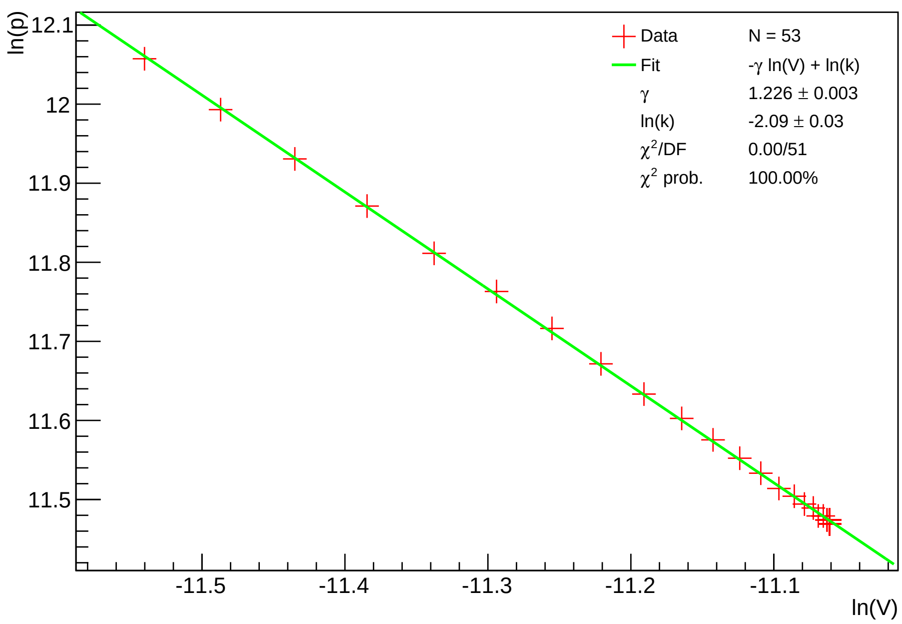


FIGURE 7.8 • DATA WITH THE PERFORMED FIT WHEN USING THE TRIMMED DATA SET, I.E ONLY DATA FROM THE INTERVAL $[0\,\text{s}, 0.26\,\text{s}]$ WHERE THERE APPEARS TO BE INSIGNIFICANT LEAKAGE, TO DETERMINE THE AMOUNT OF SUBSTANCE OF GAS IN THE CYLINDER.

It is, however, not clear how the authors of this problem intended the students to determine the amount of substance of gas n before the compression: Was it to be done by simply read off the value, fit a straight line in the interval where the trend is linear, or to the whole of the data set? It could, however, simply be that this type of problem exists to determine the students' ability to think analytical and independently, i.e show resourcefulness.

From this problem we therefore conclude that data analysis in HTX is defined as the students' ability to (1) infer the quality of the experiment or a lack hereof; (2) basic use of CAS-software to transform the data and make plots; and (3) basic fitting.

For the type of problems in connection to STX the intent is seemingly more clear. These problems are not as "open" as them from HTX and the amount of data is quite small, i.e six pairs of measurements. Therefore the handling of data is here not as prioritized as in HTX but focuses more on abstract theory; this is probable also the case, in regards of the degree to which the students should be able to use CAS-software, in data handling. On the other hand the small amount of data given here may be an indication of some basic straight line fitting algorithm is part of data analysis in STX; most likely the least squares method. We note that the students are asked to plot the data, however, since they are indirectly asked to fit some straight line it is believed that a plot is also expected.

The common denominator between HTX and STX would be the transformation, plotting, and fitting of data.

7.2 Implementation of error analysis

In seems apparent from the selected types of written exam problems that error analysis is not part of the physics curriculum in either STX or HTX. This is considered to be a major flaw as error analysis directly investigates whether the data supports the theory. In addition this field is a very big part of being a physicist; as such it should at least be part of the curriculum in a very basic form.

The topics that covers "very basic" are here meant to be (1) mean, spread, and standard deviation of the mean; (2) histograms; (3) the connection of (1) and (2) to the gaussian distribution; and, lastly, (4) errors and their use in physics experiments. As a basis the underlying statistical theory is not to be prioritized but simply how errors are used and what can be inferred by them.

In this section we will provide some suggestions as to how (1), (2), (3), and (4) could be taught in both HTX and STX. In addition discuss whether error analysis should be part of the written exams, i.e be a core subject.

7.2.1 "Error analysis bootcamp"

To introduce error analysis in HTX and STX can either be done over a number of discontinuous periods or could actually be done over some number of full school days if possible. Regardless, this could be branded as *"Error analysis bootcamp"*. The goal of this bootcamp is to teach the before mentioned topics (1), (2), (3), and (4); also how the calculations can be done in a CAS-software. The idea is that when the bootcamp has been completed the students have to apply what they have learned in future reports of experiments, i.e compute and report the uncertainties of the measured quantities.

We can of course not deny that a certain level of mathematical acquaintance and confidence are required: If the introduction of error analysis is introduced when, say, the students are in the early stages of their first year the math could end up being too difficult and thus be a barrier for actually absorbing the intended. Therefore a suitable time for the bootcamp could perhaps in the early stages of their second year; however, we hereby automatically exclude it from being held during physics level C as this has a duration of one year - typically their first. The best way to determine when to place the bootcamp would of course simply be to do it and observe where the students encounter problems; at present an educated guess would be in the early stages of their second year.

(1) Mean, spread, and standard deviation of the mean

Basically, these three concepts revolve around their mathematical definitions. We will denote the mean as $\overline{x}$, spread as s_x, and standard deviation of the mean as $s_{\overline{x}}$. First part is to get the students familiar with simply computing the values and their meaning are later specified.

Firstly, formulas for $\overline{x}$, s_x, and $s_{\overline{x}}$ are to be presented. Through a series of examples shown on the blackboard the students then have a sort of recipe on how to compute these values. The data sets used in the examples could either be prepared or improvised. By improvising the data sets opens the possibility of engaging some of the students. For instance, five or six students could be selected to measure the length of the teacher's desk with a folding ruler, and then write their measurements on separate pieces of paper, i.e independent measurements. If another data set is desired another selection of students could, in the same manner, measure the length of one of the legs of the teacher's, or their own, chair.

When the examples have been worked on the blackboard the students are then to try on their own by working through some prepared problems or, if time allows it, split them into to groups and ask them to go an collect some length measurements of a quantity of their own choosing.

In the end of this part the goal is that the students have learned that $\overline{x}$, s_x, and $s_{\overline{x}}$ are simply values, and how they are computed.

(2) Histograms

The introduction of histograms is also thought should be done in a way which engages the students. It could be done in the following way: Say that there are N students; on N pieces of paper the teacher has, before the start of the class, written an integer between, say, 1 and 10. These pieces of paper are then put in a bowl, or similar container, from which the students are asked to draw on piece of paper. When every student has drawn a number they are, by a show of hands, to indicate which number they drew. From the resulting data set of N integers a histogram is constructed on the blackboard. This data set could be used later to introduce the gaussian if it of course was chosen to be approximately normally distributed.

When the histogram has been constructed on the blackboard the students are then to form groups in which they are to discuss, without calculating anything yet, what they think the mean of the data set is approximately using only the constructed histogram. After some minutes they share the individual groups' results with the rest of the class. Without the teacher commenting they are then to actually calculate the mean of the data set. The point is, assuming that the data set is chosen to be gaussian-like, they should realize that the mean is in the neighbourhood of where the bin hits are greatest.

Following the students are then to work in groups to figure out the interpretation of spread by (1) computing the spread of the data set, (2) again but adding the integer 15, (3) adding 30, and lastly (4) adding both 15 and 30. From their results they are to discuss what they think the trend of the spread would be if more relatively large integers were to be added to the data set. Following the exact same thought they are then following the same procedure as before to compute the spread when integers close to the mean are added. The point is that they learn that the spread of a data set increases when measurements are far from the mean and vice versa.

(3) The gaussian distribution

The point of introducing the gaussian distribution is not such that the students should be able to understand the underlying mathematics, but simply that in basic error analysis we assume that errors are normally distributed. In addition it is also to demonstrate how the computed mean and spread for a data set is connected to this distribution.

To illustrate this connection the teacher could use the data from the before mentioned bowl experiment; draw a gaussian with the calculated mean and spread and overlay on the constructed histogram electronically of course. In addition it should also be illustrated how the shape of the gaussian depends on the spread and how the horizontal offset depends on the mean. Figure 3.10 and 3.11 in appendix A are good examples of such illustrations.

(4) Errors and their use in physics experiments

In this step the students are to learn how the mean, spread and standard deviation of the mean are used in physics experiments. Given that the DO for both educations states that they are to be "able to conduct simple physics experiments" it is a reasonable assumption that the mathematical models involved are either zero or first degree polynomials. The previously treated data analysis problem types, i.e STX - and HTX type 0 and 1, supports this claim.

On this basis the relevant expressions for the spread of a data set are

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2}, \quad \text{and} \quad \sigma_y = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}\left(y_i - ax_i - b\right)^2}, \tag{7.2}$$

where a and b are the slope and intercept of the best straight line y. In this context it should also be shown how to determine the values of a and b. These expressions can be seen in the box *"Fitting af ret linje"* on page 47 in the book in appendix A. The point here to be made is that in connection to physics experiments the mean is the best estimate of the true value and the error of this is the standard deviation of the mean, i.e $\overline{x} \pm s_{\overline{x}}$; also that the spread is the average error in an individual measurement.

Here the training in using a selected CAS-software should take place: Computing the mean, spread, and standard deviation of the mean, plotting of data with error bars, and fitting the relevant mathematical model to the data. The degree of programming skills required depends on the selected CAS-software, but they should be able load the data and knowing a small set of functions after instruction. Maple for instance has built-in functions for all of the above; some even in Danish via the widely known package called *Gympakken*.

Some examples of simple experiments which could be used for the students to use the learned concepts should be conducted in connection to the boot camp. The subjects to be used in the experiments of course depend on which subjects the students have been taught prior to the bootcamp. As we have suggested that the bootcamp should be held minimum in the early stages of their second year the subject of thermodynamics is most likely completed; as such we give an example of an experiment hereof.

A simple experiment within the field of thermodynamics could be to determine the power of an electric kettle P simply by heating up water with mass m. The data will be collected electronically, e.g via Vernier's LoggerPro using their sensor which measures temperature. LoggerPro will collect data pairs consisting of time t and temperature of the water $T(t)$.

Plotting the energy that the water has received $Q = cm\Delta T$, where c is the specific heat capacity of water, vs. the time t the data pairs should according to the theory follow a straight line whose slope is P as $P = Q/t \Rightarrow Q = Pt$. From the data the students would of course compute σ_Q and create an errorbar plot wherein the data with errors and the fitted line. To keep things simple it should be assumed that the time measurements have negligible errors, i.e $\sigma_t \approx 0\,\text{s}$. Surely, the line will not go through some of the error interval of the points which the students most likely will interpret as being bad. This could then be explained by looking at the percentiles of the gaussian distribution; the point being that approximately 68% of the points will most likely lie outside the fitted line as the fundamental assumption is that the errors are normally distributed.

As a wrap up of the experimental part the students should write a report of this experiment using their newly gained knowledge in simple error analysis.

Error analysis is not only restricted to physics but is applicable in every field of natural science; as such the error analysis bootcamp could be implemented as a joint venture between the commonly taught fields, i.e the periods reserved for every field are used for the bootcamp.

7.2.2 Error analysis in the written exams

The question remains whether the error analysis skills should be part of the written exams of HTX and STX. If implementing error analysis as a core subject the answer is most definitely yes: Error analysis is a big part of the field of physics.

This does, however, raise some questions: Should there then be problems which specifically are on error analysis? If so, should the students then assume, unless otherwise stated, that error analysis is to be part of the problem solution and arguments? Answers to these questions are at present not known but this could be investigated by creating and conducting a bootcamp on a selected secondary school. This is, however, out of the scope of this thesis but could form the basis for a future investigation.

An example of a problem which explicitly could test the students in some of the basics of error analysis could be the following:

Five students have measured the acceleration due to gravity g. Their results were as follows:

Student	1	2	3	4	5
g [m/s^2]	9.7	9.9	10.0	10.1	10.3

(a) *Compute the mean, spread, standard deviation of the mean.*

(b) *What is the uncertainty in a single measurement given the data?*

(c) *What is the best estimate of g and its uncertainty given the data?*

On the other hand, if error analysis was to be used indirectly, i.e applied in problems where data is given, the data sets should not be too large as this would most likely result in the students consuming a large fraction of the allotted time by simply calculating the necessary values. This would most likely be the case of an HTX type 0 problem but not as much in a STX type 0 problem.

Chapter 8 Conclusions

In the course of this thesis we found for the three given tests, i.e the self authored Lab Test, the DHD, and the FCI that the students did overall quite well. The mean test scores, i.e the mean of correct answers, are summarized in table 8.1. Even though the mean test score for the DHD is only 10.10 ± 0.23 out of 23 questions in total we concluded that it constituted a good score: This was due to (1) the fact that the DHD was written in English; (2) had very limited knowledge of data analysis and English terms used in connection hereof. The mean test score of the DHD was compared to those in 2012 and 2011: Statistically, the students body of Mek1 had neither any advantage or disadvantage in regards to knowledge of data analysis in the beginning of Mek1.

After the course curriculum of Mek1 had been taught we found, on the basis of the 136 students who were given the Lab test, that 80.88% did learn the meaning of the concept of the spread of a data set but did not know how to compute SDOM; as such we concluded that more attention to this is warranted in future courses. Further attention to the meaning of the arithmetic mean and its computation is on the other hand not needed: 89.71% had learned that the mean is the best estimate of a quantities true value and how to compute it given some measurements.

We also found that learning to use two different CAS-softwares most likely have confused the students: When asked how to define a variable in Matlab roughly a third submitted the procedure for doing so but in Maple! The remaining, however, submitted the correct Matlab procedure. A suggested solution to this problem was in the end not found.

Luckily, we did conclude that Mek1 had significantly improved the students' skills in Newtonian mechanics, i.e the students have learned something: The FCI was used as a measure hereof and we found a gain of $g = 3.82 \pm 1.01$ which was significantly greater than zero. However, there was no significant difference in gain if the students preferred any media; as such the main conclusion is that preferring books, screencasts/videos, or notes on the internet in general does not yield a higher gain in FCI.

The question whether use of a certain general media, here books, screencasts/videos, or notes on the internet, have been answered: There was no significant difference in mean test scores of either the given tests which lead to the simple conclusion that there in this case is none. Also the number of correct and wrong answers of either tests where not found to be dependent on which general media the students preferred. This was also mostly the case when investigating differences in test scores for students who found either the book *"Grundlæggende Matlab og dataanalyse"*, the screencasts, or the internet most educational.

When we looked at responses before and after the Mek1 we found that use of the book had been a statistically significant contributor in regards to the topic precision of a physical quantity. However, in regards to the use of the error propagation formula, we found that the teachings of Mek1 here was a significant

	Lab Test	**DHD**	**PRE FCI**	**POST FCI**
Mean test score	14.92 ± 0.17	10.10 ± 0.23	20.60 ± 0.59	24.42 ± 0.44

Table 8.1 • Shows the mean of correct answers in every given test during Mek1 in the school year 2013/14.

	G-test	ANOVA	McNemar-test	Student's paired t-test
Minimum power [%]	5.08	5.02	32.30	92.21
— *effect size*	0.007078	0.003358	0.195283	0.755237
Maximum power [%]	57.63	19.57	99.36	100.00
— *effect size*	0.047839	0.111554	0.406202	–

TABLE 8.2 • THIS IS TABLE 6.3 REPRODUCED.

contributor in the students learning this topic even though we found no mentioning of the formula's use herein: Whether the students had used or not used either the book or screencasts resulted in a statistically improvement in answering correctly but 19 of the 31 screencasts users also used the book. In addition those who found the book most educational, of which almost all also have used it, the paired independence test was most significant with $9.67 \cdot 10^{-2}$%. The computed odds ratios were here in every single test less than 1, i.e more students answered correctly after Mek1 relative to before regardless of the outcome of the paired independence test. In the end we concluded that the data pointed to the book actually being the most effective media, combined with the teachings, in learning the use of the formula.

The tools we used to draw the above conclusions were the performed relevant statistical tests and, where possible, the computed errors. To gain some insight as to the reliability of the results of these test their power was computed post hoc. We found that a large fraction of the performed tests resulted in low statistical power; especially the G-test of independence and ANOVA. However, according to [Hoenig, 2001] post hoc power analysis is basically flawed. Therefore it was in the end somewhat unclear whether the computed powers are a reliable measure for the quality of the tests. The minimum and maximum achieved powers of all performed tests are summarized in table 8.2.

In regards to the quality of the data, i.e the responses from the three tests, was in general good. Therefore we had no reason to suspect that the data contained misleading aspects; especially not when looking at the response and understood rates, which in every test all were above 80.00% with one exception in question 3 of the DHD where the understood rate was only 57.10% out of 175 students. The reason for this low rate was simply that only a few students ever had heard of the concept of mode which question 3 of the DHD asked. However, the media preference questions which asked the students if they had used, not used or never heard of either the book or screencasts, i.e Q.1 and Q.2, were argued to perhaps be somewhat misleading: Logically, if the students actually had never heard of either media the present formulation of the response possibilities implied that both *no* and *never heard of it* would be valid responses. The extent, if any, to which this logical flaw reflects on the inferred results is not known. Therefore, if future use of the media preference questions are to be used we suggest they should be revised to compensate for this logical flaw as a minimum.

The quality and reception of the created teaching materials was found to be quite good in the case of the book but somewhat poor for the screencasts. This conclusion was based on the fact that 48.89% of 135 students had used the book and of these 65.15% voted it most educational; only 28.68% of 136 students had used the screencasts and of these a mere 25.64% found them most educational. An argument was made that the low usage and most educational percentages of the screencasts were most likely due to eight of eleven screencasts having a duration above 17 minutes. Therefore we concluded if the screencasts in the future are to be an effective resource they seriously need trimming and revision. The book on the other hand had been received very well; as such the effort to shape it into a viable resource in terms of content and layout certainly paid off. In the future it can easily be expanded and perhaps even serve as main literature on introductory Matlab and data analysis.

Lastly, we investigated what the term data analysis entails in the perspective of the relevant Danish secondary school educations HTX and STX. The departmental orders in connection to the subject of physics were found to be quite inconclusive in regard to this definition. Therefore we reviewed problems in the written exams from both educations to gain some insight. Here we concluded that in problems where some data are given the common denominators were transformation, plotting, and fitting of the data. The sizes of the data sets given varied greatly: In STX only six data pairs were given whereas a downloadable file containing relatively much more data was given in HTX.

We argued, that since error analysis is a big part of field of physics, it is not satisfactory that this subject seemingly is not a part of the educations. A suggested solution to remedy this was a rough sketch on a course, we branded “Error analysis bootcamp”. The point of the bootcamp is to teach the topics (1) mean, spread, and standard deviation of the mean; (2) histograms; (3) the connection of (1) and (2) to the normal distribution; and (4) errors and their use in physics experiments. In addition we provided some examples as inspiration to introduce the various concepts in such a way secondary school students would understand them. Whether the suggested inspirations actually would be understandable for the students is not known. This could, however, be basis for a future investigation wherein both the book and screencasts perhaps could be adapted accordingly to serve as teachings materials for secondary school students.

References

[FCI, 2014] (2014). Fci and mbt | american modeling teachers association. www: http://modelinginstruction.org/researchers/evaluation-instruments/fci-and-mbt/.

[Agresti, 2007] Agresti, A. (2007). *Introduction to Categorical Data Analysis.* John Wiley & Sons, 2 edition.

[Angell, 2011] Angell, C. (2011). *Fysikkdidaktik.* Høyskoleforlaget Norwegian Academic Press, 1 edition.

[Barlow, 1999] Barlow, R. J. (1999). *Statistics A Guide to the Use of Statistical Methods in the Physical Sciences.* John Wiley & Sons.

[Bearden, 2013a] Bearden, I. G. (2013a). Force concept inventory test. (Hard copy of the Force Concept Inventory test).

[Bearden, 2013b] Bearden, I. G. (2013b). Private communication.

[Bearden, 2013c] Bearden, J. (2013c). Chi-square test of independence. http://www.geneseo.edu/~bearden/socl211/chisquareweb/chisquare.html.

[Blumann, 2012] Blumann, A. G. (2012). *Elementary Statistics - A Step By Step Approach.* McGraw-Hill, 8 edition.

[Burton, 1997] Burton, D. (1997). *The History of Mathematics: An Introduction.* McGraw-Hill, 6 edition.

[Cohen, 1988] Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Psychology Press, 2 edition.

[Damberg, 2011] Damberg, E. (2011). *Gymnasiepædagogik: En grundbog.* Hans Reitzels Forlag, 1 edition. Additional authors: Jens Dolin, Gitte Holten Ingerslev, et al.

[EdPER, 2013] EdPER (2013). Data handling diagnostic test. (Edinburgh Physics Education Research Group).

[Fagerland, 2013] Fagerland, M. W. (2013). The mcnemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. Coauthors: Stian Lydersen and Petter Laake. www: http://www.biomedcentral.com/1471-2288/13/91.

[Finnich, 2013] Finnich, A. (2013). Grundlæggende matlab og dataanalyse.

[Fleiss, 2003] Fleiss, J. L. (2003). *Statistical Methods for Rates and Proportions.* John Wiley & Sons, 3 edition.

[Forbes, 2011] Forbes, C. (2011). *Statistical Distributions.* John Wiley & Sons, 4 edition.

[Google, 2014] Google (2014). Visualization: Histogram. https://google-developers.appspot.com/chart/interactive/docs/gallery/histogram.

[Hald, 2004] Hald, A. (2004). *A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713 to 1935.* Department of Applied Mathematics and Statistics University of Copenhagen.

[Halloun, 2013] Halloun, I. (2013). Force concept inventory test. (Arizona State University).

[Hansen, 2009] Hansen, E. (2009). *Measure Theory.* Department of Mathematical Sciences University of Copenhagen, 4 edition.

[Hitchcock, 2009] Hitchcock, D. B. (2009). Yates and contingency tables: 75 years later. University of South Carolina, www: http://www.stat.sc.edu/~hitchcock/yates75tech.pdf.

[Hoenig, 2001] Hoenig, J. M. (2001). The abuse of power: The pervasive fallacy of power calculations for data analysis. *The American Statistician.* coauthor: Dennis M. Heisey.

[Jyllands-Posten, 2014] Jyllands-Posten (2014). Råd: Vi skal alle have nye cpr-numre. http://jyllands-posten.dk/indland/ECE6067843/rad-vi-skal-alle-have-nye-cpr-numre/.

[KU, 2013a] KU (2013a). Design guide det natur- og biovidenskabelig fakultet. http://designguide.ku.dk/science/om_design/logo_og_farve.

[KU, 2013b] KU (2013b). Nfyb10013u kvantemekanik 1 (km1) 2013/2014. http://kurser.ku.dk/course/nfyb10013u/2013-2014.

[KU, 2013c] KU (2013c). Nmab10007u lineær algebra i naturvidenskab (linalgnat) 2013/2014. http://kurser.ku.dk/course/nmab10007u/2013-2014.

[KU, 2014a] KU (2014a). Nfyb10004u indledende mekanik og relativitetsteori (mek1) 2013/2014. http://kurser.ku.dk/course/nfyb10004u/2013-2014.

[KU, 2014b] KU (2014b). Nfyb10012u videregående klassisk mekanik (mek2) 2013/2014. http://kurser.ku.dk/course/nfyb10012u/2013-2014.

[Mathworks, 2013] Mathworks (2013). Source code of the command *histfit* in matlab installation located in ⟨installation folder⟩\toolbox\stats\stats.

[Mickey, 2004] Mickey, R. M. (2004). *Applied Statistics: Analysis of Variance and Regression.* John Wiley & Sons, 3 edition.

[Retsinformation, 2013a] Retsinformation (2013a). Bekendtgørelsen for hf-uddannelsen. https://www.retsinformation.dk/Forms/R0710.aspx?id=152579#Bil18.

[Retsinformation, 2013b] Retsinformation (2013b). Bekendtgørelsen for hhx-uddannelsen. https://www.retsinformation.dk/Forms/R0710.aspx?id=152537.

[Retsinformation, 2013c] Retsinformation (2013c). Bekendtgørelsen for htx-uddannelsen. https://www.retsinformation.dk/Forms/R0710.aspx?id=152550.

[Retsinformation, 2013d] Retsinformation (2013d). Bekendtgørelsen for stx-uddannelsen. https://www.retsinformation.dk/Forms/R0710.aspx?id=152507.

[Shalabh, 2009] Shalabh, H. T. (2009). *Statistical Analysis of Designed Experiments.* Springer, 3 edition.

[Sheshgol, 2013] Sheshgol, A. P. (2013). Private communication.

[Skovgaard, 1999] Skovgaard, I. (1999). *Basal Biostatistik del 1.* DSR Forlag, 1 edition.

[Sørensen, 2013] Sørensen, A. S. (2013). Private communication.

[Taylor, 1997] Taylor, J. R. (1997). *An Introduction to Error Analysis.* University Science Books, 2 edition.

[Trosset, 2008] Trosset, M. W. (2008). *An Introduction to Statistical Inference and Its Applications with R.* Department of Statistics Indiana University.

[Undervisningsministeriet, 2011] Undervisningsministeriet (2011). Exam problem set htx physics a december 2011. http://static.uvm.dk/udd/gym/eksamensopgaver/2011/fysik_A_htx_dec13/.

[Undervisningsministeriet, 2012] Undervisningsministeriet (2012). Exam problem set htx physics a may 2012. static.uvm.dk/udd/gym/eksamensopgaver/2012/fysik_A_htx_maj25/.

[Undervisningsministeriet, 2013] Undervisningsministeriet (2013). Exam problem set stx physics a may 2013. http://www.uvm.dk/~/media/UVM/Filer/Udd/Gym/PDF13/Proever%20og%20eksamen/Opgavesaet/130527%20Opgaver_FYS_stx_A.ashx.

[Undervisningsministeriet, 2014] Undervisningsministeriet (2014). Prøver og eksamen. http://uvm.dk/Uddannelser/Gymnasiale-uddannelser/Proever-og-eksamen. Via this website the examinations of some subjects, of all the four main secondary educations, are published.

[US, 2013] US (2013). Nist/sematech e-handbook of statistical methods: 1.3.5.10. levene test for equality of variances. http://www.itl.nist.gov/div898/handbook/eda/section3/eda35a.htm. (U.S. Commerce Department).

[Weiers, 1999] Weiers, R. M. (1999). *Introduction to Business Statistics.* Thomson Higher Education, 6 edition.

[Wikipedia, 2013] Wikipedia (2013). Pascal's calculator.

[WikiQuote, 2013] WikiQuote (2013). Quote from ronald fisher. http://en.wikiquote.org/wiki/Ronald_Fisher.

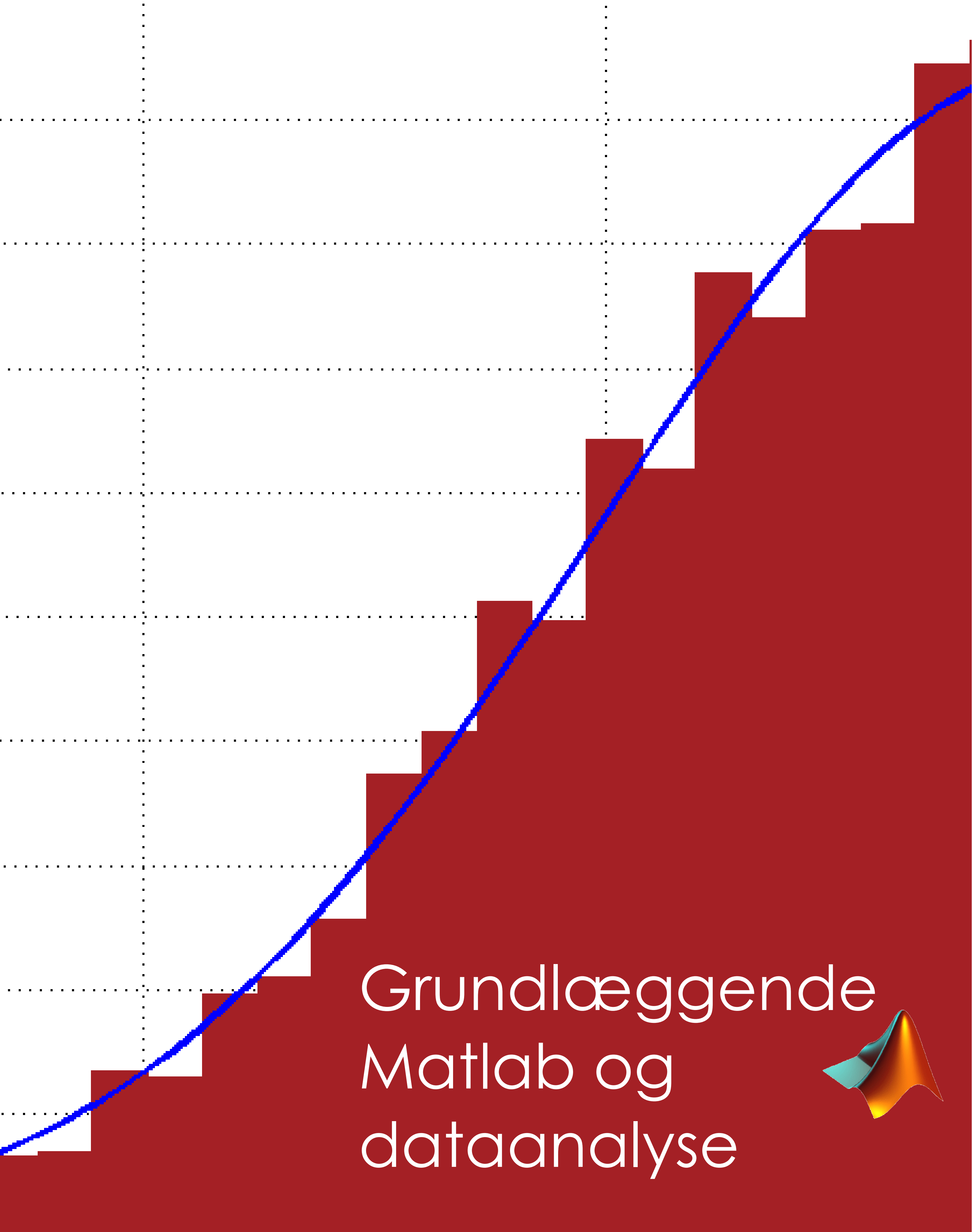
Appendix

A The book "Grundlæggende Matlab og data analyse"

This appendix contains the book "Grundlæggende Matlab og data analyse" which was written for the students of Mek1 in 2013. It treats basics use of Matlab and data analysis in the context of physics experiments. A standalone pdf of the book can be found on the enclosed DVD: The file name is *Basic_Matlab_And_Data_Analysis.pdf*. Be aware that page numbers in this appendix is not of the thesis but of the book.

