

Nio tumregler och två kungsvägar för att behärska mätosäkerhet

Thomas Svensson
Fraunhofer-Chalmers Centrum för Industrimatematik



Fraunhofer **CHALMERS**
Research Centre
Industrial Mathematics

© Fraunhofer-Chalmers Research Centre for Industrial Mathematics 2005

All rights reserved. No part of this publication may be reproduced or transmitted to any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from FCC.

FCC reports are distributed by

FCC
Chalmers Science Park
SE-412 88 Göteborg
Sweden

Telephone:+46 (0)31 7724000
Telefax:+46 (0)31 827035
E-mail:info@fcc.chalmers.se
Internet:www.fcc.chalmers.se

Innehåll

1	Inledning	1
1.1	Bantning	1
1.2	Olika ändamål för mätningar	3
1.3	Mätfel	5
1.4	Hur används mätosäkerheten?	9
1.5	Systematiska och slumpmässiga fel	9
1.5.1	Olika syften och horisonter för mätosäkerhetsangivelse	11
1.6	Sammanfattning	14
2	Statistiska grunder	15
2.1	Matematiska formler	15
2.2	Stokastisk variabel	17
2.2.1	Räkneregler för väntevärden och varianser	24
2.3	Jämförelse av mätvärden	25
2.3.1	Variansen är lika och känd	25
2.3.2	Variansen är lika men okänd	29
2.3.3	Varianserna är olika och okända	30
2.3.4	Endast konfidensintervallen är kända	31
2.4	Sammanvägd standardavvikelse	34
2.5	Linjär regression	35
2.5.1	Konfidensintervall för linjen	36
2.5.2	Prediktionsintervall för linjen	37
2.5.3	Approximativt prediktionsintervall	38
2.5.4	Bantningsexemplet	38
2.6	Sammanfattning	42
3	Bestämning av mätosäkerhet	45
3.1	Relativa osäkerheter	47
3.2	Approximationsnivå	47
3.2.1	Täckningsfaktor	49
3.2.2	Relevanta osäkerhetskomponenter	50
3.2.3	Systematiska fel	51
3.2.4	Avrundning	52

3.3	Bestäm ändamålet med mätningen	53
3.4	Bestäm alla källor till mätfel	54
3.5	Observationsmetoden	56
3.5.1	Referens	56
3.5.2	Statistiskt försök	57
3.5.3	Bantning	57
3.5.4	Variansanalys	61
3.6	Analysmetoden	68
3.6.1	Modell för mätsituationen	69
3.6.2	Arbetsgång	69
3.6.3	Bestäm standardosäkerheterna	71
3.6.4	Bestäm känslighetskoefficienter	73
3.6.5	Sammanvägning av relevanta osäkerhetskomponenter .	78
3.7	Förbättrad noggrannhet	78
3.7.1	Begränsa populationer	78
3.7.2	Medelvärdesbildning	79
3.8	Sammanfattning	81
4	Kalibrering	85
4.1	Spårbar kalibrering	85
4.1.1	Kalibrering med linjär regression	88
4.1.2	Shuntkalibrering	93
4.1.3	Mätosäkerhet	94
4.2	Sammanfattning	97
5	Speciella problem	101
5.1	Felets väntevärde är skilt från noll	101
5.2	Kvalitetsstyrning	103
5.2.1	Mätsystemsanalys (MSA)	104
5.2.2	Toleranser	105
5.3	Korrelation	108
5.3.1	Behandling av korrelerade komponenter	110
5.3.2	Korrelation på grund av gemensam normal	110
5.4	Mätosäkerhet efter sortering	112
5.4.1	Problemställning	112
5.4.2	Modell	112
5.4.3	Sammanfattning	114
6	Exempel	115
6.1	Tippningsprovning av skrivbordsstol	115
6.1.1	Observationsmetoden	115
6.1.2	Analysmetoden	117
6.2	Skjutmått	120
6.3	Kraftkalibrering	123

INNEHÅLL	7
6.3.1 Inversproblemet	127
6.3.2 Sammanfattning	128
Svar till övningsuppgifterna	129

Förord

Denna handbok är avsedd som en instruktion för bestämning av mätosäkerheter i industriella tillämpningar. Den förenar de två angreppssätten: “Top down” representerad av kvalitetssystem för laboratorier och “Bottom up” representerad av kvalitetssystem för produktion. Detta görs genom: 1. Mätosäkerhetsberäkningar görs utgående från aktuellt behov och kan därmed ge olika resultat beroende på den tillämpning man avser. 2. Statistiska experiment sätts i förgrunden för bestämning av mätosäkerheter, medan bedömningar undviks i görligaste mån. 3. Metodiken syftar till att förbättra mätosäkerheten, dvs. finna de dominerande källorna till mätosäkerheten och ange metoder för att minska dessa. Handboken innefattar olika nivåer avseende svårighetsgrad; den enklaste nivån ger en receptinriktad handfast metodik, baserad på ett antal tumregler och standardmetoder och kan användas direkt av praktiserande mättingenjörer med teknisk gymnasiekompetens, den dominerande nivån motiverar regelboken, ger en övergripande förklaring av mätosäkerhetsbegreppet ur ett statistiskt perspektiv och kräver en matematisk vana motsvarande den som ges i civilingenjörsutbildningens grundkurser, en tredje högre nivå kan slutligen urskiljas i några tillämpningsexempel som kräver en större matematisk utbildning och vana för en fullständig förståelse.

Utarbetandet av denna handbok har utförts inom ett projekt finansierat av STM, Stiftelsen för Tillämpad Matematik. Styrgruppen för projektet har varit till stor hjälp vid skrivandet, Jacques de Maré, Matematisk Statistik, Chalmers, Hans Andersson, SP, Sveriges Provnings- och Forskningsinstitut, Martin Svennungsson, Volvo 3P, samt Anders Göthberg, SKF. Dessutom har ytterligare personer varit behjälpliga i utarbetandet, varav speciellt tack riktas till Sara Lorén, FCC och Bertil Magnusson, SP.

Kapitel 1

Inledning

1.1 Inledande exempel: bantning

En person betraktar sig som för tung och gör följande åtgärder för att minska i vikt: cyklar till jobbet istället för att ta bilen, samt byter ut lättöl till lunchen mot mineralvatten. Han kontrollerar därefter viktutvecklingen genom att väga sig på en badrumsvåg efter uppstigandet varje morgon. Målet med bantningen är tvåfaldigt: 1) att minska i vikt och 2) att uppnå ett acceptabelt BMI (Body Mass Index), vilket enligt friskvårdsexpertis bör understiga 25.

$$BMI = \frac{m}{l^2} < 25 ,$$

där m är vikt i kg och l är längd i meter. I det aktuella fallet är längden 1.9 m och kravet på vikten är alltså

$$m < 25 \cdot 1.9^2 = 90.25 \text{ kg}$$

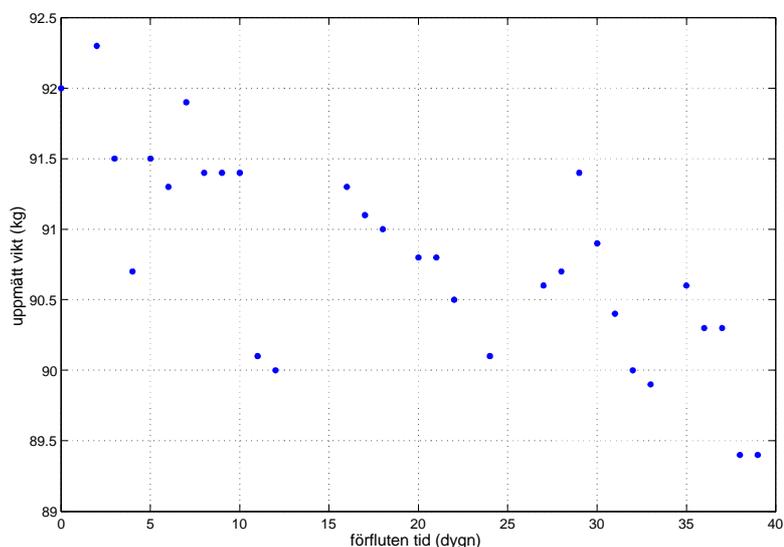
Resultatet av vägningarna ses i figur 1.1. I figuren kan man observera att vikten varierar åtskilligt från dag till dag och man kan fråga sig om denna variation beror på mätfel i vågen eller på variation i kroppsvikten.

Vår bantare har emellertid förutsett detta problem och har därför upprepat vägningen varje morgon tre gånger. Därmed får han en uppskattning av vågens tillfälliga mätfel. Resultaten av dessa vägningar kan ses i figur 1.2 där också medelvärdena av de tre dagliga vägningarna är markerade.

De upprepade mätningarna visar att vågens repeterbarhetsfel är litet i förhållande till kroppsviktens variation. En statistisk beräkning ger uppskattningarna

$$s_{rep} = 0.15 \text{ kg} \qquad s_{tot} = 0.5 \text{ kg},$$

där s betecknar standardavvikelsen för respektive variationskälla. I avsnitt 3.5.3 visas hur standardavvikelserna beräknats. Vägningarna visar en så stor

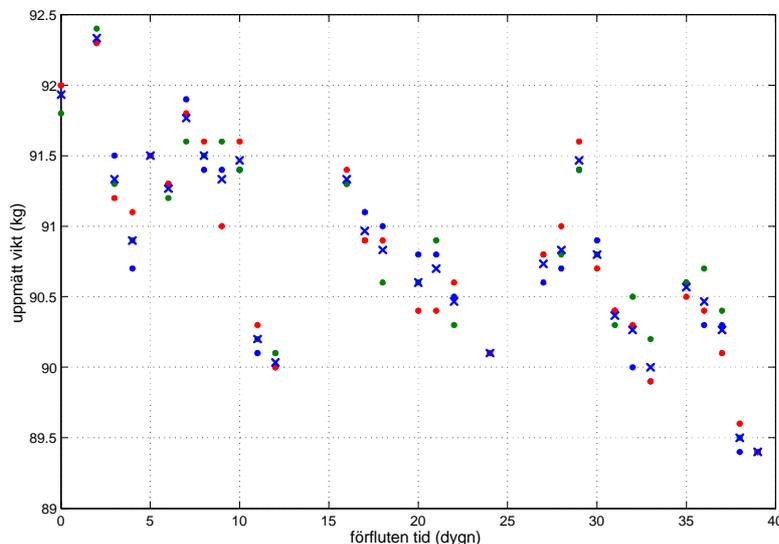


Figur 1.1: Resultat av daglig vägning på badrumsvåg.

daglig variation i kroppsvikten att det är svårt att avgöra om det första målet uppfyllts; om vikten verkligen har minskat. Figurerna antyder dock en minskande tendens, vilken möjligen förstärks något om man studerar medelvärdena. Detta ger oss två slutsatser om mätsystemet: 1) kraven på mätsystemets variation måste relateras till kravet på upplösning i mätobjektet. I vårt fall är mätsystemets repeterbarhet ca 30% av den totala variationen, vilket kan bedömas vara tillräckligt i detta fall. Faktiskt är det just denna gräns som kvalitetsstandarden QS9000 anger som acceptabelt i dess manual om mätsystemsanalys. 2) Vid repeterbarhetsfel i mätsystemet kan man genom att ta medelvärdet av upprepade mätningar minska variationen i mätresultaten. I det aktuella fallet är repeterbarhetsfelet tillräckligt litet, men kan göras mindre om det behövs genom medelvärdesbildning.

För att avgöra om det andra målet uppfyllts, om acceptabelt BMI uppnåtts, är inte de genomförda mätningarna tillräckliga. Här tillkommer nämligen det systematiska felet hos den använda badrumsvågen. Friskvårdsexpertisens angivna rekommendationer refererar förstås till längd och vikt enligt SI-systemet och vi vet ingenting om hur väl den använda vågen överensstämmer med den definierande internationella normalvikten. Badrumsvågen måste alltså kalibreras mot en referens eller bedömas utifrån eventuell specifikation om maximal felvisning.

Mätosäkerheten i bantningsproceduren härrör alltså från två typer av mätfel. Det första är ett slumpmässigt fel som ändrar sig till sin storlek för



Figur 1.2: Resultat av tre dagliga vägningar, samt deras respektive medelvärden.

varje mätning och därmed alltid kan minskas genom medelvärdesbildning. Detta representerar det vi skall kalla **repeaterbarhet** i fortsättningen. Det andra är ett systematiskt fel som är detsamma för varje mätning. Det kan inte minskas genom medelvärdesbildning, utan kan bara reduceras genom kalibrering mot referens. Detta representeras tillsammans med det slumpmässiga felet av begreppet **reproducerbarhet**. Vid jämförande mätningar däremot kan systematiska fel elimineras och endast mätosäkerhet som härrör från repeaterbarhetsfel är då intressant.

Vi ser att redan i en enkel vardagssituation är det viktigt att ha kontroll över mätosäkerheten och att osäkerheten måste betraktas utifrån vilket ändamål man har med sina mätningar.

1.2 Olika ändamål för mätningar

Mätningar genomförs i oräkneliga sammanhang och ändamål. I det inledande bantningsexemplet angavs två olika syften som leder till olika synsätt på mätosäkerheten, nämligen 1) att minska sin vikt eller 2) att uppnå ett bestämt BMI. På samma sätt måste varje mätsituation betraktas utifrån sitt ändamål för att mätosäkerheten skall kunna behandlas på ett relevant sätt. Man kan dela upp mätsituationerna i följande grupper (se figur 1.3):

- Vetenskapliga mätningar. Dessa mätningar syftar till att kontrollera en

hypotes avseende någon egenskap och är oftast möjlig att relatera till någon jämförande mätning. Bantning är ett exempel på en vetenskaplig undersökning, där hypotesen är att en ändring i levnadsvillkoren skall resultera i en minskad vikt. Mätningarna relateras till mätningar före åtgärden. Vid arbetet på utveckling av produkter inom ett företag kan egenskaperna för en nyutvecklad produkt jämföras med tidigare produkttyper. Här är alltså jämförande mätningar dominerande och systematiska fel kan elimineras genom att hålla mätförhållandena konstanta; repeterbarheten bestämmer osäkerheten.

- Kvalitetskontroll. Inom tillverkande företag kontrolleras produkterna i olika utsträckning genom mätningar. Ofta är mätningarna av företags-specifik natur med lokala referenser, och jämförelser med SI-enheter eller andra globala referenser är inte relevanta. Man intresserar sig i första hand för att kontrollera variationer över tid, mellan operatörer, maskiner, eller mellan olika underleverantörer. Här är därmed jämförande mätningar inom företaget dominerande och mätosäkerheten måste betraktas i förhållande till det aktuella syftet. Åter är repeterbarheten det dominerande begreppet.
- Globala jämförelser. Vid bedömning av produkter från olika tillverkare avseende någon jämförbar mätstorhet är situationen annorlunda. Skillnaden mellan två produkters angivna egenskaper kan orsakas av systematiska skillnader i respektive tillverkares mätsystem. En skillnad mellan tillverkarens angivna egenskap och köparens kontrollmätning kan också bero på olikheter i mätsystemen. Angivna hållfasthetsegenskaper hos en produkt måste motsvaras av de krafter som produkten kommer att utsättas för och endast en gemensam referens kan säkerställa detta. Vid dylika globala jämförelser måste de olika mätsystemens fel i förhållande till internationella standarder och referenser inkluderas i mätosäkerheten; reproducerbarhet är avgörande.
- Gränsvärden. Vid mätning mot gränsvärden, bestämda av myndigheter eller andra organisationer, måste förstas gränsvärdessättaren och tillverkaren ha samma referens och samma standardiserade mätmetod. Jämförelser med gränsen måste alltså inkludera osäkerheten i respektive mätning, vilka ger underlag för relevanta säkerhetsmarginaler. Reproducerbarheten bestämmer mätosäkerheten.

Denna uppdelning visar hur mätosäkerheten bör betraktas i olika sammanhang. Den mätsoäkerhetsangivelse som är etablerad i ISO-standarderna GUM är relaterad till globala jämförelser, med SI-enheter och internationella provningsstandarder som referenser. Detta angreppssätt ger med nödvändighet stora mätosäkerheter som vid vetenskapliga mätningar eller kvalitetskontroll kan visas sig vara alltför stora för att kunna användas. I dessa

Olika mättingsändamål

<ul style="list-style-type: none"> • Vetenskapliga mätningar • Kvalitetskontroll 	<p><i>Prova hypotes genom att mäta före och efter åtgärd.</i></p> <p><i>Företagsspecifika jämförelser</i></p>	<p><i>Fokus på skillnader, systematiska fel kan elimineras, slumpmässiga fel kan minskas genom medelvärdesbildning</i></p>
<ul style="list-style-type: none"> • Globalt jämförbara mätningar • Jämförelse mot fasta gränser 	<p><i>Konkurrerande produkter, global anpassning</i></p> <p><i>Miljökrav, säkerhetsgränser</i></p>	<p><i>Fokus på absoluta värden. Jämförelser över stora avstånd i rum och tid. Behov av globala referenser, systematiska fel är kritiska.</i></p>

Figur 1.3: *Fokusering på skillnader eller absoluta värden har inflytande på mätosäkerheten.*

fall måste speciella utredningar genomföras där man drar nytta av den fördelaktiga situationen med en lokal referens och möjligheter till minskning av mätsystemets variation genom lokala åtgärder.

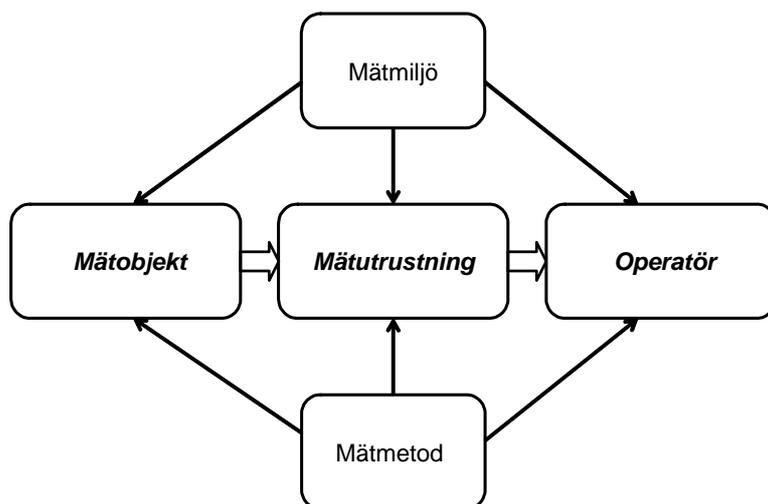
1.3 Mätfel

Vid mätning eller provning vill man bestämma någon karaktäristika hos ett objekt antingen i form av ett numeriskt värde eller i form av ett godkännande/ej godkännande. Överensstämmelsen mellan mätresultatet och karaktäristikan hos objektet är aldrig hundra procentig, utan mätresultatet är behäftat med mätfel som är mer eller mindre kända. Ett sätt att beskriva mätprocessen presenteras i [8] och är här illustrerad genom figur 1.4, där olika komponenter i mätprocessen illustreras.

- *Mätobjektet* är det objekt vars karaktäristika man skall bestämma. Det kan självt vara en källa till mätfel: Vid kemisk analys kan bestämningen av en halt av en kemisk substans påverkas av vilken lösning det befinner sig i; vid diametermätning av ett objekt kommer ovalitet att införa ett mätfel, vid vägning kan främmande smutspartiklar ge ett falskt tillskott till vikten.
- *Mätutrustningen* kan t. ex. bestå i ett skjutmått för längdmätning eller i en hel utrustning med mätton, omvandlare, förstärkare, grafiskt

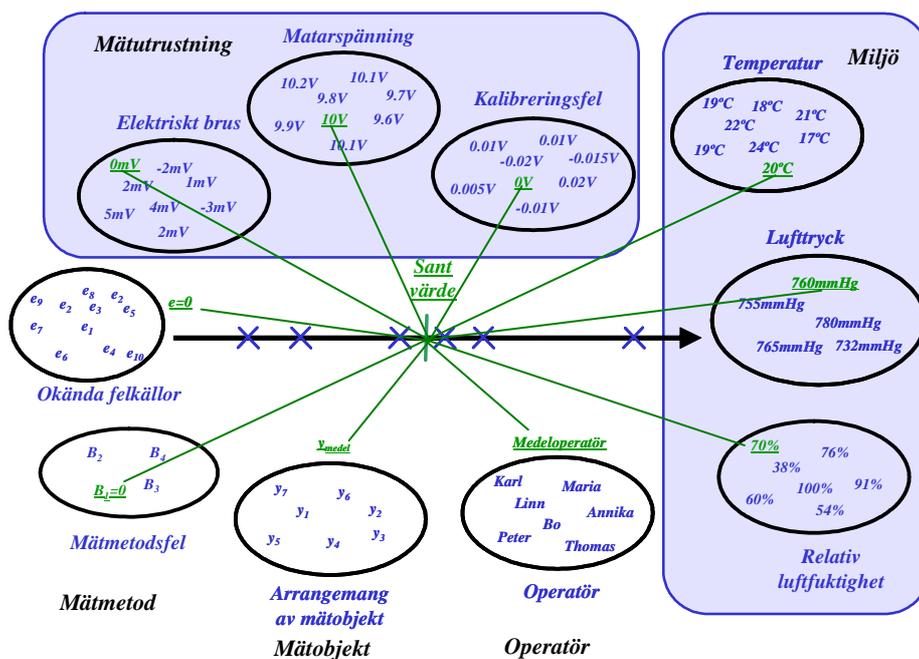
gränssnitt o. dyl. Här finns oftast flera källor till mätfel, såsom kalibreringsfel, elektriskt brus, ofullständig rikriktning och avrundning.

- *Operatören* är den eller de personer som genomför mätningen. I fallet med ett skjutmått påverkar operatören direkt mätningen, medan i andra fall mer eller mindre automatiska mätprocedurer lämnar operatörens inverkan till kontroll av mätmetodens uppfyllande. Vid direkt påverkan kommer förstås den mänskliga faktorn att ge ett mätfel, specifikt för varje operatör, medan operatörsfel också kan inträffa i det indirekta fallet genom riggning av mätutrustning och avläsning av såväl mätvärde som miljövariabler.
- *Mätmiljön* är ett uttryck för alla de störningar som kan påverka mätningen, såsom temperatur, luftfuktighet, drag, lufttryck, ljud och vibrationer. Varje avvikelse från nominellt värde på dessa variabler påverkar såväl mätobjekt, mätutrustning och operatör och ger upphov till mätfel.
- *Mätmetoden* är proceduren för själva utförandet, riggning av utrustning, preparering av mätobjekt och fastställande av nominella miljövariabler och gränser för dessa. Därmed kan den påverka mätobjekt, val av mätutrustning och operatör. T ex. anges ofta i en mätinstruktion att mätningar skall utföras i rumstemperatur under normalt lufttryck och luftfuktighet utan att exakta nominella värden är angivna för dessa miljövariabler. Vaga definitioner och instruktioner kan vara källor till mätfel då tolkningen varierar mellan olika mättillfällen.



Figur 1.4: En översiktlig bild av ett mätsystem.

Vi ser att mätfel kan uppstå i alla de delar i mätsystemet som är illustrerade i figur 1.4. Många av dessa fel kan elimineras genom att man håller miljövariabler konstanta, genom att man tillgodoser en stor noggrannhet i tillämpningen av mätmetoden och genom att se till att mätutrustningen är ordentligt kalibrerad. Fel som upptäcks vid kalibrering kan man justera för i det slutgiltiga mätvärdet och kända avvikelser från nominella miljövariabler kan eventuellt kompenseras. Men, även efter sådana elimineringsåtgärder återstår oftast åtskilliga fel som inte kan justeras p.g.a. okunskap eller bristande resurser. En enskild operatörs felbidrag är t. ex. inte känt, mätfel som uppstått i kalibreringen är begränsat men okänt till sin storlek. Personella och ekonomiska resurser begränsar möjligheterna till kalibrering och justering av instrument. Alla sådana **felbidrag** som är **okända till sin storlek** ger upphov till en **mätosäkerhet** och det är denna mätosäkerhet som vi i fortsättningen kommer att söka kvantifiera för att, trots felen, behärska mätningarna.

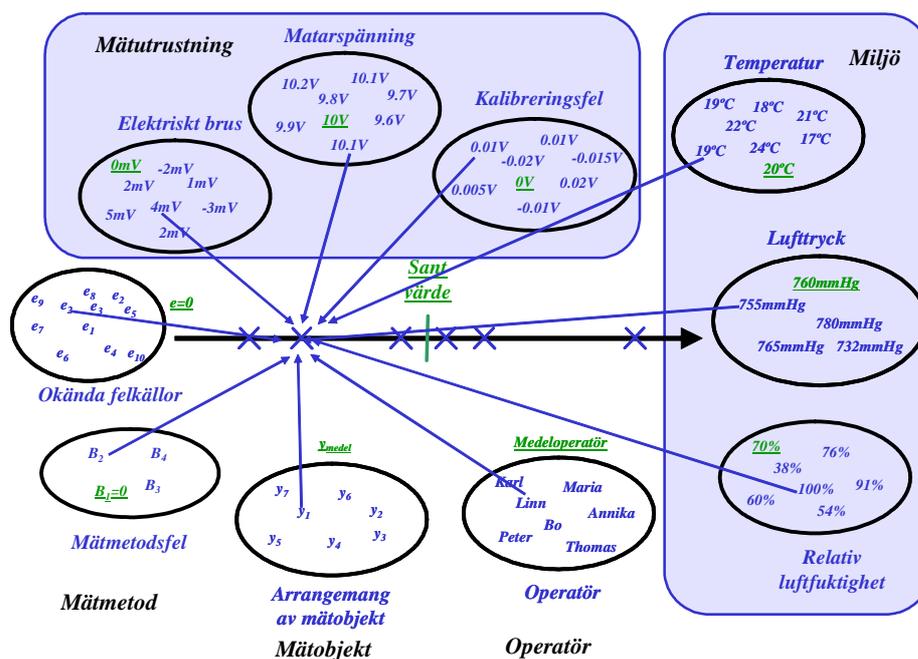


Figur 1.5: Illustration av mätsituationen. Olika felkällor ses som populationer av inflytelsevariabler. Sanningen ges av nominella värden eller medelvärden.

I figurerna 1.5 och 1.6 illustreras ett statistiskt synsätt på mätfel. Varje felkälla (ellips i figurerna) ses som en population av olika möjliga utfall av felkällan, i temperaturfallet t. ex. kommer temperaturerna vid mätningar vid olika tillfällen att variera och kan därmed ses som ett slumpmässigt

val ur populationen av tillåtna temperaturer. Populationens storlek bestäms ibland av mätmetoden eller speciella instruktioner. Det värde i populationen som ger ett felfritt bidrag till mätvärdet är antingen ett definierat nominellt värde (t. ex. rumstemperatur, 20°C , elektriskt brus, 0V) eller medelvärdet av den definierade populationen (t. ex. medelbidraget från alla operatörer med en viss accepterad utbildning). Figur 1.5 visar hur valet av nominella värden respektive medelvärden ger det sanna värdet. Figur 1.6 visar hur ett slumpmässigt val ur de olika populationerna ger ett värde behäftat med mätfel. Förutom identifierbara källor så fås också felbidrag från okända källor, vars värde är lika med noll för en sann mätning och vars medelvärde man strävar efter att göra till noll genom kalibreringar och justeringar.

Varje mätsituation representerar alltså en dragning från en **population av mätförhållanden**. Det sanna mätvärdet är det som förväntas av populationen och mätosäkerheten är ett mått på spridning orsakad av populationen.



Figur 1.6: Illustration av mätsituationen. Olika felkällor ses som populationer av inflytelsevariabler. Ett slumpmässigt urval ur var och en av populationerna ger ett slumpmässigt mätfel.

1.4 Hur används mätosäkerheten?

Syftet med en mätning är nästan alltid att man vill jämföra karakteristiken för ett objekt med något annat objekt. Mätningar i produktion görs t. ex. för att jämföra med specifikationer eller ritningar, mätningar av miljöpåverkan görs för att jämföra mot myndighetsbestämda gränsvärden, mätningar av produkter görs kanske för att jämföra med konkurrerande produkter av samma eller annorlunda typ, mätningar i utvecklingsarbete används för att jämföra nya egenskaper med gamla. Jämförelser av mätvärden handlar då oftast om att svara på frågan: är x större än y ? Men, om mätvärdena för x respektive y är behäftade med mätosäkerhet kan frågan inte besvaras säkert; x kan vara lika med y även om mätvärdena skiljer sig åt.

Vi kommer i fortsättningen att modellera mätvärdet \hat{x} som en summa av det sanna värdet x och ett mätfel e ,

$$\hat{x} = x + e .$$

Med hjälp av en osäkerhetsutredning kan vi uppskatta variationen hos felet e och bestämma ett konfidensintervall för det sanna värdet x :

$$x_L \leq x \leq x_H , \text{ med } 95\% \text{ konfidens}$$

eller

$$x = \hat{x} \pm U ,$$

där U kallas **den utvidgade standardosäkerheten** och konfidensgränserna alltså är $x_L = \hat{x} - U$ och $x_H = \hat{x} + U$. Denna utvidgade mätosäkerhet är i sin tur baserad på **standardosäkerheten** u , som motsvarar det statistiska begreppet standardavvikelse¹.

För att svara på frågan om x är större än y måste man alltså känna såväl mätvärdena \hat{x} och \hat{y} som deras respektive utvidgade mätosäkerheter U_x och U_y . I avsnitt 2.3 beskrivs hur man utifrån dessa fyra värden avgör om skillnader föreligger; är x signifikant större än y och i kapitel 3 beskrivs hur man bestämmer U_x och U_y .

1.5 Systematiska och slumpmässiga fel

Det inledande bantningsexemplet visade vikten av att skilja på slumpmässiga och systematiska fel. Skillnaden definieras på följande sätt: **Slumpmässiga fel förändras vid upprepade mätningar oberoende av varandra medan systematiska fel är konstanta mellan mätningar**. Detta innebär t. ex. att vid mätning med ett skjutmått kommer en permanent förslitning på skjutmättet att orsaka ett systematiskt fel som är detsamma

¹Begreppen standardosäkerhet och utvidgad standardosäkerhet med beteckningarna u resp. U är hämtade från ISO-guiden GUM[1].

vid upprepade mätningar, medan ett slumpmässigt fel kan uppstå genom att operatören vinklar skjutmättet något olika vid upprepade mätningar.

Utifrån den statistiska bilden i figur 1.6 av en allmän mätsituation förstår man att man kommer att generera slumpmässiga fel om man vid upprepade mätningar väljer slumpmässigt ur var och en av de olika källornas populationer. Om man istället vid upprepade mätningar gör samma val ur en eller flera populationer så får man ett systematiskt fel. Ett systematiskt fel fås också om man gör upprepade mätningar ur andra populationer än de som mätmetoden avser. Ett laboratorium i Sverige kommer t. ex. inte att ha samma populationsmedelvärde för temperatur och lufttryck som ett laboratorium i ett tropiskt klimat. Varje laboratorium har också en unik population av operatörer, vars medelvärde knappast sammanfaller med det globala medelvärdet. Alltså har varje laboratorium, varje företag dvs. varje grupp ett systematiskt fel.

Definitionerna av slumpmässighet resp. systematik ger två viktiga metoder för att minimera mätosäkerheter:

Två kungsvägar för minskad mätosäkerhet:

Om upprepade mätningar görs kommer slumpmässiga fel att variera och genom att medelvärdesbilda kan mätosäkerheten minskas.

Om två eller flera mätobjekt skall jämföras med avseende på någon mätstorhet med additiva systematiska fel så kommer felen att elimineras om man studerar differenserna.

Fel på kontrollerbar nivå elimineras alltså genom att hålla variabler konstanta, fel på ej kontrollerbar nivå genom att slumpmässigt variera variablerna!

Då dessa båda metoder är oerhört enkla och kraftfulla så är det av största vikt att man beaktar om mätfel är slumpmässiga, systematiska eller sammansatta av båda typerna av fel.

Om man t. ex. vill genomföra en jämförande mätning som är känslig för temperaturpåverkan och operatörens handlag, så kan man se till att kontrollera temperaturen speciellt och låta samma operatör genomföra mätningarna. Skillnaden i jämförelsen kommer då att ha mindre mätosäkerhet då två felkällor eliminerats.

Om man istället måste genomföra en noggrann mätning utan möjligheten till jämförande mätning, så kan man anstränga sig för att hålla nominell temperatur (om den är känd) och låta flera operatörer göra oberoende mätningar för att sedan ta medelvärdet. Mätvärdet kommer då att ha en mindre osäkerhet än normalt då de två dominerande felkällorna minskats, i temperaturfallet genom minskning av temperaturpopulationen, i operatörsfallet genom medelvärdesbildning.

1.5.1 Olika syften och horisonter för mätosäkerhetsangivelse

Om en osäkerhetskomponent är systematisk eller slumpmässig enligt ovanstående definition bestäms av från vilken horisont som mätvärdet betraktas. Om mätningen betraktas från en global horisont, så kommer mätningen att referera till en överenskommen referens, t. ex. till en väldefinierad SI-enhet eller genom någon internationell standard. Upprepade mätningar skall då tolkas som mätningar utförda av olika mätlaboratorier inom en viss klass av sådana. Laboratoriefelet kan då betraktas som ett slumpmässigt fel, som en global intressent kan uppleva om en mätning genomförs hos såväl producent som kund. Mätosäkerhet på denna globala nivå definieras vi som ett mått på **reproducerbarhet** i enlighet med de begrepp som definieras i standarden ISO 5725 [10].

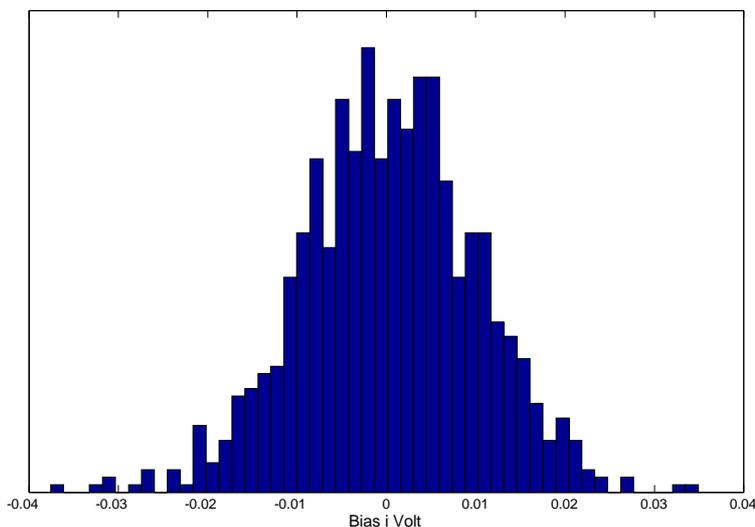
Om mätningen istället betraktas inom ett enskilt företag eller laboratorium eller någon annan population av mätförhållanden och syftar till jämförelser inom denna population så kan man betrakta medelvärdet av mätningar från mätsystemet som en referens och upprepade mätningar betyder mätningar utförda inom populationen. Denna mätosäkerhet definieras vi som mått på **repeterbarhet**, vilken måste förknippas med en population av mätförhållanden samt en referens.

Mätosäkerheten kan därmed vara olika stor i olika sammanhang och av dessa sammanhang skiljer vi ut ett, nämligen det globala, och ger mätosäkerheten här ett särskilt namn: reproducerbarhet. Repeterbarhet kan däremot definieras för olika sammanhang och därmed ge möjligheter till avvägning mellan noggrannhet och generalitet.

Ett systematiskt fel hos ett företags medelvärde blir ett slumpmässigt fel från en global horisont. Om företaget säljer sin produkt så måste dess angivna mätosäkerhet betraktas från den globala horisonten, medan man vid interna jämförelser har en mätosäkerhet som är mindre genom att det egna medelvärdesfelet kan elimineras vid jämförelser.

Exempel 1.5.1 *Betrakta världens voltmeterar som en statistisk population (se figur 1.7). Varje specifikt instrument har en systematisk avvikelse, en bias. Ett laboratorium L köper en voltmeter, dvs. väljer slumpmässigt ur populationen. Denna voltmeter har bias B_L . I ett globalt perspektiv måste en utomstående intressent kunna jämföra sin provning hos laboratoriet L med någon annan provning och ser därför en mätosäkerhet som inbegriper populationen av systematiska avvikelser B , reproducerbarheten. Vid jämförande mätningar med samma voltmeter kan däremot det systematiska felet elimineras och mätosäkerheten består endast i det slumpmässiga felet hos voltmeteren, repeterbarheten.*

Mätosäkerhetsangivelse måste alltså vara kopplad till en **referens** och en **population**, där referensen definieras utifrån vilket syfte mätosäkerhe-



Figur 1.7: *Hypotetiskt histogram över den systematiska avvikelsen hos världens voltmetrar.*

ten anges och populationen bestämmer den varians som kan förväntas i mätningar.

Den mätosäkerhet som krävs av kvalitetsstandarder är i allmänhet reproducerbarheten, dvs. den som definerar sanningen genom en internationell standard och som population betraktar alla kvalitetsgodkända mätsystem. Den internationella standard som oftast används för bestämning av mätosäkerhet är utgiven av ISO: ISO Guide to the Expression of Uncertainty in Measurement (GUM)[1]. Denna utgår just från det internationella perspektivet och betraktar mätosäkerheten med SI-enheternas internationella definition som referens och världens provningslaboratorier som population.

För att praktiskt dra nytta av sina mätosäkerhetsanalyser är det dock viktigt att också kunna definiera andra referenser och populationer. Vid jämförande undersökningar med höga krav på att kunna detektera små skillnader kan man begränsa populationen genom att välja samma instrument och operatör, hålla miljöbetingelserna konstanta och genomföra alla mätningar under begränsad tid. Referensen kan här tänkas som ett väntevärde av mätresultatet under de angivna betingelserna, dvs. ett viktat medelvärde av mätningar genomförda vid alla angivna betingelser. Vid kvalitetskontroll av produktion är ofta endast avvikelser från ett medelvärde av primärt intresse, populationen består här i alla involverade mätsystem och det angivna medelvärdet betraktas som referensvärde. Här intresserar man sig i första hand för mätningar som ger repeterbara mätvärden inom den definierade

populationen och i förhållande till den lokala referensen.

Exempel 1.5.2 *Vid tillverkning av kullager måste inner- och ytterdiametrar vara relaterade till en global referens, eftersom kunden skall passa ihop lagret med en axel och ett lagerhus tillverkade av någon annan. Här måste toleranserna hållas med en mätosäkerhet som är relaterad till den internationella meter-referensen. Däremot, vid passningen av kulor eller rullar till varandra och till ytter- och innerring är endast den interna jämförelsen relevant och här kan man därmed klara avsevärt snävare toleranser då mätosäkerheten är relaterad till någon intern referens.*

Exempel 1.5.3 *Vid tillverkning av vissa hydraulventiler måste sliden passa i ett hål med ett glapp på bara ett par tusendels millimeter. För att klara av att hålla en sådan tolerans krävs individuell passning och sortering av tillverkade slider efter håldiametrar. Vid tillverkningen kan man alltså tolerera en mätosäkerhet på kanske någon hundradels millimeter. Vid passningen krävs bättre noggrannhet, men endast jämförelsen är intressant, vilket innebär att bara den lokala repeterbarheten inverkar och en hög noggrannhet är möjlig att uppnå.*

Exempel 1.5.4 *I reklamen för en bilmodell anges kanske att bränsleförbrukningen vid allmän körning är 6.5 l på 10 mil. Denna uppgift skall tåla granskning på alla marknader och måste därmed innefatta en mätosäkerhet som är relaterad till globala referenser avseende volym, längd och definition av "allmän körning". Kanske har man mätt förbrukningen till 6 l och lagt till en halv liter på grund av mätosäkerheten. Vid utvecklingsarbetet för att uppnå denna bränsleförbrukning kan man kanske för varje enskild åtgärd bara uppnå en bråkdel av den osäkerhet på en halv liter som finns globalt och för att förändringar skall kunna detekteras krävs en mycket mindre mätosäkerhet lokalt.*

Exempel 1.5.5 *I miljörelaterade sammanhang ställs lagkrav på att vissa gränsvärden inte överskrids, t. ex. vid förbränning av avfall. Här måste såväl den gränssättande myndigheten och den ansvarige för förbränningen komma överens om hur gränsen är relaterad till mätosäkerheten för att inte hamna i rättstvister angående tolkning av mätresultat. En global referens är nödvändig och mätfelsvariationen måste beaktas över alla de mätlaboratorier som kan tänkas bli inblandade i kontroller.*

1.6 Sammanfattning

Varje mätning består i en sann egenskap störd av en rad olika felbidrag. De specifika felbidragen vid varje mätning kan ses som en slumpmässig dragning från ett antal populationer av felbidrag. Varje sådan felbidragskällas population har ett nominellt värde, definierat i mätinstruktionen eller givet av medelvärdet av populationen. **Det sanna mätvärdet för en egenskap är det mätvärde man skulle ha erhållit om samtliga störningsvariabler kunde hållas vid sina nominella värden.** Alla avvikelser från dessa nominella värden ger upphov till mätfel. Mätfel kan elimineras genom justering för kända avvikelser och avvikelser som inte är kända ger upphov till **mätosäkerhet**. Denna osäkerhet kan ges en statistisk tolkning genom att betrakta den som representerad av de **populationer av störningsvariabler** som är representerade vid mättillfället.

