

# Undersøgelse af analysemetode for bestemmeselse af enzym i dyrefoder

Rikke Stubbe Hansen

Kongens Lyngby 2008  
IMM-Eksamensprojekt-2008-4. januar

Danmarks Tekniske Universitet  
Informatik og Matematisk Modellering  
Bygning 321, 2800 Kongens Lyngby  
Telefon 45253351, Fax 45882673  
[reception@imm.dtu.dk](mailto:reception@imm.dtu.dk)  
[www.imm.dtu.dk](http://www.imm.dtu.dk)

# Resumé

---

Dette projekt undersøger en af Novozymes nyudviklede analysemetode for bestemmelse af mængden af enzym i dyrefoder. Analysemetoden kan anvendes på både tilsætning af flydende og faste enzymer til dyrefoder. Der er blevet udført to forskellige forsøg, A og B. Forsøg A anvendes til bestemmelse af laboratoriernes repeterbarhed, mens forsøg B anvendes til undersøgelse af dag til dags variationen.

Der er blevet foretaget en grafisk og beregnings undersøgelse af resultaterne for begge forsøg. De grafiske billeder beskriver hvordan resultaterne ligger, for hvert af laboratorierne, for hvert af de specifikke udvalgte materialer, og i forhold til alle materialerne. Beregningerne beskriver materialernes gennemsnit, median, varians, spredning, konfidensinterval, repeterbarhed og reproducerbarhed. Beregningerne anvendes til sammenligning i de efterfølgende analyser.

Der er blevet anvendt to forskellige slags metoder til undersøgelse af resultaterne.

Den første metode fjerner outliers fra resultaterne ved hjælp af Cochran's test og Grubbs' test. Anvendelse af outlier analyse, betyder for forsøg A, at alle resultaterne for et af laboratorierne fjernes fra resultaterne, mens det for forsøg B betyder, at nogle af resultaterne for flere af materialerne fjernes. Der dannes hermed to nye datasæt, for hveraf forsøgene, som ikke indeholder outliers. Datasættene anvendes senere i en mixed model analyse. Den anden metode, en robusthedsanalyse, fjerner ikke nogle af resultaterne fra datasættene, men ændrer på nogle af resultaterne.

Udfra medianen dannes der en øvre og nedre grænseværdi. De resultater som ligger udenfor grænseværdierne, ændres til den nærmeste grænseværdi. Der dannes hermed igen to nye datasæt for hvert af forsøgene.

Der opbygges herefter en passende mixed model for hvert af forsøgene. Modellen anvendes til analyse af de rå datasæt, data uden outlier og data efter en robusthedsanalyse. Resultatet efter en mixed model analyse giver for hvert af datasættene usikkerheden på repeterbarheden og reproducibarheden. Usikkerhederne kan af Novozymes anvendes i deres løbende kontrol af enzym mængden i dyrefoder.

Der er til sidst givet et bud på en ny måde at dimensionere Novozymes fremtidige ringtest. Diminsioneringen tager udgangspunkt i forsøg B for hvert materiale. Resultaterne for spredningen for reproducibarheden svinger for de enkelte materialer, hvilket betyder, at diminsioneringen for nogle materialer giver få laboratorier med få dage og mange gentagelser og nogle materialer giver mange laboratorier med middel antal dage og få gentagelser.

# Summary

---

This project investigates a newly developed by Novozymes analytical method for estimation of the quantity of enzyme in animal feed. The analytical method can be applied for both fluid and solid enzyme added to the animal feed. Two different tests, A and B, are carried out where A is used for estimation of the repeatability of laboratories and B is used for estimation of the day-to-day variance.

There has been made a graphical investigation and calculations of the results both tests. The graphical investigation shows how the results are spread for each laboratory, for each of the specific selected materials and in relation to all the materials. The calculations give the average, median, variance, standard deviation, confidence interval, repeatability and reproducibility of the materials. These calculations are used for comparison in the following analysis. Two different types of methods are applied for investigating the results. The first method removes outliers from the results by using Cochran's and Grubbs' test. The application of outlier analysis have the impact for test A that all the results for one laboratory are removed, while it for test B means that some of the results for a number of laboratories are removed. Consequently, two new data sets, without outliers, are created and these data sets are used in a mixed model analysis later on. The second method, the robust methods for data analysis, doesn't remove any results from the data sets, but changes some of the results. A robust methods for data analysis is used for (by iteration) finding the optimal median. From the median is created an upper and a lower limiting value and those results being outside these limiting values are changed to the nearest limiting value. This implies that two new data sets are created - one for each of the tests.

An adequate mixed model for each of the tests is build. The model is used for the analysis of the raw data, data without outliers and the resulting data from the robust methods for data analysis. The results of the mixed model analysis for each of the data sets give the standard deviation on the repeatability and the reproducibility. The standard deviation can be used by Novozymes in their ongoing quality assurance of the quantity of enzyme in animal feed.

Lastly is given a new proposal to dimension Novozymes future ring test and this dimension uses test B as the basis. The results of the standard deviation of the reproducibility is dependent on the materials, which means that the dimension for some material gives few laboratories with few days and many repetitions and for some materials gives many laboratories with a limited number of days and few repetitions.

# Forord

---

Dette er et eksamensprojekt som afslutning på Civilingeniørstudiet på Danmarks Tekniske Universitet, DTU. Eksamensprojektet er gennemført i samarbejde med Novozymes, Project Management og Information ved Peter Thyregod og Informatik og Matematisk Modellering, IMM-DTU, ved docent Helle Rootzen. Jeg vil gerne takke docent Henrik Splid fra IMM-DTU, som har bidraget til kapitlet om redimensionering af forsøgsplanen.

**Januar 2008, Rikke Stubbe Hansen**





# Indhold

---

<b>Resumé</b>	<b>i</b>
<b>Summary</b>	<b>iii</b>
<b>Forord</b>	<b>v</b>
<b>1 Indledning</b>	<b>1</b>
1.1 Baggrund . . . . .	2
1.2 Formål . . . . .	2
1.3 Metode . . . . .	3
1.4 Rapportens opbygning . . . . .	4
1.5 Væsentlige resultater . . . . .	5
<b>2 Præsentation af ISO 5725 og data</b>	<b>7</b>
2.1 Formål med præsentation af data . . . . .	8
2.2 Data . . . . .	8

2.2.1	Forsøg A . . . . .	9
2.2.1.1	Statistiske beregninger . . . . .	10
2.2.1.2	Varianshomogenitet test . . . . .	15
2.2.1.3	Variansanalyse . . . . .	18
2.2.2	Forsøg B . . . . .	19
2.2.2.1	Statistiske beregninger . . . . .	21
2.2.2.2	Varianshomogenitet test . . . . .	21
2.2.2.3	Variansanalyse . . . . .	21
2.2.3	Sammenligning af materialerne fra forsøgene . . . . .	28
2.3	Konklusion . . . . .	28
<b>3</b>	<b>Outlier undersøgelse</b>	<b>31</b>
3.1	Formål . . . . .	31
3.2	Outlier . . . . .	32
3.3	Cochran´s test . . . . .	33
3.4	Grubbs´test . . . . .	35
3.5	Forsøg A . . . . .	38
3.5.1	Grafisk afbildning af forsøg A . . . . .	38
3.5.2	Beregninger for forsøg A . . . . .	40
3.5.2.1	Statistiske beregninger . . . . .	46
3.6	Forsøg B . . . . .	47
3.6.1	Grafisk afbildning af forsøg B . . . . .	49
3.6.2	Beregninger for forsøg B . . . . .	51

---

3.6.2.1	Statistiske beregninger . . . . .	55
3.7	Sammenligning af forsøg A og B . . . . .	55
3.8	Konklusion . . . . .	55
<b>4</b>	<b>Robusthedsanalyse</b>	<b>61</b>
4.1	Formål . . . . .	61
4.2	Robusthedsanalyse . . . . .	61
4.3	Algoritme A . . . . .	64
4.4	Algoritme S . . . . .	65
4.5	Resultater . . . . .	67
4.5.1	Resultater fra forsøg A . . . . .	67
4.5.2	Resultater fra forsøg B . . . . .	74
4.6	Konklusion . . . . .	82
<b>5</b>	<b>Mixed model analyse</b>	<b>85</b>
5.1	Formål . . . . .	85
5.2	Mixed model metode . . . . .	85
5.3	Fixed, random og mixed faktor analyse . . . . .	86
5.4	Mixed model teori . . . . .	87
5.5	Model optimeringsmetoder . . . . .	90
5.6	Hierarkisk inddeling . . . . .	93
5.7	Resultater for forsøg A . . . . .	98
5.8	Resultater for forsøg B . . . . .	100

5.9	Usikkerheden af repeterbarheden . . . . .	102
5.10	Konklusion . . . . .	103
<b>6</b>	<b>Diminsionering af ny ringprøvning</b>	<b>105</b>
6.1	Formål . . . . .	105
6.2	Teoretisk implementering . . . . .	105
6.3	Implementering af metoden . . . . .	108
6.4	Reultater . . . . .	110
6.5	Konklusion . . . . .	114
<b>7</b>	<b>Konklusion</b>	<b>117</b>
<b>A</b>	<b>Appendiks</b>	<b>121</b>
A.1	Metode validering . . . . .	121
A.2	ISO 5725 . . . . .	122
A.3	Statistiske beregninger . . . . .	124
A.4	Varianshomogenitettest . . . . .	127
<b>B</b>	<b>Appendiks</b>	<b>129</b>
B.1	Liste af anvendte laboratorier . . . . .	130
B.2	Resultater for forsøg A . . . . .	145
B.2.1	Tabel med rådata . . . . .	145
B.2.2	Resultater vist grafisk . . . . .	145
B.2.3	Beregnete resultater . . . . .	145

B.3	Resultater for forsøg B . . . . .	145
B.3.1	Tabel med rådata . . . . .	145
B.3.2	Resultater vist grafisk . . . . .	145
B.3.3	Beregnete resultater . . . . .	145



## KAPITEL 1

# Indledning

---

Novozymes er verdens førende producent af enzymer og mikroorganismer til industriel brug, og de tilbyder biologiske løsninger på industriens problemer.

Enzymer er katalysatorer. Det betyder, at de får biokemiske reaktioner til at ske hurtigere, end de ellers ville have gjort. Visse livsvigtige reaktioner ville slet ikke finde sted uden hjælp fra enzymer. At enzymer er katalysatorer betyder også, at de ikke udgør en del af det endelige produkt. De får altså ting til at ske. Når arbejdet er udført, er enzymerne klar til at katalysere en ny biokemisk reaktion.

Novozymes beskæftigelse indefor enzymer dækker over, tekniske enzymer til vaskemidler, fødevarerenszymer, mikroorganismer og foderenszymer. Dette eksamensprojekt anvender data fra enzymer i dyrefoder.

Tilsætning af enzymer til dyrefoder øger dets næringsværdi og forbedrer dyrenes fosforabsorption. Enzymtilsætningen betyder for dyrene, at de vokser hurtigere og udskiller mindre fosfor via gødningen, hvilket gavner miljøet.

## 1.1 Baggrund

Novozymes har i samarbejde med andre enzymproducenter udviklet en ny industristandard til bestemmelse af enzymmængden i dyrefoder <sup>1</sup>. Den ny standard kan anvendes på alle flydende og faste enzymer, der findes på markedet i dag. Standarden undersøges gennem en ringprøve med 14 laboratorier, hvor der foretages to slags forsøg, A og B. De 14 laboratorier har først gennemgået en oplæringsperiode og efterfølgende afprøvet standarden hjemme i deres eget laboratorium. De 14 laboratorier er derfor alle bekendt med standarden og de er blevet godkendt til at analysere efter den nye standard. De to forsøg A og B er opbygget på følgende måde:

- Forsøg A: Forsøget følger standarden ISO 5725-2, Grundlæggende metode til bestemmelse af repeterbarhed og reproducerbarhed for en standardiseret målemetode. I forsøget analyseres fem enzymmaterialer, som er tilsat til dyrefoderet. Enzymet er i to af dyrefoderprodukterne flydende (materiale 1 og 2) og i de sidste tre dyrefoderprodukter faste. Hvert af laboratorierne analyserer alle materialerne på en dag. Hvert materiale er blindt repræsenteret to gange. Der gennemføres således ti forsøg pr. laboratorium.
- Forsøg B: Der undersøges otte dyrefoderprodukter seks gange over tre dage, hvor af fem af materialerne er de samme som i forsøg A. Forsøget følger standarden ISO 5725-3, Mellemliggende mål for præcision af en standardiseret målemetode. Enzymet er i fire af dyrefoderprodukterne flydende (materiale 2, 4, 5 og 7) og i de sidste fire dyrefoderprodukter faste. Hvert af laboratorierne analyserer alle analyserne på tre uafhængige dage, svarende til i alt 48 forsøg pr. laboratorium.

## 1.2 Formål

Formålet med projektet er, at lave en dybdegående undersøgelse af resultaterne for begge forsøg. Undersøgelsen vil give en grundig beskrivelse af resultaterne. Novozymes fjerner i dag outlierer fra deres forsøgesresultater, men det skal undersøges om denne metode kan erstattes af en robusthedsanalyse, som bibeholder alle forsøgsresultaterne.

Resultaterne fra den nye analysemetode skal undersøges for nøjagtighed, ved bestemmelse af repeterbarheden og reproducerbarheden.

Repeterbarheden beskriver hvor nøjagtig et laboratorium bestemmer deres resultater for et materiale.



Reproducerbarheden beskriver hvor nøjagtig alle laboratorierne samlet bestemmer deres resultater for et materiale.

Analysen af repeterbarheden og reproducerbarheden undersøges for de rå data, for data uden outlierer, for data efter en robusthedsanalyse og ved anvendelse af en mixed model analyse.

I en robusthedsanalyse undersøges resultaterne, om de ligger indenfor nogle beregnede grænseværdier. De resultater som ligger udenfor grænseværdierne fjernes ikke, men ændres til den pågældende grænseværdi.

Usikkerheden for standardafvigelsen for repeterbarheden findes efter en mixed model analyse. Usikkerheden skal anvendes til løbende kvalitetskontrol hos Novozymes.

Der beregnes et konfidensinterval for hvert af materialerne, ud fra materialets gennemsnit indenfor alle laboratorier, i hvert af de tre datasæt og de sammenlignes løbende med hinanden. Konfidensintervallet anvendes til at vise, hvordan resultaterne evt. påvirkes af fjernelse af outlier og robusthedsanalysen.

Standardafvigelsen for laboratorium, gentagelser, dag og repeterbarheden anvendes til redimensionering.

## 1.3 Metode

Der dannes i alt tre datasæt. Et datasæt med de rå data, som er resultaterne fra forsøgene, et uden outlier og et efter robusthedsanalysen.

Konfidensintervallet for materialets gennemsnit og standardafvigelsen for repeterbarheden og reproducerbarheden for de tre datasæt beregnes og sammenlignes. Sammenligningen beskriver hvordan de rå data er blevet påvirket, ved fjernelse af outlier fra datasættet og ændringer i datasættet ved anvendelse af robusthedsanalysen.

Efter de indledende undersøgelser opbygges der en model, som beskriver data ved anvendelse af mixed model. Mixed model giver mulighed for, at beskrive data fixed og random effekter i en model. Fixed effekten er materialer, som ikke er tilfældig udtaget. Random effekten er laboratorier, som fremover skal anvendes på et tilfældig laboratorier, og ikke et af de 14 anvendte laboratorier. Udfra mixed model modellen, beregnes standardafvigelsen for repeterbarheden og reproducerbarheden og deres usikkerheder.

Undersøgelse af dag til dag variation af materialerne fra de tre datasæt gennemføres ved anvendelse af mixed model. Dag til dag variationen er blevet undersøgt på tre tilfældige dage, med i alt seks prøver.

Der er ud fra standardafvigelsen for laboratorium, gentagelser, dage og repeterbarheden beregnet den optimale dimension for det pågældende materiale.

## 1.4 Rapportens opbygning

- Kapitel 2. Indeholder en beskrivelse af de to udførte forsøg i henhold til ISO standarden.  
Der foretages nogle statistiske beregninger af begge forsøg, som beskriver hvilke resultater der vil blive arbejdet med i projektet.  
Der introduceres hvilke statistiske beregninger som vil blive foretaget i de næste kapitler. De statistiske beregninger anvendes til sammenligning af resultaterne efter en statistisk databehandling.
- Kapitel 3. I dette kapitel laves der en outlier test på forsøgene. Outlier testen skal undersøge om nogle af resultaterne kan klassificeres som værende en outlier. Hvis et resultat bliver klassificeret som værende en outlier, vælges der i dette projekt at fjerne resultatet fra forsøget.  
De statistiske undersøgelser der blev præsenterede i kapitel 2, anvendes til analyse af data efter fjernelse af outlier fra datasættet.  
Konfidensintervallerne for materialerne og repeterbarheden og reproducerbarheden, sammenlignes med resultaterne fra de rå data.
- Kapitel 4. I dette kapitel laves der en robusthedsanalyse på forsøgene. I en robusthedsanalyse fjernes der ingen data, men data ændres til de beregnede øvre og nedre grænseværdier.  
Fordelen ved anvendelse af en robusthedsanalyse i forhold til outlier testen er, at alle data bibeholdes.  
De statistiske undersøgelser anvendes på data efter de er blevet ændret ved anvendelse af robusthedsanalysen. Resultaterne for konfidensintervallerne, repeterbarhed og reproducerbarhed, sammenlignes med resultaterne for de rå data og data uden outlier.
- Kapitel 5. I dette kapitel opbygges der en statistisk model for hver af forsøgene. Modellen opbygges ved anvendelse af mixed model teorien. En mixed model indeholder både fixed og random faktorer i en model.  
De modeller som beskriver forsøgene bedst, anvendes til beregning af repeterbarheden og reproducerbarheden for de tre datasæt, de rå data, data uden outlier og robusthedsdata.  
Resultaterne sammenlignes med de beregninger, som er foretaget i kapitel 2, 3 og 4.  
Usikkerheden for repeterbarheden og reproducerbarheden beregnes, og skal anvendes hos Novozymes i deres løbende kvalitetskontrol af enzymprodukter.
- Kapitel 6. I dette kapitel redimensioneres forsøgene ud fra de funde resultater. Redimensioneringen skal gerne give en mindre forsøgsmodel, hvormed der kan spares tid og penge på de fremtidige ringprøver.
- Kapitel 7. Konklusion på projektet.

## 1.5 Væsentlige resultater

- De resultater der er opnået ved anvendelse af robusthedsanalysen, sammenlignet med de resultater der er opnået efter fjernelse af de påviste outlierer, giver næsten de samme resultater. Robusthedsanalysen giver lidt bedre resultater end fjernelse af outlier.
- Der beregnes usikkerheden af repeterbarheden og reproducerbarheden til anvendelse i fremtidige kvalitetskontrol af Novozymes.
- Der er blevet foretaget en detaljeret undersøgelse af begge forsøg, hvor anvendelse af robusthedsanalyse og mixed model er nyt for Novozymes. De har hidtid anvendt outlier test i deres analyser.
- Diminsioneringen af en ny ringprøvning ud fra de udførte forsøg A og B. Der er blevet anvendt spredningen for laboratoriet, dage, gentagelser og reproducerbarheden for de anvendte materialer.



## KAPITEL 2

# Præsentation af ISO 5725 og data

---

I det følgende kapitel beskrives kort rammerne for projektet, terminologien i ISO 5725, samt en simpel statistisk gennemgang af data.

Novozymes har i samarbejde med andre enzymproducenter udviklet en ny industristandard metode til bestemmelse af enzymmængden i dyrefoder <sup>1</sup>. Det vigtige ved den nye metode, sammenlignet med de eksisterende metoder er dens bedre anvendelse på flere enzymprodukter.

Forsøgsresultaterne er blevet evalueret af European Committee for Standardisation (CEN) for at undersøge om den nye analytiske metode kan godkendes til at være en CEN standard.

### **Design af undersøgelse**

Arbejdet er en del af et validerings projekt udarbejdet af European Association of Feed Additive Manufacturers (FEFANA), hvor det er Institute for Reference Materials and Measurement (IRMM) der har designet undersøgelsen.

Der er blevet udvalgt nogle laboratorier til testning af den nye metode, de har først være involveret i et træningsforløb udarbejdet af FEFANA. Træningen har indholdt specifikke øvelser i metoden på ukendte materialer. Derefter har

laboratorierne haft en periode, hvor de selv har kunne afprøve metoden i deres eget laboratorium. De har som afslutning være med i et interlaboratorium forsøg. De i alt 14 laboratorier som bestod træningsperioden og den efterfølgende test, har været med til undersøgelse af den nye metode. De 14 laboratorier kan ses bagerst i appendisk B.

## 2.1 Formål med præsentation af data

Formålet med den indledende beskrivelse af datasættene er, at det skal fungere som en introduktion til de udførte forsøg og resultater. Der vil i dette kapitel blive lavet en gennemgang af hvert forsøg og den tilhørende variansanalyse tager udgangspunkt i, at alle faktorerne er fixed. Det at en faktor er fixed betyder, at i dette tilfælde at der analyseres for de specifikke materialer og laboratorier. Resultaterne fra hvert materiale og alle materialer tilsammen fra hvert forsøg vil blive vist grafisk. Der vil blive udregnet materialernes gennemsnit, varians, spredning, repeterbarhed og reproducerbarhed. Konfidensintervallerne beregnes for hvert af materialerne, de flydende og faste materialer og alle materialerne tilsammen. Konfidensintervallet anvendes til beskrivelse af resultaterne af enzyemmængden i dyrefoder.

## 2.2 Data

Der er foretaget to forsøg, forsøg A og forsøg B. I begge forsøg er det de samme 14 laboratorier, der har analyseret efter den tilsendte analysemetode, foderprodukternes enzymindehold.

Forsøg A følger standarden ISO 5725-2, Grundlæggende metode til bestemmelse af repeterbarhed og reproducerbarhed for en standardiseret målemetode <sup>1</sup>. Forsøget er udført som et double blind testforsøg, hvor der er blevet tilsendt ti materialer bestående af fem foderprodukter. Hvert af laboratorierne skulle analysere alle materialerne på en dag.

Forsøg B følger standarden ISO 5725-3, Mellemliggende mål for præcision af en standardiseret målemetode <sup>1</sup>. I forsøget er der blevet analyseret otte foderprodukter. Hver af laboratorierne skulle analysere alle materialerne på tre uafhængige dage, men alle otte materialer på samme dag. Der er blevet tilsendt otte materialer til laboratorierne, hvor de selv skulle udtage to prøver til analyse den pågældende dag.

Resultaterne fra forsøgene kan ses bagerst i appendiks B.

Enzymet, der er blevet tilsat fodet, har enten været på flydende eller fast form. Tabel 2.1 viser hvilken form enzymet har været før tilsætningen til fodet og

Forsøg A	Forsøg B	Form
Materiale 1	Materiale 5	Flydende
Materiale 2	Materiale 2	Flydende
Materiale 3	Materiale 1	Fast
Materiale 4	Materiale 8	Fast
Materiale 5	Materiale 6	Fast
	Materiale 3	Fast
	Materiale 4	Flydende
	Materiale 7	Flydende

Tabel 2.1: Hvilke materialer der er ens for forsøgene og hvilken form enzymet har været

hvilke enzymer der er ens for forsøg A og B.

De 14 laboratorier har haft en træningsperiode af FEFANA med efterfølgende træning i deres eget laboratorium. Derefter blev de rigtige materialer sendt til laboratorierne. De 14 laboratorier vil automatisk være mere omhyggelige med analysen, eftersom de skal anvende en ny analysemetode. De opnåede resultater kan derfor antages at være gode resultater.

Der er blevet taget logaritmen til alle resultaterne i begge forsøg. Det skyldes, at det giver nogle bedre resultater at analysere og de efterfølgende beregnede spredninger for repeterbarheden og reproducerbarheden bliver mere anvendelige.

## 2.2.1 Forsøg A

Forsøg A er et balanceret forsøg, der anvendes som basisforsøg. Det betyder, at forsøget er opbygget af  $q$  materialer, og i dette tilfælde er  $q = 5$ .  $j$  anvendes som tællevariabel for  $q$ . Forsøget er lavet af  $p$  laboratorier, der hver har analyseret præcis  $n$  gentagende test resultater under de samme omstændigheder for hver af de  $q$  materialer. Der er i forsøget anvendt  $p = 14$  laboratorier, med  $i$  som tællevariabel.

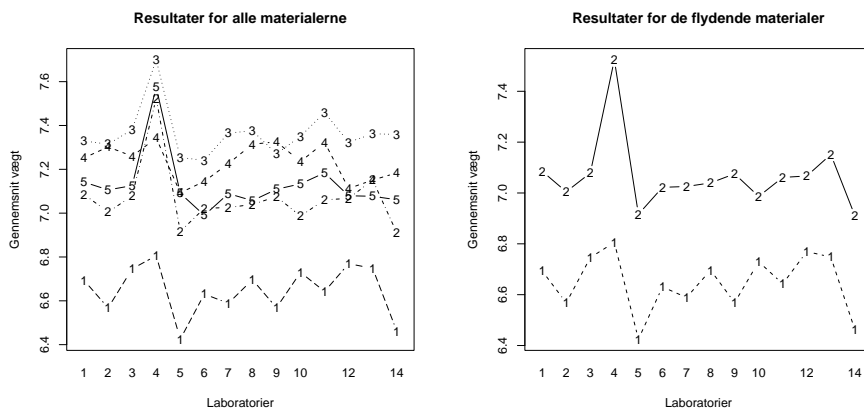
Tabel 2.2 viser hvordan forsøg A er opbygget.

Resultaterne af forsøget ses i figur 2.1, som viser alle materialerne. Det kan ud fra figuren ses, at laboratorium 4 generelt har et højere resultat for hvert af materialerne, end for de andre laboratorier. Figur 2.2 viser resultaterne for de to

<i>Lab. \ Materiale</i>	1	2	j	$q = 5$
1	$n = 2$			
2				
i			$(y_{ij1}, y_{ij2})$	
$p = 14$				$y_{ijk}$

Tabel 2.2: Opbygning af forsøg A

flydende enzymer, og de antager begge en vandret linie, men nogle udsving. Det gælder specielt for laboratoium 4 for materiale 2. Figur 2.3 viser resultaterne for de faste enzymer. Hvor der gælder det samme, som for de flydende enzymer. I tabellen ses gennemsnit og varianser.



(a) Figur 2.1, Alle materialer for hvert laboratorium for forsøg A

(b) Figur 2.2, De flydende enzym materialer for forsøg A

### 2.2.1.1 Statistiske beregninger

De fem anvendte materialer er specielt udvalgt til at dække de enzymmaterialer, der er på markedet i dag. Der er derfor ingen forventning om ens resultater for hvert af materialerne. Resultatet kan desuden afhænge af materialernes type (flydende eller fast) og hvor godt det kan blandes med foderproduktet. Det gælder for alle beregningerne, at alle laboratorieresultater indenfor et materiale er blevet lagt sammen. Det skyldes, at det skal undersøges, hvor godt hvert



6.192362	6.421622	6.428105	6.452049	6.533789	6.551080
6.586172	6.593045	6.626718	6.630683	6.630683	6.633318
6.643790	6.650279	6.651572	6.683361	6.692084	6.706862
6.722630	6.731018	6.736967	6.738152	6.753438	6.758095
6.766192	6.841615	6.887553	6.904751		

Tabel 2.3: Resultater for materiale 1, sorteret i faldende orden

materiale kan bestemmes, og der skal beregnes en variation indenfor materialet som afhænger af laboratorier.

Tabel 2.4 viser de statistiske beregninger for hvert af materialerne, og alle materialerne tilsammen. De statistiske beregninger for standardafvigelsen for repeatabiliteten, standardafvigelsen for mellem laboratorier og standardafvigelsen for reproducerbarheden kommer alle fra ISO 5725.

Eksempelet viser hvordan de enkle værdier er blevet beregnet for materiale 1.

### Gennemsnit

$$\bar{y}_1 = \frac{\sum_{i=1}^{14} 2 \cdot 6.693 + 2 \cdot 6.567 + \dots + 2 \cdot 6.461}{\sum_{i=1}^{14} 2 + 2 + \dots + 2} = 6.648$$

### Median

Resultaterne for materiale 1 er opstillet i stigende orden i tabellen. Medianen er det midterste tal, hvis antallet af resultater er ulige. Hvis antallet af resultater er et lige tal, beregnes gennemsnittet af de to midterste tal.

I dette tilfælde har vi 28 resultater, derfor beregnes medianen som følgende.

$$median = \frac{6.650279 + 6.651572}{2} = 6.650926$$

### Spredning

$$s_1 = \sqrt{\frac{1}{(14 \cdot 2) - 1} \sum_{i=1}^{14} (6.633 - 6.648)^2 + (6.753 - 6.648)^2 + \dots + (6.731 - 6.648)^2} = 0.1624 = 16.24\%$$

### Varians

$$s_1^2 = \frac{1}{(14 \cdot 2) - 1} \sum_{i=1}^{14} (6.633 - 6.648)^2 + (6.753 - 6.648)^2 + \dots + (6.731 - 6.648)^2 = 0.0264 = 2.64\%$$

### Standardafvigelsen for repeatabiliteten

$$s_{r1} = \sqrt{\frac{\sum_{i=1}^{14} ((2-1) \cdot 0.0849) + ((2-1) \cdot 0.1636) + \dots + ((2-1) \cdot 0.3809)}{\sum_{i=1}^{14} (2-1) + (2-1) + \dots + (2-1)}} = 0.1379 = 13.79\%$$

### Standardafvigelsen for mellem laboratorier

$$s_{d1}^2 = s_1^2 = 0.0226$$

$$\bar{n}_1 = \frac{1}{14-1} \cdot \left[ \sum_{i=1}^{14} (2+2+\dots+2) - \frac{\sum_{i=1}^{14} (2^2+2^2+\dots+2^2)}{\sum_{i=1}^{14} (2+2+\dots+2)} \right] = 2$$

$$s_{L1} = \sqrt{\frac{0.0226 - 0.1379^2}{2}} = 0.0606 = 6.06\%$$

### Standardafvigelsen for reproducerbarheden

$$s_{R1} = \sqrt{0.1379^2 + 0.0606^2} = 0.1507 = 15.07\%$$

### 95%-konfidensintervallet

$$\left[ 6.648 - 1.96 \cdot \frac{0.1624}{\sqrt{28}}; 6.648 + 1.96 \cdot \frac{0.1624}{\sqrt{28}} \right] = [6.59; 6.71]$$

Hvor 6.648 er gennemsnittet for materialet, 1.96 er værdien for z-fordelingen ved et 95% konfidensinterval, 0.1624 er spredningen for materialet og 28 er antallet af resultater for materialet.

Box- og Q-Q plottene af hvert materiale viser, at nogle af resultaterne ligger langt fra de andre. De resultater som ligger langt fra de andre, vil i næste kapitel blive undersøgt, om de kan klassificeres som værende outlierer.

### Materiale 1

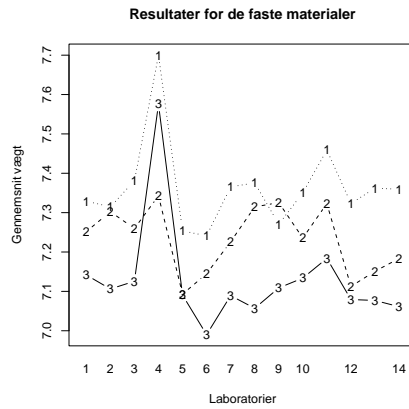
Figur 2.4 viser box plottet for materiale 1, og der ligger et resultat udenfor den nederste grænseværdi. På figur 2.5, ses ligeledes, at det første resultat ligger uden for de andre resultater, der ligger omkring linien.

De beregnede resultater for materialet viser, at det har en høj varians og standardafvigelse for repeterbarheden i forhold til de andre materialer. Mens standardafvigelsen for mellem laboratorierne er lille i forhold til de andre materialer. Standardafvigelsen for reproducerbarheden ligger på samme niveau som de andre materialer.

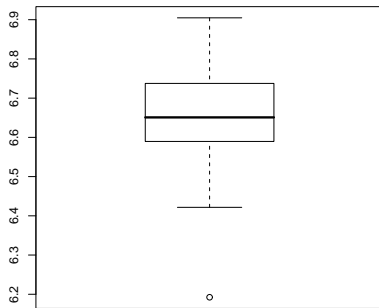
### Materiale 2

Figur 2.6 og figur 2.7 viser, at tre af resultaterne ligger udenfor de andre resultater.

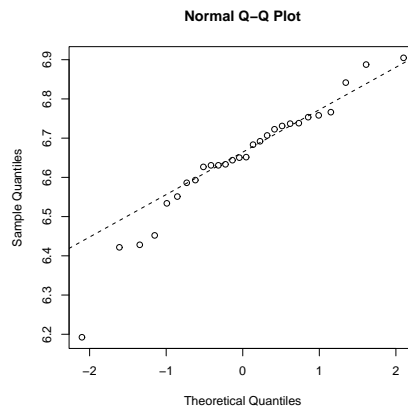
Den beregnede varians ligger høj i forhold til andre materialer, mens standard-



(a) Figur 2.3, De faste enzym materialer for forsøg A



(a) Figur 2.4, Materiale 1 for forsøg A



(b) Figur 2.5, Materiale 1 for forsøg A

	Materiale 1	Materiale 2	Materiale 3	Materiale 4	Materiale 5	Alle materialer
Gennemsnit, $\hat{m}$	6.648	7.068	7.363	7.233	7.130	7.088
Median	6.651	7.04	7.35	7.24	7.10	7.103
Spredning, $s_r\%$	16.24	14.93	12.53	10.59	13.55	27.65
Varians, $s^2\%$	2.64	2.22	1.57	1.12	1.84	7.64
Spredningen for reperaturbarhed, $s_r\%$	13.79	5.65	8.29	9.20	3.16	6.54
Spredningen for mellem laboratorier, $s_L\%$	6.06	14.08	9.57	5.36	13.43	26.96
Spredningen for reproducerbarhed, $s_R\%$	15.07	15.17	12.66	10.64	13.79	27.74
n	28	28	28	28	28	140
Konfidensinterval	(6.59;6.71)	(7.01;7.12)	(7.32;7.41)	(7.19;7.27)	(7.08;7.18)	(7.04;7.13)

Tabel 2.4: Statistisk beregninger for hvert materiale og alle materialerne tilsammen

afvigelsen for repeterbarheden ligger lav. Standardafvigelsen for mellem laboratorier og reproducerbarheden er høje i forhold til de andre materialer.

### Materiale 3

Figur 2.8 og figur 2.9 viser, at der ligger to resultater udenfor den øvregrænseværdi.

Den beregnede varians er lille og det samme gælder for standardafvigelsen for repeterbarheden og mellem laboratorier. Standardafvigelsen for reproducerbarheden er ligeledes lille i forhold til de andre materialer.

### Materiale 4

Figur 2.10 og figur 2.11 viser, at der ikke ligger nogle resultater udenfor grænseværdierne.

De beregnede værdier er de laveste for alle materialerne og det har ingen resultater som ligger udenfor de andre resultater.

### Materiale 5

Figur 2.12 og figur 2.13 viser, at der ligger to resultater langt fra den øvregrænseværdi og at der ligger et resultat udenfor den nedregrænseværdi.