KØBENHAVNS
UNIVERSITET

ALLAN FINNICH



Grundlæggende Matlab og dataanalyse

Forord

Disse noter er skrevet i forbindelse med mit speciale ved Københavns Universitet, som professor MSO Ian Bearden har vejledt.

Grunden til, at jeg har skrevet disse noter er min egen personlige frustration, da jeg selv i 2009 påbegyndte mit fysikstudie. Frustrationen ligger i, at vi allerede ved første laboratorie-gang skulle anvende grundlæggende usikkerhedsberegninger, hvilket til dags dato stadig ikke er omfattet pensum i gymnasiet. Endvidere skulle man, uden den helt store undervisning i det, benytte CAS-værktøjet Matlab. Nu, fire et halvt år efter, vil jeg med mit speciale - blandt andet - rette op på dette ved at skrive disse noter omkring, hvordan man grundlæggende benytter Matlab, og anvender statistik til at foretage usikkerhedsberegninger i forbindelse med fysikeksperimenter - altså grundlæggende dataanalyse.

For at det hele ikke skal være en for stor mundfuld er indholdet kun basalt, men dækker de emner, som skal benyttes i forbindelse med laboratorie-gangene til kurserne *Indledende mekanik og relativitetsteori (Mek1)* og *Videregående klassisk mekanik (Mek2)* - og mere til. Endvidere er anvendeligheden af noterne ikke kun begrænset til de to førnævnte kurser, da det er en del af kurset *Kvantemekanik 1 (Kvant1)*[1] at foretage numeriske beregninger i Matlab af den tidsuafhængige Schrödinger-ligning. Dog vil disse noter ikke være nok, men der skal suppleres med håndtering af flerdimensionelle arrays (matricer).

Det skal på det kraftigste understreges, at det *ikke* er nok blot at læse noterne og derved forvente, at man nu kan stoffet! Måden I lærer at benytte Matlab og håndtere usikkerhedsberegninger på er ganske simpelt ved at bruge det i forbindelse med fysikeksperimenter - altså det er erfaringen, *såvel som jeres eget engagement*, der på sigt uddanner jer! Derfor skal noterne her ses som værende et relativt letlæseligt opslagsværk.

Noterne vil i kapitel 1 vise grundlæggende funktioner/kommandoer i Matlab, og kapitel 2 fortsætter med hvordan man grafisk fremstiller sine data i plots. I kapitel 3 introduceres statistik, hvilket er en nødvendighed, da det skal benyttes til at snakke om usikkerheder efterfølgende i kapitlet. Efterfølgende vises det, hvordan man foretager fitting af en ret linje til sine data. I alle kapitler er der vist eksempler på anvendelse af stoffet (formler, kommandoer etc.), som gennemgås. Eksemplerne er forsøgt at være så virkelighedsnære som muligt.

Vi kan ikke komme uden om, at vi i de naturvidenskabelige fag ikke kommer særlig langt, hvis vi kommer til kort på den matematiske front. Undervejs kan det være, at I i noterne støder på matematiske begreber og symboler, som I ikke er bekendt med. Begreberne vil højst sandsynligt være betydningen af Σ-symbolet og begrebet partiel differentiering, da de ikke er omfattet af gymnasiepensumet. Disse matematiske begreber bliver I naturligvis undervist I i forbindelse med introduktionskurset i matematik, som I følger sideløbende med *Mek1*. Men om I har lært begreberne i kurset når I går i gang med disse noter er uvist, og derfor er de basalt beskrevet i appendiks A.

Jeg vil afslutte forordet med lidt reklame for nogle kurser, som jeg mener har stor relevans for at kunne blive en virkelig god og reflekterende eksperimental fysiker, samt et godt råd:

- EKSPERIMENTAL FYSIK (EF)[2]: Jeg fulgte dette bachelorkursus i skoleåret 2010-11. Underviseren var lektor Kim Lefmann. Kursets titel afspejler direkte, hvad det omhandler - nemlig eksperimentel fysik. I vil i kurset skulle trække på jeres - til den tid - opnåede erfaring i laboratoriet, og selv tage styringen for hvordan i udfører et forsøg som efterviser en konkret teori. Undervejs vil der blive afholdt forelæsninger omhandlende statistiske fordelinger og begreber, som senere knyttes sammen til usikkerhedsberegninger i forbindelse med eksperimenter. Endvidere har kurset også til formål at styrke jeres evne til at skrive rapporter og artikler. Det var i forbindelse med dette kursus jeg følte, at jeg fik et rigtig godt overblik omkring statistik i forbindelse med eksperimenter.

[1] Se `http://kurser.ku.dk/course/nfyb10013u/2013-2014`
[2] Se `http://kurser.ku.dk/course/nfya09025u/2013-2014`

- ANVENDT STATISTIK: FRA DATA TIL RESULTATER[3]: Kurset her er et kandidatkursus undervist af lektor Troels Christian Petersen. Dette kursus fulgte jeg i skoleåret 2012-13. Indholdet af kurset er ganske enkelt, hvordan man benytter statistik til mange - nogle gange de mest mærkelige - ting. Kurset bygger for det meste oven på de begreber man lærte omkring i EF, men har fokus på anvendelse. Derfor er graden af bevisførelse stort set lig nul. Som standard benytter kurser ikke Matlab, men ROOT som er det program benyttet og udviklet af fysikerne ved LHC i CERN. Også dette kursus vil jeg varmt anbefale, da det netop har fokus for anvendelse af statistik i forbindelse med fysikeksperimenter såvel som andre grene af naturvidenskaben og diverse erhverv.

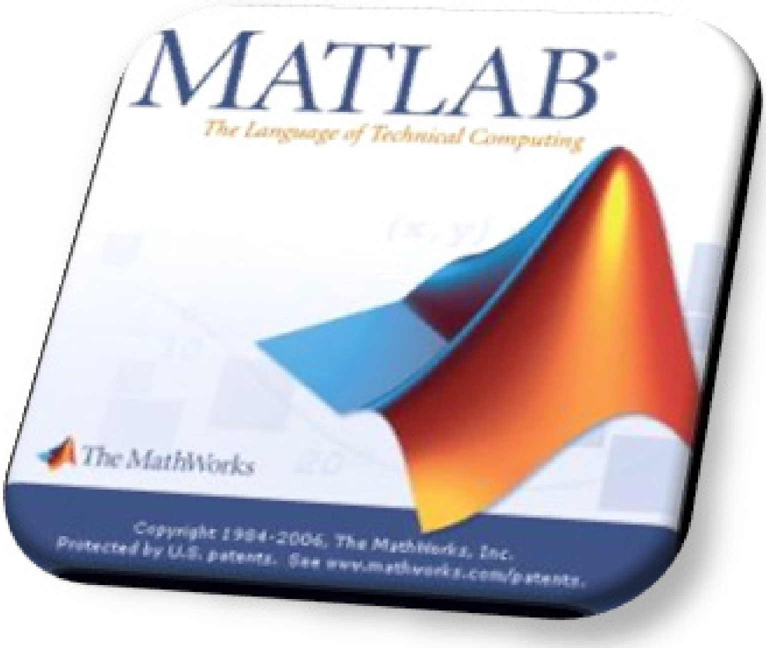
- ET GODT RÅD: Hvis I ikke allerede har tilegnet jer en basal anvendelse i LaTeX så er det bare med at komme i gang, for Word er død! Dette faktum gælder stort set inden for alle grene af naturvidenskaben. Der benyttes så mange matematiske og andre symboler, som vil tage en evighed af skrive ind med Word's equation editor. Standarden er LaTeX , og benyttes til indskrivning af videnskabelige artikler på et globalt plan[4]. Føler man det er for uoverskueligt, at skulle i gang med selv udbydes der et LaTeX kursus i mellemugen mellem blok 1 og 2. Jeg har ikke selv taget kurset, men jeg har kun hørt gode ting om kurset fra mine medstuderende. Har I allerede nu fået blod på tanden til at gå i gang med det selv kan jeg anbefale bogen *The Not So Short Introduction to LaTeX*. Bogen kan findes på `http://tobi.oetiker.ch/lshort/lshort.pdf`.

Allan Finnich
11. november 2013

[3] Se `http://kurser.ku.dk/course/nfyk13011u/2013-2014`

[4] Disse noter er skrevet med LaTeX . Men kompleksiteten af design og layout kan ikke laves i LaTeX med mindre man har erfaring; fem år for at være helt præcis. Så tålmodighed skal udvises.

Indholdsfortegnelse



1 Basic Matlab

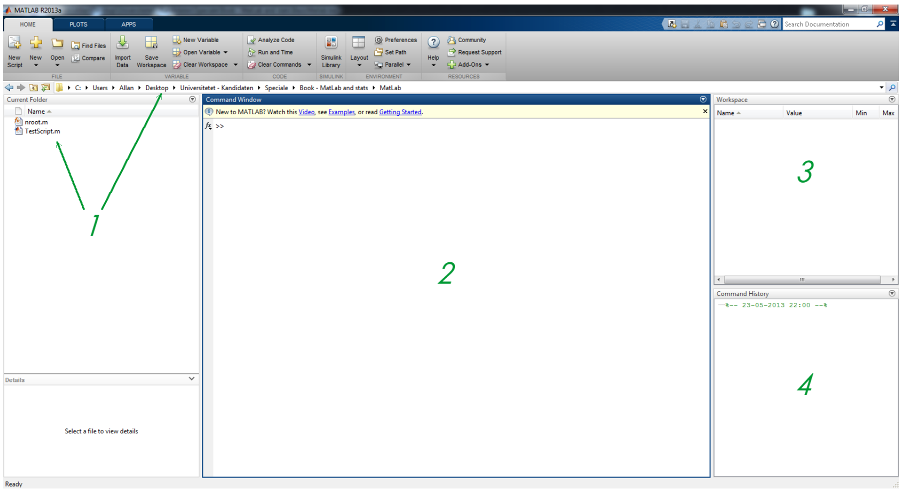
Kapiteloversigt

I dette kapitel skal vi tage de første skridt til at blive bekvemte med Matlab. Først ser vi hvordan Matlabs brugerflade ser ud. Herefter betragter vi, hvad jeg mener er god programmeringspraksis, og hvorfor dette er vigtigt. Efterfølgende ser vi på basale programmeringsbegreber herunder kommantarer, variabler og arrays, og i samme åndedrag vises det hvordan man laver scripts. Klausuler og løkker introduceres som det næste, og til sidst ser,vi hvordan man i Matlab laver sine helt egne funktioner. Der gives eksempler på simple og avancerede funktioner.

1.1 Matlabs brugerflade

Inden vi kommer i gang med at se, hvordan man benytter Matlab skal vi lige se hvordan brugerfladen fungerer helt grundlæggende.

Første gang du åbner Matlab vil din skærm se omtrent ud som i vist i figur 1.1.



FIGUR 1.1 • NÅR DU ÅBNER MATLAB FOR FØRSTE GANG VIL DU SE OMTRENT DET SAMME SOM VIST I FIGUREN. 1. CURRENT FOLDER, 2. COMMAND WINDOW, 3. WORKSPACE OG 4. COMMAND HISTORY.

1. CURRENT FOLDER: Dette vindue viser hvad der ligger i biblioteket som stien henviser til næstøverst på skærmen. De filer, som ikke umiddelbart kan åbnes i Matlab vil fremstå med en lavere opacitet/være fadet ud.

2. COMMAND WINDOW: I dette vindue gives input i form af kommandoer, og det er også her Matlab viser output. I venstre side af vinduet vil du se symbolet "> >", hvilket kaldes for prompten. Som et eksempel kan vi bede Matlab om at udregne det simple regnestykke 1+2. Hertil skal vi blot ud for prompten skrive 1+2 og trykke `enter`. Command vinduet vil nu se ud som i kodeblok 1.1.

```
>> 1+2
ans =
     3
```

KODEBLOK 1.1 • UDREGNING AF DET SIMPLE REGNESTYKKE $1+2$ I COMMAND WINDOW.

Outputtet er `ans = 3`, hvilket er Matlabs måde at give resultatet på. `ans` er en forkortelse af *answer*, hvilket betyder *svar*. I tilfældet af, at man ikke ønsker at se outputtet kan blot for skrive et semikolon. Se kodeblok 1.2.

```
>> 1+2;
```

KODEBLOK 1.2 • UDREGNING AF DET SIMPLE REGNESTYKKE $1+2$, HVOR MAN IKKE ØNSKER AT SE OUTPUTTET. DE VIL SIGE AT MATLAB IKKE VIL OUTPUTTE *ans = 3*.

Ønsker man at rydde op i Command vinduet benyttes kommandoen `clc`. Derved vil du se, at alt output nu er væk.

3. WORKSPACE: Workspace er helt præcist en oversigt over, hvad der ligger i hukommelsen. Eksempelvis, efter vi førhen udregnede `1+2` er der nu dukket en linje op i workspace vinduet. I kolonnen *Name* står der nu `ans` og i value `3`. Dette skyldes, at resultatet af udregningen har Matlab automatisk lagt i en variabel ved navn `ans`, som netop indeholder tallet `3`. Så hvis du på nuværende tidspunkt skriver `ans` vil Matlab outputte det, som ligger i variablen - nemlig `3`. Foretager du nu i stedet udregningen af regnestykket `2+3` vil du bemærke, at `ans`-variablen har nu ændret sig til `5`. Helt præcist indeholder `ans`-variablen resultatet af den sidst udførte kommando. Vi kommer lidt senere til at høre mere omkring, hvad en variabel er.

4. COMMAND HISTORY: Matlab holder styr på, hvornår du har udført hvilke kommandoer. Denne oversigt over udførte kommandoer kan du se i Command History-vinduet. Dette vindue er især nyttig hvis du skal foretage et eller andet, som du måske ikke kan huske kommandoen på. Dog ved du, at du på et eller andet tidspunkt for nyligt har brugt den, så kan du finde den i Command History-vinduet. Ønsker du at udføre kommandoen igen kan du dobbeltklikke på den i Command History-vinduet og Matlab vil eksekvere den igen.

1.2 God programmeringspraksis

Når man sidder og programmerer uanset om det er i Matlab eller ej - kan man hurtigt miste overblikket. Dette gælder især hvis det man gerne vil opnå er mere kompliceret rent kodemæssigt. Hvordan man bedst bevarer overblikket er individuel. Men personligt følger jeg et bestemt regelsæt, som hjælper mig til at bevare overblikket langt hen ad vejen.

KOMMENTARER: Underkend aldrig værdien i at skrive kommentarer alle steder i koden. Hvis man kort skriver en kommentar til hvad den efterfølgende del af koden gør er dette en fordel, da man efter efter noget tid, eksempelvis et år eller to, kan åbne koden og via kommentarene kan se, hvad koden gør.

OPRET SCRIPTS: Når udførelser af mange kommandoer skal til for, at resultatet kan opnås er indtastning i Command Window på ingen måde optimalt! Derfor er det langt mere hensigtsmæssig, at lave såkaldte *scripts*. Når scriptets køres eksekveres al dets indhold.

GEM OFTE!: Jeg kan ikke understrege dette ofte nok: ***gem ofte!*** Gør det til en vane at hver gang I minimerer Matlab, at trykke `ctrl + s` eller `cmd + s` alt efter om det er en Windows eller MacOS platform. Bemærk, at `ctrl + s` betyder, at `ctrl`-tasten skal holdes nede samtidig med at der trykkes på `s`. Det samme er gældende for `cmd + s`. Genvejstasten til at gemme, hvis platformen er Linux er jeg ret sikker på er den samme som i Windows.

GEM FLERE VERSIONER AF SCRIPTS: Har I en fungerende kode/script, men skal ændre i den fordi det skal optimeres eller andet, så husk at gemme det fungerende script, som et andet filnavn *inden* der foretages ændringer! I kan risikere, at efter ændringerne er foretaget, at koden slet ikke virker mere, og det er ekstremt svært at holde styr på hvad det nu var man ændrede. Derfor gem den fungerende kode som eksempelvis det oprindelige filnavn, men vedhæft `working`. Eksempelvis `MinKode_Working.m`. For en god ordens skyld så hold dig til det engelske alfabet når du navngiver dine filer.

Disse tre simple regler har personligt hjulpet mig rigtig meget, og som udgangspunkt kan de måske også virke for jer. Lad os derfor fremhæve regelsættet.

God programmeringspraksis

- Skriv koden i scripts.
- Lav kommentarer i scriptet.
- GEM OFTE! `ctrl + s`/`cmd + s` (Windows/MacOS).
- Gem et backup af den fungerende script, hvis der skal foretages radikale ændringer eller gøre tingene på en anden måde.

Som tiden skrider frem, og I bliver mere vant til at programmere vil I begynde, at danne jeres helt eget regelsæt. Måske fordi de førnævnte ikke virker for jer eller af andre årsager. Men dette regelsæt er et godt sted at starte.

1.3 Grundlæggende brug af Matlab

Vi kommer her til at se hvordan man programmerer i Matlab gennem en række eksempler. Først kigger vi på hvordan man laver kommentarer i sin kode, og herefter kigger vi på de mest gængse løkker og klausuler.

Oprettelse af scripts

Inden vi opretter vores første script skal vi først have oprettet et bibliotek på computerens harddisk og bede Matlab om kun at kigge i dette. Opret derfor biblioteket, hvor I nu ønsker angiv stien til dette. Det kan, eksempelvis, placeres på `C:`-drevet med navnet `Matlab_Scripts`. Ergo skal I angive stien `C:\Matlab_Scripts`. Vi er nu klar til at oprette vores første script. Tryk på *New Script*-ikonet . Der popper nu et nyt vindue op. Dette er Matlabs editor som vi vil benytte til at skrive vores scripts i. Bemærk dog, at man sagtens kunne gøre det i NotePad eller andre editorer, men fordelen ved Matlabs editor er, at denne har syntaks-highlighting.

Gem dette tomme script som `Matlab_Basic.m`. Matlab scripts har `.m` extensionen. Man behøver ikke angive extensionen når man gemmer sit script via den indbyggede editor, da Matlab sørger for at vedhæfte `.m` på filnavnet du angiver. Derfor kan du blot som navn angive `Matlab_Basic` og den vil automatisk blive gemt som `Matlab_Basic.m`.

Kommentarer

Som regelsættet for god programmeringspraksis foreskrev er det vigtigt at angive kommentarer som fortæller hvad koden gør. Kommentarer i koder bliver ikke eksekveret, dvs. at Matlab ignorerer dem. Kodeblok 1.3 viser, hvordan man laver en kommentar i sin Matlab kode.

```
%En kommentar startes med et procenttegn og vil blive fremhævet med
    grønt i Matlab.
```

KODEBLOK 1.3 • KØRES DENNE KODE I MATLAB VIL INTET SKE, DA DET BLOT ER EN KOMMENTAR.

Man skal simpelthen blot starte med et procenttegn. Alt på linjen efter procenttegnet vil Matlab opfatte som en kommentar og ignorerer.

Hvis man vil lave en kommentar med linjeskift, dvs. du har trykket på `enter`, og derved startet en ny linje, vil Matlab ikke længere opfatte det efterfølgende som en del af kommentaren.

```
%{
Ønsker man en større kommentar,

hvori der er linjeskift kan

man udkommentere hele stykket

på denne måde.
%}
```

KODEBLOK 1.4 • ER DER LINJESKIFT I EN KOMMENTAR ER DET IKKE NOK BLOT AT ANGIVE ET %-TEGN. DERIMOD SKAL KOMMENTAREN INDKRANSES AF TEGNENE `%{` OG `}%`.

For at overkomme denne problemstilling kan man angive et start- og sluttegn, hvilket medfører at alt, som står mellem tegnene opfatter Matlab som kommentarer. Kigger du i kodeblok 1.4 kan du se, at starttegnet er `%{` og sluttegnet er `%}`.

Variabler

Vi har allerede hørt om `ans`-variablen i sektion 1.1. I den lå altid det output fra den sidst udførte kommando. Men hvad er en variabel egentlig? Jo, en variabel kan opfattes helt analogt til en opbevaringskasse. Der ikke er de store grænser for, hvad man kan putte i kassen, og man ændre og benytte indholdet efter behov. I sektion 1.1 udførte vi det simple regnestykke `1+2` og her så vi, at Matlab lagde resultatet - tallet `3` - i `ans`-variablen. Så i dette tilfælde lægges der et tal i en kasse, hvorpå der står påskrevet `ans`.

Lad os nu skrive et lille script til at udføre nogle simple regnestykker. Vi ønsker at udregne `1+2` og `2+3` og lægge resultaterne i hver deres variabel hhv. `a` og `b`. Dette gøres som vist i kodeblok 1.5.

```
%Første variabel kaldes for a og skal indeholde resultatet af 1+2
a=1+2

%Anden variabel kaldes for b og skal indeholde resultatet af 2+3
b=2+3
```

KODEBLOK 1.5 • EN VARIABELS INDHOLD DEFINERES GANSKE SIMPELT VED FØRST AT SKRIVE NAVNET PÅ VARIABLEN, ET LIG-MED-TEGN OG TILSIDST INDHOLDET. HER DEFINERES VARIABLERNE `a` OG `b` TIL AT INDEHOLDE RESULTATET AF HHV. `1+2` OG `2+3`.

Kopieres koden i kodeblok 1.5 ind i det nyoprettede script `Matlab_Basic.m`, og klikker på *Run*-ikonet  eller trykker `F5` køres scriptet. Kigger du nu i Command Window kan du direkte se, at scriptet er kørt, da der nu vises output. Indvidere vil du måske lægge mærke til at, Workspace også er ændret. Variablerne `a` og `b` ligger nemlig nu i hukommelsen og derved også i Workspace.

Definition af variabler

En variabels indhold i Matlab defineres på følgende måde:

$\langle$*variablens navn*$\rangle$ = $\langle$*variablens indhold*$\rangle$

VIGTIGT!: Navngiv altid dine variabler så du hurtigt kan se hvad de indeholder! Dette vil spare tid!

Variablerne `a` og `b`, som nu ligger i hukommelsen, repræsenterer nu hhv. tallene `3` og `5`. Vi vil nu benytte disse to variabler til at vise hvordan man udfører matematiske regneoperationerne med dem.

```
%Summen af a og b
Sumab=a+b
%Differencen af mellem b og a
Diffba=b-a
%Differencen af mellem a og b
Diffab=a-b
%Produktet af a og b
Prodab=a*b
%Forholdet mellem a og b
Divab=a/b
%Forholdet mellem b og a
Divba=b/a
%a opløftet i b
Expab=a^b
```

KODEBLOK 1.6 • HVIS VARIABLER INDEHOLDER TAL KAN DER UDFØRES MATEMATISKE REGNEOPERATIONER MED DEM. HVORDAN DETTE GØRES ER IKKE ANDERLEDES END PÅ EN LOMMEREGNER.

I kodeblok 1.6 er det vist, hvordan man udfører de matematiske operationer `+`,`-`,`*`,`/` og `^` - altså sum, difference, multiplikation, division og potens. I koden er det valgt, at lægge resultaterne i hver deres respektive variabel. Bemærk, at fremgangsmetoden stort set er den samme som på en lommeregner.

Matematiske regneoperationer

Matematiske regneoperationer med tal eller variabler indeholdende tal udføres med symbolerne `+`,`-`,`*`,`/` og `^` på følgende måde:

Operation	Symbol	Eksempel
Sum	+	3+5
Difference	-	3-5
Multiplikation	*	3*5
Division	/	3/5
Potens	^	3^5

Matlab indeholder en række indbyggede matematiske funktioner. Eksempelvis kvadratrod og numerisk værdi. Tabel 1.1 viser nogle af de mest gængse. Vi skal senere se, hvordan man definerer sin helt egen funktion.

Det er også vigtigt, at kunne fjerne variabler fra hukommelsen. For hvis der bliver lagt mere og mere i hukommelsen vil systemet gå hen og blive trægt. Netop af denne grund er det vigtigt, at man i alle sine scripts starter med at rydde hukommelsen! Dette gøres ved kommandoen `clear all`. Som navnet antyder rydder kommandoen simpelthen alt, og man vil se, at Workspace efterfølgende er tomt. I enkelte tilfælde er man måske kun interesseret at slette en specifik variabel. Dette kan opnås på to forskellige måder:

1. Benytte kommandoen `clear` ⟨*navn på variabel*⟩.
2. Med musen højre-klikkes på variablen man ønsker slettet i Workspace og klikker `Delete`, eller blot markerer variablen og trykker på `Delete`-tasten.

Arrays

Vi så i sektion 1.3, hvordan man definerer variabler. Det foregik ved hjælp af =-tegnet. Dog betragtede vi kun yderst simple eksempler, hvor variablerne blev defineret til at indeholde et enkelt tal, hvilket - som fysikker samt andre naturvidenskabelige analytikere - ikke er tilstrækkeligt. Vi sidder med en masse data, hvilket er en nødvendighed for, at vi inden for rammerne sat af statistik - hvilket vi kommer til **senere** - kan drage

Kommando	Beskrivelse	Eksempel
`sqrt(x)`	Kvadratroden af `x`	`sqrt(81)`
`exp(x)`	e^x	`exp(2)`
`abs(x)`	Numerisk værdi af `x`	`abs(3)`
`log(x)`	Naturlig logaritme af `x` ($\ln(e^x) = x$)	`ln(4)`
`log10(x)`	Logaritmen af `x` ($\log(10^x) = x$)	`log10(5)`
`sin(x)`	sinus til `x` (x i radianer)	`sin(6)`
`cos(x)`	cosinus til `x` (x i radianer)	`cos(6)`
`tan(x)`	tangens til `x` (x i radianer)	`tan(6)`
`asin(x)`	arcsinus til `x` (Vinkelmål=radianer)	`asin(6)`
`acos(x)`	arccosinus til `x` (Vinkelmål=radianer)	`acos(6)`
`atan(x)`	arctangens til `x` (Vinkelmål=radianer)	`atan(6)`
`sind(x)`	sinus til `x` (x i grader)	`sind(6)`
`cosd(x)`	cosinus til `x` (x i grader)	`cosd(6)`
`tand(x)`	tangens til `x` (x i grader)	`tand(6)`
`asind(x)`	arcsinus til `x` (Vinkelmål=grader)	`asind(6)`
`acosd(x)`	arccosinus til `x` (Vinkelmål=grader)	`acosd(6)`
`atand(x)`	arctangens til `x` (Vinkelmål=grader)	`atand(6)`

TABEL 1.1 • EN RÆKKE AF DE OFTEST BENYTTEDE INDBYGGEDE MATEMATISKE FUNKTION I MATLAB.

konklusioner. Alt denne data vil vi i stort set alle tilfælde lægge i et eller flere arrays.

Et array er ganske simpelt blot en liste bestående af en række tal. Eksempelvis tallene fra 1 til 5. For at lave et array skal symbolerne `[` og `]` benyttes. Lad os se hvordan det helt præcist fungerer ved at lave et array `MyArray` indeholdende tallene fra 1 til 5. Koden er vist i kodeblok 1.7.

```
%Vi definerer variablen MyArray til at indeholde et array med
    tallene fra 1 til 5
MyArray = [1 2 3 4 5]
```

KODEBLOK 1.7 • KODEN VISER, HVORDAN MAN DEFINERER ET ARRAY. ARRAYETS ELEMENTER SKAL ADSKILLES AF ET MELLEMRUM, OG ER AFGRÆNSET AF FIRKANTEDE PARENTESER.

Når et array er defineret kan vi gøre med det hvad vi vil. Ønsker vi eksempelvis at få vist, hvad det tredje element er i `MyArray` - altså hvad der står på plads nummer tre - så kan vi blot skrive `MyArray(3)`, hvilket i vores tilfælde ville outputte tallet 3.

Antallet af elementer i et array kaldes også for dets dimension. `MyArray` er således 5-dimensionel. Ønsker vi at kende et arrays dimension kan vi gøre dette ved at benytte kommandoen `length`. I kodeblok 1.8 demonstreres det, hvordan dimensionen af et array bestemmes og hvordan man ændre et element i et array.

```
%Vi definerer variablen MyArray til at indeholde et array med
    tallene fra 1 til 5
MyArray = [1 2 3 4 5]

%Dimensionen af MyArray bestemmes ved brug af kommandoen length
length(MyArray)

%Man kan ændre et arrays elementer. Eksempelvis kan vi ændre
    MyArrays fjerde element til 10
MyArray(4)=10
```

KODEBLOK 1.8 • NÅR ET ARRAY ER DEFINERET KAN VI BESTEMME DETS DIMENSION - ALTSÅ HVOR MANGE ELEMENTER DET INDEHOLDER. DETTE SKER VED BRUG AF KOMMANDOEN *`length`*. ENDVIDERE KAN ET ARRAYS ELEMENTER SNILDT ÆNDRES EFTER BEHOV.

Definition af arrays

Et array i Matlab er en række af elementer bestående af tal. Et array defineres på følgende måde:

⟨*navn på array*⟩ = [⟨*indhold angivet med et mellemrum mellem hvert element*⟩]

Arrayets i'te element kan ændres på følgende måde:

⟨*navn på array*⟩`(i)` = ⟨*det i'tes elements nye værdi*⟩

Dimensionen af et array kan bestemmes ved at benytte kommandoen `length`:

`length(`⟨*navn på array*⟩`)`

Ligesom med variabler indeholdende tal kan der udføres matematiske regneoperationer på variabler indeholdende arrays af tal. Ønsker man at bestemme summen af to arrays benyttes symbolet + og for differencen mellem dem -. Det er dog også nyttigt blot at bestemme summen af et arrays elementer, hvilket hurtigt gøres med kommandoen `sum`. Se kodeblok 1.9.

```
%Vi definerer to arrays: MyArray1 og MyArray2
MyArray1 = [1 2 3 4 5]
MyArray2 = [6 7 8 9 10]

%Summen af MyArray1 og MyArray2
MyArray1+MyArray2

%Differensen mellem MyArray1 og MyArray2
MyArray1-MyArray2

%Summen af MyArray1
sum(MyArray1)

%Summen af MyArray2
sum(MyArray2)
```

KODEBLOK 1.9 • SUMMEN OG DIFFERENCEN MELLEM TO ARRAYS BESTEMMES PÅ SAMME MÅDE, HVIS DET BLOT HAVDE VÆRET TO VARIABLER INDEHOLDENDE ET TAL. ENDVIDERE ER DET VIST HVORDAN MAN HURTIGT KAN BESTEMME SUMMEN AF DE INDIVIDUELLE ARRAYS MED KOMMANDOEN *`sum`*.