Mätningar görs ytterst i syfte att jämföra egenskaper. Sådana jämförelser kan göras i ett **globalt** eller i något lämpligt valt **lokalt** perspektiv. I det globala perspektivet kommer mätningen i ett specifikt fall att jämföras med mätningar genomförda vid en rad andra mätlaboratorier över hela världen och populationerna av mätfelsvariabler är stora. I ett lokalt perspektiv kan man bestämma graden av variation mellan mätsituationer och därmed minska populationernas storlek. På detta sätt kan mätosäkerheten avvägas mot ekonomiska, tidsmässiga och praktiska överväganden. **Ändamål och perspektiv måste definieras för varje mätosäkerhetsutredning.**

Det statistiska betraktelsesättet tydliggör två viktiga metoder för att minska mätosäkerheten:

Kungsväg 1: Om flera upprepade mätningar görs kommer slumpmässiga fel att variera och genom medelvärdesbildning kan mätosäkerheten minskas. Detta innebär att man gör flera **slumpmässiga val** ur de olika populationerna av felbidragsvariabler. Medelvärdesbildning kommer då att flytta tyngdpunkten i respektive population närmare det nominella värdet.

Kungsväg 2: Om två eller flera mätobjekt skall jämföras med avseende på någon mätstorhet med additiva systematiska fel så kommer felen att elimineras om man studerar differenserna. Detta innebär att man gör **samma val** i någon eller några av populationerna av felvariabler vid mätningarna av de två objekten och genom att endast studera skillnaden eliminerar felbidraget.

Kapitel 2

Statistiska grunder

2.1 Matematiska formler

Denna handbok innehåller en hel del matematiska formler som vid första påseendet ser synnerligen kryptiska ut för den oinvidde. De är viktiga för att kunna beskriva samband och beräkningar på ett kompakt sätt och kunna förstå ideerna bakom olika metoder för mätosäkerhet, men kan vid praktisk användning ersättas med färdiga funktioner på miniräknare eller datorprogram. Större delen av formlerna består av olika former av kvadratsummor, vilka här skall beskrivas utifrån följande tabell med mätvärden:

Tabell över mätvärden, vikt i g

prod.	1	2	3	4	5	6	7	8	9	10	medel
A	0.65	1.00	0.85	0.85	0.55	1.00	0.95	0.85	1.00	0.60	
A	0.60	1.00	0.80	0.95	0.45	1.00	0.95	0.80	1.00	0.70	0.83
B	0.55	1.05	0.80	0.80	0.40	1.00	0.95	0.75	1.00	0.55	
B	0.55	0.95	0.75	0.75	0.40	1.05	0.90	0.70	0.95	0.50	0.77
C	0.50	1.05	0.80	0.80	0.45	1.00	0.95	0.80	1.05	0.85	
C	0.55	1.00	0.80	0.80	0.50	1.05	0.95	0.80	1.05	0.80	0.83
medel	0.57	1.01	0.80	0.82	0.46	1.02	0.94	0.78	1.01	0.67	0.81

Tabellen innehåller resultatet av vägningar av tio olika produkter, var och en vägd med tre olika mätsystem, A, B och C och varje vägning har dessutom upprepats för att ge två mätvärden per mätsystem och produkt. För att matematiskt beteckna ett mätvärde behövs därmed tre bestämningar, produkt, mätsystem och upprepningstillfälle, vilket här görs med index genom beteckningen X_{ijk} . Alltså är i detta fall $X_{2,3,1} = 1.05$ g och

$X_{9,1,2} = 1.00$ g. Medelvärden är ofta intressanta såväl över samtliga värden, som över produkter och mätsystem: Medelvärden betecknas allmänt med ett streck över variabelbeteckningen, \bar{X} , men måste ibland preciseras för att skilja på olika medelvärden. Det index som man medelvärdesbildar över ersätts då med en prick och följaktligen betyder $\bar{X}_{.2}$ = 0.77 g här att medelvärdet är taget över alla mätningar som gjorts med mätsystem B. $\bar{X}_{6..}$ = 1.02 g betyder att medelvärdet för alla mätningar på produkt nr. 6 är 1.02 g och medelvärdet för samtliga 60 mätningar betecknas $\bar{X}_{...}$ = 0.81.

Ett sätt att beskriva totalvariationen i mätvärdestabellen är att beräkna summan av alla kvadratiske avvikelser från totalmedelvärdet:

$$\begin{aligned} SS_T &= (0.65 - 0.81)^2 + (1.00 - 0.81)^2 + (0.85 - 0.81)^2 + \dots + \\ &\quad + (0.60 - 0.81)^2 + (1.00 - 0.81)^2 + \dots + (0.80 - 0.81)^2 \\ &= (X_{1,1,1} - \bar{X}_{...})^2 + (X_{2,1,1} - \bar{X}_{...})^2 + (X_{3,1,1} - \bar{X}_{...})^2 + \dots + \\ &\quad + (X_{1,1,2} - \bar{X}_{...})^2 + (X_{2,1,2} - \bar{X}_{...})^2 + \dots + (X_{10,3,2} - \bar{X}_{...})^2 . \end{aligned}$$

Denna summa kan man med hjälp av indexeringstekniken mer kompakt skriva med summatecken:

$$\begin{aligned} SS_T &= \sum_{i=1}^{10} (X_{i,1,1} - \bar{X}_{...})^2 + \sum_{i=1}^{10} (X_{i,2,1} - \bar{X}_{...})^2 + \sum_{i=1}^{10} (X_{i,3,1} - \bar{X}_{...})^2 + \\ &\quad \sum_{i=1}^{10} (X_{i,1,2} - \bar{X}_{...})^2 + \sum_{i=1}^{10} (X_{i,2,2} - \bar{X}_{...})^2 + \sum_{i=1}^{10} (X_{i,3,2} - \bar{X}_{...})^2 , \end{aligned}$$

och ännu mer kompakt genom

$$\begin{aligned} SS_T &= \sum_{i=1}^{10} \sum_{j=1}^3 (X_{i,j,1} - \bar{X}_{...})^2 + \sum_{i=1}^{10} \sum_{j=1}^3 (X_{i,j,2} - \bar{X}_{...})^2 = \\ &\quad \sum_{i=1}^{10} \sum_{j=1}^3 \sum_{k=1}^2 (X_{i,j,k} - \bar{X}_{...})^2 . \end{aligned}$$

Detta skrivsätt kommer att användas vid upprepade tillfällen i denna handbok för att beskriva olika beräkningstekniker.

Förutom strecket över variabeln som betecknar medelvärde används ibland följande beteckningar över variabeln: \hat{X} betyder uppskattat värde när man behöver skilja på det sanna och kanske okända värdet X och en uppskattning av det. \tilde{x} betyder ett enskilt observerat värde när man behöver skilja detta från det sanna eller uppskattade värdet. Övrig nomenklatur:

σ	standardavvikelse för en stokastisk variabel
s	skattad standardavvikelse
μ	väntevärde för en stokastisk variabel
ν	antal frihetsgrader för en variansskattning
i, j, k	index för variabler
m, n	antal observationer
c	känslighetskoefficient
e, ε	mätfel
u	standardosäkerhet (standardavvikelse för mätfel)
U	utvidgad standardosäkerhet (halvt konfidensintervall)
k	täckningsfaktor för konfidensintervall
\hat{X}	skattat värde av X
\bar{X}	medelvärde
\tilde{x}	observerat värde

2.2 Stokastisk variabel

Vi skiljer på deterministiska storheter och stokastiska variabler. En stokastisk variabel betecknas vanligen med stor bokstav och är slumpmässig i något avseende. Vi kallar variabeln t. ex. för X och den har då inget bestämt värde utan är snarast en symbol för en population av värden. Vi kan emellertid observera variabeln vid olika tillfällen och erhåller då reella tal x_1, x_2, \dots som är **utfall** av den stokastiska variabeln X .

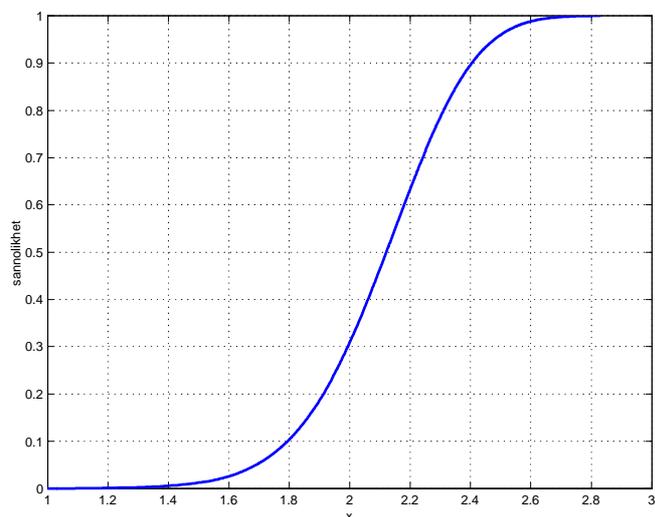
Den stokastiska variabeln X kan beskrivas på olika sätt. Den är fullständigt beskriven av sin statistiska fördelning, som för alla tänkbara numeriska värden x beskriver hur troligt det är att ett utfall skall bli just x . Fördelningen kan beskrivas i matematiska symboler som dess **fördelningsfunktion**

$$F_X(x) = P(X \leq x) \quad (2.1)$$

där P står för sannolikhet och indexeringen X på F_X oftast utlämnas. Denna funktion går från noll till ett när x växer från sitt minimala till sitt maximala värde. Se illustration i figur 2.1.

Exempel 2.2.1 *Om man betraktar temperaturen vid mätningar som en stokastisk variabel, så kan man ange en lägsta gräns till den absoluta nollpunkten -273°C och kanske en högsta gräns till säg 10000°C . Vi kan då konsta-*

tera att sannolikheten för att understiga lägsta gränsen är lika med noll och $F_T(-273) = P(T \leq -273) = 0$ och $F_T(10000) = P(T \leq 10000) = 1$.



Figur 2.1: *Fördelningsfunktion.*

Man kan också beskriva fördelningen genom dess **täthetsfunktion**, vilket är derivatan av fördelningsfunktionen med avseende på x och därmed visar hur fort fördelningsfunktionen växer. Se figur 2.2.

$$f_X(x) = F'_X(x)$$

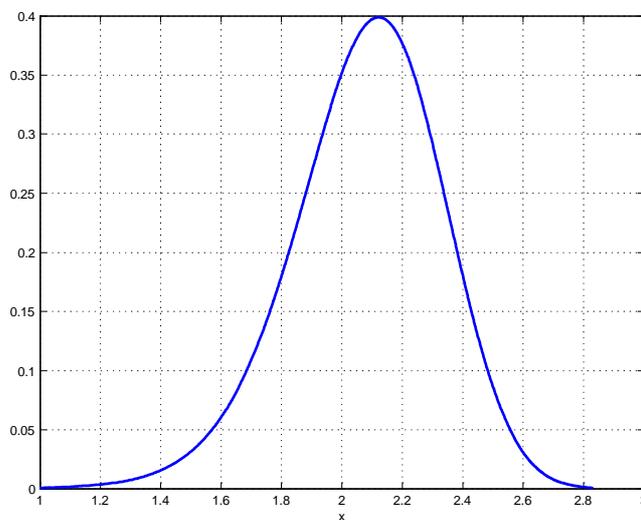
I den fysiska verkligheten är ovanstående funktioner aldrig kända för alla x och i de flesta fall skulle de inte ens kunna beskrivas exakt utan att man använder oändligt många parametrar. Ofta är det tillräckligt att sammanfatta en stokastisk variabels egenskaper i några parametrar, varav de viktigaste är **väntevärdet** som beskriver fördelningens läge,

$$E[X] = \int x f_X(x) dx$$

och **variansen**, som beskriver dess spridning

$$\text{Var}[X] = E[(X - E[X])^2] = \int (x - E[X])^2 f_X(x) dx \quad (2.2)$$

Roten ur variansen, som ofta är ett lämpligt mått på spridning då den har samma dimension som den stokastiska variabeln X , kallas **standardavvikelse**. Väntevärdet motsvarar därmed medelvärdet i populationen, medan



Figur 2.2: Täthetsfunktion.

variansen motsvarar medelvärdet av de kvadrerade avvikelserna från detta medelvärde. I mätosäkerhetssammanhang betraktar vi de okända mätfelen som stokastiska variabler, vars standardavvikelser betecknas med bokstaven u och kallas standardosäkerheter.

Om man vill studera en statistisk fördelning från verkliga mätningar så utgår man från den **empiriska fördelningen**, vilket helt enkelt är fördelningen över olika numeriska värden av den datamängd man har till sitt förfogande. Den empiriska fördelningsfunktionen är motsvarigheten till (2.1) när populationen består av de uppmätta värdena. Om man har n st observationer på $X : x_1, x_2, \dots, x_n$, så är den empiriska fördelningsfunktionen

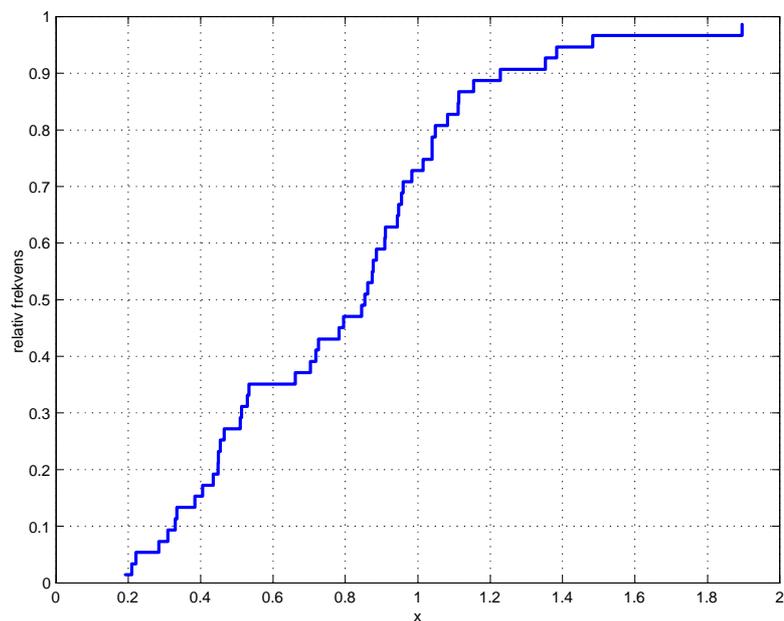
$$\hat{F}(x) = \frac{\#\{x_i \leq x\}}{n}$$

där symbolen $\#$ står för antalet. Se exempel i figur 2.3.

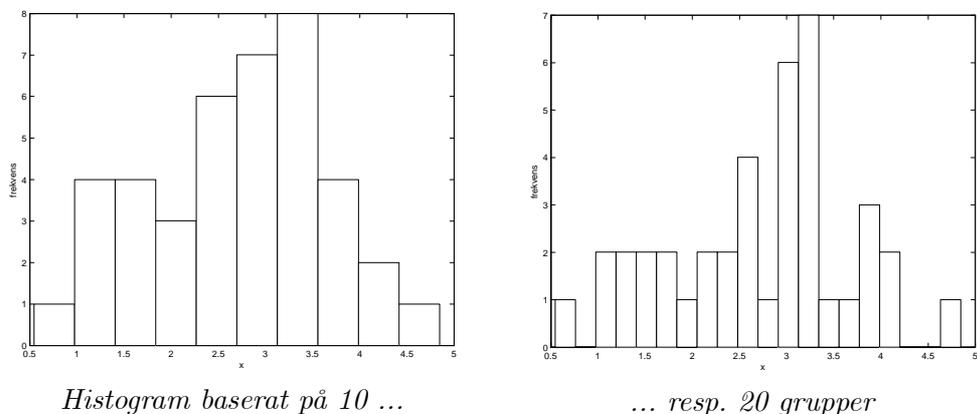
Den empiriska täthetsfunktionen kan kallas ett normerat histogram och definieras genom att de observerade datapunkterna delas in i grupper. Om dessa grupper bestäms som likformiga intervall kring värdena $x_0, x_0 + d, x_0 + 2d, \dots$ så blir det normerade histogrammet

$$\hat{f}(x_0 + jd) = \frac{\#\{x_0 + jd - \frac{d}{2} < x_i \leq x_0 + jd + \frac{d}{2}\}}{nd}$$

vilket kommer att få olika utseende beroende på valet av intervallbredden d . I figurerna nedan illustreras denna effekt på ett histogram när intervallbredden halveras. Observera att ökad upplösning ger större slumpmässigt men mindre systematiskt fel.



Figur 2.3: *Empirisk fördelningsfunktion.*



Histogram baserat på 10 ...

... resp. 20 grupper

För en empirisk fördelning kan den bakomliggande sanna fördelningens läge skattas genom att beräkna dess medelvärde eller median,

$$\widehat{E}[X] = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{eller} \quad \tilde{X} \quad (2.3)$$

där \tilde{X} här står för medianen av observationerna, dvs. det mittersta värdet och taket på väntevärdesuttrycket betecknar att det är en skattning vi räknar ut och inte det sanna väntevärdet. Fördelningens spridning bestäms oftast

genom dess stickprovsvarians,

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (2.4)$$

Att man här dividerar med $n - 1$ istället för n beror på att man annars får en viss underskattning av variansen då variationen kring medelvärdet alltid är mindre än eller lika med variationen kring det sanna väntevärdet, som ju inte är känt. Antalet $n - 1$ definieras som antalet **frihetsgrader** i skattningen. Antalet frihetsgrader i försöket är lika med n men i formeln för variansskattningen utnyttjas medelvärdet, som ju också är en skattning ur försöket och därmed förloras en frihetsgrad. Allmänt skattas alltså variansen genom formeln

$$s^2 = \frac{1}{\nu} \sum_{i=1}^n (X_i - \bar{X})^2, \quad (2.5)$$

där den grekiska bokstaven ν betecknar antalet frihetsgrader i skattningen.

Ett annat mått på fördelningens spridning är **variationsvidden** R ,

$$R = \max(X_i) - \min(X_i)$$

Variationsvidden beror på antalet observationer n , men kan justeras med en n -beroende konstant för att ge en skattning av standardavvikelsen. Detta spridningsmått används i vissa kvalitetsstandarder, t. ex. den amerikanska bilindustristandarden QS9000. I följande tabell ges exempel på justeringsfaktorer k , som ger en väntevärdesriktig skattning av standardavvikelsen genom

$$s = c \cdot R$$

om R är bestämd ur ett stickprov av storleken n .

n	2	3	5	10	20	50
c	0.89	0.59	0.43	0.33	0.27	0.22

Dessa konstanter är bestämda utifrån antagande om normalfördelning och metoden är känslig för avvikelser från detta antagande. Det är därför alltid bättre att bestämma standardavvikelsen genom formeln (2.4), som är generell för alla fördelningstyper och inte kräver någon justeringskonstant.

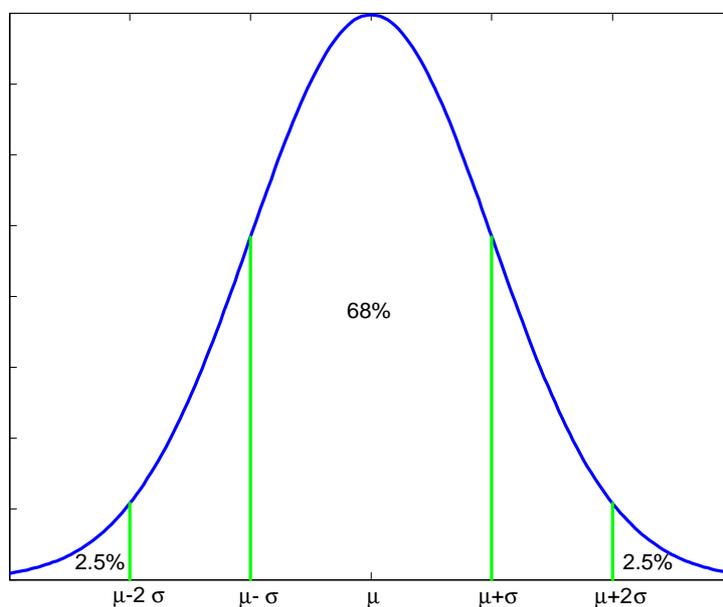
För att kunna behandla de statistiska fördelningarna matematiskt så är det lämpligt att beskriva dem med ett begränsat antal parametrar. Man har konstruerat ett stort antal dylika **parametriserade fördelningar**, som är specificerade dels genom sin typ, dels via vanligtvis två eller tre parametrar. Den vanligaste i våra sammanhang är **normalfördelningen**, vars frekvensfunktion är symmetrisk kring väntevärdet och vars parametrar brukar betecknas med de grekiska bokstäverna μ och σ . Parametrarna är i normalfördelningsfallet identiska med väntevärde respektive standardavvikelse och

en normalfördelad stokastisk variabel betecknas $X \sim N(\mu, \sigma^2)$. I figur 2.4 illustreras täthetsfunktionen för en normalfördelad stokastisk variabel, där vi också markerat de områden som täcks av 68 % resp. 95 % av fördelningen. I normalfördelningsfallet är det nämligen så att sannolikheten P för att variabeln ligger inom \pm en standardavvikelse är 68%:

$$P(\mu - \sigma < X < \mu + \sigma) = 0.68$$

och sannolikheten för att den ligger inom ± 1.96 standardavvikelser är 95%:

$$P(\mu - 1.96 \cdot \sigma < X < \mu + 1.96 \cdot \sigma) = 0.95$$

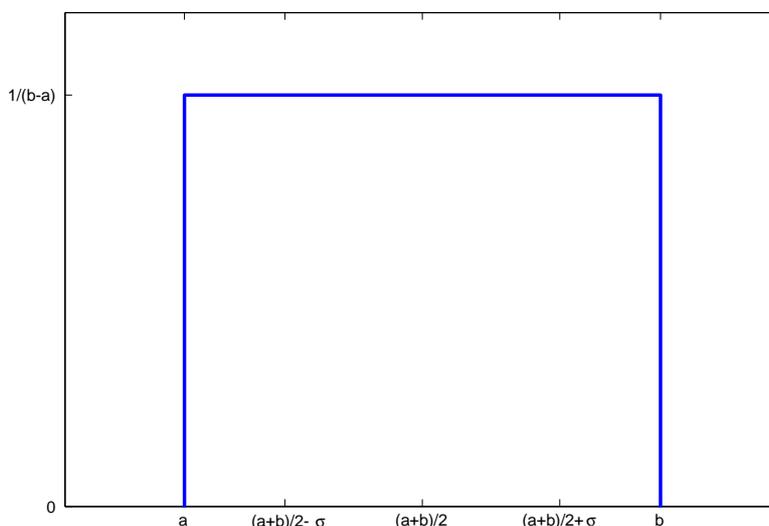


Figur 2.4: Normalfördelningens frekvensfunktion.

Eftersom normalfördelningens parametrar är precis väntevärdet och standardavvikelsen så skattas de ur mätdata enligt (2.3) och (2.4), dvs.

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$



Figur 2.5: Frekvensfunktion för en likformig fördelning med gränserna a och b .

Problem 2.2.1 För att avgöra mätosäkerheten vid bestämning av brottförlängning mättes samma förlängda provkropp upprepade gånger av två olika operatörer med följande resultat

Op. 1 116.5 115.8 116.0 117.3 117.7 117.0

Op. 2 116.0 115.5 115.7 117.1 117.3 116.4

Bestäm medelvärde och standardavvikelse för operatör 1, samt konstruera den empiriska fördelningsfunktionen. Konstruera ett histogram och beräkna variationsvidden för hela datamängden (båda operatörerna).

En annan parametrisk fördelning som används mycket i mätosäkerhets-sammanhang är den **likformiga fördelningen**, vars frekvensfunktion är illustrerad i figur 2.5. Den kallas, efter sin frekvensfunktionens utseende, också för rektangelfördelning och karaktäriseras av att den begränsas av en undre och en övre gräns och att fördelningen av variabelvärdena uppträder med precis samma frekvens för alla värden däremellan.

En likformig fördelning mellan gränserna a och b har väntevärdet lika med medelvärdet av de två gränserna och variansen lika med en tredjedel av kvadraten på halva vidden:

$$\mu_U = \frac{a+b}{2}, \quad \sigma_U^2 = \frac{(b-a)^2}{2^2 \cdot 3}, \quad (2.6)$$

där indexet U är valt efter engelskans "uniform". I fallet då fördelningen har väntevärdet noll kommer nedre gränsen att vara lika med $-b$ och standardavvikelsen blir:

$$\sigma_{U_0} = \frac{b}{\sqrt{3}}.$$

Notera att höjden på frekvensfunktionen blir $\frac{1}{b-a}$ i enlighet med kravet att ytan under hela frekvensfunktionen skall vara lika med ett. Det är också lätt att räkna ut hur stor andel av fördelningen som $\mu_U \pm \sigma$ täcker, nämligen

$$\frac{2\sigma_U}{b-a} = \frac{2(b-a)}{2\sqrt{3}(b-a)} = \frac{1}{\sqrt{3}} = 57\%$$

och för motsvarigheten till normalfördelningens 95%-iga konfidensintervall får vi för $\mu_U \pm 1.96\sigma$:

$$\frac{3.92 \cdot \sigma_U}{b-a} = \frac{3.92}{2\sqrt{3}} = 113\%,$$

dvs. hela fördelningen täcks av plus/minus två standardavvikelser.

2.2.1 Räkningregler för väntevärden och varianser

Eftersom väntevärdet är definierat som en integral, vilket i sin tur kan ses som en summa (se approximationen 2.3) så är också väntevärdet av linjärkombinationer av stokastiska variabler lätta att räkna ut:

$$E[aX + bY] = E[aX] + E[bY] = aE[X] + bE[Y] \quad (2.7)$$

Motsvarande regel gäller däremot inte för medianer.

För spridningsmättet gäller att variansen är definierad som integral (2.2) och följaktligen adderas varianser och inte standardavvikelser,

$$\text{Var}[aX + bY] = \text{Var}[aX] + \text{Var}[bY] + 2\text{Kov}[aX, bY] = \quad (2.8)$$

$$a^2 \text{Var}[X] + b^2 \text{Var}[Y] + 2ab\text{Kov}[X, Y]$$

Konstanter kvadreras och det tillkommer en term som beskriver samvariationen mellan de båda stokastiska variablerna, **kovariansen**. Skattningsformeln för denna storhet är om man har m st observationspar (X_i, Y_i) :

$$\widehat{\text{Kov}}[X, Y] = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})(Y_i - \bar{Y}) \quad (2.9)$$

Om de båda stokastiska variablerna samverkar, dvs har en tendens att vara stora respektive små samtidigt så blir denna summa positiv, medan den blir negativ om variablerna tenderar att motverka varandra. Om de är oberoende tar de olika bidragen ut varandra och summan tenderar att bli lika med noll.

Med dessa regler kan vi t. ex. räkna ut variansen för medelvärdet av ett antal oberoende observationer på en stokastisk variabel med variansen σ^2 :

$$\text{Var} [\bar{X}] = \text{Var} \left[\frac{1}{n} \sum_{i=1}^n X_i \right] = \frac{1}{n^2} \sum_{i=1}^n \text{Var} [X_i] = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n} \quad (2.10)$$

vilket är den matematiska formuleringen av den första viktiga metoden för förbättrad mätnoggrannhet som anges på sidan 10. Alla kovarianstermer blir lika med noll då vi antar att de olika observationerna är oberoende.

Dessa räkneregler kan också användas för att göra en praktisk omskrivning av normalfördelningen. Eftersom normalfördelningens parametrar är lika med dess väntevärde resp. standardavvikelse så får man:

$$\text{Om } X \sim N(\mu, \sigma) \text{ så gäller att } X - \mu \sim N(0, \sigma)$$

Om $X \sim N(\mu, \sigma)$ så gäller att

$$\frac{X - \mu}{\sigma} \sim N(0, 1) . \quad (2.11)$$

Dessutom kan man visa att summan av ett antal normalfördelade variabler också är normalfördelad, vilket innebär att man lätt kan räkna ut fördelningen för varje linjärkombination av normalvariabler.

Dessa egenskaper hos normalfördelningen gör den mycket enkel att använda vid mätosäkerhetsanalys och det visar sig också att den oftast är en utmärkt approximation till verkliga mätfel. Detta beror på att mätfel oftast är en kombination av flera felkällor, vilket gör summan mer normalfördelad. Ytterligare en anledning till normalapproximationens användbarhet i mätosäkerhetssammanhang är att man här betraktar 95%-iga konfidensintervall och inte kräver noggrannhet i fördelningens mer extrema svansar.

I nästa avsnitt beskrivs hur man använder mätosäkerheten för att avgöra om skillnader mellan mätvärden endast är orsakade av slumpen eller avspeglar en systematisk skillnad mellan mätobjekten. I enlighet med ovanstående resonemang utgår vi ifrån att mätfelen är normalfördelade.

2.3 Jämförelse av mätvärden

2.3.1 Variansen är lika och känd

Jämför två mätvärden

Vi betraktar situationen att vi vill jämföra två objekt med avseende på någon egenskap som vi kan mäta. Varje objekt mäts en gång och ger resultaten X resp. Y . Objektens egenskaper jämförs genom att betrakta differensen mellan mätvärdena, $X - Y$. Vi antar först att mätosäkerheten är densamma

för de båda objekten, samt att vi känner dess standardosäkerhet, u . Vi antar vidare att mätfelet är normalfördelat med väntevärdet noll och variansen $\sigma^2 = u^2$.

Väntevärdet för den intressanta differensen är enligt räkneregel (2.7):

$$E[X - Y] = E[X] - E[Y]$$

och variansen för differensen enligt (2.8),

$$\text{Var}[X - Y] = \text{Var}[X] + \text{Var}[Y] = \sigma^2 + \sigma^2 = 2\sigma^2$$

där vi antagit att de båda mätningarna är okorrelerade, dvs. kovariansen är lika med noll. Med hjälp av (2.11) bildar vi följande normaliserade variabel,

$$Z_0 = \frac{X - Y - (E[X] - E[Y])}{\sqrt{2}\sigma} \sim N(0, 1) . \quad (2.12)$$

En sådan normaliserad variabel kallas ibland för **statistika**. Vi vill nu kontrollera om den uppmätta differensen $X - Y$ är statistiskt signifikant, dvs. att den troligtvis inte beror på slumpen, vilket representerar mätosäkerheten i detta fall. Detta kan göras genom att beräkna ett **konfidensintervall** för den uppmätta differensen. Om konfidensintervallet täcker nollan så betraktar vi inte differensen som signifikant, utan konstaterar att de båda väntevärdena kan vara lika.

Eftersom vi känner fördelningen för den normaliserade variabeln Z_0 så kan man bestämma ett tal z_0 som uppfyller

$$P(-z_0 \leq Z_0 \leq z_0) = p \quad (2.13)$$

där p är en lämpligt vald stor sannolikhet, t. ex. 95%. Om vi väljer $p = 0.95$ så söker vi det intervall som är symmetriskt kring noll och som täcker 95% av fördelningen. Det innebär att 2.5% av fördelningen skall understiga den undre gränsen och i en normalfördelningstabell finner man att sannolikheten 0.024998 motsvarar värdet -1.96 . Alltså kommer valet av $z_0 = 1.96$ att ge ett 95%-igt konfidensintervall,

$$P(-1.96 \leq Z_0 \leq 1.96) \approx 0.95$$

Genom att nu sätta in uttrycket för Z_0 (2.12) i (2.13) så får vi följande resultat,

$$X - Y - 1.96\sqrt{2}\sigma \leq E[X] - E[Y] \leq X - Y + 1.96\sqrt{2}\sigma$$

eller enklare

$$E[X] - E[Y] = X - Y \pm 1.96\sqrt{2}\sigma \quad (\text{med konfidensen } 95\%) \quad (2.14)$$

Om detta konfidensintervall nu inte täcker nollan, så betyder det att den verkliga differensen mellan väntevärdena är signifikant skild från noll med **signifikansnivån** 5%. Om det däremot täcker nollan så **kan** den uppmätta differensen vara ett utslag av mätosäkerheten och man kan inte uttala sig om det finns någon skillnad eller inte.

Problem 2.3.1 *Två stålrör skall jämföras med avseende på deras ytterdiametrar. Rören mäts med skjutmått till 37.22 resp. 37.31 mm. Mätosäkerheten vid skjutmåttmätning under de rådande förhållandena har tidigare bestämts till $u = 0.05$ mm. Bestäm ett konfidensintervall för differensen av diametrarna och svara på frågan: Har rören signifikant olika diametrar?*

Problem 2.3.2 *Kontrollera vilken sannolikhet som motsvaras av $z_0 = 2$ i (2.13) och diskutera vad en sådan avrundning av $z_0 = 1.96$ egentligen innebär. Använd en normalfördelningstabell eller något datorprogram. I Excel (2002) finns normalfördelningen representerad i funktion/statistik/normsförd.*

Förbättring av noggrannhet genom upprepade mätningar Om ett konfidensintervall täcker nollan kan ingen statistiskt signifikant skillnad konstateras. Man kan dock inte vara säker på att ingen skillnad föreligger; en noggrannare undersökning skulle kunna visa annorlunda resultat. En sådan mer precis undersökning kan genomföras genom att minska standardavvikelsen σ . Detta kan göras på två sätt i mätsammanhang, antingen genom att förbättra mätsystemet och därmed minska u , eller genom att genomföra upprepade mätningar och jämföra medelvärdena. Notera emellertid att detta förutsätter att mätosäkerheten är av slumpmässig natur. Om n upprepade mätningar görs på X med mätosäkerheten u så blir variansen för medelvärdet enligt (2.10)

$$\sigma_{\bar{X}}^2 = \frac{u^2}{n}$$

och vid m upprepningar av Y -mätningar blir

$$\sigma_{\bar{Y}}^2 = \frac{u^2}{m}$$

Variansen för differensen blir

$$\text{Var} [\bar{X} - \bar{Y}] = \text{Var} [\bar{X}] + \text{Var} [\bar{Y}] = \frac{u^2}{n} + \frac{u^2}{m} = u^2 \left(\frac{1}{n} + \frac{1}{m} \right)$$

och

$$E[X] - E[Y] = \bar{X} - \bar{Y} \pm 1.96\sigma \sqrt{\frac{1}{n} + \frac{1}{m}} \quad (\text{med konfidensen } 95\%) \quad . \quad (2.15)$$

Om mätosäkerheten vid varje mätning är given och känd som i ovanstående fall ges precisionen i jämförelsen alltså helt av antalet upprepningar. Man kan därmed lägga upp ett test så att en föreskriven minsta skillnad E kan detekteras med en önskad signifikans. Detta görs genom att välja antalet upprepningar så att

$$1.96\sigma\sqrt{\frac{1}{n} + \frac{1}{m}} \leq E$$

dvs.

$$\sqrt{\frac{1}{n} + \frac{1}{m}} \leq \frac{E}{1.96\sigma}$$

vilket i det enklare fallet då $n = m$ ger följande kriterium

$$n \geq 7.8 \left(\frac{\sigma}{E}\right)^2$$

Problem 2.3.3 *I exemplet från problem 2.3.1: Anta att det första röret mätts vid ett tidigare tillfälle med åtta upprepningar. Hur många mätningar behöver man göra av det andra röret för att kunna detektera en skillnad som är 0.06 mm med 95% konfidens?*

Jämför väntevärde med gränsvärde

En liknande situation som ovanstående uppstår då man provar mot godkännande med avseende på något gränsvärde g . Vi antar igen att ett objekt provas eller mäts så att mätosäkerheten bidrar till slumpmässighet i resultatet. Vi har alltså resultatet X och antar att vi känner variansen σ^2 . Vi bestämmer då ett konfidensintervall för mätvärdet X och kontrollerar om detta täcker det aktuella gränsvärdet.

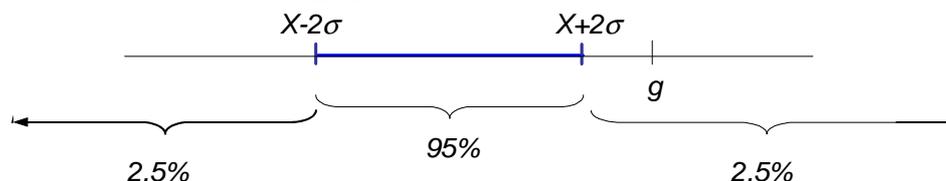
Vi bildar följande normaliserade variabel

$$Z_0 = \frac{X - E[X]}{\sigma} \sim N(0, 1)$$

och får konfidensintervallet

$$E[X] = X \pm 1.96\sigma \quad (\text{med konfidensen } 95\%) \quad (2.16)$$

Jämförelsen med gränsvärdet g kan beskrivas grafiskt, vi antar att kravet gäller att X inte får överstiga gränsen¹:



¹I illustrationen har den exakta gränsen 1.96 ersatts av det avrundade värdet 2. Vi skall senare allmänt införa denna avrundning.

Vi ser att om hela konfidensintervallet ligger under gränsvärdet så är sannolikheten för att det sanna värdet överstiger gränsen högst 2.5%, dvs. konfidensgraden blir i ett sådant ensidigt fall högre än i det allmänna differensfallet (2.14). Detta beror på att vi mycket väl kan tillåta det sanna värdet att ligga under den lägre konfidensgränsen och att detta undre intervall därmed kan inkluderas i konfidensområdet.

Problem 2.3.4 *Ett stålrör skall jämföras mot en minsta diametertolerans som är 37.15 mm. Röret mäts med skjutmått fyra gånger och medelvärdet beräknas till 37.22 mm. Mätosäkerheten vid skjutmåttsmätning är som i tidigare exempel $\sigma = 0.05$ mm. Bestäm ett 95%-igt konfidensintervall för diametern och svara på frågan: Är röret godkänt på signifikansnivån 2.5%?*

2.3.2 Variansen är lika men okänd

I ovanstående avsnitt antog vi att mätosäkerheten var känd, $\sigma = u$. Ofta är emellertid mätosäkerheten skattad ur en begränsad mängd mätningar eller bedömningar och den är själv osäker till sin storlek. För att i sådana fall få ett konfidensintervall att täcka 95% av mätfallen måste intervallet göras något större för att kompensera för osäkerheten i u . I fall med normalfördelade fel och osäkerhetsskattningar genom statistiska försök har man faktiskt full kontroll över detta med hjälp av t -fördelningen som behandlas härnäst.

Jämför två väntevärden

Den tidigare härledningen av väntevärde och varians för skillnaden i väntevärden utförs på liknande sätt när variansen är okänd och statistikan i detta fall blir snarlik den tidigare angivna (2.12)

$$T = \frac{X - Y - (E[X] - E[Y])}{\sqrt{2}S_p} \sim t(\nu) \quad (2.17)$$

Standardavvikelsen σ är här ersatt av den skattade motsvarigheten S_p som är den sammanvägda standardavvikelsen, skattad ur upprepade mätningar på samma objekt. Statistikan kallas här för T och den är nu inte längre normalfördelad utan följer en annan parametrisk fördelning, nämligen t -fördelningen. Parametern ν motsvaras av antalet frihetsgrader i skattningen av S_p vilket kommer att behandlas i det följande. Vi antar först att mätosäkerheten är densamma för de båda populationerna men eftersom väntevärdena kan vara olika så kan inte formeln (2.4) användas. Stickprovsvariansen antas ha bestämts för varje mätsystem för sig och sammanvägs sedan till en gemensam skattning. Anta att man i tidigare referensmätningar gjort n upprepade mätningar på x och m mätningar på y . De två mätsystemens respektive mätosäkerheter skattas var för sig enligt:

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, \quad S_Y^2 = \frac{1}{m-1} \sum_{i=1}^m (Y_i - \bar{Y})^2$$

och eftersom vi här antar att de verkliga varianserna är lika kan de båda skattningarna sammanvägas till en skattning av den gemensamma variansen:

$$S_p^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2} \quad (2.18)$$

Man viktar alltså varje enskild skattning med dess antal frihetsgrader vid sammanvägningen och slutresultatets antal frihetsgrader är summan av stickproven minus två, dvs vi har förlorat en frihetsgrad per medelvärdesskattning. Konfidensintervall för differensen mellan väntevärdena bildas nu helt i enlighet med (2.14), men med den sammanvägda stickprovsvariansen och kvantil $t_{0.025}$ från t -fördelningen.

$$E[X] - E[Y] = X - Y \pm \sqrt{2}S_p t_{0.025} \quad (\text{med konfidensen } 0.95) \quad (2.19)$$

I följande tabell anges 2.5%-kvantilerna från t -fördelningen vid olika antal frihetsgrader².