De beregnede værdier for variansen, standardafvigelsen for mellem laboratorier og reproducerbarheden ligger som de andre materialer. Mens standardafvigelsen for repeterbarheden er meget lav i forhold til de andre materialer.

Det kan ud fra box og Q-Q plottene for de fem materialer antages, at det kun er materiale 4, som ikke indeholder nogle outlierer. Det vil blive undersøgt i det næste kapitel, om denne antagelse holder.

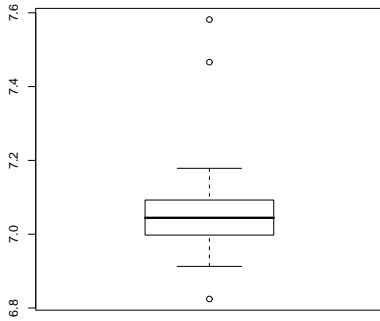
## 2.2.1.2 Varianshomogenitet test

I figur 2.14 ses spredningerne for hvert materiale. Gennemsnittene ligger ikke samme sted, eftersom materialerne ikke har ens gennemsnit. Materialernes spredninger er næsten ens, og det kan derfor ud fra figuren antages, at der forefindes varianshomogenitet mellem de fem materialer.

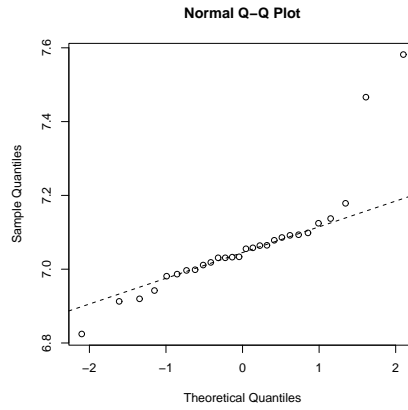
### Bartletts test for materialer

*Bartlett's K-squared = 4.1835, df = 4, p-value = 0.3817*

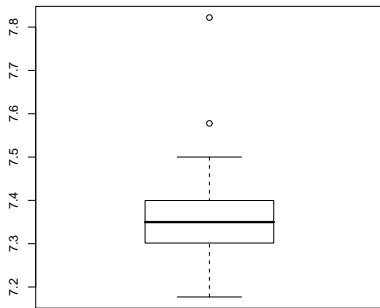
Resultaterne for Bartletts test for materialer viser, at der er varianshomogenitet blandt materialerne. Ved tabelopslag <sup>7</sup> kan værdierne for  $X^2$  – fordeligen aflæses. Den aflæste værdi for materialer med fire frihedsgrader,  $X^2_{(0.05,4)} = 9.49$ .



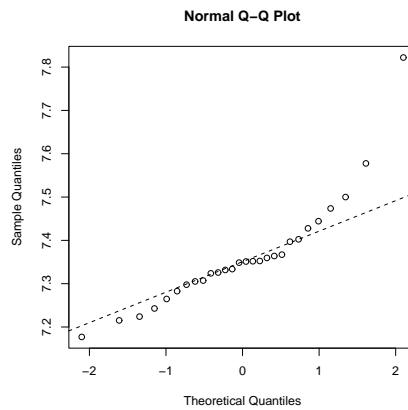
(a) Figur 2.6, Materiale 2 for forsøg A



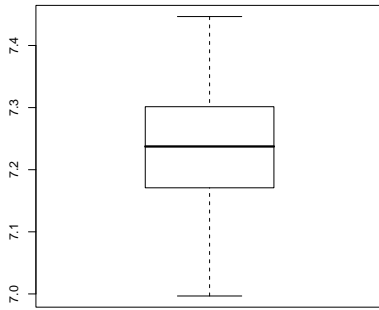
(b) Figur 2.7, Materiale 2 for forsøg A



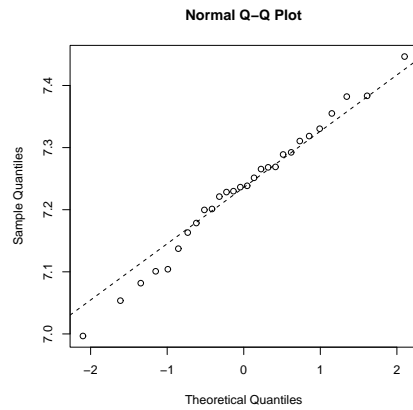
(a) Figur 2.8, Materiale 3 for forsøg A



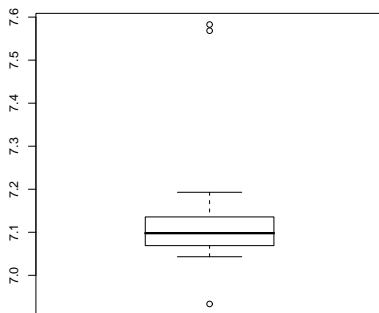
(b) Figur 2.9, Materiale 3 for forsøg A



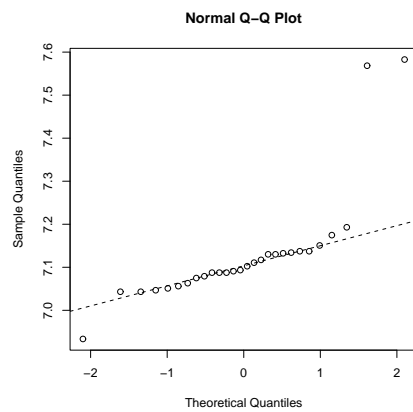
(a) Figur 2.10, Materiale 4 for forsøg A



(b) Figur 2.11, Materiale 4 for forsøg A



(a) Figur 2.12, Materiale 5 for forsøg A



(b) Figur 2.13, Materiale 5 for forsøg A

### 2.2.1.3 Variansanalyse

De materialer der er anvendt i forsøgene er indenfor hvert forsøg forskellige. Det er interessant at undersøge, om der er sammenhæng mellem materialerne og deres typer (fast eller flydende). Til sammenligning anvendes en standard p-test, dvs. en variansanalyse.

I de indledende undersøgelser sættes alle faktorerne til at være fixed faktorer, dvs. materialer, laboratorier og materiale typer.

#### Variansanalyse af modellen

Forsøg A har som udgangspunkt følgende model, som indeholder de to faktorer og deres interaktion.

$$Y_i = \mu + \alpha(\text{materiale}_i) + \beta(\text{laboratorium}_i) + \gamma(\text{materiale}_i, \text{laboratorium}_i) + \epsilon_i$$

Hvor  $\mu$  er gennemsnit,  $\alpha(\text{materiale}_i)$  er påvikningen fra hvert af materialerne,  $\gamma(\text{materiale}_i, \text{laboratorium}_i)$  er påvirkningen fra interaktionen og  $\epsilon$  er påvikningen fra fejlen.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	4	8.19	2.05	265.60	< 2.2e-16 ***
Laboratorim	13	1.29	0.10	12.92	6.924e-14 ***
Materiale:Laboratorium	52	0.60	0.01	1.49	0.0584 .
Residuals	70	0.54	0.01		

Tabel 2.5: Variansanalyse med interaktion for forsøg A

Variansanalysen, tabel 2.5, viser at interaktionen mellem materiale og laboratorium ikke er signifikant og kan dermed fjernes fra modellen. Variansanalysen viser, at både materialer og laboratorier er meget signifikante. Det betyder at indenfor de valgte materialer, vil der være en stor variation. Det samme er gældende for afhængigheden mellem de valgte laboratorier, hvilket viser at der er en stor variationen mellem dem.

Variansanalysen, tabel 2.6, viser resultatet efter fjernelse af interaktionen mellem materialer og laboratorier.



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	4	8.19	2.05	219.37	< 2.2e-16 ***
Laboratorium	13	1.29	0.10	10.67	7.796e-15 ***
Residuals	122	1.14	0.01		

Tabel 2.6: Variansanalyse uden interaktion for forøg A

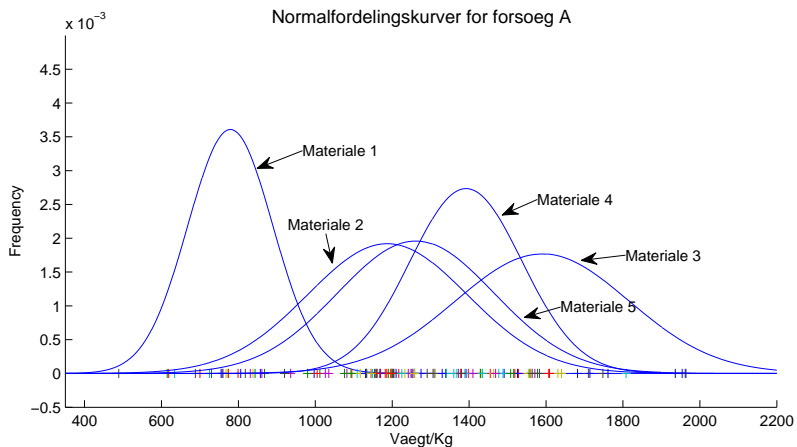
Lab.	Dag	Materiale			
		1	2	j	q
1	1	x x			
	2	x x			
	3	x x			
2	1				
	2				
	3				
i	k			$(y_{ijk1}, y_{ijk2})$	
p	1				
	2				
	3				$y_{ijkl}$

Tabel 2.7: Opbygning af forsøg B

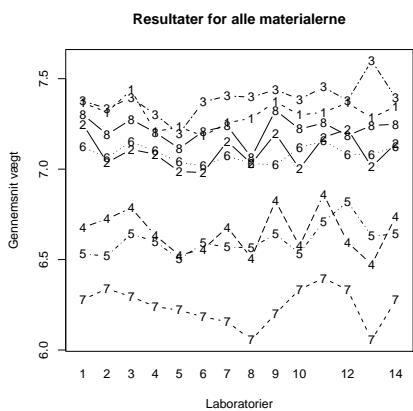
## 2.2.2 Forsøg B

Forsøg B er et krydset forsøg mellem materialer, laboratorier og dagene er nested med laboratorierne. Det betyder at dagene er underlagt det enkle laboratorium. Forsøget er blevet gentaget to gange for hvert af de 3 dage. Dagene i forsøget er uafhængige af hinanden, hvilket betyder, at dag 1 for laboratorium 1 ikke er den samme dag som dag 1 for laboratorium 2.

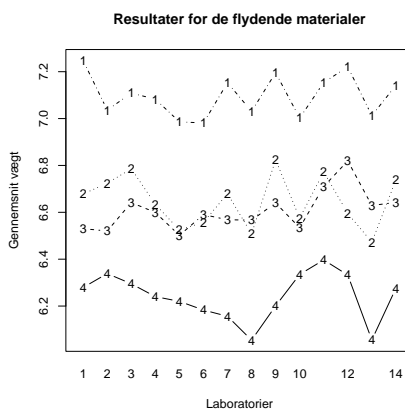
Resultaterne af forsøget ses i figur 2.15, som viser alle materialerne. Figur 2.16 og figur 2.17 viser resultaterne opdelt efter flydende og fast enzym materiale. I tabel 2.8 ses resultaterne af den statistiske undersøgelse for hvert af materialerne, de flydende, de faste og alle materialerne til sammen.



(a) Figur 2.14, Normalfordeling af hvert materiale i forsøg A



(a) Figur 2.15, Alle materialer for hvert laboratorium for forsøg B



(b) Figur 2.16, De flydende enzym materialer for forsøg B

### 2.2.2.1 Statistiske beregninger

De flydende materialer har lige som i forsøg A et lavere gennemsnit end de faste materialer. Materiale 4 og 7 som er flydende har de højeste spredninger, hvor efter det er ligelig fordelt mellem de flydende og faste materialer. De flydende materiale (2, 4, 5 og 7) har igen de største spredninger indenfor repeterbarhed, reproducerbarhed og spredningen mellem laboratorier.

Sammenlignes de beregnede resultater med de grafiske afbilledninger af resultaterne ses, at det for alle materialer gælder, at de ikke indeholder mange outlierer. Resultaterne ligger på en pæn linie i Q-Q plottene, med undtagelse af materiale 5. Materiale 5 indeholder seks resultater som ligger udenfor grænseværdierne i box plottet.

Det kan antages at resultaterne i forsøg B vil kunne indeholde outlierer for materialerne 4, 5, 6, 7 og 8.

### 2.2.2.2 Varianshomogenitet test

I figur 2.34 ses spredningerne for hvert materiale. Normalfordelingskurverne ligger ikke samme sted, eftersom materialerne ikke har ens gennemsnit. Materialernes spredninger er ikke ens, og derfor kan det ud fra figuren antages, at der ikke forefindes varianshomogenitet mellem de otte materialer. De flydende materialer er repræsenterede med tre af de fire høje og smalle fordelingskurver. De faste materials fordelingskurver ligger tæt på hinanden og med samme spredning.

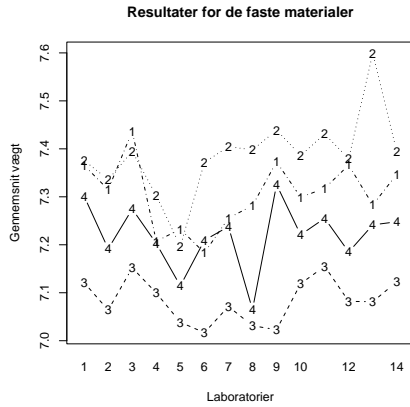
#### Bartlett's test for materialer

*Bartlett's K-squared = 76.0645, df = 7, p-value = 8.716e-14*

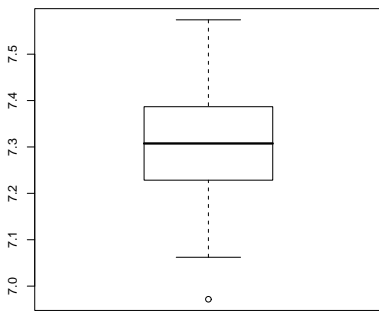
Resultaterne for Bartlett's test for materialer viser, at der ikke er varianshomogenitet blandt materialerne. Ved tabelopslag kan værdierne for  $X^2$ -fordelingen aflæses. Den aflæste værdi for materialer med syv frihedsgrader,  $X^2_{(0.05,7)} = 14.1^5$ .

### 2.2.2.3 Variansanalyse

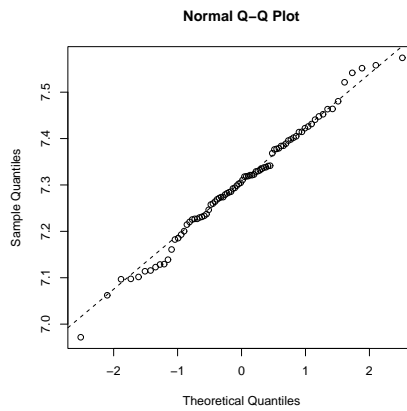
De materialer der er anvendt i forsøgene er indenfor hvert forsøg forskellige. Det er interessant at undersøge om der er nogle sammenhæng mellem materialerne og deres tilstand (fast eller flydende).



(a) Figur 2.17, De faste enzym materialer for forsøg B



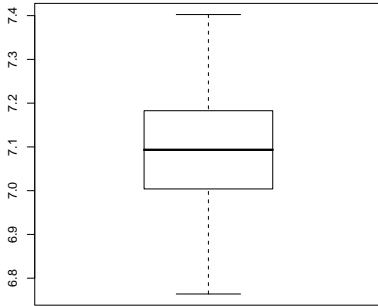
(a) Figur 2.18, Materiale 1 for forsøg B



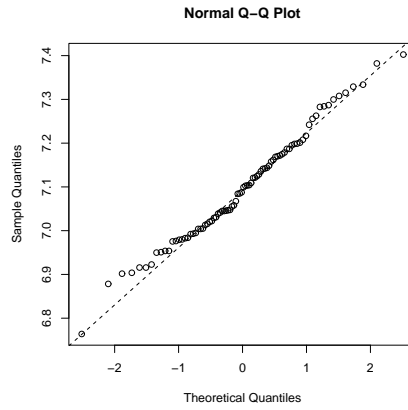
(b) Figur 2.19, Materiale 1 for forsøg B

	Materiale 1	Materiale 2	Materiale 3	Materiale 4	Materiale 5	Materiale 6	Materiale 7	Materiale 8	Alle materialer
Gennemsnit	7.304	7.098	7.388	6.654	6.607	7.084	6.24	7.219	6.949
Varians, $s^2\%$	1.48	1.62	1.7	2.43	1.45	0.34	2.21	1.64	15.85
Spredning, $s\%$	12.18	12.73	13.05	15.57	12.03	5.86	14.88	12.83	39.82
Median	7.308	7.093	7.384	6.648	6.607	7.079	6.236	7.234	7.068
Repeterebarhed, $s_r\%$	10.99	9.92	10.82	11.07	9.58	4.13	11.99	11.94	33.83
Mellem laboratorier, $s_L\%$	5.43	8.22	7.53	11.30	7.51	4.30	9.09	4.82	21.08
Reproducerbarhed, $s_R\%$	12.26	12.89	13.18	15.82	12.17	5.96	15.04	12.88	39.86
n	84	84	84	84	84	84	84	84	672
Konfidensinterval	(7.29; 7.32)	(7.08; 7.10)	(7.38; 7.40)	(6.63; 6.67)	(6.60; 6.62)	(7.08; 7.09)	(6.22; 6.26)	(7.21; 7.23)	(6.90; 6.99)

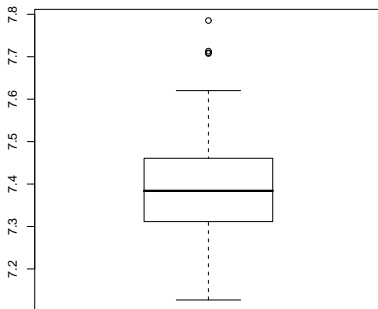
Tabel 2.8: Statistiske beregninger for hvert materiale og alle materialerne tilsammen



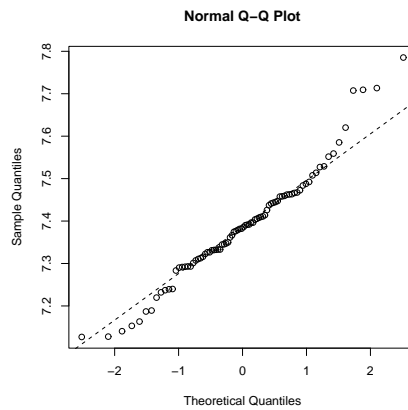
(a) Figur 2.20, Materiale 2 for forsøg B



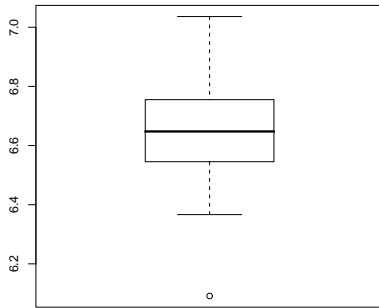
(b) Figur 2.21, Materiale 2 for forsøg B



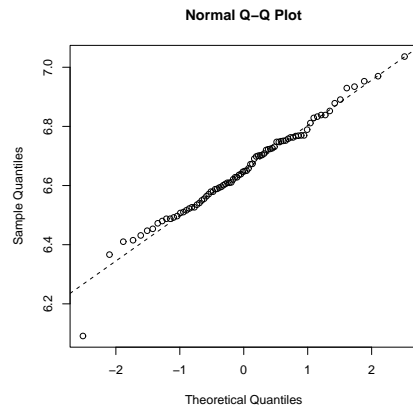
(a) Figur 2.22, Materiale 3 for forsøg B



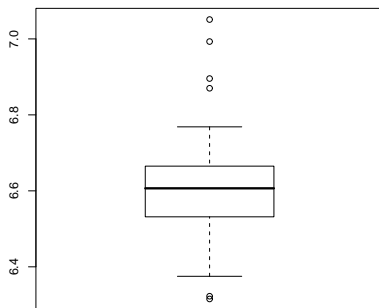
(b) Figur 2.23, Materiale 3 for forsøg B



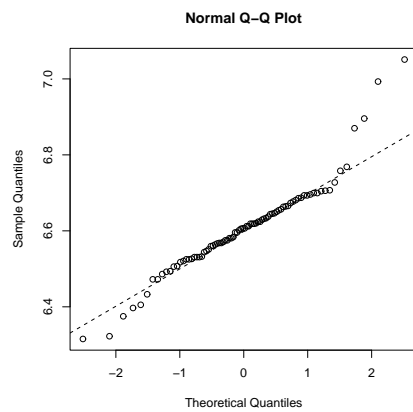
(a) Figur 2.24, Materiale 4 for forsøg B



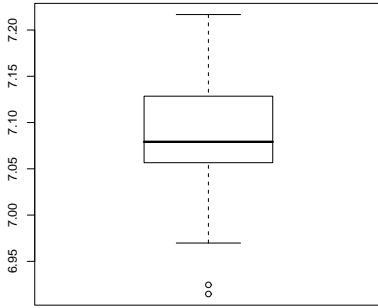
(b) Figur 2.25, Materiale 4 for forsøg B



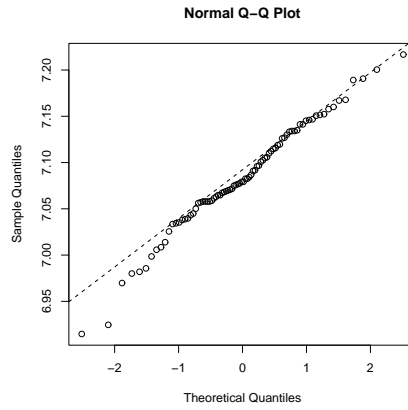
(a) Figur 2.26, Materiale 5 for forsøg B



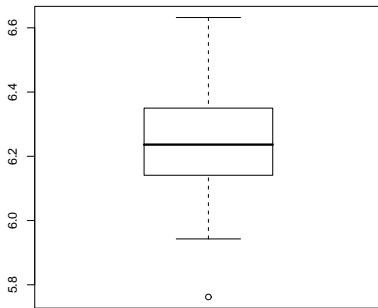
(b) Figur 2.27, Materiale 5 for forsøg B



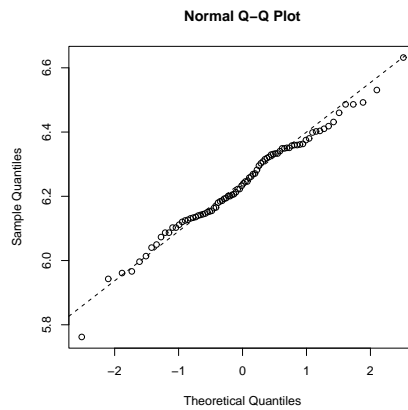
(a) Figur 2.28, Materiale 6 for forsøg B



(b) Figur 2.29, Materiale 6 for forsøg B

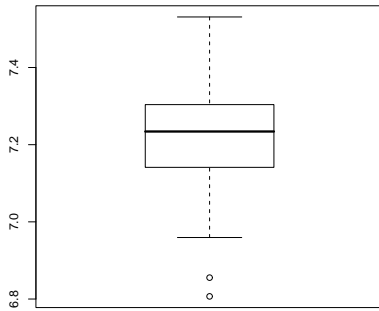


(a) Figur 2.30, Materiale 7 for forsøg B

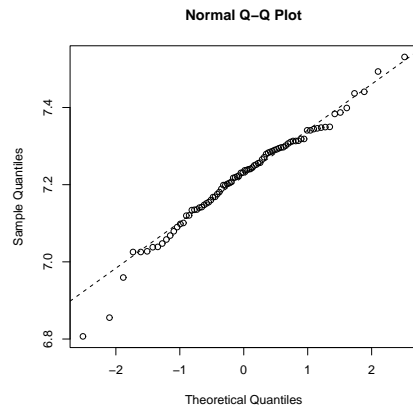


(b) Figur 2.31, Materiale 7 for forsøg B

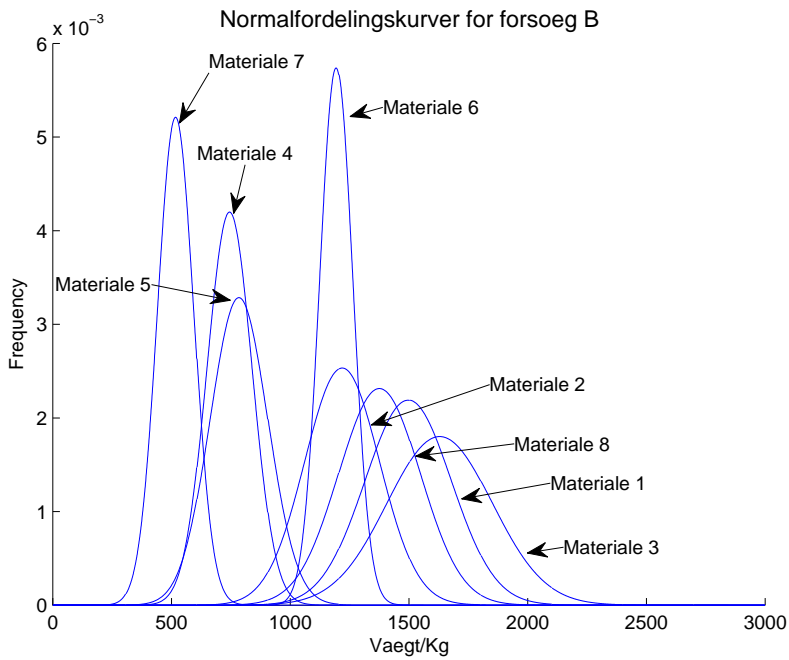




(a) Figur 2.32, Materiale 8 for forsøg B



(b) MFigur 2.33, ateriale 8 for forsøg B



(a) Figur 2.34, Normalfordeling af hvert materiale i forsøg B

I de indledende undersøgelser sættes alle faktorerne til at være fixed faktorer.

### Variansanalyse af modellen

Den fulde model, som beskriver forsøg B indeholder de to nestede faktorer laboratorium og dage og hvor dagene er underlagt laboratorium.

$$y_i = \mu + \alpha(\text{materiale}_i) + \beta(\text{dag}_i(\text{laboratorium}_i)) + \delta(\text{materiale}_i, \text{dag}_i(\text{laboratorium}_i)) + \epsilon_i$$

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	7	95.68	13.67	1906.85	< 2.2e-16 ***
Laboratorium	13	2.04	0.16	21.84	< 2.2e-16 ***
Laboratorium/Dag	28	1.12	0.04	5.56	1.436e-15 ***
Materiale:Laboratorium	91	2.68	0.03	4.11	< 2.2e-16 ***
Materiale:(Laboratorium/Dag)	196	2.45	0.01	1.75	3.870e-06 ***
Residuals	336	2.41	0.01		

Tabel 2.9: Variansanalyse for forsøg B

Variansanalysen, tabel 2.9, viser at alle faktore og interaktioner er signifikante. Det betyder, at modellen ikke kan reduceres. Det betyder at alle faktorerne påvirker hinanden og resultaterne.

### 2.2.3 Sammenligning af materialerne fra forsøgene

Tabel 2.10 viser de materialer som indgår både i forsøg A og B. Materialerne er samlet i grupper, det er derfor nemmere at sammenligne de enkle materialer med hinanden.

## 2.3 Konklusion

De grafiske afbilledninger af resultaterne for de to udførte forsøg, for hvert af materialer, viser at ikke alle resultater ligger omkring normalfordelingens centrum/toppunkt. De resultater vil i næste kapitel undersøges, om de kan antages at være outlier.

	M1A	M5B	M2A	M2B	M3A	M1B	M4A	M8B	M5A	M6B	Alle i A	Alle i B
Gennemsnit, $\hat{n}$	6.648	6.607	7.068	7.098	7.363	7.304	7.233	7.219	7.130	7.084	7.088	6.949
Madian	6.651	6.607	7.04	7.093	7.35	7.308	7.24	7.234	7.10	7.079	7.103	7.068
Spredning, $s\%$	15.02	12.03	14.93	12.73	12.53	12.18	10.59	12.83	13.55	5.86	27.65	39.82
Varians, $s^2\%$	2.26	1.45	2.22	1.57	1.62	1.12	1.48	1.64	1.84	0.34	7.64	15.85
Spredningen for re- terbarhed, $s_r\%$	19.51	9.58	7.99	9.92	11.72	10.99	13.0	11.94	4.47	4.13	26.44	39.82
Spredningen for mellem laboratorier, $s_L\%$	3.58	7.51	3.61	8.22	0.59	5.43	4.57	4.82	0.57	4.30	5.7	0.65
Spredningen for reprodu- cerbarhed, $s_R\%$	19.83	12.17	7.55	12.89	11.74	12.26	12.18	12.88	4.43	5.96	25.82	39.82
Konfidens- interval	(6.59; 6.70)	(6.60; 6.62)	(7.06; 7.08)	(7.08; 7.10)	(7.34; 7.38)	(7.29; 7.32)	(7.21; 7.25)	(7.21; 7.23)	(7.127; 7.132)	(7.08; 7.09)	(7.05; 7.13)	(6.90; 6.99)

Tabel 2.10: Statistiske beregninger for forsøg A og B for de materialer som er ens

Resultaterne for forsøg A, viser at repeterbarheden for hvert materiale varierer meget. Gennemsnittet for de materialer der kan antages at indeholde outlier giver ikke den sande værdi, mens det gælder at medianen bedre kan anvendes som et bud på den sande værdi.

Variansanalysen viser en lille signifikans for interaktion mellem materialer og laboratorier, mens der er en høj signifikant indenfor materialer og laboratorier. Konfidensintervallet for hvert af materialer er smalle og ligger fra 6.6 til 7.38. Konfidensintervallerne for alle materialerne er smalt og det er kun materiale 2, som ligger indenfor det samlede konfidensintervallet. Materiale 1 ligger udenfor i bunden, mens materialerne 3, 4, 5 og de flydende og faste ligger udenfor i den høje ende.

Resultaterne for forsøg B, viser en næsten ens repeterbarhed for hvert af materialerne. Reproducerbarheden for hvert af materialerne er næsten ens, dog med undtagelse fra materiale 6, hvor den er markant lavere.

Variansanalysen viste en stor interaktion mellem alle faktorerne. Det betyder at analyseringen over flere dage har en betydning for resultatet af materialerne og det påvirker ligeledes laboratoriet. Det er kan derfor ikke udelukkes, at den fremtidige kvalitetskontrol skal udføres over flere dage, for at minimere en eventuel dag til dag variation.

Konfidensintervallerne for alle materialerne er smalle og ligger fra 6.2 til 7.4. Intervallet er bredere end i forsøg A. For de materialer som er ens i begge forsøg, ligger konfidensintervallerne ens.

## KAPITEL 3

# Outlier undersøgelse

---

Data i forsøg A og B vil først blive undersøgt for outlier ved brug af Cochran's test og efterfølgende Grubb's test. Cochran's test anvendes til analyse af repeatabiliteten og Grubb's test anvendes til analyse af reproducerbarheden. De forkastede resultater indenfor hvert materiale og laboratorium må ikke overstige  $2/9$  (22%)<sup>1</sup> af henholdsvis materiale og laboratorium.

Der vil i undersøgelserne af data blive anvendt en estimeret værdi for standardafvigelsen,  $s$ , eftersom vi ikke kender den sande værdi af standardafvigelsen,  $\sigma$ . Gennemsnittet vil løbende blive beregnet ud fra resultaterne og det vil afhænge af antallet af laboratorier for hvert materiale.

### 3.1 Formål

Formålet med en outlierundersøgelse er at undersøge om der er resultater, der ligger langt fra de andre resultater i hver materiale. Et resultat kan enten klassificeres som værende outlier eller det skal forsat medtages i analysen. En outlier undersøgelse fjerner de resultater, som ligger udenfor Cochran's og Grubbs testens grænseværdier.

	Sorterede data					Median	Gennemsnit	Varians	95% konfidensinterval
Real data	1	3	5	9	12	5.0	6.0	20.0	[-11.53 ; 23.53]
Data med fejl	1	3	5	9	120	5.0	27.6	2676.8	[-2318.72 ; 2373.92]

Tabel 3.1: Datasæt med fejl

## 3.2 Outlier

En outlier defineres som et datapunkt som ligger langt fra resten af data og kan betragtes som værende en ikke repræsentativ observation i forhold til resten af datapunkterne. I de fleste analyser af data, vil det kunne opleves, at nogle datapunkter ligger langt fra den forventede værdi. Det kan skyldes en systematisk fejl, fejl i analysemetoden, eller det kan være at nogle af observationerne skal ligge langt fra resten af data, som følge af den fordeling de antager. Outlier punkter kan derfor indikere mangelfuld data, fejlagtig procedurer, eller områder hvor den anvendte teori ikke er dækkende. Det er forventet, at der indenfor normalfordelingen, ofte vil være nogle data som kan betragtes som outliere.

Hvornår en outlier skal fjernes fra ens data, er der delte meninger om. Der skal som udgangspunkt ikke fjernes nogle data, eftersom alle data er måleresultater. Hvis et data punkt der betragtes som en outlier, skal fjernes fra data, må der findes en forklaring på datapunktets resultat. Hvis der ikke findes nogle forklaring, skal den resterende analyse foregå både med outlier og uden outlier. Derefter kan det undersøges om outlierne har nogle påvirkning på det samlede resultat. Hvis outlierne har en lille påvirkning, fortsættes der med outlierne. Har outlierne derimod en stor betydning, skal de fjernes fra data, og det skal forklares hvor de ligger i forhold til de blivende data.

Der kan grafisk dannes et overblik over eventuelle outlier i data. Det kan gøres ved brug af boks plot og scatter plot. De data som bliver undersøgt i dette forsøg, er de logaritmiske data, og det antaget at de er normalfordelte. Det betyder at outliere vil ligge i yderkanterne af normalfordelingskurven, mens den største mængde data vil ligge omkring dens center. Outlier kan bestemmes ved anvendelse af fx Grubbs' test, som anvendes i ISO 5725<sup>2</sup>.

Hvad er problemet med outliere ?. Der vil fremkomme et gennemsnit, varians og konfidensinterval med de forkerte værdier.

Tabel 3.1 viser et eksempel, hvor datasættet indeholder en fejl. Der står 120 i stedet for 12. Medianen er ens i begge datasæt, mens gennemsnit, varians og konfidensinterval er meget forskellige.

Ved en outlier undersøgelse skal der altid overvejes følgende:

- Det bør kun være efter grundig overvejelse, at man forkaster en observation, og det vil altid være en handling der kan diskuteres.
- Der findes ingen statistisk korrekt metode til at afgøre om en bestemmelse i et sæt med få data skal beholdes eller forkastes. Den eneste universelt gyldige grund til at forkaste et eksperimentelt resultat er at man med sikkerhed ved at det er behæftet med en grov fejl.

### 3.3 Cochran´s test

Cochran´s test er den første af to test til bestemmelse af outlier. Cochran´s test skal anvendes først og derefter Grubbs´test. Cochran´s test undersøger om variansen indenfor laboratorierne er lille. Variansen indenfor laboratorierne skal helst være lille, hvormed data kan antages at være repeterbar og der kan opnås homogenitet.

Testen kræver, at der forefindes minimum to målinger pr. analyse.

Cochran´s test anvendes til at bestemme ensartetheden af laboratoriernes enkelte bestemmelser på prøve par under repeterbare betingelser.

Cochran beregnes for hver materiale og sammenlignes med dens kritiske værdi. De resultater som overskrider den angivende Cochran værdi fjernes fra datasættet.