Når enten summen eller differencen mellem to arrays ønskes bestemt vil det resultere i et nyt array med samme dimension, men hvor det m'te element vil bestå af summen/differencen mellem m'te element i de to arrays. Eksempelvis, hvis vi tager udgangspunkt i koden i kodeblok 1.9 sker der følgende:

$$\texttt{MyArray1}\pm\texttt{MyArray2} = [1\pm 6\ \ 2\pm 7\ \ 3\pm 8\ \ 4\pm 9\ \ 5\pm 10].$$

Indtil videre har vi set, hvordan man adderer og subtraherer to arrays. Man skulle umiddelbart tro, at multiplikation og division fungerer på samme måde. Men dette er ikke tilfældet! Dette bunder i, at Matlab behandler arrays som vektorer, og som I muligvis har lært i gymnasiet så er produktet mellem to vektorer jo ikke en ny vektor med samme dimension, men blot et tal - nemlig skalarproduktet[5]. Endvidere er division af to vektorer for det første intetsigende og for det andet slet ikke defineret. Det vi her er interesseret i er,

[5] Vi vil her ikke komme nærmere ind på hvordan man udregner skalarproduktet mellem to vektorer i Matlab. Grunden hertil er, at man skal kende til regnereglerne for matricer og dette kommer du ikke til at lære om før i kurset *Lineær Algebra* i blok 2. Men blot bid mærke i, at det er muligt.

at to arrays elementer parvist enten skal multipliceres eller divideres med hinanden. Dette kaldes for en *element-by-element* operation.

Som vi så i kodeblok 1.9 er dette intet problem for addering og subtrahering, da dette sker automatisk ifølge regnereglerne for vektorer - altså element-by-element. Vi tager igen udgangspunkt i `MyArray1` og `MyArray2` i kodeblok 1.9. Det vi gerne vil opnå er følgende operationer:

$$\texttt{MyArray1} \cdot \texttt{MyArray2} = [1\cdot 6 \quad 2\cdot 7 \quad 3\cdot 8 \quad 4\cdot 9 \quad 5\cdot 10]$$
$$\texttt{MyArray1}/\texttt{MyArray2} = [1/6 \quad 2/7 \quad 3/8 \quad 4/9 \quad 5/10].$$

Proceduren for at opnå dette er heldigvis ikke særlig kompliceret. Der skal blot skrives et punktum foran den matematiske operator. Se kodeblok 1.10.

```
%Vi definerer de to arrays: MyArray1 og MyArray2
MyArray1 = [1 2 3 4 5]
MyArray2 = [6 7 8 9 10]

%Her laves et nyt array som vi kalder ProdMyArray1And2, hvis i'te element indeholder MyArray1 og MyArray2's i'te element multipliceret
ProdMyArray1And2 = MyArray1.*MyArray2

%Her laves et nyt array som indeholder MyArray1 og MyArray2's elementer divideret. Altså det i'te element i DivMyArray1And2 er det i'te element i MyArray1 divideret med det i MyArray2
DivMyArray1And2 = MyArray1./MyArray2
```

Kodeblok 1.10 • Koden viser, hvordan man multiplicerer og dividerer to arrays elementer med hinanden - kort sagt element-by-element operation.

Helt analogt foregår det, hvis man ønsker at eksponere et array element-by-element. Se kodeblok 1.11. Det som koden i kodeblok 1.11 gør er følgende:

$$\texttt{MyArray3.\^{}MyArray4} = [1^{0.1} \quad 2^{0.2} \quad 3^{0.3} \quad 4^{0.4} \quad 5^{0.5}]$$

```
%Vi definerer to arrays: MyArray3 og MyArray4
MyArray3 = [1 2 3 4 5]
MyArray4 = [0.1 0.2 0.3 0.4 0.5]

%Her laves et nyt array som vi kalder PwrMyArray1And2, hvis i'te element indeholder MyArray3's i'te element opløftet til MyArray2's i'te element
PwrMyArray3And4 = MyArray1.^MyArray2
```

Kodeblok 1.11 • Koden viser, hvordan man opløfter et array til et andet array element-by-element.

De indbyggede matematiske funktioner nævnt i tabel 1.1 er heldigvis programmeret således, at de som standard udfører regneoperationen element-by-element (EBE) på et array. Det vil sige, at hvis vi eksempelvis gerne vil tage kvadratroden af et array kan man blot skrive `sqrt(`⟨*navn på array*⟩`)`, da det automatisk sker EBE. Se kodeblok 1.12.

```
%Vi definerer et array:
NewArray = [4 9 16 25 36]

%Her laves et nyt array som vi kalder sqrtNewArray, hvis i'te element indeholder kvadratroden af NewArray's i'te element
sqrtNewArray = sqrt(NewArray)
```

Kodeblok 1.12 • Hvis man som input angiver et array i en af de indbyggede matematiske funktioner i MatLab sker operationen som standard element-by-element.

EBE operationer på arrays

Skal der udføres EBE operationer mellem to arrays skal følgende syntaks benyttes:

$$\langle \textit{Første array} \rangle \,\texttt{.}\otimes \langle \textit{Andet array} \rangle,$$

hvor $\otimes$ er en af de tre regneoperationer `*`, `/` eller `^`.

1.4 Klausuler og løkker

Der vil komme tilfælde, hvor en bestemt del af jeres kode skal eksekveres, hvis en bestemt betingelse er opfyldt. For at kunne håndtere denne problemstilling skal vi snakke om `if`-klausuler. Det vil måske også være sådan, at en del af koden skal eksekveres et stort antal gange, hvilket kan løses med en `for`-løkke. Både `if`-klausulen og `for`-løkken vil vi her se hvordan virker, og eksempler på anvendelse.

if-klausulen

Som navnet antyder er `if`-klausulen et stykke kode der eksekveres, hvis en bestemt betingelse er opfyldt. Vi skal her se et eksempel på en `if`-klausul. I forbindelse med eksemplet kommer vi også ind på hvad en tekststreng/string er og hvordan disse outputtes. `if`-klausulen er nok ikke det I i starten kommer til at beskæftige jer mest med, men er stadig vigtig nok til, at I skal vide hvordan den er opbygget.

Vi starter ganske kort med, at se hvad en string er. En string kan måske bedst forestilles som sætninger; eksempelvis *Jeg er fysiker*. Strings vil du kunne genkende i Matlab's editor som værende farvede lilla. De skal omkranses af en apostrof `'`. Se kodeblok 1.13.

```
%Vi definerer en variablen MyFirstString til at indeholde stringen
    Jeg er fysiker.
MyFirstString = 'Jeg er fysiker'
```

KODEBLOK 1.13 • KODEN VISER HVORDAN EN STRING LÆGGES I EN VARIABEL. EN STRING OMKRANSES EN APOSTROF `'` I MATLAB

Ved at køre koden i kodeblok 1.13 kan du se nu se der står `MyFirstString` i Workspace. Bemærk, det lille ikon til venstre for `MyFirstString`. I den står der `abc`, hvilket er Matlab's måde at vise, at denne variabel indeholder en string. I value vil du se, at der står `'Jeg er fysiker'`.

Lad os nu se, hvordan man programmerer en `if`-klausul. Helt præcis vil vi gerne have Matlab til at fortælle os om et tal liggende i en variabel kaldet `k` er positiv eller negativ. Hertil skal vi netop benytte os af en `if`-klausul:

- hvis tallet er negativt skal Matlab outputte stringen `'Tallet er negativt'`,
- hvis tallet er positivt skal Matlab outputte stringen `'Tallet er positivt'`.

Kodeblok 1.14 viser koden til at opnå netop dette.

```
%Først definerer variablen k til at indeholde tallet 9
k = 9;

%Nu skal vi så få Matlab til at checke om tallet i variablen k er positivt eller negativt

if (k>0)
%hvis k er stoerre end 0 eksekveres denne del
    disp('Tallet er positivt')
else
%hvis k er mindre eller lig med 0 eksekveres denne del
    disp('Tallet er negativt')
end
```

KODEBLOK 1.14 • FØLGENDE KODE TJEKKER OM TALLET ILAGT VARIABLEN k ER POSITIVT ELLER NEGATIVT.

Det første vi gør er at definere variablen `k` til at indeholde et tal; her tallet ni. Herefter hvis betingelsen `k>0` er opfyldt vil `disp('Tallet er positivt')` blive eksekveret. Er `k` imidlertid mindre eller lig med nul vil `else`-delen blive eksekveret - altså `disp('Tallet er negativt')`. Bemærk, at en `if`-klausul altid sluttes med `end`.

Kommandoen `disp(x)` gør blot at indholdet af `x` vises i Command vinduet.

Vi kan dog være en smule mere eksotiske, for hvis `k` er lig med nul er det jo hverken positivt eller negativt. Det vil altså sige, at vi skal have endnu en betingelse. Dette kan gøres ved at tilføje en `elseif`-del. Se kodeblok 1.15

```
%Først definerer variablen k til at indeholde eksempelvis tallet 0
k = 0;

%Nu skal vi så få Matab til at tjekke om tallet i variablen k er positivt eller negativt

if (k>0)
%hvis k er større end 0 eksekveres denne del
    disp('Tallet er positivt')
elseif (k==0)
%hvis k er lig med 0 eksekveres denne del
    disp('Tallet er hverken positivt eller negativt')
else
%hvis k er mindre end 0 eksekveres denne del
    disp('Tallet er negativt')
end
```

KODEBLOK 1.15 • FØLGENDE KODE TJEKKER OM TALLET ILAGT VARIABLEN k ER POSITIVT ELLER NEGATIVT, OG BEHANDLER TILFÆLDET, HVOR k ER LIG MED NUL SOM VÆRENDE HVERKEN POSITIVT ELLER NEGATIVT VED AT TILFØJE EN *`elseif`*-DEL.

Bemærk, at når man gerne vil have Matlab til at tjekke om en variabel er lig med en bestemt værdi skal der angives to lig-med-tegn! Et enkelt lig-med-tegn er reserveret til at definere indholdet af variabler. Derfor står der i `elseif`-delen i kodeblok 1.15 `k==0`, da vi netop gerne vil tjekke om variablen `k` præcis er lig med nul.

Lad os indramme, hvordan man generelt programmerer en `if`-klausul i Matlab.

if-klausul

Syntaksen for en `if`-klausul i MatLab er følgende:

```
if (⟨betingelse 1⟩)
%Kode som skal eksekveres, hvis betingelse 1 er
    opfyldt
elseif (⟨betingelse 2⟩)
%Kode som skal eksekveres, hvis betingelse 2 er
    opfyldt
elseif (⟨betingelse 3⟩)
%Kode som skal eksekveres, hvis betingelse 3 er
    opfyldt
...
elseif (⟨betingelse m⟩)
%Kode som skal eksekveres, hvis betingelse m er
    opfyldt
else
%Kode som skal eksekveres, hvis ingen af
    betingelserne er opfyldt
end
```

Bemærk der kan indgå lige så mange `elseif`-del som der ønskes[a].

[a] `If`-klausuler kan gøres langt mere kompakte ved at benytte logiske operatorer. Dette er ikke noget vi her vil komme ind på, men interesserede sjæle kan se `http://www.mathworks.se/help/matlab/matlab_prog/operators.html` for information om logiske operatorer.

for-løkken

`for`-løkken er en løkke som du stort set altid skal benytte i forbindelse med databehandling. Før vi kommer til `for`-løkken skal vi dog lige se en ret vigtig og smart funktionalitet i Matlab - nemlig `:`-operatoren.

Denne operator er speciel nyttig, hvis man, eksempelvis, skal definere et array, som indeholder alle hele tal fra 1 til 1000. Indvidere har vi jo kun set eksempler, hvor arrays er blevet defineret manuelt! Kodeblok 1.16 viser blandt andet, hvordan førnævnte array snildt kan defineres ved brug af `:`-operatoren.

```
%Et array skal indeholde alle hele tal fra 1 til 1000
BigArray=1:1:1000;

%Et andet array skal indeholde tallene mellem 0 og 1, hvor forø
    gelsen af hvert element er det forrige plus 0.1, kort sagt i
    skridt af 0.1.
AnotherBigArray=0:0.1:1;
```

KODEBLOK 1.16 • `:`-OPERATOREN ER NYTTIG NÅR STORE ARRAYS SKAL DEFINERES, OG NÅR DEN NUMERISKE AFSTAND MELLEM ARRAYETS ELEMENTER ER KONSTANT.

Bemærk, at vi slet ikke behøver at benytte `[` og `]`, da `:`-operatoren altid vil outputte et array.

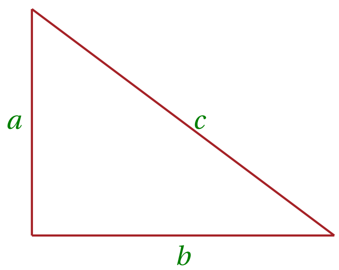
:-operatoren

Syntaksen for Matlab's indbygget :-operator er følgende:

⟨*startværdi*⟩ : ⟨*forøgelse*⟩ : ⟨*slutværdi*⟩

:-operator outputter altid et array. Derfor behøves `[` og `]` ikke blive angivet. Angives disse vil outputtet stadig blot være et array - altså der gør ingen forskel.

Vi er nu klar til at se, hvordan en `for`-løkke er opbygget. Dette gør vi ved at tage udgangspunkt i en konkret problemstilling: En matematikstuderende kommer med en række katetelængder a og b (se figur 1.2) af en række retvinklede trekanter. Han skal bestemme hypotenusen c af alle de retvinklede trekanter på baggrund af hans målte katetelængder. Katetelængderne er angivet i figur 1.2.



a	b
1 cm	1 cm
1 cm	2 cm
2 cm	3 cm
4 cm	1 cm
2 cm	2 cm

FIGUR 1.2 • EN MATEMATIKSTUDERENDE HAR MÅLT KATETELÆNGDER OG ØNSKER I HVERT TILFÆLDE, AT DU SKAL UDREGNE HYPOTENUSEN c. HERTIL BENYTTES MEST EFFEKTIVT EN `for`-LØKKE.

Lad os skitserer, hvordan vi med en `for`-løkke kan løse dette problem:

1. Definere to arrays som respektivt indeholder værdierne for katetelængderne a og b.

2. Via en `for`-løkke definere et array, hvis i'te element indeholder hypotenusen c udregnet ved formlen $\sqrt{a^2+b^2}$.

```
%Først defineres værdierne for a og b i hver deres array
a = [1 1 2 4 2];
b = [1 2 3 1 2];

%Nu påbegyndes for-løkken. Løkken skal gentages ligeså mange gange som der er elementer i arrays a og b - altså fra 1 i skridt af 1 til length(a). Men inden da defineres arrayet c til at indeholde kun nuller altså [0 0 0 0 0]
c = zeros(1,length(a));
for i=1:1:length(a)
c(i) = sqrt(a(i)^2+b(i)^2);
end
```

KODEBLOK 1.17 • KODEN UDREGNER HYPOTENUSEN PÅ BAGGRUND AF DE MÅLTE KATETELÆNGDER VIST I FIGUR 1.2.

I kodeblok 1.17 kan koden ses som løser den matematikstuderendes problem med en `for`-løkke. Bemærk, at lige som med en `if`-klausul afsluttes løkken med `end`.

En vigtig ting her at bemærke er, at hvis der konstrueres et array ved hjælp af en `for`-løkke - som i kodeblok 1.17 - kan tiden det tager for at gennemløbe løkken nedsættes ved at definere arrayet inden selve løkken. Dette er muligt, da man i de fleste tilfælde ved hvad arrayets dimension skal være efter løkken er gennemløbet. Man vil ikke kunne mærke nogen forskel i kørselstid ved små gentagelser, men skal løkken gennemløbes 10000 gange ville man formentlig begynde at kunne mærke forskel. Derfor har vi i linje 6 i kodeblok 1.17 netop defineret arrayet `c` til et array, med samme dimension som `a`, bestående af kun nuller - altså `[0 0 0 0 0]`. Det er det som kommandoen `zeros` gør. En lignende funktion som i stedet ville lave et array med lutter et-taller er `ones`.

`for`-løkke

Syntaksen for en `for`-løkke i MatLab er følgende:

```
for i=⟨startværdi⟩:⟨forøgelse⟩:⟨slutværdi⟩
%Kode som skal eksekveres ved hver værdi for i
end
```

Bemærk, at det ikke er nødvendigt at benytte `i` som tællevariabel. Man kan benytte hvilken som helst variabel. De mest almindelige er dog bogstaverne `i`, `k`, `m` og `n`.

Hvis et array defineres via en `for`-løkke kan kørselstiden nedsættes ved at definere arrayet inden løkken via kommandoen `zeros` eller `ones`. Disse har syntaksen

`zeros(`⟨*antal rækker*⟩`,`⟨*antal kolonner*⟩`)`
`ones(`⟨*antal rækker*⟩`,`⟨*antal kolonner*⟩`)`

1.5 Funktioner

Vi har tidligere nævnt en række af indbyggede matematiske funktioner i Matlab (se tabel 1.1). Men der kan snildt opstår situationer, hvor en funktion, som ikke er indbygget, skal benyttes. Hertil kan man i Matlab definere helt egne funktioner som man kan anvende.

Klikkes der I Matlab på *New* og herefter på *function* oprettes et nyt function-script. I vil nu se koden som vist i kodeblok 1.18 i editoren.

```
function [ output_args ] = Untitled( input_args )
%UNTITLED Summary of this function goes here
%   Detailed explanation goes here


end
```

KODEBLOK 1.18 • NÅR DU KLIKKER *New->function* I MATLAB VIL DEN HER VISTE KODE KUNNE SES I EDITOREN.

Som det fremgår i koden i kodeblok 1.18 starter et function-script altid med `function`, hvilket fortæller Matlab at der her defineres en ny funktion. `output_args` er en forkortelse af *output arguments*, hvilket er det funktionen skal returnere når funktionen kaldes. `input_args` er en forkortelse af *input arguments*, og er det input man skal give funktionen for at få outputtet. Der, hvor der står `Untitled` er navnet på den nye funktion. `output_args`, `input_args` og navnet på funktionen kan vælges frit, men den må ikke have sammen navn som en allerede eksisterende funktion. Det er vigtigt, at filnavnet skal være det samme som det navn man giver funktionen. Så, hvis man eksempelvis vil lave en funktion ved navn `MyFunction` skal function-filen gemmes som `MyFunction.m`! Endvidere skal der gøres opmærksom på, at når man skal kalde funktionen direkte via Command Window skal Matlabs sti være sat til det bibliotek, hvor function-filen er placeret. Kalder man den via et script skal Matlabs sti være sat til biblioteket, hvor både function-filen *og* scriptet er placeret.

Lad os som et eksempel se, hvordan man i Matlab definerer den matematiske funktion $(a \cdot x)^{1/n}$, hvor n, a og x er vilkårlige tal. Se kodeblok 1.19.

```
function [f] = nroot(a,x,n)
f=(a.*x).^(1/n);
end
```

KODEBLOK 1.19 • KODEN DEFINERER DEN MATEMATISKE FUNKTION $(a \cdot x)^{1/n}$, SOM HER NAVNGIVES *nroot*. BEMÆRK, AT DER ER ANGIVET ET PUNKTUM OG DERVED KAN DENNE FUNKTION OGSÅ HÅNDTERE TILFÆLDET HVIS ENTEN *a* ELLER *x* ER ET ARRAY.

Vi kan se, at a,x og n er input argumenterne og f er output argumentet. Bemærk, at f ikke vises, da der er angivet et semikolon! Dette skyldes, at Matlab automatisk lægger outputtet over i ans-variablen. Funktionen har vi valgt at kalde nroot, da det netop er den n'te rod funktionen beregner. Lad os via vores nye funktion nroot udregne $(1 \cdot 4)^{1/2}$, hvilket meget gerne skulle give 2 som resultat[6]. Det vil sige, at vi skal kalde funktionen nroot, hvor a=1, x=4 og n=2. Vi kalder nroot i kodeblok 1.20 fra Command Window.

```
>>nroot(1,4,2)
    ans =

         2
```

KODEBLOK 1.20 • DEN NYOPRETTEDE FUNCTION *nroot* UDREGNER HER $(1 \cdot 4)^{1/2}$ - ALTSÅ *a=1*, *x=4* OG *n=2*. *nroot* KALDES HER FRA COMMAND WINDOW.

Bemærk, at rækkefølgen af tallene i kodeblok 1.20 **ikke** er ligegyldig! Det første tal svarer nemlig til a det næste til b og det sidste til n!

Avancerede funktioner

Der er ikke grænser for, hvor avancerede man kan lave sine funktioner ud over sine egne evner. Vi giver her et bud en mere lidt mere avanceret udgave af den tidligere oprettede nroot funktion (se kodeblok 1.19).

Fra gymnasietiden bør det være jer bekendt, at man ikke må dividerer med nul. Noget som du muligvis ikke har haft om i gymnasiet er komplekse tal. Det forholder sig således, at hvis man tager kvadratroden af et negativt tal vil resultatet være et kompleks tal. Et kompleks tal har formen $a+bi$, hvor a er realdelen og b den imaginære del. i kaldes den imaginære enhed, for hvilket der gælder $i^2 = -1$. Dette vil du lære om i kurset MatIntro, hvilket I som nystartede fysikstuderende har sideløbende til Mek1.

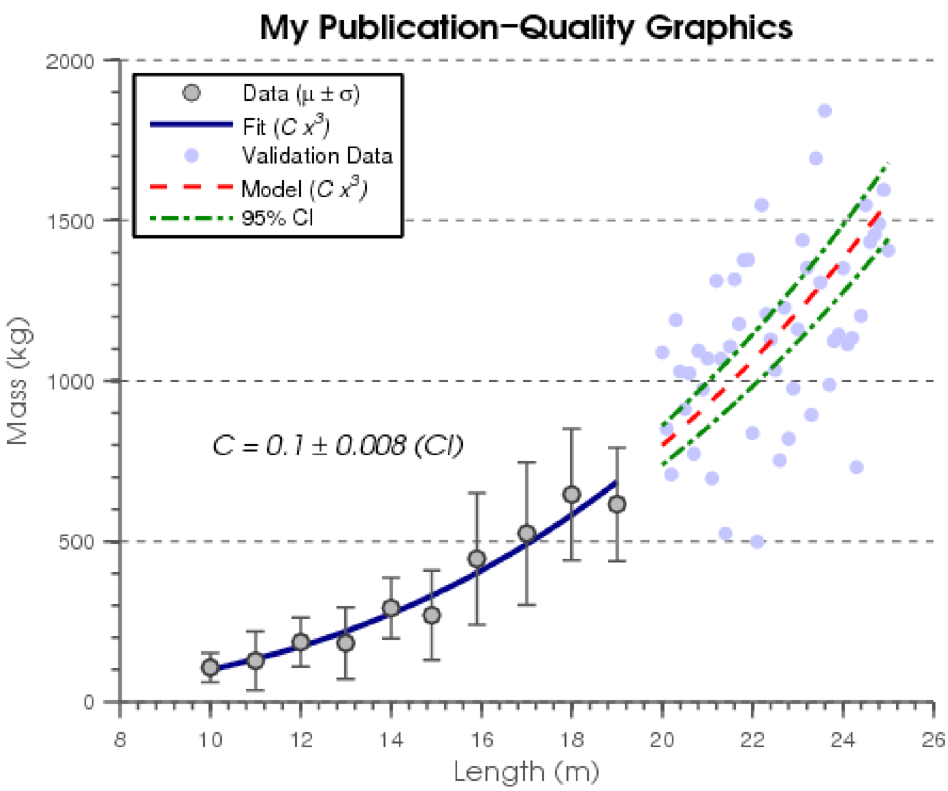
Vi er som udgangspunkt ikke interesseret i tilfældet, hvor resultatet af nroot er et kompleks tal - altså tilfældet hvor $a \cdot x < 0$ og ej heller hvis $n = 0$. Derfor vil nu modificere nroot funktionen således, at der allerførst tjekkes om de tallene, a, x og n, opfylder kravene til, at $a \cdot x < 0$ og $n = 0$. Dette gøres ved at tilføje en if-klausul. Se hvordan i kodeblok 1.21.

[6] $(1 \cdot 4)^{1/2} = 4^{1/2} = \sqrt{4} = 2$.

```
function [f]=nroots( a,x,n )

%Først tjekkes om a*x er mindre end nul
if (a*x<0)
%Hvis a*x er mindre en nul så outputtes stringen 'Kvadranten er negativ!' til Command Window
    disp('Kvadranten er negativ!')
elseif (n==0)
%Hvis n er lig med nul så outputtes stringen 'Der deles med nul i eksponenten!' til Command Window
    disp('Der deles med nul i eksponenten!')
else
%hvis både n er forskellig fra nul og a*x er positiv udregnes resultatet
f=(a*x).^(1/n);
end
```

KODEBLOK 1.21 • KODEN VISER, HVORDAN MAN MED EN `if`-KLAUSUL MODIFICERER DEN TIDLIGERE OPRETTEDE `nroot` FUNKTION SÅLEDES, AT HVIS RESULTATET ER ET KOMPLEKS TAL ELLER HVIS DER DIVIDERES MED NUL SÅ OUTPUTTES DER EN STRING TIL COMMAND WINDOW, SOM FORTÆLLER NETOP DETTE.

Som en test kan du definere funktionen `nroot` som i kodeblok 1.21 og eksempelvis prøve at skrive først `nroot(1,2,0)` og derefter `nroot(1,-4,2)`. Men inden da så overvej selv, hvad Matlab vil outputte!



2 Grafisk fremstilling af data

Kapiteloversigt

Vi skal i dette kapitel se, hvordan man fremstiller sine data grafisk i Matlab. Med andre ord vi skal plotte dem. Men det er ikke bare nok at plotte sine data, for et plot skal også være pænt og der skal være de nødvendige informationer så læseren ikke bliver forvirret. Vi kommer derfor til at gennemgå kommandoer i forbindelse med at style sit plot. Endvidere vil det også blive behandlet, hvordan man fremstiller flere dataserier i et og samme plot, men også separat. Tilsidst afslutter vi med, hvad et scatter plot er, og hvordan man kan få en fornemmelse for om der er en afhænighed mellem to målte størrelser.

2.1 At plotte sine data

Fremgangsmetoden er, at vi vil tage udgangspunkt i en konkret problemstilling og undervejs se, hvordan man plotter sine data. Problemstillingen vi vil tage udgangspunkt i er et fiktivt udført forsøg. Til forsøget har vi en relativ lufttæt beholder, der har et volumen $V = 0.5\,\mathrm{m}^3$ og indeholder atmosfærisk luft[7]. Til beholderen er tilsluttet et elektronisk termometer og trykmåler. Når beholderen opvarmes noteres temperaturen og trykket løbende. Data kan ses i tabel 2.2. Vi skal benytte Matlab til at plotte disse data.

Temperatur [K]	Tryk [Pa]
299.68	49862.66
310.83	52091.57
319.77	53397.68
330.26	54936.52
339.10	56836.74
347.84	58462.12
360.09	59023.77
369.05	61503.39
381.17	62943.02
391.74	63800.99
399.63	65815.72

TABEL 2.2 • DATA FRA ET (FIKTIVT) FORSØG. EN TÆTSLUTTENDE BEHOLDER MED VOLUMEN $V = 0.5\,\mathrm{m}^3$, INDEHOLDENDE ATMOSFÆRISK LUFT, OPVARMES. LØBENDE ER TEMPERATUREN OG TRYKKET I BEHOLDEREN NOTERET.

Det første vi skal gøre er, at indtaste data i to separate arrays, hvilket på nuværende tidspunkt ikke skulle være det store problem. Se kodeblok 2.22.

```
%Rydder hukommelsen
clear all

%Data lægges i hver deres respektive arrays
Temperatur = [299.68  310.83  319.77  330.26  339.10  347.84  360.09
      369.05  381.17  391.74  399.63];

Tryk = [49862.66  52091.57  53397.68  54936.52  56836.74  58462.12
    59023.77  61503.39  62943.02  63800.99  65815.72];
```

KODEBLOK 2.22 • DATA FRA FORSØGET (SE TABEL 2.2) LÆGGES I ARRAYS.

I vores plot ønsker vi, at de målte størrelser for tryk er op ad y-aksen og dem for tryk ud af x-aksen. For at plotte de nu indlæste data skal vi benytte `plot`-kommandoen. Denne har syntaksen

`plot(`⟨*array indeholdende x-akse værdier*⟩`,`⟨*array indeholdende y-akse værdier*⟩`).`

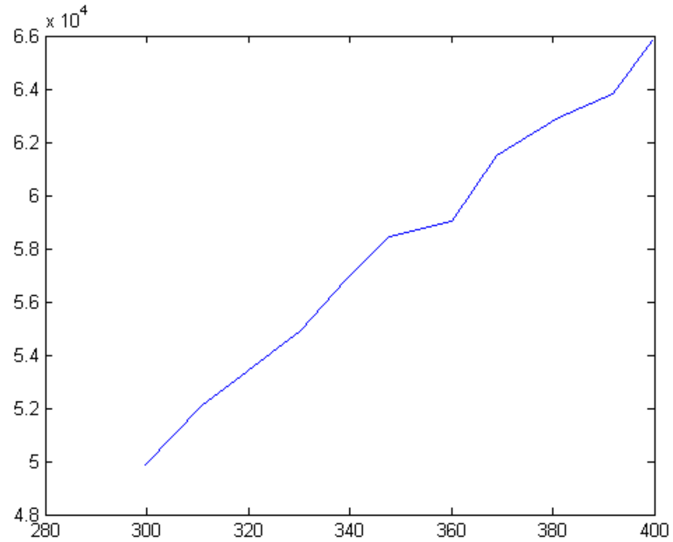
Vi tilføjer derfor en linje til vores script som indeholder `plot(Temperatur,Tryk)`. Samlet vil det se ud som i kodeblok 2.23.