Tabell 2.3.2

frihetsgrader	2	3	4	5	6	7	8	9	10	11	18	30
$t_{0.025}$	4.3	3.2	2.8	2.6	2.4	2.4	2.3	2.3	2.2	2.2	2.1	2.0

Notera att vid 30 frihetsgrader är t -faktorn i det närmaste densamma som i normalfördelningsfallet, dvs. variansskattningen kan betraktas som säker i sådant fall.

Problem 2.3.5 *Stålrör från två olika tillverkningsskift (A resp. B) skall jämföras med avseende på deras ytterdiameter. Åtta rör från vardera skift mäts med skjutmått med följande resultat i mm:*

A	37.33	37.32	37.38	37.32	37.37	37.13	37.25	37.23
B	37.09	37.28	37.14	37.39	37.17	37.30	37.27	37.16

Bestäm ett konfidensintervall för differensen av diametrarna och svara på frågan: Har rören signifikant olika diametrar?

2.3.3 Varianserna är olika och okända

Ibland är varianserna i två populationer inte alls lika och då fallerar de hittills visade metoderna för jämförelser av väntevärden. Den naturliga normaliserade variabeln i detta fall är

$$T^* = \frac{X - Y - (E[X] - E[Y])}{\sqrt{S_1^2 + S_2^2}} \sim t(\nu) \quad (2.20)$$

²Excel (2002) ger dessa kvantiler med kommandot TINV(0.05, ν).

Denna statistika har ingen känd fördelning, men är approximativt t -fördelad. Speciellt måste antalet frihetsgrader bestämmas. Detta kan göras enligt Welch-Satterwaites approximationsformel

$$\nu_{eq} = \frac{(S_1^2 + S_2^2)^2}{\frac{S_1^4}{\nu_1} + \frac{S_2^4}{\nu_2}} \quad (2.21)$$

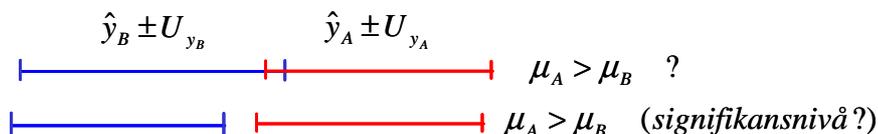
och konfidensintervallet blir i detta fall

$$E[X] - E[Y] = X - Y \pm t_{0.025} \sqrt{S_1^2 + S_2^2} \quad (\text{med konfidens } 95\%), \quad (2.22)$$

där $t_{0.025}$ då tas ur tabell med antalet frihetsgrader lika med ν_{eq} .

2.3.4 Endast konfidensintervallen är kända

I många fall har man själv inte kontroll över mätosäkerhetsberäkningarna, men vill ändå jämföra någon mätt egenskap mellan olika produkter, säg A och B . Om respektive produkttegenskap y är angiven med ett mätosäkerhetsintervall enligt standardmetodik betyder det att man har tillgång till de utvidgade mätosäkerheterna U_{y_A} och U_{y_B} och det enklaste underlaget för att avgöra om skillnaden är signifikant bör vara att kontrollera om konfidensintervallen täcker varandra eller inte. Situationen illustreras nedan



Vi skall nu beräkna vilken signifikansnivå man får vid en sådan förenklad metodik. Anta först att båda produkternas mätosäkerheter är bestämda med ett så stort antal frihetsgrader att normalfördelningsapproximationen har kunnat användas vid bestämning av U_{y_A} och U_{y_B} . Vi utgår från att vi endast intresserar oss för händelsen att A är större än B och skall nu beräkna sannolikheten för att konfidensintervallen precis sammanfaller i sina respektive yttervärden, givet att väntevärdena är lika. Gränsfallet betyder att vi observerat följande situation:

$$\begin{aligned} \hat{y}_B + U_{y_B} &= \hat{y}_A - U_{y_A} \\ \hat{y}_B + 1.96 \cdot u_{y_B} &= \hat{y}_A - 1.96 \cdot u_{y_A} \end{aligned} \quad (2.23)$$

Detta gränsfall illustreras här:



Om vi antar att egenskaperna är lika så gäller att skillnaden mellan väntevärdena är noll:

$$E[\hat{y}_A] - E[\hat{y}_B] = 0$$

I enlighet med ovanstående teori skall man då betrakta följande normalfördelade statistika

$$Z = \frac{\hat{y}_A - \hat{y}_B}{\sqrt{u_A^2 + u_B^2}} \sim N(0, 1)$$

I gränsfallet med sammanfallande konfidensgränser (2.23) blir denna statistika:

$$Z_{\text{lim}} = \frac{\hat{y}_A - \hat{y}_B}{\sqrt{u_A^2 + u_B^2}} = \frac{1.96(u_A + u_B)}{\sqrt{u_A^2 + u_B^2}}.$$

För att förenkla detta uttryck antar vi härnäst att mätosäkerheterna i de båda fallen är proportionella med proportionalitetskonstanten c , $u_{y_A} = c \cdot u_{y_B}$

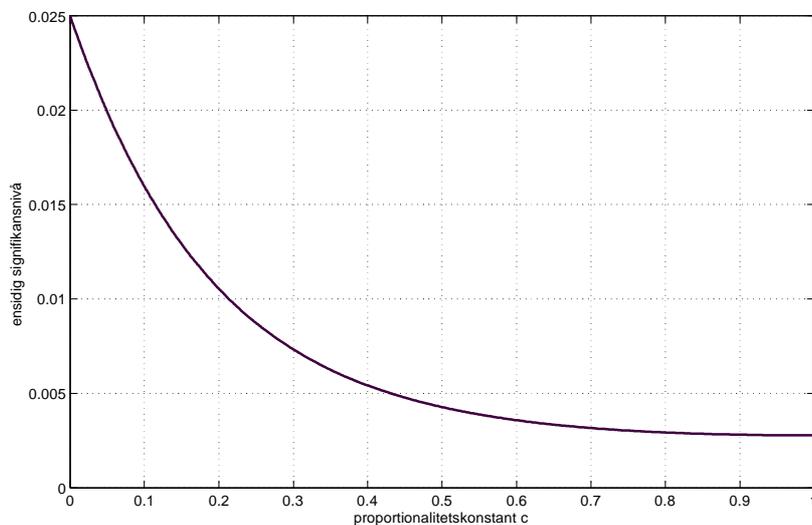
$$Z_{\text{lim}} = \frac{1.96(u_A + u_B)}{\sqrt{u_A^2 + u_B^2}} = \frac{1.96u_B(c + 1)}{u_B\sqrt{c^2 + 1}} = \frac{1.96(c + 1)}{\sqrt{c^2 + 1}} = \begin{cases} 1.96\sqrt{2} & \text{för } c = 1 \\ 1.96 & \text{för } c = 0 \end{cases}$$

I fallet med $c = 0$ har y_A ingen varians, dvs. den är en bestämd gräns och vi ser att resultatet sammanfaller med resultatet (2.16). Vi har därmed en signifikansnivå på 2.5%, eftersom vi endast ser till det ensidiga fallet precis som i (2.16). I fallet med $c = 1$ är de båda mätosäkerheterna lika stora och vi får en mycket lägre signifikansnivå, nämligen 0.2 %. I figur 2.6 illustreras signifikansnivån för olika proportionalitetskonstanter. Figuren visar att den verkliga signifikansnivån beror kraftigt på relationen mellan de båda osäkerheterna. En omedelbar slutsats tycks vara följande: Om konfidensintervallen inte täcker varandra och är av ungefär samma storlek så är signifikansnivån högst 0.5%, dvs. det är bara 0.5% sannolikhet för att mätosäkerheten skulle ge ett så extremt resultat vid lika egenskaper. Notera dock att detta gäller om antalet frihetsgrader är stort.

Vi antog i denna första beräkning att antalet frihetsgrader var tillräckligt stort för att använda normalfördelningsapproximationen. Men, om endast konfidensintervallen är kända så vet man ingenting om den bakomliggande beräkningen, som mycket väl kan vara grundad på ett fåtal försök i kombination med t -fördelningen. Om inte dessa förhållanden är kända så är en möjlig praktisk metod för att avgöra skillnader att ändå anta normalfördelning och därmed skatta varje standardosäkerhet som halva den utvidgade osäkerheten, $u = \frac{U}{2}$, och använda standardmetodiken (2.22).

$$E[y_A] - E[y_B] = \hat{y}_A - \hat{y}_B \pm 1.96\sqrt{\frac{U_{y_A}^2}{2^2} + \frac{U_{y_B}^2}{2^2}}. \quad (2.24)$$

Hur bra blir denna approximation? Det finns inte något enkelt sätt att bestämma detta matematiskt, men man kan kontrollera genom att simulera tänkbara situationer, vilket vi gjort för att få en grund för en tumregel. Följande procedur har genomförts: För standardavvikelse som skiljer sig



Figur 2.6: Signifikansnivå för detekterad skillnad vid jämförelse av två konfidensintervall som sammanfaller i övre resp. nedre gräns. Angiven som funktion av relationen mellan konfidensintervallens storlek.

maximalt en faktor 5 har vi gjort följande: 1) simulering av n_A A -mätningar och skattning av standardavvikelsen ger ett simulerat u_A , 2) simulering av n_B B -mätningar och skattning av standardavvikelsen ger ett simulerat u_B , 3) beräkning av konfidensintervall för differensen mellan en A -mätning och en B -mätning bildas enligt (2.24). 4) Simulering av 10000 nya mätningar A och B ger 10000 differenser och 5) kontroll av hur stor andel av dessa differenser som överskrider konfidensintervallets övre gräns ger den verkliga signifikansnivån. Olika kombinationer av n_A och n_B , båda i intervallet $[3, 40]$ ger resultat som visar att den verkliga signifikansnivån varierar för olika kombinationer av frihetsgrader och osäkerheter, men att den hamnar i intervallet $[0.015\%, 0.035\%]$, vilket motsvarar den acceptabla approximationsnivå som behandlas i avsnitt 3.2. Slutsatsen är att approximationen (2.24) tycks vara acceptabel om de angivna konfidensintervallen ursprungligen är beräknade med hjälp av t -fördelningen. Den kan alltså användas för att bestämma konfidensintervall för skillnader utifrån konfidensintervallen för respektive mätvärde. Speciellt kan man i fall med icke överlappande konfidensintervall konstatera att skillnad föreligger med mycket stor sannolikhet, nämligen med minst 99.5 % sannolikhet om intervallen är ungefär lika stora.

Tumregel 1

Vid jämförelse av två mätvärden som är angivna med 95%-iga mätosäkerhetsintervall fås ett approximativt 95%-igt konfidensintervall för deras differens genom formeln

$$y_A - y_B = \hat{y}_A - \hat{y}_B \pm 2\sqrt{\frac{U_{y_A}^2}{4} + \frac{U_{y_B}^2}{4}}$$

där täckningsfaktorn 1.96 i 2.24 har ersatts med 2 i enlighet med konventionen som presenteras nästa kapitel.

Tumregel 2

Vid jämförelse av två mätvärden som är angivna med 95%-iga mätosäkerhetsintervall av ungefär samma storlek gäller:

Det större mätvärdet representerar ett verkligt större väntevärde med minst konfidensen 99.5% om intervallen inte överlappar.

2.4 Sammanvägd standardavvikelse

Den sammanvägda standardavvikelsen som bestämdes för två stokastiska variabler i (2.18) kan generaliseras till ett godtyckligt antal stokastiska variabler. Detta är en viktig metod i mätosäkerhetssammanhang eftersom man ofta kan anta att mätfelen har samma varians för mätningar av objekt med olika nominella mått. För att skatta en sådan gemensamma varians måste man då skatta avvikelserna från respektive nominella mått för sig och sedan väga samman resultaten till en gemensam standardavvikelse. I en sådan sammanvägning är det förstås naturligt att lägga större vikt vid skattningar grundade på flera mätningar och sammanvägningsformeln för mätningar av m stycken objekt blir:

$$s_{SV}^2 = \frac{\nu_1 s_1^2 + \nu_2 s_2^2 + \dots + \nu_m s_m^2}{\nu_1 + \nu_2 + \dots + \nu_m}, \quad (2.25)$$

där $\nu_1, \nu_2, \dots, \nu_m$ betecknar antalet frihetsgrader i respektive s -skattning och antalet frihetsgrader i sammanvägningen blir det som ges i nämnaren, nämligen summan av antalet enskilda frihetsgrader.

I fallet med n_j upprepade mätningar på ett objekt med okänt nominellt värde blir $\nu_j = n_j - 1$. Skattningen av respektive standardavvikelse är som

tidigare:

$$s_j = \sqrt{\frac{1}{n_j - 1} \sum_{i=1}^{n_j} (x_i - \bar{x})^2}.$$

Problem 2.4.1 *Hur ser sammanvägningsformeln ut om lika många mätningar gjorts på varje variabel?*

I andra fall kan de olika ingående standardavvikelserna vara skattade på något annat sätt och antalet frihetsgrader bestäms enligt just den skattningen. Som vi skall se i nästa avsnitt om linjär regression så kommer standardavvikelsen för mätfelet där att ha endast $n - 2$ frihetsgrader.

2.5 Linjär regression

En statistisk metodik som är mycket användbar i kalibreringssammanhang är linjär regression. Allmänt behandlar den situationer när en företeelse kan ses som en linjärkombination av ett antal kända variabler, samt en slumpmässig störning. I kalibreringssammanhang är situationen typiskt sådan att man har ett instrument som skall kalibreras inom ett givet mätområde och dess utsignal ses som en linjär funktion av referensens signal störd av ett slumpmässigt mätfel; ett mätbrus. Vi skall i ett senare kapitel behandla denna kalibreringssituation, men behandlar först kort den allmänna statistiska teorin bakom linjär regression.

Antag att det finns ett linjärt samband mellan storheterna x och y , där x antas vara känd

$$y = \beta_0 + \beta_1 x$$

Linjär regression går ut på att bestämma koefficienterna β_0 och β_1 ur ett antal mätta värden. Vid observationer av y fås en summa av det verkliga värdet $y(x)$ och en slumpmässig störning ε ,

$$Y(x) = \beta_0 + \beta_1 x + \varepsilon \quad (2.26)$$

Vi antar nu att vi gjort n observationer av Y vid olika x -värden. Vi har alltså tillgång till tal-paren

$$(x_1, Y_1), (x_2, Y_2) \dots (x_n, Y_n)$$

Störningen modelleras som en normalfördelad slumpvariabel med väntevärdet noll och en viss okänd varians, $\varepsilon \sim N(0, \sigma^2)$. Regressionen går nu ut på att bestämma uppskattningar $\hat{\beta}_0$ och $\hat{\beta}_1$ av koefficienterna β_0 och β_1 så att felet, dvs avståndet mellan den uppskattade räta linjen och de uppmätta värdena, blir så litet som möjligt. Det vanligaste sättet att genomföra detta

är att definiera felet E , som skall minimeras, som summan av kvadraterna på avstånden till den räta linjen,

$$E = \sum_{i=1}^n \left(Y_i - [\hat{\beta}_0 + \hat{\beta}_1 x_i] \right)^2$$

och regressionen kallas då för minsta-kvadratmetoden. Lösningen ges av

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n Y_i (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}$$

där \bar{x} och \bar{Y} anger medelvärdena av de använda x -värdena resp. observerade Y -värdena.

För att bestämma egenskaperna hos störningen ε kan man bilda s.k. residualer, dvs differenserna mellan de observerade Y -värdena och motsvarande värden från den linjära anpassningen.

$$e_i = Y_i - [\hat{\beta}_0 + \hat{\beta}_1 x_i] \quad (2.27)$$

Genom att studera dessa residualer kan man bedöma rimligheten i modellantagandet (2.26), t. ex. : Är sambandet linjärt och är störningarna oberoende och normalfördelade?

2.5.1 Konfidensintervall för linjen

Om modellen bedöms som rimlig kan störningens varians skattas från residualerna,

$$s^2 = \frac{1}{n-2} \sum_{i=1}^n e_i^2 = \frac{1}{n-2} \sum_{i=1}^n \left(Y_i - [\hat{\beta}_0 + \hat{\beta}_1 x_i] \right)^2, \quad (2.28)$$

där $n-2$ i nämnaren anger antalet frihetsgrader i skattningen, som är lika med antalet observationer i summan minus antalet parametrar som skattats ur samma datamaterial. Med hjälp av denna skattade varians kan man beräkna ett antal olika intressanta storheter, nämligen osäkerheten i de skattade koefficienterna, osäkerheten i den skattade regressionslinjen, samt prediktionsosäkerheten vid en ny beräknad Y för ett givet x .

Vi beräknar först varianserna för de skattade koefficienterna,

$$\begin{aligned} \text{Var} [\hat{\beta}_1] &= \text{Var} \left[\frac{\sum_{i=1}^n Y_i (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] = \frac{\text{Var} [\sum_{i=1}^n Y_i (x_i - \bar{x})]}{\left(\sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} = \\ &= \frac{\sum_{i=1}^n \text{Var} [Y_i (x_i - \bar{x})]}{\left(\sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \text{Var} [Y_i]}{\left(\sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} = \end{aligned}$$

$$\frac{\sigma^2 \sum_{i=1}^n (x_i - \bar{x})^2}{\left(\sum_{i=1}^n (x_i - \bar{x})^2\right)^2} = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \approx \frac{s^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.29)$$

eftersom variansen för Y_i är densamma som för ε_i , nämligen σ^2 . Variansen för lutningskoefficienten minskar om summan i nämnaren ökar. Notera att detta kan ske på två sätt, antingen genom att öka antalet mätpunkter eller att öka spridningen av mätpunkterna.

Skattningen av den andra koefficienten har variansen:

$$\begin{aligned} \text{Var} [\hat{\beta}_0] &= \text{Var} [\bar{Y} - \hat{\beta}_1 \bar{x}] = \text{Var} \left[\frac{1}{n} \sum_{i=1}^n Y_i \right] + \text{Var} [\hat{\beta}_1 \bar{x}] = \\ &= \frac{1}{n^2} \text{Var} \left[\sum_{i=1}^n Y_i \right] + \bar{x}^2 \text{Var} [\hat{\beta}_1] = \frac{1}{n^2} \sum_{i=1}^n \text{Var} [Y_i] + \frac{\bar{x}^2 \sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \\ &= \frac{n\sigma^2}{n^2} + \frac{\bar{x}^2 \sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \approx s^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) \end{aligned}$$

där det andra likhetstecknet gäller eftersom kovariansen mellan \bar{Y} och $\hat{\beta}_1 \bar{x}$ visar sig vara lika med noll.

Vi ser att den första termen i variansen för denna koefficient är omvänt proportionell mot antalet mätvärden i regressionen medan den andra liknar variansuttrycket för $\hat{\beta}_1$.

På grundval av dessa resultat beräknas ett konfidensintervall för den skattade linjen,

$$\hat{y}(x) = \hat{\beta}_0 + \hat{\beta}_1 x$$

Variansen för denna skattning blir genom summering av varianserna av parameterskattningarna och deras kovarians

$$\text{Var} [\hat{y}(x)] \approx s^2 \left(\frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{1}{n} \right) \quad (2.30)$$

Notera att variansen för linjen är summan av två termer, den andra är konstant över hela linjen, medan den första beror på x , dvs var på linjen man befinner sig. Ju närmare regressionsmedelvärdet desto mindre varians.

2.5.2 Prediktionsintervall för linjen

Vid användning av linjär regression för kalibrering intresserar vi oss för mätosäkerheten i nya avläsningar. Detta betyder att vi vill veta variansen för ett predikerat värde, y^* . Gör man en prediktion av y^* från ett givet x och den skattade linjen så gäller för detta

$$y^*(x) = \hat{\beta}_0 + \hat{\beta}_1 x + \varepsilon = \hat{y}(x) + \varepsilon$$

och variansen blir,

$$\begin{aligned} \text{Var} [y^*(x)] &= \text{Var} [\hat{y}(x) + \varepsilon] = \\ \text{Var} [\hat{y}(x)] + \text{Var} [\varepsilon] &= \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \sigma^2 + \frac{\sigma^2}{n} + \sigma^2 \approx \\ s^2 &\left(\frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{1}{n} + 1 \right) \end{aligned}$$

Variansen för den nya slumpmässiga störningen ε måste alltså tas med i beräkningen. Ett prediktionsintervall för ett predikerat $y^*(x)$ blir

$$y^*(x) = \hat{\beta}_0 + \hat{\beta}_1 x \pm s \cdot t \sqrt{\frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{1}{n} + 1} \quad (2.31)$$

där t är tagen från t -fördelningen med $n - 2$ frihetsgrader. Notera att om antalet avläsningar vid regressionen är stort så domineras variansen för prediktionen av den nya störningen och de första två termerna får litet inflytande. Detta ger en grundval för en förenklad approximativ skattning av prediktionsintervallet.

2.5.3 Approximativt prediktionsintervall

Beräkning av prediktionsintervall enligt formeln (2.31) kräver kunskap, inte bara om den skattade variansen för slumpfelet, utan också om samtliga referensvärden som använts i referensprovet och är ibland allför komplicerat för behovet. Om antalet referensförsök är tillräckligt stort så kommer de två första termerna i rotuttrycket att vara försumbara och rotuttrycket reduceras till en etta. För att avgöra hur många referensförsök som krävs för en sådan förenkling krävs en approximationsregel, vilken vi återkommer till i kalibreringskapitlet.

2.5.4 Bantningsexemplet

I det inledande bantningsexemplet, avsnitt 1.1, kan man skönja en trend av minskande vikt. Vi skall studera denna trend genom att modellera den som en rät linje. Då vågens slumpmässiga variation inte innehåller någon information om trenden, så är det fördelaktigt att minimera denna variation genom att göra regressionen på det dagliga medelvärdet. Modellen blir då följande:

$$\bar{m}(t) = \beta_0 + \beta_1 (t - t_0) + \varepsilon ,$$

där $\bar{m}(t)$ alltså anger det dagliga medelvärdet för dagen t , t_0 anger den inledande dagen och ε är det slumpmässiga bidraget. En minsta kvadratanpassning enligt ovanstående metod med hjälp av de 31 dagsmedelvärdena ger skattningarna

$$\hat{\beta}_0 = 91.6 , \quad \hat{\beta}_1 = -0.0412 .$$

Utgångsvikten skattas alltså till 91.6 kg och den genomsnittliga viktminskningen till 41 g/dag, vilket innebär ca 1.2 kg/månad. Variansen för slumpbidraget ε skattas genom formeln (2.28) till

$$s^2 = \frac{1}{29} \sum_{i=1}^{31} (\bar{m}(t_i) - [91.6 - 0.0412(t_i - t_0)])^2 = 0.225 .$$

Den anpassade linjen ses i figur 2.7 tillsammans med konfidensintervall och prediktionsintervall beräknade med (2.30) och (2.31). Residualerna i regressionen, beräknade enligt (2.27), representerar de slumpmässiga bidragen och om modellen är rimlig bör dessa vara slumpmässigt fördelade över de olika dagarna. I figur 2.8 ses en plot av dessa residualer, vilken inte antyder någon systematik och därmed kan ses som en verifikation av modellantagandet.

Det kan i detta fall vara intressant att studera ett konfidensintervall för viktminskningen, vilket lätt kan beräknas med hjälp av (2.29):

$$Var[\hat{\beta}_1] = \frac{s^2}{\sum_{i=1}^n (t_i - \bar{t})^2} = \frac{0.225}{4599} = 4.89 \cdot 10^{-5} .$$

Ett 95%-igt konfidensintervall för viktminskningen blir

$$\begin{aligned} \beta_1 &= \hat{\beta}_1 \pm ts = -0.0413 \pm 2\sqrt{4.89 \cdot 10^{-5}} \\ &= -0.0413 \pm 0.014 \text{ kg/dag, (med 95\% konfidens),} \end{aligned}$$

där vi använt $t = 2$ eftersom antalet frihetsgrader i variansskattningen är 29.



Konfidensintervallet täcker inte nollan och man kan därmed betrakta minskningen som signifikant. Denna slutsats förstärks vid studier av konfidensintervallet i figur 2.7, som beskriver de gränser som innehåller den sanna linjen med 95 % sannolikhet.

Regressionen utförs lätt i Excel genom att använda “dataanalys”, “regression”. Värdena för x resp. y placeras i varsin kolumn i Excelbladet och i dialogrutan för regression anges dessa kolumner. Kryssrutorna “Etiketter”, “Konfidensnivå” och “Konstanten är noll” lämnas ofyllda. Genom att kryssa i lämpliga rutor av de övriga alternativen kan man bland annat få anpassningen illustrerad, samt få residualerna utskrivna och plottade (Microsoft Excel 2002). Resultatet ses i nedanstående tabeller, där skattningen av s^2 återfinns i ANOVA-tabellen som MKv för Residual. Standardavvikelsen för lutningen ges i den nedre tabellen som standardfel för X-variabel, vilket är precis roten ur den varians som angavs ovan:

$$\sqrt{Var[\hat{\beta}_1]} = \sqrt{4.89 \cdot 10^{-5}} = 0.00699 .$$

Konfidsensintervallet för denna koefficient anges i nedre tabellens sista två kolumner.

UTDATASAMMANFATTNING

<i>Regressionsstatistik</i>	
Multipel-R	0.738220685
R-kvadrat	0.54496978
Justerad R-kvadrat	0.529279083
Standardfel	0.474635911
Observationer	31

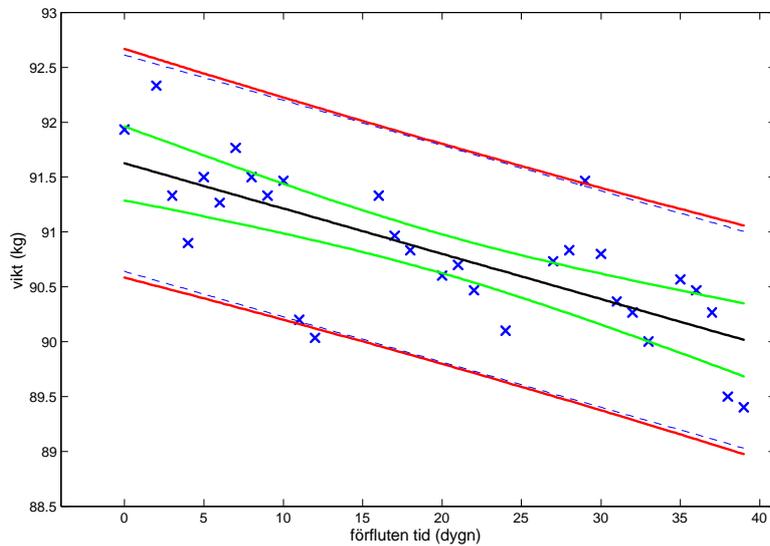
ANOVA

	<i>fg</i>	<i>KvS</i>	<i>Mkv</i>	<i>F</i>	<i>p-värde för F</i>
Regression	1	7.824405802	7.824406	34.73203	2.13479E-06
Residual	29	6.533098198	0.225279		
Totalt	30	14.357504			

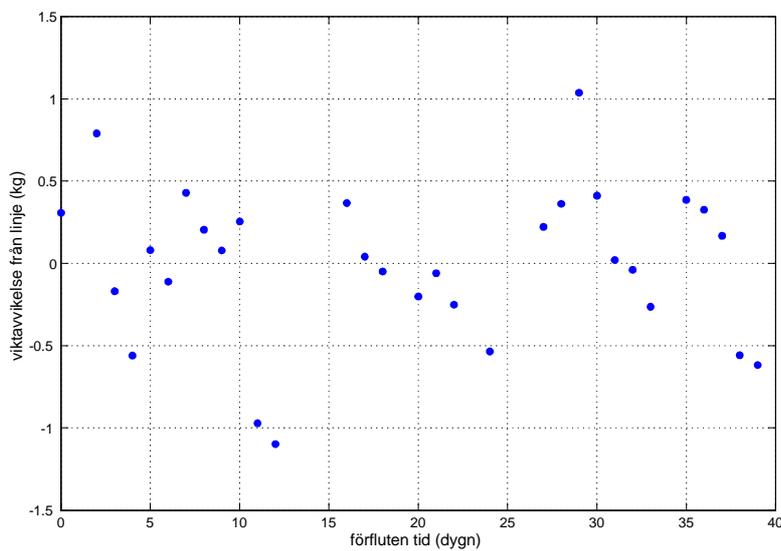
	<i>Koefficienter</i>	<i>Standardfel</i>	<i>t-kvot</i>	<i>p-värde</i>	<i>Nedre 95%</i>	<i>Ovre 95%</i>
Konstant	91.62565745	0.161971614	565.6896	3.52E-60	91.29438813	91.956693
X-variabel 1	-0.041248166	0.006999057	-5.893389	2.13E-06	-0.05556285	-0.026933

Prediktionsintervallen ges inte av Excel, men kan lätt beräknas genom formeln (2.31) och dess gränser är utritade i figur 2.7. Det representerar det intervall som man kan förvänta sig att 95% av framtida medelvärden av tre mätningar skall täcka och en kontroll av de uppmätta värdena visar två observationer utanför intervallet vilket betyder att $100 \cdot \frac{2}{31} = 6\%$ hamnat utanför. Intervallet täcker därmed 94% av värdena; ett rimligt resultat.

I figuren är också approximativa gränser angivna som negligerar osäkerheten i skattningarna, dvs. gränserna i (2.31) har satts lika med $\pm ts$. Dessa ses som streckade linjer och avviker ytterst lite från de fullständiga i detta fall, vilket torde betyda att 31 observationer är tillräckligt för att göra ett förenklat prediktionsintervall. I kapitlet om kalibrering skall vi ange villkor för att använda ett sådant enkelt intervall. I avsnitt 3.5.3 skall vi fortsätta analysen av bantningsdata för att dela upp den slumpmässiga variationen i mätosäkerhet och kroppsvariation.



Figur 2.7: Resultat av dagliga vägningar, anpassad rät linje, 95%-iga konfidens- (gröna inre linjer) och prediktionsintervall (röda yttre linjer), samt approximativt 95%-igt prediktionsintervall (streckat).



Figur 2.8: Residualer från regressionen av vägningresultat.

2.6 Sammanfattning

Det viktigaste statistiska mått som används vid mätosäkerhetsangivelser är standardavvikelsen. Denna är ett mått på en stokastisk variabels spridning kring sitt väntevärde och används såväl för populationerna av mätfelsbidrag som för den slutliga mätosäkerheten. I mätosäkerhetssammanhang betecknas standardavvikelsen med bokstaven u och kallas standardosäkerhet. Om man gjort n_j observationer på en stokastisk variabel x kan dess standardavvikelse s_j beräknas genom formeln:

$$s_j = \sqrt{\frac{1}{n_j - 1} \sum_{i=1}^{n_j} (x_i - \bar{x})^2},$$

där \bar{x} är medelvärdet av observationerna. Om man har flera observationsserier med samma standardavvikelse men med möjligen olika väntevärden skattas standardavvikelsen genom sammanvägning av de olika observationsseriernas enskilda skattningar:

$$s_{SV}^2 = \frac{\nu_1 s_1^2 + \nu_2 s_2^2 + \dots + \nu_m s_m^2}{\nu_1 + \nu_2 + \dots + \nu_m},$$

där $\nu_1, \nu_2, \dots, \nu_m$ betecknar antalet frihetsgrader i respektive s -skattning och antalet frihetsgrader i sammanvägningen blir det som ges i nämnaren, nämligen summan av antalet enskilda frihetsgrader. I fallet med n_j upprepade mätningar på ett objekt med okänt väntevärde blir $\nu_j = n_j - 1$.

Med hjälp av standardavvikelseskattningen s kan man bestämma ett **konfidensintervall** om man känner den statistiska fördelningen för variabeln. I fallet med normalfördelning ges ett 95%-igt konfidensintervall av $\pm 1.96 \cdot s$, vilket i mätosäkerhetssammanhang kallas för den **utvidgade mätosäkerheten** och betecknas med bokstaven U .

Konfidensintervallet för skillnaden mellan två objekts egenskaper blir:

$$E[Y_A] - E[Y_B] = \hat{y}_A - \hat{y}_B \pm t_{\nu_{eq}} \sqrt{s_A^2 + s_B^2} \quad (\text{med konfidens } 95\%)$$

Den sanna skillnaden $E[Y_A] - E[Y_B]$ är lika med den uppmätta skillnaden $\hat{y}_A - \hat{y}_B$ med en viss osäkerhet som bestäms dels av summan av standardavvikelserna för de ingående mätningarna $\sqrt{s_A^2 + s_B^2}$, dels av en t -faktor $t_{\nu_{eq}}$ motsvarande 95% konfidens och vars storlek beror på noggrannheten i skattningen av standardavvikelsen $\sqrt{s_A^2 + s_B^2}$. t -faktorn ges i följande tabell för olika antal frihetsgrader:

Tabell 2.3.2

frihetsgrader	2	3	4	5	6	7	8	9	10	11	18	30
$t_{0.025}$	4.3	3.2	2.8	2.6	2.4	2.4	2.3	2.3	2.2	2.2	2.1	2.0

Det ekvivalenta antalet frihetsgrader för standardavvikelsesumman $\sqrt{s_A^2 + s_B^2}$ ges av Welch-Satterwaite-formeln:

$$\nu_{eq} = \frac{(s_A^2 + s_B^2)^2}{\frac{s_A^4}{\nu_A} + \frac{s_B^4}{\nu_B}},$$

där ν_A och ν_B är antalet frihetsgrader i respektive standardavvikelseskattning.

Om antalet frihetsgrader bakom angivna osäkerhetsintervall är okända så kan ett approximativt konfidensintervall för skillnaden bestämmas utående från de utvidgade mätosäkerheterna U_{y_A} och U_{y_B} genom tumregel 1:

Tumregel 1

Vid jämförelse av två mätvärden som är angivna med 95%-iga mätosäkerhetsintervall fås ett approximativt 95%-igt konfidensintervall för deras differens genom formeln

$$y_A - y_B = \hat{y}_A - \hat{y}_B \pm 2\sqrt{\frac{U_{y_A}^2}{4} + \frac{U_{y_B}^2}{4}}$$

Kapitel 3

Bestämning av mätosäkerhet

För mätstorheten y vill vi etablera följande osäkerhetsintervall

$$y = \hat{y} \pm U_y, \quad (95\%) \quad (3.1)$$

där \hat{y} är det uppmätta värdet som är behäftat med osäkerheter och U_y är den utvidgade mätosäkerheten som antas täcka ca 95% av alla mätvärden. Statistisk teori, samt antagande om ett normalfördelat totalt mätfel ger den utvidgade mätosäkerheten via följande samband

$$U_y = k \cdot u_y,$$

där u_y är standardosäkerheten för mätvärdet \hat{y} , vilket helt enkelt är standardavvikelsen för slumpvariabeln \hat{y} och konstanten k , **täckningsfaktorn**, är det tal som man måste multiplicera standardosäkerheten med för att få den önskade täckningsgraden 95%. Om mätfelet är en normalfördelad stokastisk variabel så skall täckningsfaktorn vara lika med 1.96, men med tanke på andra approximationer (se nästa avsnitt) är detta en onödigt noggrann angivelse och istället används följande allmänna regel: Om standardosäkerheten anses vara känd så kan konstanten $k = 2$ användas, medan man i andra fall använder en större täckningsfaktor, bestämd enligt en t -fördelningstabell, såsom behandlats i avsnitt 2.3.

För att få fram en osäkerhetsangivelse behöver man alltså skatta u_y , samt hålla reda på antalet ekvivalenta frihetsgrader i skattningen. Denna standardosäkerhet är oftast ett resultat av många olika felkällor i mätningen, vilket kan modelleras på följande sätt,

$$\hat{y} = y + f(e_1, e_2, \dots, e_m) \quad (3.2)$$

där alltså det sanna värdet symboliseras med beteckningen y , mätvärdet med \hat{y} och mätfelet som en funktion av de olika felbidragen e_i med f .

Man kan identifiera två olika angreppssätt för att ta reda på mätosäkerheten hos \hat{y} , nämligen 1) att analysera varje felkälla för sig, med avseende på

deras variation, samt med avseende på deras påverkan på mätvärdet \hat{y} , vilket vi här kallar **analysmetoden**, eller 2) att genomföra upprepade mätningar med variation i de olika felkällorna och direkt skatta variationen hos \hat{y} , vilket vi kallar **observationsmetoden**. Det första angreppssättet representeras av den nämnda ISO-Guiden GUM [1] och tillämpas i kvalitetsstandarder för laboratorier såsom EN45000. Det kallas ibland på engelska för “the bottom up approach” medan det andra kallas “the top down approach”, vilket representeras av ISO 5725 [10] och typiskt tillämpas inom kvalitetsstandarder som QS9000.

Analysmetoden kan sammanfattas i formeln

$$u_{\hat{y}}^2 = \sum_{i=1}^m c_i^2 u_{e_i}^2, \quad (3.3)$$

där känslighetskoefficienterna c_i mäter respektive felbidrags inflytande på \hat{y} och u_{e_i} är standardosäkerheten för felbidrag e_i . Om de olika felbidragen samverkar så måste formeln också kompletteras med kovarianstermer.

Observationsmetoden representeras av följande formel

$$u_{\hat{y}}^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2, \quad (3.4)$$

där man i en referensmätning gjort n upprepade observationer på y , då de olika felkällorna tvingats variera inom sina respektive populationer. Medelvärde \bar{y} betraktas här som en approximation av det sanna värdet och divisionen med $n-1$ beror just på att denna approximation av sanningen hämtats ur det aktuella datamaterialet.

I det första angreppssättet ligger svårigheterna i

- att identifiera alla källor till mätfel. Här visar ofta den mänskliga faktorn sina svagheter och viktiga källor till mätfel kan lätt förbises.
- att skatta felkällornas väntevärden och varianser. Här handlar det först om att definiera vad man menar med sanning eller referens, t. ex. vad det nominella värdet är för en felkälla, sedan att fastställa hur stor variation man kan förvänta sig kring detta nominella värde i den aktuella mätsituationen.
- att beräkna felkällornas inflytande på y . Koefficienterna c_i i (3.3) skattas med hjälp av känslighetsanalys, numeriskt, experimentellt eller analytiskt.

I det andra angreppssättet ligger svårigheterna i

- att arrangera ett urval av mätsituationer som representerar alla populationer av mätfel och vars medelvärde representerar den intressanta referensen.

3.1 Relativa osäkerheter

I modellen (3.1) utgick vi ifrån att osäkerheten skall adderas till mätvärdet. Vi förutsätter då att osäkerheten är densamma för olika värden på mätvärdet y , vilket inte alltid är någon bra modell. Ofta är felet istället proportionellt mot mätvärdet och man bör då använda relativa mätosäkerheter,

$$y = \hat{y} \pm \frac{U_y^{\%} \cdot \hat{y}}{100} = \hat{y} \left(1 \pm \frac{U_y^{\%}}{100} \right), \quad (95\%)$$

där nu osäkerheten är given som en procentuell storhet. En tumregel att ha i åtanke är att man för storheter som brukar plottas i logaritmisk skala bör använda relativa osäkerheter.

Vid storheter där mätosäkerheten är proportionell mot mätvärdet är ofta kvoter intressanta vid jämförelser istället för differenser och vid användning av kungsväg 2 på sidan 10 skall då kvoter mellan intressanta mätvärden studeras. Därmed kommer multiplikativa fel att elimineras (förkortas bort) på samma sätt som additiva fel elimineras vid differensbildning.