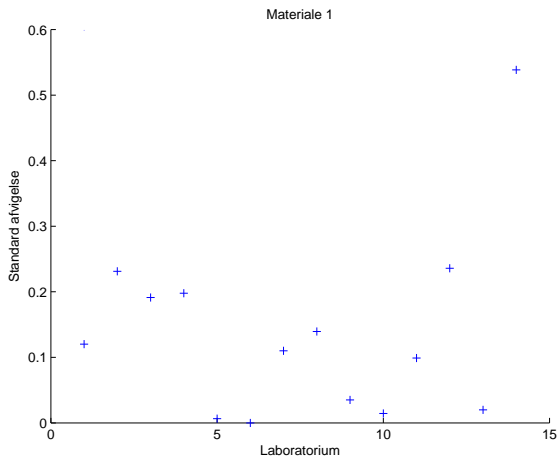
Det betyder for forsøg A, at ved fjernelse af en værdi vil, det pågældende laboratorium ikke længere være repræsenteret ved det pågældende materiale. For forsøg B vil det betyde, at laboratoriet ikke længere er repræsenteret for det pågældende materiale for alle dagene.

Figur 3.1 viser standardafvigelserne for materiale 1 i forsøg A. Figuren viser, at laboratorium 14 har en langt højere standardafvigelse end de andre laboratorier. I Cochran´s test sættes værdien af standardafvigelsen for laboratorium 14 til at repræsentere værdien  $s_{max}$ .

Cochran´s teststørrelse,  $C$ , er følgende <sup>2</sup>

$$C = \frac{s_{max}^2}{\sum_{i=1}^p s_i^2}$$

Der er givet et sæt af  $p$  varianser,  $s_i^2$ , og de er alle sammen beregnet med det samme antal,  $n$ , af gentagede resultater.  $s_{max}^2$  er den højeste varians for et af



Figur 3.1: Standard afvigelser for materiale 1 i forsøg A

materialerne, og  $\sum_{i=1}^p s_i^2$  er summen af alle varianser for ét materiale for alle laboratorier. Der indgår  $p$  laboratorier, som løbende bliver nummeret med  $i$ .

Hvis en celle  $ij$  kun indeholder to test resultater, kan cellens standardafvigelse beregnes, som den absolutte afstand mellem de to resultater <sup>2</sup>.

$$s_{ij} = |y_{ij1} - y_{ij2}|$$

hvor  $y_{ij1}$  er det ene resultat i cellen for materiale,  $j$ , og laboratorium,  $i$ , mens  $y_{ij2}$  er det andet resultat i cellen.

I forsøg A er der anvendt denne metode, hvorimod der for forsøg B er anvendt den generelle metode <sup>2</sup>.

$$s_{ij} = \sqrt{\frac{1}{n_{ij}-1} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij})^2}$$

Hvor  $\bar{y}_{ij}$  er gennemsnittet for alle resultaterne i cellen, for materiale,  $j$ , og laboratorium,  $i$ .  $y_{ijk}$  er det enkelte resultat i cellen og  $n_{ij}$  er antallet af resultater i cellen.

Cochran tester kun for den højeste værdi af standardafvigelsen, dermed er den en



en-siddet outlier test. Hvis en værdi bliver klassificeret som værende en outlier, bliver værdien fjernet fra data og data undersøges igen.

- Hvis test værdien  $C$  er mindre end eller lig med den 5% kritiske værdi, er værdien godkendt.
- Hvis test værdien  $C$  er større end den 5% kritiske værdi og mindre end eller lig med den 1% kritiske værdi, kaldes værdien for en straggler og markes med en stjerne.
- Hvis test værdien  $C$  er større end den 1% kritiske værdi, kaldes værdien for en outlier og markes med to stjerner.

De kritiske værdier for Cochran´s test kan aflæses i ISO 5725 <sup>2</sup>.

Cochran´s test har den ulæmpe, at den kan klassificere en værdi som værende en outlier pga. af der findes mange lave resultater. De lave resultater betyder, at summen af standardafvigelser bliver lille, og dermed en stor Cochran værdi.

### 3.4 Grubbs´ test

Grubbs´ test (Grubbs 1969 and Stefansky 1972). Grubbs´ test er en enkel måde at bestemme om data indeholder nogle outlier. Grubbs´ test hedder også ESD method (extreme studentized deviate) og den tester for outlier ved anvendelse af materialets gennemsnit. Grubbs´ test kan beskrive hvilke data, som ikke er ønskelige at have med i den videre data analyse. Det første trin er at bestemme hvor langt fra data ligger outlieren. Det kan gøres ved beregning af Grubb´s teststørrelse,  $G_p$ , som er differencen mellem gennemsnittet og outlieren, divideret med standardafvigelsen. Gennemsnittet og standardafvigelsen indeholder outlieren i deres beregninger. Hvis  $G_p$  er stor betyder det, at outlieren ligger lagt fra de andre data. Ved analyse af forsøgsdata vil standardafvigelsen være ukendt. I stedet beregnes standardafvigelsen ud fra data. Hvis der er en outlier i data betyder det, at standardafvigelsen stiger, og afstanden mellem gennemsnittet og outlier værdien vil ligeledes stige. Det betyder, at  $G_p$  ikke bliver særlig stor.  $G_p$  kan ikke blive større end  $\frac{p-1}{\sqrt{p}}$ , hvor  $p$  er antal af laboratorier der stadig er repræsenteret i analysen, dvs. eventuelle outliere har betydet at det pågældende laboratorium ikke tæller med længere. Den kritiske værdi for  $G_p$  stiger, når antallet af  $p$  stiger. Der kan ud fra historiske data fastsættes et gennemsnit og en standardafvigelse, og der kan der laves et kontrol kort, som anvendes i kvalitets kontrol.

Grubb's test anvendes som enkelt og dobbelt test, på de laboratorier som ikke indeholder outlier i henhold til Cochran's test. Det gøres for at vurdere ensartetheden i middelværdien på prøve par mellem laboratorier. Dobbelt outlier test udføres på de to mest ekstreme (højeste og/eller laveste) prøve par. Der er udført Grubb's test for enkelt henholdsvis dobbelt outlier tests for laboratorier som er kvalificeret til videre statistisk analyse i henhold til Cochran's test. Grubbs' test er kendt som maksimum normaliserede residual test. Testen kræver at data er normalfordelte, derfor er det vigtigt at data er kontrolleret for normalitet før testen laves. Grubbs' test bestemmer en outlier adgangen og proceduren forsættet til der ikke findes flere outliere.

Grubbs' test kan defineres som et test af hypotese:

H0: Der findes ingen outliere i data sættet

H1: Der er mindst en outlier i data sættet

### En outlier observation

Data ordnes i stigende rækkefølge afhængig af celle gennemsnittet,  $x_p$ , derefter kan gennemsnit og standardafvigelse beregnes. Det højeste og laveste data punkt kan testes for værende en outlier. Test for det højeste data punkt er en outlier <sup>2</sup>:

$$G_p = \frac{x_p - \bar{x}}{s}$$

hvor

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

og

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

Hvor  $\bar{x}$  er gennemsnittet for et materiale for alle laboratorierne og  $s$  er standard afvigelse for et materiale over alle laboratorier.  $x_i$  er celle gennemsnittet for det  $i$ 'th laboratorium op til i alt  $p$  laboratorier.

Test for det laveste data punkt er en outlier <sup>2</sup>:

$$G_1 = \frac{\bar{x} - x_1}{s}$$

### To outlier observationer

De materialer som ingen outlier har, vil blive testet om de to højeste og to laveste resultater kan være outlier. Der beregnes Grubbs´ teststørrelse  $G$  for de to højeste og to laveste resultater.

De to højeste resultater tjekkes for outlier, ved beregning af Grubbs´ test statistisk,  $G$ .  $G$  er forholdet mellem variansen,  $s_{p-1,p}^2$  og  $s_0^2$ .  $s_{p-1,p}^2$  er variansen for alle resultaterne for et materiale og alle laboratorier, undtagen de to højeste celle gennemsnit og deres varians.  $s_0^2$  er variansen for alle resultaterne for et materiale og alle laboratorierne. <sup>2</sup>

$$G = \frac{s_{p-1}^2}{s_0^2}$$

hvor

$$s_0^2 = \sum_{i=1}^p (x_i - \bar{x})^2$$

og

$$s_{p-1,p}^2 = \sum_{i=1}^{p-2} (x_i - \bar{x}_{p-1,p})^2$$

og

$$\bar{x}_{p-1,p} = \frac{1}{p-2} \sum_{i=1}^{p-2} x_i$$

Hvor  $\bar{x}_{p-1,p}$  er gennemsnittet af resultaterne undtagen de to højeste celle gennemsnit, for et materiale og for alle anvendte laboratorier.

De to laveste resultater beregnes på næsten samme måde. Grubbs´ test statistisk,  $G$  er forholdet mellem varianserne,  $s_{1,2}^2$  og  $s_0^2$ .  $s_{1,2}^2$  er variansen for alle resultaterne med undtagelse af de to laveste celle gennemsnit, indenfor et materiale og alle de anvendte laboratorier.  $s_0^2$  er den samme varians som der er anvendt for test af de to højeste resultater. Det er variansen for alle resultaterne for et materiale og alle anvendte laboratorier.

$$G = \frac{s_{1,2}^2}{s_0^2}$$

hvor

$$s_{1,2}^2 = \sum_{i=3}^p (x_i - \bar{x}_{1,2}^2)^2$$

og

$$\bar{x}_{1,2} = \frac{1}{p-2} \sum_{i=3}^p x_i$$

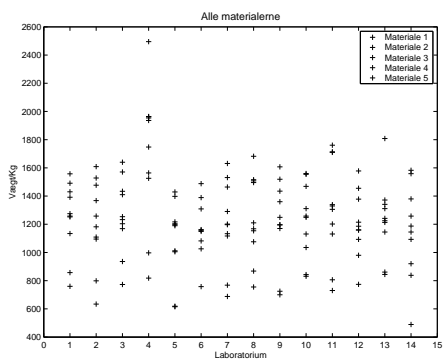
Hvor  $\bar{x}_{1,2}$  er gennemsnittet for alle resultaterne med undtagelse af de to laveste celle værdier, indenfor et materiale og alle de anvendte laboratorier.

Hvis resultatet er større end den 1% kritiske værdi betyder det, at nul hypotesen er afvist og konklusionen er, at data punktet er en outlier, som markeres med to stjerner. Hvis resultatet ligger mellem 1% og 5% for den kritiske værdi betyder det, at nul hypotesen er afvist og konklusionen er, at datapunktet er en strejfer, som markeres med én stjerne. Data kan vises grafisk for at se evt. outliers, det kan gøres med et boksplot eller histogram.

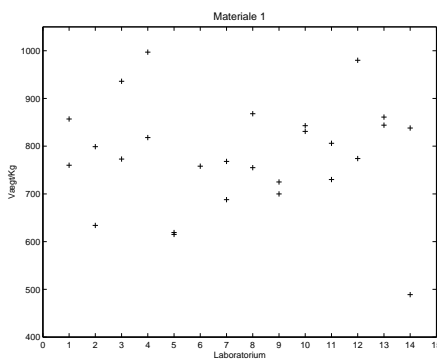
## 3.5 Forsøg A

Resultaterne for forsøg A ses grafisk på figur 3.2-3.7. De viser hvert laboratories resultater i forhold til hinanden. Figur 3.8-3.12 viser hvert materiales normalfordelingskurver. Resultaterne er efterfølgende undersøgt for outliers ved anvendelse af Cochran´s test og Grubbs´test.

### 3.5.1 Grafisk afbildning af forsøg A

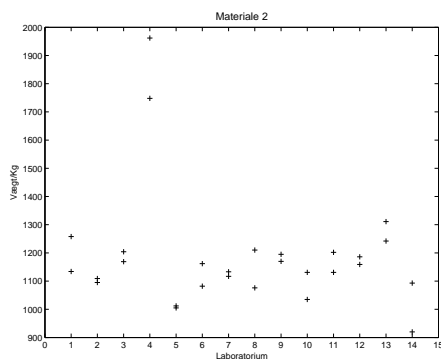


(a) Figur 3.2, Alle materialer for hvert laboratorium for forsøg A

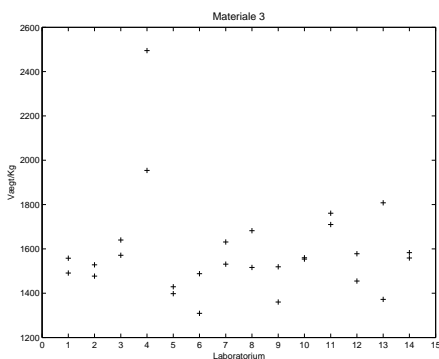


(b) Figur 3.3, Materiale 1 i forsøg A

Figur 3.3 viser resultaterne for materiale 1. Alle resultaterne ligger tæt på gennemsnit. Der er ikke nogle resultater som falder udenfor og dermed indeholder materialet sikkert ikke nogle outlier. Nogle af laboratorierne har en lille repeterbarhed og andre en meget stor repeterbarhed i deres resultater.



(a) Figur 3.4, Materiale 2 i forsøg A



(b) Figur 3.5, Materiale 3 i forsøg A

Figur 3.4 viser resultaterne for materiale 2. Resultaterne for laboratorium 4 ligger langt fra de andres og vil i Grubbs´ test blive klassificerede som værende en outlier. Med fjernelse af laboratorium 4 vil der kunne dannes et gennemsnit, som ligger tættere på de restende resultater.

Figur 3.5 viser resultaterne for materiale 3. Her gælder det samme som for materiale 2, at resultaterne for laboratorium 4 ligger langt fra de andres og vil i Grubbs´ test blive klassificerede som værende en outlier.

Figur 3.6 viser resultaterne for materiale 4. Resultaterne svinger meget fra gennemsnit, og der er ikke nogle resultater, som umiddelbart vil blive klassificerede som værende outlier. Variationen mellem resultaterne indenfor hvert laboratorium svinger for de fleste.

Figur 3.7 viser resultaterne for materiale 5. Her gælder det samme som for materiale 3 og 4. Resultaterne for laboratorium 4 ligger her igen langt fra de andres resultater.

Materiale 1	Materiale 2	Materiale 3	Materiale 4	Materiale 5
0.5445 * (lab. 14)	0.3320 (lab. 14)	0.3957 (lab. 13)	0.2575 (lab. 11)	0.4593 (lab. 6)

Tabel 3.2: Resultater af Cochran´s test for forsøg A

Normalfordelingskurverne for de fem materialer, viser hvor tæt resultaterne ligger på gennemsnittet. De resultater som bliver klassificerede som værende outlier, vil i en normalfordelingskurve ligge i yder kanterne af kurven eller uden/inden for graffen, men ikke på toppen af kurven.

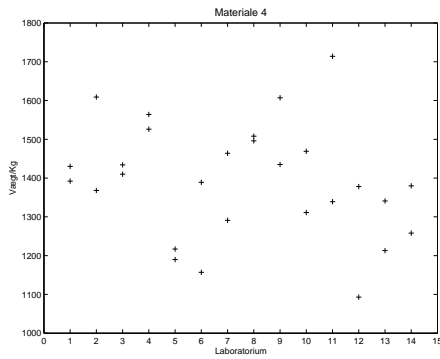
Figur 3.8-3.12 viser normalfordelingskurverne for hvert materiale. De antagelser der er blevet beskrevet for hvert materiale på figurerne 3.3-3.7, kan ses mere tydelig. På figur 3.9, 3.10 og 3.12 ligger outlierne i yder kanten af normalfordelingskurven. På figur 3.11 ligger resultaterne spredt, de er ikke koncentrerede om kurvens centrum. Det kan udfra grafferne antages at materiale 4 har en lille reproducerbarhed.

### 3.5.2 Beregninger for forsøg A

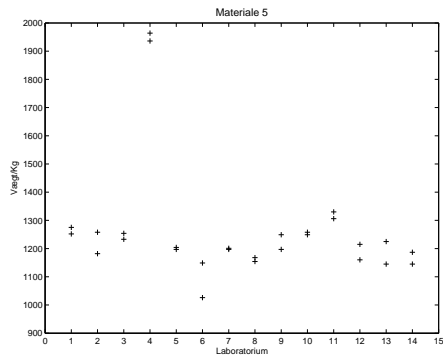
#### Cochran´s test

Test for outlier i forsøg A giver ved brug af Cochran´s test følgende resultater. Tabel 3.2 viser resultatet for hvert materiale, når alle 14 laboratorier tæller med i beregningerne. Tallet der vises for hvert materiale er Cochran´s resultatet efter beregning af Cochran´s teststørrelse,  $C$ . En stjerne betyder, at resultatet ligger over den kritiske 5% værdi, mens to stjerner betyder, at resultatet ligger over den kritiske 1% værdi. Tallet i parrentesen referere til det laboratoriums varians som anvendes, som  $s_{max}^2$  i beregningerne.

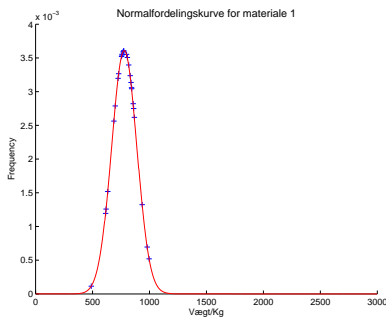
De kritiske værdier ved Cochran´s test er for 1% 0,599 og ved 5% er den 0,492. Der er igen af materialerne som har en værdi højere end de kritiske værdier.



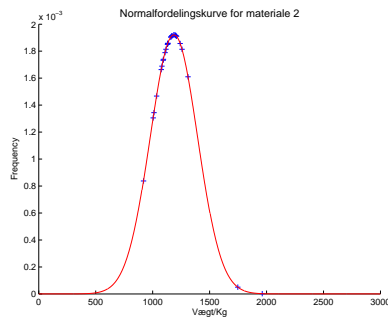
(a) Figur 3.6, Materiale 4 i forsøg A



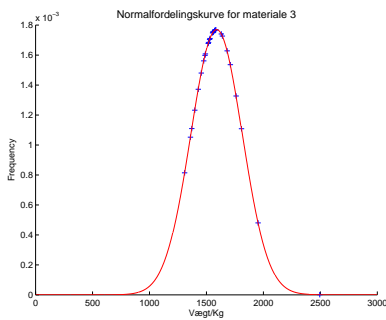
(b) Figur 3.7, Materiale 5 i forsøg A



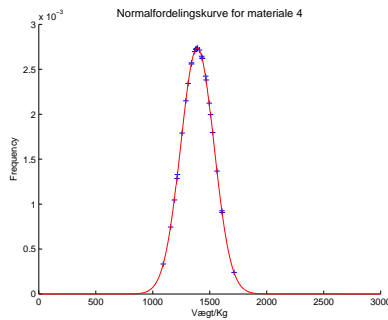
(a) Figur 3.8



(b) Figur 3.9



(a) Figur 3.10



(b) Figur 3.11

Materiale 1 ligger meget tæt på 1% grænsen, som betyder at laboratorium 14 har en dårlig repeterbarhed på det pågældende produkt.

Figur 3.13 viser resultatet af Cochran's test for alle materialerne og hvordan de ligger i forhold til 1% og 5% grænse værdierne.

### Grubbs' test

Resultatet af Grubbs' test for, en og to outlier observationer kan ses i tabel 3.3. Tabel 3.3 viser alle resultaterne for hvert materiale. Resultatet for "Enkel lav" referere til det mindste celle gennemsnit, som bliver testet og  $G_1$  bliver beregnet. Resultatet for "Enkel høj" referere til det højeste celle gennemsnit, som bliver testet og  $G_p$  bliver beregnet. De efterfølgende "Enkel lav" og "Enkel høj", anvendes hvis der har været nogle outlier i et materiale. "Double lave" og "Double høj" anvendes kun for de materialer der ingen outlier er fundet for "Enkel lav" eller "Enkel høj". Strejfer og Outlier værdierne referere i dette tilfælde, til tabelværdierne<sup>2</sup> for 14 og 13 laboratorier.

Grubb's test viser, at materialerne 2, 3 og 5 har én outlier hver, som kommer fra den højeste værdi for hvert af materialerne. Det gælder for alle tre materialer, at resultaterne kommer fra laboratorium 4. Det betyder at laboratorium 4 overstiger de 22%, der er den anbefalede grænse for fjernelse af outliere indenfor et materiale<sup>1</sup>.

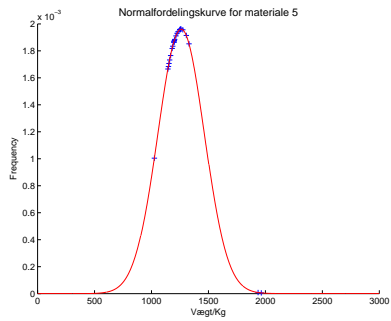
Ved undersøgelse af resultaterne for materiale 1 og 4, gælder det, at det højeste resultat kommer ligeledes fra laboratorium 4. Laboratorium 4 er derfor for alle materialerne repræsenterede med de højeste analyse resultater. I de efterfølgende analyser hvor outliere er fjernet, skal resultaterne fra laboratorium 4 helt fjernes for alle materialerne.

Ved analyse af de laveste resultater, ses det i tabel 3.3, at de ikke indeholder nogle outlier. Det kunne i forlængelse heraf være interessant, at undersøge fra hvilke laboratorier de laveste resultater kommer fra.

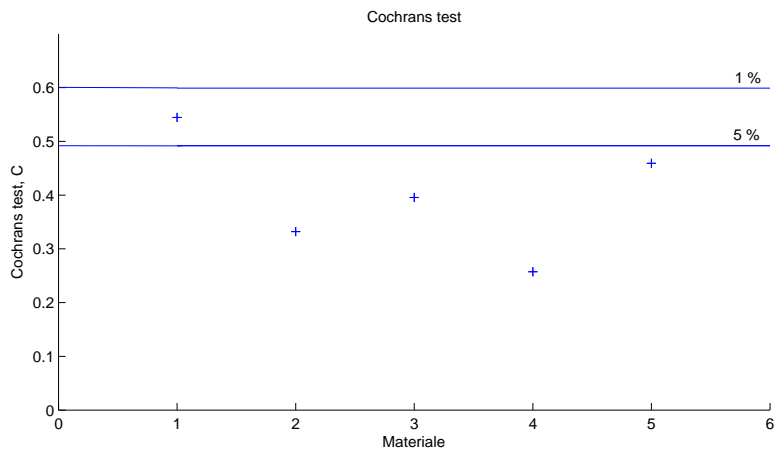
De laveste resultaterne kommer fra følgende laboratorier:

- Materiale 1 fra laboratorium 5
- Materiale 2 fra laboratorium 14
- Materiale 3 fra laboratorium 6
- Materiale 4 fra laboratorium 5
- Materiale 5 fra laboratorium 6





(a) Figur 3.12



(a) Figur 3.13, Cochrans test værdier for materialerne i forsøg A

Materiale	Enkel lav	Enkel høj	Enkel lav	Enkel høj	Doubl lave	Doubl høje
1	2.2483	1.3561			0.5055	0.6777
2	1.0741	3.1166 **	1.8136	1.7649		
3	1.0884	2.9994 **	1.6327	2.0625		
4	1.6657	1.2979			0.5963	0.7225
5	1.0280	3.2736 **	2.2258	1.8569		
Strejfer, 5%	2.507	2.507	2.462	2.462	0.3112	0.3112
Outlier, 1%	2.755	2.755	2.699	2.699	0.2280	0.2280

Tabel 3.3: Resultater af Grubbs' test for forsøg A

Laboratorium 5 er repræsenterede med de højeste Grubb's test værdier for de lave analyseresultater. Det har de næstlaveste analyseresultater for materiale 2 og 3.

Grubb's testen udføres en gang mere for materialerne 2, 3 og 5, eftersom de har en outlier. Outlieren er for alle materialerne fjernet, og de restende resultater er analyseret igen. Resultaterne fra Grubb's testen for anden runde ses i tabel 3.3 med Grubb's resultaterne. Der er ingen af analyseresultaterne, som her kan betragtes som værende outlier, uanset hvilket laboratorium analyserne kommer fra.

Grubb's test værdier er blevet beregnet igen, for de tre materialer som indeholdt en outlier. De høje analyseresultater kommer for:

- Materiale 2 fra laboratorium 13
- Materiale 3 fra laboratorium 11
- Materiale 5 fra laboratorium 11

Grubb's test for to resultater, er kun blevet undersøgt for de materialer som ingen outliers har, ved undersøgelse af ét resultat. Det er materiale 1 og 4. Double Grubb's test viser for begge materialer, at der er outliers for både de lave og høje analyse resultater. De lave analyse resultater kommer for:

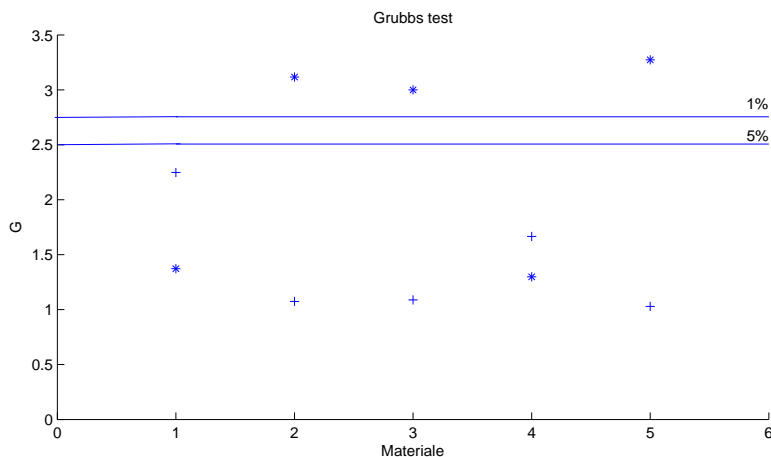
- Materiale 1 fra laboratorium 5 og 14
- Materiale 4 fra laboratorium 5 og 12

De højeste analyse resultater kommer for:

- Materiale 1 fra laboratorium 4 og 12
- Materiale 4 fra laboratorium 4 og 11

Resultaterne af Grubbs'test for hvert af materialerne er vist på figur 3.14. Resultaterne for den mindste værdi er vist med et kryds og den højeste værdi med en stjerne. De kritiske værdier for Grubbs' test er vist på figur 3.14 som to vandrette linier.

Det ses, at der er fire laboratorier som adskiller sig fra de andre laboratorier. Laboratorium 5 og 6 har analyseresultater som ligger lavere end gennemsnittet, og laboratorium 4 og 11 har analyseresultater som ligger højere end gennemsnittet.



(a) Figur 3.14, Grubbs test værdier for materialerne i forsøg A

Reproducerbarheden for laboratorium 4, 5, 6 og 11 er ud fra Grubb's test lille. Reproducerbarheden for laboratorium 4 er meget lille, og laboratories resultater burde være gennemgået for fejl. Det kan tyde på at laboratoriet ikke har fulgt den medfølgende analysevejledning præcis, eller at der har været andre påvirkninger på deres resultater. Ses der på laboratoriumets repeterbarhed, har det nogle gode værdier. Repeterbarheden ligger på højde med de andre laboratorier, derfor kan det antages at laboratoriet ikke har sjusket med analyserne, men derimod må have misforstået nogle informationer. Reproducerbarheden for de andre tre laboratorier er ikke kritiske, eftersom de er ikke blevet klassificerede som havende nogle outliere.

De efterfølgende analyser for forsøg A, vil bestå af to data sæt. Det ene data sæt vil indeholde alle 14 laboratorier og det andet data sæt vil indeholde 13 laboratorier.

Tabel 3.4 viser de samlede resultater af Cochran's test og Grubbs' test for forsøg A.

### 3.5.2.1 Statistiske beregninger

Resultaterne for data uden outlier sammenlignes med de rå data. Resultaterne for data uden outlier ses i tabel 3.5. Det er de samme beregninger, der er blevet lavet som for de rå data i det foregående kapitel.

Det gælder for alle materialerne, at resultaterne for laboratorium 4 er blevet

	Materiale	Laboratorium
Strejfer ved Cochran´s test	1	14
Outlier ved Cochran´s test		
Strejfer ved Grubbs´test		
Outlier ved Grubbs´test	2 3 5	4 4 4

Tabel 3.4: Resultater fra Cochran´s test og Grubbs´test for forsøg A

fjernet. De tre materialer, 2, 3 og 5, som hver indeholdt en outlier har fået deres spredning halveret og deres gennemsnit ligger tæt på deres median. De to andre materialer har ligeledes fået et lavere gennemsnit, som ligger tæt på deres median og en lavere spredning.

Repeterbarheden og reproducerbarheden er for alle materialerne blevet lidt mindre end for de rå data. Fjernelsen af outlier fra de rå data har for alle materialerne betydet nogle mere anvendelige og sigende resultater.

Variansanalysen, tabel 3.6, viser at fjernelsen af laboratorium 4, har betyder at interaktionen mellem materialer og laboratorier ikke længere er signifikant. Det betyder, at laboratorie resultaterne ikke længere påvirkes af materiale valget. Materialerne har samme P-værdi som de rå data, mens P-værdien for laboratorierne er blevet større. Det skyldes, at der er blevet større forskel mellem laboratorierne.

## 3.6 Forsøg B

I forsøg B er der anvendt otte materialer i analysen. Resultaterne for materialerne kan ses i nedenstående figurer. Resultaterne for materialerne viser at der kan antages at der en dag til dag variation.

	Materiale 1	Materiale 2	Materiale 3	Materiale 4	Materiale 5	Alle materialerne
Gennemsnit, $\hat{m}$	6.636	7.033	7.337	7.225	7.096	7.065
Median	6.647	7.033	7.341	7.233	7.092	7.096
Spredning, $s_i$ %	14.65	7.66	7.74	10.53	5.21	25.83
Varians, $s_i^2$ %	2.14	0.59	0.60	1.11	0.27	6.67
Spredningen for reperaturbarhed, $s_r$ %	13.78	5.41	7.14	9.53	3.26	6.43
Spredningen for mellem laboratorier, $s_L$ %	5.07	5.54	3.04	4.56	4.15	25.12
Spredningen for reproducerbarhed, $s_R$ %	14.68	7.74	7.76	10.57	5.28	25.93
n	26	26	26	26	26	130
Konfidensinterval	(6.58;6.69)	(7.00;7.06)	(7.31;7.37)	(7.18;7.27)	(7.08;7.12)	(7.02;7.11)

Tabel 3.5: Statistiske beregninger for hvert materiale og alle materialer tilsammen for forsøg B

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	4	7.43	1.86	249.89	< 2.2e-16 ***
Laboratorium	12	0.32	0.03	3.57	0.0004 ***
Materiale:Lab.	48	0.38	0.01	1.05	0.4173
Residualer	65	0.48	0.01		

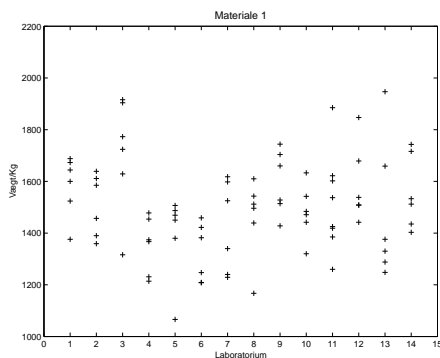
Tabel 3.6: Variansanalyse for forsøg B

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	4	7.43	1.86	244.27	< 2.2e-16 ***
Laboratorium	12	0.32	0.03	3.49	0.0002 ***
Residualer	113	0.86	0.01		

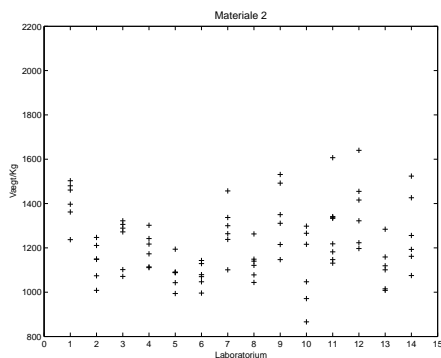
Tabel 3.7: Variansanalyse for forsøg B

### 3.6.1 Grafisk afbildning af forsøg B

Der er i forsøg B analyserede otte materialer, hvor det er de samme materialer som i forsøg A incl. tre andre materialer.



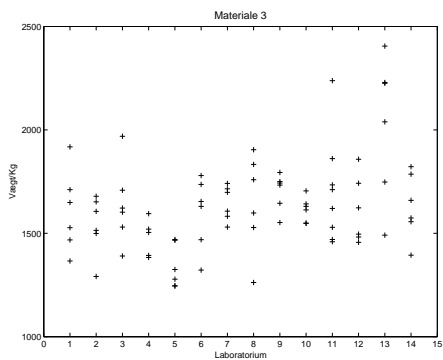
(a) Figur 3.15, Materiale 1 i forsøg B



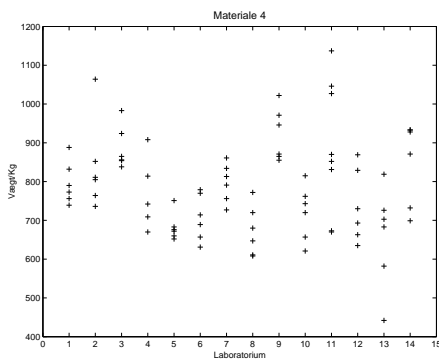
(b) Figur 3.16, Materiale 2 i forsøg B

Alle resultaterne for materialerne ligger tæt omkring gennemsnit. Der er ikke nogle laboratorier som adskiller sig markant fra de andre, som det var gældende i forsøg A med laboratorium 4. Der kan på grafferne ses, at nogle af laboratorierne har en lille repeterbarhed. Det gælder for følgende:

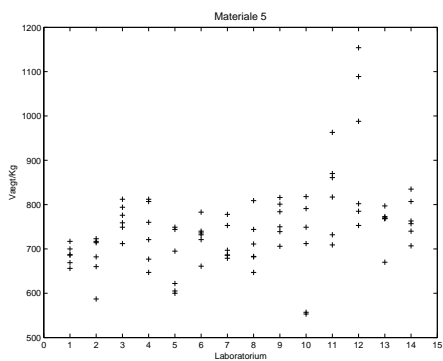
- Materiale 3 for laboratorium 13



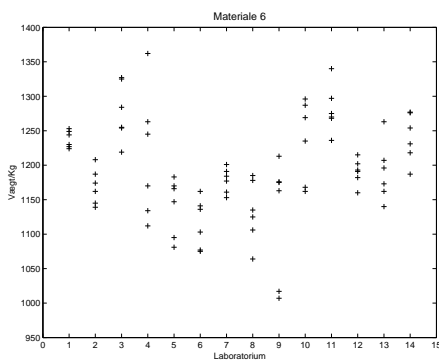
(a) Figur 3.17, Materiale 3 i forsøg B



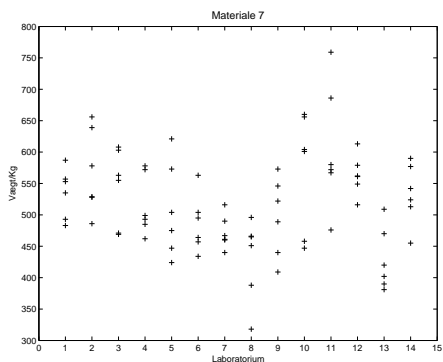
(b) Figur 3.18, Materiale 4 i forsøg B



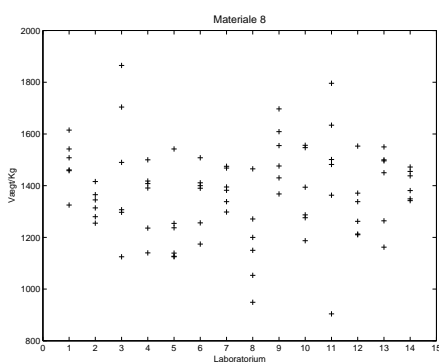
(a) Figur 3.19, Materiale 5 i forsøg B



(b) Figur 3.20, Materiale 6 i forsøg B



(a) Figur 3.21, Materiale 7 i forsøg B



(b) Figur 3.22, Materiale 8 i forsøg B



- Materiale 5 for laboratorium 12
- Materiale 6 for laboratorium 4
- Materiale 8 for laboratorium 11

Resultaterne for repeterbarheden af hvert materiale beregnes ved brug af Cochran's test. Resultaterne kan ses i det næste afsnit.