[7] Senere kommer vi til at se, hvordan vi ved brug af grundlæggende statistik og dataanalyse kan bestemme stofmængden af den indesluttede atmosfæriske luft.

```
%Rydder hukommelsen
clear all

%Data lægges i hver deres respektive arrays
Temperatur = [299.68  310.83  319.77  330.26  339.10  347.84  360.09
      369.05  381.17  391.74  399.63];

Tryk = [49862.66  52091.57  53397.68  54936.52  56836.74  58462.12
    59023.77  61503.39  62943.02  63800.99  65815.72];

%Her plottes vores data
plot(Temperatur,Tryk)
```

KODEBLOK 2.23 • NÅR DATA ER LAGT I ARRAYS PLOTTES DE MED `PLOT`-KOMMANDOEN.

Nu kører vi scriptet i kodeblok 2.23, og resultatet vil være plottet som kan ses i figur 2.3.



FIGUR 2.3 • DETTE PLOT ER ET RESULTAT AF AT KØRE SCRIPTET I KODEBLOK 2.23, HVILKET IKKE ER DET KØNNESTE.

2.2 Styling af plots

Som det måske allerede er gået op for jer er plottet ikke særlig kønt og det er heller ikke muligt at se datapunkterne, eller hvad der er ud af akserne. Det er der heldigvis råd for!

Som standard viser Matlab ikke datapunkterne, men binder dem sammen med en ret linje mellem hvert datapunkt. Dette er ikke acceptabelt! Så lad os style plottet! Så vi udvider scriptet således, at vi

- giver plottet en titel,
- labels så man kan se hvad hhv. x- og y-aksen repræsenterer,
- kan se datapunkterne som røde ringe,
- ikke forbinder en ret linie mellem hvert datapunkt,
- sørger for en mindre afstand fra datapunkterne til randen af plottet.

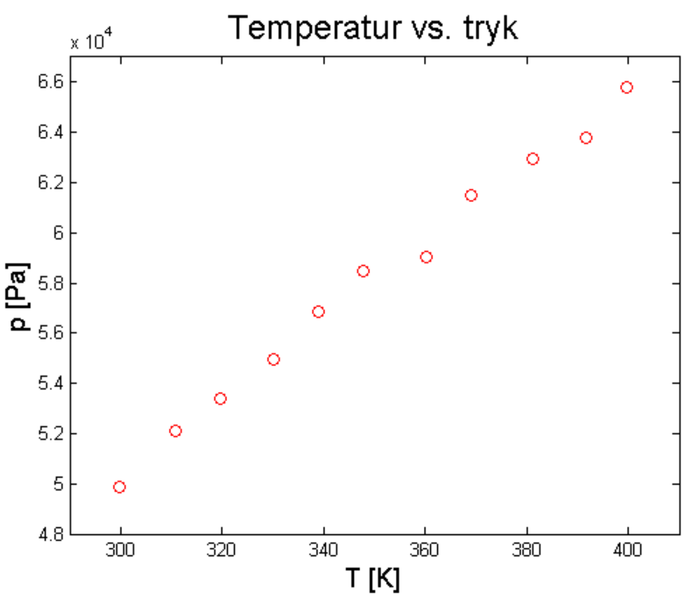
Koden til ovenstående styling kan ses i kodeblok 2.24.

```
%Rydder hukommelsen
clear all

%Data lægges i hver deres respektive arrays
Temperatur = [299.68  310.83  319.77  330.26  339.10  347.84  360.09
      369.05  381.17  391.74  399.63];

Tryk = [49862.66  52091.57  53397.68  54936.52  56836.74  58462.12
    59023.77  61503.39  62943.02  63800.99  65815.72];

%Her plottes vores data hvor datapunkter vises røde ringe
plot(Temperatur,Tryk,'ro')

%plottes titel med skriftstørrelse 16
title('Temperatur vs. tryk','fontsize',18);

%Labels på x- og y-akserne med skriftstørrelse 14
xlabel('T [K]','fontsize',14); %
ylabel('p [Pa]','fontsize',14);

%Angiv i hvilket interval der skal vises på akserne
xlim([290 410]);
ylim([4.8*10^4 6.7*10^4]);
```

KODEBLOK 2.24 • NÅR DATA ER LAGT I ARRAYS PLOTTES DE MED PLOT-KOMMANDOEN. HEREFTER STYLES PLOTTET SÅLEDES, AT DET FÅR EN TITEL, LABELS PÅ AKSERNE OG VISNINGSINTERVALLET ÆNDRES SÅLEDES, AT INGEN DATAPUNKTER LIGGER I RANDEN AF PLOTTET.

Kører vi det nye script i kodeblok 2.24 får vi et plot som vist i figur 2.4, hvilket må siges at være en stor forbedring! Nu er der angivet en titel så læseren ved, hvad plottet indeholder. Endvidere er det klart, hvad der er ud af de respektive akser, og der er ingen af datapunkterne som ligger i randen af plottet.

Der er naturligvis et hav af muligheder for styling af plots. Lad os indramme nogle af dem.



FIGUR 2.4 • DETTE PLOT ER ET RESULTAT AF AT KØRE SCRIPTET I KODEBLOK 2.24. DET FREMGÅR NU KLART HVAD PLOTTET VISER.

Styling af plots

Data fremstilles grafisk i MatLab via `plot`-kommandoen, der har følgende syntaks:

`plot(`⟨*array indeholdende x-akse værdier*⟩`,`⟨*array indeholdende y-akse værdier*⟩`,`⟨*option*⟩`)`,

hvor ⟨*option*⟩ kan antage følgende værdier[a].

⟨*option*⟩	Beskrivelse
`'o'`	Datapunkter vises som ringe.
`'x'`	Datapunkter vises som krydser.
`'+'`	Datapunkter vises som plus-tegn.
`'-'`	Forbinder punkter med linje.
`'-'`	Forbinder punkter med stiplet linje.
`':'`	Forbinder punkter med prikket linje.
`'r'`	Farver punkter og linjer rød.
`'b'`	Farver punkter og linjer blå.
`'k'`	Farver punkter og linjer sort.

Options kan sammensættes! Eksempelvis vil optionen `'-kx'` resulterer i, datapunkterne vises som sorte krydser forbundet med sort en stiplet linje.

Plottes titel kan angives ved `title`-kommandoen:

`titel(`⟨*Plottes titel som string*⟩`,`⟨*StyleOption*⟩`,`⟨*værdi*⟩`)`,

hvor ⟨*StyleOption*⟩`,`⟨*værdi*⟩ kan antage følgende værdier[b]:

⟨*StyleOption*⟩`,`⟨*værdi*⟩	Beskrivelse
`'fontsize',a`	Skriftstørrelsen sættes til tallet *a*.
`'fontweight','bold'`	Titelteksten er fed.
`'fontweight','italic'`	Titelteksten er kursiv.

Eksempel på brug: `title('fontsize',13,'fontweight','bold')` resulterer i titelteksten har skriftstørrelsen 13 og skrifttypen er fed.

Aksernes respektive labels kan angives ved hhv.:

`xlabel(`⟨*Aksens label som string*⟩`,`⟨*StyleOption*⟩`,`⟨*værdi*⟩`)`
`ylabel(`⟨*Aksens label som string*⟩`,`⟨*StyleOption*⟩`,`⟨*værdi*⟩`)`.

Endvidere kan aksernes visningsinterval sættes ved hhv.:

`xlim(`⟨*array indeholdende start- og slutværdi for intervallet*⟩`)`
`ylim(`⟨*array indeholdende start- og slutværdi for intervallet*⟩`)`.

[a] For flere mulige options så se MathWorks hjemmeside: `http://www.mathworks.se/help/matlab/ref/linespec.html`.
[b] For flere mulige options så se MathWorks hjemmeside: `http://www.mathworks.se/help/matlab/ref/title.html`.

2.3 Flere plots i det samme plot

En medstuderende har også udført det førnævnte fiktive forsøg, og har foretaget målingerne vist i tabel 2.3. I er nu alle interesseret i, hvordan jeres data ligger i forhold til hinanden. Hertil vil det være oplagt, at plotte jeres og den medstuderendes data i samme plot. Dette er muligt i Matlab.

Temperatur [K]	Tryk [Pa]
303.03	50128.83
310.73	52438.94
319.94	52054.58
330.71	55277.09
339.80	56667.38
349.88	57516.16
361.49	59615.20
371.41	61665.31
381.42	64945.20
390.67	66202.72
398.79	65805.06

TABEL 2.3 • DATA FRA DET TIDLIGERE NÆVNT (FIKTIVE) FORSØG SOM EN AF DINE MEDSTUDERENDE OGSÅ HAR UDFØRT. OPSTILLINGEN ER IDENTISK TIL DIN, HVILKET VIL SIGE EN TÆTSLUTTENDE BEHOLDER MED VOLUMEN $V = 0.5\,\mathrm{m}^3$, INDEHOLDENDE ATMOSFÆRISK LUFT, OPVARMES.

Ønsker man at plotte flere dataserier i samme plot skal man benytte sig af `hold on`-kommandoen. Denne kommando låser et eksisterende plot og gengiver det næste plot i samme plot. Plots gengives i det samme plot indtil kommandoen `hold off` er blevet givet.

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

Temperatur = [299.68  310.83  319.77  330.26  339.10  347.84  360.09  369.05  381.17  391.74  399.63];

Tryk = [49862.66  52091.57  53397.68  54936.52  56836.74  58462.12  59023.77  61503.39  62943.02  63800.99  65815.72];

%Nu indlæsses den medstuderendes data i arrays
TemperaturStudent = [303.03 310.73 319.94 330.71 339.80 349.88 361.49 371.41 381.42 390.67 398.79];

TrykStudent = [50128.83 52438.94 52054.58 55277.09 56667.38 57516.16 59615.20 61665.31 64945.20 66202.72 65805.06];

%Her plottes vores data hvor datapunkter vises som røde ringe
plot(Temperatur,Tryk,'ro')
%Nu låses plottet med hold on-kommandoen. Ergo vil det næste plot blive gengivet i det samme plot
hold on
%Nu plottes den medstuderendes data som blå krydser
plot(TemperaturStudent,TrykStudent,'bx')
%Der skal ikke plottes flere data i dette plot derfor angives hold off kommandoen
hold off
%plottes titel med skriftstørrelse 16
title('Temperatur vs. tryk','fontsize',18);
%Labels på x- og y-akserne med skriftstørrelse 14
xlabel('T [K]','fontsize',14);
ylabel('p [Pa]','fontsize',14);

%Angiv hvilket interval der skal vises på akserne
xlim([290 410]);
ylim([4.8*10^4 6.7*10^4]);
```

KODEBLOK 2.25 • ØNSKES FLERE PLOTS I DET SAMME PLOT SKAL `hold on`-KOMMANDOEN BENYTTES. MATLAB GENGIVER IKKE EFTERFØLGENDE PLOT I DET SAMME NÅR `hold off`-KOMMANDOEN ER BLEVET GIVET.

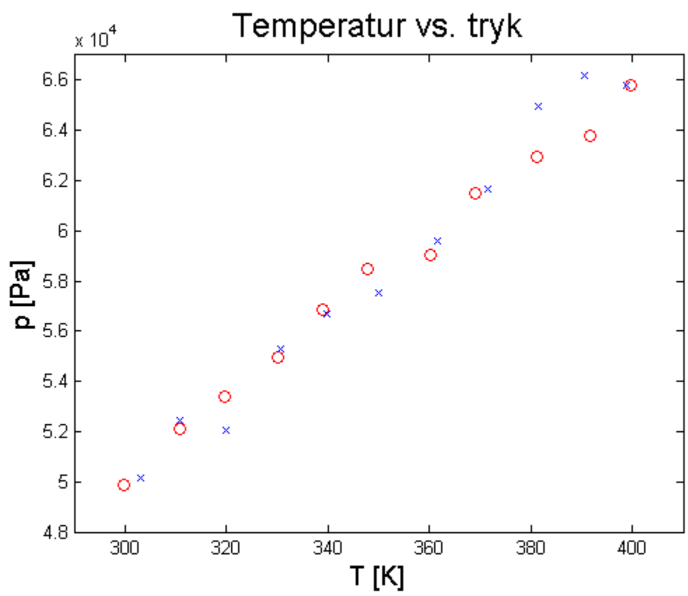
Lad os se, hvordan vi gengiver både jeres og den medstuderendes plot i et og samme. Hvordan dette gøres kan ses i kodeblok 2.25. Køres scriptet i kodeblok 2.25 vil Matlab outputte et plot som vist i figur 2.5.

Vi indrammer proceduren for at fremstille flere dataserier i samme plot.

> **Flere plots i det samme**
>
> I Matlab skal kommandoen `hold on` angives for at låse et plot. Plots som herefter genereres il blive gengivet i det plot som blev genereret før `hold` on blev angivet. For at låse plottet op angives kommandoen `hold off`. Se kodeblok 2.25 for eksempel på anvendelse.

2.4 Legends i plots

Vi er dog ikke helt færdige, for det fremgår ikke af plottet i figur 2.5 hvad de forskellige datapunkter betyder. Hertil skal vi lave en legend som netop beskriver dette. Dette kan gøres i Matlab ved at benytte `legend`-kommandoen.



FIGUR 2.5 • DETTE PLOT ER ET RESULTAT AF AT KØRE SCRIPTET I KODEBLOK 2.25. VI SER, AT JERES OG DEN MEDSTUDERENDES DATA NU ER FREMSTILLET I DET SAMME PLOT.

Legends i plots

For at angive legends i MatLab skal `legend`-kommandoen benyttes. Denne har følgende syntaks:

`legend(`⟨*Navn på første dataserie*⟩`,`⟨*Navn på anden dataserie*⟩`,'location',`⟨*Placerings option*⟩`).`

⟨*Placerings option*⟩ er en option som angiver hvor legenden skal placeres. De mest almindelige placerings options[a] er som følger:

⟨***Placerings option***⟩	**Beskrivelse**
`'NorthEast'`	Legend placeres øverst til højre i plottet.
`'NorthWest'`	Legend placeres øverst til venstre i plottet.
`'SouthEast'`	Legend placeres nederst til højre i plottet.
`'SouthWest'`	Legend placeres nederst til venstre i plottet.
`'Best'`	Legend placeres, hvor det er mest optimalt.

[a] For flere options se `http://www.mathworks.se/help/matlab/ref/legend.html`.

Lad os se, hvordan vi laver en legend til vores plot i figur 2.5. Dette hvilket er vist i kodeblok 2.26. Det resulterende plot er vist i figur 2.6, og det fremgår nu tydelig, hvad de forskellige typer af datapunkter betyder.

Det skal nævnes, at det lige som et plots titel og aksernes label er muligt at ændre skriftstørrelse, -farve og mange andre ting i legender. Vi vil ikke komme meget mere ind på dette. Men der henvises til MathWorks hjemmeside[8] for mere dybdegående information omkring dette, da det er relativt teknisk.

[8] `http://www.mathworks.se/help/matlab/ref/legend.html`

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

%Dine data indlæses i arrays
Temperatur = [299.68  310.83  319.77  330.26  339.10  347.84  360.09  369.05  381.17  391.74  399.63];

Tryk = [49862.66  52091.57  53397.68  54936.52  56836.74  58462.12  59023.77  61503.39  62943.02  63800.99  65815.72];

%Nu indlæses den medstuderendes data i arrays
TemperaturStudent = [303.03 310.73 319.94 330.71 339.80 349.88 361.49 371.41 381.42 390.67 398.79];

TrykStudent = [50128.83 52438.94 52054.58 55277.09 56667.38 57516.16 59615.20 61665.31 64945.20 66202.72 65805.06];

%Her plottes vores data hvor datapunkter vises som røde ringe
plot(Temperatur,Tryk,'ro')
%Nu låses plottet med hold on-kommandoen. Ergo vil det næste plot blive gengivet i det samme plot
hold on
%Nu plottes den medstuderendes data som blå krydser
plot(TemperaturStudent,TrykStudent,'bx')
%Der skal ikke plottes flere data i dette plot derfor angives hold off kommandoen
hold off
%plottes titel med skriftstørrelse 18
title('Temperatur vs. tryk','fontsize',18);
%Labels på x- og y-akserne med skriftstørrelse 14
xlabel('T [K]','fontsize',14);
ylabel('p [Pa]','fontsize',14);

%Angiv hvilket interval der skal vises på akserne
xlim([290 410]);
ylim([4.8*10^4 6.7*10^4]);

%For at vise læseren hvad de to dataserier betyder laver vi en legend. Vi vælger at placere den i øverst venstre hjørne
legend('Dine data','Medstuderendes data','location','NorthWest');
```

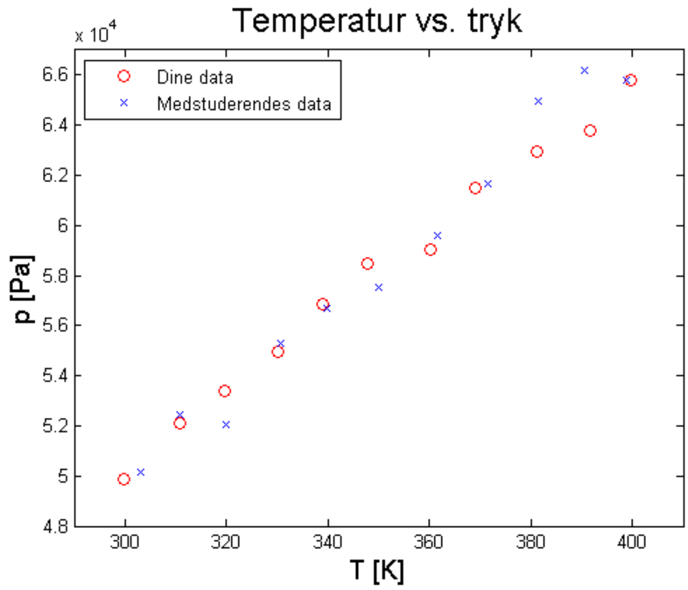
KODEBLOK 2.26 • NÅR DET I PLOTTET SKAL ANGIVES, HVAD DE FORSKELLIGE DATASERIER BETYDER BENYTTES LEGENDS. DETTE ANGIVES VIA KOMMANDOEN `legend`.

2.5 Flere plots separat

Alt afhængig af det foreliggende eksperiment vil behovet for flere forskellige plots måske være en realitet. Vi har indtil videre set, hvordan man laver pæne og overskuelige plots. Når man skal lave flere plots separat er det vigtig, at man fortæller Matlab, at den skal gøre et nyt plot-vindue klart og putte det nye plot i det.

Som en lille demonstration kan vi som eksempel betragte det tilfælde, at vi gerne vil have vist dine data fra det fiktive forsøg (se tabel 2.2) og din medstuderendes (se tabel 2.3) i hvert sit plot. Koden for at opnå netop dette er vist i kodeblok 2.27.

Som det fremgår af kodeblok 2.27 gives kommandoen `figure` lige inden et plot skal plottes i et nyt vindue. Bemærk dog, at hvert plots styling-options nu skal angives ved hvert plot! Et illustrativt eksempel, hvad der sker hvis `figure`-kommandoen ikke gives får I selv lov til at prøve. Derfor kør scriptet i kodeblok 2.27, men hvor de to `figure`-kommandoer er udkommenterede, og se hvad resultatet er. Hvor mange plots viser MatLab? Hvilket plot viser Matlab; jeres eller den medstuderendes?



Figur 2.6 • Dette plot er resultatet af at køre scriptet i kodeblok 2.26. Det fremgår nu, hvad de to forskellige dataserier betyder i legenden placeret i øverste venstre hjørne.

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

%Dine data indlæses i arrays
Temperatur = [299.68  310.83  319.77  330.26  339.10  347.84  360.09
      369.05  381.17  391.74  399.63];

Tryk = [49862.66  52091.57  53397.68  54936.52  56836.74  58462.12
    59023.77  61503.39  62943.02  63800.99  65815.72];

%Nu indlæses den medstuderendes data i arrays
TemperaturStudent = [303.03 310.73 319.94 330.71 339.80 349.88
    361.49 371.41 381.42 390.67 398.79];

TrykStudent = [50128.83 52438.94 52054.58 55277.09 56667.38 57516.16
     59615.20 61665.31 64945.20 66202.72 65805.06];

%Åbn det plot-vindue MatLab skal plotte dit plot i.
figure

%Her plottes dine data hvor datapunkterne vises som røde ringe
plot(Temperatur,Tryk,'ro')
%For dit plot skal styling-options nu angives
title('Temperatur vs. tryk','fontsize',18);
xlabel('T [K]','fontsize',14);
ylabel('p [Pa]','fontsize',14);
xlim([290 410]);
ylim([4.8*10^4 6.7*10^4]);

%Åbn et nyt plot-vindue MatLab skal plotte den medstuderendes plot i
figure

%Nu plottes den medstuderendes data som blå krydser
plot(TemperaturStudent,TrykStudent,'bx')
%Hans plots styling-options
title('Temperatur vs. tryk','fontsize',18);
%Labels på x- og y-akserne med skriftstørrelse 14
xlabel('T [K]','fontsize',14);
ylabel('p [Pa]','fontsize',14);
xlim([290 410]);
ylim([4.8*10^4 6.7*10^4]);
```

Kodeblok 2.27 • Ved flere plots separat skal man direkte fortælle Matlab, at det er det resultat man gerne vil opnå. Dette gøres ved give kommandoen `figure` lige inden `plot`-kommandoen.

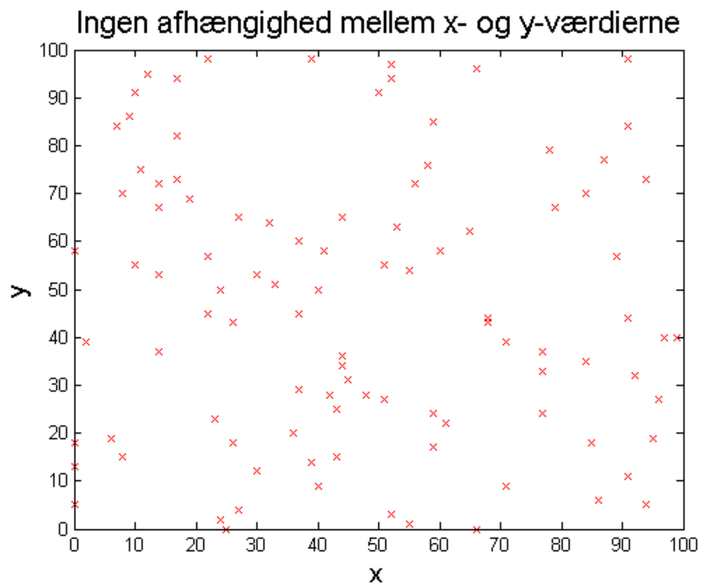
Vi indrammer proceduren for `figure`-kommandoen:

Flere plots separat

Når man skal genererer flere plot separat skal kommandoen `figure` gives lige inden man anvender `plot`-kommandoen. Bemærk, at hvert plot skal styles separat.

2.6 Scatter plots

Scatter plots - spredningsdiagram på dansk, men ordet scatter plot anvendes oftest også på dansk - er et diagram, som har til formål at klarlægge hvorvidt der er tale om afhængighed mellem to størrelser eller ej. Egentlig har vi faktisk allerede lavet scatter plots i form af de hidtil viste plots. Eksempelvis, er plottet i figur 2.6 et scatter plot, da denne via de målte værdier viser den matematiske sammenhæng mellem trykket p og temperaturen T. Det fremgår, at der højst sandsynligt er tale en ret liniet sammenhæng i begge tilfælde. Havde datapunkterne i stedet ligget vilkårligt i plottet er der ingen sammenhæng mellem trykket og temperaturen. Et eksempel på et scatter plot som viser, at der er ingen sammenhæng mellem værdierne på x- og y-aksen er vist i figur 2.7. Datapunkterne i plottet er ikke ægte data fra et forsøg eller lignende, men simuleret med Matlab.

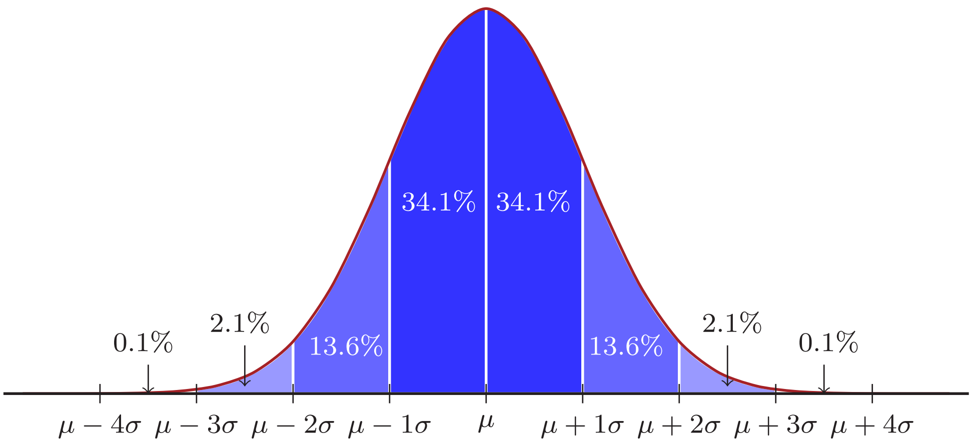


FIGUR 2.7 • FIGUREN VISER ET EKSEMPEL PÅ, HVORDAN ET SCATTER PLOT KAN SE UD, NÅR DER IKKE ER EN SAMMENHÆNG MELLEM TO STØRRELSER. BEMÆRK DATAPUNKTERNE ER SIMULERET!

Scatter plot

Et scatter plot er et matematisk visuelt redskab til at klarlægge, om der er en sammenhæng mellem to størrelser. Den størrelse, som er repræsenteret ud ad x-aksen, kaldes den *uafhængige variabel* og den op ad y-aksen den *afhængige variabel*.

I Matlab er proceduren for at lave et scatter plot det samme som man skal lave et plot - altså med den allerede behandlede `plot`-kommando. Hvis der er en sammenhæng vil man se datapunkterne følge en bestemt type matematisk funktion, og hvis punkterne ligger vilkårligt er der ingen sammenhæng. Se hhv. figur 2.6 og 2.7.



3 Grundlæggende dataanalyse

Kapiteloversigt

For at kunne analysere data er statistik en nødvendighed. Derfor skal vi i dette kapitel introducere basale statistiske begreber, og hvordan man i Matlab anvender disse begreber i forbindelse med dataanalyse. Først snakker vi kort om middelværdi og spredning. Herefter introduceres histogrammer og i forlængelse heraf normalfordelingen. Derefter snakker vi om det vi - som fysikere - går rigtig meget op i - nemlig usikkerheder i målinger og hvordan disse fremstilles grafisk sammen med målingerne i et plot i Matlab. Gennem et større eksempel ser så hvordan fitting af en ret linje til data gøres. Til sidst ser vi hvordan man indlæser data fra en ekstern fil i Matlab.

3.1 Middelværdi og spredning

Når man som fysiker, eller anden person med anderledes naturvidenskabelig baggrund, har foretaget et eksperiment skal data analyseres. I analysen forsøger man at påvise en eller anden form for hypotese. Dette kunne eksempelvis være, om sammenhængen mellem to målte størrelser er eksponentiel. Der er ingen vej uden om: Den eneste måde, at påvise sin hypotese er gennem statistik!

Selve den matematiske teori omkring statistik er yderst langhåret, og - for at være helt ærlig - gabende kedsommelig! Anvendelsen af statistik i forbindelse med eksperimenter derimod, synes jeg, er ret spændende, og det er præcis det vi skal arbejde os frem imod - nemlig anvendelsen. Så den matematiske teori bekymrer vi os ikke om, og lader matematikkerne om det.

Når vi udfører eksperimenter forholder det sig naturligvis således, at jo mere data desto mere præcision. Derfor for at være 100% præcise ville vi skulle udføre vores eksperiment uendeligt mange gange! Dette er naturligvis ikke muligt. Men sæt nu for et øjeblik, at vi rent faktisk har udført forsøget uendeligt mange gange, så kalder vi middelværdien af dette uendelige store datasæt for *den sande middelværdi* og betegnes μ_s. Helt analogt kaldes spredningen af datasættet blive benævnt *den sande spredning*, og benævnes σ_s.

Problemet er hvordan vi så skal forholde os, da det ikke er muligt at udføre eksperimenter uendeligt mange gange. Svaret hertil er, at vi ser det datasæt som vi får fra vores eksperiment som værende en stikprøve fra det førnævnte uendelig store datasæt. Denne stikprøve er naturligvis endelig, og det kan vi forholde os til!

MIDDELVÆRDI: Det forholder sig således, at det bedste estimat - det vil sige det bedste bud - for den sande middelværdi μ_s er givet ved det aritmetiske gennemsnit $\overline{x}$. Denne er defineret ved

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{x_1 + x_2 + x_3 + \cdots + x_N}{N}, \tag{3.1}$$

hvor N er antallet er datapunkter.

SPREDNING: Datasættes spredning - standard deviation på engelsk og mange gange forkortet std - er formenlig et begreb du endnu ikke er stødt på. Spredningen af datasættet tolkes som værende et mål for, hvor stor afvigelse hvert datapunkt gennemsnitlig ligger fra middelværdien. Denne betegnes σ_x, og er defineret ved

$$\sigma_x = \sqrt{\frac{1}{N-1}\cdot\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2}, \tag{3.2}$$

hvor N igen er antallet af datapunkter i datasættet. σ_x er det bedste estimat for spredningen af det førnævnte uendelige store datasæt σ_s.

Der findes et hav af andre statistiske formler, men de eneste vi her kan nøjes med er blot middelværdi og spredning givet i hhv. (3.1) og (3.2).

Lad os nu se, hvordan vi i Matlab udregner middelværdien og spredningen af et datasæt. Hertil benytter vi et lille datasæt bestående af fem målinger[9] af tyngdeaccelerationen - nemlig $9.9\,\mathrm{m/s^2}$, $9.6\,\mathrm{m/s^2}$, $9.5\,\mathrm{m/s^2}$,

[9] Bemærk, at de datasæt I kommer til at sidde med formentlig aldrig vil være så små i og med, at de statistiske konklusioner drages på baggrund af et datasæt med dårlig opløsning. Tænk på det som netop opløsningen af en computerskærm: Hvis jeg viser jer et billede, som har en opløsning på 5px·5px, og beder jer om at fortælle mig hvad billedet ligner vil svaret sikkert blot være nogle prikker. Vises samme billede med en opløsning på eksempelvis 600px·600px vil man være bedre i stand til, at se hvad billedet viser. Det er præcist det samme med datasæt; jo flere datapunkter desto bedre opløsning.

$9.7\,\mathrm{m/s^2}$ og $9.8\,\mathrm{m/s^2}$. Vi ønsker at bestemme middelværdi og spredning i Matlab. Kommandoerne, som hertil skal benyttes er indbygget i Matlab, og kan nemt huskes hvis man kender de engelske betegnelser for middelværdi (mean) og spredning (standard deviation forkortet std). De er nemlig hhv. `mean` og `std`. Kodeblok 3.28 viser, hvordan datasættets - det indeholdende fem målinger for tyngdeaccelerationen - middelværdi og spredning udregnes i Matlab.

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

%Data indlæses i et array
Data = [9.9 9.6 9.5 9.7 9.8];

%Middelværdien udregnes vha. kommandoen mean
Snit = mean(Data)

%og spredningen udregnes vha. kommandoen std
Spredning = std(Data)
```

Kodeblok 3.28 • Kommandoerne *`mean`* og *`std`* udregner hhv. middelværdien og spredningen af et datasæt. Her er datasættet lagt i et array kaldet `Data`.