3.2 Approximationsnivå

Rekommendationerna i denna handbok grundar sig på att man intresserar sig för 95%-iga konfidensintervall. Detta betyder att den statistiska teorin blir speciellt användbar och avgörande förenklingar kan göras i beräkningar och skattningsmetoder. Om högre konfidens är nödvändig måste varje steg i mätosäkerhetsberäkningen göras med större precision, vilket inte inför några principiella svårigheter om den statistiska angreppssättet används, men kräver mer noggrann analys, samt mer precisa kunskaper om osäkerhetskomponenternas statistiska fördelningar.

För att göra kontrollerade approximationer för tumregler och andra förenklingar behövs en regel för hur mycket fel vi kan acceptera i beräkningarna av mätosäkerhet. En enkel regel är följande:

En approximation av ett 95%-igt konfidensintervall betraktar vi som acceptabelt om täckningen ligger i intervallet 93-97%.

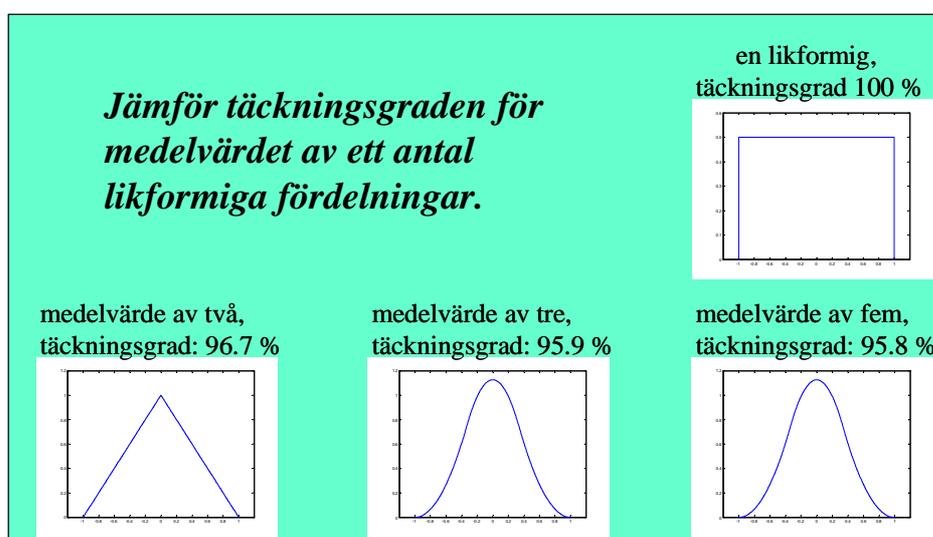
Detta kan man motivera med hänvisning till vad som egentligen menas med 95%-iga konfidensintervall, vilket kunde uttryckas som "rätt så säkert" eller med något liknande vagt uttryck, som lika gärna kunde betyda 93% eller 97%. Ytterligare ett argument för en dylik approximation är att man i statistiska sammanhang i allmänhet talar om intervall motsvarande 90%, 95% eller 99% konfidens, medan värden däremellan sällan förekommer.

En typisk egenskap hos ett mätfel är att det ofta är sammansatt av flera oberoende källor. Denna egenskap i kombination med att vi intresserar oss för 95% konfidens gör att man oftast kan anta normalfördelning för det sammanvägda felet. Normalfördelningsapproximationen kan teoretiskt

motiveras genom den centrala gränsvärdessatsen i statistiken, som säger att man vid summering av oändligt många slumpvariabler får en normalfördelad summa oavsett vilken typ av fördelning de enskilda delarna har. I mätosäkerhetssammanhang har vi förstås inte oändligt många felbidrag av samma storleksordning, men vi kräver inte heller en korrekt normalfördelning, utan nöjer oss med en approximation som är tillräckligt bra kring 95% täckning. I figur 3.1 illustreras hur den centrala gränsvärdessatsen fungerar vid medelvärdet av likformigt fördelade slumpvariabler med samma variationsvidd. Standardosäkerheten för en likformigt fördelad variabel med variationsvidd $\pm a$ är $\frac{a}{\sqrt{3}} \approx 0.58 \cdot a$ (2.6) och om man använder $k = 2$ enligt ovanstående rekommendation så blir den utvidgade mätosäkerheten för ett likformigt fördelat mätfel

$$U_U = 2 \cdot 0.58 \cdot a = 1.16 \cdot a$$

och hamnar därmed utanför variationsvidden $\pm a$. Den verkliga täckningsgraden blir då 100%. Om man beräknar motsvarande täckningsgrader för medelvärdet av två, tre, resp. fem likformiga fördelningar så fås täthetsfunktioner och täckningsgrader enligt figur 3.1. Vi ser att redan efter summering



Figur 3.1: *Demonstration av centrala gränsvärdessatsen.*

av tre variabler blir fördelningen klockliknande och redan vid två (triangel-fördelning) uppfylls vårt krav på att 95% skall betyda 93-97%. Den likformiga fördelningen betraktas ofta som den mest konservativa fördelningstypen som uppträder i normala mätosäkerhetssammanhang vilket i kombination

med resultaten i figuren motiverar följande tumregel.

Tumregel 3

Om standardosäkerheten u_y är bestämd av minst 2 felkällor av samma storleksordning, så kan normalfördelningsapproximationen användas,

$$U_y = 2 \cdot u_y$$

Om en enda felkälla dominerar bör konfidensintervallet för denna felkälla beräknas direkt ur dess skattade fördelning.

Man måste också se upp med extrema fördelningar som kan uppkomma i vissa sammanhang. Ett exempel är fördelningar efter kvalitetssortering, då det kan förekomma att produktpopulationer som är klassade som lågkvalitetsprodukter helt kan sakna komponenter med litet fel, då dessa sorterats bort för en högre klass. En dylik fördelning blir elakare än den likformiga och kräver en mer noggrann analys.

3.2.1 Täckningsfaktor

Om standardosäkerheten för ett mätvärde är bestämd med tillräckligt många frihetsgrader så kan täckningsfaktorn 2 användas. I tabell 2.3.2 har vi sett att detta klart gäller om antalet frihetsgrader är minst 30. Med ovanstående approximationsregel kan detta emellertid modifieras ytterligare.

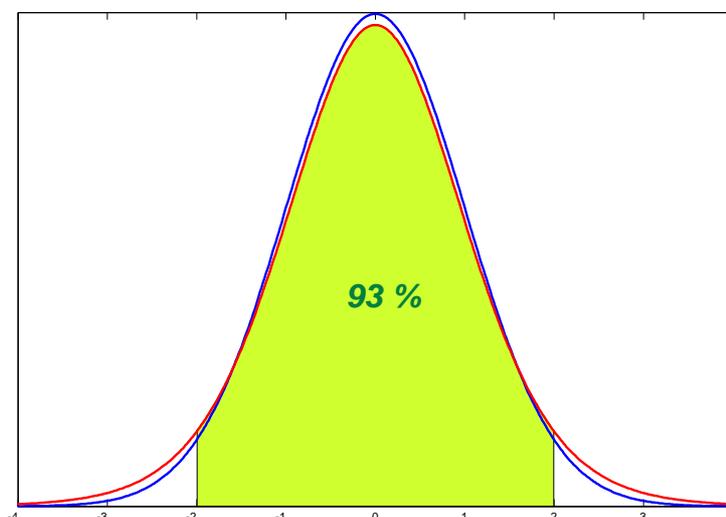
Istället för att använda t -fördelningen för att bestämma täckningsfaktorn k vid olika antal frihetsgrader bestämmer vi nu istället hur många frihetsgrader som krävs för att täckningsfaktorn 2 skall ge en täckningsgrad som överstiger 93%. Det visar sig att denna gräns går vid 11 frihetsgrader och vi kan formulera följande approximationsregel för mätosäkerheten:

Tumregel 4

Om standardosäkerheten u_y är bestämd med minst 11 frihetsgrader så kan täckningsfaktorn $k = 2$ användas

$$U_y = 2 \cdot u_y$$

t -fördelningen med elva frihetsgrader illustreras i figur 3.2 genom att dess frekvensfunktion jämförs med normalfördelnings. Ytan under t -fördelningens frekvensfunktion mellan -2 och 2 standardavvikelse täcker 93 % av fördelningen.



Figur 3.2: Frekvensfunktioner för 1) normalfördelningen med vänteevärde noll och variansen ett (blå) och 2) t-fördelningen med elva frihetsgrader.

3.2.2 Relevanta osäkerhetskomponenter

Felbidragen i en mätning är oftast många, men domineras normalt av några få komponenter. En mätosäkerhetsanalys enligt analysmetoden kan därmed bli alltför omfattande om varje möjligt felbidrag analyseras i detalj. Det angivna approximationskriteriet kan emellertid användas för att bedöma vilka bidrag i (3.3) som skall ingå i det slutliga analysen.

Anta att termen i (3.3) för det största felbidraget är $\{|cu|\}_{\max}$. Anta vidare att ytterligare fem felbidrag har termer som är lika med 20% av det största, $\{|cu|\}_2 = \{|cu|\}_3 = \{|cu|\}_4 = \{|cu|\}_5 = \{|cu|\}_6 = 0.2 \{|cu|\}_{\max}$. Vi får då den sammanvägda mätosäkerheten till

$$u_y^2 = \sum_{i=1}^m c_i^2 u_{e_i}^2 = \sum_{i=1}^6 c_i^2 u_{e_i}^2 = \{c^2 u^2\}_{\max} (1 + 5 \cdot 0.2^2) = 1.2 \cdot \{c^2 u^2\}_{\max}$$

Ett 95%-igt konfidensintervall skulle därmed vara $\pm 2\sqrt{1.2 \cdot \{c^2 u^2\}_{\max}}$. Hur mycket fel blir resultatet om vi bortser från alla osäkerhetskällor utom den maximala? En sådan approximation innebär alltså att konfidensintervallet istället beräknas till $\pm 2\sqrt{\{c^2 u^2\}_{\max}}$, vilket är detsamma som att använda k -faktorn $\frac{2}{\sqrt{1.2}}$ på den korrekta standardavvikelsen

$$2\sqrt{\{c^2 u^2\}_{\max}} = \frac{2}{\sqrt{1.2}} \sqrt{1.2 \cdot \{c^2 u^2\}_{\max}} = 1.83 \sqrt{1.2 \cdot \{c^2 u^2\}_{\max}}$$

Denna täckningsfaktor ger täckningsgraden 93.25% enligt normalfördelningstabell och uppfyller därmed vårt approximationsvillkor. Slutsatsen av denna räkning kan sammanfattas i en tumregel:

Tumregel 5

Om den näst största standardosäkerhetskomponenten i en varianssumma är högst 20% av den största, $\{|cu|\}_{(2)} \leq 0.2 \cdot \{|cu|\}_{\max}$ så kan alla komponenter utom den största försummas

$$u_y = \{|cu|\}_{\max}$$

3.2.3 Systematiska fel

Systematiska fel hålls på en låg nivå genom kalibreringar mot erkända referenser, men något systematiskt fel återstår alltid på grund av exempelvis mätosäkerhet i kalibreringarna, drift mellan kalibreringstillfällena, eller olinjära transformationer. Som tidigare har påpekats så kan dessa återstående avvikelser utifrån en global horisont betraktas som en slumpmässig population av bias och behandlas med statistiska metoder, men på samma sätt som i föregående avsnitt om relevanta osäkerhetskomponenter kan en känd bias negligeras om den är tillräckligt liten till sitt belopp.

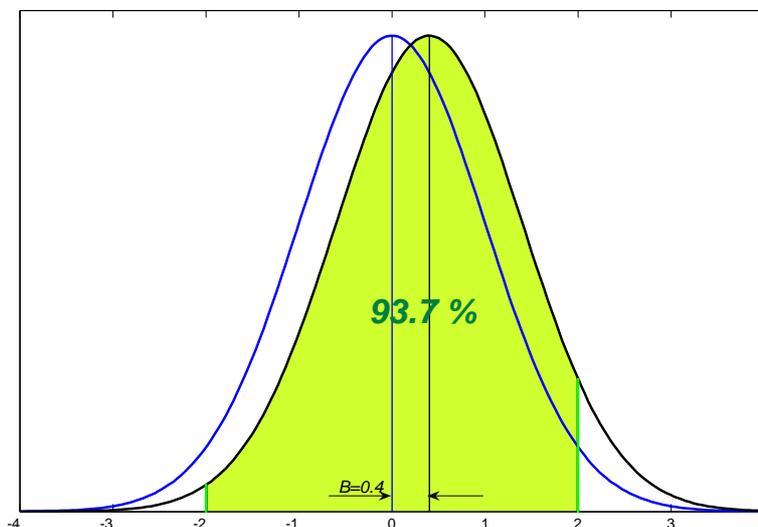
Betrakta ett mätvärde som dels har en mätosäkerhet, u_y och dessutom har en känd bias gentemot det sanna värdet. Om man vid en mätning negligerar denna bias så kommer man att ange mätvärdet som

$$y = \hat{y} \pm 2u_y \quad (\text{med ca } 95 \% \text{ sannolikhet}),$$

där mätvärdet egentligen består av två komponenter $\hat{y} = \tilde{y} + B$, där \tilde{y} har väntevärdet noll och B är värdets bias. Man kan då fråga sig för vilka B som det approximativa uttrycket ger minst 93 % täckningssannolikhet. I figur 3.3 illustreras situationen när man har en positiv bias lika med 0.4 och slumpfelet är normalfördelat med variansen ett. Motsvarande situation uppstår vid negativ bias och en kontrollberäkning med normalfördelning visar att man faktiskt kan ha en bias så stor som $0.47 \cdot u_y$ utan att understiga täckningsgraden 93%. Att acceptera en så stor bias kan dock vara riskfyllt då varje annan approximation kan ge alltför stora fel. Vi formulerar istället en tumregel i samma storlek som för osäkerhetskomponenter.

Tumregel 6

Om ett systematiskt fel är högst 20% av det sammanvägda slumpmässiga felet så kan det försummas.



Figur 3.3: Begränsad bias ger acceptabel täckningsgrad utan justering.

Denna tumregel är speciellt användbar vid regelbundna kalibreringar av instrument, där man kan sätta ett villkor på det systematiska felet vid kalibreringen för godkännande utan justering.

3.2.4 Avrundning

När man anger ett mätvärde med sin mätosäkerhet är det viktigt att ange mätvärdet med en precision som står i proportion till mätosäkerheten, dvs. att inte ange fler siffror än de som är signifikanta. Vi kan använda approximationsregeln för att avgöra ett lagom antal decimaler för såväl mätosäkerheten som för mätvärdet.

För mätosäkerhetsangivelsen u_y gäller att om denna är beräknad på ett riktigt sätt så ger $\pm t_{0.025} \cdot u_y$ ett 95%-igt konfidensintervall. Om man nu avrundar u_y till två signifikanta siffror så kommer detta att införa ett fel i mätosäkerheten på maximalt fem hundradelar. Konfidensintervallet ändras då till $\pm t_{0.025} \cdot u_y \left(1 + \frac{5}{100}\right)$. Denna förändring ändrar täckningsgraden med klart mindre än två procent och uppfyller därmed vår approximationsregel.

Exempel 3.2.1 Om $u_y = 0.1445$ och vi avrundar till $u'_y = 0.14$ så är avrundningsfelet $\frac{0.0045}{0.1445} = 0.03$. Om vi antar att $t_{0.025} = 2$ så skulle det 95%-iga konfidensintervallet vara: $\pm 2 \cdot 0.1445 = 0.289$. Efter avrundning till två signifikanta siffror blir konfidensintervallet istället $\pm 2 \cdot 0.14 = 0.28$. Detta

motsvarar en ändring av täckningfaktorn från 2 till $2 \frac{0.28}{0.289} = 1.94$, vilket ger täckningsgraden 94.2%.

För avrundning av mätvärdet \hat{y} väljer vi en rekommendation i enlighet med Wheeler [6]: Avrunda mätresultatet till samma decimalplats som den mest signifikanta siffran hos standardosäkerheten delad med 1.35, dvs. för $\frac{u_y}{1.35} = 0.21$ eller $\frac{u_y}{1.35} = 0.98$ avrunda till en decimal, för $\frac{u_y}{1.35} = 1.2$ eller $\frac{u_y}{1.35} = 8.1$ avrunda till närmaste heltal. Denna regel ger i gränsfallet med största avrundningen ett avrundningsfel som har standardavvikelsen $\frac{u_y}{2 \cdot 1.35 \cdot \sqrt{3}} = 0.21 \cdot u_y$, eftersom avrundningsfel är likformigt fördelade. Denna avrundning betyder alltså att vi adderar ytterligare en mätosäkerhet till den sammanvägda mätosäkerheten u_y och får

$$u'_y = u_y \sqrt{1 + 0.21^2} = 1.02 \cdot u_y,$$

och om denna negligeras innebär det att täckningsgraden sjunker från 95% till 94.5%. Detta är ett genomsnittligt resultat för alla avrundningar, vi skall i följande exempel se hur avrundningen verkar i det mest extrema fallet med maximal avrundning i förhållande till standardosäkerheten.

Exempel 3.2.2 *Ett extremt fall inträffar t. ex. om mätvärdet är 30.5 och mätosäkerheten är 1.35. Vi får då att $\frac{u_y}{1.35} = 1.0$ och enligt avrundningsregeln skall man då avrunda till närmaste heltal, dvs. 31. Vi anger resultatet som 31 ± 2.7 med täckningfaktorn 2. Det verkliga konfidensintervallet skulle dock vara 30.5 ± 2.7 . En kontroll med normalfördelningen visar att man som resultat av avrundningen får täckningsgraden 93.4%.*

Tumregel 7. Avrundning

Den utvidgade mätosäkerheten U_y avrundas alltid till två värdesiffror.

Det angivna mätvärdet avrundas till samma decimalplats som den mest signifikanta siffran hos standardosäkerheten delad med 1.35.

Exempel: $u_y/1.35 = 0.21$: avrunda till en decimal.

$u_y/1.35 = 8.1$: avrunda till heltal.

3.3 Bestäm ändamålet med mätningen

I enlighet med resonemanget i inledningen av denna handbok måste man relatera en mätosäkerhetsangivelse till det ändamål som är för handen. Om

mätosäkerheten avser reproducerbarhet så skall den kunna användas av en global betraktare vid jämförelser med mätvärden från andra globala aktörer. Den måste då inkludera felkällor som i det egna laboratoriet är konstanta, men i ett globalt perspektiv kan variera. Om mätosäkerheten avser jämförelser inom laboratoriet eller företaget eller om den avser jämförande mätningar i tid eller rum så är det istället repeterbarheten inom den aktuella sfären som avgör mätosäkerheten.

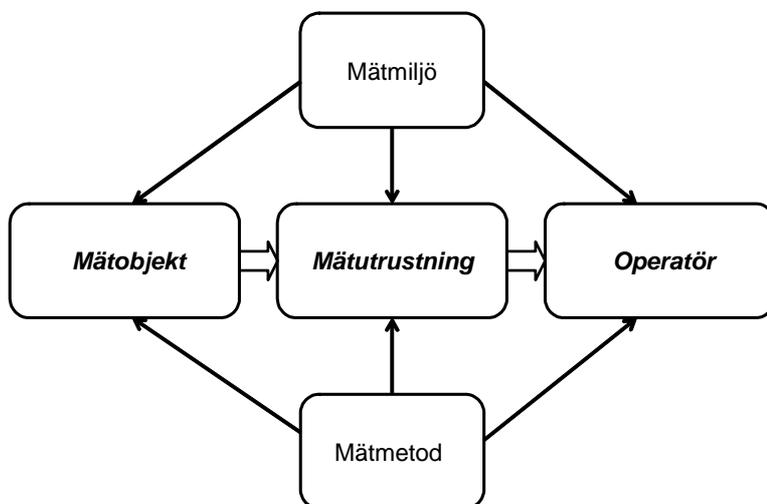
Denna identifiering av ändamål eller betraktelsehorisont påverkar såväl sanningsvärdet som spridningen kring detta. Sanningen har vi definierat som det värde man uppnår då alla variabler som påverkar mätningen är lika med sina önskvärda värden (nominella eller medelvärden). Den beror därför på vilka önskvärda värden man väljer i varje situation. I det globala perspektivet är mätvärdet relaterat till någon standardiserad SI-enhet, nominella värden definierade i någon standard, och medelvärden av andra variabler bestämda av den population av acceptabla laboratorier som definierats. I det lokala perspektivet kan mycket väl en lokal referens, lokala nominella värden och medelvärdet av lokala övriga variationer bestämma sanningen.

3.4 Bestäm alla källor till mätfel

Denna del i mätosäkerhetsutredningen är tveklöst den svåraste att genomföra. Den kräver inte kunskap i matematik, eller statistik, utan i första hand en gedigen erfarenhet av mätningen ifråga. Om en enda väsentlig källa till mätfel försummas, så kan inga avancerade matematiska procedurer rädda situationen, approximationsreglerna kan förlora sin relevans och kalibreringsrutiner bli ett spel för galleriet. Det är därför av största vikt att alla källor till mätfel kan kartläggas, speciellt gäller detta vid användning av analysmetoden om ingen kontroll av den totala mätosäkerheten görs experimentellt. Vid användning av observationsmetoden är det naturligtvis också av största vikt att man är medveten om alla felkällor, då skattningen av mätosäkerheten här förutsätter att försöken är gjorda med ett slumpmässigt val av olika möjliga felbidrag.

Det är svårt att ange ett generellt tillvägagångssätt för att försäkra sig om att få med alla felkällor i utredningen. En möjlig väg är att utgå från figuren på sidan 6, som vi upprepar här i figur 3.4:

För de övre fyra rutorna i denna bild kan man söka identifiera de källor till mätfel som kan uppträda i den aktuella mätsituationen. I mätmetoden kan man finna underlag för att definiera det sanna värdet genom nominella värden på miljövariabler, samt få underlag för hur mycket de olika felkällorna kan bidra genom sina respektive tillåtna intervall. För felkällor som inte är specificerade i mätmetoden eller i mätsituationer utan given mätmetod måste man bedöma såväl nominella värden som tillåtna intervall utifrån egen och kollegors erfarenhet.



Figur 3.4: En översiktlig bild av ett mätsystem.

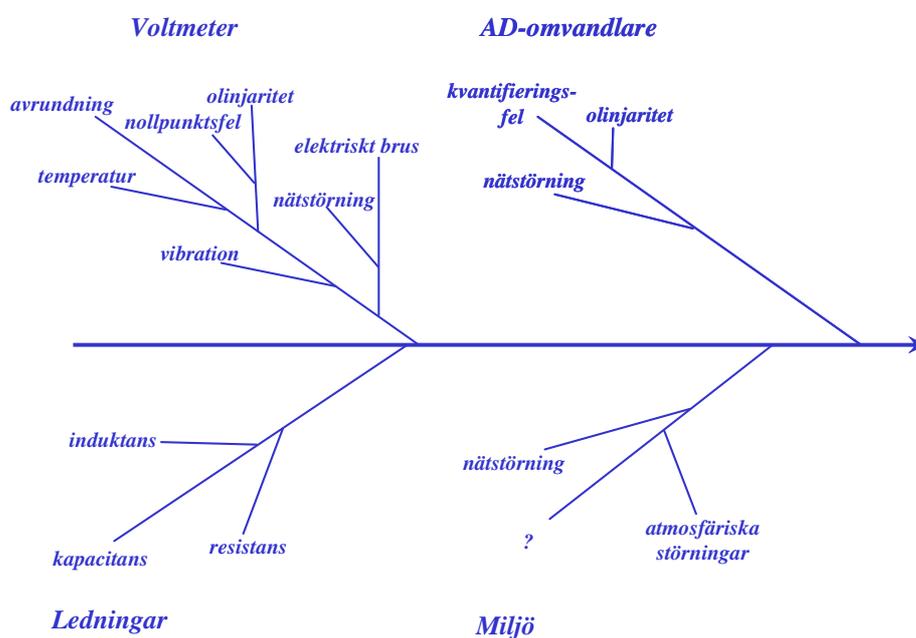
Mätutrustningens bidrag till mätosäkerheten kan oftast hämtas från kalibreringar. Det kan emellertid tillkomma felkällor vid användningen av instrumentet då ju mätsituationen kan skilja sig från den ideala som använts vid kalibrering.

Mätmiljöns påverkan kan ibland uppskattas genom längdutvidgningskoefficienter (temperatur) och liknande fysikaliska karakteristika, men måste i många fall kontrolleras genom experiment om inte välkontrollerade fysikaliska samband är tillgängliga.

Mätobjektet bidrar till mätosäkerheten om mätstorheten inte är tillräckligt väl definierad, vilket ofta är fallet. Diametern för en cylinder är exempelvis inte definierad om den är oval, utan något medelvärde för yttermättet får betraktas som det nominella värdet. Längden på ett rektangulärt bord är på samma sätt endast väldefinierat om bordet verkligen är rektangulärt, vilket aldrig inträffar exakt.

Ett annat sätt att systematiskt lista alla källor till mätfel är att använda ett s.k. fiskbensdiagram, vilket hjälper till att "bena" ut olika delar i mätprocessen. I figur 3.5 ses ett exempel på ett sådant fiskbensdiagram för en mätning av elektrisk spänning över något objekt, placerat på ett visst avstånd från mätutrustningen. Varje identifierbart objekt i mätkedjan analyseras med avseende på olika felkällor.

När alla källor till mätfel har identifierats skall deras bidrag till den totala mätosäkerheten bestämmas. Som tidigare nämnts kan detta ske genom observationsmetoden eller genom analysmetoden, vilka behandlas nedan. Oftast har man i en mätosäkerhetsanalys användning för båda grundmeto-



Figur 3.5: Exempel på ett fiskenbendiagram för listning av felkällor.

derna och krav på kvalitet i kombination med tillgång på uppgifter, tid och pengar avgör hur fördelningen sker.

3.5 Observationsmetoden

Den överlägset säkraste metoden för att bestämma mätosäkerhet är att

1. definiera ett referensvärde för mätningen,
2. genomföra ett statistiskt försök över den population av mätförhållanden som osäkerheten skall avse.

3.5.1 Referens

Som tidigare har framhållits beror referensen på vilket användningsområde som avses vid mätosäkerhetsanalysen.

Om användningen avser en global mätosäkerhet gentemot en internationellt definierad enhet så måste mätutrustningen vara kalibrerad mot denna referens. Detta sker t. ex. genom en spårbar serie av kalibreringar lokalt - nationellt - internationellt, där varje kalibrering görs med en viss kontrollerad mät noggrannhet och resulterar i en slutlig angiven mätosäkerhet, u_{ref} , som

är ett mått på hur mätinstrumentet under kalibreringsförhållanden överensstämmer med referensen. Denna osäkerhetskomponent måste inkluderas i den slutgiltiga osäkerhetsangivelsen.

Om användningen avser en lokal mätosäkerhet kan man i vissa fall använda en gemensam referens vid det statistiska försöket och mätosäkerheten är då helt bestämd av det statistiska försöket.

Om användningen endast avser jämförelser kan man definiera medelvärdet i det statistiska försöket som referens.

3.5.2 Statistiskt försök

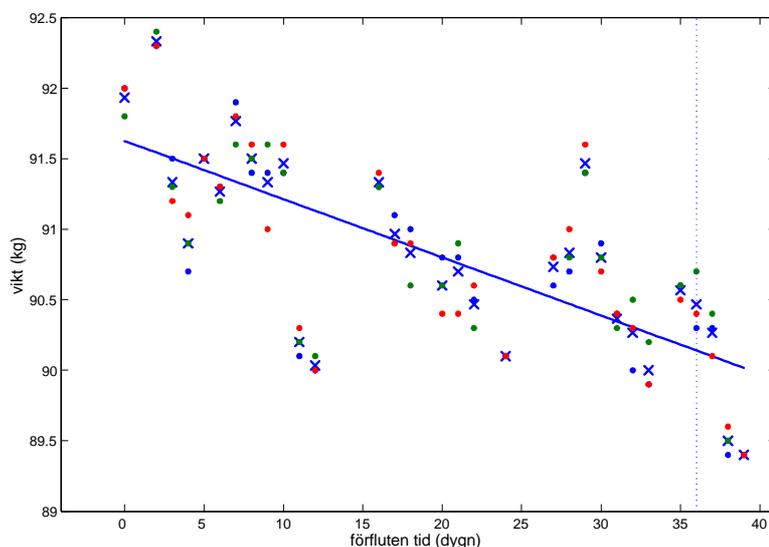
För att få en korrekt skattning av mätosäkerheten inom ett arbetsområde bör ett försök genomföras med mätningar som är slumpmässigt valda ur hela det användningsområde som osäkerheten skall avse. Om t. ex. det handlar om ett globalt användningsområde så bör man använda ett slumpmässigt urval av mätlaboratorier i världen, där variationer i t. ex. temperatur, luftfuktighet, hantering, instrumentkvalitet osv. representerar alla de förhållanden som är accepterade enligt mätinstruktion eller standard. Om användningsområdet är ett enskilt företag så bör man välja ett korrekt urval av instrument, operatörer och relevanta miljövariabler för ett statistiskt försök. Detta kan ske antingen genom att välja samtliga representanter ur populationen, t. ex. för operatörer och instrument, eller genom att välja slumpmässigt ur populationerna.

Det övergripande statistiska verktyget för att skatta osäkerhetskomponenter ur ett statistiskt försök är variansanalysen, ofta förkortad ANOVA enligt dess engelska benämning ANalysis Of VAriance. Bakgrunden till metoden kommer att förklaras här och användningen i olika mätosäkerhetssituationer kommer att expemplieras. Själva utförandet av beräkningarna kan lätt programmeras i något lämpligt programmeringspråk, men finns också tillgängligt i många etablerade program. Vi börjar med att analysera det inledande bantningsexemplet.

3.5.3 Bantning

I det inledande bantningsexemplet i avsnitt 1.1 finns flera variationskällor, varav viktens naturliga variation över tiden tycktes dominera variationen, se figur 3.6.

Mätosäkerheten representeras här framför allt av två felkällor: 1) den slumpmässiga variationen i varje mätning som kan bero på hur man ställer sig på vågen, med vilken hastighet det sker och andra orsaker som t. ex. elektriskt brus i vågens elektriska komponenter och 2) den systematiska avvikelserna från den sanna vikten, representerad av det internationella referensskilot i Paris. Övriga mätfel som instabilitet i tiden, miljöpåverkan o. dyl. bedöms som försumbara i sammanhanget.



Figur 3.6: Resultat av vägningar med en skattad linjär trend.

För att analysera mätresultaten antar vi följande matematiska modell för mätresultaten

$$m_{ik} = m_0 + c(t_k - t_0) + e_k + B + \varepsilon_{ik}, \quad \{i = 1, 2, 3\}, \quad \{k = 1, 2, \dots, 31\} \quad (3.5)$$

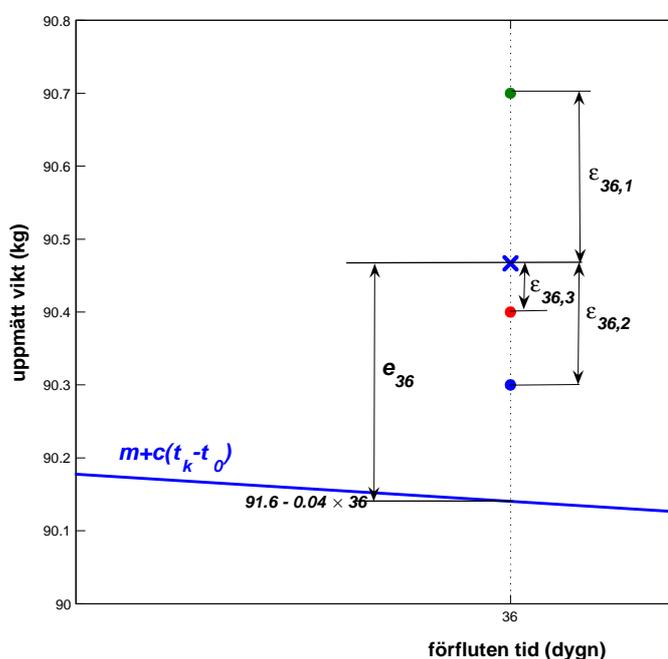
där k indikerar vid vilken dag vägningen skett och index i anger vilken av de tre dagliga vägningarna som avses. Begynnelsevikten betecknas med m_0 , viktökning per dag med konstanten c , dagsnumret med t_k , kroppsviktsvariationen med e_k , vågens systematiska avvikelse från referensvikt med B och det slumpmässiga felet vid mätning nr i , dag t_k med ε_{ik} . För att illustrera den matematiska modellen (3.5) har vi förstörat upp figur (3.6) kring dag 36 i figur 3.7.

Eftersom vi gjort upprepade mätningar vid varje vägningstillfälle så kan vi uppskatta det slumpmässiga mätfelets varians. Bilda först medelvärdena för varje dag:

$$\bar{m}_{.k} = m_0 + c(t_k - t_0) + e_k + B + \bar{\varepsilon}_{.k}, \quad \{k = 1, 2, \dots, 31\},$$

där punkten som ersätter indexet i anger att medelvärdet är taget över detta index. Här är de flesta termerna i modellen konstanta för varje k och om man bildar differenserna mot dagsmedelvärdet så får man

$$m_{ik} - \bar{m}_{.k} = \varepsilon_{ik} - \bar{\varepsilon}_{.k}, \quad \{k = 1, 2, \dots, 31\}.$$



Figur 3.7: Illustration av komponenterna i den statistiska modellen för linjär regression.

Vi kan därmed skatta variansen för det slumpmässiga felet enligt (2.5):

$$\widehat{\text{Var}}[\varepsilon_{ik}] = \frac{1}{3 \cdot 31 - 31} \sum_{k=1}^{31} \sum_{i=1}^3 (\varepsilon_{ik} - \bar{\varepsilon}_{.k})^2 = \frac{1}{62} \sum_{k=1}^{31} \sum_{i=1}^3 (m_{ik} - \bar{m}_{.k})^2.$$

Antalet frihetsgrader i denna skattning är det totala antalet observationer, $3 \cdot 31 = 93$, minskat med det antal skattningar i kvadratsumman som bestäms ur samma observationer. Dessa är alla medelvärdena $\bar{m}_{.k}$, dvs. 31 stycken. Resultatet av exemplets observationer blir därmed

$$u_\varepsilon = \sqrt{\widehat{\text{Var}}[\varepsilon_{ik}]} = \sqrt{\frac{1}{62} \sum_{k=1}^{31} \sum_{i=1}^3 (m_{ik} - \bar{m}_{.k})^2} = 0.15 \text{ kg}$$

Notera att denna bestämning av standardosäkerheten kan ses som en sammanvägd standardavvikelse enligt (2.18), där standardavvikelserna för dagarnas slumpmässiga fel vägs samman till en gemensam skattning.

$$\frac{1}{62} \sum_{k=1}^{31} \sum_{i=1}^3 (m_{ik} - \bar{m}_{.k})^2 =$$

$$\begin{aligned} \frac{1}{62} \left[\sum_{i=1}^3 (m_{i1} - \bar{m}_{.1})^2 + \sum_{i=1}^3 (m_{i2} - \bar{m}_{.2})^2 + \dots + \sum_{i=1}^3 (m_{i31} - \bar{m}_{.31})^2 \right] = \\ \frac{1}{62} \left[\frac{3-1}{3-1} \sum_{i=1}^3 (m_{i1} - \bar{m}_{.1})^2 + \frac{3-1}{3-1} \sum_{i=1}^3 (m_{i2} - \bar{m}_{.2})^2 + \dots \right] = \\ \frac{\nu_1 s_1^2 + \nu_2 s_2^2 + \dots + \nu_{31} s_{31}^2}{\sum_{k=1}^{31} \nu_k} \end{aligned}$$

där frihetsgraderna för varje daglig skattning är som vid den vanliga variansskattningen (2.4), $\nu_k = n - 1 = 2$.

För att kunna vidare analysera variationerna i mätresultaten måste först den deterministiska viktutvecklingen elimineras från resultaten. Då vi i den enkla modellen antog en linjär trend kunde vi skatta viktminskningstakten med hjälp av linjär regression (i avsnitt 2.5). Modellen, medelvärdesbildad över de upprepade vägningarna, skrivs då

$$\bar{m}_{.k} = a + c(t_k - t_0) + E_k ,$$

där konstanten a ses som en summa av initialvikt och bias, $a = m_0 + B$, och felbidraget är en summa av viktvariation och slumpfel, $E_k = e_k + \bar{\varepsilon}_{.k}$. Regressionsanalys gav parameterskattningarna $\hat{a} = 91.6$ kg och $\hat{c} = -0.041$ kg/dag (se avsnitt 2.5). Med hjälp av dessa kan vi vidare analysera variationskomponenterna. Om den skattade genomsnittliga viktminskningen tas bort från varje observation så får man en skattning av just variationskomponenterna:

$$\hat{E}_k = \bar{m}_{.k} - \hat{a} - \hat{c}(t_k - t_0)$$

och man kan skatta dess varians med den vanliga skattningsformeln

$$\widehat{\text{Var}}[E_k] = \frac{1}{29} \sum_{k=1}^{31} \left(\hat{E}_k - \overline{\hat{E}_k} \right)^2 = \frac{1}{29} \sum_{k=1}^{31} (\bar{m}_{.k} - \bar{m}_{..})^2 = 0.225,$$

där antalet frihetsgrader ges av det totala antalet observationer i kvadratsumman minus antal parametrar skattade ur samma resultat, vilket är två, \hat{a} och \hat{c} . Om vi nu antar att viktvariationen är oberoende av den slumpmässiga vägningvariationen så gäller enligt (2.8) att

$$\text{Var}[E_k] = \text{Var}[e_k + \bar{\varepsilon}_{.k}] = \text{Var}[e_k] + \text{Var}[\bar{\varepsilon}_{.k}] .$$

Variansen för slumpfelet har vi skattat tidigare till u_ε^2 och variansen för medelvärdet av tre sådana slumpfel är enligt (2.10) lika med $u_\varepsilon^2/3$. Därmed kan vi skatta viktvariationen σ_e^2 med

$$\widehat{\text{Var}}[e_k] = \widehat{\text{Var}}[E_k] - u_\varepsilon^2/3 ,$$

vilket i vårt fall resulterar i standardavvikelsen för viktvariationen

$$\sqrt{\widehat{Var}[e_k]} = \sqrt{0.225 - 0.15^2/3} = 0.47 \text{ kg.}$$

Vi har alltså nu delat upp variationen i observationerna i en deterministisk trend, en slumpmässig viktvariation över dagar och en slumpmässig mätvariation. För syftet att kontrollera just viktförändring över tid representeras då mätosäkerheten av repeterbarheten i mätningarna, $u_\varepsilon = 0.15$ kg. För att kontrollera BMI mot expertisens hälsokriterier däremot, tillkommer en mätosäkerhetskomponent, nämligen det systematiska felet hos den aktuella badrumsvågen. Detta fel skulle kunna skattas genom att man genomförde en kalibrering mot en normal, som i sin tur är kalibrerad mot det internationella kilot genom en kontrollerad kalibreringskedja. Alternativt kan man bedöma denna felkomponent genom vågens specifikation eller allmän erfarenhet. I det aktuella fallet finns varken någon kalibrering eller någon specifikation tillgänglig och vi bedömer att dylika vågars systematiska fel har standardavvikelsen 0.25 kg. Den sammanvägda mätosäkerheten för en enstaka vägning är då ur ett globalt perspektiv,

$$u_m = \sqrt{u_\varepsilon^2 + u_B^2} = \sqrt{0.15^2 + 0.25^2} = 0.29 \text{ kg}$$

3.5.4 Variansanalys

Beräkningarna i bantningsexemplet illustrerar idén bakom variansanalysen, nämligen att man försöker dela upp variationen i ett mätresultat i olika komponenter, relaterade till olika intressanta företeelser. Då variansanalysen är så användbar har den standardiserats för vissa enkla modeller och kan därmed genomföras med hjälp av vanliga datorprogram. Vi skall här ge standardformuleringarna och visa på hur de kan användas i mätosäkerhets-sammanhang. Vi utgår från ett exempel på vägning av olika detaljer.

Urskiljning av en faktor

Vi utgår först från att vägningarna har utförts på tio olika produkter med tre olika mätsystem. Vi antar nu att de tre mätsystemen är ett slumpmässigt urval av laboratoriets utrustning och ser därmed försöket som ett statistiskt experiment för att med observationsmetoden bestämma mätosäkerheten vid vägning. Resultatet av försöket ger alltså tre vägningar på varje produkt och mätvärdena ses i tabellen nedan.