### 3.6.2 Beregninger for forsøg B

#### Cochran's test

Tabel 3.8 viser resultatet af Cochran's test. I dette forsøg er der for nogle af materialerne blevet lavet flere beregninger. Det skyldes, at de indeholder en outlier. Det er vist efter Cochran's resultatet, for det pågældende materiale med to stjerner. For de materialer, hvor der vises kun en stjerne indeholder materialet en strejfer.

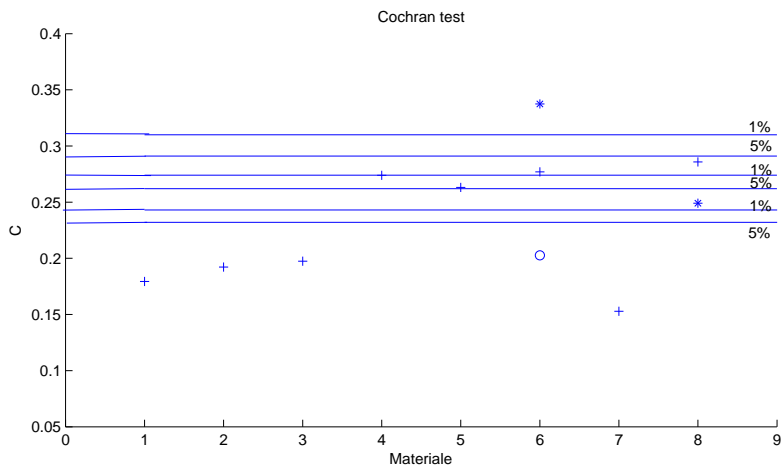
De kritiske værdier ved Cochran's test for de 14 laboratorier er for 1% 0,274 og ved 5% er den 0,232. Materialerne 4, 5, 6 og 8 har en Cochran værdi højere end den kritiske værdi, og for materiale 6 og 8 er den højere end 1% grænsen. Det betyder, at de har et laboratorium med en lille repeterbarhed. Efter fjernelse af de resultater, som giver en Cochran værdi højere end de kritiske værdier, analyseres resultaterne igen. De kritiske værdier ved Cochran's test for 13 laboratorier er for 1% 0,291 og ved 5% er den 0,243. Laboratorier med en Cochran værdi højere end den kritiske værdi:

- Materiale 4 for laboratorium 13
- Materiale 5 for laboratorium 12
- Materiale 6 for laboratorium 9
- Materiale 6 for laboratorium 4
- Materiale 8 for laboratorium 11
- Materiale 8 for laboratorium 3

Figur 3.23 viser resultaterne efter en Cochran's test for alle materialerne, og grænseværdierne for anvendelse af 12, 13 og 14 laboratorie resultater. De to nederste linier gælder for grænseværdierne for 12 laboratorier. De næste to linier for 13 laboratorier og de to øverste linier for 14 laboratorier.

Material 1	Material 2	Material 3	Material 4	Material 5	Material 6	Material 7	Material 8
0.1794	0.1922	0.1974	0.2739 *	0.2631 *	0.2770 **	0.1528	0.2858 **
					0.3374 **		0.2491 *
					0.2027		

Tabel 3.8: Resultater af Cochran´s test for forsøg B



(a) Figur 3.23, Cochran's test for forsøg B

### Grubbs' test

Resultaterne for forsøg B er ligeledes blevet undersøgt for eventuelle outlier. I tabel 3.9 ses resultaterne for Grubbs test.

Der er ingen outlier for forsøg B ved anvendelse af Grubbs' test, da de er blevet fjernet ved Cochran's test. Materiale 5 har en værdi højere end den kritiske 5% for strejfer, men den bliver ikke fjernet fra datasættet. Der er ikke nogle laboratorier, som skiller sig markant ud med hensyn til hvem der har de højeste eller laveste test resultater.

De efterfølgende analyser for forsøg B, vil bestå af to datasæt, som i forsøg A. Det ene datasæt er for de rå data og det anden datasæt indeholder ingen outlier. Det er i forsøg B kun de to faste materialer 6 og 8 der har fået fjernet nogle resultater, de andre materialer har ikke fået klassificerede nogle outlier.

Tabel 3.10 viser de samlede resultater af Cochran's test og Grubbs' test for forsøg B.

	Enkel lav	Enkel høj	Doubel lav	Doubel høj
Materiale 1	1.7197	1.8808		
Materiale 2	1.2755	1.6227		
Materiale 3	2.1904	2.4233		
Materiale 4	1.4912	1.6960		
Materiale 5	1.2410	2.5136 *		
Materiale 6	1.5373	1.4516		
Materiale 7	1.7988	1.5169		
Materiale 8	2.1504	1.5411		
Strejfer, 5% ved 14 lab.	2.507	2.507	0.3112	0.3112
Outlier, 1% ved 14 lab.	2.755	2.755	0.2280	0.2280
Strejfer, 5% ved 13 lab.	2.462	2.462		
Outlier, 1% ved 13 lab.	2.699	2.699		
Strejfer, 5% ved 12 lab.	2.412	2.412		
Outlier, 1% ved 12 lab.	2.636	2.636		

Tabel 3.9: Resultater af Grubbs´ test for forsøg B

	Materiale	Laboratorium
Straggler ved Cochran´s test	4	13
	5	12
	8	3
Outlier ved Cochran´s test	6	4
	6	9
	8	11
Straggler ved Grubbs´test	5	
Outlier ved Grubbs´test	–	–

Tabel 3.10: Resultater for Cochran´s test og Grubbs´test for forsøg B

### 3.6.2.1 Statistiske beregninger

Resultaterne for data uden outlier sammenlignes med de rå data. Resultaterne for data uden outlier ses i tabel 3.11. Det er de samme beregninger, der er blevet lavet som for de rå data.

Variansanalysen, tabel 3.12 viser, at modellen ikke kan reduceres, eftersom alle interaktioner stadig er signifikante. P-værdien for interaktionerne sammenlignes med P-værdierne for variansanalysen for de rå data. Interaktionen mellem materialer og laboratorier underlagt dagene er blevet mindre. Det samme er gældende for laboratorier underlagt dagene.

## 3.7 Sammenligning af forsøg A og B

I forsøg A blev der anvendt 5 materialer nummererede med 1 - 5, og i forsøg B blev der anvendt 8 materialer som ligeledes er nummererede med 1 - 8. For at kunne sammenligne resultaterne for de to typer forsøg, kan det i tabel 3.13 ses hvilke materialer der er ens.

Ved sammenligning af resultaterne fra forsøg A og B er det interessant, at se på de outliere som er fundet via Grubb's test. Ved sammenligning af materiale B1 og A3 kan det ses, at den fundne outlier i forsøg A for laboratorium 4, ligeledes afviger meget fra forsøg B resultaterne. Laboratoriet har i forsøg B en klar dag til dag variation i deres resultater. Det samme er gældende for materiale B2 og A2 og B6 og A5. Ved sammenligning af materiale B5 og A1 kan det ses, at den fundne outlier i forsøg B for laboratorium 12, skyldes en stor spredning i resultaterne. Ved sammenligning af materiale B8 og A4, følger de fint hinanden.

## 3.8 Konklusion

Der er blevet udført to slags forsøg. Til undersøgelse af reproducerbarheden er der anvendt Grubb's test, som ligeledes bestemmer eventuelle outliere, og til

	Materiale 1	Materiale 2	Materiale 3	Materiale 4	Materiale 5	Materiale 6	Materiale 7	Materiale 8	Alle materialer
Gennemsnit	7.304	7.098	7.388	6.654	6.607	7.088	6.240	7.217	6.944
Median	7.308	7.093	7.384	6.648	6.607	7.083	6.236	7.231	7.067
Varians, $s^2\%$	1.48	1.62	1.70	2.43	1.45	0.278	2.21	1.39	16.12
Spredning, $s\%$	12.18	12.73	13.05	15.57	12.03	5.29	14.88	11.80	40.15
Repetitionsbarhed, $s_r\%$	10.99	9.92	10.82	11.07	9.58	3.14	11.99	10.56	9.38
Mellem laboratorier, $s_L\%$	5.43	8.22	7.53	11.30	7.51	4.41	9.09	5.45	39.19
Reproducerbarhed, $s_R\%$	12.26	12.89	13.18	15.82	12.17	5.42	15.04	11.88	40.30
n	84	84	84	84	84	72	84	78	654
Konfidensinterval	(7.29; 7.32)	(7.08; 7.10)	(7.38; 7.40)	(6.63; 6.67)	(6.60; 6.62)	(7.08; 7.10)	(6.22; 6.26)	(7.19; 7.24)	(6.91; 6.97)

Tabel 3.11: Statistiske beregninger for forsøg B

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	7	94.97	13.57	1933.88	< 2.2e-16 ***
Laboratorium	13	2.13	0.16	23.35	< 2.2e-16 ***
Lab.:Dag	28	1.10	0.04	5.62	1.150e-15 ***
Materiale:Lab.	88	2.55	0.03	4.13	< 2.2e-16 ***
Materiale:Lab.:Dag	190	2.23	0.01	1.68	2.263e-05 ***
Residualer	327	2.29	0.01		

Tabel 3.12: Variansanalyse for forsøg B

undersøgelse af repeterbarheden er der anvendt Cochran´s test. Alle resultaterne er blevet afbilledet for hvert materiale i de to forsøg, hvormed det kan vises hvordan resultaterne ligger i forhold til hinanden.

I forsøg A viste det sig, at laboratorium 4 har nogle meget højere analysereultater i forhold til de andre laboratorier. Konklusionen af outlier testen for forsøget blev, at den videre analyse af resultaterne skal foregå på to datasæt, et med de rå data og et datasæt uden outlierer. Den videre undersøgelse skal foregå i det ene datasæt med alle 14 laboratorier, og i det andet datasæt med 13 laboratorier, som er uden laboratorium 4. I forsøg A er der for laboratorium 4 sket en konsekvent fejl, som har betydet de meget højere analyse resultater. Resultaterne indenfor hvert materiale har en god repeterbarhed, som betyder at det har været den samme laborant der har udført analyserne og der er ikke blevet sjusket med resultaterne.

I forsøg B har laboratorierne 4, 9 og 11 en lille reproducerbarhed hvilket betyder, at efterfølgende analyser for forsøg B skal ske med to datasæt. Et datasæt med alle 14 laboratorier og et datasæt, hvor analyse resultaterne for materiale 6 for laboratorierne 4 og 9 fjernes og for materiale 8 for laboratorierne 11 fjernes. Cochran´s test viste i forsøg B, at det kan antages, at laboratorierne har en dag til dag variation. Dags variationen skal derfor inddrages i den videre analyse til bestemmelse af det optimale gennemsnit og spredning.

Ved sammenligning af forsøg A og B følger deres analyse resultater pænt hinanden. Det kan hermed konkluderes at analyse metoden som laboratorierne har anvendt har en god reproducerbarhed.

Outlier undersøgelsen for forsøg A og B har betydet, at hvert forsøg har fået tilføjet et reduceret datasæt, som ikke indeholder outlierer. De statistiske undersøgelser af data uden outlier, viser for begge forsøg, at de berørte materialer får et lidt lavere gennemsnit. Gennemsnittet ligger lidt tættere på deres median. Spredningen er for nogle af materialerne, blevet kraftig reducerede som følge af

	M1A	M5B	M2A	M2B	M3A	M1B	M4A	M8B	M5A	M6B	Alle i A	Alle i B
Gennemsnit, $\hat{\mu}$	6.648	6.607	7.068	7.098	7.363	7.304	7.233	7.219	7.130	7.084	7.088	6.949
Median	6.651	6.607	7.04	7.093	7.35	7.308	7.24	7.234	7.10	7.079	7.103	7.068
Spredning, $s_r\%$	15.02	12.03	14.93	12.73	12.53	12.18	10.59	12.83	13.55	5.86	27.65	39.82
Varians, $s^2\%$	2.26	1.45	2.22	1.57	1.62	1.12	1.48	1.64	1.84	0.34	7.64	15.85
Spredningen for reprobearbarhed, $s_r\%$	19.51	9.58	7.99	9.92	11.72	10.99	13.0	11.94	4.47	4.13	26.44	39.82
Spredningen for mellem laboratorier, $s_L\%$	3.58	7.51	3.61	8.22	0.59	5.43	4.57	4.82	0.57	4.30	5.7	0.65
Spredningen for reproducibarhed, $s_R\%$	19.83	12.17	7.55	12.89	11.74	12.26	12.18	12.88	4.43	5.96	25.82	39.82
Konfidensinterval	(6.59; 6.70)	(6.60; 6.62)	(7.06; 7.08)	(7.08; 7.10)	(7.34; 7.38)	(7.29; 7.32)	(7.21; 7.25)	(7.21; 7.23)	(7.127; 7.132)	(7.08; 7.09)	(7.05; 7.13)	(6.90; 6.99)

Tabel 3.13: Sammenligning af forsøg A og B for ens materialer



fjernelse af outlier. Repeterbarheden og reproducerbarheden er ligeledes blevet mindre, end for de rå data.



## KAPITEL 4

# Robusthedsanalyse

---

Data er nu blevet undersøgt for outlier vha. Cochran's og Grubbs's test. Resultatet er et datasæt uden outlier, hvorpå der er foretaget statistiske beregninger.

I en robusthedsanalyse fjernes data, der ligger langt fra andre data, ikke fra datasættet. Derimod konverteres og fastholdes de i datasættet.

## 4.1 Formål

Formålet med robusthedsanalysen i dette projekt, er at få "rettet" data i yderkanterne, hvormed de kommer til at ligge pænere indenfor normalfordelingskurven. Der beregnes nye varianser på de ændrede data. De forventes, at være mere realistiske for de pågældende materialer, end for de rå data.

## 4.2 Robusthedsanalyse

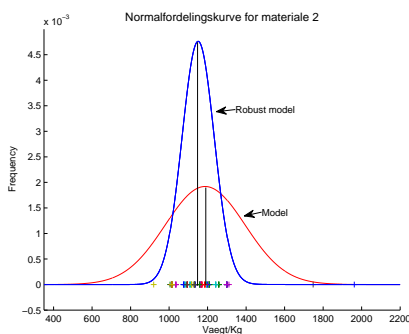
Heavy-tailed (data med tunge haler) fordelingen er betegnelsen for, data som følger en normalfordeling, hvor lidt for mange i forhold til, hvad der forventes af

data, ligger i ydrekanterne. Data kan betragtes som værende outlier, eller det kan være de tilhører en anden fordeling. En blanding af normalfordelingen og data som potentielt tilhører en anden fordeling, analyseres via en robusthedsanalyse <sup>4</sup>.

Robusthedsanalysen anvendes for ikke, at skulle udelade informationer fra forsøgene. Ved fjernelse af outlier fra data, fjernes måske vigtig information. Outlier i data kan fremkomme af flere årsager, og de er dannet på samme grundlag som de andre data. Medtagelsen af outlier data betyder, at det ikke kun er de gode resultater som kommer til at påvirke den videre analyse og resultaterne. Ideen med robusthedsanalysen er, at give nogle bedre resultater på baggrund af alle data, men med en justering, af de data som ligger udenfor den øvre og nedre grænse.

Robusthedsanalysen er en moderne metode, at justere resultater når det formodes, at de indeholder en lille mængde outliere. De fleste estimater for det beregnede gennemsnit og beregnede standardafvigelse afhænger af, at data består af tilfældige observationer for en normalfordeling. Det er kendt, at analytiske data ofte afviger fra denne model. De er ofte heavy tailed (indeholder en koncentration af data andet sted end ved det beregnede gennemsnit), som indeholder outliere <sup>6</sup>.

Figur 4.1 viser en normalfordelingskurve for rå data og data efter robusthedsanalysen. Normalfordelingen for rå data har et gennemsnit med en stor standardafvigelse. Normalfordelingskurven for data efter robusthedsanalysen er blevet smallere, med et ændret gennemsnit og en mindre standardafvigelse.



Figur 4.1: Normalfordeling for model og robust model

Der er blevet udviklet mange metoder indenfor robusthedsanalyse. I denne opgave anvendes den metode der er beskrevet i ISO 5725-5. Robusthedsanalysen anvendes i standarden kun på de datasæt som indeholder outlier fundet ved

Cochran's test eller Grubbs' test.

Standardafvigelsen påvirkes let af data som ligger langt fra de andre data. Der kan vælges at beregne standardafvigelserne uden outlier, men det kan give nogle standardafvigelser der bliver for små. Derfor kan det være en fordel, at anvende en robusthedsanalyse.

Robusthedsanalysen, der anvendes i 5725-5 ændrer betydningen af de resultater, som ligger udenfor den nedre og øvre grænseværdi. Det gøres ved brug af to algoritmer. Algoritme A, der anvendes til at beskrive variationen mellem laboratorierne og reproducerbarheden. Algoritme S, der anvendes til at beskrive variationen indenfor laboratorierne og repeterbarheden.

I denne sammenhæng anvendes kun algoritme A det skyldes, at det kun er ændringerne af resultaterne der er interessant i dette tilfælde. Der beregnes ikke på celle gennemsnittene, men direkte på celleværdierne. Det skyldes, at data efterfølgende skal analyseres ved brug af mixed model modeller, hvor det er vigtig at have hver celleværdi og ikke kun et gennemsnit for celleværdien.

Der skal i algoritme A bestemmes en startværdi for gennemsnittet og standardafvigelsen. Her anvendes medianen som repræsentant for gennemsnit og standardafvigelsen er ligeledes beregnet på baggrund af medianen. Medianen og standardafvigelsen itereres indtil den nye værdi og den foregående værdi er ens.

I et eksperimentelt forsøg har vi ingen konkret viden, om data kommer fra en heavy-tailed fordeling eller om den vil indeholde outlier. Det ideale vil derfor være, at anvende et estimat til bestemmelse af gennemsnittet. Derfor anvendes medianen til bestemmelsen af gennemsnittet.

Medianen er defineret som den midterste værdi i et ordnet talsæt med et ulige antal talstørrelser. Ved et lige antal talstørrelser tages gennemsnittet af de to midterste størrelser. Medianen er betegnelsen på 50%-fraktilen.

Lad

$$x_1, x_2, \dots, x_n$$

være komponenterne af en stikprøveværdi opstillet efter størrelse, således at  $x_i \leq x_{i+1}$ ,  $i = 1, 2, \dots, n - 1$ . Medianen defineres da ved <sup>5</sup>:

$$\tilde{x} = \begin{cases} x_{(\frac{n+1}{2})} & \text{for } n \text{ ulige} \\ \frac{1}{2}(x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}) & \text{for } n \text{ lige} \end{cases}$$

Medianen er forholdsmæssig upåvirket overfor tilstedeværelsen af ekstreme (af-

vigende) værdier.

### 4.3 Algoritme A

Algoritme A anvendes til at bestemme robuste værdier for gennemsnit og standardafvigelsen for hvert forsøgsresultat.

Alle forsøgsresultaterne indenfor et materiale opstilles i stigende orden.  $p$  er antallet af laboratorier indenfor materialet.

$$x_1, x_2, \dots, x_i, \dots, x_p$$

Det robuste gennemsnit og standardafvigelse beregnes ud fra hvert analyse resultat.  $x^*$  er det robuste gennemsnit,  $s^*$  er den robuste standardafvigelse og  $i$  repræsenterer det  $i$ 'ne laboratorium udaf de  $p$  laboratorier.

Startværdien for det robuste gennemsnit er medianen af analyse resultaterne.

$$x^* = \text{median af } x_i$$

For at kunne beregne den robuste standardafvigelse. Beregnes først differancen mellem analyse resultatet og medianen (det robuste gennemsnit). Differencerne ordnes i stigende rækkefølge og medianen findes af disse værdier. Værdien kaldes median absolute deviation (MAD). MAD ganges med en scale faktor på 1.483, hvorved man opnår det bedste resultat for normalfordelingsdata <sup>3</sup>.

$$MAD = \text{median}[|x_i - x^*|]$$

$$s^* = 1.483 * MAD$$

Det beregnede robuste gennemsnit og standardafvigelse anvendes til iteration af det optimale robuste gennemsnit og standardafvigelse. Huber's metode anvendes til iterationen af den øvre og nedre grænseværdi for analyseresultaterne. Hvis et analyseresultat ligger udenfor den øvre eller nedre grænseværdi, bliver det sat til grænseværdien. De analyseresultater som ligger uden for grænserne bliver derfor ikke fjernet fra data, som det er gældende for outlier testen. De bliver rettet til og medtages i de efterfølgende beregninger <sup>3</sup>.

Der beregnes en kalkulator:  $\varphi = 1.5 * s^*$ , som anvendes til dannelse af grænseværdierne. Faktoren på 1.5 er den mest anvendte for en winsorisation proces <sup>3</sup>. Den øvre grænseværdi er det robuste gennemsnit plus kalkulatoren og den nedre grænseværdi er det robuste gennemsnit fratrukket kalkulatoren.

øvre grænseværdi =  $x^* + \varphi$   
 nedre grænseværdi =  $x^* - \varphi$

Analyse resultaterne,  $x_i^*$  ( $i=1, 2, \dots, p$ ), hvor  $p$  er antallet af laboratorier, undersøges om de ligger uden for grænseværdierne.

$$x_i^* = \begin{cases} x^* - \varphi & \text{if } x_i < x^* - \varphi \\ x^* + \varphi & \text{if } x_i > x^* + \varphi \\ x_i \text{ellers} & \end{cases}$$

Der beregnes en opdateret værdi for det robuste gennemsnit.

$$x^* = \sum_{i=1}^p x_i^* / p$$

Der beregnes ligeledes en opdateret værdi for den robuste standardafvigelse. Den opdaterede værdi for den robuste standardafvigelse tager udgangspunkt i standardafvigelsen for en normalfordeling, hvor der ganges med en konstant <sup>3</sup>.

$$[\hat{\sigma}^j]^2 = \frac{1}{n\beta} \sum (\tilde{x}_i - \mu)^2$$

Hvor der anvendes en  $\beta$  værdi på 0.778 og  $n$  er 1 <sup>3</sup>.

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

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

$$s = \sqrt{\frac{1}{0.778}} \sqrt{\frac{\sum_{i=1}^p (x_i - \bar{x})^2}{(p-1)}} \iff$$

$$s^* = 1.134 \sqrt{\sum_{i=1}^p (x_i^* - x^*)^2 / (p-1)}$$

## 4.4 Algoritme S

Algoritme S beregner en robust standardafvigelse indenfor variationen af laboratorier. Den giver en robust pooled værdi af celle standardafvigelsen eller celleafstanden, hvis der kun er to gentagelser.

Celleafstanden eller standardafvigelsen for et materiale opstilles i stigende orden for alle laboratorierne,  $p$ .

Antal af frihedsgrader $v$	Grænsefaktor $\eta$	Tilføjelse faktor $\xi$
1	1.645	1.097
2	1.517	1.054
3	1.444	1.039
4	1.395	1.032
5	1.359	1.027
6	1.332	1.024
7	1.310	1.021
8	1.292	1.019
9	1.277	1.018
10	1.264	1.017

Tabel 4.1: Grænsefaktorer

$$w_1, w_2, \dots, w_i, \dots, w_p$$

Den robuste værdi,  $w^*$  beregnes ved, at der bruges median for de beregnede celleafstandene eller celle standardafvigelser for hvert af alle laboratorierne.

$$w^* = \text{median of } w_i \text{ (} i = 1, 2, \dots, p \text{)}$$

Den robuste værdi skal herefter opdateres ved iterering. Under iterationen opdateres celleafstanden eller celle standardafvigelserne, med en ny korrigeret værdi. Der beregnes en psi værdi, hvor der er ganget en grænsefaktor til den robuste værdi. Grænsefaktoren aflæses fra tabel 4.1, som tager udgangspunkt i antallet af frihedsgrader for en celle. Hvis  $w_i$  er en afstand, betyder det at antallet af frihedsgrader,  $v = 1$ . Hvis  $w_i$  er en standard afvigelse med  $n$  resultater er antallet af frihedsgrader,  $v = n-1$ .

Psi beregnes:

$$\psi = \eta \cdot w^*$$



Psi anvendes som den øvre grænseværdi, som hver af de robuste celle værdier højest kan antage. Hvis en celle værdi er højere end grænseværdien, bliver den sat til grænseværdien. På den måde bliver spredningen mellem celle værdierne mindre og der opnåes en bedre varians.

$$w_i^* = \begin{cases} \psi & \text{if } w_i > \psi \\ w_i & \text{otherwise} \end{cases}$$

Der kan derefter beregnes en ny robust værdi.

$$w^* = \xi \sqrt{\sum_{i=1}^p (w_i^*)^2 / p}$$

## 4.5 Resultater

Resultaterne fra robusthedsanalysen kan ses bagerst i appendisk.

### 4.5.1 Resultater fra forsøg A

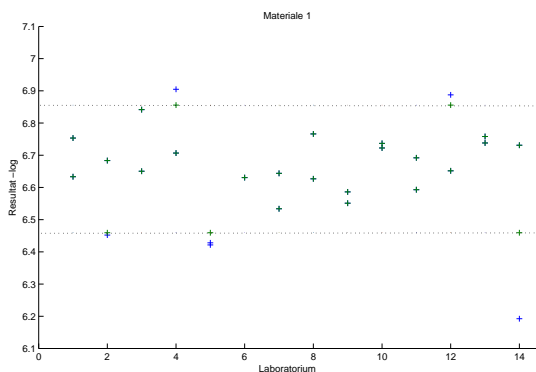
#### Materiale 1:

Laboratorium 14 der har fået ændret et resultat til den nedre grænseværdi, hvor den oprindelig ligger langt fra grænseværdien. Resultatet er ikke blevet klassificeret som værende en outlier fra Cochrans eller Grubbs´ test.

Der er ingen af resultater fra materialet, der er blevet klassificerede som værende outlier for én værdi i Grubbs´ test, men ved double testen fandt Grubbs´test outlier for laboratorium 4, 5, 12 og 14. Det er de samme laboratorier som har fået deres resultater ændret til grænseværdierne.

- Laboratorium 2 ændret til nedre grænseværdi
- Laboratorium 4 ændret til øvre grænseværdi
- Laboratorium 5 ændret til nedre grænseværdi
- Laboratorium 5 ændre til nedre grænseværdi
- Laboratorium 12 ændret til øvre grænseværdi
- Laboratorium 14 ændret til nedre grænseværdi

Figur 4.2 viser resultatet af robusthedsanalysen for resultaterne. Grænseværdierne er indtegnnet med en stiplede linie. De resultater som ligger udenfor grænseværdierne ser med blå, er blevet ændret til henholdsvis øvre/nedre grænseværdi, som ligger på den stiplede linie.



Figur 4.2: Robusthedsresultat for materiale 1

### Materiale 2:

Resultaterne for robustsanalysen viser hvordan laboratorium 4, som i Grubbs' test er klassificeret som værende outlier, adskiller sig markant fra grænseværdien. De andre resultater, som er blevet ændret ligger tæt på grænseværdierne.

- Laboratorium 4 ændret til øvre grænseværdi
- Laboratorium 4 ændret til øvre grænseværdi
- Laboratorium 5 ændret til nedre grænseværdi
- Laboratorium 5 ændret til nedre grænseværdi
- Laboratorium 13 ændret til øvre grænseværdi
- Laboratorium 14 ændret til nedre grænseværdi

Figur 4.3 viser resultatet af robusthedsanalysen.

**Materiale 3:**

Resultaterne for robustsanalysen viser hvordan laboratorium 4, som i Grubbs´ test er klassificeret som værende outlier, adskiller sig markant fra grænseværdien. De andre resultater, som er blevet ændret ligger tæt på grænseværdierne.

- Laboratorium 4 ændret til øvre grænseværdi
- Laboratorium 4 ændret til øvre grænseværdi
- Laboratorium 6 ændret til nedre grænseværdi
- Laboratorium 13 ændret til øvre grænseværdi

Figur 4.4 viser resultatet af robusthedsanalysen.

**Materiale 4:**

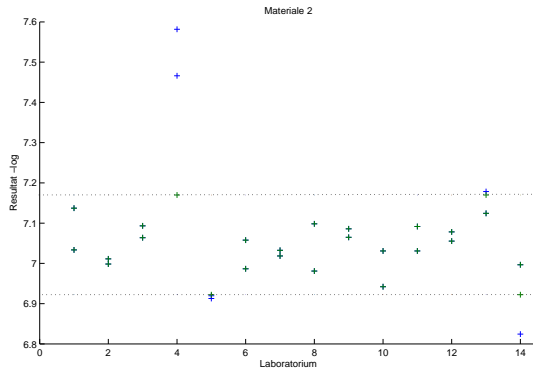
Materialet havde ingen outlier i Grubbs´ test og det kan på resultaterne. De få resultater som er blevet ændret ligger tæt på grænseværdierne. Variansen er stadig meget stor i forhold til de andre materialer.

- Laboratorium 6 ændret til nedre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi
- Laboratorium 12 ændret til nedre grænseværdi

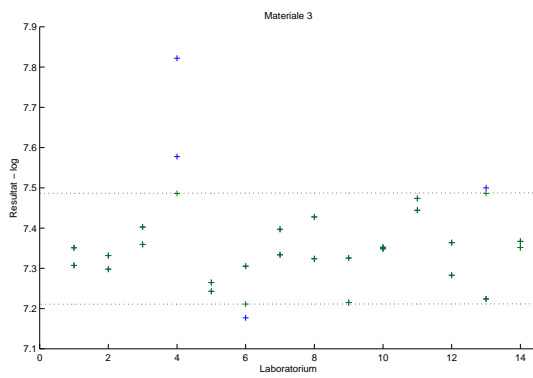
Figur 4.5 viser resultatet af robusthedsanalysen.

**Materiale 5:**

Materialet har fået klassificerede en outlier fra laboratorium 4. Outlier resultaterne ligger langt fra den øvre grænseværdi. De andre resultater som er blevet ændret ligger derimod meget tæt på deres grænseværdi.



Figur 4.3: Robusthedsresultat for materiale 2



Figur 4.4: Robusthedsresultat for materiale 3

- Laboratorium 4 ændret til øvre grænseværdi
- Laboratorium 4 ændret til øvre grænseværdi
- Laboratorium 6 ændret til nedre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi

Figur 4.6 viser resultatet af robusthedsanalysen.

### Konklusion

Resultaterne af robusthedsanalysen for forsøg A viser, at der er flere resultater der biver ændret til de beregnede grænseværdier, end der bliver klassificerede som værende outlier efter Cochrans og Grubb's test.

Materiale 1 og 2 som er flydende materialer har begge 6 resultater som er blevet ændret til deres grænseværdi, mens de faste materialer kun har 3 til 4 resultater som er ændret til deres grænseværdier.

Der er ikke nogen forskel på spredningen om det er et flydende eller fast materiale.

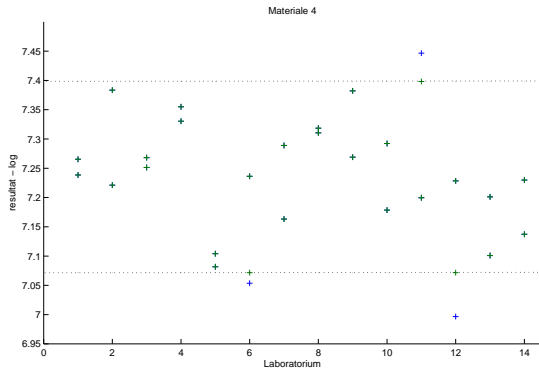
### Statistisk undersøgelse

De statistiske resultater af resultaterne efter de er blevet ændret ved anvendelse af robusthedsanalysen, ses i tabel 4.2. Ved sammenligningen med de rå resultater og data uden outlier, ses en lavere spredning i forhold til de rå data, mens resultaterne er næsten det samme for data uden outlier. Konfidensintervallerne er meget ens med de rå data, mens data uden outlier har et smallere konfidensinterval.

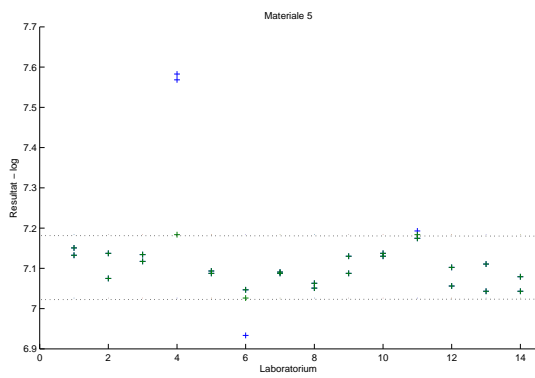
Spredningen af repeterbarheden og reproducerbarheden er fladet i forhold til de rå data, og meget ens med data uden outlier.

Variansanalysen, tabel 4.3, viser at interaktionen mellem materialer og laboratorier kan fjernes.

Variansanalysen, tabel 4.4, uden interaktionen viser, at begge faktorer er meget signifikante. Det var forventet, eftersom de har været det for både de rå data og data uden outlier. Materialerne har i alle tre tilfælde (rå data, data uden outlier og robustheds data) samme F-værdi. Det betyder, at signifikans niveauet for materialerne ikke ændres ved fjernelse af outlier eller anvendelse af robusthedsanalysen. Signifikantniveauet for laboratorierne er blevet halverede i forhold til de rå data og meget større end for data uden outlier.



Figur 4.5: Robusthedsresultat for materiale 4



Figur 4.6: Robusthedsresultat for materiale 5

	Materiale 1	Materiale 2	Materiale 3	Materiale 4	Materiale 5	Alle materialer	Flydende materialer	Faste materialer
Gennemsnit, $\hat{m}$	6.658	7.046	7.349	7.235	7.105	7.078	6.852	7.230
Madian	6.651	7.04	7.35	7.24	7.10	7.103	6.889	7.226
Spredning, $s\%$	11.65	7.28	8.08	9.60	4.62	25.06	21.84	12.59
Varians, $s^2\%$	1.36	0.53	0.65	0.92	0.24	6.28	4.77	1.59
Spredningen for reperaturbarhed, $s_r\%$	10.04	4.27	6.48	8.01	2.32	5.06	6.30	4.72
Spredningen for mellem laboratorier, $s_L\%$	6.03	6.01	4.92	5.38	4.08	24.64	21.11	11.74
Spredningen for reproducerbarhed, $s_R\%$	11.71	7.37	8.14	9.65	4.69	25.15	22.03	12.66
n	84	84	84	84	84	140	56	84
Konfidensinterval	(5.61;6.70)	(7.02;7.07)	(7.32;7.38)	(7.20;7.27)	(7.09;7.12)	(7.04;7.12)	(6.79;6.91)	(7.20;7.26)

Tabel 4.2: Statistiske beregninger for hvert materiale og alle materialerne tilsammen

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	4	7.74	1.94	419.61	< 2.2e-16 ***
Laboratorium	13	0.38	0.03	6.31	1.094e-07 ***
Materiale:Lab.	52	0.29	0.01	1.22	0.2222
Residualer	70	0.32	0.00		

Tabel 4.3: Variansanalyse med interaktion for forsøg A

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	4	7.74	1.94	384.38	< 2.2e-16 ***
Laboratorium	13	0.38	0.03	5.78	3.394e-08 ***
Residualer	122	0.61	0.01		

Tabel 4.4: Variansanalysen uden interaktion for forsøg A

## 4.5.2 Resultater fra forsøg B

### Materiale 1:

Figur 4.7 viser resultatet af robusthedsanalysen.