Middelværdi og spredning

Matlab har indbyggede funktioner, som outputter et arrays middelværdi og spredning via formlerne angivet i hhv. eq. (3.1) og (3.2). Funktionerne er hhv. `mean` og `std`, og har følgende syntaks:

`mean(`⟨*Navn på array indeholdende dataserie*⟩`)`

`std(`⟨*Navn på array indeholdende dataserie*⟩`)`.

Se kodeblok 3.28 for eksempel på anvendelse.

3.2 Histogrammer

Vi har nu set, hvordan man både i hånden og i Matlab kan udregne middelværdien og spredningen af et datasæt, som var de bedste estimater for hhv. den sande middelværdi μ_s og spredning σ_s. Nu skal vi koble disse to begreber sammen med statistiske fordelinger.

Det første vi dog er nødsaget til at få på plads er hvad et histogram er, og hvordan man i Matlab genererer sådan et. Forestil jer, at vi har følgende tal

$$7,5,6,4,2,3,6,9,8,4,5,6,2,1,3,4,5,7,8,5,10. \tag{3.3}$$

Disse tal vil vi gerne genererer et histogram af. Fremgangsmåde er følgende

- Identificer hhv. den laveste og højeste værdi (I dette tilfælde er det hhv. 1 og 10).
- Inddel det lukkede interval afgrænset af hhv. den laveste og højeste værdi i N lige store delintervaller (da vi har med hele tal at gøre har vi i dette tilfælde valgt 10 delintervaller hver med længden 1). Antallet af lige store delintervaller kaldes for *bins* (se tabel 3.4). Hele dette skridt kaldes *binning af data*.
- Identificer nu antallet af datapunkter som ligger i hver bin (Se tabel 3.4 for dette konkrete tilfælde). Dette kaldes for *hits*.

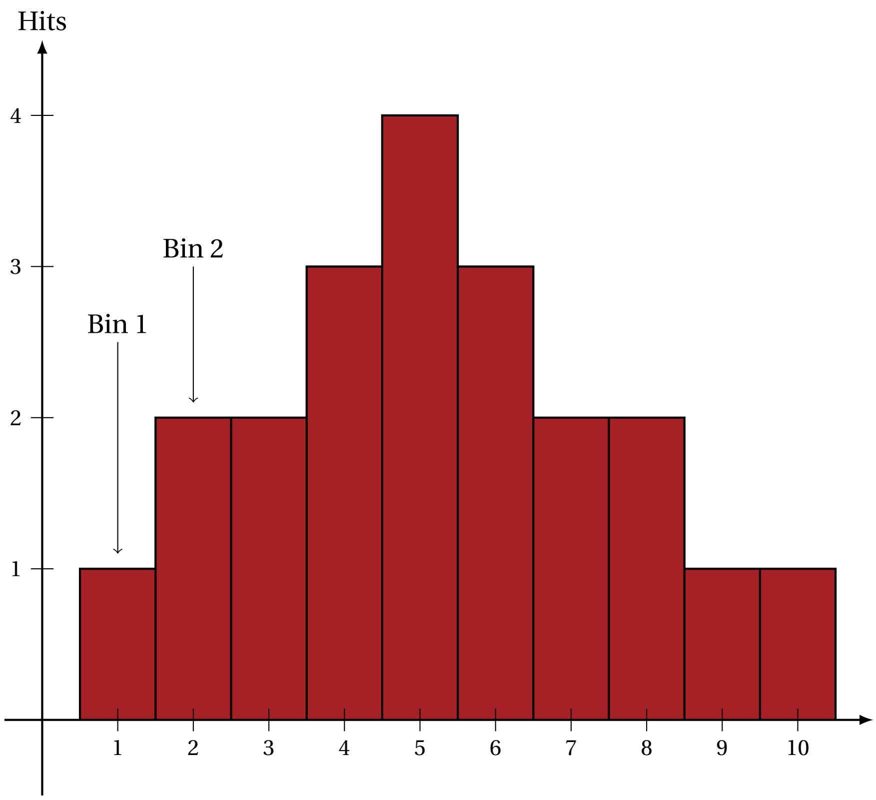
- Indtegn nu binsne i et koordinatsystem, hvor hits er op ad y-aksen og x-aksen er som vanelig - altså $1,2,3,4,5,6,...osv.$. Binsne repræsenteres af rektangler med bredden svarende til længden af delintervallet. De individuelle bins har højden svarende dens hits. Se figur 3.8.

Bin	Interval	Hits
1	$[0.5, 1.5[$	2
2	$[1.5, 2.5[$	2
3	$[2.5, 3.5[$	3
4	$[3.5, 4.5[$	3
5	$[4.5, 5.5[$	4
6	$[5.5, 6.5[$	3
7	$[6.5, 7.5[$	3
8	$[7.5, 8.5[$	2
9	$[8.5, 9.5[$	2
10	$[9.5, 10.5]$	2

TABEL 3.4 • TABELLEN VISER, HVORDAN VI HAR INDDELT x-AKSEN I BINS OG HITS I HVER BIN.

Bemærk, at hvis man summer hitsne fra hver bin er resultatet antallet af målinger. I dette tilfælde havde vi 21 tal (se eq. (3.3)). Summer vi nu hitsne fra hver bin ser vi

$$\sum \text{Hits} = 1+2+2+3+4+3+2+2+1+1 = 21. \tag{3.4}$$

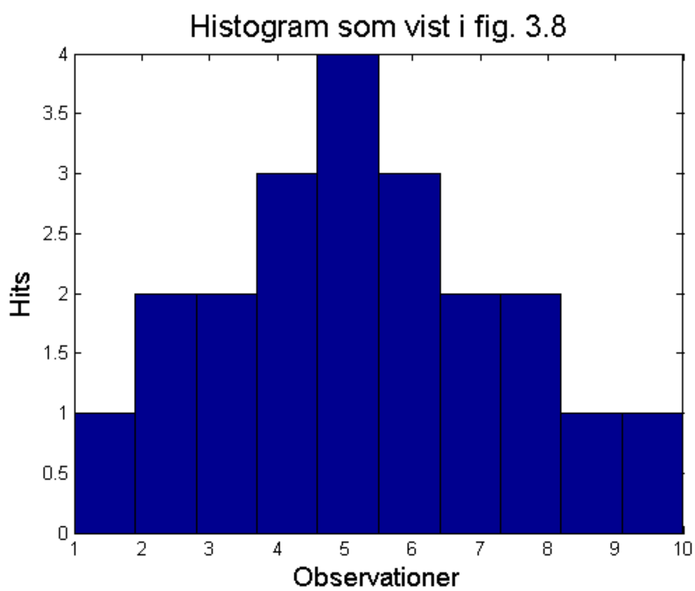


FIGUR 3.8 • HISTOGRAM OVER MÅLINGER FRA EQ. (3.3). HVER BIN HAR LÆNGDEN 1. HVER BINS HØJDE AFGØRES ALT EFTER DERES RESPEKTIVE HITS - ALTSÅ ANTALLET AF OBERSERVATIONER I HVER BIN.

Jeg vil anbefale, at prøve at lave et par histogrammer i hånden for at få en føling med fremgangsmetoden. Når dette er sagt så lad os se, hvordan vi på ingen tid kan gøre det samme i Matlab. I Matlab skal kommandoen `hist` benyttes for at generere et histogram over en dataserie indlæst i et array. Kodeblok 3.29 viser hvordan og figur 3.9 viser det histogram Matlab vil outputte.

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

%Data indlæses i et array
Dataserie = [7 5 6 4 2 3 6 9 8 4 5 6 2 1 3 4 5 7 8 5 10];

%Åbn plotvindue til det kommende histogram
figure
%Histogram genereres med hist-kommandoen og der angives specifikt,
    at vi ønsker ti bins
hist(Dataserie,10);
%histogram styling
title('Histogram som vist i fig. 3.8','fontsize',16);
ylabel('Hits','fontsize',14);
xlabel('Observationer','fontsize',14);
```

KODEBLOK 3.29 • HISTOGRAMMER AF SINE DATASERIE ER HURTIGT LAVET I MATLAB! MAN SKAL BLOT LÆGGE SINE MÅLINGER I ET ARRAY OG DEREFTER BENYTTE KOMMANDOEN *hist*. HISTOGRAMMER I MATLAB KAN STYLES MED TITEL, X- OG Y-LABELS SOM ANGIVET I BOKSEN PÅ SIDE 28.



FIGUR 3.9 • DETTE HISTOGRAM ER RESULTATET AF AT KØRE SCRIPTET I KODEBLOK 3.29.

Sammenligner vi histogrammet som vi selv genererede (figur 3.8) med det, som er genereret med Matlab bemærker vi (figur 3.9) ses det, at Matlab ikke benytter samme delintervaller for binsne som vi gjorde (se tabel 3.4), men formen af histogrammerne er den samme. Hvis det ønskes kan man bede Matlab om at centrere binsne omkring bestemte værdier, men dette vil vi ikke her gennemgå[10].

[10] Se `http://www.mathworks.se/help/matlab/ref/hist.html` for nærmere information omkring dette.

Histogrammer i Matlab

Histogrammer over en dataserie i Matlab kan genereres ved at benytte kommandoen `hist`, som har følgende syntaks:

`hist(`⟨*Navn på array indeholdende data*⟩,⟨*antal bins*⟩`)`.

Bemærk, at der skal klargøres et plotvindue som histogrammet skal ilægges - altså kommandoen `figure` skal angives inden man kalder `hist`. Angives ⟨*antal bins*⟩ ikke vil Matlab som standard sætte ⟨*antal bins*⟩ = 10.

Histogrammet kan styles med titel koordinatakse-labels som beskrevet i boksen på side 28. For mere dybdegående information omkring `hist` se `http://www.mathworks.se/help/matlab/ref/hist.html`.

Bemærk, at valget af antallet af bins er en personlig sag - altså der findes ikke den helt store gyldne regel. Det eneste der er at sige hertil er, at de hverken må være for eller for små. Sagt på en anden måde: Du må prøve dig frem.

3.3 Normalfordelingen

Vi er nu kommet dertil, hvor begrebet statistiske fordelinger kan introduceres. Der findes et hav af fordelinger: Poisson-, binomial-, beta-, χ^2-, normal-fordelingen og mange mange andre[11]. Den eneste fordeling som i denne sammenhæng er relevant er den vigtigste af dem alle - nemlig normalfordelingen eller Gaussfordelingen[12] som den også kaldes.

Normalfordelingen er givet ved den kontinuerte funktion

$$f(x;\sigma,\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-(x-\mu)^2/(2\sigma^2)}, \quad x \in \mathbb{R}, \tag{3.5}$$

hvor σ er spredningen og $x = \mu$ er der hvor funktionen har sit maksimum. Se figur 3.10. Det er ikke et tilfælde at σ og μ er benyttet, for hvis man laver et histogram over sine målinger og kan se, at histogrammet tilnærmelsesvis har samme form som en normalfordeling (hvilket de med ret stor sandsynlighed har i jeres tilfælde) så siger vi, at målingerne er normalfordelte med spredningen $\sigma \approx \sigma_x$ og middelværdi $\mu \approx \overline{x}$ (se ligning (3.1) og (3.2)). Normalfordelingens middelværdi $x = \mu$ er der, hvor den er symmetrisk omkring og også har sit maksimum. Se figur 3.10. Hvordan formen af fordelingen afhænger af σ er illustreret i figur 3.11. Det fremgår tydeligt, at jo større σ er desto mere bred er kurven - altså kurven har større spredning.

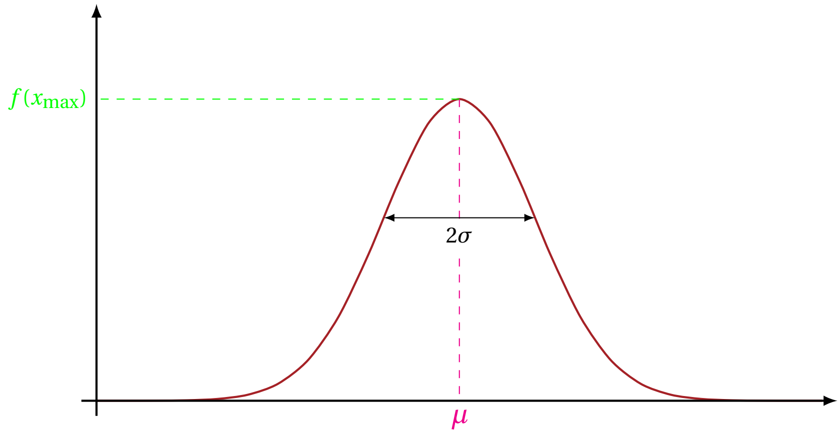
I boksen Histogrammer i Matlab øverst nævnes det, at valget af antallet af bins i histogrammer er en personlig sag, og man må derfor prøve med forskellige antal. En hjælp til at finde det optimale antal af bins består i, at der samtidig med generering af et histogram fittes[13] en normalfordeling. Når histogrammet fylder mest muligt af arealet under kurven af den fittede normalfordeling er det optimale antal bins nået.

Denne procedure er mulig at foretage i Matlab ved brug af kommandoen `histfit`. Denne kommando genererer et histogram over data indlæst i et array og samtidig fitter en valgt fordeling til histogrammet og viser denne sammen med histogrammet.

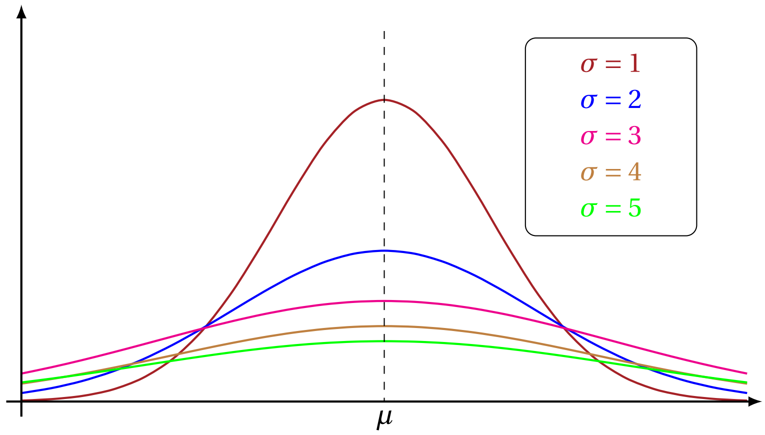
[11] Information om mange af disse fordelinger kan findes i både [Taylor, 1997] og [Barlow, 1999].

[12] Efter den tyske matematikker Carl Friedrich Gauss som udledte den. Dog har andre uafhængig af Gauss også udledt den. Eksempelvis, den franske matematikker Pierre-Simon Laplace.

[13] Et fit betyder blot den bedste kurve ud fra den funktion som man har valgt skal udgøre kurven. Her er det så normalfordelingen.



FIGUR 3.10 • FIGUREN VISER NORMALFORDELINGEN. DEN ER SYMMETRISK OMKRING $x = \mu$, OG HAR OGSÅ SIT MAXIMUM HER.



FIGUR 3.11 • FEM NORMALFORDELINGER MED FORSKELLIG SPREDNING. BEMÆRK, AT JO STØRRE SPREDNINGEN ER DESTO BREDERE ER FORDELINGEN. FORDELINGERNE PEAKER DOG STADIG DET SAMME STED - NEMLIG VED $x = \mu$ SOM ER FORDELINGERNES MIDDELVÆRDI μ. ERGO DISSE FEM FORDELINGER HAR SAMME MIDDELVÆRDI MEN FORSKELLIG SPREDNING

Histogram med fit

I Matlab kan der genereres et histogram, hvori en fittet normalfordeling samtidig vises. Dette opnås ved brug af kommandoen `histfit`, der har syntaksen

`histfit(`⟨*Array indeholdende data*⟩,⟨*antal bins*⟩,⟨*fordeling*⟩`)`.

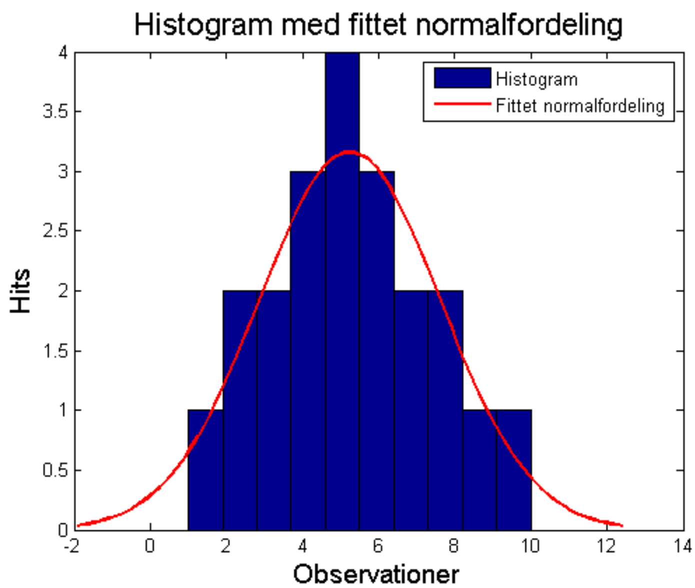
For at fitte en normalfordeling til histogrammet skal ⟨*fordeling*⟩ = `'normal'`, men der kan fittes andre statistiske fordelinger[a]. Antallet af bins *skal* her angives og kan ikke som i kommandoen `hist` undlades. Bemærk, at der skal klargøres et figurvindue med `figure`. Se kodeblok 3.30 for et eksempel på anvendelse.

[a] For at se hvilke statistiske fordelinger der kan fittes og for nærmere information omkring `histfit`-kommandoen se `http://www.mathworks.se/help/stats/histfit.html`.

Kodeblok 3.30 viser, med vi udgangspunkt i tallene i eq. (3.3), hvordan kommandoen `histfit` benyttes, og det resulterende histogram med den fittede normalfordeling kan ses i figur 3.12.

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

%Data indlæses i et array
Dataserie = [7 5 6 4 2 3 6 9 8 4 5 6 2 1 3 4 5 7 8 5 10];

%Åbn plotvindue til det kommende histogram med fittet
    normalfordeling
figure
%Histogram med ti bins og fittet normalfordeling genereres med
    histfit-kommandoen
histfit(Dataserie,10,'normal');
%histogram styling
title('Histogram med fittet normalfordeling','fontsize',16);
ylabel('Hits','fontsize',14);
xlabel('Observationer','fontsize',14);
legend ('Histgram','Fittet normalfordeling','location','NorthEast');
```

KODEBLOK 3.30 • FOR AT GENERERE ET HISTOGRAM MED FITTET NORMALFORDELING BENYTTES KOMMANDOEN ***histfit***. KODEN VISER ET EKSEMPEL PÅ BRUGEN AF KOMMANDOEN. DEN RESULTERENDE FIGUR MATLAB OUTPUTTER VED KØRSEL AF DENNE KODE KAN SES I FIGUR 3.12.



FIGUR 3.12 • DETTE HISTOGRAM ER RESULTATET AF AT KØRE SCRIPTET I KODEBLOK 3.30.

Som I måske nok har bemærket får vi vist histogrammet med den fittede normalfordeling når man benytter `histfit`-kommandoen, men hvad er så den fittede normalfordelings middelværdi og spredning? Se, `histfit`-kommandoen er så "smart" programmeret, at den slet ikke giver mulighed for at få adgang til detaljerne af den fittede normalfordeling! Heldigvis er der råd for dette, for `histfit` benytter nemlig en anden kommando til at fitte normalfordelingen, hvilket er `fitdist`. Derfor for at få adgang til detaljerne af den fittede normalfordeling skal denne anvendes.

Adgang til normalfordelingens fit

For at få adgang til det fit som `histfit` laver skal kommandoen `fitdist` benyttes. Denne har syntaksen

`fitdist(`⟨*array indeholdende data*⟩,⟨*fordeling*⟩`)`.

Ved fit af en normalfordeling skal ⟨*fordeling*⟩ = 'normal'. Lægges fittet af normalfordelingen i en variabel - altså

⟨*variabel*⟩ = `fitdist(`⟨⟨*(række)array indeholdende data*⟩`(:)`,'normal'`)`,

så kan dens middelværdi og spredning udtrækkes således

middelværdi: ⟨*variabel*⟩.mu og spredning: ⟨*variabel*⟩.sigma

Se kodeblok 3.31 for eksempel på anvendelse.

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

%Data indlæses i et (række)array
Dataserie = [7 5 6 4 2 3 6 9 8 4 5 6 2 1 3 4 5 7 8 5 10];

%Åbn plotvindue til det kommende histogram med fittet
    normalfordeling
figure
%Histogram med ti bins og fittet normalfordeling genereres med
    histfit-kommandoen
histfit(Dataserie,10,'normal');
%histogram styling
title('Histogram med fittet normalfordeling','fontsize',16);
ylabel('Hits','fontsize',14);
xlabel('Observationer','fontsize',14);
legend ('Histgram','Fittet normalfordeling','location','NorthEast');