Tabell 3.5.4.1 Vikt i g

produkt	1	2	3	4	5	6	7	8	9	10
	0.60	1.00	0.85	0.85	0.55	1.00	0.95	0.85	1.00	0.60
	0.55	1.05	0.75	0.75	0.40	1.00	0.90	0.70	0.95	0.50
	0.55	1.05	0.80	0.80	0.50	1.05	0.95	0.80	1.05	0.85
mean	0.57	1.03	0.80	0.80	0.48	1.02	0.93	0.78	1.00	0.65
std.	0.03	0.03	0.05	0.05	0.08	0.03	0.03	0.08	0.05	0.18

För att få mått på mätosäkerheten måste man göra en matematisk analys av datamängden. Vi betraktar först mätresultaten som resultat av följande modell

$$y_{ij} = m_i + e_{ij}, \quad i = 1, 2, \dots, 10, \quad j = 1, 2, 3, \quad (3.6)$$

där m_i är väntevärdet för produkten nr i och e_{ij} är mätfelet för produkt nr. i vid mätning nr. j . Medelvärdena i tabellen är då skattningar av m_i och standardavvikelseerna är mått på variationen i mätfelet e_{ij} .

Uppdelning av variabiliteten Variabiliteten i datamängden beskrivs här matematiskt som kvadratiska avvikelser från medelvärdet och man kan visa att denna variabilitet kan delas upp i två kvadratsummor:

$$\sum_{i=1}^{10} \sum_{j=1}^3 (y_{ij} - \bar{y}_{..})^2 = 3 \sum_{i=1}^{10} (\bar{y}_i - \bar{y}_{..})^2 + \sum_{i=1}^{10} \sum_{j=1}^3 (y_{ij} - \bar{y}_i)^2$$

$$SS_T = SS_D + SS_E$$

där punkterna i medelvärdesbeteckningarna som tidigare betyder det eller de index som respektive medelvärde har tagits över och de förkortade beteckningarna för de tre termerna i summan är de standardbeteckningar som ofta används i variansanalytlitteratur. Anledningen till just denna uppdelning är att de här bildade kvadratsummorna har mycket lämpliga matematiska egenskaper som underlättar skattningen av de ingående variationskomponenterna. Det är nämligen så att de olika kvadratsummorna här är statistiskt oberoende, vilket gör att man har bra kontroll över deras egenskaper.

Speciellt vet man vad antalet frihetsgrader är för de olika komponenterna. Totala antalet frihetsgrader ν_T är lika med totala antalet mätningar minus ett, dvs. antalet kvadrater i kvadratsumman SS_T minus antalet parametrar som skattats med samma datamaterial. Detta antal frihetsgrader summeras också på högersidan i formeln, där frihetsgraderna för SS_E bestäms av det totala antalet mätningar minus antalet medelvärden i dess

kvadratsumma och slutligen frihetsgraderna för SS_D bestäms som återstoden,

$$\nu_T = \nu_D + \nu_E .$$

I programmet Microsoft Excel ger en datanalys med verktyget “ANOVA - en faktor” variansanalystabellen:

ANOVA						
Variationsursprung	KvS	fg	MKv	F	p-värde	F-krit
Mellan grupper	1.008667	9	0.112074	20.3771	3.33E-08	2.392817
Inom grupper	0.11	20	0.0055			
Totalt	1.118667	29				

Här visar den första kolumnen KvS uppdelningen av kvadratsumman i komponenter där “Mellan grupper” avser SS_D , “Inom grupper” avser SS_E och “Totalt” avser SS_T . Den andra kolumnen fg visar antalet frihetsgrader i de olika kvadratsummorna och med hjälp av dessa kan man beräkna de två medelkvadratsummor som står i tredje kolumnen MKv ,

$$MS_D = \frac{SS_D}{\nu_D} , MS_E = \frac{SS_E}{\nu_E} .$$

Variabiliteten i mätresultatet har därmed delats upp i två delar, en som representerar variationen mellan detaljerna och en som representerar variationen i mätfel. Mätosäkerheten skattas i detta fall direkt ur medelkvadratsumman för mätfelet:

$$u_y = \sqrt{MS_E}$$

och denna skattning har ν_E frihetsgrader. I det aktuella fallet kan standardosäkerheten fås ur variansanalystabellen

$$u_y = \sqrt{0.0055} = 0.074 \text{ g} \quad (3.7)$$

och antalet frihetsgrader är lika med 20, vilket är tillräckligt stort för att använda täckningsfaktorn 2 och därmed ge den utvidgade standardosäkerheten

$$U_y = 2 \cdot u_y = 0.15 \text{ g}$$

Resultatet av variansanalysen kan också användas för att bedöma om mät-systemet är tillräckligt bra för sitt användningsområde, genom att jämföra mätosäkerheten med produktvariationen. Detta behandlas i avsnitt 5.2.1.

Urskiljning av två faktorer

De tre olika mätsystemen som användes i förra exemplet är behäftade med olika felkällor, varav man kanske bedömer operatörsinflytandet som dominerande. För att undersöka inverkan från denna källa kan man genomföra

ett mer omfattande experiment där man håller reda på inverkan av de olika operatörerna. I tabell 3.5.4.2 redovisas resultaten från ett sådant experiment, där tre olika operatörer vägt de tio detaljerna vardera två gånger. För att studera mätosäkerhetens olika komponenter delar vi nu upp mätfelet i modellen (3.6) i två delar:

$$y_{ijk} = m_i + B_j + \varepsilon_{ijk}, \quad i = 1, 2, \dots, 10, \quad j = 1, 2, 3, \quad k = 1, 2 \quad (3.8)$$

där vi infört ett operatörsspecifikt mätfel, B_j . I tabell 3.5.4.2 visas vilka mätningar som gjorts med respektive operatör och operatörsmedelvärdena.

Tabell 3.5.4.2 Vikt i g

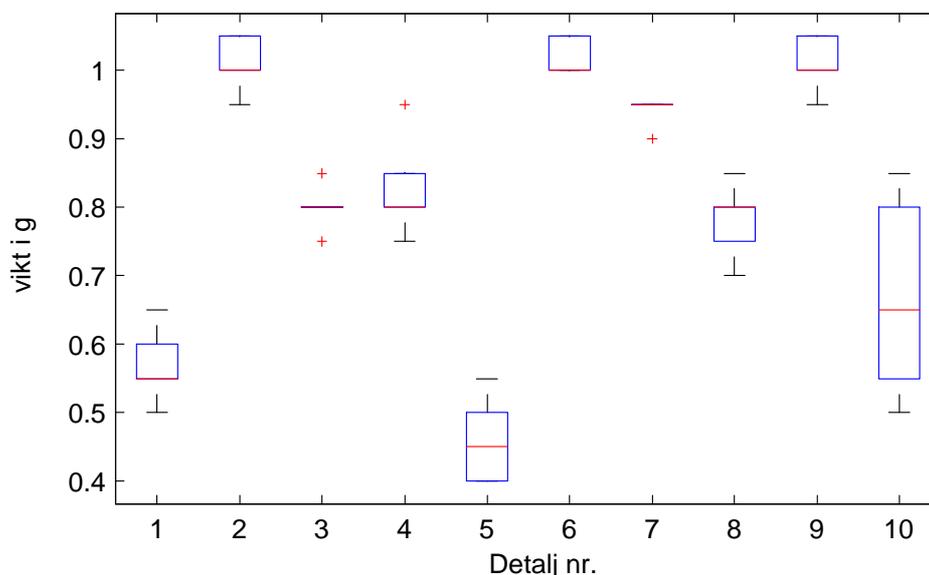
prod.	1	2	3	4	5	6	7	8	9	10	medel
A	0.65	1.00	0.85	0.85	0.55	1.00	0.95	0.85	1.00	0.60	
A	0.60	1.00	0.80	0.95	0.45	1.00	0.95	0.80	1.00	0.70	0.83
B	0.55	1.05	0.80	0.80	0.40	1.00	0.95	0.75	1.00	0.55	
B	0.55	0.95	0.75	0.75	0.40	1.05	0.90	0.70	0.95	0.50	0.77
C	0.50	1.05	0.80	0.80	0.45	1.00	0.95	0.80	1.05	0.85	
C	0.55	1.00	0.80	0.80	0.50	1.05	0.95	0.80	1.05	0.80	0.83
medel	0.57	1.01	0.80	0.82	0.46	1.02	0.94	0.78	1.01	0.67	
std.	0.05	0.04	0.03	0.07	0.06	0.03	0.02	0.05	0.04	0.14	

I figur 3.8 visas resultaten i en **box-plot**. I en box-plot illustreras en datamängd så att varje box representerar värden från en delgrupp i datamängden, där boxen omsluter hälften av värdena, med 25 % av värdena under nedre gränsen och 75 % under övre gränsen. Mittstreckket representerar medianen och de båda svansarnas yttergränser visar det minsta resp. största värdet med undantag av enstaka extremer¹ som visas som kryss. Box-plotten ger en mycket klar bild över mätsituationens väsentliga delar; man kan se att variationen mellan olika detaljer är mycket större än mätvariationen, samt att mätvariationen är ungefär i samma storleksordning för alla produkter utom den tionde, där en mycket större variation kan konstateras. Denna bild av situationen är av stor betydelse, framför allt när man vill analysera mätosäkerheten för att finna medel att minska den.

Ett annat sätt att illustrera mätresultaten är trädidiagrammet i figur 3.9. Här ses totalmedelvärdet som roten i det upp-och-ned-vända trädet, med produktmedelvärdena i nästa nivå. De olika operatörernas bidrag illustreras

¹Bedömningen om vad som är ett extremt värde kan variera mellan olika implementeringar och skall här bara ses som en grov hjälp för bedömning av datamängden.

i nästa nivå, där medelvärdena av varje upprepad mätning är placerad. På lägsta nivån är samtliga mätvärden ritade.



Figur 3.8: *Box-plot som beskriver vägningsresultat av tio olika detaljer.*

Kvadratsummeuppdeleningen blir i detta fall:

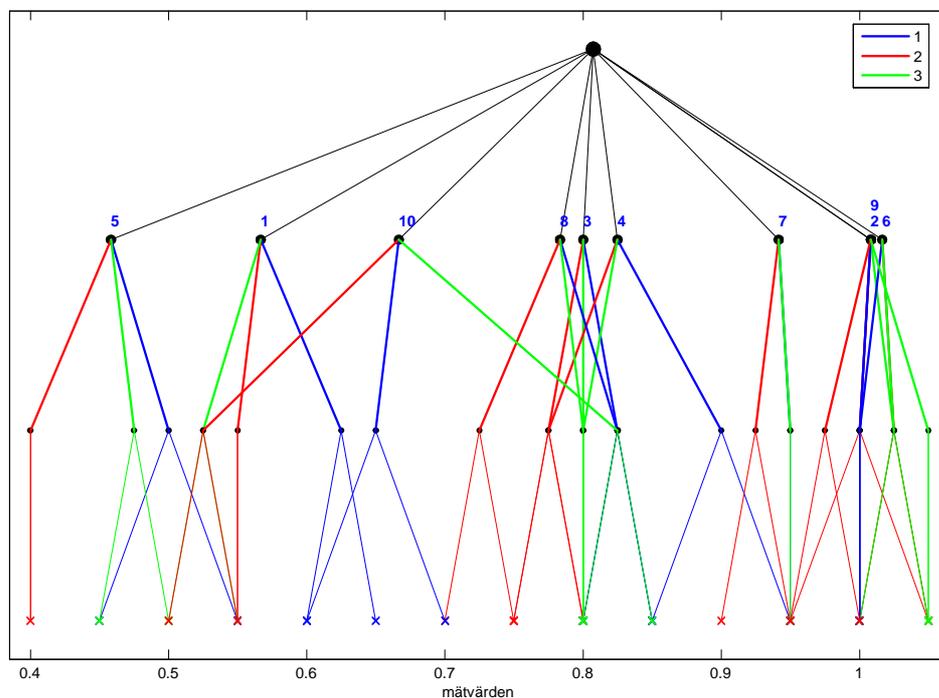
$$\sum_{i=1}^{10} \sum_{j=1}^3 \sum_{k=1}^2 (y_{ijk} - \bar{y}_{...})^2 =$$

$$6 \sum_{i=1}^{10} (\bar{y}_{i..} - \bar{y}_{...})^2 + 20 \sum_{j=1}^3 (\bar{y}_{.j.} - \bar{y}_{...})^2 + \sum_{i=1}^{10} \sum_{j=1}^3 \sum_{k=1}^2 (y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$$

$$SS_T = SS_D + SS_O + SS_E$$

Vi ser att de två kvadratsummorna SS_D och SS_O bildas genom att man a) gör en variansskattning för respektive medelvärden utan att dividera med antalet frihetsgrader och b) multiplicerar med en konstant som gör att antalet kvadrattermer i summan blir densamma som det totala antalet mätningar.

Vid användning av Excels "Anova- två faktorer med reproducering" fås följande tabell:



Figur 3.9: Träddiagram som beskriver datamaterialet från vägning av tio produkter av tre operatörer.

ANOVA						
Variationsursprung	KvS	fg	Mkv	F	p-värde	F-krit
Sampel	0.048	2	0.024	18.58065	5.62E-06	3.315833
Kolumner	2.058708	9	0.228745	177.0932	1.71E-23	2.210697
Interaktion	0.103667	18	0.005759	4.458781	0.000156	1.960117
Inom	0.03875	30	0.001292			
Totalt	2.249125	59				

Här motsvaras operatörernas kvadratsumma SS_O av raden "Sampel", detaljens SS_D av raden "Kolumner" och resterande variation SS_E av två rader, nämligen "Interaktion" och "Inom". Denna ytterligare uppdelning av restsumman motsvarar en något mer komplicerad modell än (3.8), som också inkluderar en eventuell samverkans effekt mellan detalj och mätfel. Oftast är denna effekt försumbar och den enklare modellen (3.8) är fullt tillräcklig för att analysera mätosäkerhetskomponenterna.

På samma sätt som i variansanalysen med en effekt så ges antalet frihetsgrader i andra kolumnen och medelkvadratsummorna bildas genom division av respektive kvadratsumma med dess antal frihetsgrader. Speciellt så är

alltså

$$MS_O = \frac{SS_O}{\nu_O} = \frac{0.048}{2} = 0.024, \quad MS_E = \frac{SS_E}{\nu_E} = \frac{0.103667 + 0.03875}{18 + 30} = 0.0030,$$

där samverkans-effekten inkluderats i den slumpmässiga mätvariationen.

Den näst sista kolumnen i variansanalystabellen, "p-värde", används för att bedöma om de olika effekterna är signifikanta eller ej. Statistisk teori visar nämligen att om inga effekter från våg eller detalj existerar så kommer kvoter mellan de olika medelkvadratsummorna att ha en känd statistisk fördelning: F -fördelningen. Genom att jämföra varje variationskomponents medelkvadratsumma med "inom"-variationen så kan man få en uppfattning om de olika effekterna kan vara ett utslag av slumpen eller ej. Detta betyder i det aktuella fallet att om ingen variation existerar i detaljernas vikt så är sannolikheten för det uppnådda mätresultatet så liten som $1.7 \cdot 10^{-23} \approx 0$. Om operatörernas medelavvikelse inte existerar så är sannolikheten för det uppnådda resultatet $5.6 \cdot 10^{-6}$, dvs en chans på 200,000. Variationskomponenternas inverkan är därmed i högsta grad signifikanta; man brukar ofta betrakta inverkan som signifikant om "p-värdet" är mindre än $0.05 = 5\%$. Kolumnen "F" anger värdet på den aktuella kvoten, t. ex. $F_D = \frac{MS_D}{MS_{Inom}} = \frac{0.228745}{0.001292} = 177$, och kolumnen "F-krit" anger det kritiska värdet på "F" för den angivna signifikansnivån (normalt 5%). Man kan också konstatera en signifikant samverkans-effekt i det aktuella fallet.

Om vi nu betraktar operatörerna i experimentet som ett slumpmässigt urval ur en population av operatörer så kan vi ur vår analys skatta variansen för denna population. Betrakta kvadratsumman

$$SS_O = 20 \sum_{j=1}^3 (\bar{y}_{.j} - \bar{y}_{...})^2 = 20 \cdot 2 \left[\frac{1}{3-1} \sum_{j=1}^3 (\bar{y}_{.j} - \bar{y}_{...})^2 \right]$$

där uttrycket inom hakparentesen faktiskt är en vanlig variansskattning av variabeln $\bar{y}_{.j} - \bar{y}_{...}$. Om man använder modellen (3.8) så ser man att denna variabel lika med

$$\begin{aligned} \bar{y}_{.j} - \bar{y}_{...} &= (\bar{m} + B_j + \bar{\varepsilon}_{.j}) - (\bar{m} + \bar{B} + \bar{\varepsilon}_{...}) = B_j - \bar{B} + \bar{\varepsilon}_{.j} - \bar{\varepsilon}_{...} = \\ &= (B_j - \bar{B}) + (\bar{\varepsilon}_{.j} - \bar{\varepsilon}_{...}), \end{aligned}$$

dvs. är en summa av två slumpvariabler minskade med sina respektive medelvärden. Om man nu antar att den slumpmässiga mätvariationen ε är oberoende av variationen i det systematiska mätfelet i operatörerna B så betyder det att

$$Var[\bar{y}_{.j} - \bar{y}_{...}] = Var[B] + Var[\bar{\varepsilon}_{.j}]$$

och vi kan skatta operatörsvariansen med

$$s_B^2 = Var[\widehat{\bar{y}_{.j} - \bar{y}_{...}}] - Var[\widehat{\bar{\varepsilon}_{.j}}] = \frac{SS_O}{20 \cdot 2} - \frac{MS_E}{20} = \frac{MS_O - MS_E}{20}.$$

Här är variansen för medelvärdet av alla slumpfel för en viss operatör lika med variansen för slumpfelet dividerat med antalet som medelvärdesbildats, nämligen 20. Allmänt gäller att om en variationskomponent är representerad av a varianter skattas man variansen för denna komponent genom att dra ifrån det slumpmässiga felets medelkvadratsumma från den aktuella medelkvadratsumman och sedan dividera med talet n/a , där n är det totala antalet försök.

I det aktuella fallet skattas vi operatörsvariationen till

$$s_B^2 = \frac{0.024 - 0.003}{60/3} = 0.00105$$

och motsvarande standardavvikelse blir $\sqrt{0.00105} = 0.032$ g.

Den totala mätosäkerheten består i detta fall av en sammanvägning av operatörs- och slumpvariationen,

$$\begin{aligned} u_y &= \sqrt{\widehat{Var}[B] + \widehat{Var}[\varepsilon]} = \sqrt{\frac{MS_O - MS_E}{20} + MS_E} = \\ &= \sqrt{\frac{MS_O + 19MS_E}{20}} = \sqrt{\frac{0.024 + 19 \cdot 0.003}{20}} = 0.064. \end{aligned}$$

För att använda denna standardosäkerhet för konfidensintervall behöver vi också beräkna dess antal frihetsgrader. Det finns i detta fall inget exakt sätt att beräkna frihetsgraderna, men en approximation som motsvarar formeln (2.21) är följande

$$\nu_{eq} = \frac{u_y^4}{\left(\frac{1}{20}\right)^2 \frac{(MS_V)^2}{\nu_V} + \left(\frac{19}{20}\right)^2 \frac{(MS_E)^2}{\nu_E}} = \frac{0.064^4}{\left(\frac{1}{20}\right)^2 \frac{0.024^2}{2} + \left(\frac{19}{20}\right)^2 \frac{0.003^2}{48}} = 18,$$

där vi avrundat nedåt till närmaste heltal. Vi ser att approximationsformeln här är i enlighet med (2.21), fast man här viktar varje variationskomponent i enlighet med viktningen i sammanvägningen.

Exempel på användning av observationsmetoden finns i exemplet skjutmått, avsnitt 6.2 och i avsnitt 5.2.1 där mätsystemsanalysen "MSA" i kvalitetsstandarden QS9000 behandlas.

3.6 Analysmetoden

I detta avsnitt behandlas den metod för mätosäkerhetsanalys som rekommenderas i ISO-standarden GUM. Först sammanfattas metoden i en arbetsgång för bestämning och angivande av mätosäkerheten i en mätstorhet i form av en process i tio steg, sedan beskrivs de ingående stegen mer detaljerat.

Då vi relaterar till den metodik som rekommenderas i GUM för globala mätosäkerhetsangivelser betraktar vi två olika typer av osäkerheter, typ A resp. B, där typ A betyder att standardosäkerheten bestämts med hjälp

av statistiska försök, medan typ B anger att man använt andra metoder för bestämningen, t. ex. bedömningar. När alla osäkerhetskomponenter tagits fram betraktas de som statistiska standardavvikelser och vägs samman kvadratiskt till den slutliga mätosäkerheten u_y .

3.6.1 Modell för mätsituationen

Vid analys av mätosäkerhet bör man först formulera en modell för mätsituationen, där inverkan av olika felkällor på den aktuella mätstorheten framgår. Vid vägning, t. ex. kommer oavsiktlig smuts på vågen att ge ett direkt tillskott av vikt och denna källa till mätfel kan ses som en ren additiv komponent i mätningen. Om vägningen utförs i en dragig lokal så måste man ha ett samband mellan lufthastighet i meter per sekund och mätfel i gram för att kunna analysera denna källas påverkan på vägningsfelet. Vi kan skriva ner en matematisk modell för mätstorheten,

$$y = \hat{y} + e_1 + e_2 + \dots = f(x_1 + \varepsilon_1, x_2 + \varepsilon_2, \dots) + e_1 + e_2 + \dots,$$

där e_1, e_2, \dots är fel som direkt adderas till mätningen, x_1, x_2, \dots är underliggande nominella värden på variabler som i sin tur är behäftade med fel $\varepsilon_1, \varepsilon_2, \dots$ och $f(\cdot)$ beskriver sambandet mellan mätvärdet y och de ingående variablerna. I vägningsexemplet är smutsen på vågen kanske e_1 , den nominella lufthastigheten i rummet, $x_1 = 0$ och dess eventuella avvikelse från detta är ε_1 , vars standardosäkerhet måste bestämmas. Den nominella relativa luftfuktigheten är kanske $x_2 = 60\%$ och avvikelser från denna representeras av ε_2 . Här måste man alltså dels bestämma standardosäkerheterna för respektive ingående felkälla, dels bestämma dess inverkan på mätstorheten.

3.6.2 Arbetsgång

1. Bestäm alla källor till osäkerhet i mätvärdet \hat{y} , se avsnitt 3.4.
2. Definiera det sanna värde y som mätvärdet \hat{y} skall representera, se avsnitt 3.3.
3. Bestäm standardosäkerheterna för alla mätvariabler som ingår i osäkerhetskällorna,

$$u_i, i = 1, 2, \dots$$

- Gör statistiska försök, kalibrering (typ A)
 - Om detta inte är möjligt, gör preliminär konservativ bedömning (typ B)
4. För varje mätvariabel som inte påverkar y direkt (i samma enhet), bestäm dess känslighetskoefficient c_i . För övriga variabler sätt känslighetskoefficienten lika med 1.

5. Sortera beloppen av alla bidrag, $|c_i| u_i$, i fallande storleksordning. Försumma alla bidrag som understiger 20% av det största, dvs. vars osäkerhetsbidrag uppfyller

$$|c_i| u_i < 0.2 \cdot \max [|c_i| u_i].$$

6. Om osäkerhetskomponenterna kan antas vara oberoende, bestäm den sammanvägda standardosäkerheten för det aktuella mätvärdet,

$$u_y = \sqrt{c_1^2 u_1^2 + c_2^2 u_2^2 + \dots}$$

7. Granska kritiskt alla de mätosäkerheter av typ B som inte har försummats. Planera framtida statistiska försök för att kunna ersätta dem eller omorganisera mätrutiner för att eliminera dem.

8. Beräkna ekvivalent antal frihetsgrader för den slutgiltiga standardosäkerheten genom ev. sammanvägning av bidrag från de olika typ A-osäkerheterna,

$$\nu_{eq} = \frac{u_y^4}{\frac{u_1^4}{\nu_1} + \frac{u_2^4}{\nu_2} + \dots}$$

Om ν_{eq} är minst 11 så använd täckningsfaktorn 2, om inte bestäm täckningsfaktorn k ur t -fördelningstabell.

9. Beräkna den utvidgade mätosäkerheten,

$$U_y = k \cdot u_y.$$

10. Beräkna ett jämförande mått för att bedöma de olika källornas inflytande på den totala mätosäkerheten, $|c_i| u_i t_{\nu_i}$.

Tabell 1. Osäkerhetstabell

Källa	u_i	typ	ν_i	$ c_i $	$u_i c_i $	$t_{\nu_i} u_i c_i $
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
y	$u_y = \sqrt{\sum c_i^2 u_i^2}$	-	ν_{eq}	-	-	-

Anmärkning 3.6.1 Osäkerheter av typ B kan i de flesta fall endast accepteras som preliminärt resultat. En möjlighet att göra den försumbar är att avrunda mätvärdet så att avrundningsfelets osäkerhetsbidrag blir så stort att det största typ B-bidraget kan försummas. En annan möjlighet kan vara att arrangera mätningen på ett sådant sätt att osäkerheten kan elimineras.

Anmärkning 3.6.2 Upprepade mätningar ger en minskad mätosäkerhet för medelvärdet:

$$u_{\bar{y}} = \sqrt{\frac{u_{y_1}^2}{n} + u_{y_2}^2}$$

där $u_{y_1}^2$ är summan av de kvadrerade osäkerhetsbidragen som motsvarar repeterbarhet, dvs. fel som varierar slumpmässigt mellan de upprepade mätningarna, medan $u_{y_2}^2$ är summan av de kvadrerade osäkerhetsbidragen som måste betraktas som konstanta mellan de upprepade mätningarna.

3.6.3 Bestäm standardosäkerheterna

Då alla komponenter i en mätosäkerhetsutredning betraktas som slumpvariabler är det naturligt att den bästa metoden för att uppskatta deras storlek är att göra statistiska försök. Detta innebär alltså att man bör sträva efter att endast använda osäkerhetsskattningar av typ A. I detta avsnitt kommer vi att ta upp de grundläggande statistiska metoderna för skattning av osäkerhetskomponenter av typ A, samt hur man bedömer motsvarande osäkerheter när man tvingas att acceptera typ B.

Bestämning av u_i , typ A

Detta innebär bestämning genom observationsmetoden på en viss osäkerhetskomponent. För att bestämma dessa osäkerheter måste man följaktligen först definiera aktuell population. Populationen vid det statistiska försöket måste representera den population som mätosäkerheten skall avse. Vill man t. ex. bestämma mätosäkerhetsbidraget från olika operatörer inom ett företag väljer man som försökspopulation alla operatörer eller ett slumpmässigt urval av tänkbara framtida operatörer. Populationen får sedan genomföra en mätning på samma objekt och mätosäkerhetskomponenten och dess frihetsgradsantal skattas som:

$$u_1 = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2} \quad , \quad \nu_1 = n - 1 \quad ,$$

där n är antalet försöksoperatörer, x_j är mätvärdet för operatör nr j och \bar{x} är medelvärdet av de n mätvärdena och därmed representerar den referens som mätosäkerheten avser. Genom att göra flera mätserier kan man i samma statistiska försök skatta olika mätosäkerhetskomponenter, se det tidigare avsnittet om variansanalys, avsnitt 3.5.4.

Kalibrering av systematiska fel Kalibrering är en typ av statistiska försök som används för att bestämma systematiska fel hos instrument eller

lokala normaler. Resultatet kan användas för justering eller, om man väljer att avstå från justering, för skattning av det systematiska felets varians. Se kapitel 4.

Förstörande mätning Ibland kommer man vid mätning av ett objekt att förstöra det eller väsentligen förändra dess egenskaper med avseende på den intressanta mätstorheten. Det är då inte möjligt att skatta mätosäkerheten med statistiskt experiment eftersom man tvingas byta produkt vid varje mätning. Spridningen mellan produkttegenskaperna blandas då med mätosäkerheten och den skattade osäkerhetskomponenten blir

$$\sqrt{u_1^2 + u_{prod}^2}$$

Detta är en överskattning av den eftersökta mätosäkerheten och man måste därmed använda denna skattning under konstaterandet att

$$u_1 < \sqrt{u_1^2 + u_{prod}^2}$$

Notera att varje ansträngning att minska variationen hos produkterna vid det statistiska försöket kommer att ge en skattning som närmar sig den sanna.

Bestämning av u_i , typ B

Bedömning av mätosäkerheter kan göras i följande fall:

- vid avrundningsfel
- när det inte är möjligt att göra ett statistiskt försök
- när man kan kosta på sig en grov överskattning av den totala mätosäkerheten
- när man ännu inte hunnit genomföra ett statistiskt försök

Avrundning Vid avrundning av ett mätvärde till ett bestämt antal decimaler bör man använda avrundningsregeln: avrunda till närmaste högre siffra i sista decimalen om de följande decimalerna är större än eller lika med 5, annars nedre. Denna regel innebär under vissa förutsättningar att avrundningsfelet får en likformig fördelning och om man avrundar till p decimaler:

$$u_r = \frac{10^{-p}}{2\sqrt{3}}$$

Vid avrundning av talet 5.6789 till två decimaler får man alltså 5.68 och osäkerheten för det avrundade värdet är $\frac{10^{-2}}{2\sqrt{3}} = 0.0028 = 0.003$. Notera att

detta resultat förutsätter att populationen av mätvärden har en större variation än avrundningsintervallet 10^{-p} . Om man t. ex. vill ange avrundningsfelets fördelning för ett mätinstrument med mätområdet 1-10 så gäller den likformiga fördelningen om man avrundar till en decimal eftersom $10^{-1} \ll 9$. Här utgörs alltså populationen av alla mätvärden mellan 1 och 10. I ett fall då man gör upprepade mätningar, och vill förbättra noggrannheten genom medelvärdesbildning, gäller att avrundningsfelet inte får överstiga det slumpmässiga felet före avrundning. Om man vid upprepade mätningar t. ex. får värdena 5.23, 5.20, 5.24, och 5.21 så kommer avrundning till en decimal att ge värdena 5.2, 5.2, 5.2, resp. 5.2 och medelvärdesbildning skulle inte förbättra precisionen. Här är alltså populationens variation i storleksordningen ± 0.02 , medan avrundningen är fem gånger större.

Specifierade osäkerhetsintervall Ibland är det inte möjligt att göra ett statistiskt försök. Om man exempelvis har inhandlat ett instrument med angiven "accuracy 1%" så vet man inte om tillverkaren bestämt denna osäkerhet med statistiskt försök eller på något annat sätt; man måste göra en bedömning.

Relevans i osäkerhetsangivelse I vissa sammanhang är spridningen i produkter så stor och mätosäkerheten så liten att även en grov överskattning av den är tillräcklig för att bedöma skillnader i produkter. I sådana fall är det naturligtvis fullt tillräckligt att göra en bedömning utifrån erfarenhet eller specifikationer.

Tillfällig bedömning I övriga fall kan man tvingas till en bedömning på grund av brist på data, men man bör då sträva efter att genomföra statistiska försök. Internt genom att organisera mätningar under kontrollerad population, externt genom att delta i Round-Robin-provningar/mätningar.

3.6.4 Bestäm känslighetskoefficienter

Om den intressanta mätstorheten inte är direkt påverkad av de ingående osäkerhetskomponenterna måste deras inverkan kvantifieras. Detta görs genom en känslighetsanalys, antingen experimentellt eller numeriskt genom differensanalys, eller analytiskt genom derivering.

Antag att mätstorheten y beror på storheterna x_1, x_2, \dots, x_m via någon funktion $y = f(x_1, x_2, \dots, x_m)$. Denna funktion kan vara känd i form av en formel, t. ex. bestäms brottgränsen hos ett materialprov som

$$R_m = \frac{F_{\max}}{A}$$

där F_{\max} är den maximalt registrerade kraften i ett dragprov och A är tvärsnittsarean i provet. Här är alltså

$$y = R_m \quad \text{en funktion av två variabler,}$$

$$f(x_1, x_2, \dots, x_m) = f(x_1, x_2) = f(F_{\max}, A) = \frac{F_{\max}}{A}.$$

Här är båda de ingående storheterna källor till mätosäkerhet, men deras påverkan på osäkerheten i R_m beror på det aktuella funktionssambandet. I sådana fall kan känslighetskoefficienterna bestämmas genom numerisk differensanalys eller analytisk derivering. I andra fall är funktionen gömd i något datorprogram och kan inte skrivas explicit och då är endast den numeriska differensanalysen användbar. För vissa osäkerhetskällor känner man kanske inte alls sambandet, t. ex. hur en temperaturförändring påverkar en volt-meters mätvärde, och ett experiment måste genomföras för att bestämma känsligheten. Gemensamt för de olika metoderna är att de utgår från Gauss approximationsformel

$$u_y^2 = \sum_{i=1}^m c_i^2 u_i^2 + \sum_{i=2}^m \sum_{j=1}^{i-1} c_i c_j u_i u_j \text{Korr}(x_i x_j),$$

där dubbelsumman innehåller alla korrelationskoefficienter mellan de ingående osäkerhetskällorna. Dessa korrelationer är oftast lika med noll och formeln förenklas då till

$$u_y^2 = \sum_{i=1}^m c_i^2 u_i^2, \quad (3.9)$$

där känslighetskoefficienterna c_i är lika med derivatorna för den aktuella funktionen $f(x_1, x_2, \dots, x_m)$,

$$c_i = \frac{\partial f}{\partial x_i}.$$

Om korrelationer bedöms vara väsentliga måste dessa tas med i beräkningen, se avsnitt 5.3.2, si. 110.

Differensmetod

Känsligheten kan bestämmas med hjälp av ett experimentellt eller ett numeriskt försök, vilket vi kallar differensmetoden. Detta innebär att derivatorna bestäms genom differenskvoter:

$$c_i = \frac{f(\tilde{x}_1, \dots, \tilde{x}_i + u_i, \dots, \tilde{x}_m) - f(\tilde{x}_1, \dots, \tilde{x}_i - u_i, \dots, \tilde{x}_m)}{2u_i}, \quad (3.10)$$

där \tilde{x}_i är det värde på variablerna som mätosäkerheten skall gälla för. I denna differenskvot hålls alltså alla variabler konstanta förutom den som

man skall bestämma känslighetskoefficienten för. Denna variabel störs istället med precis en standardosäkerhet, positiv resp. negativt och differensen mellan funktionsvärdena divideras med differensen mellan variabelvärdena. Med denna definition av känslighetskoefficienten blir den totala mätosäkerheten helt enkelt

$$u_y^2 = \sum_{i=1}^m c_i^2 u_i^2 = \frac{1}{4} \sum_{i=1}^m [f(\tilde{x}_1, \dots, \tilde{x}_i + u_i, \dots, \tilde{x}_m) - f(\tilde{x}_1, \dots, \tilde{x}_i - u_i, \dots, \tilde{x}_m)]^2.$$

Störningarna $\tilde{x}_i + u_i$ respektive $\tilde{x}_i - u_i$ görs antingen numeriskt genom datorberäkningar eller experimentellt genom försök vid olika värden på de ingående variablerna. Vid många ingående osäkerhetskällor vars känslighet skall bestämmas genom experiment kan man genomföra försöken effektivare genom ett reducerat tvåfaktorförsök, varvid man dessutom kan få skattningar utav eventuella korrelationskoefficienter, se t. ex. Montgomery[7].

Exempel 3.6.1 *Brottgränsbestämning.* Anta att den maximala kraften vid ett dragprov är uppmätt till $\hat{F}_{\max} = 100 \text{ kN}$ för en cylindrisk provstav med den uppmätta diametern $\hat{d} = 10 \text{ mm}$ och att standardosäkerheterna för de ingående storheterna är $u_{F_{\max}} = 1.2 \text{ kN}$ och $u_d = 0.02 \text{ mm}$. Tvärsnittsarean för provstaven är då $\pi d^2/4$ och formeln för brottgränsen är

$$R_m = \frac{4F_{\max}}{\pi d^2} \quad (3.11)$$

Övergång till SI-enheter ger

$$\begin{aligned} \hat{F}_{\max} &= 100 \cdot 10^3 \text{ N}, \quad \hat{d} = 10 \cdot 10^{-3} \text{ m} \\ u_{F_{\max}} &= 1.2 \cdot 10^3 \text{ N}, \quad u_d = 0.02 \cdot 10^{-3} \text{ m} \end{aligned}$$

och resultatet av mätningen blir

$$\tilde{R}_m = \frac{4\tilde{F}_{\max}}{\pi\tilde{d}^2} = \frac{4 \cdot 100 \cdot 10^3}{\pi (10 \cdot 10^{-3})^2} = 1.27 \cdot 10^9 \text{ Pa} = 1270 \text{ MPa}$$

Känslighetskoefficienterna bestäms enligt (3.10):

$$\begin{aligned} c_1 = c_{F_{\max}} &= \frac{f(\tilde{x}_1 + u_1, \tilde{x}_2) - f(\tilde{x}_1 - u_1, \tilde{x}_2)}{2u_1} = \\ &= \frac{f(\hat{F}_{\max} + u_{F_{\max}}, \hat{d}) - f(\hat{F}_{\max} - u_{F_{\max}}, \hat{d})}{2u_{F_{\max}}} = \\ &= \frac{\frac{4(\hat{F}_{\max} + u_{F_{\max}})}{\pi\hat{d}^2} - \frac{4(\hat{F}_{\max} - u_{F_{\max}})}{\pi\hat{d}^2}}{2u_{F_{\max}}} = \end{aligned}$$

$$\frac{4 \cdot 10^3 (100 + 1.2) - 4 \cdot 10^3 (100 - 1.2)}{2 \cdot 1.2 \cdot 10^3 \cdot \pi \cdot (10 \cdot 10^{-3})^2} = 12732$$

$$c_2 = c_d = \frac{f(\tilde{x}_1, \tilde{x}_2 + u_2) - f(\tilde{x}_1, \tilde{x}_2 - u_1)}{2u_2} =$$

$$\frac{f(\widehat{F}_{\max}, \widehat{d} + u_d) - f(\widehat{F}_{\max}, \widehat{d} - u_d)}{2u_d} =$$

$$\frac{\frac{4\widehat{F}_{\max}}{\pi(\widehat{d}+u_d)^2} - \frac{4\widehat{F}_{\max}}{\pi(\widehat{d}-u_d)^2}}{2u_d} =$$

$$\frac{\frac{4 \cdot 100 \cdot 10^3}{\pi(10+0.02)^2 \cdot 10^{-6}} - \frac{4 \cdot 100 \cdot 10^3}{\pi(10-0.02)^2 \cdot 10^{-6}}}{2 \cdot 0.02 \cdot 10^{-3}} = -2.5 \cdot 10^{11}$$

Den totala osäkerheten blir under antagande om oberoende mellan diame-
termätningen och kraftmätningen:

$$u_{R_m}^2 = \sum_{i=1}^m c_i^2 u_i^2 =$$

$$c_1^2 u_1^2 + c_2^2 u_2^2 = (12732)^2 \cdot (1.2 \cdot 10^3)^2 + (-2.5 \cdot 10^{11})^2 \cdot (0.02 \cdot 10^{-3})^2 =$$

$$2.58 \cdot 10^{14}$$

$$u_y = \sqrt{2.58 \cdot 10^{14}} = 1.6 \cdot 10^7 \text{ Pa} = 16 \text{ MPa}$$

Partiell derivering

När funktionen $y = f(x_1, x_2, \dots, x_m)$ kan beskrivas som ett matematiskt uttryck så kan känslighetskoefficienterna bestämmas genom partiell derivering. I fallet med brottgräns (3.11) får man då resultatet

$$c_1 = c_{F_{\max}} = \frac{\partial f}{\partial x_1} = \frac{\partial R_m}{\partial F_{\max}} = \frac{4}{\pi d^2} \quad , \quad c_d = \frac{\partial R_m}{\partial d} = -\frac{8F_{\max}}{\pi d^3} \quad (3.12)$$

och den totala mätosäkerheten blir

$$u_{R_m}^2 = \sum_{i=1}^m c_i^2 u_i^2 = c_{F_{\max}}^2 u_{F_{\max}}^2 + c_d^2 u_d^2 = \frac{16}{\pi^2 \widetilde{d}^4} u_{F_{\max}}^2 + \frac{64 \widetilde{F}_{\max}^2}{\pi^2 \widetilde{d}^6} u_d^2$$

där åter tildetecknet $\widetilde{}$ betecknar att osäkerhetsangivelsen avser just de bestämda värdena på den uppmätta kraften respektive arean.