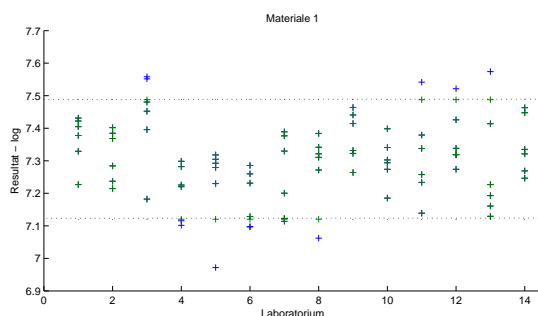
Det faste materiale indeholder ingen klassificerede outlier, og de resultater der er blevet ændret til grænseværdierne, ligger tæt ved grænseværdierne.

Ved sammenligning med forsøg A (materiale 3), er spredningen mindre i forsøg A end forsøg B. Den øvre grænseværdi er ens for de to forsøg, mens den nedre grænseværdi for forsøg B er større end for forsøg A. Der er flere laboratorier som kommer udenfor grænseværdierne i forsøg B. Der gælder for 9 laboratorier for forsøg B og kun 3 laboratorier for forsøg A.

- Laboratorium 3 ændret til øvre grænseværdi
- Laboratorium 4 ændret til nedre grænseværdi
- Laboratorium 5 ændret til nedre grænseværdi
- Laboratorium 6 ændret til nedre grænseværdi
- Laboratorium 7 ændret til nedre grænseværdi
- Laboratorium 8 ændret til nedre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi
- Laboratorium 12 ændret til øvre grænseværdi



- Laboratorium 13 ændret til øvre grænseværdi



Figur 4.7: Robusthedsresultat for materiale 1

### Materiale 2:

Figur 4.8 viser resultatet af robusthedsanalysen.

Det flydende materiale har ingen klassificerede outliers og de resultater, der er blevet ændret ligger tæt på grænseværdierne.

Ved sammenligning med forsøg A (materiale 2) er spredningen størst for forsøg B. De har næsten samme nedre grænseværdi, mens forsøg B har en markant højere øvre grænseværdi. Det er næsten samme antal laboratorier som har fået ændret deres værdier. Det er 6 for forsøg B og 4 for forsøg A.

- Laboratorium 1 ændret til øvre grænseværdie
- Laboratorium 9 ændret til øvre grænseværdi
- Laboratorium 10 ændret til nedre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi
- Laboratorium 12 ændret til øvre grænseværdi
- Laboratorium 14 ændret til øvre grænseværdi

### Materiale 3:

Det faste materiale har ingen klassificerede outlier.

Figur 4.9 viser, at laboratorium 5 og 13 kun har 2 ud af de 6 resultater liggende indenfor grænseværdierne. De andre laboratoriers resultater ligger tæt ved grænseværdierne. Materialet har ikke været inkluderet i forsøg A.

- Laboratorium 1 ændret til øvre grænseværdi
- Laboratorium 2 ændret til nedre grænseværdi
- Laboratorium 3 ændret til øvre grænseværdi
- Laboratorium 5 ændret til nedre grænseværdi
- Laboratorium 6 ændret til nedre grænseværdi
- Laboratorium 8 ændret til nedre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi
- Laboratorium 13 ændret til øvre grænseværdi

**Materiale 4:**

Figur 4.10 viser resultatet af robusthedsanalysen.

Det flydende materiale har i Cochran's test en straggler for laboratorium 13. Det kan ligeledes ses på robusthedsanalysen, at den har halvdel af sine resultater udenfor den øvre grænseværdi. Laboratorium 13 har to resultater udenfor grænseværdien, hvor det ene ligger langt fra grænseværdien. De resultater som ligger indenfor grænseværdierne ligger fundt i middel, med en lille spredning. De andre laboratorier ligger tæt på grænseværdierne. Materialet har en lille spredning, med en god repeterbarhed. Materialet er ikke blevet undersøgt i forsøg A.

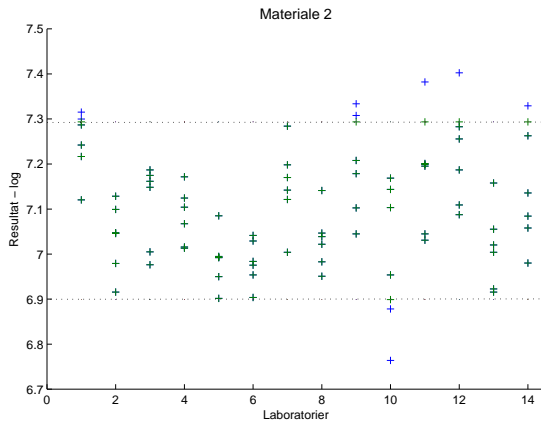
- Laboratorium 2 ændret til øvre grænseværdi
- Laboratorium 3 ændret til øvre grænseværdi
- Laboratorium 8 ændret til nedre grænseværdi
- Laboratorium 9 ændret til øvre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi
- Laboratorium 13 ændret til nedre grænseværdi

**Materiale 5:**

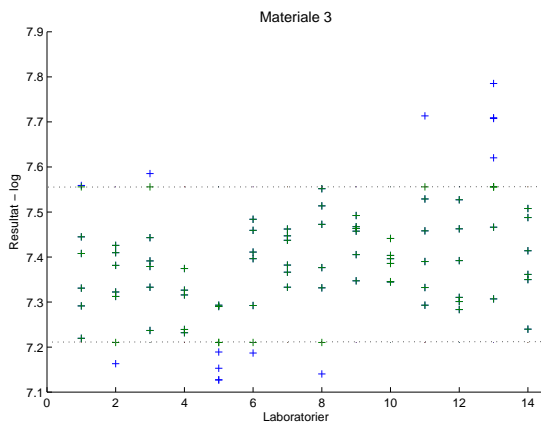
Figur 4.11 viser resultatet af robusthedsanalysen.

Det flydende materiale har en strejfer for laboratorium 12 i Cochran's test. Strejferen kan ligeledes ses i robusthedsanalysen, ved at dens resultater ligger langt fra grænseværdierne. Laboratorium 5 og 11 har begge kun halvdelen af deres resultater indenfor grænseværdierne, og deres resultater udenfor grænseværdierne ligger ikke tæt på grænseværdierne.

Materialet er blevet undersøgt i forsøg A (materiale 5), og ved sammenligning



Figur 4.8: Robusthedsresultat for materiale 2



Figur 4.9: Robusthedsresultat for materiale 3

af resultaterne har forsøg B en mindre spredning end forsøg A. Laboratorierne i forsøg A har en stor spredning i forhold til forsøg B, med samme nedre grænseværdi.

- Laboratorium 2 ændret til nedre grænseværdi
- Laboratorium 5 ændret til nedre grænseværdi
- Laboratorium 10 ændret til nedre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi
- Laboratorium 12 ændret til øvre grænseværdi

#### **Materiale 6:**

Figur 4.12 viser resultatet af robusthedsanalysen.

Det faste materiale har outlier fra laboratorium 4 og 9 i Cochran's test. Laboratorium 4 har kun et resultat udenfor den øvre grænseværdi, men dens resultater ligger meget spredt. Laboratorium 9 har to resultater liggende langt fra den nedre grænseværdi, mens de andre resultater ligger pænt indenfor grænseværdierne. De andre resultater ligger tæt på grænseværdier.

Ved sammenligning med forsøg A (materiale 5), har 7 laboratorier i forsøg B resultater liggende udenfor grænseværdierne, hvorimod det for forsøg A kun er gældende for 3 laboratorier. Spredningen er i forsøg B lidt mindre end i forsøg A.

- Laboratorium 3 ændret til øvre grænseværdi
- Laboratorium 4 ændret til øvre grænseværdi
- Laboratorium 5 ændret til nedre grænseværdi
- Laboratorium 6 ændret til nedre grænseværdi
- Laboratorium 8 ændret til nedre grænseværdi
- Laboratorium 9 ændret til nedre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi

#### **Materiale 7:**

Figur 4.13 viser resultatet af robusthedsanalysen.

Det flydende materiale har ingen outlier, men laboratorium 8 har to resultater liggende udenfor den nedre grænseværdi. Hvor det ene ligger langt fra grænseværdien. Resultaterne for de andre laboratorier ligger pænt tæt ved grænseværdierne. Materialet har ikke været inkluderet i forsøg A.



- Laboratorium 2 ændret til øvre grænseværdi
- Laboratorium 8 ændret til nedre grænseværdi
- Laboratorium 9 ændret til nedre grænseværdi
- Laboratorium 10 ændret til øvre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi
- Laboratorium 13 ændret til nedre grænseværdi

### Materiale 8:

Figur 4.14 viser resultatet af robusthedsanalysen.

Det faste materiale har en outlier for laboratorium 11 i Cochran's test. Det skyldes laboratoriet har en meget stor spredning på resultaterne. Der er et resultat, der ligger langt fra den nedre grænseværdi, samtidig er der et resultat udenfor den øvre grænseværdi. Laboratorium 8 har to resultater som ligger langt fra den nedre grænseværdi, mens de andre ligger pænt indenfor grænseværdierne. De andre resultater ligger pænt ved grænseværdierne.

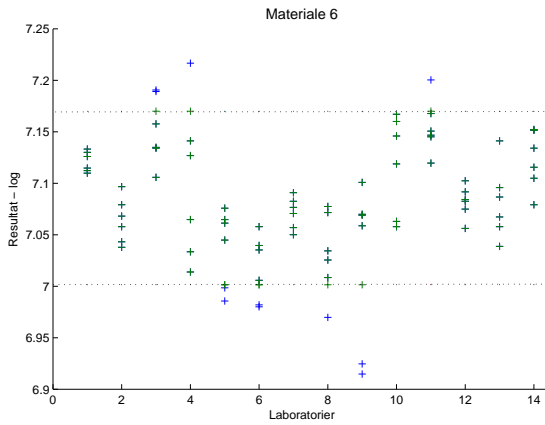
Ved sammenligning med forsøg A (materiale 4), er der 6 laboratorier i forsøg B med resultater udenfor grænseværdierne, og kun det halve i forsøg A. Spredning er ens i begge forsøg, med ens grænseværdier.

- Laboratorium 3 ændret til øvre grænseværdie
- Laboratorium 3 ændret til nedre grænseværdi
- Laboratorium 4 ændret til nedre grænseværdi
- Laboratorium 5 ændret til nedre grænseværdi
- Laboratorium 8 ændret til nedre grænseværdi
- Laboratorium 9 ændret til øvre grænseværdi
- Laboratorium 11 ændret til øvre grænseværdi
- Laboratorium 11 ændret til nedre grænseværdi

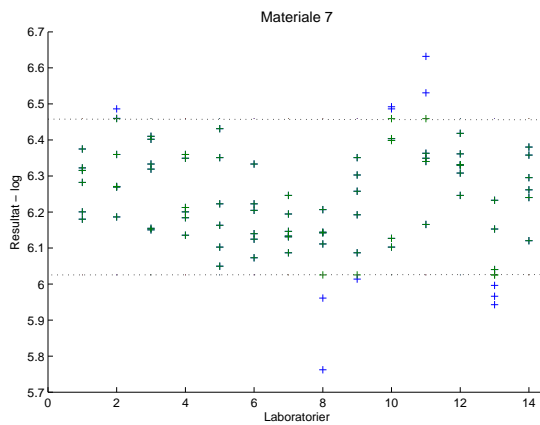
### Statistiske undersøgelser

Resultaterne af de statistiske beregninger for data efter anvendelse af robusthedsanalysen for forsøg B ses i tabel 4.5.

Tabellen viser, at gennemsnittet ligger mellem de rå data og data uden outlier. Gennemsnittet er i dette tilfælde nærmest deres median som betyder, at der er opnået et optimalt gennemsnit.



Figur 4.12: Robusthedsresultat for materiale 6



Figur 4.13: Robusthedsresultat for materiale 7

Spredningen er den laveste for de tre datasæt, og det samme gør sig gældende for repeterbarheden og reproducerbarheden.

Konfidensintervallerne for de enkle materialer ligger tæt på de andre resultater, mens konfidensintervallerne for alle materialerne, de flydende og faste materialer er ens.

Variansanalysen, tabel 4.6, viser at interaktionen mellem materialer og laboratorier ikke er signifikant. Interaktionen kan ikke fjernes, eftersom interaktionen mellem materialer og laboratorier med dagene underlagt er lidt signifikant. Laboratorier med dagene underlagt er blevet mindre signifikant i forhold til de rå data og data uden outlier.

Det er den bedste variansanalyse, der er opnået for de tre datasæt.

## 4.6 Konklusion

De rå data er blevet analyseret ved anvendelse af en robusthedsanalyse. Resultatet af metoden har betydet, at de resultater som lå udenfor de beregnede grænseværdier er blevet ændret til grænseværdierne. Ændringen har betydet, at gennemsnittet for et materiale ligger meget tæt på deres median og dens spredning er blevet mindre.

Anvendelsen af robusthedsanalysen i forhold til de rå data, har betydet at det ikke er alle interaktioner der længere er signifikante ved en variansanalyse. Det er derfor en fordel at anvende en robusthedsanalyse på de rå data, inden resultaterne skal anvendes.

Ved sammenligning med de resultater der blevet lavet ved en outlier undersøgelse, ligger de meget tæt på hinanden. Problemet ved outlier undersøgelsen er, at der bliver fjernet nogle af resultaterne, mens der for robusthedsanalysen bliver ændret nogle af resultaterne.

Ved at fjerne resultaterne fra forsøgene betyder det, at de ingen påvirkning har på de statistiske resultater. Mens de resultater som bliver ændret, stadig har en betydning for de statistiske resultater. Det er derfor en fordel at anvende en robusthedsanalyse, eftersom alle resultaterne bibeholdes og giver næsten det samme resultat, som ved anvendelse af en outlier undersøgelse.



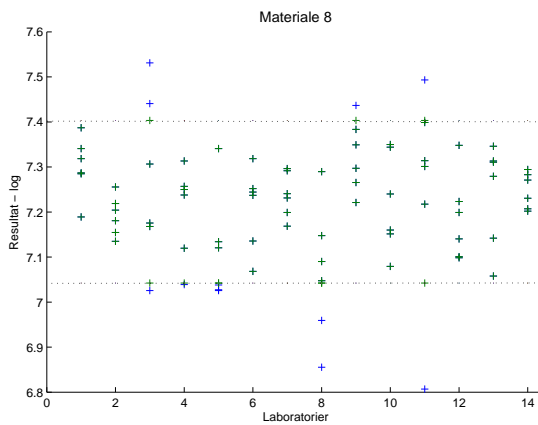
## 4.6 Konklusion

	Materiale 1	Materiale 2	Materiale 3	Materiale 4	Materiale 5	Materiale 6	Materiale 7	Materiale 8	Alle materialer	Flydende materialer	Faste materialer
Gennemsnit	7.339	7.120	7.364	6.6952	6.5413	7.0966	6.316	7.229	6.963	6.668	7.2571
Median	7.373	7.110	7.380	6.695	6.542	7.108	6.317	7.212	7.102	6.603	7.246
Varians, $s^2\%$	0.83	1.66	0.69	0.41	0.17	0.10	0.83	0.65	14.22	9.40	1.68
Spredning, $s\%$	9.09	12.87	8.29	6.39	4.10	3.21	9.10	8.04	37.71	30.66	12.97
Repetitbarhed, $s_r\%$	9.75	11.19	8.55	6.86	4.35	2.78	9.54	7.76	7.43	29.40	12.35
Mellem laboratorier, $s_L\%$	3.64	6.56	2.17	2.59	1.50	1.66	2.95	2.16	37.11	8.79	3.97
Reproducerbarhed, $s_R\%$	9.05	12.97	8.27	6.35	4.08	3.24	9.07	8.05	37.84	30.68	12.98
n	84	84	84	84	84	84	84	84	672	336	336
Konfidensinterval	(7.32; 7.36)	(7.09; 7.15)	(7.35; 7.38)	(6.68; 6.71)	(6.53; 6.55)	(7.09; 7.10)	(6.30; 6.34)	(7.21; 7.25)	(6.93; 6.99)	(6.64; 6.70)	(7.24; 7.27)

Tabel 4.5: Statistiske beregninger for forsøg B

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	7	90.99	13.00	2309.38	< 2.2e-16 ***
Laboratorium	13	0.17	0.01	2.33	0.0056 **
Lab.:Dag	28	0.34	0.01	2.15	0.0008 ***
Materiale:Lab.	91	0.62	0.01	1.21	0.1174
Materiale:Lab.:Dag	196	1.40	0.01	1.27	0.0294 *
Residuals	336	1.89	0.01		

Tabel 4.6: Variansanalyse for forsøg B



Figur 4.14: Robusthedsresultat for materiale 8

# Mixed model analyse

---

Mixed model anvendes til analyse af blandede analyser. Variable kan antage en af to typer af faktorer. De kan være fixed eller random faktorer og ved analyse af begge typer af faktorer i samme model anvendes en "mixed model".

## 5.1 Formål

Der vil i dette kapitel blive beskrevet hvordan forsøgsresultaterne kan analyseres via en mixed model analyse. Analysen vil foregå på de rå data, data uden outlier og data efter robusthedsanalysen. Varianserne for de to forsøg med de tre datasæt vil blive beregnet og sammenlignet.

## 5.2 Mixed model metode

De tidligere analyser har taget udgangspunkt i, at alle faktorerne har været fixed. Der vil i dette kapitel blive set på, hvis nogle af faktorer er fixed og andre af faktorerne er random.

I dette kapitel vil der først være en gennemgang af fixed, random og mixed faktor analyse. I den forbindelse vises forskellene på varianskomponenterne når de opfattes som henholdsvis fixed og random.

Herefter beskrives teorien for opbygningen af en mixed model, samt den tilhørende kovarians. Herefter beskrives model fitting metoder, der angiver hvordan random effekterne i en model kan reduceres.

En model kan være hierarkisk inddelt, som beskrives i det efterfølgende afsnit.

### 5.3 Fixed, random og mixed faktor analyse

En faktor kan være enten fixed eller random.

**Fixed faktor analyse**, som foreligger, hvis alle faktorens niveauer er bevidst udvalgte (kontrollerede). Det kan f.eks. være køn hvor det kan antage kvinde eller mand eller det kan være alder. I en fixed faktor analyse er vi interesseret i at vurdere virkningerne af de i variansanalysen konkret betragtede faktorniveauer.

**Random faktor analyse**, som foreligger, hvis niveauer for hver faktor er fremkommet som tilfældigt valgte blandt de mulige niveauer. I en random analyse er vi interesseret i en vurdering af faktorens generelle virkning, dvs. virkningerne for alle mulige niveauer. Der er af praktiske grunde udvalgt nogle af disse niveauer og foretaget analysen på basis af observationer svarende til de således fremkomne niveauer.

**Mixed faktor analyse**, som foreligger, hvis nogle faktorerers niveauer er bevidst udvalgte, medens de øvrige faktorerers niveauer er tilfældigt valgte. En mixed model er en blanding af fixed og random faktorer. Strukturen for en mixed model kan beskrives vha. en lineær model. Hvor observationer, residualer og fejl skal følge normalfordeling. Observationer, residualer og fejl skal ligeledes være uafhængige og variationen skal være homogen.

Som tidligere nævnt kan den samme faktor betragtes som enten værende fixed eller random. Det afhænger af hvilken analyse vi ønsker der skal udføres.

I det følgende eksempel betragtes nogle udvalgte materialer svarende til en fixed faktor. Der kan være tale om en række forskellige materialetyper, i det man af hensyn til den fremtidige produktion ønsker at undersøge, hvilken materialetype der er mest velegnet.

Antag i stedet for, at virksomheden har et stort antal næsten ens materialer, og der ønskes udført et forsøg med det formål, at undersøge den variation, som materialerne tilfører produktionen. Det er da hensigtsmæssigt at udføre forsøget med et antal materialer, som udtages tilfældigt, og den pågældende

variansanalyse bliver i så fald en random faktor analyse.

## 5.4 Mixed model teori

En mixed model tager udgangspunkt i en lineær model. Den generelle lineære model tager udgangspunkt i, at alle faktorerne er fixed. Med udgangspunkt i, at sætte både materialer og laboratorier til at være fixed faktorer, vises deres modelopbygning. Materialer er behandlinger og laboratorier blokke med følgende model.

$$y_{ij} = \underbrace{\mu}_{\text{gennemsnit}} + \underbrace{\alpha_i}_{\text{behandlingseffekt}} + \underbrace{\beta_j}_{\text{blokeffekt}} + \underbrace{\varepsilon_{ij}}_{\text{analysestøj}}, \varepsilon_{ij} \sim N(0, \sigma^2)$$

hvor  $i = 1, 2, 3, 4, 5$   $j = 1, 2, \dots, 14$  og  $\varepsilon_{ij}$  er uafhængige. Modellen for hver enkel observation vises udfra en mere simpel model, med kun  $i = 1, 2$  og  $j = 1, 2, 3$ .

$$\begin{array}{rcllcl} y_{11} & = & \mu & + & \alpha_1 & + & \beta_1 & + & \varepsilon_{11} \\ y_{21} & = & \mu & & + & \alpha_2 & + & \beta_1 & + & \varepsilon_{21} \\ y_{12} & = & \mu & + & \alpha_1 & & + & \beta_2 & + & \varepsilon_{12} \\ y_{22} & = & \mu & & + & \alpha_2 & & + & \beta_2 & + & \varepsilon_{22} \\ y_{13} & = & \mu & + & \alpha_1 & & & + & \beta_3 & + & \varepsilon_{13} \\ y_{23} & = & \mu & & + & \alpha_2 & & + & \beta_3 & + & \varepsilon_{23} \end{array}$$

Modellen viser hvordan hver parameter bliver beregnet udfra observationerne. Observationen i blok 3 med behandling 1,  $y_{13}$ , bliver beregnet ved at tillægge gennemsnittet,  $\mu$ , behandlingsparameteren for den første behandling,  $\alpha_1$ , og blok parameter for den tredje blok,  $\beta_3$ . Modellen kan skrives på følgende matrix notation

$$\underbrace{\begin{pmatrix} y_{11} \\ y_{21} \\ y_{12} \\ y_{22} \\ y_{13} \\ y_{23} \end{pmatrix}}_{\mathbf{Y}} = \underbrace{\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}}_{\mathbf{X}} \underbrace{\begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{21} \\ \varepsilon_{12} \\ \varepsilon_{22} \\ \varepsilon_{13} \\ \varepsilon_{23} \end{pmatrix}}_{\boldsymbol{\varepsilon}}$$

$\mathbf{X}$  betegnes som designmatrixen og  $\boldsymbol{\beta}$  er en vektor.

Matrix notationen for en mixed model, minder om matrix notationen for en lineær model, med kun fixed faktorer. Den store forskel er i stedet for én designmatrix, som vist i den lineære model, for kun fixed faktorer, indeholder en mixed model to designmatrixer. Der er en designmatrix  $X$ , som beskriver fixed effekten i modellen, og en designmatrix  $Z$  der beskriver random effekten i modellen. Fixed effekt designmatrixen  $X$  er konstrueret på samme måde som designmatrixen i den lineære model, men kun for de fixed effekter. Den har dimensionerne  $n \times p$ , hvor  $n$  er antal af observationer i datasættet og  $p$  er antallet af fixed effekt parametre i modellen. Random effekt design matrixen  $Z$  er konstrueret på samme måde som designmatrixen, men kun for random effekterne. Matrixen har dimensionerne  $n \times q$ , hvor  $q$  er antallet af random effekt koefficienter i modellen.

Eksemplet anvendes igen til visning af random blok effekt. Der anvendes stadig to behandlinger og tre blokke.

$$y_{ij} = \mu + \alpha_i + b_j + \varepsilon_{ij}$$

hvor  $b_j \sim N(0, \sigma_B^2)$  og  $\varepsilon_{ij} \sim N(0, \sigma^2)$ . Her er  $i = 1, 2$ ,  $j = 1, 2, 3$  og random effekten  $b_j$  og  $\varepsilon_{ij}$  er alle uafhængige. Matrix notationen for modellen er

$$\underbrace{\begin{pmatrix} y_{11} \\ y_{21} \\ y_{12} \\ y_{22} \\ y_{13} \\ y_{23} \end{pmatrix}}_{\mathbf{Y}} = \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}}_{\mathbf{X}} \underbrace{\begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{pmatrix}}_{\boldsymbol{\alpha}} + \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}}_{\mathbf{Z}} \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}}_{\mathbf{b}} + \underbrace{\begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{21} \\ \varepsilon_{12} \\ \varepsilon_{22} \\ \varepsilon_{13} \\ \varepsilon_{23} \end{pmatrix}}_{\boldsymbol{\varepsilon}}$$

hvor  $b \sim N(0, G)$  og  $\varepsilon \sim N(0, \sigma^2 I)$ . Kovariansmatrixen  $G$  for random effekter er i dette tilfælde en  $3 \times 3$  diagonal matrix med diagonal elementer  $\sigma_B^2$ . Matrix modellen skrives ofte via notationen

$$y = X\alpha + Zb + \varepsilon$$

hvor  $b \sim N(0, G)$  og  $\varepsilon \sim N(0, R)$ .

Vektoren  $\alpha$  er en samling af alle fixed effekt parametre, og vektoren  $b$  er en samling af alle random effekt koefficienter. Analysefejls kovarians matrix  $R = \text{var}(\varepsilon) = \sigma^2 I$  har dimensionen  $n \times n$  og random effekt koefficienten  $G = \text{var}(b)$  har dimensionen  $q \times q$ , hvor  $q$  er antallet af random effekt koefficienter. Hvis alle random effekt koefficienterne er uafhængige, betyder det, at  $G$  er en diagonal matrix med diagonal elementerne, som er lig med variansen af random effekt

koefficienterne.

### Kovariansmatrix $V$

Kovariansmatrixen beskriver kovariansen mellem to observationer for positionen  $i, j$  i en matrix. Matrixen vil altid være symmetrisk omkring diagonalen ( $V_{i,j} = V_{j,i}$ ), og diagonal elementerne er varianser af observationerne  $cov(y, y) = var(y)$ . For en fixed effekt model er kovariansen matrix altid givet ved  $V = \sigma^2 I$ , hvis alle observationer er uafhængige, mens der for mixed modeller vil fremkomme nogle mere interessante strukturer. Det betyder, at for en fixed effekt model, at alle observationer har en ens varians og observationerne er ukorrelerede. Dette gælder ikke for en model med random effekt. Der bliver observationerne fittet, hvilket betyder, at der kommer korrelerede observationer.

Ved fremstilling af kovarians matrixen for en mixed model, kan den opbygges direkte fra en matrix notation.

Kovariancen for  $y$ ,  $var(y) = V$  kan skrives som

$$y = X\alpha + Z\beta + \varepsilon \Leftrightarrow$$

$$var(y) = var(X\alpha + Z\beta + \varepsilon) \Leftrightarrow$$

$$V = var(X\alpha + Z\beta + \varepsilon)$$

Siden det antages, at random effekterne og residualerne er ukorrelerede, kan modellen splittes op til følgende.

$$V = var(X\alpha) + var(Z\beta) + var(\varepsilon)$$

Siden  $\alpha$  beskriver de fixed effekt parameter, hvor det gælder, at  $var(X\alpha) = 0$ . Hvor  $Z$  er en matrix for konstanterne. Derfor,

$$V = var(Z\beta) + var(\varepsilon) \Leftrightarrow$$

$$V = Zvar(\beta)Z' + var(\varepsilon)$$

Kovarians matrix  $V$  kan blive beregnet for modellen

$$V = ZGZ' + R$$

$$V = \text{var}(y) = ZGZ' + \sigma^2 I$$

$$V = \begin{pmatrix} \sigma^2 + \sigma_B^2 & \sigma_B^2 & 0 & 0 & 0 & 0 \\ \sigma_B^2 & \sigma^2 + \sigma_B^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2 + \sigma_B^2 & \sigma_B^2 & 0 & 0 \\ 0 & 0 & \sigma_B^2 & \sigma^2 + \sigma_B^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 + \sigma_B^2 & \sigma_B^2 \\ 0 & 0 & 0 & 0 & \sigma_B^2 & \sigma^2 + \sigma_B^2 \end{pmatrix}$$

Kovariansmatrixen  $V$  viser variansstrukturen for modellen. Den første observation  $y_{11}$  har variansen  $V_{11} = \sigma^2 + \sigma_B^2$ . Kovariansen mellem  $y_{11}$  og  $y_{21}$  er  $V_{11}(= V_{21}) = \sigma_B^2$  og kovariansen mellem  $y_{11}$  og  $y_{12}$  er nul.

Kovariansmatrixen  $V$  viser, at observationer for den samme random effekt bliver korrelerede, mens observationer for forskellige random effekter er ukorrelerede.

## 5.5 Model optimeringsmetoder

Der findes nogle numeriske metoder til optimering af en mixed model, og de vil blive beskrevet i dette afsnit. En mixed model kan ikke blive reduceret udelukkende ved anvendelse af en variansanalyse, som for en model med kun fixed faktorer, men der skal også anvendes maximum likelihood.

### Likelihood funktion

En Likelihood funktion anvendes til bestemmelse af de optimale observationer for en mixed model. Model optimeringen har tre forskellige komponenter: estimering af fixed effekter,  $\alpha$ , estimering af random effekter,  $b$  og estimering af varians parameterne (varianskomponenterne eller kovariansbetingelser).

Likelihood for varians parameterne og fixed effekterne kan defineres ved brugen af normalfordelingen.

Random effekt værdierne kan forventes at være nul, og derfor ingen påvirkning af gennemsnittet. Normalfordelingens gennemsnit kommer fra de fixed effekter,  $X\alpha$  og har en kovariansmatrixen  $V$ .

Likelihood funktionen er derfor defineret ved <sup>8</sup>

$$L_n(\theta) = \prod_{i=1}^n f(X_i; \theta)$$



Log-Likelihood funktionen er defineret ved <sup>8</sup>.

$$l_n(\theta) = \log L_n(\theta)$$

I praksis anvendes Log-likelihood funktionen mere end Likelihood funktionen, siden den er mere simpel at arbejde med og eftersom maximum værdierne er ens for de to funktioner. Her vises hvordan Likelihood og log-likelihood funktionen er for en normalfordeling <sup>8</sup>.

$$L_n(\mu, \sigma) = \prod_i \frac{1}{\sigma} \exp\left\{-\frac{1}{2\pi\sigma^2}(X_i - \mu)^2\right\} \Leftrightarrow$$

$$L_n(\mu, \sigma) = \sigma^{-n} \exp\left\{-\frac{1}{2\pi\sigma^2} \sum_i (X_i - \mu)^2\right\}$$

$$l_n(\mu, \sigma) = -\frac{1}{2}n \log(2\pi) - n \log \sigma - \frac{\sum_i (X_i - \mu)^2}{2\sigma^2}$$

Log-likelihood funktionen for en mixed model, der antager normalfordeling kan beskrives som følgende <sup>9</sup>.

$$l(y, \alpha, \gamma) = \frac{1}{2}\{n \log(2\pi) + \log |V(\gamma)| + (y - X\alpha)'(V(\gamma))^{-1}(y - X\alpha)\}$$

hvor  $\alpha$  er en vektor for fixed effekt parameterne og  $\gamma$  er en vektor for parameterne som anvendes i de to kovariansmatrixer G og R og ligeledes i V.  $\sigma = V|\gamma|$ ,  $\mu = X\alpha$ ,  $X_i = y$  og n = antal observationer.

$n \log(2\pi)$  er en konstant, som kan blive ignoreret i en maximisation proces. Det betyder, at Log-likelihood funktionen kan reduceres til

$$l(y, \alpha, \gamma) = \frac{1}{2}\{\log |V(\gamma)| + (y - X\alpha)'(V(\gamma))^{-1}(y - X\alpha)\}$$

### Maximum Likelihood estimation (MLE)

Maximum Likelihood estimation er baseret på maximering af Log-likelihood funktionen. Maximum likelihood udvælger de værdier i en model parameter, som gør data mest optimalt i forhold til hvad de andre værdier af en parameter ville have kunnet. I dette tilfælde, hvor vi har at gøre med en normalfordeling, betyder det at det giver et entydigt resultat. Parameterestimeringen findes ved <sup>9</sup>

$$(\hat{\alpha}, \hat{\gamma}) = \arg \min l(y, \alpha, \gamma)$$

Minimumet beregnes i tre step

1. Det forventes, at estimeringen af fixed effekt parameter  $\alpha$  er en funktion af random effekt parameterne  $\gamma$ . Det er uanset hvilke værdier for model para-

meterne der minimeres  $l$  og det betyder at  $\hat{\alpha}(\gamma) = (X'(V(\gamma))^{-1}X)^{-1}X'(V(\gamma))^{-1}y$ .

2. Estimatet af random effekt paramaterne er fundet ved minimering af  $l(y, \hat{\alpha}(\gamma), \gamma)$  som en funktion af  $\gamma$ .
3. Estimatet af fixed effekt parameterne er beregnet ud fra  $\hat{\alpha} = \hat{\alpha}(\hat{\gamma})$ .

Varians parameterne beregnes først, og derefter beregnes estimatet for fixed effekterne.

Ved analyse af varians parameterne som er fixed, findes værdierne for  $\alpha$ , som maximere log-likelihood. Fordelen ved denne metode er estimeringen af varians paramaterne. Det kan illustreres ved følgende eksempel. Vi har en simpel random analyse,  $x_1, x_2, \dots, x_n$  og ønsker at estimer gennemsnittet og variansen. Hvis  $\hat{\mu}$  er analysens gennemsnittet, det betyder at estimeringen af variansen ved ML er  $\sum_i (x_i - \hat{\mu})^2/n$ , i stedet for den ikke biased estimering  $\sum_i (x_i - \hat{\mu})^2/(n-1)$ . Ved estimering af varians parameterne er bias stort når antallet af frihedsgrader er få <sup>9</sup>.

### Residual maximum likelihood (REML)

Residual maximum likelihood bliver også betegnet som begrænset maximum likelihood. I dette tilfælde bliver parameteren  $\alpha$  elimineret fra log likelihood, hvormed den kun er defineret ved udtrykket for varians parameterne. REML er en modificeret udgave af maximum likelihood metoden. Minimeringen gennemføres, som for ML. Trin 1 og 3 er identiske, mens trin 2 ikke minimerer den negative log likelihood funktion  $l$  med derimod log likelihood for residual maximum likelihood,  $l_{re}$ .