%Få adgang til den fittede normalfordeling middelværdi og spredning
%Da Dataserie er et rækkearray skal det laves om til et kolonnearray
     (vektor) dette gøres ved at skrive Dataserie(:)
MyFit = fitdist(Dataserie(:),'normal')
%Lægger fittes middelværdi og spredning over i variabler
middelfit = MyFit.mu;
spreadfit = MyFit.sigma;
```

KODEBLOK 3.31 • NÅR DEN FITTEDE NORMALFORDELINGS MIDDELVÆRDI OG SPREDNING ØNSKES BESTEMT SOM ***histfit*** LAVER SKAL KOMMANDOEN ***fitdist*** BENYTTES. KODEN VISER ET EKSEMPEL PÅ BRUGEN AF KOMMANDOEN.

3.4 Usikkerheder

Indtil videre har der ikke været den store anvendelse involveret, og ej heller hvad vi dog skal bruge alt det foregående til! Dette kommer vi til nu. For nu skal vi benytte det til at snakke om usikkerheder i målinger.

Uanset, hvordan vi vender og drejer det så forholder det sig altså således, at vi lever i en virkelighed, hvor det på ingen måde er, og heller aldrig vil være, muligt at foretage en måling med 100% præcision[14]. Dette skyldes blandt andet, at vores måleinstrumenter har en begrænset opløsning. Valget af måleinstrumenter

[14] På det kvantemekaniske plan fortæller Heisenbergs usikkerhedsprincip, at det er umuligt!

har dog stor betydning for præcisionen! Eksempelvis, hvis vi skulle måle længden af en genstand ville en lineal med millimeter opdeling være at foretrække frem for en med kun centimeter opdeling!

Eksperimentelt resultat

Sæt, at alle N første års studerende har foretaget en måling x_i, alle med samme måleinstrument, af et objekts længde, eksempelvis længden af kateteret i auditoriet. Målingerne er foretaget således, at ingen af de N første års studerende ved, hvad alle de andre har målt længden til. Dette kaldes for *uafhængige målinger*[15]! Vi har således et datasæt bestående N uafhængige målinger. Det kan med god samvittighed antages, at disse er normalfordelte. På baggrund af disse normalfordelte målinger vil det bedste estimat for den sande middelværdi μ_s - det vil sige den helt præcise længde af kateteret - af målingerne x_i være $\overline{x}$. Men der må være en usikkerhed forbundet hermed, da alle N målinger ikke resulterede i lige præcis $\overline{x}$, hvilket i øvrigt er statistisk umuligt, da de er uafhængige. Det er her vores tidligere antagelse kommer i spil: De N målinger er normalfordelte med spredning $\sigma = \sigma_x$ (og middelværdi $\mu = \overline{x}$), og husk, at spredningen er mål for, hvor meget hver af de N målinger gennemsnitlig afviger fra middelværdien $\overline{x}$. Derfor kan vi benytte σ_x som et mål for den gennemsnitlige usikkerhed på en enkelt måling. Af konventionelle årsager skrives *usikkerheden i x* med symbolerne δx - altså $\sigma_x = \delta x$. Så et indledende eksperimentelt resultat på kateterets længde på baggrund af de N målinger er $\overline{x} \pm \delta x$.

Det er vigtigt at bemærke, at $\overline{x} \pm \delta x$ *ikke* betyder, at alle målinger - både dem foretaget og eventuelle fremtidige - kun vil ligge i intervallet $[\overline{x} - \delta x, \overline{x} + \delta x]$! Det skal tolkes som, at cirka 68% af alle målinger vil ligge i intervallet $[\overline{x} - \delta x, \overline{x} + \delta x]$. Men hvorfor lige 68%? Det kommer af, at målingerne er normalfordelte og derfor vil chancen for, at en måling vil ramme et givent interval være dikterede af normalfordelingen. Se figur 3.13.

Det er tidligere nævnt, at alle de målinger vi foretager os anses for at være en stikprøve af et uendeligt stort datasæt med spredning σ_s og middelværdi μ_s. Men vi kan aldrig komme til kende hverken σ_s eller μ_s! Heldigvis er det muligt for os gennem statistik, at kunne sige noget om i hvilket interval μ_s ligger, og dette vil netop være vores usikkerhed på $\overline{x}$. Det viser sig, at sålænge målingerne i stikprøverne er normalfordelte vil der være cirka 68% chance for, at μ_s ligger inden for $\overline{x} \pm \delta\overline{x}$ - altså i intervallet $[\overline{x} - \delta\overline{x}, \overline{x} + \delta\overline{x}]$, hvor

$$\delta\overline{x} = \frac{\sigma_x}{\sqrt{N}}. \tag{3.6}$$

Derfor kan vi i tilfældet med kateterets længde nu konkludere, at dens sande længde μ_s ligger i det førnævnte interval - altså $\overline{x} \pm \delta\overline{x}$. Dette er det endelige eksperimentelle resultat. Det skal tolkes som, at kateterets sande længde μ_s har cirka 68% chance for at ligge i intervallet $[\overline{x} - \delta\overline{x}, \overline{x} + \delta\overline{x}]$. Vi har således bestemt kateterets længden *inden for* 1σ, hvilket er den præcision I for det meste kommer til at opgive jeres eksperimentelle resultater med.

Værdien, som udregnes ved brug af eq. (3.6), kaldes for *standard afvigelsen af middelværdien* (engelsk: standard deviation of the mean, forkortet SDOM). Vi vil ikke komme ind på det teoretiske fundament for påstanden. Ergo må I blot - indtil videre - tage den som værende sand.

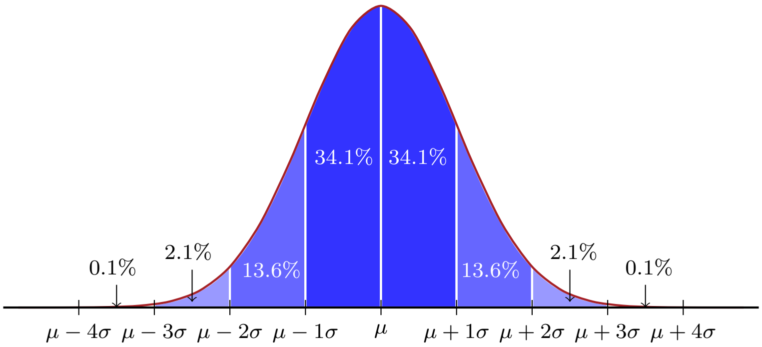
Det er værd at bemærke, hvad eq. (3.6) egentlig fortæller os! Den fortæller helt præcist, at jo flere målinger desto større præcision og derfor kommer vi samtidig tættere på den sande middelværdi μ_s! Eksempelvis, hvis I har foretaget N målinger, og bagefter bliver bedt om at gentage eksperimentet (på præcis den samme måde og med samme udstyr!), men skal være to gange så præcise som før. Hvor mange målinger M skal der nu udføres? Dette er simpelt at bestemme ved brug af SDOM! Det første eksperiment gav jer $\overline{x} \pm \delta\overline{x}$. Det nye

[15] Det er vigtigt når I i grupper skal foretage målinger, at det sker uafhængigt. Eksempelvis, hvis I skal måle vinklen et plan hælder så sig ikke til hinanden, hvad I har målt vinklen til, da det ville farve jeres mening om målingen. Dermed vil det ikke være uafhængige målinger, og det kan vi ikke lide!

eksperiment vil give jer $\overline{x}_{\text{ny}} \pm \delta\overline{x}_{\text{ny}}$. Da vi skal være to gange så præcise svarer dette til $\delta\overline{x}_{\text{ny}} = 1/2\delta\overline{x}$. Ergo

$$\delta\overline{x}_{\text{ny}} = \frac{1}{2}\delta\overline{x} = \frac{\sigma_x}{2\sqrt{N}} = \frac{1}{2}\delta\overline{x} = \frac{\sigma_x}{\sqrt{4N}}. \tag{3.7}$$

Altså for at være to gange så præcis som før skal vi foretage fire gange så mange målinger! Hvis vi vil være tre gange så præcis som før finder vi, at der skal foretages ni gange så mange målinger. Helt generelt forholder det sig, at hvis vi gerne vil øge præcisionen med en faktor q skal der udføres q^2 så mange målinger. Så hvis man gerne vil være mere præcis i sit eksperiment kan det i det lange løb bedre betale sig at bruge lidt ekstra tid til at sørge for godt eksperimentelt udstyr, og at de er kalibreret korrekt, da det klart vil tage længere tid at foretage flere målinger.



FIGUR 3.13 • FIGUREN VISER EN NORMALFORDELING MED SPREDNING σ OG MIDDELVÆRDI μ. SANDSYNLIGHEDEN FOR, AT EN MÅLING VIL RAMME I ET AF DE ANGIVNE INTERVALLER ER ANGIVET I DE FARVEDE OMRÅDER. DET SKAL TOLKES PÅ DEN MÅDE, AT DET EKSEMPELVIS ER 34.1% SANDSYNLIGT, AT DEN VIL RAMME I INTERVALLET $[\mu - 1\sigma, \mu]$, ELLER 68.2% SANDSYNLIGT DEN VIL RAMME I $[\mu - 1\sigma, \mu + 1\sigma]$.

Vi vil som det sidste gennemgå nogle eksempler på oplyste eksperimentelle resultater og se, hvordan man skal og *ikke* skal oplyse disse. Betragt følgende eksperimentelle resultater:

1. $6051.78\,\text{m/s} \pm 30\,\text{m/s}$
2. $9.82\,\text{m/s}^2 \pm 0.02385\,\text{m/s}^2$
3. $55.5\,\Omega \pm 0.3\,\Omega$

Af ovenstående eksperimentelle resultater er der kun en af dem som er oplyst korrekt. Lad os gennemgå dem en ad gangen, og se hvorfor det enten er rigtigt eller forkert.

1. Udsagnet $6051.78\,\text{m/s} \pm 30\,\text{m/s}$ fortæller os, at cifret 5 kan være enten 8 eller 2 grundet usikkerheden. Vi har derfor på ingen måde den nødvendige præcision til at kunne påstår noget om de resterende cifre 1, 7 og 8. En afrunding er derfor påkrævet op til det ciffer med usikkerhed. Så resultatet oplyses korrekt når vi oplyser det som $6050\,\text{m/s} \pm 30\,\text{m/s}$.
2. Udsagnet $9.82\,\text{m/s}^2 \pm 0.02385\,\text{m/s}^2$ siger, at usikkerheden er $0.02385\,\text{m/s}^2$. Sagt på en anden måde: Usikkerheden er oplyst med fire *betydende cifre*[16]. Dette er yderst præcist, hvilket næsten er umuligt! Men lad os lege med tanken, at denne usikkerhed er reel nok, se er resultatet alligevel opgivet forkert. For, at det skal oplyses korrekt skal antallet af decimalpladser stemme overens - altså det korrekte ville være $9.82000\,\text{m/s}^2 \pm 0.02385\,\text{m/s}^2$. Helt generelt når I oplyser eksperimentelle resultater så er en gylden regel, at usikkerheden skal have mindst et, men højst to, betydende cifre. Se [Taylor, 1997, side 14-16] for en mere uddybende diskussion.

[16] Hvis et tal, eksempelvis 32.05056, skulle opgives med to betydende cifre ville det være 32. Med tre betydende cifre 32.05, og med fire 32.0506. Bemærk, at nul ikke tæller som et betydende ciffer.

3. Resultatet $55.5\,\Omega \pm 0.3\,\Omega$ er oplyst helt korrekt. Antallet af decimalpladser stemmer overens og usikkerheden er ikke oplyst med mere end et betydende cifre.

Kombination af usikkerheder

De hypoteser vi som fysikere vil be- eller afkræfte eller de lovmæssigheder vi vil eftervise i forbindelse med et eksperiment er i stort set alle tilfælde givet ved en matematisk model - altså et funktionsudtryk. Eksempelvis, Ohm's lov

$$V = RI, \tag{3.8}$$

hvor V er spændingen målt i V, R er modstanden målt i Ω og I er strømmen målt i A. Eller Galileis faldlov for et objekt, som fra start er i hvile,

$$s = \frac{1}{2} g t^2, \tag{3.9}$$

hvor s er strækningen i m, g er tyngdeaccelerationen i m/s^2, t er tiden i s.

I de førnævnte modeller indgår størrelser som vi kan måle: V, R, I, g og t. Men i begge tilfælde kan der være tale om usikkerheder i flere af de i modellen indgående størrelser. Eksempelvis, vil usikkerheden i spændingen V afhænge af usikkerhederne i både modstanden R og strømmen I. Så hvordan kan vi bestemme usikkerheden i V? Opfyldelsen af en række krav sætter os i stand til at kunne besvare netop dette. Disse er, at usikkerhederne skal være

1. normalfordelte,
2. uafhængige,
3. meget mindre relativt til målingerne.

Punkt 1 er ret svær at vise gyldigheden af eftersom dette ville kræve, at alle amperemetre og modstande i hele universet skal indsamles, og så skal der genereres et histogram over deres usikkerheder. At usikkerhederne, som punkt 2 udsiger, skal være uafhængige betyder, at usikkerheden i amperemeteret ikke afhænger af den i modstanden og omvendt. Her vil et simpelt argument for uafhængigheden være, at fordi usikkerheden i modstanden er stor vil det ikke have indflydelse på den usikkerhed i amperemetret[17]. Punkt 3 kan først bekræftes når man har foretaget nogle målinger. Et eksempel ville være, at vi målte modstanden til $20.2\,\Omega$ og dens usikkerhed påtrykt er $0.2\,\Omega$. Her ville det være klart at usikkerheden er meget lille relativt til målingen af den. Dette skrives $\delta R << R$.

Gør altid overvejelser i jeres rapporter omkring, hvorvidt punkt 2 og 3 er overholdt. Selvom punkt 1 er meget svær at påvise så nævn alligevel, at det i de fleste tilfælde er en rimelig antagelse.

De førnævnte tre antagelser medfører, at loven om kombinerede usikkerheder gælder:

[17] Bemærk dog, at ved store værdier for R og I kan usikkerhederne godt være afhængige af hinanden. Men det ser vi her bort fra.

Loven om kombinerede usikkerheder

Er de respektive usikkerheder normalfordelte, uafhængige og meget mindre relativt til målingerne gælder der for usikkerheden i en størrelse givet ved en funktion $f(z_1, z_2, ..., z_M)$, at

$$\delta f(z_1, z_2, ..., z_M) = \sqrt{\sum_{i=1}^{M} \left(\frac{\partial f}{\partial z_i} \cdot \delta z_i\right)^2}. \tag{3.10}$$

Ovenstående kaldes også for *loven om fejlpropagering* eller *ophobningsloven* (engelsk *law of error propagation*).

Et eksempel på anvendelsen af loven om kombinerede usikkerheder kunne være, at vi opstiller et kredsløb som illustreret i figur 3.14, og vi ønsker at bestemme hvilken spænding $\mathscr{E}$ spændingskilden leverer samt usikkerheden i denne $\delta\mathscr{E}$. Det er blevet målt, at $I = 1.2\,\mathrm{A} \pm 0.1\,\mathrm{A}$ og $R = 10.0\,\Omega \pm 0.3\,\Omega$. Usikkerheden i $\mathscr{E}$ bestemmes via loven om kombinerede usikkerheder i eq. 3.10, hvor $f = \mathscr{E} = RI$. Ergo er $\mathscr{E} = 10.0\,\Omega \cdot 1.2\,\mathrm{A} = 12.0\,\mathrm{V}$, og $\delta\mathscr{E}$ er givet ved

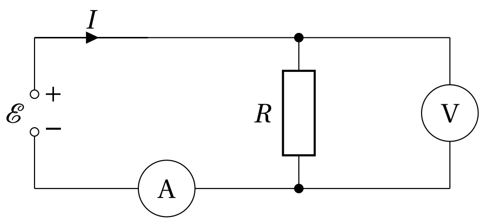
$$\delta\mathscr{E} = \sqrt{\left(\frac{\partial\mathscr{E}}{\partial I} \cdot \delta_I\right)^2 + \left(\frac{\partial\mathscr{E}}{\partial R} \cdot \delta_R\right)^2} = \sqrt{\left(\frac{\partial}{\partial I}(RI) \cdot \delta_I\right)^2 + \left(\frac{\partial}{\partial R}(RI) \cdot \delta_R\right)^2} \tag{3.11}$$

$$= \sqrt{(R \cdot \delta_I)^2 + (I \cdot \delta_R)^2} \tag{3.12}$$

$$= \sqrt{(10.0\,\Omega \cdot 0.1\,\mathrm{A})^2 + (1.2\,\mathrm{A} \cdot 0.3\,\Omega)^2} \tag{3.13}$$

$$= 1.1\,\mathrm{V}. \tag{3.14}$$

Konklusionen er således, at spændingskilden leverer en spænding på $12.0\,\mathrm{V} \pm 1.1\,\mathrm{V}$.



FIGUR 3.14 • ET KREDSLØB BESTÅENDE AF EN SPÆNDINGKILDE, DER LEVERER EN SPÆNDING $\mathscr{E}$, EN MODSTAND MED MODSTANDEN R. SPÆNDINGSFALDET OVER MODSTANDEN MÅLES MED ET VOLTMETER OG ER ÆKVIVALENT MED $\mathscr{E}$. STRØMMEN MÅLES MED ET AMPEREMETER.

Usikkerheder med Matlab

Så langt så godt: σ_x er usikkerheden i en enkelt måling, $\overline{x}$ er det bedste estimat på den sande middelværdi og usikkerheden på denne er givet ved SDOM - altså $\delta\overline{x}$. Vi skal her se, hvordan vi i Matlab kan plotte vores målinger med tilhørende usikkerheder.

Målingerne vi vil benytte er fra den første forelæsning i kurset *Anvendt Statistik: Fra Data til Resultater* jeg fulgte i efteråret 2012. Forelæseren - Troels C. Petersen - bad de studerende om, en efter en, at komme op og måle længden af kateteret i auditorium A på NBI, hvor forelæsning den pågældende dag blev afholdt, med en 30 cm lineal[18]. Ligesom i tilfældet med det førnævnte fiktive kateter vidste ingen af de studerende

[18] Kateteret skulle også måles med en tommestok, men målingerne med lineal betragtes kun her.

hvad de andre havde målt længden til. Derfor er målingerne uafhængige. Vi vil nu med disse data visse, hvordan man kan plotte sine målinger med tilhørende usikkerhed i Matlab. Målinger kan ses i tabel 3.5.

Fremgangsmetoden er egentlig, at vi skal have udregnet $\overline{x}$, σ_x og SDOM (se tabel 3.5). Herefter plottes målingerne med tilhørende usikkerhed, hvilket gøres med kommandoen `errorbar`. I plottet med målingerne vil vi også gerne vise middelværdien $\overline{x}$. Hvordan dette opnås i Matlab kan ses i kodeblok 3.32.

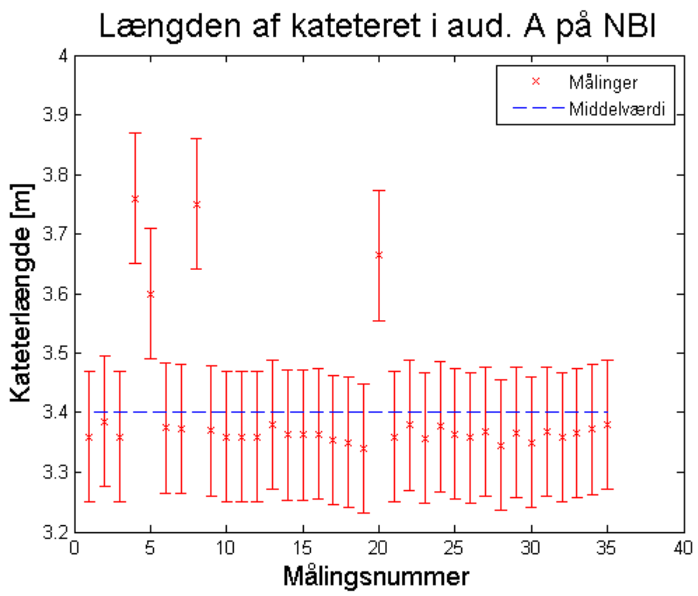
Målinger x_i [m]	$\overline{x}$ [m]	σ_x [m]	SDOM [m]
3.360, 3.385, 3.360, 3.760, 3.600, 3.375, 3.373, 3.750, 3.370, 3.360, 3.360, 3.360, 3.380, 3.363, 3.363, 3.364, 3.354, 3.350, 3.340, 3.664, 3.360, 3.379, 3.357, 3.377, 3.364, 3.358, 3.368, 3.345, 3.367, 3.350, 3.368, 3.359, 3.366, 3.372, 3.380	3.402	0.109	0.018

TABEL 3.5 • TABELLEN VISER UAFHÆNGIGE MÅLINGER FORETAGET AF STUDERENDE I FORBINDELSE MED DEN FØRSTE FORELÆSNING AF TROELS C. PETERSEN I KURSET ANVENDT STATISTIK: FRA DATA TIL RESULTATER (2012). MÅLINGERNE ER LÆNGDEN AF KATETERET I AUDITORIUM A PÅ NBI. MÅLEINSTRUMENTET ER EN 30cm LINEAL.

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

%Data lægges som altid i et array
Data = [3.360 3.385 3.360 3.760 3.600 3.375 3.373 3.750 3.370 3.360
    3.360 3.360 3.380 3.363 3.363 3.364 3.354 3.350 3.340 3.664
    3.360 3.379 3.357 3.377 3.364 3.358 3.368 3.345 3.367 3.350
    3.368 3.359 3.366 3.372 3.380];

%antallet af målinger N
N = length(Data);

%Hver måling har usikkerheden sigma_x
Error=std(Data);

%For hvert datapunkt knyttes usikkerheden Error.
ErrorArray = ones(1,N)*Error;

figure
%Her plottes vores data med tilhørende usikkerhed. Datapunkterne og
%usikkerhederne vises som røde krydser
errorbar(1:N,Data,ErrorArray,'rx')
%Fastlås plottet så vi kan plotte middelværdien oven i. Middelvæ
    rdien plottes som en stiplet blå streg
hold on
MeanOfData = ones(1,N)*mean(Data);
plot(1:N,MeanOfData);
%Plottes styling
legend('Målinger','Middelværdi','location','NorthEast');
title('Længden af kateteret i aud. A på NBI','fontsize',18);
xlabel('Målingsnummer','fontsize',14);
ylabel('Kateterlængde [m]','fontsize',14);
```

KODEBLOK 3.32 • DENNE KODE VISER, HVORDAN MAN BENYTTER KOMMANDOEN *errorbar* TIL AT PLOTTE MÅLINGER MED TILHØRENDE USIKKERHED. I PLOTTET ER MIDDELVÆRDIEN AF MÅLINGER OGSÅ REPRÆSENTERET. DETTE ER GJORT VED AT BENYTTE *hold on*.

Det plot vi får ud af at køre scriptet i kodeblok 3.32 kan ses i figur 3.15.



FIGUR 3.15 • PLOTTET VISER MÅLINGERNE FRA TABEL 3.5 MED DERES TILHØRENDE USIKKERHED σ_x. DETTE PLOT ER LAVET MED KOMMANDOEN `errorbar` OG HEREFTER *`hold on`* FOR AT KUNNE PLOTTE MIDDELVÆRDIEN OVEN I. KODEBLOK 3.32 VISER HVORDAN.

Plot med usikkerheder

Ønskes der plots af målinger, hvor de tilhørende usikkerheder skal vises benyttes kommandoen `errorbar`. Denne har følgende syntaks:

`errorbar(`⟨*array med x-akse værdier*⟩`,`⟨*array med y-akse værdier*⟩`,`⟨*array med usikkerheder på y-værdier*⟩`)`.

`errorbar`-plottet kan styles med de gængse styling options. Se boks på side 28. For mere information om *`errorbar`* se da `http://www.mathworks.se/help/matlab/ref/errorbar.html`.

3.5 Fitting

Det sidste som vi skal kigge på i forbindelse med dataanalyse er fitting. Dette skyldes at vi, som førnævnt, beskæftiger os med matematiske modeller i naturvidenskaben. Vi skal se, hvordan man kan bestemme forskriften af den funktion, som bedst beskriver sine data. Dette kaldes for *fitting*.

Vi vil her begrænse os til, hvordan man bestemmer den bedste rette linje til sine data, men det skal dog nævnes at der også findes metoder for mere komplicerede funktionstyper. Eksempelvis, anden- og tredjegrads polynomier. Dette behandler [Taylor, 1997, side 193-196].

Som det første præsenteres formlerne, og herefter ser vi hvordan de anvendes. Som eksempel tager vi udgangspunkt i det fiktive forsøg med den lufttætte beholder. Se tabel 2.2. Selve den teoretiske behandling og udledning af de kommende formlerne vil vi ikke komme nærmere ind på, men interesserede sjæle kan se [Taylor, 1997, side 181-192] eller [Barlow, 1999, side 100-102].

Fitting af ret linje

Hvis der er foretaget N målingspar (x_i, y_i), hvor der på y_i er en usikkerhed δ_y, og man ønsker at fitte en ret linje - det vil sige en funktion af typen $mx + c$, hvor $m, c \in \mathbb{R}$ - til målingerne udregnes konstanterne m og c ved

$$m = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\overline{x^2} - \overline{x}^2} \tag{3.15}$$

og

$$c = \overline{y} - m\overline{x}. \tag{3.16}$$

Usikkerheden på hhv. m og c er givet ved

$$\delta_m = \frac{\delta_y}{\sqrt{N\left(\overline{x^2} - \overline{x}^2\right)}}, \quad \delta_c = \delta_y \cdot \sqrt{\frac{\overline{x^2}}{N\left(\overline{x^2} - \overline{x}^2\right)}}, \tag{3.17}$$

hvor

$$\delta_y = \sqrt{\frac{1}{N-2} \cdot \sum_{i=1}^{N} \left(y_i - (m \cdot x_i + c)\right)^2}. \tag{3.18}$$

Bemærk, at $\overline{x^2}$ *ikke* betyder gennemsnittet af målingerne x_i kvadreret, men gennemsnittet af kvadratet af x_i.

Der er på nuværende tidspunkt måske ophav til en smule forvirring, da usikkerheden δ_y ikke er givet ved den sædvanlige formel:

$$\sigma_y = \sqrt{\frac{1}{N-1} \left(y_i - \overline{y}\right)^2}, \tag{3.19}$$

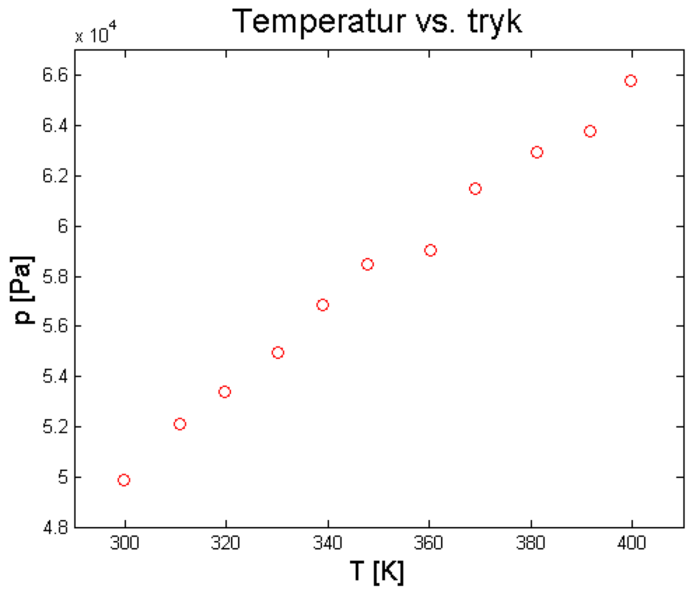
som er spredningen for målingerne af y. Hvorfor ikke? Husk på, at spredningen var den gennemsnitlige afvigelse til middelværdien! Men vi er her interesseret i den gennemsnitlige afvigelse til den bedste rette linje $mx + c$, fordi vi netop har inspiceret målingerne i et plot og har gjort os den hypotese, at der sandsynligvis er en retlinjet sammenhæng mellem p og T! Derfor, hvis I kigger på formlen for δ_y i eq. (3.18) og erstatter leddet $m \cdot x_i + c$ med $\overline{y}$ får vi næsten formlen for σ_y i eq. (3.19) frem. Dog dividerer vi med $N-2$ frem for $N-1$. Grunden hertil skal findes i det vi kalder for *frihedsgrader* (engelsk degrees of freedom) forkortet NDOF. Ganske kort fortalt er NDOF defineret ved antallet af målinger minus antallet af estimerede parametre. Her er m og c de estimerede parametre. Ergo er NDOF $= N-2$, da vi har N målinger. Havde der været tre estimerede parametre ville NDOF $= N-3$ osv. Så antallet af frihedsgrader falder jo flere parametre skal estimeres. Dette kaldes for *Bessels korrektion* efter den tyske matematikker og astronom Friedrich Wilhelm Bessel. Mere dybdegående information om frihedsgrader kan findes i [Taylor, 1997, side 188] og [Barlow, 1999, side 107-108].

Som lovet i kapitel 2 vil vi nu i det fiktive forsøg med den lufttætte beholder bestemme stofmængden n af atmosfærisk luft i beholderen. Forsøget var følgende: En tætsluttende beholder med volumen $V = 0.5\,\mathrm{m}^3$, indeholdende en stofmængde atmosfærisk luft n, opvarmes. Løbende blev temperaturen og trykket i beholderen noteret. I forrige afsnit bestemte vi usikkerheden i målingerne af trykket p, og vi antog at der ikke var nogen nævneværdig usikkerhed i målingerne af temperaturen - altså $\delta_T \approx 0\,\mathrm{K}$. Målingerne kan ses i tabel

2.2.

Det første som man *altid* skal gøre er at inspicere sine målinger - det vil sige plotte dem! Dette gjorde vi på side 23 i kapitel 2. Plottet er vist igen i figur 3.16.

Som det fremgår af plottet i figur 3.16 ligger målingerne på en tilnærmelsesvis ret linje. Dette burde ikke



FIGUR 3.16 • PLOTTET VISER MÅLINGERNE FRA TABEL 2.2. PLOTTET ER DET SAMME SOM I FIGUR 2.4. DET FREMGÅR AT MÅLINGERNE LIGGER RELATIVT PÆNT PÅ EN RET LINJE.

komme som den store overraskelse, da atmosfærisk luft kan antages som værende en idealgas i det målte temperaturinterval. Ergo har vi den matematiske model

$$pV = nRT, \tag{3.20}$$

hvor p er trykket i Pa, V er beholderens volumen, n er stofmængden af gas i mol, R er gaskonstanten som er $8.31\,\frac{\text{Pa}\cdot\text{m}^3}{\text{mol}\cdot\text{K}}$, og T er temperaturen i K. Vi antager, at der ingen usikkerhed er i hverken beholderens volumen og gaskonstanten. Opgaven er at bestemme n samt usikkerheden på denne δ_n.

Det første vi skal gøre er, at få idealgasloven - eq. 3.20 - omskrevet til en funktion $p(T)$, da vi i vores plot har trykket p op ad y-aksen og temperaturen T ud ad x-aksen. Kort sagt: Vi skal have isoleret p

$$p(T) = \frac{nR}{V}T. \tag{3.21}$$

Ergo kan vi nu se, at $p(T)$ vil være en ret linje med hældningen nR/V - altså $m = nR/V$. Så ifølge vores matematiske model går den rette linje igennem punktet $(0,0)$, men betyder det så, at vi slet ikke behøver, at bekymre os om skæringen med y-aksen? Svaret hertil er JO! Det er med vilje der benyttes store bogstaver for det kan ikke siges oftest nok: Selvom din matematiske model går i gennem et specifikt punkt eller skal have en bestemt hældning osv. må parametre ALDRIG fastlåses/fikseres! Det vil sige i vores tilfælde, at vi stadig skal tage højde for c, da vi på ingen måde kan være sikre på, at vores målinger vil give $c = 0\,\text{Pa}$! Det er endda så vigtig, at det får en boks.

Fiksering af parametre

Når der foretages fitting af en bestemt matematisk model til målinger vil en fornuftig eksperimental fysiker ALDRIG fiksere en parameter, selvom den matematiske model forudsiger det!

Derfor bestemmer vi både m og c. Hvordan man kan gøre dette i Matlab er vist i kodeblok 3.33.

Køres scriptet i kodeblok 3.33 finder vi først

$$m = 153.9\,\mathrm{Pa/K} \pm 4.5\,\mathrm{Pa/K}, \quad \text{og} \quad c = 4200\,\mathrm{Pa} \pm 1600\,\mathrm{Pa}. \tag{3.22}$$

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

%Data lægges som altid i arrays
Temperatur = [299.68  310.83  319.77  330.26  339.10  347.84  360.09  369.05  381.17  391.74  399.63];

Tryk = [49862.66  52091.57  53397.68  54936.52  56836.74  58462.12  59023.77  61503.39  62943.02  63800.99  65815.72];

%antallet af målinger N, gaskonstanten R og volumen V
N = length(Tryk);
R = 8.31;
V = 0.5;
%Gennemsnit af temperaturen og trykket
snittemp = mean(Temperatur);
snittryk = mean(Tryk);
%Udregning gennemsnittet af xy og x^2
snittemptryk = mean(Temperatur.*Tryk);
snittempsquare = mean(Temperatur.^2);
%Beregning af hældningen m=nR/T
m = (snittemptryk-snittemp*snittryk)/(snittempsquare-snittemp^2);
%Beregning af skærningen med y-aksen
c = snittryk - m*snittemp;
%Usikkerheden i trykket p - altså deltap
residualssquare = zeros(1,N);
for l=1:N
residualssquare(l) = (Tryk(l) - m*Temperatur(l)-c)^(2);
end
deltap = sqrt(sum(residualssquare)/(N-2));
deltapArray = deltap * ones(1,N);
%Usikkerhederne på m og c
deltam = deltap/sqrt(N*(snittempsquare-snittemp^2));
deltac = deltap*sqrt(snittempsquare/(N*(snittempsquare-snittemp^2)));
%Udregn n
n=m*V/R;
%Usikkerheden i n - altså deltan udregnes ved brug af loven om kombinerede usikkerheder
deltan=deltam/abs(R/V);
%Åbn figur vindue
figure
%Her plottes vores data med tilhørende usikkerhed. Datapunkterne og
%usikkerhederne vises som røde krydser
errorbar(Temperatur,Tryk,deltapArray,'rx')
%Oven i plottet ønsker vi at vise den bedste rette linje
hold on
%Udregn m*T + c og plot den resulterende rette linje
yfitvalues = m*Temperatur + c;
plot(Temperatur,yfitvalues)

%Intervaller for x og y vi ønsker at zoome ind på
xlim([290 410]);
ylim([4.9*10^4 6.7*10^4]);
%Plottes styling
legend('Dine data','Fit: p=mT+c','location','NorthWest');
title('Temperatur vs. tryk','fontsize',18);
xlabel('T [K]','fontsize',14);
ylabel('p [Pa]','fontsize',14);
```

KODEBLOK 3.33 • KODEN VISER, HVORDAN MAN KAN BESTEMME PARAMETRENE m OG c. ENDVIDERE PLOTTES MÅLINGERNE SAMMEN MED DEN ENDELIGE FITTING AF DEN RETTE LINJE $p = mT + c$

Da vi nu har beregnet hældningen m er det muligt at bestemme stofmængden af atmosfærisk luft n, da $m = nR/V$ og derfor må $n = mV/R$. Vi finder, at $n = 9.3\,\text{mol}$.

Nu mangler vi blot, at bestemme usikkerheden δ_n. Hertil benytter vi loven om kombinerede usikkerheder givet i eq. (3.10), hvor $f = n = mV/R$. Ergo har vi

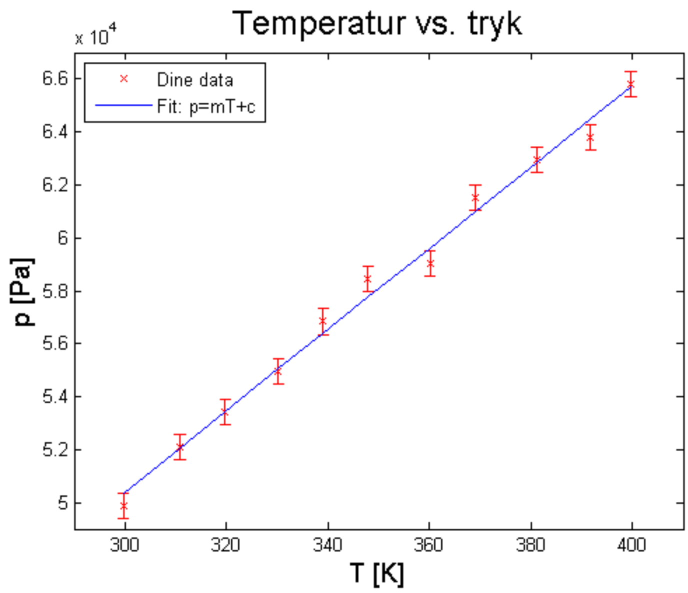
$$\delta_n = \sqrt{\left(\frac{\partial n}{\partial m}\cdot\delta_m\right)^2 + \left(\frac{\partial n}{\partial V}\cdot\delta_V\right)^2 + \left(\frac{\partial n}{\partial R}\cdot\delta_R\right)^2}, \tag{3.23}$$

men der er ingen usikkerhed på hverken R eller V. Derfor har vi det noget simplere udtryk

$$\delta_n = \sqrt{\left(\frac{\partial n}{\partial m}\cdot\delta_m\right)^2} = \sqrt{\left(\frac{V}{R}\cdot\delta_m\right)^2} = \left|\frac{V}{R}\right|\delta_m = \frac{0.5\,\text{m}^3}{8.31\,\frac{\text{Pa}\cdot\text{m}^3}{\text{mol}\cdot\text{K}}}\cdot 4.5\,\text{Pa/K} = 0.3\,\text{mol}. \tag{3.24}$$

Ergo indeholder den lufttætte beholder $9.3\,\text{mol} \pm 0.3\,\text{mol}$. Plottet som kodeblok 3.33 resulterer i kan ses i figur 3.17.

Hvad med parameteren c? Det eneste vi ønskede, at bestemme var stofmængden af gas i beholderen og dennes usikkerhed. Så parameteren c er teknisk set i denne sammenhæng ikke interessant. Men reflekterer vi dog alligevel lidt over parameteren $c = 4200\,\text{Pa} \pm 1600\,\text{Pa}$ vil vi hurtigt opdage, at $0\,\text{Pa}$ ikke ligger inden for usikkerhedsintervallet! Hvis idealgas loven skulle holde stik, så burde vi igennem vores målinger kunne bekræfte dette, men det er nu engang ikke tilfældet. Derfor kan der være tale om *systematiske fejl*. Systematiske fejl består i, at vi eksempelvis ikke ikke har nulstillet måleinstrumenterne korrekt eller et eller flere af dem er defekte. Hvis jeg kom ud for sådan en uoverensstemmelse ville jeg, med den erfaring jeg i skrivende stund har opnået, gøre følgende i den nævnte rækkefølge:



FIGUR 3.17 • PLOTTET VISER MÅLINGERNE FRA TABEL 2.2 SAMMEN MED DEN FITTET RETTE LINJE $mT + c$ (BLÅ). BEMÆRK, AT DEN FITTEDE RETTE LINJE GENNEMLØBER $(11-3)/11\cdot 100 \approx 73\%$ AF MÅLINGERNES USIKKERHEDSINTERVAL ALTSÅ $p_i \pm \delta_p$, HVILKET ER MEGET TÆT PÅ DE 68 %!

- Gennemgå koden benyttet til dataanalysen for fejl.
- Tjekke, at måleinstrumenterne nulstilles korrekt (hvilket egentlig bør tjekkes inden målingerne foretages!).
- Gentage forsøget. Eventuelt med flere målinger.

- Udskifte måleinstrumenter en ad gangen og gentage forsøget.

Fejlfinding i et eksperiment er ikke nemt, og det er kun noget som man bliver bedre til med tiden.

3.6 Indlæsning af data fra fil

I alle de koder vi hidtil har set er data bliver lagt manuelt i arrays. Virkeligheden forholder sig dog meget anderledes. Det er ganske normalt, at man har flere tusinder af målinger af forskellige art. Vi vil aldrig blive færdige, hvis vi manuelt skulle sidde at taste dem ind i arrays! Heldigvis kan vi indlæse data direkte ind i Matlab fra filer.

De extensions som Matlab kan læse er blandt andre *.dat, *.csv, *.txt og *.dlm. Den sidstnævnte har jeg dog endnu ikke stødt på. Derimod er de tre førstnævnte extensions meget brugt.

Sæt, at vi har udført et forsøg, og måleinstrumentet har outputtet en *.dat, *.csv eller *.txt, som indeholder nedenstående:

```
[Data]
Datatype 1, Datatype 2, Datatype 3, Datatype 4, Datatype 5, Datatype 6, Datatype 7, Datatype 8,
Datatype 9, Datatype 10
 0.1834  -0.6980   0.6699  -1.1114  -2.6411  -0.0281   1.4213  -0.2256   0.6370   1.7044
-0.2597   0.3041   0.3758   0.4682  -0.4861   0.1404   1.4076  -0.9660   0.0749   0.0236
-1.3548   1.1011   0.7667   0.3226   0.1960   0.2376  -1.0290   0.0950   1.1233   0.2900
-1.1574  -0.4466   1.5235   0.1000   0.9597  -0.6715   0.2065  -0.2567  -0.0331  -1.4199
-0.7620  -0.4667   0.5768   0.3014   1.3803  -1.0450   1.3411   2.3101  -0.0977   0.4753
-0.2503  -1.4358   0.4260   0.0238  -0.9415   0.9658   1.3327   0.1901  -0.5566  -1.4473
-1.6488  -0.9777  -0.0429  -0.0221   0.7609  -0.2198  -1.2849  -0.1734  -0.6155  -0.9883
-0.1820   0.6059  -1.2423  -0.0088   0.2379   1.4145   1.6184  -0.0140   1.6046   0.9494
 0.6157  -0.1137  -0.5050   0.9295  -0.2084  -0.9242   0.6616  -0.6127   0.7685   0.3512
-0.3767   0.7646   0.3259  -0.0904   0.0247  -0.5941   0.2273   2.0718   0.0869  -0.8723
```

Data skal læses som, at den første kolonne indeholder målinger af datatype 1, anden kolonne datatype 2 osv.

Det vi ønsker at importere i Matlab er naturligvis kun data og ikke teksten. Det vil sige, at vi ikke er interesseret i de tre første linjer. For, at kunne importere data i Matlab skal vi benytte en lille cocktail af funktionerne `cell2mat` og `textscan`. `textscan` gør - som navnet antyder - læser de rækker vi angiver og `cell2mat` konverterer de læste rækker til et array. Er der mere en en række vil det resulterer i et flerdimensionalt array også kaldet en matrix.

Vi viser nu koden til at importere al data i et array - altså alle rækker og kolonner og så gennemgår vi de forskellige aspekter ved koden. Koden er vist i kodeblok 3.34.

```
%Rydder hukommelsen
clear all
%Lukker alle vinduer med plots
close all

%Åbn filen så Matlab kan læse den
file = fopen('TestImport.txt');

%Indlæs data med textscan
DataScan = textscan(file,'%n %n %n %n %n %n %n %n %n %n','headerlines',3)

%Data er nu læst og vi kan lukke filen
fclose(file);

%Konverter det som ligger i variablen DataScan til et stort array - altså et 10 gange 10 array
DataBigArray = cell2mat(DataScan);