Exempel 3.6.2 Tillämpning av de numeriska värdena från förra exemplet ger

$$u_{R_m}^2 = \frac{16}{\pi^2 \tilde{d}^4} u_{F_{\max}}^2 + \frac{64 \tilde{F}_{\max}^2}{\pi^2 \tilde{d}^6} u_d^2 =$$

$$\frac{16}{\pi^2 (10 \cdot 10^{-3})^4} (1.2 \cdot 10^3)^2 + \frac{64 (100 \cdot 10^3)^2}{\pi^2 (10 \cdot 10^{-3})^6} (0.02 \cdot 10^{-3})^2 =$$

$$2.6 \cdot 10^{14}$$

$$u_{R_m} = \sqrt{2.6 \cdot 10^{14}} = 1.6 \cdot 10^7 \text{ Pa} = 16 \text{ MPa}$$

Vi ser att resultatet blir detsamma med de olika metoderna.

Förenkling vid relativa osäkerheter Om man arbetar med relativa osäkerheter kommer de partiella derivatorna att bli avsevärt förenklade om funktionen f är ett rent multiplikativt samband,

$$f(x_1, x_2, \dots, x_m) = C \cdot x_1^{p_1} \cdot x_2^{p_2} \cdot \dots \cdot x_m^{p_m},$$

där C är någon deterministisk konstant. Det visar sig att om man genomför den partiella deriveringen enligt ovanstående på en dylik form så blir den resulterande relativa osäkerheten:

$$\frac{u_y^2}{y^2} = p_1^2 \frac{u_1^2}{x_1^2} + p_2^2 \frac{u_2^2}{x_2^2} + \dots + p_m^2 \frac{u_m^2}{x_m^2} \quad (3.13)$$

Exempel 3.6.3 Om vi återtar vårt brottgränsexempel så blir resultatet

$$\frac{u_{R_m}^2}{\tilde{R}_m^2} = p_1^2 \frac{u_1^2}{x_1^2} + p_2^2 \frac{u_2^2}{x_2^2} + \dots + p_m^2 \frac{u_m^2}{x_m^2} = \frac{u_{F_{\max}}^2}{\tilde{F}_{\max}^2} + (-2)^2 \frac{u_d^2}{\tilde{d}^2} =$$

$$\frac{(1.2 \cdot 10^3)^2}{(100 \cdot 10^3)^2} + 4 \frac{(0.02 \cdot 10^{-3})^2}{(10 \cdot 10^{-3})^2} = 0.00016$$

$$\frac{u_{R_m}}{\tilde{R}_m} = \sqrt{0.00016} = 0.0126 = 1.26\%$$

och 1.3% av den nominella brottgränsen blir $1.26 \cdot 1270 / 100 \text{ MPa} = 16 \text{ MPa}$ i överensstämmelse med resultaten från de två tidigare metoderna.

3.6.5 Sammanvägning av relevanta osäkerhetskomponenter

Ett mätresultat kan, efter eventuell justering för kända fel, modelleras som en summa av det sanna värdet och ett antal felbidrag:

$$\hat{y} = y + f(e_1, e_2, \dots, e_m) \approx y + e$$

där de olika felbidragen summeras till ett gemensamt fel e , vars statistiska fördelning man vill bestämma. Om man justerat för alla kända systematiska fel så kommer felet e att ha väntevärdet noll och om man antar att alla felbidrag är oberoende kan variansen för felet approximeras med Gauss approximationsformel:

$$u_{\hat{y}}^2 = \sum_{i=1}^m c_i^2 u_{e_i}^2,$$

Om flera felkällor väsentligen bidrar till det gemensamma felet så kommer fördelningen för denna summa att bli approximativt normalfördelad och därmed kan man bestämma ett 95%-igt konfidensintervall för det sanna värdet y genom att multiplicera standardosäkerheten med två:

$$y = \hat{y} \pm 2 \cdot u_{\hat{y}} \quad (\text{med ca } 95\% \text{ sannolikhet})$$

3.7 Hur skall man uppnå en förbättrad mätnoggrannhet?

Att förbättra mätnoggrannheten i en mätsituation är detsamma som att

- se till att man har väntevärdesriktiga mätningar, dvs. att man i medeltal kommer att uppnå det sanna mätvärdet. Detta åstadkommer man genom **kalibrering**, vilket behandlas i nästa avnitt.
- minska variationen vid upprepade mätningar. Utifrån vårt statistiska perspektiv kan man göra detta på två olika sätt, antingen genom att **begränsa populationerna** av felbidrag eller genom använda **medelvärden av upprepade mätningar**.

3.7.1 Begränsa populationer

Vi återknyter till figur 1.6 på sidan 8. Felet i mätvärdet kan ses som ett resultat av avvikelser från önskvärt värde på en rad olika variabler, som var och en påverkar utfallet av mätningen och mätosäkerheten bestäms av hur mycket variablerna kan tänkas avvika från det önskvärda värdet. En minskad osäkerhet fås alltså genom att minska de tänkbara avvikelserna; begränsa populationerna av variabelvärden. Flera olika sätt kan tänkas för att åstadkomma en sådan begränsning av populationerna.

1. Definiera ett snävare tillämpningsområde. Vid exempelvis mätningar som är känsliga för omgivande temperatur kan man specificera snäva temperaturgränser och därmed definiera en mätmetod som ger minskad variation. Medelvärdet av det snävare området bör förstås vara lika med det nominella värdet som definierar felfria förhållanden. I en mätning där operatörsinflytandet är väsentligt för variationen kan man begränsa antalet operatörer. Genom en speciell kalibrering kan den utvalda operatörsgruppens medelvärde jämföras med sanningen och eventuell justering införs. Varje dylikt tillämpningsområde får sin egen mätosäkerhet (repetierbarhet) definierad, som endast gäller vid repetition inom tillämpningsområdet.
2. Bättre kalibreringsrutiner. Instrument vars mätosäkerhet bestäms av sin specifikation kan ges en mindre spridning genom regelbunden kalibrering och klassificering. Instrumentosäkerhet som angetts för ett stort mätområde kan minskas för specifika tillämpningar genom att kalibrering sker i ett snävare mätområde. För mätsystem som ändras med tiden kan variationen minskas genom att införa tätare kalibreringar.
3. Eliminera enskilda felkällor. Vid analys av mätsystem med hjälp av t. ex. variansanalys kan man skatta de olika faktorerna och identifiera eventuella stora avvikelser. Detta kan leda till åtgärder såsom utbyte av instrument eller utbildning av operatörer för att minska populationen av felbidrag. I tabellen på sidan 64 är medelvärdena för de olika operatörerna angivna i den högra kolumnen. Här kan man alltså se om någon enstaka operatör har en exceptionell avvikelse och åtgärda detta. I t. ex. ett trädidiagram enligt figur 3.9 kan man också söka finna extrema avvikelser som kan vara möjliga att eliminera. I dylika grafiska framställningar kan man också identifiera om t. ex. mätvariationen ändras sig med storleken på själva mätvärdet.
4. Genomföra jämförande mätningar under konstanta förhållanden. Genom att tillfälligt hålla vissa variabler konstanta kan man genomföra jämförande mätningar med mycket små mätfel. Detta är just den ena av de två kungsvägarna till förbättrad mätnoggrannhet som anges på sidan 10. Det innebär dels att man eliminerar eventuella avvikelser från det nominella värdet, dels att man undviker spridning mellan mätningar inom de variabler som hålls konstanta.

3.7.2 Medelvärdesbildning

För felkällor som ändras sig för varje mättillfälle kan man använda medelvärdet av flera mätningar för att få ett noggrannare mätvärde. I teorin gäller att om man tar medelvärdet av n st. mätningar av Y där varje mätning har

mätosäkerheten u så blir mätosäkerheten för medelvärdet

$$u_{\bar{Y}} = \frac{u}{\sqrt{n}}$$

För att detta skall vara sant i verkligheten skall felen tas slumpmässigt ur de dominerande felkällorna, dvs. ur hela populationen av dessa felbidrag. Detta är mycket viktigt att beakta och kräver därför att man i varje mätosäkerhetsutredning klart håller isär de osäkerhetskomponenter som är av systematisk natur och de som är slumpmässiga.

3.8 Sammanfattning

Mätresultatet \hat{y} är behäftat med ett okänt mätfel. Mätosäkerhetsutredningen syftar till att finna ett intervall kring mätvärdet som täcker det sanna värdet med 95% konfidens. Resultatet av en sådan utredning ger möjlighet att ange mätvärdet med sitt osäkerhetsintervall i enlighet med:

$$y = \hat{y} \pm U_y \quad (\text{med ca 95\% konfidens}).$$

Här är \hat{y} det uppmätta värdet och U_y är den utvidgade mätsosäkerheten. Denna erhålles genom att multiplicera standardosäkerheten u_y med täckningsfaktorn k . För 95% konfidens gäller att

$$k = 2$$

om standardosäkerheten u_y är sammanvägd av minst två felkällor i ungefär samma storlek och bestämd med minst 11 frihetsgrader (tumregler 3 och 4, sid. 49).

Standardosäkerheten u_y sammanfattar samtliga felkällors inflytande på mätresultatet och skall ses som den statistiska standardavvikelsen hos en normalfördelad mätvariabel \hat{y} .

En mätosäkerhetsutredning genomförs på följande sätt:

- Definiera ändamålet med mätningen. Skall mätresultatet kunna jämföras med motsvarande mätning i ett globalt perspektiv måste osäkerhetskällorna relateras till en globalt överenskommen sanning, t. ex. genom en kontrollerad kalibreringskedja till en överenskommen normal. Skall mätresultatet däremot användas endast för lokala jämförelser kan osäkerhetskällor som är systematiska eliminieras (kungsväg 1, sid. 10). Ändamålet definierar också de populationer av felvariabler som skall analyseras.
- Bestäm alla källor till osäkerhet i mätvärdet \hat{y} , se avsnitt 3.4 sid. 54.
- Definiera det sanna värde y som mätvärdet \hat{y} skall representera och de populationer av felvariabler som kan uppträda för den aktuella ändamålet, se avsnitt 3.3 sid. 53.

Det finns två grundläggande sätt att bestämma standardosäkerheten u_y , nämligen genom observationsmetoden eller genom analysmetoden.

Observationsmetoden består i att man genomför upprepade mätningar på ett objekt, där samtliga väsentliga felvariabler varierar inom sina respektive populationer mellan mätningarna **på ett slumpmässigt sätt** och

att man säkerställer att medelinflytandet av dessa populationer är lika med noll. Resultatet av dessa mätningar ger en direkt skattning av standardosäkerheten genom formeln

$$u_{\bar{y}}^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2,$$

där \bar{y} är medelvärdet av de n mätningarna och skattningen har $n-1$ frihetsgrader. Denna metod har fördelen att vara robust mot misstag såsom förbiseende av felkällor och räknefel, men kan vara svår att genomföra och kan ge felaktiga resultat om man misslyckas i det slumpmässiga urvalet ur felkällorna.

Analysmetoden består i att man analyserar felbidragens inflytande på mätresultatet, bestämmer varje källas standardosäkerhet för sig och väger samman samtliga bidrag till en standardosäkerhet i \hat{y} . Mätresultatet ses som en summa av det sanna värdet och ett antal felbidrag:

$$y = \hat{y} + e_1 + e_2 + \dots = f(x_1 + \varepsilon_1, x_2 + \varepsilon_2, \dots) + e_1 + e_2 + \dots, \quad (3.14)$$

där e_1, e_2, \dots är fel som direkt adderas till mätningen, x_1, x_2, \dots är underliggande nominella värden på variabler som i sin tur är behäftade med fel $\varepsilon_1, \varepsilon_2, \dots$ och $f(\cdot)$ beskriver sambandet mellan mätvärdet y och de ingående variablerna. Analysen går ut på att bestämma standardosäkerheter för samtliga väsentliga felvariabler $u_{e_1}, u_{e_2}, \dots, u_{\varepsilon_1}, u_{\varepsilon_2}, \dots$ och dessutom bestämma hur mycket variationen i ε -variablerna påverkar y . För överskådlighetens skull skiljer vi i fortsättningen inte på beteckningarna för direkt och indirekt verkande felvariabler utan numrerar dem i en serie.

- Bestäm standardosäkerheterna för alla mätvariabler som ingår i osäkerheteskällorna: u_1, u_2, \dots . Detta genomförs företrädesvis med hjälp av statistiska försök (typ A), dvs. med observationsmetoden, varav kalibrering är ett specialfall. Notera antalet frihetsgrader i varje osäkerhetsskattning i enlighet med observationsmetoden.

Om det inte är möjligt att genomföra statistiska försök, gör man en **preliminär konservativ bedömning** (typ B). Om variationskällorans population är känd att vara begränsad till ett visst intervall $\pm a$ (t. ex. angiven noggrannhet hos ett instrument) så är det lämpligt att bedöma osäkerheten som likformigt fördelad inom intervallet och därmed sätta standardosäkerheten hos denna komponent till

$$u = \frac{a}{\sqrt{3}}.$$

Om man istället kan bedöma gränser $\pm b$ för vilka det gäller att de innefattar 95% av populationen så sätts standardosäkerheten till

$$u = \frac{b}{2}.$$

Antalet frihetsgrader i typ B-skattningar sätts till oändligheten.

- För mätvariabler som verkar direkt på y (e i 3.14) sätts känslighetskoefficienten till ett. För varje mätvariabel som inte påverkar y direkt (ε i 3.14) bestäms dess känslighetskoefficient c_i . Detta görs företrädesvis med differensmetoden:

$$c_i = \frac{f(\tilde{x}_1, \dots, \tilde{x}_i + u_i, \dots, \tilde{x}_m) - f(\tilde{x}_1, \dots, \tilde{x}_i - u_i, \dots, \tilde{x}_m)}{2u_i},$$

där \tilde{x}_i är det värde på variablerna som mätosäkerheten skall gälla för. I denna differenskvot hålls alltså alla variabler konstanta förutom den som man skall bestämma känslighetskoefficienten för. Denna variabel störs istället med precis en standardosäkerhet, positiv resp. negativt och differensen mellan funktionsvärdena divideras med differensen mellan variabelvärdena.

Notera att denna metod är tillämplig även om funktionen f inte är känd i sin analytiska form utan bara tillgänglig i form av ett datorprogram. Den är även tillämplig om funktionen inte alls är känd, då differenskvotens två värden i täljaren bestäms genom två mätningar för varje komponent.

I fall då funktionen f är känd till sin analytiska form kan känslighetskoefficienterna bestämmas med hjälp av partiell derivering.

- Sortera beloppen av alla bidrag, $|c_i| \cdot u_i$ i fallande storleksordning. Försumma alla bidrag som understiger 20% av det största (tumregel 5, sid. 51).
- Om de olika osäkerhetskomponenterna kan betraktas som oberoende, bestäm den sammanvägda standardosäkerheten för det aktuella mätvärdet,

$$u_y = \sqrt{c_1^2 u_1^2 + c_2^2 u_2^2 + \dots}$$

Om några komponenter misstänks vara beroende av varandras värden gäller inte denna förenklade formel, utan beroendet måste analyseras speciellt, se avsnitt 5.3, sid. 108.

- Granska kritiskt alla de mätosäkerheter av typ B som inte har kunnat försummas. Planera framtida statistiska försök för att kunna ersätta dem eller omorganisera mätrutiner för att göra dem försumbara.
- Beräkna ekvivalent antal frihetsgrader för den slutgiltiga standardosäkerheten genom sammanvägning av bidrag från de olika osäkerheterna,

$$\nu_y = \frac{u_y^4}{\frac{u_1^4}{\nu_1} + \frac{u_2^4}{\nu_2} + \dots}$$

Notera att typ B-skattningarnas bidrag i nämnaren blir lika med noll eftersom deras antal frihetsgrader betraktas som oändliga.

Om ν_y är minst 11 så används täckningsfaktorn 2, annars bestäms täckningsfaktorn k ur t -fördelningstabell på sid. 29.

- Beräkna den slutgiltiga utvidgade mätosäkerheten:

$$U_y = k \cdot u_y .$$

- Beräkna slutligen ett jämförande mått för att bedöma de olika källornas inflytande på den totala mätosäkerheten, $|c_i u_i t_{\nu_i}|$, där t_{ν_i} är t -fördelningens värde enligt tabell 2.3.2 på sid. 30 och samla resultaten i en tabell:

Osäkerhetstabell						
Källa	u_i	typ	ν_i	$ c_i $	$u_i c_i $	$ c_i u_i t_{\nu_i} $
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
y	$u_y = \sqrt{\sum c_i^2 u_i^2}$	-	ν_{eq}	-	-	-

Den slutgiltiga mätosäkerheten används för att ange ett 95%-igt konfidensintervall för mätvärdet och följande avrundningsregler rekommenderas för det angivna slutresultatet (tumregel 7, sid 53): Den utvidgade mätosäkerheten U_y avrundas alltid till två värdesiffror. Det angivna mätvärdet avrundas till den decimalplats som den mest signifikanta siffran hos $\frac{U_y}{2.7}$.

- Om mätosäkerheten är för stor för det önskade ändamålet kan tabellens sista kolumn användas för att prioritera förbättringsarbetet tillsammans med rekommendationerna i avsnitt 3.7.

Upprepade mätningar ger en minskad mätosäkerhet för medelvärdet:

$$u_{\bar{y}} = \sqrt{\frac{u_{y_1}^2}{n} + u_{y_2}^2}$$

där $u_{y_1}^2$ är summan av de kvadrerade osäkerhetsbidragen som motsvarar repeterbarhet, dvs. fel som man vet varierar slumpmässigt mellan de upprepade mätningarna, medan $u_{y_2}^2$ är summan av de kvadrerade osäkerhetsbidragen som måste betraktas som konstanta mellan de upprepade mätningarna.

Kapitel 4

Kalibrering

Varje mätosäkerhetsangivelse måste som tidigare påpekats relateras dels till en population av mätförhållanden, dels till en referens som representerar sanningen. För att minimera avvikelsen från sanningen i ett mätvärde genomför man **kalibrering** av mätsystemet gentemot den aktuella referensen, vilket kan beskrivas som ett sätt att *kontrollera om mätsystemets medelvärde överensstämmer med det nominella (önskvärda) värdet*. Om mätsystemet riskerar att förändras i tid genomförs sådana kalibreringar vid upprepade tillfällen, ofta med ett fördefinierat tidsintervall. Med kalibrering menas här helt enkelt en jämförelse av mätsystemets resultat med en referens. Utfallet av kalibreringen kan eventuellt ge anledning till **justering** av mätsystemet för bättre överensstämmelse. Principen för **spårbar** kalibrering och justering behandlas i avsnitt 4.1.

Vid kalibrering av ett instrument som har en linjär överföringsfunktion används kalibrering oftast för att bestämma instrumentets **känslighet**, dvs lutningen b hos den linje

$$y = a + bx$$

som representerar sambandet mellan insignalen x och utsignalen y . Konstanten a är ibland irrelevant för användningen då instrumentet nollställs såväl vid kalibrering som vid användning. Teorin bakom kalibreringen är emellertid densamma oavsett nollställning, nämligen **linjär regression**, som presenterades i avsnitt 2.5. Det specifika problemet i kalibreringssituationen är att själva kalibreringen utförs med kända invariabler x , medan man vid användningen läser av y och vill veta x . Detta problem behandlas i avsnitt 4.1.1 och slutligen behandlas den speciella proceduren **shuntkalibrering** i avsnitt 4.1.2.

4.1 Spårbar kalibrering

Vid kalibrering av mätutrustning kontrolleras utrustningen mot en referens som i sin tur är spårbart kalibrerad mot någon referens accepterad att re-

presentera sanningen. Om man antar att det förekommer en drift hos mätutrustningen så kommer den vid varje kalibrering att ha en viss systematisk avvikelse, en bias. Denna kan skattas vid kalibreringen genom att man gör upprepade mätningar av referensen. Man kan nu fråga sig om man bör justera med denna skattade bias efter varje kalibrering eller om man skall avstå.

Vi betraktar den systematiska avvikelsen som en stokastisk variabel som är normalfördelad enligt

$$B \sim N(0, \sigma_B^2)$$

Vid varje kalibreringstillfälle antas alltså en slumpmässig bias. Vid mätning med utrustningen har vi också ett normalfördelat mätfel, repeterbarheten

$$\varepsilon \sim N(0, \sigma_\varepsilon^2)$$

Vid kalibreringstillfälle i görs m mätningar med känt sant värde μ och man får

$$X_{ij} = \mu + B_i + \varepsilon_j, j = 1, 2, \dots, m$$

Avvikelsen från referensen skattas med hjälp av medelvärdet till

$$\widehat{B}_i = \overline{X}_i - \mu$$

och variansen för denna skattning är σ_ε^2/m .

Vid efterföljande mätningar med mätsystemet fås, om det sanna värdet är y :

$$\widehat{Y}_i = y + B_i + \varepsilon,$$

dvs. en summa av det sanna värdet y , den systematiska avvikelsen från referensen B_i , samt ett slumpmässigt mätfel ε .

Reproducerbarhet utan justering

Om man inte justerar med skattad bias så används därmed följande skattning av det sanna mätvärdet

$$\widehat{y}_i^{(eJ)} = \widehat{Y}_i,$$

Väntevärdet av denna skattning över kalibreringsperioden, dvs. med avseende på ε är

$$E_\varepsilon \left[\widehat{y}_i^{(eJ)} \right] = y + B_i$$

då ju B_i är konstant under perioden. Reproducerbarheten för skattningen i form av varians är

$$Var \left[\widehat{y}_i^{(eJ)} \right] = Var_\varepsilon [y + B_i + \varepsilon] = \sigma_B^2 + \sigma_\varepsilon^2$$

Variansen σ_B^2 för bias är svår att bestämma. Det bästa sättet är att organisera en Round-Robin-mätning med ett stort antal laboratorier som

representerar den globala spridningen i mätförhållanden och skatta σ_B^2 med hjälp av variansanalys. Ett annat alternativ är att använda egna upprepade kalibreringsresultat och skatta variationen i bias ur dessa, där åter frihetsgraderna bestäms av antalet upprepningar. En sådan skattning ger förstås endast en uppfattning om variationen i den egna mätutrustningens bias, men kan vara en nödvändig approximation om inga andra uppgifter finns tillgängliga.

En metod som används i vissa sammanhang är att skatta reproducerbarheten med summan av repeterbarhet och uppmätt bias,

$$s_\varepsilon^2 + \widehat{B}_i^2.$$

Denna skattning är faktiskt en variant på skattningen med Round-Robin-mätningar, men med endast en deltagare! Skattningen av biasvariansen har då bara en frihetsgrad och Welch-Satterwaite-formeln (2.21) ger antalet ekvivalenta frihetsgrader för summan.

Reproducerbarhet vid justering

Om man istället justerar med skattad bias vid varje mätning så fås följande skattning av det sanna mätvärdet

$$\widehat{y}_j^{(J)} = \widehat{Y}_i - \widehat{B}_i$$

Väntevärdet av denna skattning över kalibreringsperioden, dvs. med avseende på ε är

$$E_\varepsilon \left[\widehat{y}_j^{(J)} \right] = y$$

då ju B_i är konstant och \widehat{B}_i är en väntevärdesriktig skattning av denna bias. Variansen för skattningen är

$$\text{Var}_\varepsilon \left[\widehat{y}_j^{(J)} \right] = \text{Var}_\varepsilon \left[y + B_i + \varepsilon - \widehat{B}_i \right] = \sigma_\varepsilon^2 + \sigma_\varepsilon^2/m = \sigma_\varepsilon^2 \left(1 + \frac{1}{m} \right).$$

Antalet frihetsgrader för denna skattning ges av antalet mätningar bakom skattningen av σ_ε^2 .

Skall man justera?

Resultaten visar att om man jämför justering/icke justering blir reproducerbarheten lika om $\sigma_B^2 = \sigma_\varepsilon^2/m$, och om bias har större varians än medelvärdet av biasskattningen så kommer regelbunden justering att öka mätosäkerheten istället för att minska den. Detta problem är jämförbart med kvalitetsstyrning i en produktionsprocess och i de fall där ett mätsystem ligger i närheten av denna gräns kan man använda styrdiagramsteknik för att endast justera vid s.k. larm, se allmänt om styrdiagram i t. ex. Bergman/Klevsjö[11].

Skattningsvariansen kan ju dock göras godtyckligt liten genom att öka antalet replikat, gränssituationen kan undvikas om upprepade mätningar är möjliga och justering ger då alltid en mindre mätosäkerhet.

Accepterat biasintervall

Den optimala metodiken för justering som beskrivs ovan är ibland inte praktiskt eller ekonomiskt genomförbar i en mätprocess. Justeringar efter varje kalibrering kan hindras av organisatoriska eller ekonomiska skäl och kan också innebära risker för feljusteringar med än större mätosäkerhet som följd. Vid små avvikelser från referensvärden kan det i dylika fall vara bättre att avstå från justering. Regelbundna kalibreringar används då endast för kontroll av att avvikelserna håller sig inom vissa gränser och dessa gränser kan användas för att skatta mätosäkerheten.

Vi bestämmer oss alltså för att acceptera en avvikelse från referensvärdet vid kalibreringen som maximalt får vara $\pm B$. Detta innebär i praktiken en sorts sortering av kalibrerade instrument i acceptabla och oacceptabla. Ett sätt att modellera situationen efter sorteringen är att betrakta det systematiska felet som likformigt fördelat inom det angivna intervallet och mätosäkerhetsbidraget blir då enligt (2.6):

$$\sigma_B^2 = \frac{B^2}{3}$$

En komplikation i detta förfarande är att vi använder den uppmätta avvikelsen från referensen i sorteringen, vilken är behäftad med ett visst fel ε , eventuellt minskat genom medelvärdesbidning. Detta fel har emellertid normalt ett försumbart inflytande, se avsnitt 5.4.

4.1.1 Kalibrering med linjär regression

Många mätinstrument ger en indirekt indikation av det mätta värdet och kalibreringen innebär då ofta att bestämma den överföringsfunktion som är vid handen mellan mätvärdet och instrumentvisningen. Sålunda är många mätinstrument byggda med elektriska kretsar och insignal i form av en mätvärdesförändring i t. ex. längd, kraft, töjning, koncentration eller tryck ger en utsignal i form av en elektrisk spänning. Sambandet mellan insignal och utsignal kontrolleras genom att kända insignaler mäts och utsignalerna avläses. Ur en mängd sådana resultat kan en generell överföringsfunktion bestämmas. För en kraftgivare kan en sådan överföringsfunktion ofta tecknas som en enkel linjär funktion: Utsignalens spänning är lika med en summa av en offset a och en term som är proportionell mot kraften F med proportionalitetskonstanten k :

$$u = a + kF .$$

Instrumentets offset kan eventuellt justeras bort genom att instrumentet nollställs såväl vid kalibrering som användning och därmed återstår bara proportionalitetskonstanten k , då ofta kallad känslighet:

$$\Delta u = k \cdot \Delta F .$$

För att bestämma konstanterna a och k används linjär regression enligt teorin i avsnitt 2.5 och ett exempel på hur det går till visas i avsnitt 6.3, där vi också visar hur regressionsmetodikens prediktionsintervall används för att bestämma mätosäkerheten. Jämfört med det generella regressionsproblemet tar vi här upp två speciella problem vid kalibrering; nämligen 1) inversproblemet och 2) vår allmänna strävan efter förenklade metoder.

Inversproblemet

När ett kalibrerat instrument skall användas så kommer kalibreringskurvan att användas i motsatt riktning mot vad fallet är vid kalibreringstillfället. Vid t. ex. kraftkalibrering så kontrollerar man vid kalibreringen vad instrumentet visar (y -värdena) när det belastas med vissa kända krafter x . När instrumentet används vill man uttala sig om en okänd kraft x utifrån vad instrumentet visar, y . Detta betyder att den skattade kalibreringslinjen måste inverteras vid användning,

$$y = \beta_0 + \beta_1 x + \varepsilon \quad \Leftrightarrow \quad x = \frac{y}{\beta_1} - \frac{\beta_0}{\beta_1} - \frac{\varepsilon}{\beta_1}$$

och vi skattar x med följande uttryck av de vid kalibreringen skattade parametrarna

$$\hat{x} = \frac{y}{\hat{\beta}_1} - \frac{\hat{\beta}_0}{\hat{\beta}_1}$$

Osäkerheten vid prediktion har då inte någon enkel lösning i enlighet med (2.31). Variansen för ett predikerat värde x^* blir

$$Var[x^*] = y^2 Var\left[\frac{1}{\hat{\beta}_1}\right] + Var\left[\frac{\hat{\beta}_0}{\hat{\beta}_1}\right] + Var\left[\frac{\varepsilon}{\hat{\beta}_1}\right] + \dots$$

där ... står för de tre kovarianstermerna. Vi känner fördelningarna för de skattade paramterarna $\hat{\beta}_0$ och $\hat{\beta}_1$ men inte för kvoterna. Man kan dock ansätta en approximativ lösning och får då konfidensintervallet

$$x = \hat{x} \pm \frac{s \cdot t}{\hat{\beta}_1} \sqrt{\frac{(\hat{x} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{1}{n} + 1}$$

där t tas ur en t -fördelningstabell med $n - 2$ frihetsgrader precis som i fallet (2.31).

Ett alternativ till denna approximativa lösning är att undvika inverteringen genom att redan vid kalibreringen göra regressionen i “fel riktning”, dvs. minimera felen i x -riktningen. Modellen blir då

$$x = \beta'_0 + \beta'_1 y + \varepsilon'$$

där man direkt skattar de parametrar som skall användas vid framtida avläsningar av x . Prediktionsintervall beräknas helt i enlighet med (2.31),

$$x^*(y) = \widehat{\beta}'_0 + \widehat{\beta}'_1 y \pm s \cdot t \sqrt{\frac{(y - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} + \frac{1}{n} + 1}$$

De två linjerna “regression av x på y ” respektive “regression av y på x ” kommer inte att sammanfalla. Vilken bör då användas? Vi kan inte ge något generellt svar på denna fråga, men i de flesta fall ger de två metoderna likvärdiga resultat och man bör välja den som bäst passar in i den praktiska tillämpningen. Följande exempel är ett starkt argument för den sistnämnda metoden:

Betrakta följande observationer på längden hos far/son i cm:

Far	170	170	180	180	190	190
Son	170	180	170	190	180	190

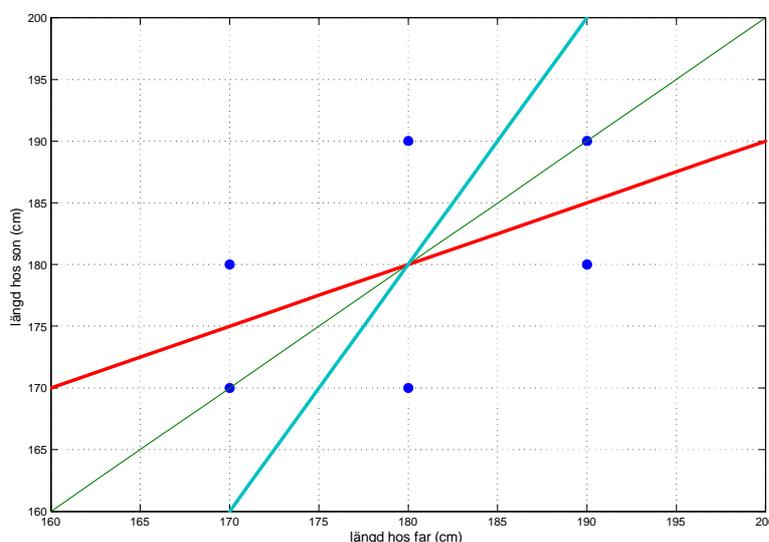
Vi ritar upp punkterna i ett diagram och ritar in tre linjära anpassningar till datapunkterna: 1) den som minimerar felen i Son-riktning (röd linje med minst lutning), 2) den som minimerar felen i Far-riktning (blå linje med störst lutning) och 3) den som man kan förvänta sig vara närmast sanningen (grön linje mellan de två övriga).

Vilken av linjerna 1) och 2) förutsäger bäst sonens längd om man känner faderns? Vilken av linjerna 1) och 2) gissar bäst faderns längd utifrån kunskap om sonens? Figuren visar att om man vill gissa sonens längd utgående från faderns så bör man minimera felen i referensdata i just Son-riktning. Om man istället ville gissa faders längd utifrån sonens så bör anpassningen göras i den andra riktningen. Båda anpassningarna avviker från den linje som borde representera det korrekta sambandet, men om framtida mätvärden motsvarar de referensvärden som anpassningen gjorts på, så är det uppenbarligen så att anpassning i den riktning som prediktioner skall göras är den som minimerar prediktionsfel.

Förenklat prediktionsintervall vid kalibrering

Enligt (2.31) i avsnitt 2.5 ges ett prediktionsintervall för en skattad linje av formeln

$$y^*(x) = \widehat{\beta}_0 + \widehat{\beta}_1 x \pm s \cdot t \sqrt{\frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{1}{n} + 1} .$$



Figur 4.1: Tre linjära anpassningar till uppmätta längder på far resp. son.

Vid kalibrering kan detta uttryck ofta vara onödigt komplicerat att använda och dessutom är det inte ens möjligt att använda om inte alla x -värden vid kalibreringen är kända. Vi skall därför härleda ett förenklat uttryck för en specifik kalibreringssituation, samt undersöka villkoren avseende antalet kalibreringspunkter för ytterligare förenkling.

För att förenkla uttrycket vill vi ersätta alla x -värden med något enklare och vi väljer då att studera det statistiska väntevärdet. Betrakta först den första termen under rottecknet:

$$\frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad . \quad (4.1)$$

Anta att man vid kalibreringen valt n stycken referensvärden jämnt fördelade inom det aktuella mätområdet:

$$x_{\min} = x_0, x_0 + \Delta, x_0 + 2\Delta, \dots, x_0 + (n-1)\Delta = x_{\max} .$$

Nämnaren i (4.1) blir då lika med

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n (x_0 + (i-1)\Delta - \bar{x})^2 = \frac{n(n^2-1)}{12} \Delta^2$$

Anta vidare att instrumentet i fortsättningen kommer att användas med lika stor frekvens inom hela mätområdet. Medelvärdet av referenspunkterna blir precis mittpunkten i mätområdet och vi betraktar framtida

mätningar som likformigt fördelade inom mätområdet. Täljaren i (4.1) ersätter vi nu med dess väntevärde över denna likformiga fördelning:

$$\begin{aligned} E[(x - \bar{x})^2] &= \text{Var}[X] = \frac{(x_{\max} - x_{\min})^2}{12} = \frac{(x_0 + (n-1)\Delta - x_0)^2}{12} = \\ &= \frac{(n-1)^2 \Delta^2}{12}, \end{aligned}$$

där vi använt variansuttrycket i (2.6). Slutligen bestämmer vi väntevärdet för hela uttrycket (4.1):

$$\begin{aligned} E\left[\frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right] &= \frac{E[(x - \bar{x})^2]}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{12\Delta^2 (n-1)^2}{12\Delta^2 n (n^2 - 1)} = \\ \frac{(n-1)^2}{n(n-1)(n+1)} &= \frac{n-1}{n(n+1)} < \frac{1}{n}. \end{aligned}$$

Med denna approximation för den komplicerade termen i uttrycket för prediktionsintervall får man följande förenklade uttryck:

$$y^*(x) = \hat{\beta}_0 + \hat{\beta}_1 x \pm s \cdot t \sqrt{\frac{2}{n} + 1},$$

där man endast behöver känna till den skattade standardavvikelsen s och antalet referensprover n . Konstanten t tas från t -fördelningen där antalet frihetsgrader i detta fall blir $n - 2$. Ofta görs flera mätningar på varje nivå, varvid ovanstående approximation står sig men antalet frihetsgrader i t -fördelningen istället blir antalet nivåer minus två, $\frac{n}{r} - 2$, där r är antalet replikat.

Om ett stort antal referensmätningar gjorts kommer förstas termen $\frac{2}{n}$ att få försumbart inflytande och det visar sig att det kritiska antalet kalibreringsavläsningar blir lika med $n = 12$. Det korrekta konfidensintervallet blir i ett sådant fall

$$\pm s \cdot t \sqrt{\frac{2}{n} + 1} = \pm s \cdot t \sqrt{\frac{2}{12} + 1} = \pm 1.08 \cdot s \cdot t.$$

Ett approximativt konfidensintervall skulle bli

$$\pm s \cdot t \sqrt{1} = \pm 1 \cdot s \cdot t$$

och motsvarande konfidensgrad kan beräknas som den som motsvarar faktorn $k = 1.96 \frac{1}{1.08} = 1.81$ i normalfördelningen, vilket ger täckningsfaktorn 93%.

Tumregel 8

Om kalibrering vid linjär regression utförts vid jämnt fördelade nivåer, med totalt n kalibreringsmätningar så är mätosäkerheten vid följande mätningar

$$s \cdot t_{0.025} \sqrt{\frac{2}{n} + 1},$$

där s är standardavvikelsen skattad vid kalibreringen och $t_{0.025}$ är tagen ur t -fördelningstabell. Om $n \geq 12$ kan mätosäkerheten istället beräknas med

$$s \cdot t_{0.025}$$

4.1.2 Shuntkalibrering

I mätsystem där den mätta storheten i någon del av mätkedjan storhet-instrument-mätvärde representeras av en elektrisk resistans så kan man använda en referensresistans för **shuntkalibrering**. Detta innebär allmänt att man i samband med kalibreringen av instrumentet ersätter insignalen med en eller flera kända referensresistanser, läser av och noterar instrumentets utsignal. Vid senare användning av instrumentet omkalibreras det genom att med hjälp av referensresistanserna jämföra utsignalen med de noterade värdena från kalibreringstillfället. Denna omkalibrering kan göras antingen endast som kontroll av att instrumentet inte är skadat eller som ett hjälpmedel för att justera instrumentet för den specifika mätsituationen. Normalt är en av referensresistanserna en kortslutning, dvs. en nollavläsning och ytterligare en resistans är tillräcklig för att kalibrera instrumentet mot alla linjära förändringar som kan skilja kalibreringstillfället mot användningen: En enkel offset och/eller en proportionell förstärkning av signalen kan alltså kontrolleras vid omkalibreringen. Flera shuntresistanser används t. ex. för tillämpning på olika mätområden eller för att kontrollera linjäritet.

Anledningar till en omkalibrering med hjälp av shuntning kan vara att man vid mätsituationen inte har samma uppsättning av instrument i mätkedjan, såsom förstärkare, ledningar, filter eller displayutrustning. Andra anledningar kan vara okontrollerad offset i mätsituationen, miljönverkan eller osäkerhet i instrumentets inställning. Justeringen av ett instrument med hjälp av en shuntresistans går då oftast till på följande sätt: Instrumentet kopplas in i aktuellt mätsystem, en nollsignal simuleras och instrumentet nollställs. Därefter kopplas referensresistansen in och mätsystemets förstärk-

ning justeras till det vid ursprungskalibreringen noterade värdet. Notera att endast linjära förändringar i mätsystemet kalibreras med denna metod.

Vid varje justering i samband med shuntkalibrering införs ett mätfel som är systematiskt till nästa justeringstillfälle. Ett fel uppstår på grund av att man vid varje shuntavläsning får fel i enlighet med den vanliga linjära modellen (2.26), ε . Då proceduren med nollställning och förstärkningsjustering innebär två sådana avläsningar så kommer mätosäkerheten hos utsignalen att få ett tillskott på $2 \cdot u_\varepsilon^2$, vilket är systematiskt. Ett annat fel kan uppstå genom den att inställningen av förstärkningen görs med sämre noggrannhet än avläsningen. Ytterligare fel uppstår om inte samma exemplar av referensresistansen används vid kalibreringen och justeringen och om antagandena om linjära system inte är uppfyllda.

4.1.3 Mätosäkerhet

Mätosäkerhetsbidraget från ett kalibrerat mätinstrument härrör sig normalt från tre olika felkällor, nämligen 1) fel i referensen, 2) slumpmässigt fel och 3) kvarstående bias och dess standardosäkerhet avseende reproducibilitet bestäms ur summan av dessa tre felkällors varianser

$$u_{instr} = \sqrt{u_{ref}^2 + u_\varepsilon^2 + u_B^2}.$$

Notera att av dessa tre felkällor måste den första och tredje ibland betraktas som systematiska, medan den andra är slumpmässig. I ett globalt perspektiv måste hela felet betraktas som slumpmässigt och summeringen ske enligt formeln, men vid jämförande mätningar lokalt kan ibland de systematiska elimineras.

Fel i referensen

Om den referens som används vid kalibreringen representerar sanningen så kommer $\sigma_{ref}^2 = 0$. Detta kan t. ex. inträffa i kvalitetsstyrningfall då endast variation inom en tillverkningsorganisation är intressant och man definierat sanningen med en vald referens.