Minimummet beregnes således i tre step <sup>9</sup>

1. Det forventes, at estimeringen af de fixed effekt parameter  $\alpha$  er en funktion af random effekt parameterne  $\gamma$ . Det er uanset hvilke værdier for model parameterne der minimeres  $l$  og det betyder at  $\hat{\alpha}(\gamma) = (X'(V(\gamma))^{-1}X)^{-1}X'(V(\gamma))^{-1}y$ .
2. Estimatet af random effekt paramaterne findes ved minimering af  $\frac{1}{2}\{\log |V(\gamma)| + (y - X\alpha)'(V(\gamma))^{-1}(y - X\alpha) + \log |X'(V(\gamma))^{-1}X|\}$ .
3. Estimatet af fixed effekt parameterne er beregnet ud fra  $\hat{\alpha} = \hat{\alpha}(\hat{\gamma})$ .

REML anvendes til optimering af likelihood for samtlige residualer. Samtlige residualer indeholder information om varians parameterne. Alle residualerne er defineret ved observationernes negative del af fixed effekterne i modellen.

Ved anvendelse af REML, hvor data er balancerede, kan random effekt parameterne estimeres uden bias. Derfor giver det god mening, at anvende REML estimering i en mixed model.

## 5.6 Hierarkisk inddeling

Den hierarkiske inddeling foretages både for fixed og random effekter. Hierarkisk inddeling betyder, at en faktor er underordnet en anden faktor.

### Hierarkisk fixed effekter

I begge forsøg gælder det for effekterne materialetype og materiale, at de er fixed. Materialetypen kan enten være på flydende eller fast form og indenfor hvert type er der udvalgt nogle materialer.

Alle materialerne i begge forsøg er nøje/bevidst udvalgt for, at de kan repræsenterer spæktret af materialer, der anvendes i dag. Derfor er materialer en fixed effekt. Hvis materialer skulle have været en random effekt, skulle de have været udtaget tilfældig blandt alle materialerne og uden krav om materiale type.

Det er undersøgt om den hierarkiske inddeling mellem materiale type og materiale har nogen påvirkning for resultat i forsøg A.

$$y_i = \mu + \alpha(\text{status}_i) + \beta(\text{materiale}_i) + \varepsilon_i$$

hvor status er materialetype,  $\varepsilon \sim N(0, \sigma^2)$  og  $\alpha(\text{status}_i)$  er materiale type.

Forsøg A	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Status	1	4.79	4.79	652.04	< 2.2e-16 ***
Status/Materiale	3	2.95	0.98	133.60	< 2.2e-16 ***
Residualer	135	0.99	0.01		

Tabel 5.1: Variansanalyse for forsøg A

Variansanalysen, tabel 5.1, viser at den hierarkiske inddeling påvirker resultaterne signifikant. Det betyder, at der er forskel på resultaterne for et flydende og fast materiale.

Det samme undersøges for forsøg B.

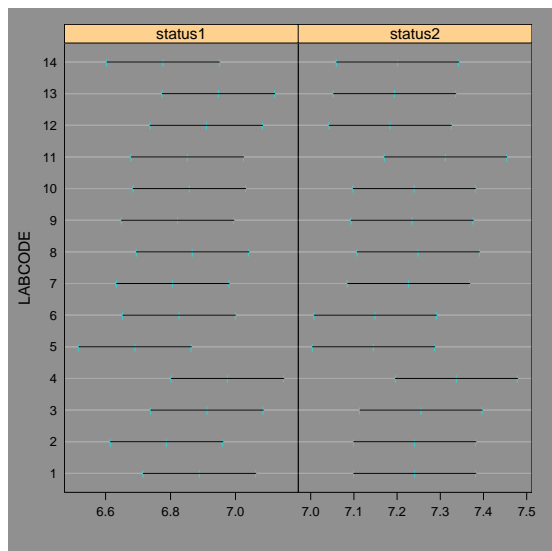
$$y_i = \mu + \alpha(\text{status}_i) + \beta(\text{materiale}_i) + \varepsilon_i$$

Variansanalysen, tabel 5.2, viser det samme som for forsøg A, at forskellen mellem flydende og fast materiale påvirker resultaterne signifikant.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Status	1	60.51	60.51	5405.61	2.2e-16 ***
Status:Materiale	6	34.93	5.82	520.14	2.2e-16 ***
Residualer	664	7.43	0.01		

Tabel 5.2: Variansanalyse for forsøg B

Den hierarkiske inddeling for fixed effekterne materiale type og materiale, betyder at, resultaterne vil ligge i to grupper. Det afspejler sig på figurerne 5.1 og 5.2.



Figur 5.1: Resultatet for de flydende (status 1) og de faste (status 2) enzymer for hvert laboratorium

Det ses på figurerne at resultaterne for de flydende materialer er mindre end for de faste materialer.

### Hierarkist random effekter

Den hierarkiske random effekt findes kun i forsøg B for dage og laboratorier. Dage er underordnet i forhold til laboratorier. De tre dage hvert laboratorium har analyserede deres forsøg over, er ikke de samme for hvert af laboratorierne. Derfor kan vi ikke undersøge om laboratorierne på 1. dagen har fået de samme analyseresultater, men kun undersøge om der er en dag til dag variation inden for

samme laboratorie. Det forventes en dag til dag variation i analyseresultaterne, som evt. kan komme fra dagens standardprøve. Den hierarkiske random effekt model er en tre faktor model. Det vises hvordan en et, to og tre faktor model opbygges og de tilhørende resultater.

### En-faktor model

Dette er den mest simple hierarkiske model, som beskrives ved følgende model.

$$y_i = \mu + \varepsilon_i, \varepsilon_i \sim N(0, \sigma^2)$$

Alle observationer er for denne model uafhængige, og derfor ikke en rigtig hierarkisk model. Variansparameteren i modellen er et estimat fra REML.

$$\hat{\sigma}^2 = \frac{1}{N-p} \sum_{i=1}^N (y_i - \hat{\mu}_i)^2$$

$N$  er det totale antal observationer,  $p$  er antallet af gennemsnit parametre og  $\hat{\mu}_i$  er det Best Linear Unbiased Estimate (BLUE) af gennemsnits værdien for den  $i$ 'th observation.

$$\hat{\mu}_i = (X\hat{\alpha})_i$$

Analysen af forsøg B efter en et faktor model, hvor der bliver set bort fra informationerne fra dag og laboratorium.

$$y_i = \mu + \alpha(\text{materiale}_i) + \varepsilon_i, \varepsilon \sim N(0, \sigma^2)$$

hvor  $i = 1, \dots, 672$  er observations nummer.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Materiale	7	90.99	13.00	1953.20	< 2.2e-16 ***
Residualer	664	4.42	0.01		

Tabel 5.3: Variansanalyse for en et faktor model for forsøg B

Modellens estimerede varians parameter  $\hat{\sigma}^2 = 28.6\%$ , to gange den negative log likelihood  $2l_{re} = -1469.325$  og AIC = -1310.

### To-faktor model

Forsøget analyseres ved en to faktor model

$$y_i = \mu + a(\text{sub}_i) + \varepsilon_i$$

hvor  $a(sub_i) \sim N(0, \sigma_a^2)$  og  $\varepsilon_i \sim N(0, \sigma^2)$  og alle  $a$  og  $\varepsilon$  er uafhængige variable. I to faktor modellen er alle observationer ikke længere uafhængige. Kovariansen mellem to observationer

$$cov(y_{i_1}, y_{i_2}) = \left\{ \begin{array}{ll} 0 & , \text{hvis } sub_{i_1} \neq sub_{i_2} \text{ og } i_1 \neq i_2 \\ \sigma_a^2 & \text{hvis } sub_{i_1} = sub_{i_2} \text{ og } i_1 \neq i_2 \\ \sigma_a^2 + \sigma^2 & i_1 = i_2 \end{array} \right\}$$

Hvor varians parameteren for emnet  $\sigma_a^2$  er variansen mellem emnerne. Residual varians parameteren  $\sigma^2$  er variansen mellem analyse resultaterne indenfor samme emne.

Varians parameterne og de fixed effekt parametre for to niveau modellen, estimeres ved brug af REML metoden. Analyse af data fra forsøg B ved brug af to faktor modellen, sker stadig uden informationer om dag, hvorimod informationerne om laboratorier medtages. To faktor modellen med random laboratorie effekter.

$$y_i = \mu + \alpha(material_i) + d(laboratorium_i) + f(material_i, laboratorium_i) + \varepsilon_i$$

hvor  $d(laboratorium_i) \sim N(0, \sigma_d^2)$ ,  $f(material_i, laboratorium_i) \sim N(0, \sigma_e^2)$  og  $\varepsilon_i \sim N(0, \sigma^2)$ .

	numDF	denDF	F-value	p-value
(Intercept)	1.00	664.00	2481689.49	<.0001
Materialer	7.00	664.00	1909.65	<.0001

Tabel 5.4: Variansanalyse for et to faktor model for forsøg B

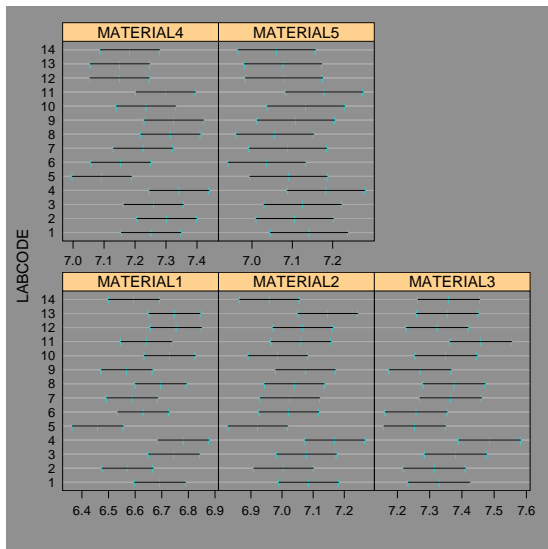
Modellens estimerede varians parameter ses i tabel 5.5.

Resultatet for den fixed effekt parameter er forskellig fra det beregnede fra et faktor modellen. Materialer havde i et faktor modellen en p-værdi på  $2.2e - 16$ , mens p-værdien i to faktor modellen er på 0.0001.

### Tre-faktor modellen

Den hierarkiske model med tre faktorer har modellen

$$y_i = \mu + a(blok_i) + b(sub_i) + \varepsilon_i$$



Figur 5.2:

AIC	-1391
BIC	-1341
LogLik	706
$\sigma_d$	1.15%
$\sigma_f$	0.74%
$\sigma$	8.1%

Tabel 5.5: Estimerte variansparameter for forsøg B

AIC	-1395
BIC	-1345
LogLik	708
$\sigma_d$	1.36%
$\sigma_f$	1.12%
$\sigma$	7.87%

Tabel 5.6: Estimerede variansparameter for forsøg B

hvor  $a(\text{blok}_i) \sim N(0, \sigma_a^2)$ ,  $b(\text{sub}_i) \sim N(0, \sigma_b^2)$  og  $\varepsilon_i \sim N(0, \sigma^2)$ .  $\sigma_a^2$  er variansen mellem laboratorierne,  $\sigma_b^2$  er variansen mellem dagene inden for samme laboratorium og  $\sigma^2$  er variansen mellem observationer fra samme dag. Parameterne estimeres ved brug af REML metoden, som ved to faktor modellen.

Analyse af forsøg B ved anvendelse af en tre faktor model. Hvor laboratorier er top niveau, dag er middel niveau og hvert resultat er det nederste niveau.

$$y_i = \mu + a(\text{mat}_i) + d(\text{dag}_i(\text{lab}_i)) + f(\text{dag}_i(\text{lab}_i), \text{mat}_i) + \varepsilon_i$$

hvor  $d(\text{dag}_i(\text{lab}_i)) \sim N(0, \sigma_d^2)$ ,  $f(\text{dag}_i(\text{lab}_i), \text{mat}_i) \sim N(0, \sigma_f^2)$  og  $\varepsilon \sim N(0, \sigma^2)$ .

Resultatet af modellen giver varianserne for alle random effekterne der ses i tabel 5.6.

## 5.7 Resultater for forsøg A

Mixed model modellen for forsøg A indeholder en fixed faktor fra materialerne og en random faktor fra laboratorierne. Interaktionen mellem materialer og laboratorier bliver analyseret som random.

Den største model for forsøget er opstillet først, og analyseret ved en variansanalyse. Det skal først undersøges om random interaktionen kan fjernes fra modellen.

Ved analyse om random interaktionen kan fjernes fra modellen, sammenlignes model 1, som indeholder random interaktionen med model 2, som ikke indeholder random interaktionen. Sammenligningen foretages ved en Pearson's chi square fordeling,  $\chi^2$ . Chi fordelingen giver som resultat en p-værdi. Hvis p-værdien er højere end 0.05, kan model 1 reduceres til den mindre model 2.



$$p - vrdi = 1 - pchisq(2 \cdot (|\log - likelihood_{model1} - \log - likelihood_{model2}|), f)$$

hvor  $f$  er antallet af frihedsgrader mellem de to modeller.

Model 1, for forsøg A analyseres.

$$y_i = \mu_i + \alpha(materiale_i) + d(laboratorium_i) + f(materiale_i, laboratorium_i) + \epsilon_i$$

	numDF	denDF	F-value	p-value
(Intercept)	1.00	135.00	70625.02	<.0001
Materiale	4.00	135.00	177.73	<.0001

Tabel 5.7: Resultat for model 1

Det undersøges om det er muligt, at reducere modellen ved fjernelse interaktionen. Model 2, viser reduktionen af model 1, ved fjernelse af interaktionen mellem materialer og laboratorier.

Model 2.

$$y_i = \mu_i + \alpha(materiale_i) + d(laboratorium_i) + \epsilon_i$$

Model 1 og model 2 sammenlignes og ses i tabel 5.8.

	call	df	AIC	BIC	logLik	Test	L.Ratio	p-value
model2	1.00	7.00	-277.75	-257.42	145.88			
model1	2.00	8.00	-276.33	-253.08	146.16	1 vs 2	0.57	0.45

Tabel 5.8: Sammenligning af model 1 og 2, om model 1 kan reduceres til model 2

Beregning af Chi sandsynligheden.

$$p - vrdi = 1 - pchisq(2 \cdot (|146.16 - 145.88|), (8 - 7)) = 0.45$$

Resultatet af sammenligningen er, at model 1 kan reduceres til model 2, hvormed modellen ikke indeholder en interaktion mellem materialer og laboratorier. Det betyder de statistiske undersøgelser for forsøg A tager udgangspunkt i model 2.

Resultat i %	Rå data		Data uden outlier		Data efter robusthedsanalyse	
	Fixed	Mixed	Fixed	Mixed	Fixed	Mixed
$S_{lab}$		9.4		4.3		4.8
$S_{lab,mat}$		4.4		14.2		2.2
$S_{residual}$		8.8		8.6		6.8
$S_{total}$		13.6		17.1		8.6
$S_{lab}$		9.5		4.4		4.9
$S_{residual}$	6.5	9.7	6.4	8.7	5.1	7.1
$S_{total}$	27.7	13.6	25.9	9.7	25.2	8.6

Tabel 5.9: Repeterbarheden og reproducerbarheden efter anvendelse af mixed model i forsøg A

$$S_{residual} = S_{reperbarhed}$$

$$S_{total} = S_{reproducerbarhed} = \sqrt{S_{lab}^2 + S_{lab,mat}^2 + S_{residual}^2}$$

Tabel 5.9 viser resultaterne for repeterbarheden og reproducerbarheden efter anvendelse af en mixed model.

Resultaterne for repeterbarheden er blevet lidt større efter anvendelse af en mixed model, det gælder for alle tre datasæt. Reproducerbarheden er derimod faldet meget i forhold til kun, at anvende en fixed model for de tre datasæt.

## 5.8 Resultater for forsøg B

Mixed model modellen for forsøg B indeholder en fixed faktor fra materialerne og to random faktorer fra laboratorierne og dage. Dagene er underlagt laboratorierne, og det vil blive vist med  $dagene_i(laboratorium_i)$ .

Model 1 for forsøget, vil være den største mulige model der kan opstilles. Den indeholder materialer som fixed faktor, laboratium nested med dagene som random og interaktionen mellem materialer og laboratium nested med dagene som random.

Model 1.

$$y_i = \mu + a(mat_i) + d(dag_i(lab_i)) + f(dag_i(lab_i), mat_i) + \varepsilon_i$$

hvor  $d(dag_i(lab_i)) \sim N(0, \sigma_d^2)$ ,  $f(dag_i(lab_i), mat_i) \sim N(0, \sigma_f^2)$  og  $\varepsilon \sim N(0, \sigma^2)$ .

Modellen undersøges for mulig reduktion, ved fjernelse af interaktionen mellem materiale og laboratorium/dagene. Model 2 ligner model 1, men uden interaktionen.

Model 2.

$$y_i = \mu + a(mat_i) + d(dag_i(lab_i)) + \varepsilon_i$$

Tabel 5.10 viser sammenligningen af model 1 og model 2, den kan ikke udføres ved beregning af Chi sandsynligheden. Det skyldes, at begge modeller har 11 frihedsgrader. Sammenligningen af modeller vil derfor tage udgangspunkt i modellernes Akaike Information Criterion (AIC) og Bayesian Information Criterion (BIC) værdier. Summen af AIC og BIC for model 1 er mindre end model 2, derfor kan modellen ikke reduceres. Det skyldes, at den bedste model er den hvor summen af AIC og BIC er mindst. De efterfølgende statistiske undersøgelser vil derfor blive foretaget med model 1.

	call	df	AIC	BIC	logLik
model2	1.00	11.00	-1397.53	-1348.05	709.77
model1	2.00	11.00	-1394.59	-1345.11	708.29

Tabel 5.10: Sammenligning af model 1 og 2 for eventuel model reduktion i forsøg B

Tabel 5.11 viser resultatet for kun to af de tre datasæt, ved anvendelse i en mixed model. Det skyldes, at datasættet uden outlier er ubalancerede og derfor ikke kan anvendes i en mixed model analyse i R.

Resultaterne for repeterbarheden (residual) er for de rå data langt mindre ved anvendelse af en mixed model.

Resultaterne for reproducerbarheden (total) er igen for de rå data langt mindre ved anvendelse af en mixed model. Der samme gør sig gældende for datasættet efter en robusthedsanalyse.

Resultat i %	Rå data		Data efter robusthedsanalyse	
	Fixed	Mixed	Fixed	Mixed
$S_{dag(lab)}$		5.30		1.36
$S_{dag(lab),mat}$		5.25		1.12
$S_{residual}$	33.83	9.02	7.43	7.87
$S_{total}$	39.86	11.71	37.84	8.06

Tabel 5.11: Repeterbarheden og reproducerbarheden efter anvendelse af mixed model i forsøg B

Datasæt	Spredning i %	laveste	middel	højeste
Rå data	laboratorier	6.22	9.50	14.52
	residual	8.52	9.66	10.95
Robust data	laboratorier	3.08	4.91	7.81
	residual	6.26	7.10	8.04

Tabel 5.12: Usikkerheden for forsøg A

## 5.9 Usikkerheden af repeterbarheden

Usikkerheden for repeterbarheden beregnes for begge forsøg. Usikkerhedsintervallet beskriver intervallet hvor resultat af repeterbarheden skal ligge indenfor for de kommende analyser.

Intervallerne kan beregnes i et statistik program. Jeg har valgt, at anvende resultatet fra mine beregninger i statistik programmet R.

Jeg har valgt ikke, at beskrive den bagved liggende teori for usikkerheds beregningen af repeterbarheden. Jeg nævner derfor kun kort resultater for min beregninger.

Intervallerne er blevet afslæst for de rå data og data efter en robusthedsanalyse.

### Forsøg A

### Forsøg B

Datasæt	Spredning i %	laveste	middel	højeste
Rå data	laboratorier/dage	3.86	5.30	7.27
	materialer,laboratorier/dage	4.33	5.25	6.36
	residual	8.39	9.02	9.68
Robust data	laboratorier/dage	0.70	1.36	2.63
	materialer,laboratorier/dage	0.33	1.12	3.86
	residual	7.37	7.87	8.40

Tabel 5.13: Usikkerheden for forsøg B

## 5.10 Konklusion

I fremtiden vil en analyse af enzym mængden i dyrefoder bestå af nogle tilfældige (kaldes også for random udtaget) laboratorier. Materialerne vil stadig være specifikt udtaget som betyder, at de ikke er udtaget tilfældigt (kaldes også for fixed udtaget). Hvis der indgår flere dage i forsøget vil det være tilfældige dage, hvor materialerne vil blive analyserede. Med udgangspunkt i dette haves en model med både fixed og random effekter. Blandingen af begge typer effekter i en model, er i dette kapitel blevet analyserede ved anvendelse af mixed model teorien.

Der er i dette kapitel blevet anvendt mixed model modeller for begge forsøg og for de rå data, data uden outlier og data efter en robusthedsanalyse. Resultatet har for alle datasættene udmåndet sig i en værdi for repeterbarheden og reproducerbarheden og deres 95% konfidensinterval også kaldt deres usikkerhed.

Resultaterne for alle datasættene giver en lavere værdi for repeterbarheden og reproducerbarheden, end ved anvendelse af en model hvor alle faktorer er fixed faktor. Det betyder, at værdien for repeterbarheden og reproducerbarheden afhænger af hvilken model der anvendes til analyse af resultaterne.

Usikkerheden for repeterbarheden og reproducerbarheden skal af Novozymes, anvendes til fremtidige kvalitetskontrol af målemetoden.



## KAPITEL 6

# Diminsionering af ny ringprøvning

---

### 6.1 Formål

I dette kapitel vil resultaterne for reproducerbarheden fra foregående kapitel blive anvendt til rediminsionering af en ny ringprøve. Der tages udgangspunkt i forsøg B, som indeholder tre dage med to gentagelser for hver dag.

Diminsioneringen af en ny ringprøve, med udgangspunkt i den beregnede reproducebarhed indenfor et 95% konfidensinterval for laboratorier og dage underlagt laboratorier, vil der kunne blive beskrevet hvor mange laboratorier, dag og gentagelser der vil være det optimale.

### 6.2 Teoretisk implementering

De materialer der er blevet anvendt i begge forsøg er som tidligere nævnt bevidst udvalgt og dermed ikke tilfældig (random) udtaget. Det skyldes, at der på forhold allerede vides, at der er forskel fra materiale til materiale. Konfidensintervallet for det beregnede reproducerbarhed, skal derfor gerne dække de materialer som ligger i yderkanter, og ikke kun de materialer som ligger i midten.

	SSQ	Antal frihedsgrader	$s^2$	EMS
Laboratorium		l-1	$s_L^2$	$\sigma_G^2 + g \cdot \sigma_D^2 + g \cdot d \cdot \sigma_L^2$
Dag		$(d-1) \cdot l$	$s_D^2$	$\sigma_G^2 + g \cdot \sigma_D^2$
Gentagelser		$(q-1) \cdot l \cdot d$	$s_G^2$	$\sigma_G^2$
Total				

Tabel 6.1: Den teoretiske variansanalyse

Det skyldes, at konfidensintervallet fremover skal anvendes til kvalitetskontrol af målemetoden.

Det beregnede 95% konfidensinterval for reproducerbarheden er blevet beregnet for alle materialerne lagsammen. Eftersom der er stor variation i resultaterne fra materiale til materiale, vil diminonering blive beregnet for hvert materiale. Hvert materiale kan følger følgende model:

$$Y_{ijk} = \mu + L_i + D(L)_{j(i)} + G(LD)_{k(ij)}$$

Hvor  $L_i$  repræsenterer et af de i alt 14 laboratorier,  $D(L)_{j(i)}$  repræsenterer en af de i alt tre dage, som er underlagt hvert laboratorium,  $G(LD)_{k(ij)}$  repræsenterer en af de to gentagelser indenfor en dag, og  $Y_{ijk}$  repræsenterer et enkelt resultat indenfor en gentagelse, en dag og et laboratorium, og alle faktorerne betragtes som værende random.

I forsøg B er der i alt 672 resultater fordelt på de otte materialer. Det betyder, at der er 84 resultater for hver materiale.

Den teoretiske variansanalyse kan ses i tabel 6.1.

Hvor SSQ er summen af afvigelse kvadrat,  $s^2$  er variansen og EMS er Expected Mean Square (forventningsværdier af  $s^2$ -værdier).  $s_L^2$ ,  $s_D^2$  og  $s_G^2$  er variansen for laboratorium, dag og gentagelser. l er antallet af laboratorier, d er antallet af dage og q er antallet af gentagelser.

Den teoretiske variansanalyse anvendes til, at beskrive variansen for reproducerbarheden.

Det kan antages, at variansen  $s^2$  kan sættes lig med EMS. Det betyder, at de estimerede varianser for laboratorium, dag og gentagelse kan beregnes. Satterhwaite metoden anvendes til, at beregne variansen for reproducerbarheden. Det gøres ved at opsætte en linear kombination, hvor EMS linierne anvendes.

$$\hat{\sigma}_G^2 = s_G^2$$



$$\hat{\sigma}_D^2 = \frac{s_D^2 - s_G^2}{g}$$

$$\hat{\sigma}_L^2 = \frac{s_L^2 - s_D^2}{g \cdot d}$$

$$\hat{\sigma}_{repro}^2 = \hat{\sigma}_G^2 + \hat{\sigma}_D^2 + \hat{\sigma}_L^2$$

Værdierne for de estimerede varianser indsættes i formelen for variansen for reproducerbarheden.

$$\hat{\sigma}_{repro}^2 = s_G^2 + \frac{s_D^2}{g} - \frac{s_G^2}{g} + \frac{s_L^2}{g \cdot d} - \frac{s_D^2}{g \cdot d} = (1 - \frac{1}{g})s_G^2 + \frac{s_D^2}{g}(1 - \frac{1}{d}) + \frac{s_L^2}{g \cdot d} \iff$$

$$\hat{\sigma}_{repro}^2 = a_1 \cdot s_G^2 + a_2 \cdot s_D^2 + a_3 \cdot s_L^2$$

Det medfører at de tre konstanter kan defineres.

$$a_1 = 1 - \frac{1}{g}$$

$$a_2 = (1 - \frac{1}{d}) \cdot \frac{1}{g}$$

$$a_3 = \frac{1}{g \cdot d}$$

Der kan nu konstrueres en F test ved anvendelse af varianserne for laboratorium, dag, gentagelser og reproducerbarheden. Testen kan en pseudo F test. Det er ikke resultatet af F test der i denne situation er det interessante, men bestemmelsen af antallet af frihedsgrader til reproducerbarhedsvariansen.

Hvor værdien af varianserne  $s_G^2$ ,  $s_D^2$  og  $s_L^2$  afhænger af antallet af frihedsgrader for hver af dem. Antallet af frihedsgrader som tilhører variansen for reproducerbarheden kan beregnes på følgende måde <sup>10</sup>.

$$f_{repro} = \frac{(\hat{\sigma}_{repro}^2)^2}{a_1^2 \left[ \frac{(s_G^2)^2}{l \cdot d \cdot (g-1)} \right] + a_2^2 \left[ \frac{(s_D^2)^2}{(d-1) \cdot l} \right] + a_3^2 \left[ \frac{(s_L^2)^2}{l-1} \right]}$$

95% konfidensintervallet for variansen for reproducerbarheden giver den øvre grænseværdi og den nedre grænseværdi <sup>10</sup>.

$$\frac{f_{repro} \cdot \hat{\sigma}_{repro}^2}{\chi_{\alpha/2, f_{repro}}^2} \leq \hat{\sigma}_{repro}^2 \leq \frac{f_{repro} \cdot \hat{\sigma}_{repro}^2}{\chi_{1-\alpha/2, f_{repro}}^2}$$

## 6.3 Implementering af metoden

Metoden er blevet implementerede i MatLab, hvor den beregner et resultat for alle kombinationerne for hvert af materialerne. Resultaterne for hvert af materialerne består af mange informationer. Informationerne er følgende:

- Antal laboratorier
- Antal dage
- Antal gentagelser
- Kostpris for et laboratorium
- Kostpris for en dag
- Kostpris for en gentagelse
- Spredningen for laboratorium for det pågældende materiale
- Spredningen for dage for det pågældende materiale
- Spredningen for gentagelser for det pågældende materiale

Antallet af laboratorier kan være mellem 2 til 20, antallet af dage kan være mellem 2 til 10 og antallet af gentagelser kan være mellem 2 til 10. Kostpriserne er blevet sat af Peter Thyregod fra Novozymes og kan til en vær tid ændres i programmet til andre kostpriser. Kostprisen for et laboratorium er sat til 4, kostprisen for en dag er blevet sat 2 og kostprisen for en gentagelse er blevet sat 1. Det betyder, at det er fire gange dyre, at tillægge et laboratorium i forhold til en gentagelse og to gange dyre, at tillægge en dag i forhold til en gentagelse. Spredningerne for laboratorium, dage og gentagelser er blevet beregnet i R for hvert af materialerne i forsøg B. Spredningerne for hvert af materialerne kan ses i tabel 6.2.

95% konfidensintervallet for spredningen for reproducerbarheden beregnes for alle 1539 kombinationer ( $19 \cdot 9 \cdot 9$ ).

Resultatet for konfidensintervallet for det gennemførte forsøg B anvendes, som de nominelle værdier i forhold til de andre beregnede konfidensintervaller. Det betyder, at konfidensintervallet for 14 laboratorier med 3 dage og 2 gentagelser er den nominelle værdi i forhold til de andre kombinationer.

Der beregnes en gevinst.

	$\sigma_{lab.}$	$\sigma_{dage}$	$\sigma_{gent.}$
Materiale 1	$6.684 \cdot 10^{-7}$	$6.028 \cdot 10^{-7}$	0.0903
Materiale 2	0.0482	0.0871	0.0802
Materiale 3	$4.544 \cdot 10^{-8}$	$8.061 \cdot 10^{-7}$	0.0823
Materiale 4	$7.978 \cdot 10^{-7}$	0.0191	0.0605
Materiale 5	$7.161 \cdot 10^{-7}$	0.0253	0.0318
Materiale 6	0.0122	0.0219	0.0196
Materiale 7	$7.207 \cdot 10^{-8}$	$7.288 \cdot 10^{-7}$	0.0904
Materiale 8	$2.038 \cdot 10^{-6}$	0.0645	0.0471
Alle materialer	$1.979 \cdot 10^{-6}$	$9.775 \cdot 10^{-6}$	0.3767

Tabel 6.2: Spredninger for hvert materiale og alle materialerne samlet

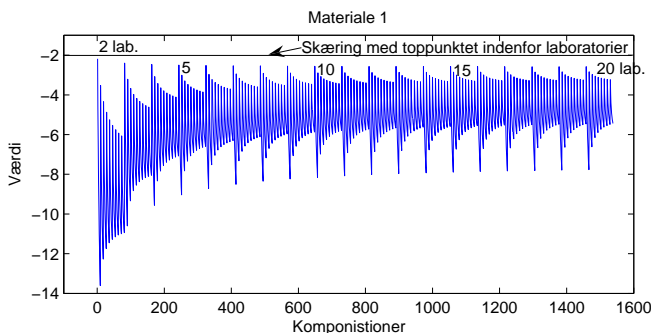
$$Gevinst = \frac{(\sigma_{repro,hj} - \sigma_{repro,lav})^2}{\sigma_{repro}^2}$$

Hvor  $\sigma_{repro,hj}$  er den højste værdi i konfidensintervallet,  $\sigma_{repro,lav}$  er den laveste værdi i konfidensintervallet og  $\sigma_{repro}^2$  er reproducerbarheden.

Resultaterne for prisen og gevinsten af forsøget bliver omregnet til nominelle værdi, dermed er det nemmere at forholde sig til deres resultater. De resultater som har en nominel pris mindre end en, er billigere end forsøg B, mens de resultater som har en højere nominel værdi end en, er dyrere end forsøg B. For gevinsten gælder det, at de nominelle værdi som er mindre end en, er bredden af konfidensintervallet mindre end i forsøg B, mens de resultater som har en højere nominel værdi end en, er bredden af konfidensintervallet større end i forsøg B. Forholdet mellem prisen og gevinsten minimeres for, at finde den optimale dimension af forsøget for det pågældende materiale.

$$Vaerdi = \frac{pris - 5 \cdot gevinst}{pris}$$

Hvor pris er prisen for en kombination og gevinst er gevinsten for en kombination. Gevinsten ganges med fem som betyder, at et bedre konfidensinterval koster 5. Det betyder, at gevinsten bliver fem gange mindre, når konfidensintervaller bliver mindre end konfidensintervallet for den nominelle værdi. Resultatet af dimensioneringen ændre sig ikke ved anvendelse af en anden kostværdi end 5. Programmet er ligeledes blevet afprøvet med kostværdierne 2 og 8. Ved anvendelse af en kostværdi på 2 betyder det, at spredningen på værdiaksen



Figur 6.1: Det optimale antal laboratorier for materiale 1

bliver mindre. Anvendelse af en kostværdi på 8 betyder det, at spredningen på værdi akse bliver større i forhold til anvendelse af en kostværdi på 5. Graffen og diminsionering bliver i alle tilfældene ens.

## 6.4 Reultater

Figur 6.1 viser hvordan antallet af laboratorier kan aflæses, og hvilket antal laboratorier som er det optimale i forhold til den valgte mængde.

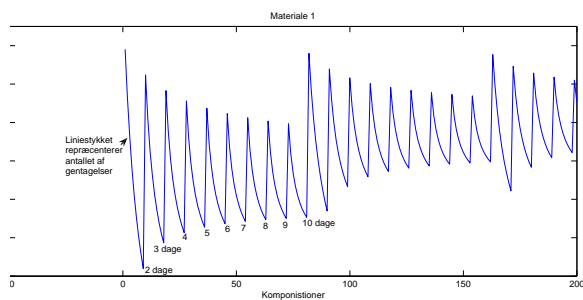
Det optimale antal laboratorier for materiale 1 er 2, mens det for materiale 2 er 20 laboratorier, som ses på figur 6.2.

Figur 6.2 viser hvordan antallet af dage og gentagelser kan aflæses, og hvilke antal dage og gentagelser som er det optimale i forhold til de valgte mængder.

Det optimale antal dage for materiale 1 er 2 og 10 gentagelser.

Det optimale antal dage aflæses på figuren efter det optimale antal laboratorier er fundet. Hvis det optimale antal dage er få, kan det ses på figuren ved, at bunden af kurven indenfor hvert mængde laboratorier stiger. Hvis det modsatte er gældende, at det optimale antal dage er det højeste, som for materiale 5 kan det ses ved, at bunden af kurven falder. For materiale 4 tilfælde gælder det, at det optimale antal dage er 4. Det ses ved, at bunden af kurven først falder til fire dage, og derefter stiger.

Det optimale antal gentagelser kan ikke direkte aflæses på figuren, men det kan ses om det optimale antal er højt eller lavt. Hvis det optimale antal gentagelser



Figur 6.2: Det optimale antal dage og gentagelser for materiale 1

er lavt ses det ved, at det øverste af kurven stiger indenfor hvert mængde af laboratorier. Det kan ses på figur 6.4 for materiale 2. Hvis det modsatte gælder, at det optimale antal gentagelser er højt, falder den øverste del af kurven. Det kan ses på figur 6.2 for materiale 1

Resultaterne for materialerne 1, 2 og 5 kan ses på figur 6.3 - 6.5. De repræcenter de tre typer resultater som dannes af de otte materialer.