%eller vi kan eksempelvis lave et array kun bestående af først kolonne
FirstColumn = cell2mat(DataScan(1))';
```

KODEBLOK 3.34 • NÅR DATA SKAL INDLÆSES FRA EN FIL TIL MATLAB SKAL FILEN FØRST ÅBNES MED `fopen`. DEREFTER BENYTTES `cell2mat` SAMMEN MED `textscan` FOR AT INDLÆSE DATA TIL ET ARRAY. TIL SIDST LUKKES FILEN MED `fclose`. BEMÆRK, AT SCRIPTET OG DATAFILEN SKAL I DETTE TILFÆLDE LIGGE I SAMME BIBLIOTEK.

Det første vi skal sikre os er, at filen indeholdende data ligger i samme bibliotek som scriptet, og selvfølgelig at i Current Folder er stien til det bibliotek, hvor scriptet ligger, angivet. Information omkring den fil vi åbner med `fopen` lægger vi i variablen `file`. Hernæst kalder vi `textscan`, hvor vi angiver filen Matlab skal læse fra - altså `file`. Herefter fortæller vi `textscan` at hver af de ti kolonner skal lægges i et array. Dette gøres ved at angive strengen `'%n %n %n %n %n %n %n %n %n %n'`. Strengen `'headerlines'` fortæller `textscan`, at det efterfølgende tal er det antal rækker den ikke skal læse. Her er det 3, da de første tre linjer er tekst. Variablen `DataScan` indeholder nu ti celler, hvor der i den første celle ligger et array indeholdende den første kolonne, den anden kolonne nummer osv. For at konvertere fra celle strukturen til array som vi kan benytte benyttes kommandoen `cell2mat`. Hvis vi eksempelvis gerne vil sætte alle de læste kolonner sammen til et stort array (her et ti gange ti array) skal vi angive `cell2mat(DataScan)`. Men måske er vi kun interesseret i den første eller anden kolonne. Hertil angiver vi hhv. `cell2mat(DataScan(1))'` eller `cell2mat(DataScan(2))'`. Apostrofen for enden gør, at outputtet er et række array og ikke et kolonne array (vektor).

For mere information omkring `textscan` og `cell2mat` se hhv. http://www.mathworks.se/help/matlab/ref/textscan.html og http://www.mathworks.se/help/matlab/ref/cell2mat.html.

$$\int_{-\infty}^{\infty} |\psi|^2 \, dx = 1$$

A Matematik appendiks

Kapiteloversigt

Dette appendiks har til formål, at grundlæggende behandle nogle af de matematiske begreber som I måske ikke endnu er stødt på. De begreber som er udvalgt er summantions-tegnet og partial differentiering.

A.1 Σ-symbolet

I naturvidenskaben har vi for vane, at alt skal være så kompakt som overhovedet muligt, hvilket helt klart gør skrivearbejdet en hel del lettere. Et yderst simpelt eksempel kunne være tallet en milliard, som er 1000000000. Men vi orker ikke at skulle skrive alle de nuller, plus det også optager unødig plads. Derfor er en langt mere kompakt måde mere hensigtsmæssig. Så i stedet for 1000000000 ville vi skrive 10^9. Matlab skriver det samme en smule anderledes for der er 10^9 gengivet som $1e+9$.

Fuldstændigt analogt til ovenstående eksempel har bestemte symboler en given betydning. Det symbol vi her skal snakke om er Σ-symbolet, som er det græske bogstav store sigma. Dette symbol benyttes til at skrive summe kompakt. Et eksempel herpå er hvis vi gerne vil lægge alle hele tal sammen fra 1 til 10 - det vil sige

$$1+2+3+4+5+7+8+9+10.$$

Ovenstående skrives ved brug af Σ-symbolet

$$\sum_{i=1}^{10} i.$$

Først starter man med at sætte $i = 1$, hvilket står under Σ-symbolet. Dette kaldes den *nedre grænse*. Når man så har sat $i = 1$ fortsætter med at øge i i skridt af 1 og slutter når tællevariablen i er nået den *øvre grænse* 10 - altså

$$\sum_{i=1}^{10} i = \underbrace{1}_{i=1} + \underbrace{2}_{i=2} + \underbrace{3}_{i=3} + \underbrace{4}_{i=4} + \underbrace{5}_{i=5} + \underbrace{7}_{i=7} + \underbrace{8}_{i=8} + \underbrace{9}_{i=9} + \underbrace{10}_{i=10}.$$

Som et andet eksempel betragter vi to vektorer $\vec{a}$ og $\vec{b}$ givet ved

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad \text{og} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

hvor vektorernes komposanter kan være vilkårlige reelle tal. Skalarproduktet, også kendt som prikproduktet, af $\vec{a}$ og $\vec{b}$ er hurtigt udregnet

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Dette kan skrives mere kompakt med Σ-symbolet således:

$$\vec{a} \cdot \vec{b} = \sum_{i=1}^{3} a_i b_i.$$

Faktisk gælder der helt generelt, at skalarproduktet mellem to n-dimensionelle vektorer, $\vec{a}$ og $\vec{b}$, er givet ved

$$\vec{a} \cdot \vec{b} = \sum_{i=1}^{n} a_i b_i.$$

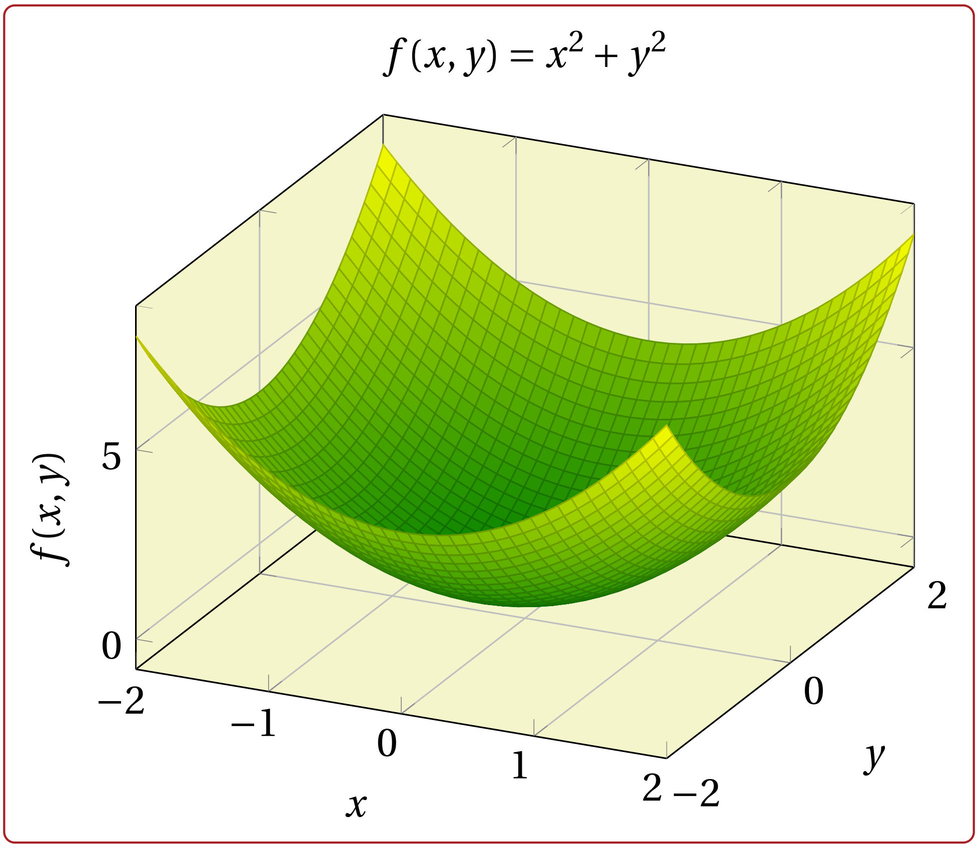
Det kan være, at denne notation på nuværende tidspunkt virker en smule overvældende, især hvis det er første gang I ser den. For en mere detaljeret beskrivelse kan Wikipedias engelske artikel “Summation” anbefales[19], eller [Lindstrøm, 2006, side 29-32].

[19] Artiklen kan findes på følgende adresse: `https://en.wikipedia.org/wiki/Summation`.

A.2 Partial differentiering

Fra gymnasietiden husker vi sikkert, at hvis vi har en funktion $f(x)$ og vi differentierer den med hensyn til x - altså vi udregner $\frac{df}{dx}$ så finder vi et udtryk, der fortæller os hvad tangentens hældning til f er i punktet x er. Dette er også helt korrekt. Men når vi har at gøre med funktioner, som afhænger af mere end en variabel, kaldes det *partial differentiering*. Vi vil ikke gennemgå det matematiske fundament for partial differentiering, men kun holde os for øje på hvordan det anvendes.

Formålet med partial differentiering er at undersøge, hvordan en funktion af flere variabler vokser relativt til de uafhængige variabler. Betragt eksempelvis den simple funktion $f(x,y) = x^2 + y^2$. Se figur A.18. Vi



FIGUR A.18 • 3D PLOT AF FUNKTIONEN $f(x,y) = x^2 + y^2$

vil nu gerne vide, hvordan $f(x,y)$ vokser ud af hhv. x- og y-aksen. Hertil skal de partielle afledet bestemme med hensyn til hhv. x og y. Dette skrives matematisk som hhv. $\frac{\partial f}{\partial x}$ og $\frac{\partial f}{\partial y}$. Tegnet ∂ nedstammer fra det lille græske bogstav delta δ.

Så langt så godt, men hvordan gør man? Jo, det man skal når man bestemmer partielle afledede af en funktion er, at man holder alle andre variabler, end den man differentierer med hensyn til, konstante. I tilfældet med den tidligere nævnte funktion $f(x,y)$ har vi

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\left(x^2 + y^2\right) = \frac{\partial}{\partial x}\left(x^2\right) + \frac{\partial}{\partial x}\left(y^2\right) = 2x + 0 = 2x \tag{A.25}$$

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\left(x^2 + y^2\right) = \frac{\partial}{\partial y}\left(x^2\right) + \frac{\partial}{\partial y}\left(y^2\right) = 0 + 2y = 2y. \tag{A.26}$$

I eq. (A.25) differentierer vi med hensyn til x, og derfor vil leddet $\frac{\partial}{\partial x}\left(y^2\right)$ være lig nul, da vi jo skal holde variablen y konstant. Samme argument er gældende for, at leddet $\frac{\partial}{\partial y}\left(x^2\right)$ er lig med nul, hvor vi dog her skal holde x konstant.

Lad os se et eksempel fra [Kro, 2003, side 71]:

EKSEMPEL: Betragt funktionen $g(x, y, z) = 2xy + z^2$. Bestem nu de partielle afledede $\frac{\partial g}{\partial x}$, $\frac{\partial g}{\partial y}$ og $\frac{\partial g}{\partial z}$. Ved udregning får vi

$$\frac{\partial g}{\partial x} = \frac{\partial}{\partial x}\left(2xy + z^2\right) = 2y + 0 = 2y \tag{A.27}$$

$$\frac{\partial g}{\partial y} = \frac{\partial}{\partial x}\left(2xy + z^2\right) = 2x + 0 = 2x \tag{A.28}$$

$$\frac{\partial g}{\partial z} = \frac{\partial}{\partial x}\left(2xy + z^2\right) = 0 + 2z = 2z. \tag{A.29}$$

Der kan læses meget mere om det matematiske fundament i [Kro, 2003, side 65-94] og endnu mere i [Asmar, 2005]. Vigtigheden af partial differentiering i forbindelse med fysik vil du komme til at mærke gradvist i løbet af din tid på studiet. Dette skyldes at mange fysiske systemer kan beskrives ved hjælp af *partielle differentialligninger*. Et par eksempler herpå er den en-dimensionelle tidsafhængige Schrödinger ligning

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = \left(\frac{-\hbar^2}{2m} + V(x,t)\right)\Psi(x,t)$$

og den en-dimensionelle bølgeligning

$$\frac{\partial^2 u(x,t)}{\partial t^2} = c^2\frac{\partial^2 u(x,t)}{\partial x^2}.$$

Schrödinger ligningen kommer du til at benytte i forbindelse med kvantemekanik kurserne, og bølgeligningen støder du første gang på i kurset *Videregående Klassisk Mekanik* og senere igen i kurset Elektrodynamik og Bølger. Flere eksempler på partielle differentialligninger kan [Asmar, 2005] anbefales. [Asmar, 2005] benyttes i øvrigt til matematikkurset *Introduction to Partial Differential Equations*.

Anvendelsen af partial differentiering kommer du nærmere ind på i kurset *Matematik for fysikere (MatF)* i blok 3.

Litteratur

*Litteratur

[Asmar, 2005] Asmar, N. H. (2005). *Partial Differential Equations with Fourier Series and Boundry Value Problems.* Pearson, 2. udgave.

[Baird, 1998] Baird, D. C. (1998). *Experimentation An Introduction To Measurement Theory and Experiment Design.* Prentice Hall, 3. udgave.

[Barlow, 1999] Barlow, R. J. (1999). *Statistics A Guide to the Use of Statistical Methods in the Physical Sciences.* John Wiley & Sons.

[Kro, 2003] Kro, T. A. (2003). *Funksjoner av flere variable.* Københavns Universitet, 1. udgave.

[Lindstrøm, 2006] Lindstrøm, T. (2006). *Kalkulus.* Universitetsforlaget, 3. udgave.

[Taylor, 1997] Taylor, J. R. (1997). *An Introduction to Error Analysis.* University Science Books, 2. udgave.

Appendix B The Lab Test

In this appendix the self authored test can be reviewed. It has been dubbed the “Lab Test”. The students of Mek1 2013 were given it at the end of the course.

Lab ”test”

Tak fordi du tager denne test. Testen har **<u>ingen</u>** indflydelse på din MEK1-karakter. Din besvarelse vil – uden navns nævnelse – indgå i den forskning mit speciale omhandler.

Der er i alt 23 spørgsmål. I de 4 første spørgsmål er jeg interesseret i din mening. Ergo er der til disse spørgsmål ingen rigtige eller forkerte svar. De resterende 19 spørgsmål, og der er til hvert spørgsmål tre svar muligheder. Der er kun en mulighed som er korrekt. Du skal markere den mulighed som **<u>du</u>** mener er korrekt ved at **<u>sætte et X i cirklen</u>** ved den mulighed du mener er korrekt. Hvis du ikke forstår et eller flere af de 23 spørgsmål kan du ud for spørgsmålets nummer skrive ***F***.

Til prøven er det kun tilladt, at benytte en lommeregner og eventuel kladdepapir. Kladdepapiret skal ikke afleveres. **Prøvens varighed er 30 minutter.**

Tusinde tak for din deltagelse
Allan Finnich

Dit fulde navn:_____________________________________

Din fødselsdato:___________

1. *Har du på noget tidspunkt benyttet dig af noterne ”Grundlæggende Matlab og dataanalyse”, som ligger på Absalon?*
   ◯ *Ja* ◯ *Nej* ◯ *Aldrig hørt om dem*

2. *Har du set nogle af screencastsne som også ligger på Absalon?*
   ◯ *Ja* ◯ *Nej* ◯ *Aldrig hørt om dem*

3. *Hvilket af de følgende tre medier fandt du var mest lærerig for dig?*
   ◯ *Noterne fra spg. 1*
   ◯ *Screencastsne fra spg. 2*
   ◯ *Internet (youtube.com, computerfysik.dk,.......)*

4. *Hvilket af de tre følgende medier føler du generelt fungerer bedst når du skal lære noget nyt?*
   ◯ *Bøger* ◯ *Screencasts/videoer* ◯ *Noter på internetsider*

5. *Hvordan ville man i Matlab definere en variabel med navnet MyFirstVar til at indeholde tallet 4?*
   ◯ *MyFirstVar = 4* ◯ *MyFirstVar := 4* ◯ *4 -> MyFirstVar*

6. *Hvordan defineres et array i Matlab til at indeholde tallene 1, 2, 3 og 4.*
   ◯ *(1 2 3 4)* ◯ *[1 2 3 4]* ◯ *{1 2 3 4}*

7. *Hvis der til aller sidst i en linje med kode angives et semikolon, hvad sker der så når Matlab eksekverer linjen?*
   ◯ *Matlab outputter en fejl*
   ◯ *Matlab eksekverer linjen men outputter intet*
   ◯ *Matlab eksekverer ikke linjen, men ignorerer den*

8. *Hvilken kommando rydder al output i Command Window i Matlab?*
   ◯ *clc* ◯ *clw* ◯ *clear*

9. *Hvilken kommando rydder Workspace, dvs. hukommelsen, i Matlab?*
   ◯ *clean all* ◯ *close all* ◯ *clear all*

10. *Et array ligger i en variabel kaldet OneArray og indeholder tallene 1, 2, 6 og 4. Hvordan ville man ændre det tredje element i OneArray til 3 i Matlab?*
    ◯ *OneArray.index(3) = 3*
    ◯ *OneArray(3) = 3*
    ◯ *index(3) -> OneArray = 3*

11. Hvordan kan man med kolon-operatoren lave arrayet [2 4 6 8 ….. 20]?

◯ *0:2:20* ◯ *1:2:20* ◯ *2:2:20*

12. To variabler FirstArray og SecondArray indeholder hver et array med data. Begge arrays har samme længde. Hvordan ville man få Matlab til lave et plot, hvor FirstArray er op ad y-aksen og SecondArray ud ad x-aksen, og hvor datapunkterne skal vises som røde ringe?

◯ *plot(SecondArray,FirstArray,'ro')*
◯ *plot(FirstArray,SecondArray,'-ro')*
◯ *plot(SecondArray,'--ro')*

13. Hvilket kommando skal man benytte til at lave en overskrift til et histogram, plot osv.?

◯ *header* ◯ *title* ◯ *main*

14. N personer har målt længden af et bord. Hvad er det bedste estimat på bordets længde på baggrund af de N målinger?
Datasættets...

◯ *median* ◯ *spredning* ◯ *middelværdi*

15. Hvad er det bedste estimat på usikkerheden i hver af de N målinger fra spg. 14?
Datasættes...

◯ *median* ◯ *spredning* ◯ *middelværdi*

16. Et histogram har blandt andet den funktion, at den visuelt...

◯ *viser hvordan dine data er fordelt*
◯ *viser, om der er foretaget for mange målinger*
◯ *viser, om forsøget er udført forkert*

17. Antag, at en række målinger af en bestemt kvantitet er normalfordelt. Indenfor hvor mange standardafvigelser (sigma: σ) vil 68.2 % af målingerne ligge?

◯ *1* ◯ *2* ◯ *3*

18. Fem målinger af tyngdeaccelerationen er følgende i m/s^2: 9.7, 9.9, 10.1, 10.0 og 10.3. Hvad er det bedste estimat på tyngdeaccelerationen på baggrund af målingerne?

◯ *9.9 m/s^2* ◯ *10.0 m/s^2* ◯ *10.1 m/s^2*

19. Hvad er usikkerheden på middelværdien i spg. 18?

◯ *0.1 m/s^2* ◯ *0.2 m/s^2* ◯ *0.3 m/s^2*

20. *Hvilken kommando skal man i Matlab benytte for at lave et plot af sine datapunkter med deres tilhørende usikkerhed?*
   - ◯ *errorbarplot*
   - ◯ *errorbardraw*
   - ◯ *errorbar*

21. *En genstand måles på en lille vægt til 20.0 g. Den lille vægts usikkerhed er 4.0 g. På en industrivægt måles et større antal af de samme genstande til 8000 g ± 20 g. Hvad er et passende estimat af antallet af genstande på industrivægten?*
   ◯ *400 ± 40* ◯ *400 ± 80* ◯ *400 ± 120*

22. *En genstand vejer 2.8476 g. Du foretager tre målinger af genstandens vægt. Hvilken af de følgende tre målinger er mest præcis når alle cifre er betydende?*
   ◯ *2.95 g* ◯ *2.2 g* ◯ *2.3654 g*

23. *Hvilket af de følgende tre for-loops er korrekt til at konstruere et array EvenNumbers der indeholder alle lige tal 0 til og med 200?*

◯	*for m=1:1:100 EvenNumbers(m) = 2*(m-1); end loop*	◯	*for m=1:1:100 do EvenNumbers(m) = 2*(m-1); end*	◯	*for m=1:1:100 EvenNumbers(m) = 2*(m-1); end*

Test slut

APPENDIX C
Statistical distributions

In this appendix we present the statistical distributions relevant in the analysis of the students' responses of the Lab test, DHD, and FCI. The statistical tests performed is the independence test, ANOVA, and Student's paired t-test. Therefore distributions presented here are the χ^2-, Fisher's F-, and Student's t-distribution. Also we present the use of their corresponding non-central distributions in computing the power of the statistical test.

C.1 χ^2-distribution

When performing either a G- or χ^2 test of independence both test statistics (see eq. (4.18) and (4.23), respectively) will be distributed according to a χ^2-distribution with $(b-1)(c-1)$ degrees of freedom where b and c are the number of rows and columns, respectively, in the constructed contingency table (see table 4.1). This also holds for the test statistic for the approximate McNemar's χ^2 test of independence (see eq. 4.25).

The definition of the χ^2-distribution is as follows: Let $Z_1, Z_2, \ldots, Z_r$ be independent standard normal distributed stochastic variables. If we define a new stochastic variable X as

$$X = \sum_{i=1}^{r} (Z_i)^2.$$

then X will be distributed according to what is simply defined as the χ^2-distribution with r degrees of freedom [Hansen, 2009, page 337]; as the number of degrees of freedom increases it will tend to a gaussian. This actually applies for all distributions, which the central limit theorem proves [Hansen, 2009, page 334-335].

The probability density function (pdf) of the χ^2-distribution is given by

$$p_{\chi^2}(x; r) = \frac{2^{-r/2}}{\Gamma(r/2)} x^{r/2-1} e^{-x/2}, \quad \text{where} \quad \Gamma(t) = \int_0^\infty x^{t-1} e^{-x}\, dx, \tag{C.1}$$

and is defined for $x \in [0, \infty]$. The mean and variance of the χ^2-distribution are $\mu = n$ and $\sigma^2 = 2n$ [Hansen, 2009, page 337,375]. The distribution is shown in figure C.1 for $r = 1, 2, 3, 6$. The trend towards a gaussian is difficult to see as the choices for the r's here are low; however, for $r \geq 15$ it is more clear.

C.2 Fisher's F-distribution

Let U_1 and U_2 be two independent χ^2-distributed stochastic variables with df_1 and df_2 degrees of freedom. If we define a new stochastic variable Q as the fraction

$$Q = \frac{U_1/\mathrm{df}_1}{U_2/\mathrm{df}_2}$$

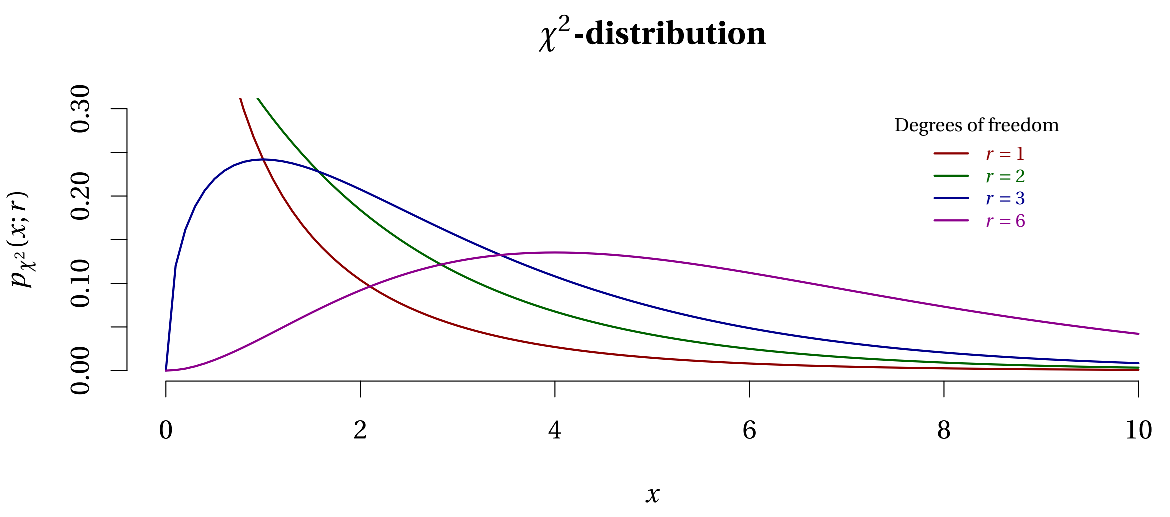


FIGURE C.1 • FOUR χ^2-DISTRIBUTIONS EACH WITH DIFFERENT DEGREES OF FREEDOM r; HERE IS SHOWN $r = 1, 2, 3, 6$. WHEN PERFORMING EITHER A G- OR χ^2-TEST OF INDEPENDENCE THE DEGREES OF FREEDOM ARE DIRECTLY GIVEN BY THE RELATION $(b-1)(c-1)$, WHERE b AND c ARE THE NUMBER OF ROWS AND COLUMNS, RESPECTIVELY IN THE CONSTRUCTED CONTINGENCY TABLE.

the resulting distribution is called a F-distribution with df_1 and df_2. [Cohen, 1988, page 414] refers to df_1 as the degrees of freedom of the numerator and df_2 as that of the denominator.

The pdf for the F-distribution is given by

$$p_F(x; \mathrm{df}_1, \mathrm{df}_2) = \frac{1}{x \cdot B(\mathrm{df}_1/2, \mathrm{df}_2/2)} \left(\frac{\mathrm{df}_1}{\mathrm{df}_2}\right)^{\mathrm{df}_1/2} \cdot x^{\mathrm{df}_1/2-1} \cdot \left(1 + \frac{\mathrm{df}_1}{\mathrm{df}_2} x\right)^{-(\mathrm{df}_1+\mathrm{df}_2)/2},$$

where

$$B(x_1, x_2) = \int_0^1 t^{1-x_1} (1-t)^{1-x_2} \, dt, \quad \text{for} \quad x_1, x_2 > 0,$$

known as the beta function. The mean and variance of the F-distribution are $\mu = \mathrm{df}_2 / (\mathrm{df}_2 - 2)$ and

$$\sigma^2 = \frac{2(\mathrm{df}_2)^2 (\mathrm{df}_1 + \mathrm{df}_2 - 2)}{\mathrm{df}_1 (\mathrm{df}_2 - 2)^2 (\mathrm{df}_2 - 4)} \text{ [Hansen, 2009, page 339, 376 – 377].}$$

The F-distribution is needed when performing an ANOVA and Levene's test as their test statistics (see eq. (4.45) and (4.36), respectively, in section 4.3) will be distributed according to it under the null hypothesis of the respective tests. Four F-distributions are illustrated in figure C.2 with various choices of df_1 and df_2.

C.3 Student's t-distribution

As the name, of this last distribution, suggests it is needed when performing a Student's paired t-test as the test statistic (see eq. (4.50)) will be distributed according to it. Historically, this distribution is named after its discoverer the English mathematician William Gossett who submitted a paper under the pseudonym *Student* wherein it was used. He worked his entire career for the Guinness Brewery [Barlow, 1999, page 135][Hansen, 2009, page 341].

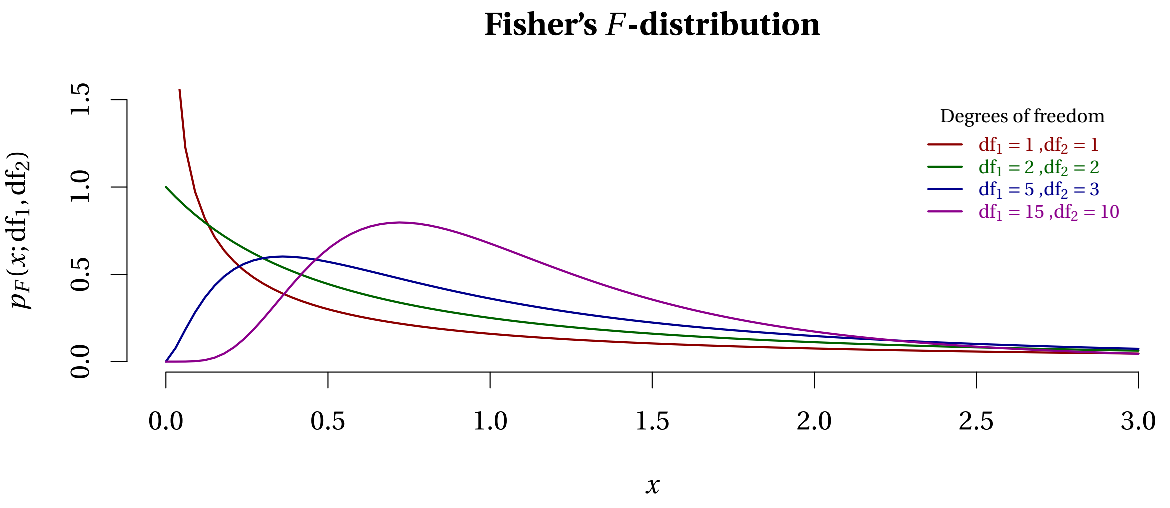


FIGURE C.2 • WHEN A STOCHASTIC VARIABLE DEFINED AS THE FRACTION OF TWO χ^2-DISTRIBUTED VARIBLES THEN IT WILL DISTRIBUTED ACCORDING TO FISHER'S F-DISTRIBUTION. THEREFORE THIS DISTRIBUTION HAS TWO PARAMETERS DF$_1$ AND DF$_2$ WHICH TOGETHER IS THE DEGREES OF FREEDOM. FOUR F-DISTRIBUTION ARE HERE ILLUSTRATED.

Say, we have a stochastic variable A which is normally distributed with mean μ and spread σ. Assume that we drew N samples from a normal distribution x_i with mean μ and spread σ: If we define the stochastic variable Y as

$$Y = \frac{\overline{x} - \mu}{\sigma_{\overline{x}}}, \quad \text{where} \quad \overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{C.2}$$

it will be distributed according to a unit gaussian, i.e zero mean and unit spread. In the case where $\sigma_{\overline{x}}$ is unknown we have to estimate the spread of the population from the sample; hereto we use

$$s_{\overline{x}} = \frac{\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}}{\sqrt{N}}. \tag{C.3}$$

$s_{\overline{x}}$ is an unbiased estimators of $\sigma_{\overline{x}}$ [Barlow, 1999, page 49,77]. The resulting distribution of a new stochastic variable T defined as

$$T = \frac{\overline{x} - \mu}{s_{\overline{x}}}, \tag{C.4}$$

is what we call Student's t-distribution with $\mathrm{df} = N - 1$.

The pdf for the Student's t-distribution is given by

$$p_t(x; k) = \frac{1}{\sqrt{k}B(k/2, 1/2)}\left(1 + \frac{x^2}{k}\right)^{-(k+1)/2} \tag{C.5}$$

with mean 0, about which it is symmetric as shown in figure C.3, and variance $k/(k-2)$ [Hansen, 2009, page 341,377]. The mean and variance are only defined for $k > 1$ and $k > 2$, respectively. It can be shown that the distribution of the transformed stochastic variable T^2 is a F-distribution with $(1, k)$ degrees of freedom [Hansen, 2009, page 342].

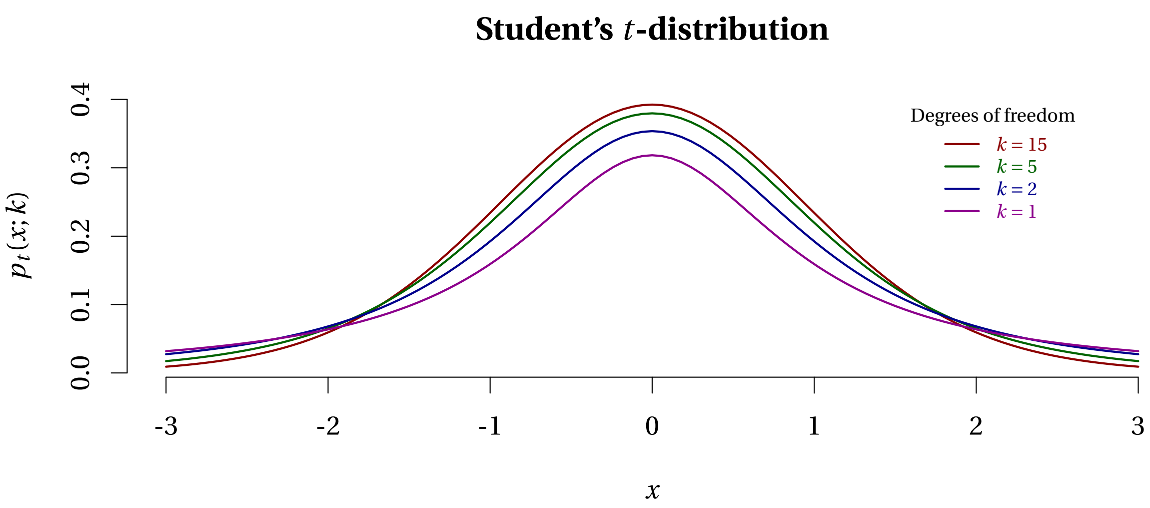


FIGURE C.3 • LIKE THE χ^2-DISTRIBUTION STUDENT'S t ONLY TAKES ONE PARAMETER FOR DEGREES OF FREEDOM. THIS DISTRIBUTION IS SYMMECTRIC AROUND $x = 0$ AND RESEMBLES THE GAUSSIAN, HOWEVER, WITH HEAVIER TAILS.

C.4 Non-central distributions

The χ^2-, F-, and Student's t-distribution, described in the previous sections, are what is called central; all are a specific case of their corresponding non-central, or generalized, distribution. These distributions takes an additional parameter called the *non-centrality parameter*, usually denoted by λ. Here we denote the non-centrality parameters as

- λ_{χ^2} for the non-central χ^2-distribution,
- λ_F for the non-central F-distribution, and
- λ_t for the non-central Student's t-distribution.

When the non-centrality parameter equals zero in the respective non-central distributions they become central, e.g a χ^2-distribution with r degrees of freedom equals a non-central χ^2-distribution with r degrees of freedom and $\lambda_{\chi^2} = 0$. The parameters are restricted as $\lambda_{\chi^2}, \lambda_F \in [0, \infty[$ and $\lambda_t \in]-\infty, \infty[$.

The use of these non-central distributions are especially important when conducting power analysis of, but not restricted to, the independence test, ANOVA, and Student's paired t-test. How the individual test statistics, under their respective null hypothesis, are distributed depends on which test is being conducted as described in chapter 4; under the alternative hypothesis, however, it will be distributed according the corresponding non-central distribution, and it is from this distribution the power of the test is computed. See page 32 for details.

We will not give a detailed explanation as to how the non-central distributions arises but instead present their respective pdf's. The non-central χ^2-distribution:

$$p_{\mathrm{NC},\chi^2}\left(x; r, \lambda_{\chi^2}\right) = \frac{1}{2^{r/2}} \cdot e^{-1/2\left(x+\lambda_{\chi^2}\right)} \sum_{j=0}^{\infty} \frac{x^{\lambda_{\chi^2}/2+j-1} \cdot \left(\lambda_{\chi^2}\right)^j}{\Gamma\left(\lambda_{\chi^2}/2+j\right) \cdot 2^{2j} \cdot j!}, \tag{C.6}$$

non-central F-distribution:

$$p_{\mathrm{NC},F}\,(x;\mathrm{df}_1,\mathrm{df}_2,\lambda_F) = \zeta e^{-\lambda_F/2}\cdot\frac{(\mathrm{df}_1)^{\mathrm{df}_1/2}\cdot(\mathrm{df}_2)^{\mathrm{df}_2/2}\cdot x^{(\mathrm{df}_1-2)/2}}{B\,(\mathrm{df}_1/2,\mathrm{df}_2/2)\cdot(\mathrm{df}_2+\mathrm{df}_1\cdot x)^{(\mathrm{df}_1+\mathrm{df}_2)/2}}, \tag{C.7}$$

where

$$\zeta = 1+\sum_{j=1}^{\infty}\left(\frac{\mathrm{df}_1\cdot\lambda_F\cdot x}{\mathrm{df}_2+\mathrm{df}_1\cdot x}\right)^j, \tag{C.8}$$

and, lastly, the non-central Student's t-distribution:

$$p_{\mathrm{NC},t}\,(x;k,\lambda_t) = e^{-(\lambda_t)^2/2}\cdot\frac{k^{k/2}}{\Gamma\,(k/2)\cdot\sqrt{\pi}\cdot\left(k+x^2\right)^{(k+1)/2}}\sum_{i=0}^{\infty}\Gamma\left(\frac{k+i+1}{2}\right)\cdot\frac{(\lambda_t\cdot x)^i}{i!}\cdot\left(\frac{2}{k+x^2}\right)^{i/2} \tag{C.9}$$

How these non-central distributions depends on their non-centrality parameter is shown in figures C.4, C.5, and C.6. For further mathematical details of these non-central distribution we suggest [Forbes, 2011, page 74,107,187].

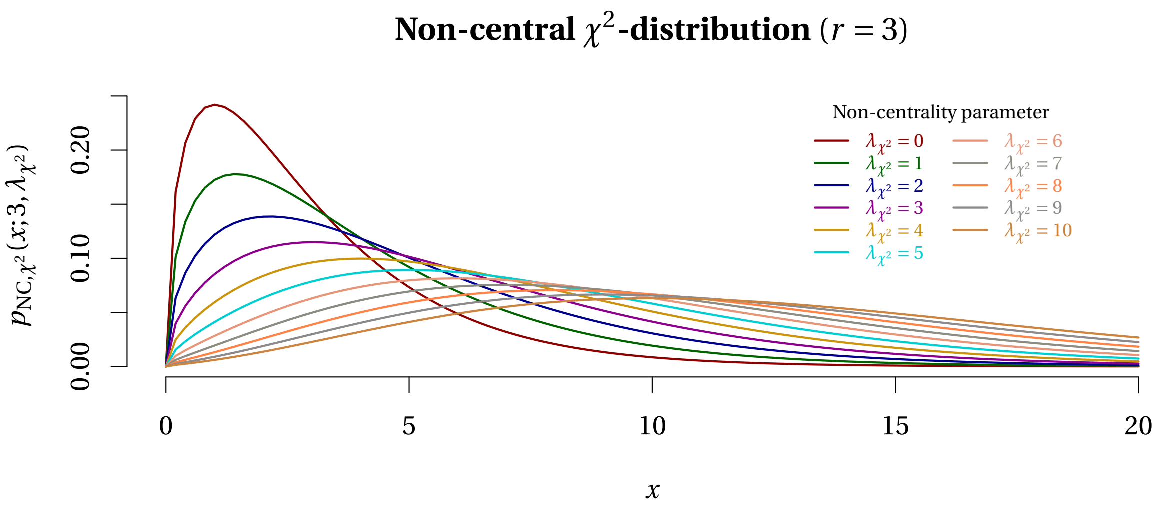


FIGURE C.4 • THE NON-CENTRAL χ^2-DISTRIBUTION WITH $r = 3$. λ_{χ^2} INCREASES HERE WITH INTEGER STEPS OF ONE FROM 0 TO 10. NOTE THAT IN THE INSTANCE WHERE $\lambda_{\chi^2} = 0$ THE DISTRIBUTION IS CENTRAL.

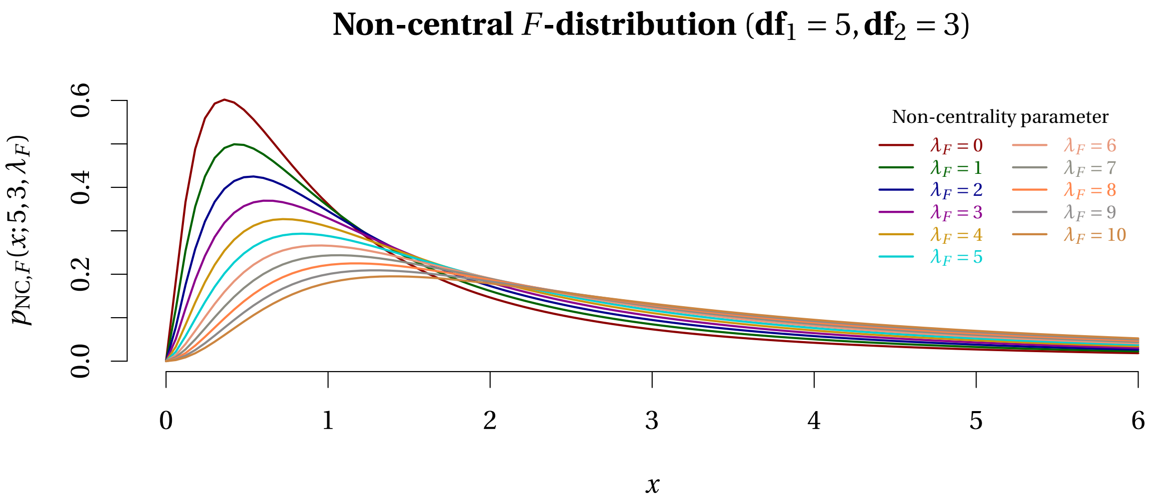


FIGURE C.5 • THE NON-CENTRAL F-DISTRIBUTION WITH $\mathrm{DF}_1 = 5$ AND $\mathrm{DF}_2 = 3$. λ_F INCREASES WITH INTEGER STEPS OF ONE FROM 0 TO 10. NOTE THAT IN THE INSTANCE WHERE $\lambda_F = 0$ THE DISTRIBUTION IS CENTRAL.

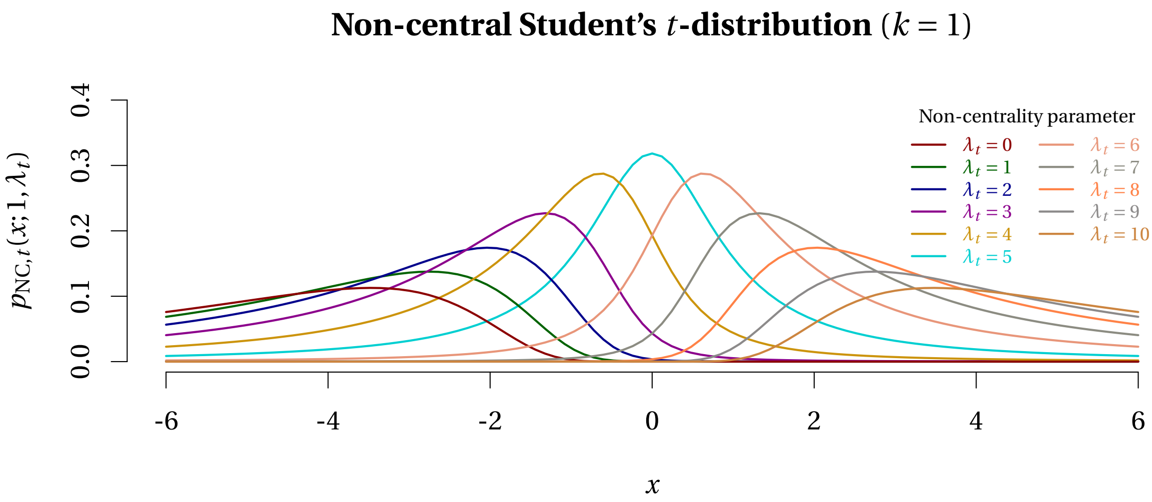


FIGURE C.6 • THE NON-CENTRAL STUDENT'S t-DISTRIBUTION WITH $k = 1$. λ_t INCREASES WITH INTEGER STEPS OF ONE FROM -5 TO 5. NOTE THAT IN THE INSTANCE WHERE $\lambda_t = 0$ THE DISTRIBUTION IS CENTRAL.

APPENDIX

D Additional tables

This appendix contains additional tables that were not included in the main text due to either space considerations or degree of importance.

CONTINGENCY TABLE

	Correct	Wrong	
Used the book ($N = 66$)	389	139	528
Not used the book ($N = 39$)	225	87	312
	614	226	840

RESULTS

	G-**test**	χ^2-**test**
Test statistic	0.241598	0.242328
Effect size	0.016959	0.016985
λ_{χ^2}	0.241598	0.242328
Probability	62.31%	62.25%
Power	7.81%	7.82%

TABLE D.1 • CONTINGENCY TABLE TEST RESULTS FOR DEPENDENCE BETWEEN THE STUDENTS' NTC IN THE LAB TEST AND WHETHER THEY EVER HAVE USED THE BOOK OR NOT, WHEN ONLY CONSIDERING THE QUESTIONS WITHIN THE DATA ANALYSIS GROUP (SEE TABLE 3.1 ON PAGE 15).

CONTINGENCY TABLE

	Correct	Wrong	
Used the book ($N = 66$)	598	128	726
Not used the book ($N = 39$)	354	75	429
	952	203	1155

RESULTS

	G-**test**	χ^2-**test**
Test statistic	0.004098	0.004096
Effect size	0.001884	0.001883
λ_{χ^2}	0.004098	0.004096
Probability	94.90%	94.90%
Power	5.05%	5.05%

TABLE D.2 • CONTINGENCY TABLE AND TEST RESULTS WHEN TESTING FOR DEPENDENCE BETWEEN THE STUDENTS' NTC IN THE LAB TEST AND WHETHER THEY EVER HAVE USED THE BOOK OR NOT, WHEN ONLY CONSIDERING THE QUESTIONS WITHIN THE MATLAB GROUP (SEE TABLE 3.1 ON PAGE 15).

CONTINGENCY TABLE

	Correct	Wrong	
Seen some of the screencasts ($N = 39$)	350	79	429
Not Seen some of the screencasts ($N = 74$)	674	140	814
	1024	219	1243

RESULTS

	G-test	χ^2-test
Test statistic	0.284614	0.286158
Effect size	0.015132	0.015173
λ_{χ^2}	0.284614	0.286158
Probability	59.37%	59.27%
Power	8.32%	8.34%

TABLE D.3 • CONTINGENCY TABLE AND TEST RESULTS FOR TESTING DEPENDENCE BETWEEN THE STUDENTS' NTC IN THE LAB TEST AND WHETHER THEY EVER HAVE SEEN THE SCREENCASTS, OR NOT, WHEN ONLY CONSIDERING THE QUESTIONS WITHIN THE MATLAB GROUP (SEE TABLE 3.1 ON PAGE 15).

CONTINGENCY TABLE

	Correct	Wrong	
Seen any of the screencasts ($N = 39$)	231	81	312
Not seen any of the screencasts ($N = 74$)	422	170	592
	653	251	904

RESULTS

	G-test	χ^2-test
Test statistic	0.778231	0.773035
Effect size	0.029341	0.029243
λ_{χ^2}	0.778231	0.773035
Probability	37.77%	37.93%
Power	14.28%	14.22%

TABLE D.4 • CONTINGENCY TABLE AND TEST RESULTS WHEN TESTING FOR DEPENDENCE BETWEEN THE STUDENTS' NTC IN THE LAB TEST AND WHETHER THEY EVER HAVE SEEN THE SCREENCASTS, OR NOT, WHEN ONLY CONSIDERING THE QUESTIONS WITHIN THE DATA ANALYSIS GROUP (SEE TABLE 3.1 ON PAGE 15).

CONTINGENCY TABLE

	Correct	Wrong	
Not used the book ($N = 33$)	331	428	759
Never heard of the book ($N = 26$)	286	312	598
	617	740	1357

RESULTS

	G-test	χ^2-test
Test statistic	2.397159	2.397852
Effect size	0.042030	0.042036
λ_{χ^2}	2.397159	2.397852
Probability	12.16%	12.15%
Power	34.05%	34.06%

TABLE D.5 • CONTINGENCY TABLE AND TEST RESULTS WHEN TESTING FOR DEPENDENCE BETWEEN THE STUDENTS' NTC OF THE DHD AND WHETHER THEY LATER DID NOT USE OR NEVER HEARD OF THE BOOK.

CONTINGENCY TABLE
(USED THE BOOK)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	13	5	18
DHD Q.17 wrong	22	19	41
	35	24	59

RESULTS

	McNemar's χ^2-test
Test statistic	10.703704
Effect size	0.425933
Odds ratio	0.227273
λ_{χ^2}	10.703704
Probability	0.11%
Power	90.52%

TABLE D.6 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF A PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 59 STUDENTS WHO USED THE BOOK, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(NOT USED THE BOOK)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	7	3	10
DHD Q.17 wrong	12	11	23
	19	14	33

RESULTS

	McNemar's χ^2-test
Test statistic	5.400000
Effect size	0.404520
Odds ratio	0.250000
λ_{χ^2}	5.400000
Probability	2.01%
Power	62.20%

TABLE D.7 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF A PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 33 STUDENTS WHO DID NOT USE THE BOOK, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(NEVER HEARD OF THE BOOK)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	7	3	10
DHD Q.17 wrong	10	6	16
	17	9	26

RESULTS

	McNemar's χ^2-test
Test statistic	3.769231
Effect size	0.380750
Odds ratio	0.300000
λ_{χ^2}	3.769231
Probability	5.20%
Power	49.27%

TABLE D.8 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF A PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 26 STUDENTS WHO CLAIMED TO NEVER HAVE HEARD OF THE BOOK, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(USED THE SCREENCASTS)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	6	2	8
DHD Q.17 wrong	14	9	23
	20	11	31

RESULTS

	McNemar's χ^2-test
Test statistic	9.000000
Effect size	0.538816
Odds ratio	0.142857
λ_{χ^2}	9.000000
Probability	0.27%
Power	85.08%

TABLE D.9 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF AN PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 31 STUDENTS WHO USED THE SCREENCASTS, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(NOT USED THE SCREENCASTS)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	16	7	23
DHD Q.17 wrong	22	22	44
	38	29	67

RESULTS

	McNemar's χ^2-test
Test statistic	7.758621
Effect size	0.340294
Odds ratio	0.318182
λ_{χ^2}	7.758621
Probability	0.53%
Power	79.54%

TABLE D.10 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF AN PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 67 STUDENTS WHO DID NOT USE THE SCREENCASTS, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(NEVER HEARD OF THE SCREENCASTS)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	6	2	8
DHD Q.17 wrong	8	5	13
	14	7	21

RESULTS

	McNemar's χ^2-test
Test statistic	3.600000
Effect size	0.414039
Odds ratio	0.250000
λ_{χ^2}	3.600000
Probability	5.78%
Power	47.51%

TABLE D.11 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF AN PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 21 STUDENTS WHO CLAIMED TO NEVER HAVE HEARD OF THE SCREENCASTS, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(FOUND INTERNET MOST EDUCATIONAL)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	14	6	20
DHD Q.17 wrong	22	17	39
	36	23	59

RESULTS

	McNemar's χ^2-test
Test statistic	9.142857
Effect size	0.393654
Odds ratio	0.272727
λ_{χ^2}	9.142857
Probability	0.25%
Power	85.63%

TABLE D.12 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF AN PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 59 STUDENTS WHO FOUND THE INTERNET MOST EDUCATIONAL, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(SCREENCASTS MOST EDUCATIONAL)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	3	1	4
DHD Q.17 wrong	3	4	7
	7	4	11

TABLE D.13 • THIS TABLE SHOWS THE CONTINGENCY TABLE FOR THE 11 STUDENTS WHO FOUND THE SCREENCASTS MOST EDUCATIONAL.

CONTINGENCY TABLE
(BOOKS MOST BENEFICIAL)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	15	5	20
DHD Q.17 wrong	20	15	35
	35	20	55

RESULTS

	McNemar's χ^2-test
Test statistic	9.000000
Effect size	0.404520
Odds ratio	0.250000
λ_{χ^2}	9.000000
Probability	0.27%
Power	85.08%

TABLE D.14 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF AN PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 55 STUDENTS WHO FOUND BOOKS IN GENERAL MOST BENEFICIAL WHEN LEARNING A NEW TOPIC, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(NOTES ON THE INTERNET MOST BENEFICIAL)

	Lab Test Q.21 correct	Lab Test Q.21 wrong	
DHD Q.17 correct	2	2	4
DHD Q.17 wrong	12	8	20
	14	10	24

RESULTS

	McNemar's χ^2-test
Test statistic	7.142857
Effect size	0.545545
Odds ratio	0.166667
λ_{χ^2}	7.142857
Probability	0.75%
Power	76.20%

TABLE D.15 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF AN PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 24 STUDENTS WHO FOUND NOTES ON THE INTERNET MOST BENEFICIAL WHEN LEARNING A NEW TOPIC, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 21 ON THE LAB TEST) AND AFTER (QUESTION 17 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(USED THE BOOK)

	Lab Test Q.22 correct	Lab Test Q.22 wrong	
DHD Q.5 correct	26	5	31
DHD Q.5 wrong	17	11	28
	43	16	59

RESULTS

	McNemar's χ^2-test
Test statistic	6.545455
Effect size	0.333076
Odds ratio	0.294118
λ_{χ^2}	6.545455
Probability	1.05%
Power	75.52%

TABLE D.16 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF A PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 59 STUDENTS WHO LATER USED THE BOOK, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 22 ON THE LAB TEST) AND AFTER (QUESTION 5 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(USED THE SCREENCASTS)

	Lab Test Q.22 correct	Lab Test Q.22 wrong	
DHD Q.5 correct	9	5	14
DHD Q.5 wrong	10	7	17
	19	12	31

RESULTS

	McNemar's χ^2-test
Test statistic	1.666667
Effect size	0.231869
Odds ratio	0.500000
λ_{χ^2}	1.666667
Probability	19.67%
Power	25.23%

TABLE D.17 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF A PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 31 STUDENTS WHO LATER USED THE SCREENCASTS, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 22 ON THE LAB TEST) AND AFTER (QUESTION 5 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(FOUND THE INTERNET MOST EDUCATIONAL)

	Lab Test Q.22 correct	Lab Test Q.22 wrong	
DHD Q.5 correct	32	3	35
DHD Q.5 wrong	11	13	24
	43	16	59

RESULTS

	McNemar's χ^2-test
Test statistic	4.571429
Effect size	0.278356
Odds ratio	0.272727
λ_{χ^2}	4.571429
Probability	3.25%
Power	57.07%

TABLE D.18 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF A PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 59 STUDENTS WHO LATER FOUND THE INTERNET WAS MOST EDUCATIONAL, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 22 ON THE LAB TEST) AND AFTER (QUESTION 5 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.

CONTINGENCY TABLE
(FOUND BOOKS IN GENERAL MOST BENEFICIAL)

	Lab Test Q.22 correct	Lab Test Q.22 wrong	
DHD Q.5 correct	26	6	32
DHD Q.5 wrong	13	10	23
	39	16	55

RESULTS

	McNemar's χ^2-test
Test statistic	2.578947
Effect size	0.216541
Odds ratio	0.461538
λ_{χ^2}	2.578947
Probability	10.83%
Power	36.18%

TABLE D.19 • THIS TABLE SHOWS THE CONTINGENCY TABLE AND RESULTS OF A PAIRED INDEPENDENCE TEST, WHEN CONSIDERING THE 55 STUDENTS WHO LATER CLAIMED THAT BOOKS IN GENERAL WAS MOST BENEFICIAL WHEN LEARNING A NEW TOPIC, TO INVESTIGATE IF THERE IS A CORRELATION BETWEEN ANSWERING THE SAME QUESTION BEFORE (QUESTION 22 ON THE LAB TEST) AND AFTER (QUESTION 5 ON THE DHD) THE CURRICULUM OF MEK1 WAS TAUGHT.