Om referensen i sin tur är bestämd genom något spårbart kalibreringsförfarande så kommer den att vara bestämd med en osäkerhetsangivelse, som normalt måste antas domineras av dess eventuella bias. Om specifikationen uttrycklingen är angiven i enlighet med etablerad standardmetodik (GUM, EAL) så ges förstås referensens osäkerhetsbidrag av dess utvidgade standardosäkerhet delat med två:

$$u_{ref} = \frac{U_{ref}}{2}$$

Om specifikationen inte är bestämd enligt denna metodik bör man vara

försiktig och anta att felet är likformigt fördelat inom intervallet:

$$u_{ref} = \frac{U_{ref}}{\sqrt{3}} = \frac{U_{ref}}{1.7}$$

Slumpmässigt fel

Det slumpmässiga felet i detta sammanhang är det som i kalibreringsprotokoll ofta kallas repeterbarhet och representerar det mätfel som ändras vid upprepade mätningar. Om man kan anta att detta felbidrag är lika stort vid efterföljande mätningar som vid kalibreringen så kan dess osäkerhetsbidrag bestämmas direkt ur kalibreringen genom att beräkna standardavvikelsen från m upprepade mätningar av referensen:

$$u_{\varepsilon} = \sqrt{\frac{1}{m-1} \sum_{j=1}^m (X_j - \bar{X})^2}$$

Antalet frihetsgrader i denna skattning är lika med $m - 1$ och skall beaktas vid den slutliga mätosäkerhetsangivelsen. Om antalet mätningar är få kan man förbättra skattningen (och få fler frihetsgrader) genom att sammanväga resultat från olika nivåer eller från olika kalibreringstillfällen enligt formeln för sammanvägd standardavvikelse (2.25).

Kvarstående bias

Efter kalibrering kan man välja att justera i enlighet med den uppmätta avvikelsen från referensen. Detta kan ske antingen genom att instrumentet justeras eller, om så inte är möjligt, genom att framtida mätvärden justeras. Justeringen görs med den bias som erhållits vid kalibreringen. Detta är förstas en skattning av sann bias och man får ett osäkerhetsbidrag för kvarstående bias lika med osäkerheten i det medelvärde som mätts upp:

$$u_B^2 = \frac{u_{\varepsilon}^2}{m},$$

där m är antalet mätvärden som medelvärdesbildats och u_{ε} är den skattade standardosäkerheten för slumpfelet. Antalet frihetsgrader i denna skattning är lika med antalet frihetsgrader i slumpvariationsskattningen, dvs. $m - 1$ om denna är skattad ur endast den aktuella mätningen.

Om man inte justerar med kalibreringsresultatet måste bias-variationen skattas på något sätt. Vi anger här två huvudmetoder för detta: a) skattning genom observationsmetoden, eller b) sortering. En tredje möjlighet är att man kan bortse från bias om den är försumbar. Kriteriet för detta ges i tumregel 6 på sidan 51.

Skattning enligt observationsmetoden bygger på att man har tillgång till upprepade observationer på avvikelser från referensen. Sådana observationer kan vara grundade på resultat från ett antal olika instrument, antingen i

den egna organisationen eller genom mätjämförelser med andra laboratorier. De kan också vara grundade på upprepade kalibreringar på ett instrument om variationen i bias är tidsberoende. Vid kalibrering med linjär regression kommer avvikelser från den rätta linjen att vara en typ av kvarstående bias. Dess inverkan på osäkerheten kommer emellertid att ingå i skattningen av variationen kring linjen vid kalibreringen, se exemplet om kraftkalibrering i avsnitt 6.3 på sid. 123. Allmänt gäller här, precis som alltid vid användning av observationsmetoden, att det urval som används måste motsvara ett slumpmässigt urval ur den population som mätosäkerheten skall avse. Skattning av u_B^2 genomförs genom variansanalys, se avsnitt 3.5.4.

Skattning genom sortering kallar vi det förfarandet att man bestämmer gränser för acceptabel avvikelse från referensen och att instrumentet justeras eller kasseras om det hamnar utanför gränserna. Ett sådant förfarande är ofta praktiskt i fall där kraven på mätosäkerheten inte är större än att man kan kosta på sig en extra osäkerhet för att underlätta rutiner. Om gränserna är satta till $\pm B$ så antar man lämpligen en likformig fördelning för de sorterade instrumenten:

$$u_B = \frac{B}{\sqrt{3}} = \frac{B}{1.7}$$

4.2 Sammanfattning

Ett mätsystems relation till sanningen bestäms genom **kalibrering**. Sanningen representeras av en referens som i globalt perspektiv kan vara härledd till de grundläggande SI-enheternas definitioner och i lokala perspektiv kan representeras av en lokal normal. Varje mätinstrument bör vara spårbart kalibrerat mot den referens som är intressant för mätändamålet, vilket betyder att en spårbar kedja av jämförelser skall vara genomförd med tillhörande mätosäkerhetsangivelse i varje steg. I varje steg har man därmed tillgång till 1) osäkerhet i referensen och 2) ett antal kalibreringsmätningar mot denna referens. En kalibrering kan bestå i jämförelse vid en enda nivå eller jämförelser inom ett bestämt mätområde.

Mätosäkerhetsbidraget från ett kalibrerat mätinstrument härrör sig från tre olika felkällor, nämligen 1) fel i referensen, 2a) slumpmässigt fel och 2b) kvarstående bias. Dess standardosäkerhet bestäms ur summan av dessa tre felkällors varianser

$$u_{instr} = \sqrt{u_{ref}^2 + u_{\varepsilon}^2 + u_B^2}.$$

Av dessa tre felkällor måste den första och tredje ibland betraktas som systematiska, medan den andra alltid är slumpmässig. I ett globalt perspektiv betraktas hela felet som slumpmässigt och summeringen sker enligt formeln, men vid jämförande mätningar lokalt kan de två systematiska felen ibland elimineras.

Fel i referensen

Om den referens som används vid kalibreringen representerar sanningen så kommer $\sigma_{ref}^2 = 0$. Detta kan t. ex. inträffa i kvalitetsstyrningfall då endast variation inom en tillverkningsorganisation är intressant och man definierat sanningen med en vald referens.

Om referensen i sin tur är bestämd genom något spårbart kalibreringsförfarande så kommer den att vara bestämd med en osäkerhetsangivelse. Om specifikationen uttryckningen är angiven i enlighet med etablerad standardmetodik (GUM, EAL) så ges referensens osäkerhetsbidrag av dess utvidgade standardosäkerhet delat med två:

$$u_{ref} = \frac{U_{ref}}{2}$$

Om specifikationen inte är bestämd enligt denna metodik antar man att felet är likformigt fördelat inom intervallet:

$$u_{ref} = \frac{U_{ref}}{\sqrt{3}}$$

Slumpmässigt fel

Det slumpmässiga felet är det som i kalibreringsprotokoll ofta kallas repeterbarhet och representerar det mätfel som ändras vid upprepade mätningar. Om man kan anta att detta felbidrag är lika stort vid efterföljande mätningar som vid kalibreringen så kan dess osäkerhetsbidrag bestämmas direkt ur kalibreringen genom att beräkna standardavvikelsen från m upprepade mätningar av referensen:

$$u_\varepsilon = \sqrt{\frac{1}{m-1} \sum_{j=1}^m (X_j - \bar{X})^2}$$

Antalet frihetsgrader i denna skattning är lika med $m - 1$ och skall beaktas vid den slutliga mätosäkerhetsangivelsen. Om antalet mätningar är få kan man förbättra skattningen (och få fler frihetsgrader) genom att sammanväga resultat från olika nivåer eller från olika kalibreringstillfällen enligt formeln för sammanvägd standardavvikelse (2.25).

Kvarstående bias

Efter kalibrering kan man välja att justera i enlighet med den uppmätta avvikelsen från referensen. Detta kan ske antingen genom att instrumentet justeras eller, om så inte är möjligt, genom att framtida mätvärden justeras. Justeringen görs med den bias som erhållits vid kalibreringen. Detta är förstas en skattning av sann bias och man får ett osäkerhetsbidrag för kvarstående bias lika med osäkerheten i det medelvärde som mätts upp:

$$u_B^2 = \frac{u_\varepsilon^2}{m},$$

där m är antalet mätvärden som medelvärdesbildats och u_ε är den skattade standardosäkerheten för slumpfelet. Antalet frihetsgrader i denna skattning är lika med antalet frihetsgrader i slumpvariations-skattningen, dvs. $m - 1$ om denna är skattad ur endast den aktuella mätningen.

Om man inte justerar med kalibreringsresultatet måste bias-variationen skattas:

1. skattning enligt observationsmetoden kan göras om man har tillgång till upprepade observationer på avvikelser från referensen. Sådana observationer kan vara grundade på resultat från ett antal olika instrument, antingen i den egna organisationen eller genom mätjämförelser med andra laboratorier. De kan också vara grundade på upprepade kalibreringar på ett instrument om variationen i bias är tidsberoende. Skattning av u_B^2 genomförs genom enkel standardavvikelse över ett antal instrument, eller genom variansanalys om andra variationskällor måste skiljas ut, se avsnitt 3.5.4.

2. skattning genom sortering görs genom att man bestämmer gränser för acceptabel avvikelse från referensen och att instrumentet justeras eller kasseras om det hamnar utanför gränserna. Ett sådant förfarande är ofta praktiskt i fall där kraven på mätosäkerheten inte är större än att man kan kosta på sig en extra osäkerhet för att underlätta rutiner. Om gränserna är satta till $\pm B$ så antar man lämpligen en likformig fördelning för de sorterade instrumenten (se tumregel 9, sid. 114:

$$u_B = \frac{B}{\sqrt{3}}$$

3. bias-variationen sätts till noll om den är försumbar. Kriteriet för detta ges i tumregel 6 på sidan 51.

Kalibrering för ett mätområde

Vid kalibrering för mer än en mätnivå kan resultatet sammanfattas i en kalibreringskurva som används i efterföljande mätningar. Vid anpassning till en rät linje används lämpligen modellen

$$x = \beta'_0 + \beta'_1 y + \varepsilon' ,$$

där y är de uppmätta resultaten, x är de inställda nivåerna på referensen, β'_0 och β'_1 är parametrarna i modellen och ε' är ett felbidrag. Detta förfarande innebär i praktiken att man justerar för en övergripande bias genom β'_0 medan nivårelaterade avvikelser från den räta linjen inkluderas i ε' . Osäkerheten i framtida mätningar avseende olinjäriteten kommer därmed att ingå i skattningen av variationen kring linjen vid kalibreringen. Mätosäkerhetsbidraget från ett instrument som kalibrerats i ett intervall ges av tumregel 8 på sid. 93: Om kalibrering vid linjär regression utförts vid jämnt fördelade nivåer, med totalt n kalibreringsmätningar så är mätosäkerheten vid följande mätningar

$$s \cdot t_{0.025} \sqrt{\frac{2}{n} + 1} ,$$

där s är standardavvikelsen skattad vid kalibreringen och $t_{0.025}$ är tagen ur t -fördelningstabell. Om $n \geq 12$ kan mätosäkerheten istället beräknas med

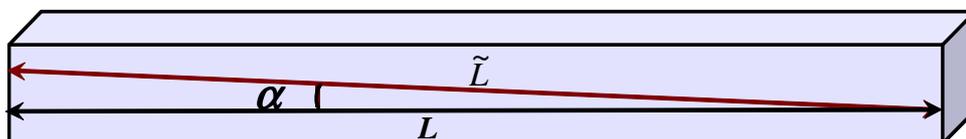
$$s \cdot t_{0.025}$$

Kapitel 5

Speciella problem vid mätosäkerhetsberäkning

5.1 Felets väntevärde är skilt från noll

Mätosäkerhetssummeringen bygger på att endast varianserna summeras och att det resulterande konfidensintervallet är symmetriskt kring mätvärdet, dvs. vi antar att väntevärdet för mätfelet är lika med noll. I vissa fall är detta inte uppfyllt, vilket kan kräva ytterligare utredning. Ett exempel är det vinkelfel som uppstår vid längdmätning av en rektangulär detalj. Se figur nedan



där α är det vinkelfel som operatören introducerar. Mätningen av längden L

blir därmed alltid längre än eller lika med den korrekta, oavsett om vinkeln α är positiv eller negativ. Matematiskt kan detta uttryckas på följande sätt

$$L = \tilde{L} \cos(\alpha) ,$$

där \tilde{L} är den uppmätta längden med vinkelfelet α . För att få ett osäkerhetsmått som är användbart i den allmänna mätosäkerhetsanalysen bör man här justera sin längdmätning med väntevärdet av felet och sedan ansätta standardosäkerheten för vinkelfelet som dess standardavvikelse. Det visar sig dock inte vara så enkelt att bestämma osäkerheten. Om vi använder Gauss approximationsformel i detta fall så skulle vinkelfelet ge känslighetskoefficienten

$$c_\alpha = \frac{\partial L}{\partial \alpha} = \tilde{L}_0 \sin(\alpha_0) ,$$

där \tilde{L}_0 och α_0 står för respektive väntevärde. Men väntevärdet för vinkelfelet är rimligen lika med noll och därmed blir $c_\alpha = 0$. Vinkelfelet bidrar inte alls till mätosäkerheten! Detta visar hur Gauss approximationsmetodik inte alltid fungerar och i detta fall måste man göra en mer detaljerad undersökning. Vi formulerar då först det relativa felet:

$$\frac{e_L}{\tilde{L}} = \frac{\tilde{L} - L}{\tilde{L}} = 1 - \cos(\alpha),$$

och gör en approximativ undersökning med hjälp av simulering i ett datorprogram. Anta att vinkelfelet är normalfördelat med väntevärde noll och variansen σ_k^2 ,

$$\alpha \sim N(0, \sigma_k^2).$$

Generera 100 000 normalfördelade slumpstal $\{\alpha_i, i = 1, 2, \dots, 100\,000\}$ som representerar vinkelfel. Beräkna för varje sådant tal det relativa längdmätningensfelet

$$\left\{ \frac{e_L}{\tilde{L}} \right\}_i = 1 - \cos(\alpha_i).$$

Beräkna medelvärde och standardavvikelse för dessa 100 000 relativa fel. Gör motsvarande simulering för ett antal olika varianser σ_k^2 . Anpassa funktioner till resultaten. Det visar sig då att man för $\sigma_k < 10^\circ$ får en utmärkt approximation genom att sätta

$$E \left[\frac{e_L}{\tilde{L}} \right] = \frac{1}{\cos(\sigma_k)} - 1, \quad \sqrt{\text{Var} \left[\frac{e_L}{\tilde{L}} \right]} = 1.4 \left(\frac{1}{\cos(\sigma_k)} - 1 \right).$$

Man bör alltså justera sin längdmätning med hjälp av $E \left[\frac{e_L}{\tilde{L}} \right]$:

$$\begin{aligned} L &= \tilde{L} - e_L = \tilde{L} - \tilde{L} \cdot \frac{e_L}{\tilde{L}} \approx \tilde{L} \left(1 - E \left[\frac{e_L}{\tilde{L}} \right] \right) = \\ &= \tilde{L} \left[1 - \left(\frac{1}{\cos(\sigma_k)} - 1 \right) \right] = \tilde{L} \left(2 - \frac{1}{\cos(\sigma_k)} \right) \end{aligned}$$

och ansätta standardosäkerheten

$$u_\alpha = 1.4 \cdot \tilde{L} \cdot \left(\frac{1}{\cos(\sigma_k)} - 1 \right)$$

Allmänt gäller alltså att man i fall då felets väntevärde inte är lika med noll justerar med felets väntevärde och sedan använder dess varians till skattning av standardosäkerheten. Naturligtvis kan man med stöd av tumregel 6 på sidan 51 försumma avvikelser från väntevärdet om den är tillräckligt liten.

5.2 Kvalitetsstyrning

Kvalitetsstyrning i produktionsprocesser är ett område där statistiska verktyg kommit till användning och där mätningar intar en central plats. T. ex. Shewharts styrdiagram [11] bygger på att man genom kontinuerliga mätningar bevakar en produktionsprocess variation. Man kan identifiera två olika angreppssätt i kvalitetsstyrningen:

- Styrning mot målvärde
- Styrning mot tolerans

Styrning mot målvärde innebär i princip att man alltid siktar på att uppnå nominellt värde för en produkt. Denna strävan i kombination med att variationen i produktionen begränsas ger en liten felsannolikhet, vilket kan uttryckas som att produktionen är under statistisk kontroll. Variationen i produktionen mäts ofta med måttet duglighet (eng: capability), som jämför variabiliteten i mätvärden med toleransintervallet,

$$C_p = \frac{|U - L|}{6\sigma}$$

där U och L representerar övre respektive undre toleransgräns och σ är standardavvikelsen för mätvärden, dvs. en kombination av produkternas variation och mätsystemets.

Ett typiskt verktyg för att bevaka att produktionen är under statistisk kontroll är ett *styrdiagram* av typ medelvärdes- och variationsviddsdiagram (average and range control charts) där mätningar registreras på ett systematiskt sätt. Mätosäkerheten i form av reproducerbarhet är här ofta irrelevant eftersom lokala referenser är vanliga. Regelbundna kalibreringar mot sådana lokala referenser ser till att styrningen riktas mot ett korrekt värde. Repeterbarheten har här betydelse endast som en komponent i totalvariationen. I fall där produktionens egen variation dominerar kommer förbättringar i mätsystemets repeterbarhet därför att ha en marginell betydelse. Instruktionen för mätsystemsanalys i kvalitetsstandarden QS9000 (MSA) är typiskt avsedd att tillämpas på denna kvalitetsstyrningsprincip, vilken behandlas i nästa avsnitt, 5.2.1.

Styrning mot tolerans innebär att produkterna skall ligga inom bestämda toleranser och kontrollmätning sker på varje produkt med resultatet: godkänd eller icke godkänd. Vid sådan kvalitetsstyrning får mätosäkerheten en avgörande betydelse; mätfel kan orsaka att vissa produkter kasseras i onödan och andra godkänns trots felaktiga egenskaper. Här är ibland reproducerbarheten avgörande, nämligen om toleranserna avser något som skall vara jämförbart globalt, men ibland är endast interna jämförelser relevanta och då gäller att mätosäkerheten avser repeterbarhet. Mätosäkerhetens inverkan

på producent- resp. konsumentrisk vid styrning mot tolerans behandlas i avsnitt 5.2.2.

Naturligtvis styrs ofta kvalitet genom någon blandning av de här två styrningsprinciperna och mätosäkerhetens betydelse i varje specifikt fall måste avgöras med hänsyn till de speciella förhållandena som gäller.

5.2.1 Mät-systemsanalys (MSA)

I kvalitetsstandarden QS9000 finns en instruktion för att genomföra en mät-systemsanalys (MSA) för mätsystem avsedda att bevaka produktionsprocesser. Ett av målen för en sådan analys är att undersöka om mätsystemet är tillräckligt noggrant i jämförelse med produkternas variation eller i jämförelse med aktuella toleranser. Ett annat mål är att finna de avgörande mät-felskällorna och åtgärda dem. Instruktionen utgår från en något föråldrad metod, byggd på variationsvidder, men anger också variansanalysen (ANOVA) som ett alternativ. Analysen genomförs då i enlighet med exemplet i avsnittet om variansanalys med två faktorer, 3.5.4. MSA-manualen använder följande beteckningar för de olika variationskomponenterna. Våra beteckningar för motsvarande storheter anges i tredje kolumnen., där för överskådlighetens skull de numeriska värdena i nämnaren är tagna ur det konkreta exemplet i avsnitt 3.5.4.

detalj (produkt)	PV	$\sqrt{\frac{MS_D - MS_E}{3 \cdot 2}}$
operatör	AV	$\sqrt{\widehat{Var}[B]} = \sqrt{\frac{MS_O - MS_E}{10 \cdot 2}}$
samverkan	IV	0
slumpfel (instrument)	EV	\sqrt{MSE}

Om samverkans-effekten visar sig vara signifikant så ges MSA-storheterna istället av

slumpfel (instrument)	EV	$\sqrt{MKv(\text{Inom})}$
detalj (produkt)	PV	$\sqrt{\frac{MKv(\text{Kolumner}) - MKv(\text{Interaktion})}{3 \cdot 2}}$
operatör	AV	$\sqrt{\frac{MKv(\text{Sampel}) - MKv(\text{Interaktion})}{10 \cdot 2}}$
samverkan	IV	$\sqrt{\frac{MKv(\text{Interaktion}) - MKv(\text{Inom})}{2}}$

där Excel-storheterna ur exemplet använts i högerkolumnen.

Dessa komponenter skattas ur variansanalystabellen och sammanvägs till en mätosäkerhet, kallad GRR (Gage Repeatability & Reproduceability):

$$GRR = \sqrt{EV^2 + AV^2 + IV^2} .$$

Problem 5.2.1 Beräkna GRR för exemplet givet i avsnitt 3.5.4 med hjälp av de båda metoderna angivna här, dvs. bortse i det ena fallet från att sam-

verkan effekten är signifikant. Jämför resulterande mätosäkerhetskattningar, samt studera hur de olika komponenterna i GRR påverkas.

MSA-manualen ger också en metod för jämförelse av mätosäkerheten med produktionens totala variation. Det mått som då används är

$$\%GRR^{(1)} = 100 \cdot \frac{GRR}{TV}, \text{ eller } \%GRR^{(2)} = 100 \cdot \frac{GRR}{|U - L|/6},$$

där i den första varianten TV är den totala variationen som kan skattas ur variansanalysen som $TV = \sqrt{GRR^2 + PV^2}$. I den andra formeln jämförs mätosäkerheten istället med toleransvidden $|U - L|$, skalad med konstanten 6.

MSA-manualens tumregel för godkännande av ett mätsystem är: godkänt om $\%GRR < 10\%$ och underkänt om $\%GRR > 30\%$, medan resultat däremellan normalt kräver speciella överväganden. Dessa gränser har utsatts för mycket kritik, då de ofta är svåra att uppnå, men då mätsystemet ändå anses duga för att bevaka produktionsvariationen. Detta kan ha flera orsaker, t. ex.: a) att man inte tar hänsyn till att man i kvalitetskontrollen medelvärdesbildar flera mätningar, vilket naturligtvis ger en mycket bättre $\%GRR$, b) att endast repeterbarheten är relevant vid bevakningen, varvid GRR är ett överdrivet mått och EV borde användas istället, eller c) att toleransintervallet $|U - L|$ i jämförelsen inte är bestämt utifrån samma filosofi som den QS9000 representerar.

5.2.2 Toleranser

Om mätsystemet skall användas för att kontrollera uppfyllandet av en tolerans ställs oftast högre krav på dess noggrannhet. Mätosäkerheten vid jämförelse med toleransgränser orsakar dels att korrekta produkter förkastas, dels att felaktiga produkter godkänns; mätosäkerheten ger en "producentrisk" resp. en "konsumentrisk". Vi skall i detta avsnitt undersöka hur dessa risker kan beräknas utifrån den beräknade mätosäkerhetens relation till toleransvidden.

Anta att produktionens resultat är normalfördelat med variansen $\left(\frac{U-L}{c}\right)^2$, där U och L är övre respektive undre toleransgränsen och c är en konstant som relaterar produktionsvariansen till toleransen. I MSA-manualen är $c = 6$. Anta vidare att produktionens väntevärde är μ .

$$X \sim N(\mu, \tau^2), \quad \tau^2 = \left(\frac{U - L}{c}\right)^2$$

Mätsystemet som används för att avgöra produkternas godkännande eller förkastande antas ha en mätosäkerhet som är en viss andel av produktionens variation. Varje mätning av x har då fördelningen

$$Y \sim N(x, \sigma^2)$$

där standardavvikelsen σ är proportionell mot produktionens standardavvikelse

$$\sigma = \gamma \cdot \tau$$

I MSA-manualen identifieras σ med begreppet GRR och proportionalitetskonstanten 100γ med %GRR.

Vi undersöker först sannolikheten för felaktigt godkännande av en produkt pga. osäkerheten i mätsystemet:

$$\begin{aligned} P\{L \leq Y \leq U \mid x < L \vee x > U\} &= \\ P\{L \leq Y \leq U \mid x < L\} + P\{L \leq Y \leq U \mid x > U\} &= \end{aligned}$$

$$\int_{-\infty}^L [F_Y(U) - F_Y(L)] f_X(x) dx + \int_U^{\infty} [F_Y(U) - F_Y(L)] f_X(x) dx \quad (5.1)$$

där $F(\cdot)$ och $f(\cdot)$ är fördelnings- resp. frekvensfunktioner för respektive variabel, givna av normalfördelningarna ovan.

Vi bestämmer härnäst sannolikheten för felaktigt underkännande av korrekta produkter:

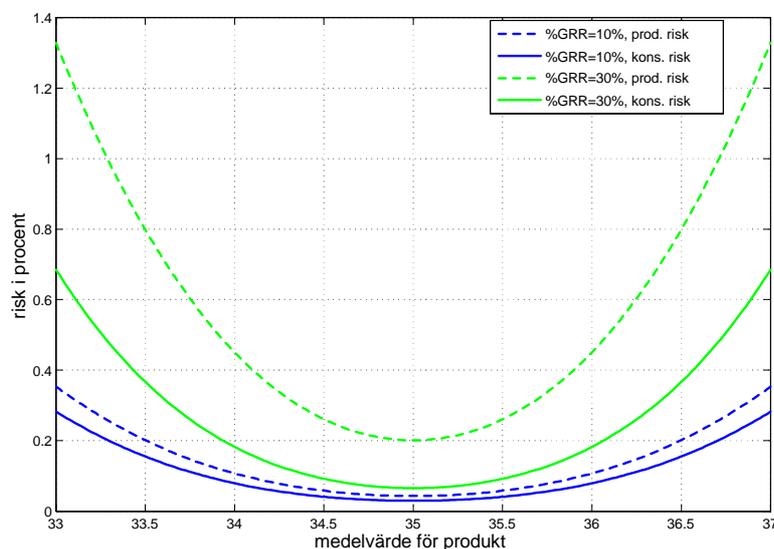
$$\begin{aligned} P\{Y < L \vee Y > U \mid L \leq x \leq U\} &= \\ P\{Y < L \mid L \leq x \leq U\} + P\{Y > U \mid L \leq x \leq U\} &= \end{aligned}$$

$$\int_L^U F_Y(L) f_X(x) dx + \int_L^U [1 - F_Y(U)] f_X(x) dx \quad (5.2)$$

Dessa båda sannolikheter kan beräknas numeriskt och vi visar i ett exempel hur resultatet kan se ut.

Exempel 5.2.1 Vi fastställer toleransintervallet till $[L, U] = [30, 40]$ och sätter $c = 6$ enligt MSA-manualen. Vidare sätter vi %GRR=10%, vilket är det som betraktas som helt godkänt enligt MSA-manualen. Vi ansätter först att väntevärdet i produktionen ligger mitt i toleransintervallet, $\mu = \frac{U+L}{2} = 35$. Sannolikheten för underkännande av en korrekt produkt enligt (5.2) blir då 0.04% medan sannolikheten för godkänd felaktig produkt enligt (5.1) blir 0.03%. Om vi istället har ett mätsystem med %GRR=30%, vilket med tvekan är godkänt enligt MSA, så blir isället sannolikheterna 0.2% resp. 0.07%.

Exempel 5.2.2 Man kan ju också undersöka fallet då väntevärdet i produktionen avviker från det mittersta värdet. I följande figur är ovanstående exempel utvidgat till de fall då väntevärdet kan avvika från mittvärdet med maximalt två enheter.



Vi ser att sannolikheten för felaktiga beslut kan bli synnerligen stor, speciellt om man inte har väntevärdesriktigt medelvärde i produktionen.

Lösningen för konstruktören kan vara att minska toleransintervallet, dvs. använda en säkerhetsmarginal. Om denna säkerhetsmarginal är 100s % av intervallets omfång så kommer sannolikheterna att bli

$$P\{L + s(U - L) \leq Y \leq U - s(U - L) \mid x < L \vee x > U\} =$$

$$\int_{-\infty}^L [F_Y(U - s(U - L)) - F_Y(L + s(U - L))] f_X(x) dx +$$

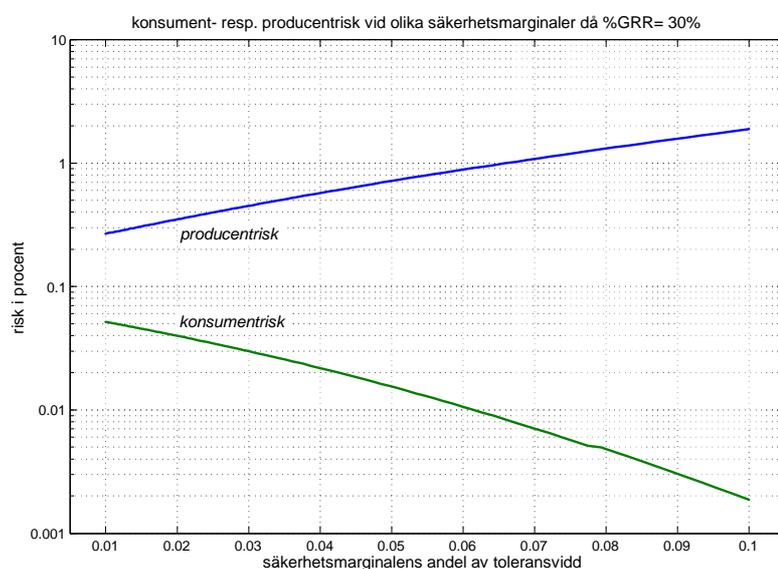
$$\int_U^{\infty} [F_Y(U - s(U - L)) - F_Y(L + s(U - L))] f_X(x) dx$$

respektive

$$P\{Y < L + s(U - L) \vee Y > U - s(U - L) \mid L \leq x \leq U\} =$$

$$\int_L^U F_Y(L + s(U - L)) f_X(x) dx + \int_L^U [1 - F_Y(U - s(U - L))] f_X(x) dx$$

Resultatet blir att sannolikheten för felaktigt godkännande (kundrisken) minskar, medan sannolikheten för felaktigt underkännande (producentrisken) ökar:



Vi ser alltså att utifrån de angivna förutsättningarna så är det enda sättet att minska såväl producent- som konsumentrisk att förbättra mätsystemet, dvs minska %GRR, och att för ett givet mätsystem så gäller att en säkerhetsmarginal på toleransintervallet kan sättas för att minska konsumentrisken och där man därmed måste avväga de båda riskerna med andra kriterier än statistiska. Vi ser också att det är mycket väsentligt att hålla produktionens väntevärde nära målvärdet. Ett sätt att genomföra en kompromiss mellan producent- och konsumentrisk utan att förbättra mätsystemet är förstås att etablera en säkerhetsmarginal i vardera riktningen och kanske ge resultat som faller inom ett intervall utan avgörande en särskild behandling.

Risken för felaktiga kassationer eller godkännanden beror alltså på flera storheters förhållande till varandra, nämligen mätosäkerheten, produktions fördelning, speciellt dess väntevärde och varians, samt toleransvidden. Eventuella säkerhetsmarginalers storlek har också stort inflytande. Det är därför svårt att ge generella rekommendationer, utan beräkningar bör genomföras i varje specifikt fall för de aktuella värdena på de kritiska storheterna. Det givna exemplet kan tjäna som vägledning för hur och när sådana beräkningar är nödvändiga att genomföra.

5.3 Korrelation

Om osäkerhetskomponenter är relaterade till varandra på något sätt så kommer den enkla summeringen av varianskomponenter enligt ovan att bli felaktig. Detta kan exempelvis inträffa om några osäkerhetskomponenter har tendens att samvariera, dvs. tendera att vara stora resp. små samtidigt. Ett

statistiskt mått på samvariation är kovarians, som presenterades i kapitel 2. Det är definierat som väntevärdet av produkten mellan två slumpvariablers avvikelse från sina respektive väntevärden:

$$Kov(x, y) = E[(x - E[x])(x - E[y])]$$

Denna storhet är ofta svårtolkad då den beror på vilken skala man använder för de två slumpvariablerna. Därför har man definierat den normaliserade storheten korrelationskoefficienten:

$$\rho[x, y] = \frac{Kov[x, y]}{\sigma_x \sigma_y}$$

Denna storhet är alltid ett tal mellan -1 och 1, där $\rho = -1$ betyder en fullständig linjär negativ samvariation, $\rho = 1$ en motsvarande positiv samvariation och $\rho = 0$ betyder att variablerna är okorrelerade, dvs. inget linjärt samband finns mellan variablerna. I vårt problem med att sammanväga olika osäkerhetskomponenter är vi intresserade av att studera summor av stokastiska variabler och då gäller:

$$Var[X + Z] = \sigma_X^2 + \sigma_Z^2 + 2 \cdot Kov(X, Z) = \sigma_X^2 + \sigma_Z^2 + 2 \cdot \sigma_X \sigma_Z \rho_{XZ}$$

Vi kan identifiera följande specialfall för korrelationerna 0, 1, resp. -1:

$$Var[X + Z] = \begin{cases} \sigma_X^2 + \sigma_Z^2, & \rho = 0 \\ \sigma_X^2 + \sigma_Z^2 + 2 \cdot \sigma_X \sigma_Z = (\sigma_X + \sigma_Z)^2 & \rho = 1 \\ (\sigma_X - \sigma_Z)^2 & \rho = -1 \end{cases}$$

Notera att vår standardmetodik beskriven ovan motsvarar fallet då korrelationskoefficienten är lika med noll. Vid full positiv korrelation får man ett uttryck som motsvarar den metod som var tradition i klassisk fysikundervisning, dvs. summering av maxfel. Den innebär att man betraktar det värsta fallet, då båda felen är störst samtidigt. Slutligen innebär en negativ korrelation att variationen för summan minskar, med totalt utsläckning om man har lika varianser och full negativ korrelation.

Exempel 5.3.1 *Osäkerhetskomponenterna temperatur, luftfuktighet och lufttryck påverkar de flesta mätningar i varierande grad. Här torde det finnas en tendens till samvariation, då högtrycksdagar, åtminstone på sommaren i Sverige, har hög temperatur och låg luftfuktighet. Beroende på hur dessa tre komponenter påverkar mätningen får man en positiv eller negativ korrelation och den kombinerade mätosäkerheten påverkas av detta.*

Exempel 5.3.2 *Vid mätning av ett stålobjekt med ett skjutmått av stål kommer temperaturen att vara en osäkerhetskomponent. Den påverkar dock*

såväl mättdon som mätobjekt, i det ena fallet positivt, i det andra negativt. Detta innebär att felen tar ut varandra vilket kan beskrivas i våra matematiska termer så att korrelationskoefficienten mellan dessa två osäkerhetskomponenter är lika med -1 . Vid mätning av en komponent av något annat material blir korrelationskoefficienten något tal mellan -1 och 0 .

Exempel 5.3.3 Vid kvalitetskontroll av produkter i en tillverkningsprocess har man ofta ett toleransintervall, som produkterna skall uppfylla. Ibland förekommer en fullständig kontroll av produktionen och produkter som faller utanför toleransen kasseras. Här kan man tänka sig en korrelation mellan produkternas egenskap och mätoperatörens mätfel: Antag att operatörerna har en god vilja och hellre friar än fäller, dvs. hellre godkänner än kasserar om produkten ligger nära gränsen. Detta skulle betyda att stora produkters mått underskattas, medan små produkters överskattas och en negativ korrelation skulle kunna detekteras.

Exempel 5.3.4 Vid upprepade mätningar kan man ibland ha en tidsberoende felkälla. Ett exempel är 50 Hz-störningar orsakade av vårt elnät pga bristfällig likriktning. Om upprepade mätningar genomförs med ett elektriskt mätinstrument under en kort tidsrymd (mindre än en period $T = 0.02$ s) så kommer störningens värde att likna det förra mätvärdet och man har en positiv korrelation.

I det sista exemplet har man korrelation mellan olika mättillfällen. Denna typ av fel är alltså konstant vid upprepningar och är därmed systematiskt enligt vår tidigare definition. Systematiska fel kan alltså modelleras med hjälp av korrelation, vilket ytterligare demonstreras i kalibreringsexemplet i avsnitt 5.3.2 nedan.

5.3.1 Behandling av korrelerade komponenter

Korrelationer ger komplikationer vid beräkning av mätsosäkerheter eftersom korrelationskoefficienten oftast är okänd och dessutom ofta är svår att uppskatta. Ofta orsakas korrelationen mellan två eller flera variabler av att de alla är funktioner av någon annan underliggande variabel. Det är därför viktigt att försöka förstå de underliggande mekanismerna och söka uppskatta de underliggande, mindre korrelerade, osäkerhetskomponenterna.

Skattning av korrelationer kan ibland ske genom fysikaliska bedömningar som i exemplen med skjutmått och 50 Hz-störningen. I andra fall, som i kvalitetskontroll-exemplet, måste statistiska försök och formel (2.9) användas.

5.3.2 Korrelation på grund av gemensam normal

Betrakta fallet då vi använder en kalibrerad vikt som har ett okänt systematiskt fel B och vår vägning tillför ett slumpmässigt fel e . Det okända

systematiska felet kan betraktas som slumpmässigt draget ur en population av viktavvikelser och vi antar att vi känner standardosäkerheten för denna population, u_B . Standardosäkerheten för det slumpmässiga felet i mätningen är u_e . Sett ur ett globalt perspektiv är då den totala mätosäkerheten, reproducerbarheten:

$$u = \sqrt{u_B^2 + u_e^2} \quad . \quad (5.3)$$

Denna osäkerhet representerar situationen att man väger med en slumpmässigt vald referensvikt. En sådan referensvikt har massan $m_j = m + B_j$, dvs. en summa av den nominella massan m och den specifika systematiska avvikelserna B_j . Vid en mätning mot referensvikten tillkommer ett slumpmässigt fel och mätning nr i ger följande resultat:

$$\widehat{m}_i = m + B_j + e_i$$

Variansen för en sådan mätning blir

$$Var(\widehat{m}_i) = Var(B) + Var(e) + 2Kov(B, e) = u_B^2 + u_e^2 ,$$

i enlighet med (5.3) eftersom kovariansen är noll då de båda slumpvariablerna B och e är oberoende. Om man nu istället använder samma referensvikt vid flera mätningar så kommer kovariansen för två mätningar att bli:

$$\begin{aligned} Kov(\widehat{m}_i, \widehat{m}_k) &= E[(\widehat{m}_i - m)(\widehat{m}_k - m)] = E[(B_j + e_i)(B_j + e_k)] = \\ &= E[B_j e_i] + E[B_j e_k] + E[e_i e_k] + E[B_j^2] = 0 + 0 + 0 + u_B^2 = u_B^2 \end{aligned}$$

dvs. om man tar summan av två dylika mätningar

$$\widehat{m}_i + \widehat{m}_k = 2m + 2B_j + e_i + e_k \quad (5.4)$$

så kommer variansen för denna summa att bli

$$\begin{aligned} Var(\widehat{m}_i + \widehat{m}_k) &= Var(\widehat{m}_i) + Var(\widehat{m}_k) + 2Kov(\widehat{m}_i, \widehat{m}_k) = \\ &= 2u_B^2 + 2u_e^2 + 2u_B^2 = 4u_B^2 + 2u_e^2. \end{aligned}$$

Samma resultat får vi om vi istället räknar direkt på mätresultatet (5.4):

$$\begin{aligned} Var(\widehat{m}_i + \widehat{m}_k) &= Var(2m + 2B_j + e_i + e_k) = Var(2B_j) + 2Var(e) = \\ &= 4Var(B) + 2Var(e) = 4u_B^2 + 2u_e^2. \end{aligned}$$

Kovariansen är alltså vid upprepade mätningar lika med den variansdel i den totala mätosäkerheten som representerar den slumpvariabel som hålls konstant, u_B^2 i (5.4). Denna kovarians kan uttryckas i form av korrelationskoefficient:

$$Korr(\widehat{m}_i, \widehat{m}_k) = \frac{Kov(\widehat{m}_i, \widehat{m}_k)}{\sqrt{Var(\widehat{m}_i) \cdot Var(\widehat{m}_k)}} = \frac{u_B^2}{u_B^2 + u_e^2}.$$

Korrelationen i upprepade mätningar är alltså kvoten mellan variansen för det systematiska felet och den totala variansen.