Materiale 1 har for både laboratorium og dage en lav spredning, mens spredningen for gentagelser er høj i forhold til spredningen for laboratorium og dage. Resultatet af diminsioneringen giver to laboratorier, to dage og 10 gentagelser. Det betyder, at laboratorium og dage vælges som det lavest mulige, mens gentagelser vælges som det højest mulige. Dette resultat gælder ligeledes for materialerne 3, 4 og 7, eftersom de har en lav spredning for både laboratorier og dage, og en høj spredning for gentagelser.

Materiale 2 har for laboratorium, dage og gentagelser en spredning på samme niveau. Det betyder, at resultatet af diminsioneringen giver 20 laboratorier, fire dage og to gentagelser. I dette tilfælde er antallet af laboratorier valgt som det højest mulige, antallet af dage er middelværdien af de mulige dage og antallet af gentagelser er det mindst mulige. Dette resultat gælder ligeledes for materiale 6.

Materiale 5 har for labpratorium en lav spredning og en høj spredning for både dage og gentagelser. Resultatet af diminsioneringen giver to laboratorier, 10 dage og tre gentagelser. Det betyder, at antallet af laboratorier bliver det mindst mulige, antallet af dage det højest mulige og antallet af gentagelser næsten det mindst mulige. Dette resultat gælder ligeledes for materiale 8, dog med to gentagelser.

	Laboratorier	Dage	Gentagelser
Materiale 1 Materiale 3 Materiale 4 Materiale 7 Alle materialer	2	2	10
Materiale 2 Materiale 6	20	4	2
Materiale 5	2	10	3
Materiale 8	2	10	2

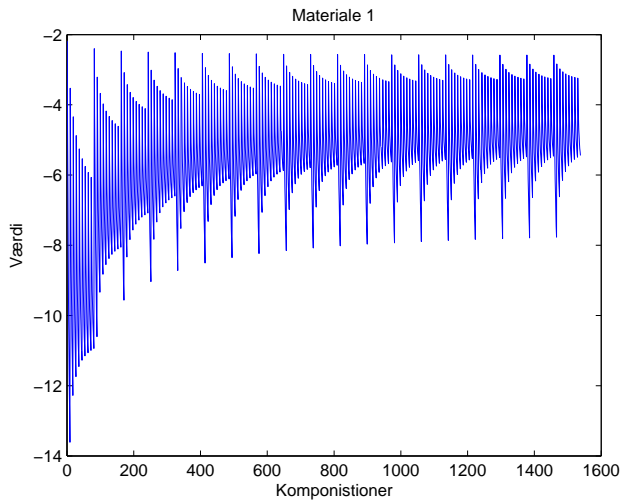
Tabel 6.3: Resultaterne af dimensionering af hvert materiale og alle materialerne samlet

Med udgangspunkt i alle otte materialer, som repræsenterer spændet af alle de enzym materialer der findes på marked i dag, er der blevet lavet en dimensionering af et nyt forsøg gældende for alle materialer. Det giver to laboratorier, to dage og 10 gentagelser. Det betyder, at laboratorier og dage er det laveste mulige og gentagelser det højeste mulige, med udgangspunkt i de værdier programmet har regnet med.

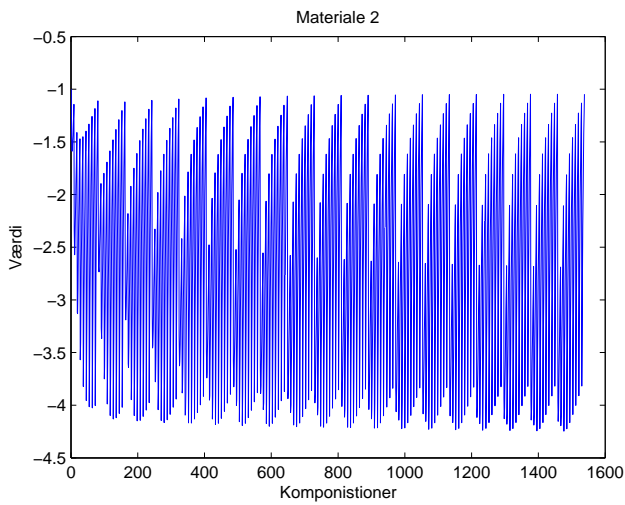
Tabel 6.3 viser resultatet af dimensioneringen af hvert materiale.

Resultaterne for dimensionering af en ny forsøgsplan for hvert af materialerne, viser at dimensioneringsmodellen er følsom overfor ændringer i spredningerne indenfor laboratorium, dage og gentagelser.

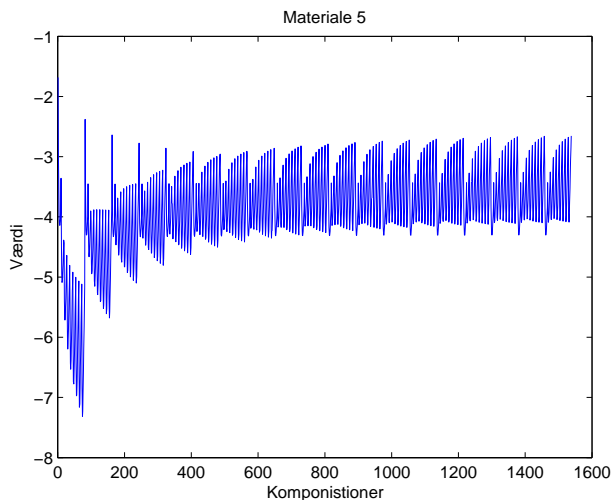
Resultatet for alle materialerne lagt sammen giver ikke en optimal dimensionering i forhold til hvert af materialerne. Dimensioneringen vil ikke kunne anvendes for alle materialerne. Anvendelsen af dimensioneringen vil derfor resulterer i, at nogle materialer ikke giver de rigtige reproducerbarheder og reproducerbarheder. SI



Figur 6.3: Diminsionering af materiale 1



Figur 6.4: Diminsionering af materiale 2



Figur 6.5: Diminsionering af materiale 5

## 6.5 Konklusion

Resultaterne af spredningerne for hvert af materialerne, giver for de fleste af materialerne en lav spredning for laboratorium og dage i forhold til spredningen for gentagelser. Det skyldes den måde hvor på forsøg B er blevet udført.

Det gælder for forsøg B, at de 14 laboratorier hver har fået tilsendt otte materialer. Den tilsendte vejledning beskriver, at laboratoriet selv skal udtage to prøver fra hvert af de otte tilsendte materialer på tre uafhængige dage. Det betyder, at laboratorierne selv kan kontrollere deres resultater, indenfor hvert materiale. De kan eventuelt have valgt, at udfører tre prøver den pågældende dag, og derefter vælge de to prøveresultater der ligger tættest på hinanden. De efterfølgende dage har laboratoriet en rettet snor for hvor resultaterne for hvert materiale gerne skal ligge indenfor. Det vil helt naturligt betyde, at variationen mellem prøveresultaterne indenfor en dag og mellem dage vil være lille.

Variationen mellem laboratorierne vil stadig være den samme, eftersom de ikke ved hvad de andre laboratorier har fået af resultater. Det kan ligeledes ses af spredningerne for gentagelser for hvert materiale, at de ligger konstant for materiale til materiale. Der er ikke de samme store svinger i spredningerne for gentagelser, som for laboratorier og dage.

Det vil have været mere hensigtsmæssigt, at udføre forsøget som et blind forsøg. Det betyder, at Novozymes skulle have tilsendt 48 prøver til hvert af laboratorierne. De 48 prøver er nummerede fra 1 til 48, hvor de første 16 prøver udføres



en dag, de næste 16 prøver en anden dag og de sidste 16 prøve en tredje dag. Indenfor hver af de 16 prøver er der to prøver fra hvert af de otte materialer. Hved at udfører forsøget på denne måde, ved laboratoriet ikke hvilke prøver der tilhører det enkle materiale, og de kan derfor ikke kontrollere deres resultater.

Resultatet af diminsionering af en ny ringprøve, taget udgangspunkt i hvert af materialerne, eftersom materialerne vejere meget mellem hinanden. Det udviklede program kan ligeledes anvendes for alle materialerne.

Resultatet for hvert af materialerne giver enten få laboratorier med få dage og mange gentagelser eller mange laboratorier med få gentagelser. Resultaterne viser, at diminsioneringen af et nyt forsøg er meget følsom overfor vejeresioner af spredningerne for laboratorier, dage og gentagelser.

Det udviklede program indeholder værdier på kostpriserne for laboratorier, dage, gentagelser og ændring i konfidensintervallet. Det er alle kostpriser, som kan ændres og de vil kunne påvirke resultatet af diminsioneringen. Det vil derfor anbefales, at Novozymes afprøver programmet med andre kostpriser.



# Konklusion

---

Der er i dette projekt blevet undersøgt forskellige metoder til at analysere resultaterne fra studiet med enzymmængde i dyrefoder. Studiet er udført af Novozymes og følger ISO 5725 standarden. Der er indledningsmæssigt givet et overblik over resultater for de to forsøg A og B. Det tager udgangspunkt i resultaterne før der er gennemført nogen form for rensning af data - altså på rådata.

Resultaterne viser, at forsøg A giver en god repeterbarhed indenfor hvert af laboratorierne, men at deres resultater har en stor variation fra laboratorium til laboratorium. Resultaterne viser også at laboratorium fire har en god repeterbarhed, men at laboratoriets resultater ligger ikke på samme niveau som de andre laboratorier.

De indledende analyser for forsøg B, giver nogle gode repeterbarheder indenfor hvert laboratorium og det samme gælder for reproducerbarheden. Forsøget viser, at der indenfor hvert laboratorium ikke er den store dags til dags variation. Resultaterne mellem laboratorierne kan som i forsøg A variere, men det har ingen direkte indflydelse på analysen.

Det gælder for begge forsøg, at de analyserede materialer ikke ligner hinanden i deres resultater. Det betyder, at det samlede materiale giver nogle signifikante afvigelser. Det ses for begge forsøg, at resultatet af variansanalysen giver en signifikant forskel på materialer og ligeledes på laboratorier. I forsøg B bliver interaktionerne mellem materiale og laboratorium, materiale og dag, samt dag og laboratorium alle signifikante. Det betyder, at resultaterne samlet for alle materialerne afhænger af laboratoriet. Ved analyse af hvert materiale uafhængig

fra de andre materialer, viser resultaterne for nogle af materialerne, at de ikke påvirkes af det enkle laboratorium. Mens resultaterne for de andre materialer, afhænger af det enkle laboratorium.

De indledningsmæssige analyser er baserede på de rå data, og er efterfølgende blevet undersøgt for outliers ved anvendelse af først Cochran´s test og efterfølgende Grubbs´ test.

Resultaterne for forsøg A, giver ingen outliers ved anvendelse af Cochran´s test. Den efterfølgende analyse med Grubbs´ test viser, at tre ud af de fem materialer har en outlier. Outlierne kommer alle fra laboratorium fire og det betyder, at alle resultaterne for laboratorium fire er fjernet fra datasættet. Det nye datasæt "data uden outlier"indeholder derfor kun resultater for de resterende 13 laboratorier.

Resultaterne for forsøg B, giver ved anvendelse af Cochran´s test nogle få outliers, og den efterfølgende anvendelse af Grubbs´ test giver ikke flere outliers. De outliers, der er fundet for forsøg B, giver ikke anledning til, at fjerne alle resultaterne for et enkelt laboratorium som ved forsøg A.

De nye datasæt for forsøgene er analyseret, herunder er repeterbarheden og reproducerbarheden er blandt andet blevet beregnet. Ved sammenligning med de rå datasæt for hvert af forsøgene, ses der et faldt i værdien for repeterbarheden og reproducerbarheden. Det skyldes, at de fjernede outliers giver et mindre spænd mellem resultaterne indenfor hvert materiale.

Anvendelse af en outlier undersøgelse med efterfølgende fjernelse af resultater, er ikke altid det mest optimale, eftersom alle resultater må antages at være vigtige for den videre analyse. Derfor anvender endvidere en anden metoden, robusthedsanalyse, til undersøgelse af resultaterne. Robusthedsanalysen fjerner ingen resultater, men ændrer nogle af resultaterne til en beregnet grænseværdi. Den robusthedsanalyse der er anvendt i dette projekt, tager udgangspunkt i den ISO standarden selv har beskrevet. Beregningerne i robusthedsanalysen tager udgangspunkt i medianen og ikke gennemsnittet som ved outlier undersøgelserne. Det betyder, at udgangspunktet er lidt anderledes end for anvendelse af gennemsnittet og outlier vægtes ikke på samme måde.

Anvendelsen af robusthedsanalysen for begge forsøg, har medført at flere resultater er blevet ændret end for outlier testen. På baggrund af de ændrede resultater er de nye datasæt blevet analyseret, herunder er repeterbarheden og reproducerbarheden blevet beregnet. De beregnede værdier for repeterbarheden og reproducerbarheden er ved sammenligning med de rådata og data uden outlier blevet mindre. Det betyder, at robusthedsanalysen har givet de bedste resultater for repeterbarheden og reproducerbarheden.

De indledende analyser har taget udgangspunkt i resultaterne for de rå data. Det betyder, at materiale, laboratorium og dag alle har være fixed faktorer. I mixed model analysen opstilles en model til fremtidige analyser af enzymmængden i dyrefoder. I denne model er materiale en fixed faktor, mens laboratorium og

dag er random faktorer. Resultaterne fra mixed model giver usikkerheden for repeterbarheden, som Novozymes kan anvende løbende i deres kvalitetskontrol.

Der er til sidst lavet et program til estimering af dimensioneringen af fremtidige forsøg, ud fra resultaterne for forsøg B. Estimeringen af dimensioneringen er blevet lavet for hvert materiale og alle materialerne tilsammen. Resultatet af dimensioneringen viste igen hvilken forskel der er på resultaterne for hvert af materialerne. Dimensioneringen giver for nogle af materialerne det mindste antal laboratorier og dag og det højeste antal gentagelser. Hvor dimensioneringen for andre materialer gav et højt antal laboratorier og dag og det laveste antal gentagelser. Dimensioneringen set over alle materialerne som et materiale, gav det mindste antal laboratorier og dag med det højeste antal gentagelser.



# Appendiks

---

## A.1 Metode validering

Valideringen af metoden arbejder med princippet at phytase enzymet frigiver foshat fra substratet myo-inositol-hexaphosphate (phytate). Det frivigete uorganiske fosfat er bestemt ved dannelse af en gul masse med syre molybdate/vanadate blanding. Den gul masse er blevet analyserede med et spektrometer ved bølglængden 415 nm, og det frivigete foshat er bestemt ved den *optiske densitet* ( $OD_{415}$ ) og en kvalitativ standard kurve for fosfat. Aktiviteten er bestemt i *phytate enheder* (U/kg) foder analyse. En phytase enhed (U) er mængden af enzym der frigiver 1  $\mu\text{mol}$  af det uorganiske fosfat fra phytatse per minut, under de angivende reaktions bestemmelser. Aktiviteten er beregnet efter følgende formel:

$$\text{Aktivitet} = \frac{\Delta OD \cdot D}{m \cdot W \cdot t}$$

Hvor  $\Delta OD = OD_{415\text{sample}} - OD_{415\text{blank}}$   $m = \text{slope of the standard curve}[OD_{415}/(\mu\text{mol} \cdot \text{ml}^{-1})]$

$D = \text{dilution factor (extraction volume} \cdot \text{dilution of the extract) [ml]}$

$W = \text{weight of the sample [kg]}$

$t = 30 \text{ [min]}$

$OD_{415}$  *sample* er resultatet af målingerne af dyrefoderet efter hele analyse proceduren, hvor  $OD_{415}$  *blank* er resultatet af betemmelsen af det uorganiske phosphat, som der allerede er blevet analyserede i den samme analyse af dyre foder, men før udførelsen af den enzymatiske reaktion.

Ved konstruktionen af standard kurven, reaktioen mellem en serie af phosphat standarder og syre molybdate/vanadte analyse er analyserede i triplicate. Side-løbende kontrolleres phytase niveauet for den validerede phytase standard. Det er lavet for at tjekke om resultaterne og reaktionerne er korrekte. Tilsidst laves endnu en analyse af test analyserne. Det gøres ved doublet udtagelse af hver analyse og de analyseres med reaktionen af syre molybdate/vanadate.

## A.2 ISO 5725

Der vil i dette afsnint blive beskrevet nogle af de termer som ISO 5725 metoden anvender. Beskrivelsen af termerne skal ses som en hjælp til læsere, som ikke er bekendt med ISO 5725.

### Præcision og nøjagtighed

Definitionen af de to begreber:

- Præcision vedrører spredningen af måleresultater
- Nøjagtighed vedrører afstanden til en accepteret referenceværdi (den "sande" værdi)

Kombinationen af præcision og nøjagtighed har fire senarier:

- God præcision og god nøjagtighed: Resultaterne ligger tæt i en klump og klumpen ligger omkring centrum i et korredinatsæt.
- God præcision og dårlig nøjagtighed: Resultaterne ligger tæt i en klump, men klumpen ligger ikke tæt på centrum i korredinatsættet.
- Dårlig præcision og god nøjagtighed: Resultaterne ligger spredt omkring centrum, men danner ingen klump.
- Dårlig præcision og Dårlig nøjagtighed: Resultaterne ligger spredt og danner ingen klump, som ikke ligger tæt på centrum.



**Bestemmelse af nøjagtigheden**

Den "sande værdi" er en uopnåelig størrelse. Den bliver hurtigt påvirket af behandlingen af prøven.

Bestemmelsen kan i praktisk ske f.eks. ved anvendelse af et kendt referance materiale og præstationsprøvning.

**Præstationsprøvninger**

Præstationsprøvninger og ringtest har samme betydning. En præstationsprøvning skal som udgangspunkt analyseres som alle andre prøver. Derfor må laboratoriet ikke tage særlige hensyn til prøven ved eventuel at,

- kalibrere apparaturet en ekstra gang
- sætte en dygtig laborant til analysen
- lave gentagende analyser

**Repetierbarhed**

Definitionen af repeterbarhed er: Den maksimale forskel mellem to målinger ved et 95% signifikansniveau, som er udført under repeterbarhedsbetingelser. Repeterbarhedsbetingelser:

- samme operatør
- samme laboratorium
- samme apparatur
- samme tidsrum
- samme reagenser

**Reproducerbarhed**

Definitionen af reproducerbarhed er: Den maksimale forskel mellem to målinger ved et 95% signifikansniveau, som er udført under reproducerbarhedsbetingelse. Reproducerbarhedsbetingelser:

- forskellige operatør
- forskellige laboratorium
- forskellige apparatur
- forskellige tidsrum

- forskellige reagenser

Høje værdier for reproducerbarheden kan skyldes stor spredning, dårlig præcision blandt de involverede laboratorier.

### Interaktionen mellem repeter- og reproducerbarhed

Repeterbarheden påvirker resultatet af reproducerbarheden, men ikke omvendt. Et forsøg kan derfor ikke have en lav reproducerbarhed og samtidig have en høj repeterbarhed.

Reproducerbarheden er i de fleste tilfælde større end repeterbarheden, hvilket betyder at,  $R > r$ . Hvis det modsatte forekommer, at reproducerbarheden er mindre en repeterbarheden,  $R < r$ , kan det skyldes at prøverne er inhomogene.

### Normalfordeling

Normalfordelingen beskriver en sandsynlighedsfordeling for en kontinuert statistisk variabel. Lad  $X$  være en statistisk variabel med tæthedsfunktionen bestemt ved:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

hvor  $\mu \in \mathbb{R}$  og  $\sigma \in \mathbb{R}_+$ . Fordelingen betegnes  $N(\mu, \sigma)$  og middelværdien,  $E(X)$ , og variansen,  $V(X)$ , er:

$$\begin{aligned} E(X) &= \mu \\ V(X) &= \sigma^2 \end{aligned}$$

## A.3 Statistiske beregninger

Der vil for resultaterne blive beregnet gennemsnit, spredning, repeterbarhed, reproducerbarhed for hvert materiale, og alle materialerne tilsammen for hvert af forsøgene. Alle beregningerne er taget fra ISO 5725<sup>2</sup>.

I afsnittet beskrives hvordan beregningerne af forsøgene er foretaget. Det kræver først en beskrivelse af de anvendte symboler.

Der er i hvert forsøg  $q$  materialer,  $p$  laboratorier med  $n$  gentagelser, og i forsøg B er der  $r$  dage. De  $q$  materialer nummereres med  $j$  ( $j = 1, 2, \dots, q$ ), hvert af de  $p$  laboratorier nummereres med  $i$  ( $i = 1, 2, \dots, p$ ), med de  $n$  gentagelser for hvert niveau, med kombinationen  $ij$ . De  $r$  dage i forsøg B nummereres med ( $l = 1, 2, \dots, r$ ).

- $n_{ij}$  antallet af resultater i en celle for laboratorium  $i$  og materiale  $j$   
 $y_{ijk}$  er et af resultaterne i en celle for laboratorium  $i$  og materiale  $j$  ( $k = 1, 2, \dots, n_{ij}$ )  
 $p_j$  antallet af laboratorier, som har rapporteret mindst et resultat for materiale  $j$

Det totale antal resultater i et forsøg vil være  $pqn$ . Det gælder for forsøg A at  $p = 14$ ,  $q = 5$  og  $n = 2$ , som betyder at det totale antal resultater er 140. Det samme kan beregnes for forsøg B,  $p = 14$ ,  $q = 8$  og  $n = 6$ , som betyder at det totale antal resultater er 672.

### Gennemsnit for materiale $j$

$$\hat{m}_j = \bar{y}_j = \frac{\sum_{i=1}^p n_{ij} \bar{y}_{ij}}{\sum_{i=1}^p n_{ij}}$$

hvor  $\bar{y}_{ij}$  er celle gennemsnittet <sup>2</sup>

$$\bar{y}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk}$$

### Spredningen for materiale $j$ og laboratorium $i$

$$s_{ij} = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij})^2}$$

### Spredningen for materiale $j$

$$s_j = \sqrt{\frac{1}{(p \cdot n) - 1} \sum_{i=1}^p (y_{ijk} - \hat{m}_j)^2}$$

### Variansen for materiale $j$

$$s_j^2 = \frac{1}{(p \cdot n) - 1} \sum_{i=1}^p (y_{ijk} - \hat{m}_j)^2$$

Der beregnes tre standardafvigelser for hvert niveau. Det er standardafvigelsen svarende til repeterbarheden, standardafvigelsen svarende til mellem-laboratorier og standardafvigelsen svarende til reproducerbarheden.

### Standardafvigelsen for repeterbarheden for materiale $j$

Standardafvigelsen for repeterbarheden er det beregnede gennemsnit for standardafvigelsen for indenfor-laboratorier,  $s_w$ . Det beregnede gennemsnit for et materiale, gælder for alle de laboratorier som har været med i forsøget. Antallet af laboratorier kan variere som følge af fjernelse af laboratorie resultater der forventes at være outlierer.

$$s_{rj} = \sqrt{\frac{\sum_{i=1}^p (n_{ij} - 1) s_{ij}^2}{\sum_{i=1}^p (n_{ij} - 1)}}$$

hvor  $s_{ij}^2$  er variansen for hvert celle, som indeholder resultater for et materiale og et laboratorium.

### Standardafvigelsen for mellem-laboratorier for materiale $j$

Standardafvigelsen for mellem-laboratorier er variansen for alle laboratorier minus variansen for repeterbarheden og divideret med gennemsnittet af resultater fra en celle. I begge forsøg er antallet af resultater for en celle, det samme for alle cellerne. Det betyder at  $n_{ij}$  vil være to i forsøg A og seks i forsøg B.

$$s_{Lj} = \sqrt{\frac{s_{dj}^2 - s_{rj}^2}{\bar{n}_j}}$$

hvor variansen  $s_{dj}^2$  og  $\bar{n}_j$  beskrives ved

$$s_{dj}^2 = \frac{1}{p-1} \sum_{i=1}^p n_{ij} (\bar{y}_{ij} - \bar{y}_i)^2$$

$$\bar{n}_j = \frac{1}{p-1} \left[ \sum_{i=1}^p n_{ij} - \frac{\sum_{i=1}^p n_{ij}^2}{\sum_{i=1}^p n_{ij}} \right]$$

### Standardafvigelsen for reproducerbarheden for materiale $j$

Standardafvigelsen for reproducerbarheden er kvadratroden af summen af variansen for mellem-laboratorier og variansen for repeterbarheden.

$$s_{Rj} = \sqrt{s_{rj}^2 + s_{Lj}^2}$$

### 95%-konfidensintervallet

Det er antaget at resultaterne følger en normalfordeling med udgangspunkt i resultaternes gennemsnit og standardafvigelse.

Ved rapportering af statistiske analyser anfører man estimerne for parametrene i modellen eller efter den seneste accepterede hypotese sammen med estimer for deres standardafvigelser. En anden måde at give et numerisk udtryk for den usikkerhed, hvormed et parameterestimat er behæftet, er ved hjælp af et konfidensinterval for den pågældende parameter.

Konfidensintervallet kan defineres som et interval omkring parameterestimatet, hvor den sande parameter for en stikprøve formodes at befinde sig.

Et konfidensinterval er opbygget som et interval, konstrueret ud fra observationerne, som indeholder de værdier af parameteren, som er overensstemmende med data på et givet signifikansniveau (testsandsynlighed).

Til beregning af konfidensintervaller anvendes en  $z$ -fordeling, ved forsøg med mange observationer, og  $t$ -fordelingen normalt for forsøg med få observationer.

Den lille mængde af observationer betyder, at den beregnede standardafvigelsen,  $s$ , bliver mindre pålidelig som estimat for standardafvigelsen  $\sigma$ .

Forskellen på anvendelse af z-fordelingen i forhold til anvendelse af t-fordelingen i dette projekt, kan dog først ses på fjerde betydende ciffer. Derfor anvendes z-fordelingen til alle beregninger af 95% konfidensintervallet for middelværdien.

$$[\bar{x} - 1.96 \cdot (\frac{\sigma}{\sqrt{n}}); \bar{x} + 1.96 \cdot (\frac{\sigma}{\sqrt{n}})]$$

hvor  $\bar{x}$  er gennemsnittet,  $\sigma$  er spredningen og  $n$  er antallet af resultater.

## A.4 Varianshomogenitettest

Varianshomogenitettestet undersøger om en variabel har samme varians for forskellige grupper af observationer. Den kaldes også for Bartletts test og er et approksimativt test. Bartletts test forudsætter, at der foreligger normalfordelte variable. Bartletts test er følsom over for afvigelser fra normalitet, og bør kun benyttes, når forudsætningen om normalitet med rimelighed kan antages opfyldt. Med Bartletts test undersøges det hvorvidt variansen for resultaterne fra  $n$  involverede parter (fx. laboratorier, laboranter, metoder eller materialer) kan antages at være ens.

Den efterfølgende variansanalyse forudsætter varianshomogenitet dvs., at residualerne kan antages at have samme varians. Variansanalysen er robust over for afvigelser fra varianshomogenitet, såfremt der foretages det samme antal observationer i hver celle. Bartletts test har derfor særlig interesse, såfremt der af en eller anden grund ikke er det samme antal observationer i hver celle.

Ved udførelse af Bartletts test betragtes hypotesen

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2,$$

og der anvendes estimaterne

$$s_1^2, s_2^2, \dots, s_n^2$$

Det antages at 1,2,...,n parter hver især har udført et antal bestemmelser ( $n_1, n_2, \dots, n_n$ ). Antallet af bestemmelser behøver ikke at være ens.

### Bestemmelse af pooleet varians <sup>5</sup>

Den poolede standardafvigelse beregnes på baggrund af

$$s_{pooleet}^2 = \frac{\sum f_i \cdot s_i^2}{\sum f_i} = \frac{f_1 \cdot s_1^2 + f_2 \cdot s_2^2 + \dots + f_n \cdot s_n^2}{f_1 + f_2 + \dots + f_n}$$

hvor  $f_1 = (n_1 - 1), f_2 = (n_2 - 1), \dots, f_n = (n_n - 1)$ .

### Bestemmelse af testværdi <sup>5</sup>

Testværdien beregnes på baggrund af

$$Ba = \frac{\ln(s_{pooled}^2) \cdot \sum f_i - \sum (f_i \cdot \ln(s_i^2))}{1 + \frac{1}{3 \cdot (m-1)} \cdot (\sum \frac{1}{f_i} - \frac{1}{\sum f_i})}$$

hvor n er antallet af parter.

### Udfald

Hvis  $Ba > x_p^2$  forkastes  $H_0$  som betyder, at der ikke er varianshomogenitet og  $H_a$  accepteres som betyder, at varianserne ikke er ens blandt de involverede parter.

BILAG B

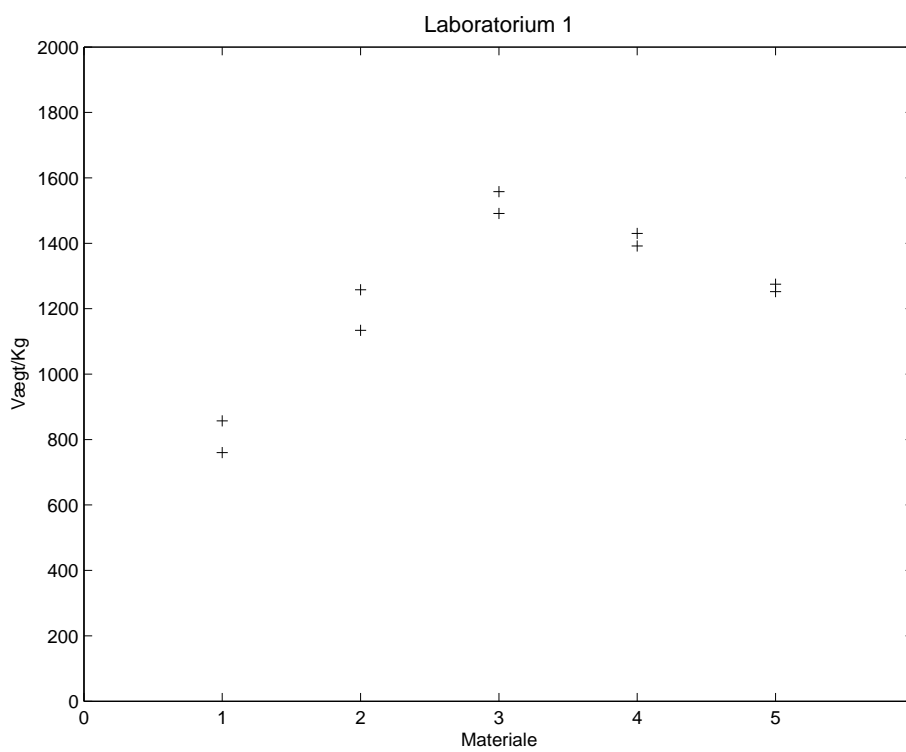
Appendiks

---

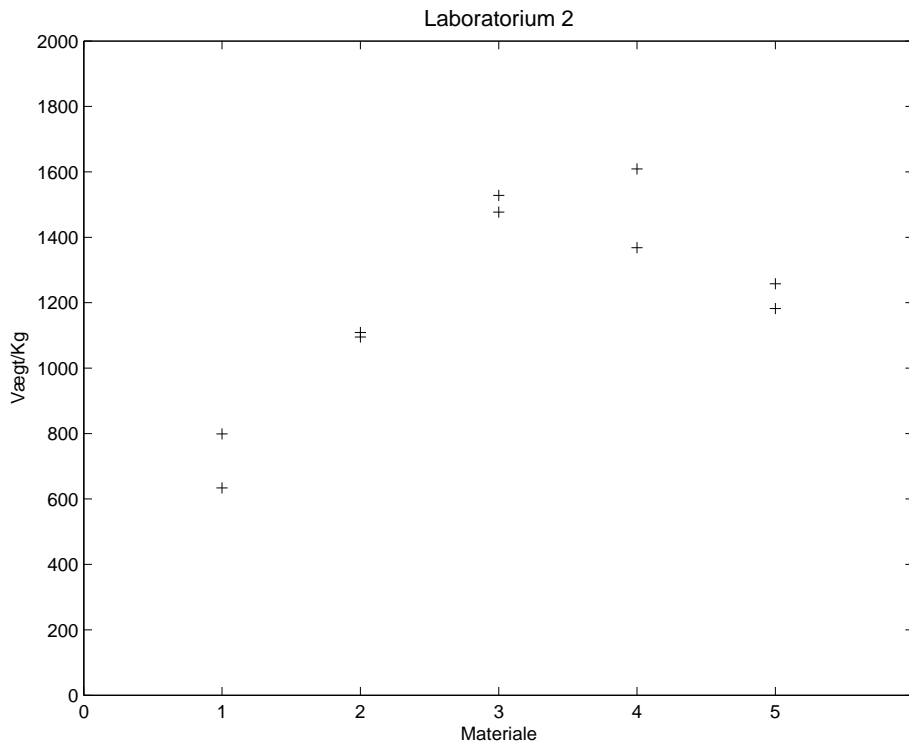
## B.1 Liste af anvendte laboratorier

Firma	Adresse	Ekspert
DSM Nutritional Products	Basel, Switzerland	Stefanie Kirchen
Novozymes	Assay and Technology Development, Bagsværd, Denmark	Betina B. Andersen
BASF	Ludwigshafen, Germany	Roland Betz
Alltech	Meath, Irland	Shauna McKelvey
Danisco	Brabrand, Denmark	Faurshou Isaksen
Forschungsanstalt für Nutztiere und Milchwirtschaft	Posieux, Switzerland	Dominique Heimo
The Danish Plant Directorate	Lyngby, Denmark	Annette Plöger
Laboratoire DGCCRF de Rennes	Rennes, France	Roger Ziebal
Institut Europeen de l'environnement de Bordeaux	Bordeaux	Olivier Locqueneux
LUFA Freising	Freising, Germany	Jürgen Danier
LUFA Leipzig	Leipzig, Germany	Jens Schönherr
Agentur für Gesundheit und Ernährungssicherheit Institut für Futtermittel	Wien, Austria	Irmengard Strnad
OMMI Central Laboratory of National Institute for Agricultural Quality Control	Budapest, Hungary	Lajos Debreczeni
Canadian Food Inspection Agency	Ottawa Ontario, Canada	Etsuko Yamamoto

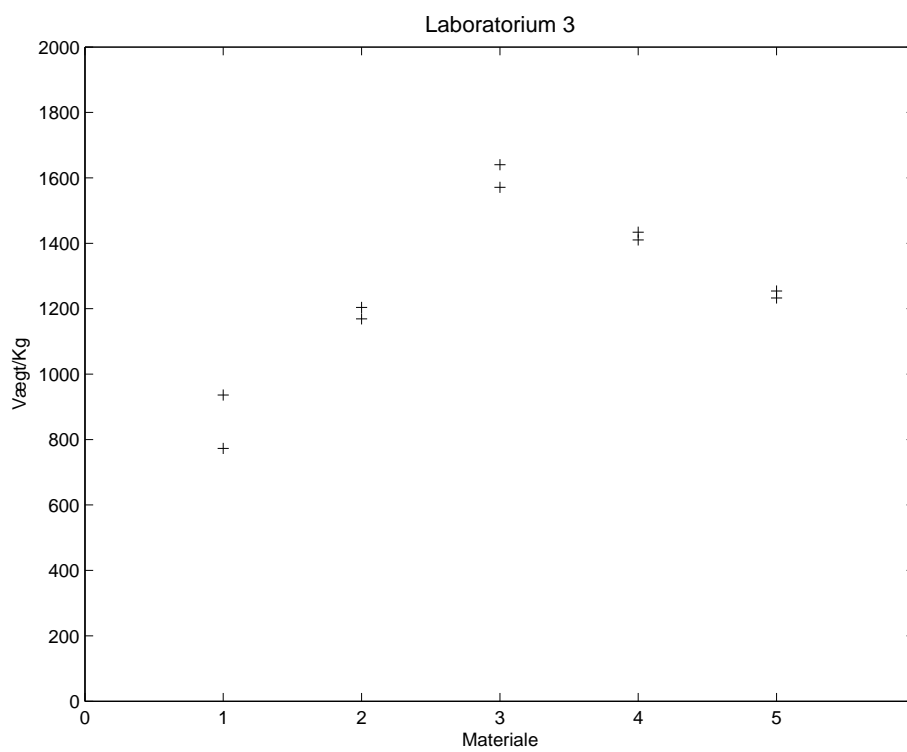




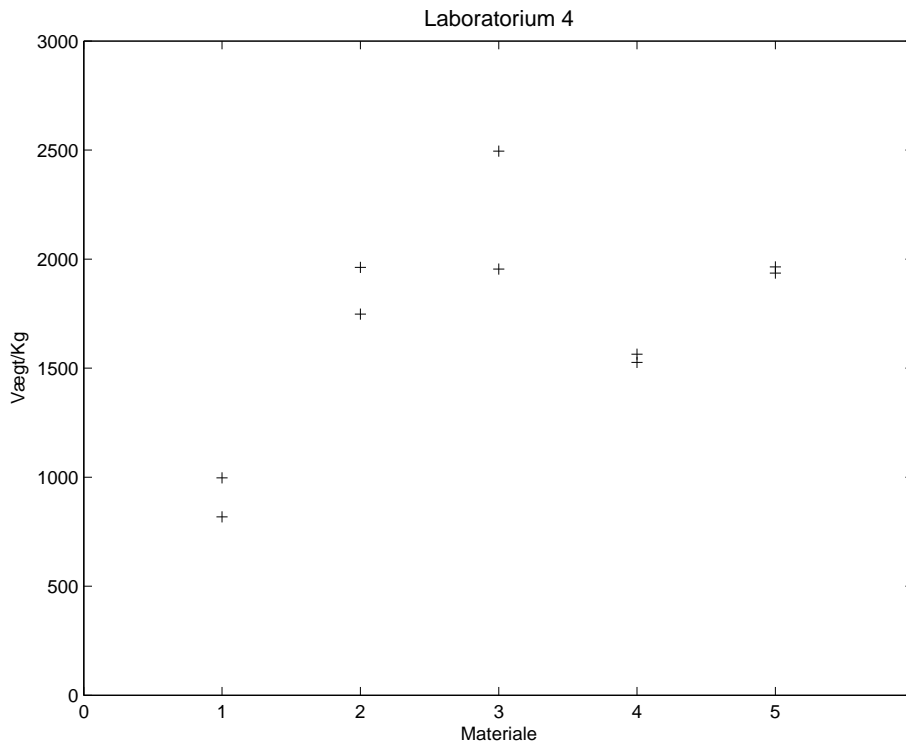
Figur B.1:



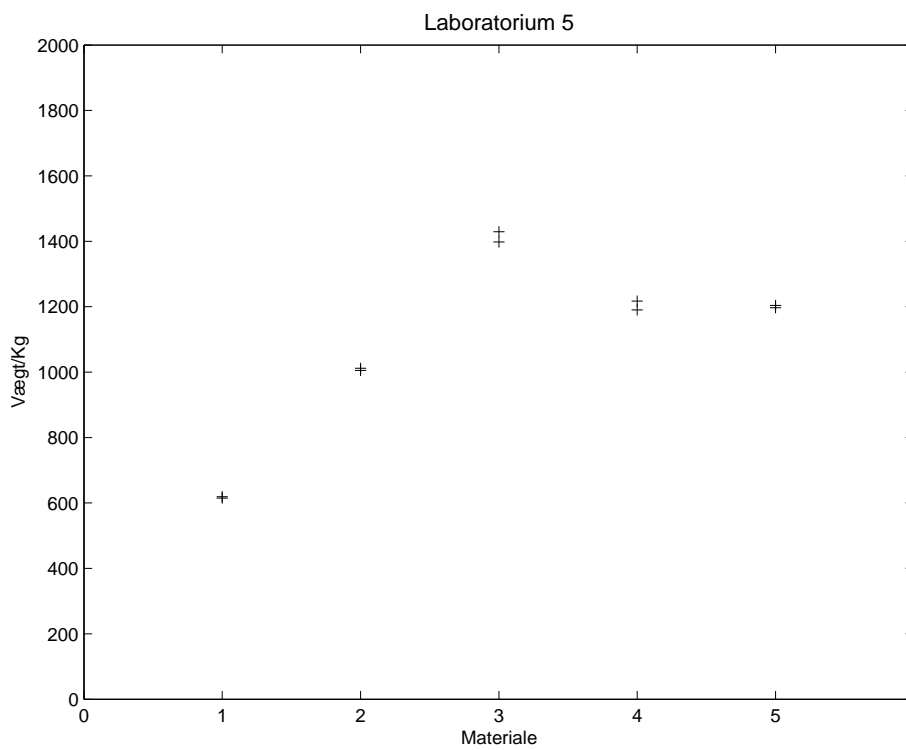
Figur B.2:



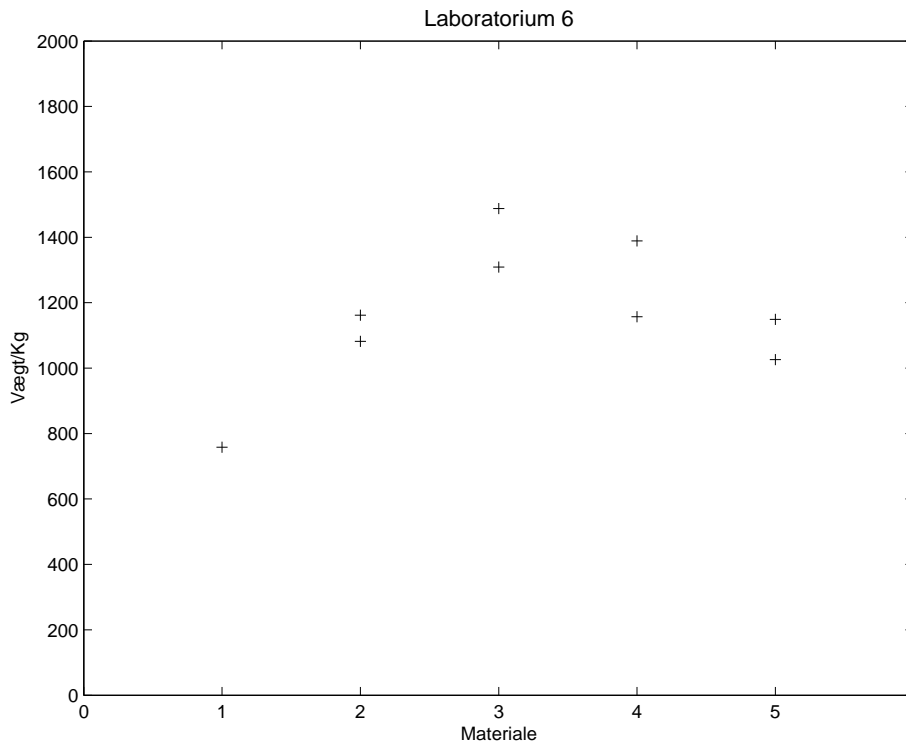
Figur B.3:



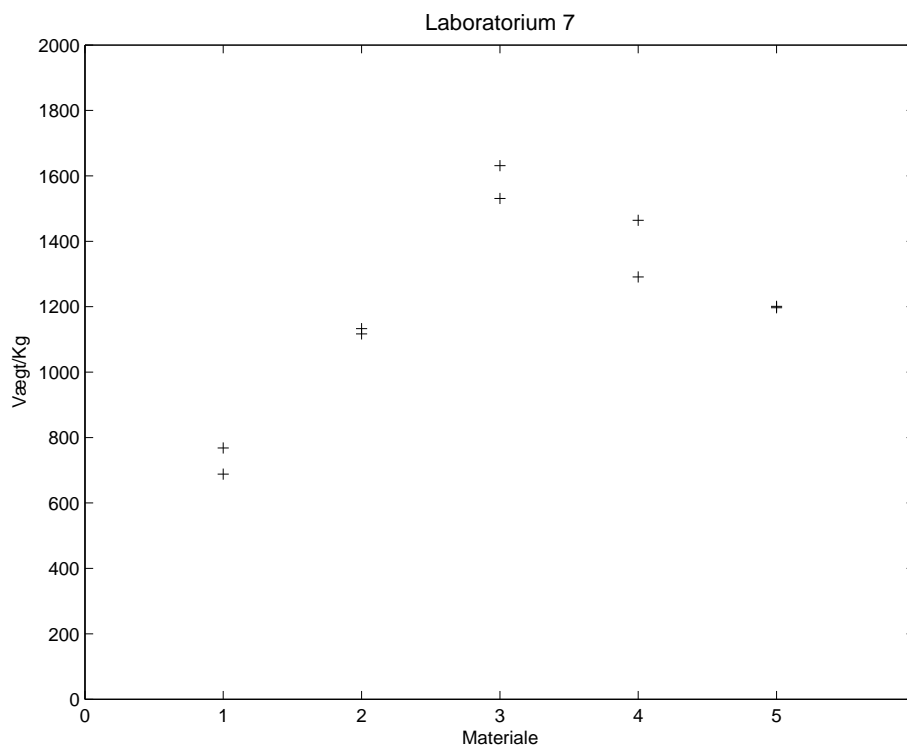
Figur B.4:



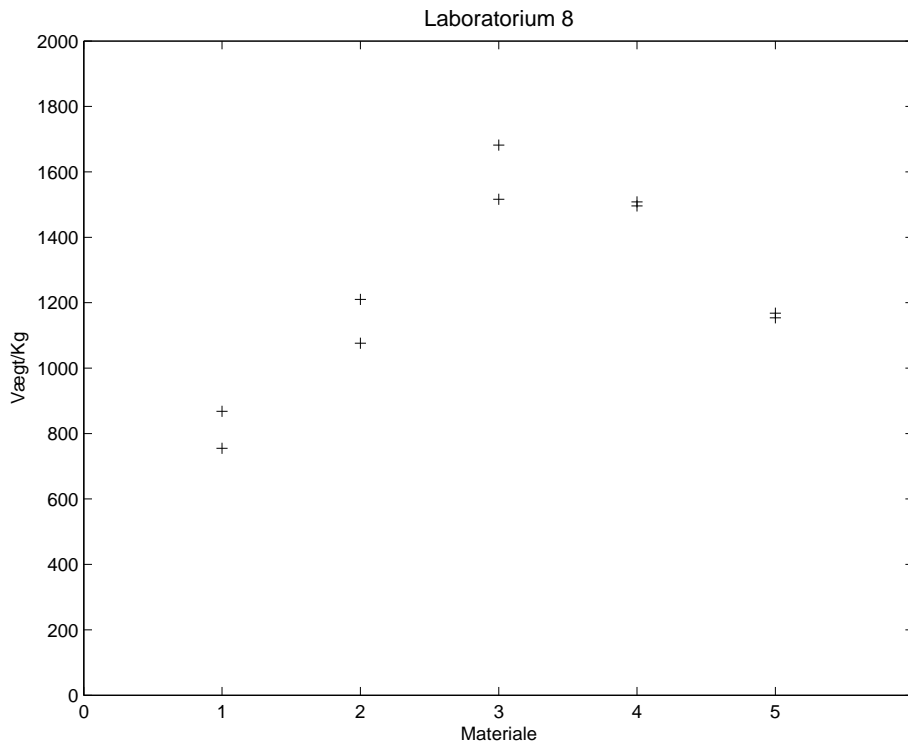
Figur B.5:



Figur B.6:

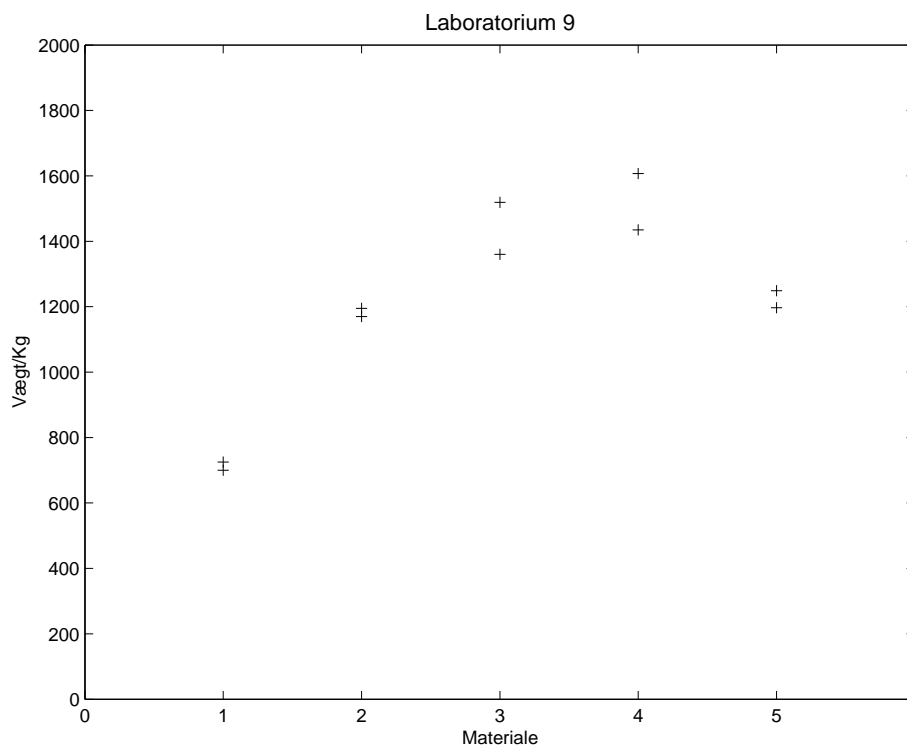


Figur B.7:

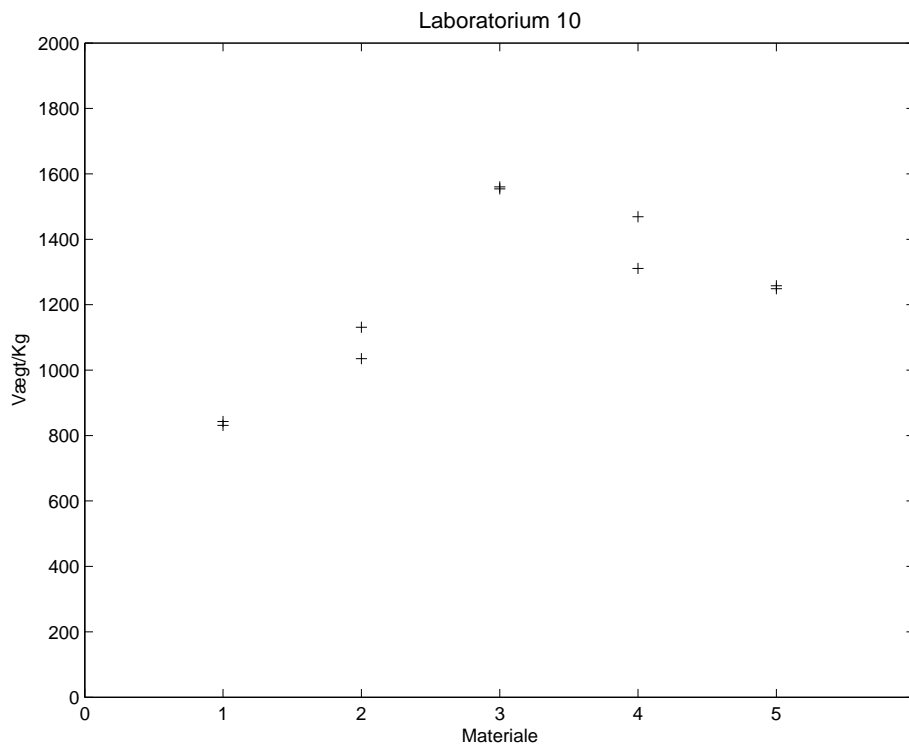


Figur B.8:

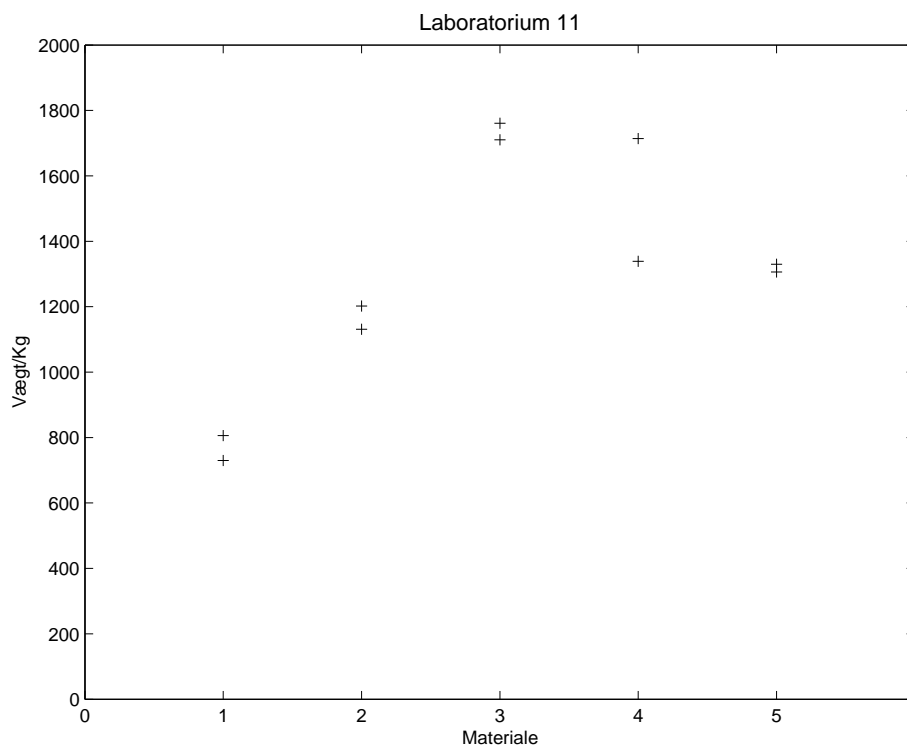




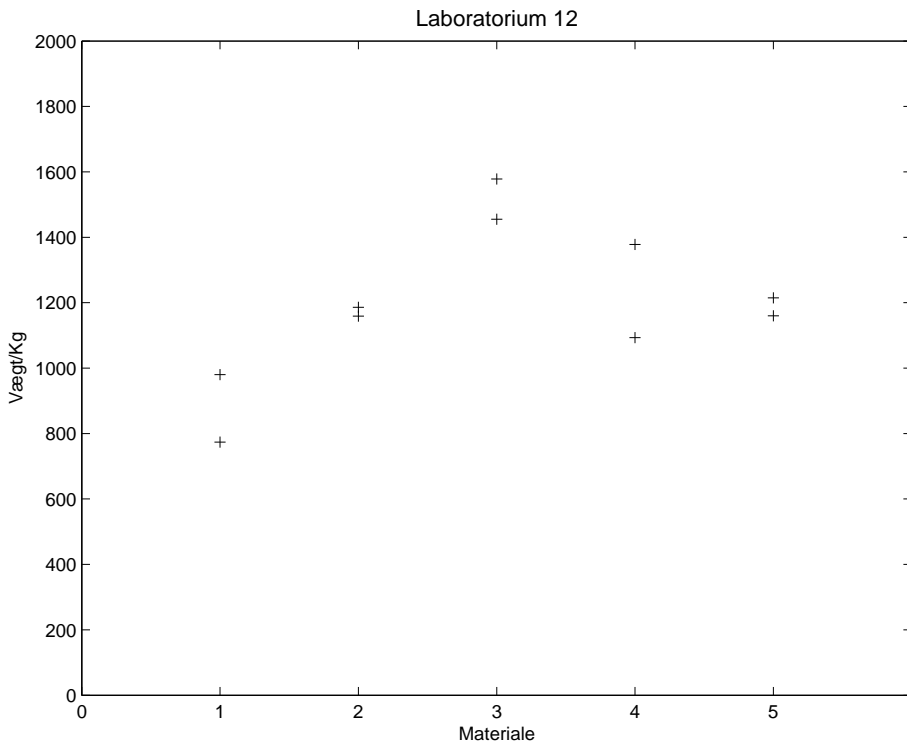
Figur B.9:



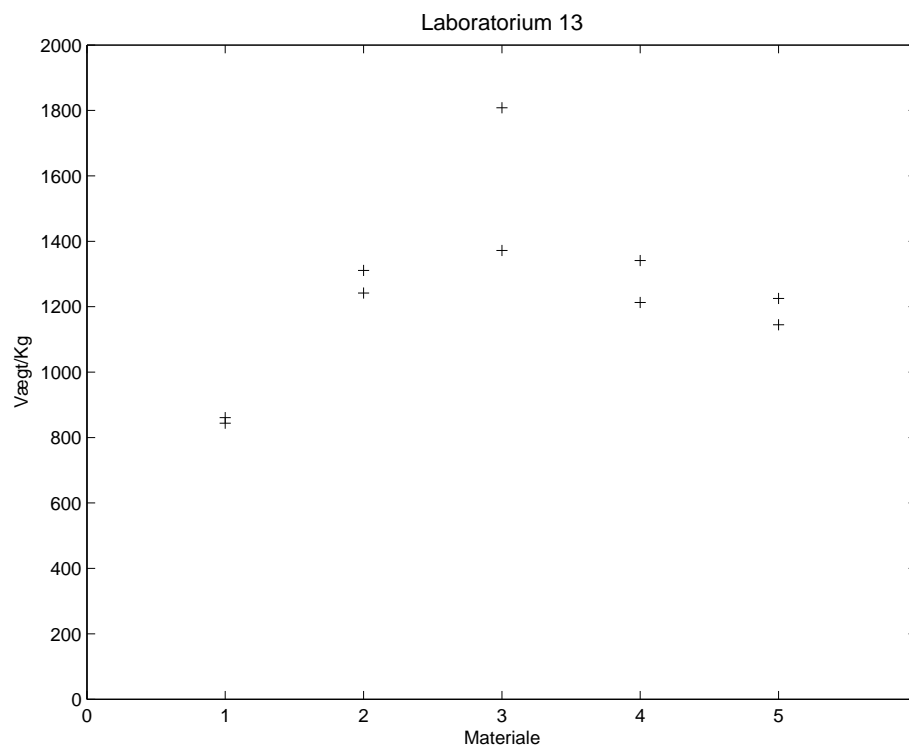
Figur B.10:



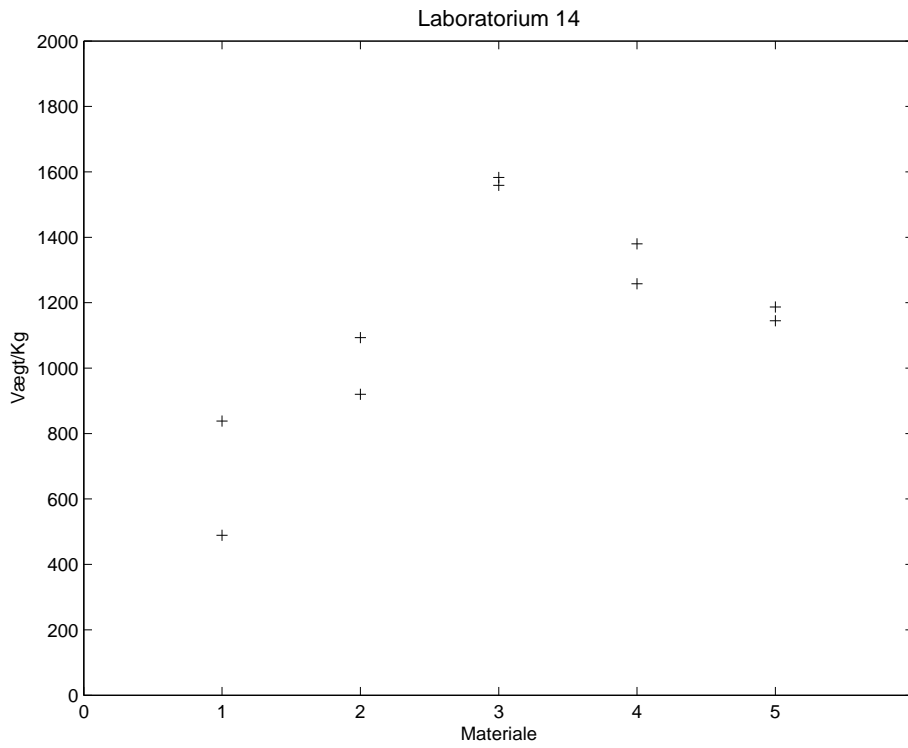
Figur B.11:



Figur B.12:



Figur B.13:



Figur B.14:

## B.2 Resultater for forsøg A

B.2.1 Tabel med rådata

B.2.2 Resultater vist grafisk

B.2.3 Beregnede resultater

## B.3 Resultater for forsøg B

B.3.1 Tabel med rådata

B.3.2 Resultater vist grafisk

B.3.3 Beregnede resultater

	MATERIAL	LABCODE	Result	Resultlog	Resultrobust
1	1	1	760.00	6.6333	6.6333
2	1	1	857.00	6.7534	6.7534
3	1	2	634.00	6.4520	6.4593
4	1	2	799.00	6.6834	6.6834
5	1	3	773.00	6.6503	6.6503
6	1	3	936.00	6.8416	6.8416
7	1	4	818.00	6.7069	6.7069
8	1	4	997.00	6.9048	6.8557
9	1	5	615.00	6.4216	6.4593
10	1	5	619.00	6.4281	6.4593
11	1	6	758.00	6.6307	6.6307
12	1	6	758.00	6.6307	6.6307
13	1	7	688.00	6.5338	6.5338
14	1	7	768.00	6.6438	6.6438
15	1	8	755.00	6.6267	6.6267
16	1	8	868.00	6.7662	6.7662
17	1	9	700.00	6.5511	6.5511
18	1	9	725.00	6.5862	6.5862
19	1	10	831.00	6.7226	6.7226
20	1	10	843.00	6.7370	6.7370
21	1	11	730.00	6.5930	6.5930
22	1	11	806.00	6.6921	6.6921
23	1	12	774.00	6.6516	6.6516
24	1	12	980.00	6.8876	6.8557
25	1	13	844.00	6.7382	6.7382
26	1	13	861.00	6.7581	6.7581
27	1	14	489.00	6.1924	6.4593
28	1	14	838.00	6.7310	6.7310
29	2	1	1134.00	7.0335	7.0335
30	2	1	1258.00	7.1373	7.1373
31	2	2	1095.00	6.9985	6.9985
32	2	2	1109.00	7.0112	7.0112
33	2	3	1169.00	7.0639	7.0639
34	2	3	1204.00	7.0934	7.0934
35	2	4	1748.00	7.4662	7.1699
36	2	4	1962.00	7.5817	7.1699
37	2	5	1005.00	6.9127	6.9222
38	2	5	1012.00	6.9197	6.9222
39	2	6	1082.00	6.9866	6.9866
40	2	6	1162.00	7.0579	7.0579
41	2	7	1117.00	7.0184	7.0184
42	2	7	1133.00	7.0326	7.0326
43	2	8	1076.00	6.9810	6.9810
44	2	8	1210.00	7.0984	7.0984
45	2	9	1170.00	7.0648	7.0648
46	2	9	1195.00	7.0859	7.0859
47	2	10	1035.00	6.9422	6.9422
48	2	10	1131.00	7.0309	7.0309
49	2	11	1131.00	7.0309	7.0309



Materiale nr.	Gennemsnit af celle værdi			Celle værdi			Lab. nr.
	Lab. nr.	Gennemsnit	Ny værdi	Lab. nr.	Celle værdi	Ny værdi	
1		617	643.177	4	997	936.6305	2
				5	615	625.8576	4
				5	619	625.8576	5
				12	980	936.6305	5
				14	489	625.8576	12
							14
2		1855	1287.767	4	1748	1290.972	4
				4	1962	1290.972	4
				5	1005	1008.692	5
				13	1311	1290.972	5
				14	920	1008.692	13
							14
3		2224.5	1717.349	4	2495	1769.408	4
		1735.5	1717.349	4	1954	1769.408	4
				6	1309	1345.193	6
				13	1808	1769.408	13
4				6	1157	1164.671	6
				11	1714	1618.689	11
				12	1093	1164.671	12
5		1950	1309.483	4	1964	1314.243	4
		1318	1309.483	4	1936	1314.243	4
				6	1026	1123.333	6
				11	1330	1314.243	11

	Site	Day	Samplenummer	Weighing	Units	Log..U.
1	1.00	1.00	1.00	1.00	1600.00	7.38
2	1.00	1.00	1.00	2.00	1644.00	7.40
3	1.00	1.00	2.00	1.00	1362.00	7.22
4	1.00	1.00	2.00	2.00	1480.00	7.30
5	1.00	1.00	3.00	1.00	1918.00	7.56
6	1.00	1.00	3.00	2.00	1649.00	7.41
7	1.00	1.00	4.00	1.00	888.00	6.79
8	1.00	1.00	4.00	2.00	756.00	6.63
9	1.00	1.00	5.00	1.00	669.00	6.51
10	1.00	1.00	5.00	2.00	687.00	6.53
11	1.00	1.00	6.00	1.00	1230.00	7.11
12	1.00	1.00	6.00	2.00	1227.00	7.11
13	1.00	1.00	7.00	1.00	553.00	6.32
14	1.00	1.00	7.00	2.00	483.00	6.18
15	1.00	1.00	8.00	1.00	1615.00	7.39
16	1.00	1.00	8.00	2.00	1461.00	7.29
17	1.00	2.00	1.00	1.00	1524.00	7.33
18	1.00	2.00	1.00	2.00	1376.00	7.23
19	1.00	2.00	2.00	1.00	1503.00	7.32
20	1.00	2.00	2.00	2.00	1237.00	7.12
21	1.00	2.00	3.00	1.00	1527.00	7.33
22	1.00	2.00	3.00	2.00	1468.00	7.29
23	1.00	2.00	4.00	1.00	832.00	6.72
24	1.00	2.00	4.00	2.00	773.00	6.65
25	1.00	2.00	5.00	1.00	700.00	6.55
26	1.00	2.00	5.00	2.00	686.00	6.53
27	1.00	2.00	6.00	1.00	1253.00	7.13
28	1.00	2.00	6.00	2.00	1244.00	7.13
29	1.00	2.00	7.00	1.00	535.00	6.28
30	1.00	2.00	7.00	2.00	493.00	6.20
31	1.00	2.00	8.00	1.00	1325.00	7.19
32	1.00	2.00	8.00	2.00	1458.00	7.28
33	1.00	3.00	1.00	1.00	1688.00	7.43
34	1.00	3.00	1.00	2.00	1673.00	7.42
35	1.00	3.00	2.00	1.00	1397.00	7.24
36	1.00	3.00	2.00	2.00	1461.00	7.29
37	1.00	3.00	3.00	1.00	1711.00	7.44
38	1.00	3.00	3.00	2.00	1366.00	7.22
39	1.00	3.00	4.00	1.00	790.00	6.67
40	1.00	3.00	4.00	2.00	739.00	6.61
41	1.00	3.00	5.00	1.00	656.00	6.49
42	1.00	3.00	5.00	2.00	717.00	6.58
43	1.00	3.00	6.00	1.00	1249.00	7.13
44	1.00	3.00	6.00	2.00	1224.00	7.11
45	1.00	3.00	7.00	1.00	587.00	6.38
46	1.00	3.00	7.00	2.00	557.00	6.32
47	1.00	3.00	8.00	1.00	1508.00	7.32
48	1.00	3.00	8.00	2.00	1542.00	7.34
49	2.00	1.00	1.00	1.00	1611.00	7.38

Materiale nr.	Celle værdi			Log celle værdi			
	Lab. nr.	Celle værdi	Ny værdi	Lab. nr.	Celle værdi	Ny værdi	
1		1916	1764.609		7.557995	7.487636	
		1904	1764.609		7.551712	7.487636	
		1214	1219.948		7.115583	7.120422	
		1066	1219.948		7.101673	7.120422	
		1208	1219.948		6.971667	7.120422	
		1209	1219.948		7.096721	7.120422	
		1167	1219.948		7.097549	7.120422	
		1887	1764.609		7.113956	7.120422	
		1847	1764.609		7.062192	7.120422	
		1947	1764.609		7.541683	7.487636	
					7.521318	7.487636	
					7.574045	7.487636	
	2		1480	1449.906		7.299797	7.293499
			1503	1449.906		7.315218	7.293499
		1461	1449.906		7.307873	7.293499	
		1457	1449.906		7.333676	7.293499	
		1492	1449.906		6.878326	6.898912	
		1531	1449.906		6.763885	6.898912	
		971	976.3		7.382124	7.293499	
		866	976.3		7.402452	7.293499	
		1607	1449.906		7.329094	7.293499	
		1455	1449.906				
		1640	1449.906				
		1524	1449.906				
3			1918	1890.965		7.559038	7.555738
			1291	1335.577		7.163132	7.210276
		1969	1890.965		7.585281	7.555738	
		1246	1335.577		7.127694	7.210276	
		1904	1890.965		7.126891	7.210276	
		1245	1335.577		7.189168	7.210276	
		1325	1335.577		7.153052	7.210276	
		1278	1335.377		7.186901	7.210276	
		1322	1335.577		7.140453	7.210276	
		1262	1335.577		7.713338	7.555738	
		2238	1890.965		7.707512	7.555738	
		2225	1890.965		7.709308	7.555738	
		2229	1890.965		7.785305	7.555738	
		2405	1890.965		7.620215	7.555738	
4		2039	1890.965				
		1064	954.6622		6.698268	6.879578	
		924	954.6622		6.890609	6.879578	
		1022	954.6622		6.410175	6.428244	
		971	954.6622		6.415097	6.428244	
		1027	954.6622		6.929517	6.879578	
		1137	954.6622		6.934397	6.879578	
		1046	954.6622		7.036148	6.879578	
		442	606.5706		6.952729	6.879578	
		582	606.5706		6.091310	6.428244	



# Litteratur

---

1. Gisele Gixxi and Christoph von Holst Validation study on a new method for the determination of phytase activity in feed: Results from an interlaboratory study conducted according to the IUPAC harmonised protocol. 8. december 2005
2. ISO 5725
3. Robust Statistics - How Not to Reject Outliers, Analyst, December 1989, VOL. 114, Royal Society of Chemistry
4. Statistics and Chemometrics for Analytical Chemistry, Miller & Miller
5. Statistik I og II, L. Brøndum og J.D. Monrad, 4. udgave, Den private ingeniørfond, 1989.
6. amc technical brief, Analytical Methods Committee, No.6 Apr 2001, Royal Society of Chemistry 2001, Robust statistics: a method of coping with outliers.
7. Statistisk forsøgsplanlægning I og II, L. Brøndum og J.D. Monrad, 3. udgave, Den private ingeniørfond, 1991.
8. All of Statistics, A Concise Course in Statistical Inference. Larry Wasserman, Springer, 2005.
9. Notater til Mixed model.
10. Fundamental Concepts in the Design of Experiments, fifth edition, Charles R. Hicks og Kenneth V. Turner.