5.4 Mätosäkerhet efter sortering

5.4.1 Problemställning

Vid kalibrering av vikter specificeras ibland ett toleransintervall som används vid senare användning av vikterna för att uppskatta den totala mätosäkerheten. Vikter som inte uppfyller toleransintervallet kasseras. Det fel som toleransintervallet representerar är den systematiska avvikelse som varje vikt har gentemot en normal. Vid kalibreringen kommer bestämningen av viktens fel att själv ha en slumpmässig variation på grund av t. ex. vågen, miljön och handhavandet och följaktligen kommer vissa vikter som skulle godkännas att underkännas och vice versa. Det specificerade toleransintervallet kan alltså inte ses som en korrekt beskrivning av den slutgiltiga variationen hos de godkända vikternas systematiska fel. Hur skall man i efterföljande mätningar modellera den mätosäkerhet som toleransintervallet representerar?

5.4.2 Modell

Vi vill beräkna fördelningen för det systematiska felet, givet att man godkännt vikten. Godkännande innebär att man vid vägning funnit att den ligger inom toleransgränserna $\pm a$. Vi antar att mätfelets standardavvikelse är proportionellt mot halva toleransvidden och är normalfördelat enligt

$$\varepsilon \sim N(0, u_\varepsilon^2),$$

där $u_\varepsilon^2 = v^2 a^2$ och proportionalitetskonstanten alltså är satt till v . Det systematiska felet är likformigt fördelat med gränser som något överstiger toleransgränserna, $B \sim U(-ca, ca)$, $c > 1$. Alltså, beräkna sannolikheten för att det systematiska felet är lika med ett visst värde givet att detta värde har godkännas vid mätning:

$$P(B = x | -a < x + \varepsilon < a) = C \int_{-a-x}^{a-x} f(\varepsilon) \frac{1}{2ca} d\varepsilon = \frac{C}{2ca} \int_{-a-x}^{a-x} f(\varepsilon) d\varepsilon =$$

$$\frac{C}{2ca} \left\{ \Phi\left(\frac{a-x}{u_\varepsilon}\right) - \Phi\left(\frac{-a-x}{u_\varepsilon}\right) \right\} = \frac{C}{2ca} \left\{ \Phi\left(\frac{1-x/a}{v}\right) - \Phi\left(\frac{-1-x/a}{v}\right) \right\},$$

där $\Phi(\cdot)$ står för fördelningsfunktionen för den standardiserade normalfördelningen $N(0, 1)$ och konstanten C skall väljas så att frekvensfunktionen blir lika med ett vid integrering över alla möjliga värden x :

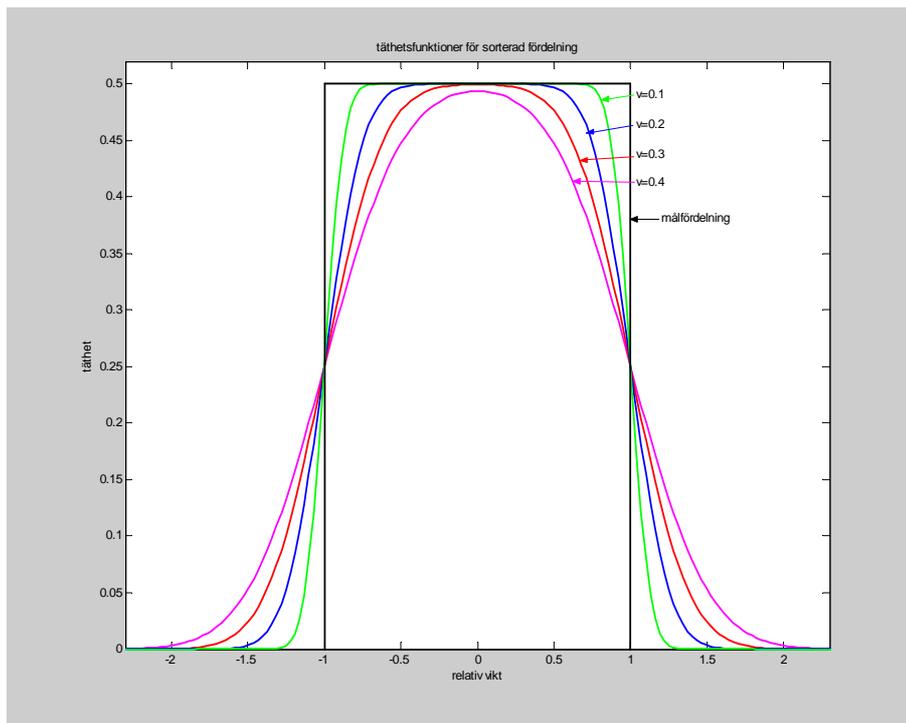
$$\int_{-ca}^{ca} \frac{C}{2ca} \left\{ \Phi\left(\frac{1-x/a}{v}\right) - \Phi\left(\frac{-1-x/a}{v}\right) \right\} dx = 1$$

Denna ekvation kan lösas numeriskt vilket visar att $C \approx ac$ om $c > 1 + 2u_\varepsilon$, dvs. om man kan anta att fördelningen före sortering har bidrag

klart utanför toleransgränserna. Frekvensfördelningen för det resulterande biasfelet blir då approximativt

$$f_{B|gokdänd}(x) = \frac{1}{2} \left\{ \Phi \left(\frac{1-x/a}{v} \right) - \Phi \left(\frac{-1-x/a}{v} \right) \right\},$$

vilken illustreras i figur 5.1 för olika relativa mätosäkerheter v .



Figur 5.1: Illustration av teoretiska frekvensfunktioner för sorterade vikter vid $a = 1$ och fyra olika mätnoggrannheter. Den ideala fördelningen är plotad som en rektangel.

Figuren visar att biasfördelningen efter sortering avviker mer och mer från den ursprungliga likformiga fördelningen vid ökande mätosäkerhet. Vid användning av de sorterade vikterna får man en standardosäkerhet enligt den approximativa formeln

$$u_B = a \left(\frac{1}{\sqrt{3}} + 0.8 \cdot v^2 \right), \text{ för } v \leq 0.4$$

Notera att om mätosäkerheten vid sorteringen är försumbar så blir den sorterade mätosäkerheten som för en likformig fördelning med vidden $2a$. Det är naturligtvis här intressant att kontrollera vilken mätosäkerhet som krävs vid

kalibreringen för att man skall kunna försumma den andra termen i parentesens ovan. Det visar sig att en lämplig tumregel är att kräva att $u_\varepsilon < 0.2a$. Detta motiveras av följande räkning: Vid försummandet approximerar man biasosäkerheten med

$$u_B \approx \frac{a}{\sqrt{3}}.$$

Ett 95%-igt konfidensintervall fås då genom $\pm 2u_B = \pm \frac{2a}{\sqrt{3}} = \pm 1.15a$. Hur mycket täcker detta intervall av den sanna fördelningen med standardavvikelse $a \left(\frac{1}{\sqrt{3}} + 0.82 \cdot v^2 \right) = a \left(\frac{1}{\sqrt{3}} + 0.82 \cdot 0.2^2 \right) = 0.61a$? Det är detsamma som att beräkna sannolikheten för att en standardiserad normalfördelad variabel Z ligger i intervallet

$$P \left(\frac{-1.15a}{0.61a} < Z < \frac{1.15a}{0.61a} \right) = P(-1.89 < Z < 1.89) = 0.94.$$

5.4.3 Sammanfattning

Vi kan alltså konstatera att med motivering av vår allmänna approximationsregel från avsnitt 3.2 att mätosäkerheten efter sortering kan sättas lika med en likformig fördelning bestämd av gränserna för godkännande om mätosäkerheten vid kontrollen är bättre än 20% av intervallens halva bredd.

Tumregel 9. Sortering

Mätosäkerheten för bias efter sortering kan skattas med

$$u_B = \frac{a}{\sqrt{3}}$$

med sorteringsintervallet $\pm a$ om mätosäkerheten vid sorteringen är högst 20% av halva intervallbredden a .

Kapitel 6

Exempel på mätosäkerhetsutredningar

6.1 Tippningsprovning av skrivbordsstol

Enligt en standardiserad metod för bestämning av stabiliteten hos en skrivbordsstol gör man en provning på följande sätt: Stolens främre hjul ställs mot ett stöd i golvet, en vikt M placeras på dynan, en lina fästs i dynans tyg och man drar manuellt i linan horisontellt. Dragkraften mäts med en dynamometer som är fäst i linan och stolen godkänns om tippning inte sker vid kraften 100 N. Provningsuppställningen är schematiskt beskriven i figur 6.1.

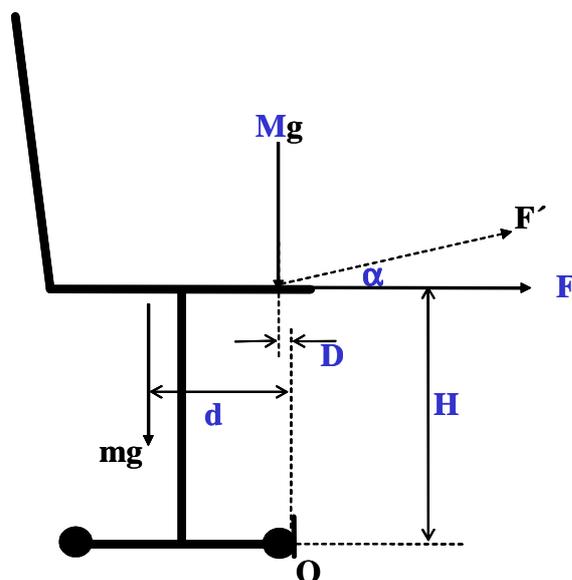
Här är höjden H lika med avståndet från stolens angreppspunkt på stödet till draglinans fästpunkt, D är avståndet från stödets angreppspunkt till massans tyngdpunkt, d är avståndet från stolens tyngdpunkt till stödet, m är stolens massa och α är den verkliga angreppsvinkel som linan har vid mätningen. För att bestämma mätosäkerheten i den horisontella kraften F måste man först identifiera de ingående källorna till mätfel, skatta deras respektive mätosäkerheter och deras inflytande på mätresultatet. En studie av problemet identifierar följande källor till mätosäkerhet: Kraftmätningen i dynamometern, \tilde{F} , den applicerade dödvikten M , angreppsvinkeln α , samt avstånden D och H .

De nominella värdena på de ingående storheterna är

$$\begin{aligned}M_0 &= 55 \text{ kg}, m_0 = 20 \text{ kg}, F_0 = 156 \text{ N}, H_0 = 0.45 \text{ m}, \\D_0 &= 0.05 \text{ m}, d = 0.22 \text{ m}, g = 9.81 \text{ m/s}^2\end{aligned}$$

6.1.1 Observationsmetoden

Osäkerheten i mätning av tippningskraft är i detta fall främst intressant i företagets interna jämförelser och populationen av mätförhållanden ges då av de i företaget tillgängliga förhållandena. För att genomföra ett försök för



Figur 6.1: Schematisk skiss av skrivbordsstol med anbringade krafter och intressanta mått.

bestämning av osäkerheten gjordes därför mätningar på samma stol med de tre operatörer som normalt genomför provningen. Angreppsvinkeln α , samt avstånden D och H är variabler som påverkas av operatören. För kraftmätningen användes samma dynamometer vid alla provningar och även vikten M var densamma. Detta innebär att de systematiska felen i dessa variabler inte inkluderas i den skattade mätosäkerheten och att man behöver göra om mätningen om nya instrument införs. Varje operatör genomförde fem provningar och resultaten ses i tabellen nedan.

Mätresultat, skrivbordsstol (N)

operatör	1	2	3
	151	148	146
	151	152	144
	149	151	150
	149	148	147
	151	150	149
medel	150	150	147
st. av.	1.0	1.6	2.5

Tabellen antyder såväl en variation mellan operatörer som en repeterbarhetsvariation. Med hjälp av variansanalysverktyget i Excel (Variansanalys - en faktor) får vi följande tabell

ANOVA						
Variationsursprung	KvS	fg	Mkv	F	p-värde	F-krit
Mellan grupper	26.53333	2	13.26667	3.940594	0.048354	3.88529
Inom grupper	40.4	12	3.366667			
Totalt	66.93333	14				

Variationerna för upprepningar representeras här av "Inom grupper" och ger osäkerhetsbidraget

$$u_{\varepsilon} = \sqrt{MKv_{\varepsilon}} = \sqrt{3.37} = 1.84 \text{ N.}$$

Variationerna mellan operatörer representeras av "Mellan grupper" och deras osäkerhetsbidrag skattas med

$$u_O = \sqrt{\frac{MKv_O - MKv_{\varepsilon}}{5}} = \sqrt{\frac{13.27 - 3.37}{5}} = 1.41 \text{ N.}$$

Mätosäkerheten blir

$$u_{\hat{F}} = \sqrt{MKv_{\varepsilon} + \frac{MKv_O - MKv_{\varepsilon}}{5}} = \sqrt{MKv_O + 4MKv_{\varepsilon}} = \sqrt{\frac{13.27 + 4 \cdot 3.37}{5}} = 2.31 \text{ N.}$$

Antalet frihetsgrader i denna mätosäkerhetsskattning är

$$\nu_{eq} = \frac{(2.31)^4}{\left(\frac{1}{5}\right)^4 \frac{13.27^4}{2} + \left(\frac{4}{5}\right)^4 \frac{3.37^4}{12}} = 6,$$

avrundat nedåt till närmaste heltal. Täckningfaktorn blir i detta fall $k = 2.4$ enligt tabell 2.3.2 på sidan 29.

$$F = \hat{F} \pm 2.4 \cdot u_{\hat{F}} = \hat{F} \pm 5.5 \text{ N.}$$

Detta är den mätosäkerhet som man kan förvänta sig vid slumpmässigt val av operatör, men med användning av samma dynamometer och dödvikt. Den representerar alltså repeterbarhet inom det aktuella företaget. Om reproducerbarheten är intressant så måste man lägga till osäkerhetskomponenter från dynamometerns och dödviktens systematiska fel, vilket vi kommer att göra nedan.

6.1.2 Analysmetoden

För att bestämma osäkerheten i den verkliga kraften vid tippning måste vi dels bestämma osäkerheten i vilken kraft som krävs för tippning, dels osäkerheten i dynamometerns visning, dels osäkerheten i kraftens angreppsvinkel. Enligt specifikationer uppfyller provningen följande:

- alla krafter ligger inom $\pm 5\%$,
- dödvikten inom $\pm 0.5\%$,
- alla dimensioner inom $\pm 1\text{mm}$,
- angreppsvinkeln inom $\pm 2^\circ$
- och positioneringen av kraften inom $\pm 5\text{mm}$.

Stolens vikt m och tyngdpunktsposition d betraktar vi som konstanter för en bestämd stol och inkluderar inte variation i dessa storheter i mätosäkerheten. Däremot beror höjden H och positionen D på den manuella infästningen av kraftens angreppspunkt och ger därmed mätfel.

Innebörden av specifikationerna är inte given. Vi antar därför att alla de givna intervallen motsvarar 95% konfidensintervall, vilket resulterar i

$$\frac{u_{\tilde{F}}}{F} = 2.5\%, \quad \frac{u_M}{M} = 0.25\%, \quad u_\alpha = 1^\circ, \quad u_H = u_D = 2.5 \text{ mm}$$

För att bestämma de ingående mätstorheternas inflytande på den horisontella kraft F som krävs för tippning ställer vi upp en momentjämvikt kring punkten O , vilken enligt mekanikens lagar skall vara lika med noll i momentet strax före tippning.

$$FH - mgd - MgD = 0$$

$$F = \frac{g}{H} (MD + md)$$

Ur denna formel kan känslighetskoefficienterna för de ingående mätfelen bestämmas, antingen genom differensmetoden (3.10) eller genom partiell derivering i enlighet med (3.12). Däremot kan man inte använda den förenklade metoden för relativa osäkerheter (3.13) eftersom funktionen här innehåller en summa. De relativa osäkerheter som är givna måste därför först räknas om till absoluta osäkerheter, vilket vi gör med de nominella värdena på kraft och dödvikt: $F_0 = 156 \text{ N}$ och $M_0 = 55 \text{ kg}$,

$$u_F = \frac{2.5 \cdot 156}{100} = 3.9 \text{ N}, \quad u_M = \frac{55 \cdot 0.25}{100} = 0.14 \text{ kg}$$

Differensmetoden ger följande känslighetskoefficient för dödvikten M ,

$$\begin{aligned} c_M &= \frac{g}{H_0} \frac{[D_0 (M_0 + u_M) + md] - [D_0 (M_0 - u_M) + md]}{2u_M} = \frac{g}{H_0} D_0 = \\ &= \frac{9.81}{0.45} 0.05 = 1.09 \text{ m/s}^2 = 1.09 \text{ N/kg} \end{aligned}$$

och motsvarande beräkningar för övriga felkällor ger känslighetskoefficienterna

$$c_D = \frac{g}{H_0} M_0 = \frac{9.81}{0.45} 55 = 1200 \text{ N/m},$$

$$c_H = -\frac{g(M_0 D_0 + m_0 d_0)}{(H_0 + u_H)(H_0 - u_H)} = -\frac{9.81(55 \cdot 0.05 + 20 \cdot 0.22)}{(0.45 + 0.0025)(0.45 - 0.0025)} = -346 \text{ N/m}.$$

Därmed har osäkerhetskomponenterna för den nödvändiga kraften F bestämts. Återstår att bestämma osäkerheten i mätningen av denna kraft. Där finns två felkällor, nämligen dynamometers visning, \tilde{F} , och angreppsvinkeln α . Dynamometers standardosäkerhet har vi bedömt från specifikationen till 3.9 N. Angreppsvinkelns inverkan behandlas i avsnitt 5.1 där vi förordar en justering av värdet och en approximativ standardosäkerhet. I det aktuella fallet blir väntevärdet av det relativa vinkelfelet lika med

$$E \left[\frac{e}{\tilde{F}} \right] = \frac{1}{\cos(\sigma_k)} - 1 = \frac{1}{\cos(1^\circ)} - 1 = 1.5 \cdot 10^{-4} = 0.015\%$$

vilket tydligt är försumbart jämfört med kraftmätningens egen osäkerhet på 2.5%. Också standardosäkerheten är då försumbar.

Osäkerhetstabell, analysmetoden

Källa	u_i	typ	ν_i	$ c_i $	$u_i c_i $	$t_{\nu_i} u_i c_i $
M	0.14 kg	B	-	1.09	0.15	0.3
D	0.0025 m	B	-	1200	3	6
H	0.0025 m	B	-	346	0.87	1.7
\tilde{F}	3.9 N	B	-	1	3.9	7.8
α	1°	B	-	-	-	-
$u_{\tilde{F}}^2 = \sum u_i^2 c_i^2$	-	-	-	-	25	-
$u_{\hat{F}}$	-	-	-	-	5 N	-

Då samtliga mätosäkerheter här är av typ B kan antalet frihetsgrader inte bestämmas och jämförelsemåttet i sista kolumnen bestäms med oändligt antal frihetsgrader. De komponenter som har signifikant inverkan på den slutgiltiga mätosäkerheten bör i framtiden bestämmas med statistiska försök. Vi ser att mätosäkerheten domineras av osäkerheterna i dynamometern och avstånden D och H . Användningen av tumregeln innebär att osäkerhetskomponenter förutom dessa kan negligeras då återstående komponenten i tabellen har jämförelsemåttet $u_M |c_M| = 0.15 < 0.2 \cdot u_{\tilde{F}} |c_{\tilde{F}}| = 0.2 \cdot 3.9 = 0.78$. Denna negligering ger ett försumbart fel i uppskattningen. Om man

hoppas att typ B-osäkerheterna är konservativt skattade blir den utvidgade mätosäkerheten

$$U_{\hat{F}} = 2 \cdot u_{\hat{F}} = 10 \text{ N.}$$

Vi kompletterar nu analysmetoden med resultaten från observationsmetoden och får följande tabell:

Osäkerhetstabell, observation+analysmetod

Källa	u_i	typ	ν_i	$ c_i $	$u_i c_i $	$t_{\nu_i} u_i c_i $
M	0.14 kg	B	-	1.09	0.15	0.3
D, H, α, \dots	2.31 N	A	6	1	2.31	5.7
\tilde{F}	3.9 N	B	-	1	3.9	7.8
$u_{\tilde{F}}^2 = \sum u_i^2 c_i^2$	-	-	89	-	20.6	-
$u_{\hat{F}}$	-	-	-	-	4.5 N	-

Här har vi olika antal frihetsgrader för de ingående osäkerhetsskattningarna och bestämmer ekvivalenta antalet frihetsgrader för slutresultatet:

$$\nu_{\hat{F}} = \frac{20.6^2}{2.31^4/6 + 0 + 0} = 89$$

Vi ser att eftersom typ B-osäkerheten i dynamometern dominerar blir antalet frihetsgrader stort och den utvidgade mätosäkerheten blir

$$U_{\hat{F}} = 2 \cdot u_{\hat{F}} = 9 \text{ N.}$$

Det statistiska försöket gjorde i detta fall att mätosäkerheten minskade något. Ytterligare undersökningar genom regelbunden kalibrering av den dominerande källan \tilde{F} skulle göra den till typ A och ge än bättre kontroll över mätosäkerheten.

6.2 Mätosäkerhet hos skjutmått

Mätfel som kan uppträda vid användning av ett analogt skjutmått är kalibreringsfel, avläsningsfel på nonieskalan, materialförändring pga. temperaturförändringar, ovalitet hos mätobjektet, vinkelfel vid mätning, samt glapp mellan skjutmåttets skänklar och mätobjektet. Att bestämma denna mätosäkerhet med hjälp av analysmetoden är behäftat med stora svårigheter då flera av felkällorna inte är så lätta att bedöma. Avläsningsfel på nonieskala, vinkelfel och glapp är felkällor som samtliga beror mycket på operatörens

skicklighet. Kalibreringsfelet är också något svårt att avgöra storleksmässigt, eftersom kalibreringen ofta begränsas till en kontroll av nollan med regelbundna mellanrum.

Här kommer observationsmetoden till sin fulla rätt. Genom att de dominerande felkällorna hos instrumentet är operatörsberoende kan man genomföra en jämförande mätning med hjälp av ett antal operatörer som representerar populationen av användare. Mätningar bör genomföras på ett eller flera objekt som är representativa för användningsområdet.

En sådan jämförande mätning genomfördes på följande sätt: Ett cylindriskt mätobjekt tillverkades i stål med flera cylindriska steg på ett sådant sätt att ovaliteten kunde betraktas som försumbar. Objektet cirkulerades tillsammans med ett analogt skjutmått i stål till 23 olika operatörer, som var och en oberoende fick mäta de olika diametrarna. I tabellen på nästa sida ses resultaten från de fyra olika diametrarna i mm, samt respektive medelvärde och standardavvikelse. Vi ser att standardavvikelsen varierar något mellan de olika måtten, men tycks inte vara proportionell mot måttet. Därmed tycks det vara mer relevant att ange en absolut mätosäkerhet och inte en relativ. Om man nu vill ha en gemensam mätosäkerhetsangivelse för mätning av ett skjutmått av denna typ är det lämpligt att väga samman de fyra resultaten enligt formeln för sammanvägd standardavvikelse (2.25). Då antalet frihetsgrader är detsamma i var och en av de fyra skattningarna blir sammanvägningen helt enkelt medelvärdet av variansskattningarna:

$$s = \sqrt{s_1^2 + s_2^2 + s_3^2 + s_4^2} = \sqrt{\frac{0.047^2 + 0.069^2 + 0.043^2 + 0.03^2}{4}} = 0.049 \text{ mm}$$

Denna standardavvikelse representerar nu operatörsinflytandet på mätosäkerheten. Av de felkällor vi listade i inledningen av detta avsnitt återstår då kalibreringsfel, materialförändring pga. temperaturförändringar och ovalitet hos mätobjektet.

Kalibreringsfelet kan, om regelbunden kontroll genomförs, möjligen betraktas som försumbart, men kunde i det aktuella fallet också kontrolleras genom en referensmätning av objektet med mikrometer.

Vid referensmätningen var såväl skjutmått som objekt tillverkat av stål och därmed kommer längdutvidgning i objektet att kompenseras av samma längdutvidgning i instrumentet. Detta gäller vid varje mätning av objekt i stål. Vid mätning av annat material tillkommer en felkälla pga temperaturavvikelse från nominell temperatur.

Problem med ovalitet hos mätobjektet kan inte betraktas som en del av mätosäkerheten hos skjutmättet, men måste beaktas vid varje mätning där ett enda diametervärde skall bestämmas. Detta kan också göras med observationsmetoden, genom att man på några objekt ur den intressanta populationen genomför mätning i flera positioner och beräknar standardavvikelsen för variationen.

Resultat från mätning med skjutmått (mm)

10.65	49.80	30.20	95.30	
10.60	49.80	30.20	95.25	
10.60	49.80	30.15	95.30	
10.65	49.80	30.25	95.30	
10.60	49.75	30.15	95.25	
10.60	49.75	30.15	95.25	
10.65	49.80	30.20	95.30	
10.60	49.75	30.15	95.20	
10.60	49.75	30.20	95.25	
10.65	49.80	30.20	95.25	
10.60	49.75	30.15	95.30	
10.70	49.80	30.30	95.30	
10.70	49.90	30.25	95.30	
10.60	49.80	30.20	95.30	
10.60	49.75	30.20	95.25	
10.60	49.80	30.20	95.25	
10.65	49.75	30.20	95.30	
10.55	49.75	30.15	95.25	
10.75	49.75	30.30	95.25	
10.70	49.80	30.20	95.25	
10.65	49.80	30.20	95.30	
10.60	49.50	30.20	95.25	
10.60	49.80	30.20	95.30	
\bar{x}	10.63	49.77	30.20	95.27
s	0.047	0.069	0.043	0.030

6.3 Kalibrering av kraftgivare

Kraftgivare kan kalibreras genom att montera givaren i en dödviktsmaskin, belasta den med ett antal olika vikter och läsa av givarens utsignal. Enligt etablerad standard skall kalibreringen utföras vid minst 5 olika nivåer, jämnt fördelade över givarens giltighetsområde. Vid varje nivå görs normalt tre oberoende mätningar. Resultatet av mätning vid 10 nivåer + noll-last redovisas t. ex. som

F(kN)	avläst, $y_{i,1}$	avläst, $y_{i,2}$	avläst, $y_{i,3}$	$\bar{y}_{i,\cdot}$	maxdiff
0	0.0108	0.0089	0.0092	0.0096	0.0019
1	0.1190	0.1207	0.1194	0.1197	0.0017
2	0.2453	0.2467	0.2455	0.2459	0.0014
3	0.3887	0.3895	0.3864	0.3882	0.0031
4	0.5364	0.5375	0.5361	0.5367	0.0014
5	0.6967	0.6988	0.6977	0.6977	0.0021
6	0.8617	0.8620	0.8611	0.8616	0.0009
7	1.0286	1.0280	1.0276	1.0281	0.0010
8	1.1943	1.1959	1.1932	1.1945	0.0027
9	1.3555	1.3540	1.3552	1.3549	0.0015
10	1.5118	1.5115	1.5100	1.5111	0.0018

där man alltså gjort tre mätserier vid 10 lastnivåer och efter varje mätserie gjort en förnyad avläsning vid noll-last.

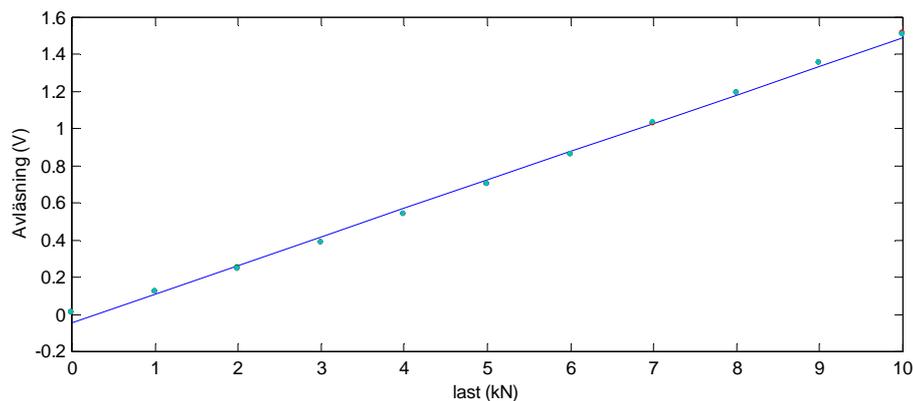
För att bestämma mätosäkerheten vid framtida användning av instrumentet finns det olika angreppssätt. Vi skall här först beskriva det traditionella sättet och demonstrera dess svagheter och sedan presentera ett angreppssätt som bättre överensstämmer med den här redovisade statistiska teorin.

För att utvärdera kalibreringsresultatet går man traditionellt tillväga på följande sätt: För varje nivå (förutom kontroll-nollnivån, som används för att bestämma ett **nollpunktsfel**) bestäms medelvärdet (\bar{y}) och den maximala differensen, $D_i = \max(y_{i,\cdot}) - \min(y_{i,\cdot})$. Den största av av dessa differenser $D = \max(D_i)$ definieras som **repeterbarhetsfel**. I vårt fall blir detta

0.0031 V. Därefter gör man en minsta-kvadrat-anpassning av medelvärdena $\{\bar{y}_i; i = 1, 2, \dots, 11\}$, vilket i det aktuella exemplet ger formeln

$$\hat{y} = 0.1536 \cdot x - 0.0456. \quad (6.1)$$

Här betecknar \hat{y} ett predikerat värde vid en ny avläsning när lasten är x . I figur 6.2 ser vi resultatet av den linjära anpassningen tillsammans med avläsningarna.



Figur 6.2: Linjär anpassning vid kalibrering.

Olinjäritetsfelet definieras sedan som den maximala avvikelserna från de uppmätta medelvärdena,

$$e_{ol} = \max_i (|\hat{y}_i - \bar{y}_{i,\cdot}|)$$

Detta blir i vårt exempel 0.0552. Genom att jämföra repeterbarhetsfelet, olinjäritetsfelet, samt nollpunktsfelet med existerande standarder kan lastcellen klassificeras och därmed godkännas eller underkännas för sitt avsedda användningsområde. Ytterligare användning av kalibreringsresultatet är att finna vilken mätosäkerhet man kommer att ha vid framtida användning av lastcellen. Den totala **mätosäkerheten** vid användning av kraftmätdonet beräknas enligt standardmetodik som

$$U = 2 \cdot \sqrt{\left(\frac{0.0552}{\sqrt{3}}\right)^2 + \left(\frac{0.0031}{\sqrt{3}}\right)^2} = 0.064, \quad (6.2)$$

där nollpunktsfelet här antas var försumbart. Man summerar alltså repeterbarhetsfelet och olinjäritetsfelet kvadratisk efter att ha normerat felen med faktorn $\sqrt{3}$. Faktorn $\sqrt{3}$ härrör sig från antagande om likformig fördelning, tvåan framför rottecknet härrör sig från antagande om normalfördelad total mätosäkerhet och den kvadratiske summeringen bygger på antagande om

oberoende fel. Notera att U blir större ju fler mätningar som görs, eftersom max-differensen D växer med antalet mätningar, dvs. ju bättre kalibrering, desto större mätosäkerhet!

Mätosäkerheten används sedan som ett approximativt 95%-igt konfidensintervall för framtida avläsningar,

$$y(x) = \hat{y}(x) \pm U = 0.1536 \cdot x - 0.0456 \pm 0.064$$

Vi skall nu studera denna metodik i ljuset av vanlig linjär regressionsteori.

En matematisk modell för kalibreringssituationen kan skrivas,

$$Y(x) = \beta_0 + \beta_1 x + \varepsilon + e(x)$$

där den nya termen $e(x)$ representerar ett fel som inte är slumpmässigt, utan beror på fel i den linjära modellen; en olinjäritet. I vår kalibreringsmätning med 11 nivåer och 3 upprepningar får vi

$$Y_{i,j}(x_i) = \beta_0 + \beta_1 x_i + \varepsilon_{i,j} + e_i \quad , \quad i = 1, 2, \dots, 11 \quad , \quad j = 1, 2, 3$$

Det slumpmässiga felet modelleras här som oberoende slumpvariabler och är alltså unikt för varje mätning. Det systematiska felet däremot följer x och är detsamma inom upprepningarna. Efter medelvärdesbildningen över upprepningarna får vi

$$\bar{Y}_i(x_i) = \beta_0 + \beta_1 x_i + \frac{1}{3} \sum_{j=1}^3 \varepsilon_{i,j} + e_i \quad , \quad i = 1, 2, \dots, 11$$

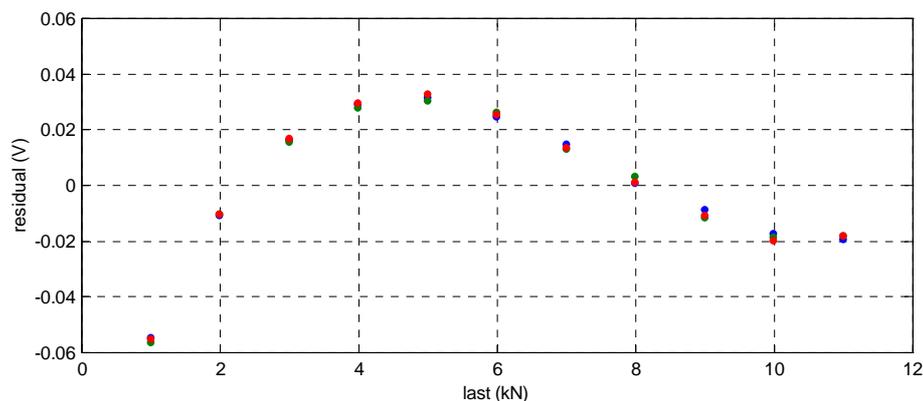
Om variansen för det slumpmässiga felet är σ^2 så kommer medelvärdesbildningen att reducera variansen med en faktor 3 till $\frac{\sigma^2}{3}$. De residualer som man erhåller som resultat av den linjära regressionen på $\bar{Y}_i(x_i)$ innehåller dock fortfarande en del av det slumpmässiga felet och överskattar därmed olinjäriteten något. Följaktligen kan man förvänta att det totala felet enligt (6.2) blir överskattat.

Varianserna för repeterbarhetsfelet resp. olinjäritetsfelet betraktas i den standardiserade proceduren som härrörda från likformigt fördelade slumpvariabler, bestämda via max-uppskattningar. Detta är en grov uppskattning, som saknar egentlig grund i statistisk teori, men som används då traditionella statistiska metoder inte är tillämpliga.

En rimlig motivering för att betrakta ett systematiskt fel som slumpmässigt är att man inte vet vid vilka x -värden den framtida användningen av mätdonet kommer att ske och att man därmed kan betrakta valet av x (och därmed storleken på det systematiska felet) som slumpmässigt. Ett sådant resonemang är analogt med den allmänna syn på slumpmässighet och systematik som denna handbok grundar sig på.

Den slutliga beräkningen av mätosäkerheten i den standardiserade metoden (6.2) bygger på antagandena 1) att det systematiska felet kan betraktas

som slumpmässigt, samt 2) att det totala felet approximativt följer normalfördelningen. Man kan då fråga sig om inte den vanliga linjär-regressionsmetodiken är direkt tillämpbar. Detta skulle innebära att vi betraktar summan $\frac{1}{3} \sum_{j=1}^3 \varepsilon_{i,j} + e_i$ som slumpmässig och normalfördelad och följer den normala statistiska proceduren. Vi betraktar alla 11 medelvärdena som oberoende observationer och gör en linjär anpassning enligt minsta-kvadratmetoden precis som i standardproceduren. Studier av residualerna ger bilden 6.3.



Figur 6.3: *Residualer vid linjär regression på kalibreringsresultat.*

Residualerna ser allt annat än slumpmässiga ut, då det systematiska felet dominerar. Notera att detta beror på att residualerna här är ordnade i enlighet med referensvärdet. En slumpmässig ordning av värden på x -axeln skulle ge en residualplot som ser slumpmässig ut, vilket precis motsvarar den situation som användaren står inför när framtida användningsnivåer är obestämbara. Vi beräknar standardavvikelsen för $\frac{1}{3} \sum_{j=1}^3 \varepsilon_{i,j} + e_i$ för att jämföra med den etablerade metodiken. Formel (2.28) ger i detta fall

$$s = 0.029 .$$

För att jämföra resultatet med den standardiserade metoden vill vi ha ett resultat som motsvarar (6.2), dvs ett approximativt 95%-igt prediktionsintervall för framtida avläsningar. Vi använder tumregel 8 på sidan 93 där t -värdet tas ur t -fördelningstabell för 9 frihetsgrader och 95% : $t = 2.26$

Vårt jämförelseintervall blir enligt (2.31),

$$y(x) = 0.1536 \cdot x - 0.0456 \pm 0.029 \cdot t_{0.025} \cdot \sqrt{\frac{2}{n} + 1} =$$

$$0.1536 \cdot x - 0.0456 \pm 0.029 \cdot 2.26 \cdot \sqrt{\frac{2}{11} + 1} =$$

$$0.1536 \cdot x - 0.0456 \pm 0.071 .$$

Vi får alltså i detta fall ett prediktionsintervall som är något större än i den standardiserade metodiken. Fördelen med den här redovisade metoden är att den tar hänsyn till antalet försöksnivåer på ett statistiskt korrekt sätt via t -fördelningen och rotuttrycket.

En komplikation är att prediktionsintervallet motsvarar en prediktion grundad på medelvärdet av tre avläsningar. Om framtida avläsningar inte är medelvärden så tillkommer ett slumpmässigt felbidrag. Det slumpmässiga felets standardavvikelse kan vi emellertid lätt skatta genom att beräkna den sammanvägda standardavvikelsen,

$$s_\varepsilon = \sqrt{\frac{\sum \nu_i s_i^2}{\sum \nu_i}} = \sqrt{\frac{\sum 2s_i^2}{11 \cdot 2}} = 0.00098 .$$

Den totala mätosäkerheten vid en enstaka prediktion skattas i detta fall som

$$u_y = \sqrt{s_e^2 + s_\varepsilon^2} = \sqrt{s^2 + 2/3 s_\varepsilon^2} = \sqrt{0.029^2 + 2/3 \cdot 0.00098^2} = 0.029.$$

Här har vi alltså bara lagt till två tredjedelar av slumpvariansen eftersom en tredjedel ingår i skattningen av residualernas varians s . I detta fall är tydligen slumpfelet försumbart och mätosäkerheten blir lika med den skattade regressionsvariansen. Antalet frihetsgrader i denna mätosäkerhet bestäms med hjälp av Welch-Satterwaite-formeln (2.21)

$$\nu_{eq} = \frac{(0.029^2 + 2/3 \cdot 0.00098^2)^2}{\frac{0.029^4}{9} + \frac{(2/3 \cdot 0.00098^2)^2}{22}} = 9$$

och blir här lika med nio eftersom slumpfelet är av försumbar storlek.

6.3.1 Inversproblemet

Vid användning av den erhållna kalibreringskurvan kommer man att läsa av instrumentet i volt (y) och vill veta vilken kraft (x) som orsakat spänningen. Man behöver alltså invertera uttrycket (6.1) till

$$x_1(y) = \frac{y + 0.0456}{0.1536} = 6.51 y + 0.297.$$

Som angivits i kapitlet om linjär regression blir prediktionsintervallet här detsamma som ovan skalat med den skattade lutningen:

$$x(y) = 6.514y + 0.297 \pm \frac{0.071}{0.1536} = 6.51 y + 0.297 \pm 0.46$$

Om vi istället gör regressionen i andra riktningen får vi följande resultat:

$$x_2(y) = 6.49y + 0.309$$

och standardavvikelsen skattas nu med hjälp av residualerna till $s = 0.190$ och konfidensintervallet blir

$$x(y) = 6.49y + 0.309 \pm 0.190 \cdot 2.26 \cdot \sqrt{\frac{2}{11} + 1} = 6.49y + 0.309 \pm 0.47.$$

Vi får snarlika uttryck med de två metoderna, men den andra är enklare att tillämpa då ingen invertering av koefficienterna behövs och standardavvikelsen fås direkt ur residualerna.

6.3.2 Sammanfattning

Vid kalibrering med linjär regression görs upprepade mätningar på ett antal bestämda nivåer. För varje nivå beräknas medelvärdet av avläsningarna och kalibreringskurvan bestäms med hjälp av linjär regression i omvänd riktning, dvs. med de avlästa medelvärdena som x -värden och de inställda nivåerna som y -värden. Därmed kommer olinjäritetsfel att betraktas som slumpmässigt dragna och den resulterande mätosäkerheten att avse instrumentets hela mätområde.

Bilaga

Svar till övningsuppgifterna

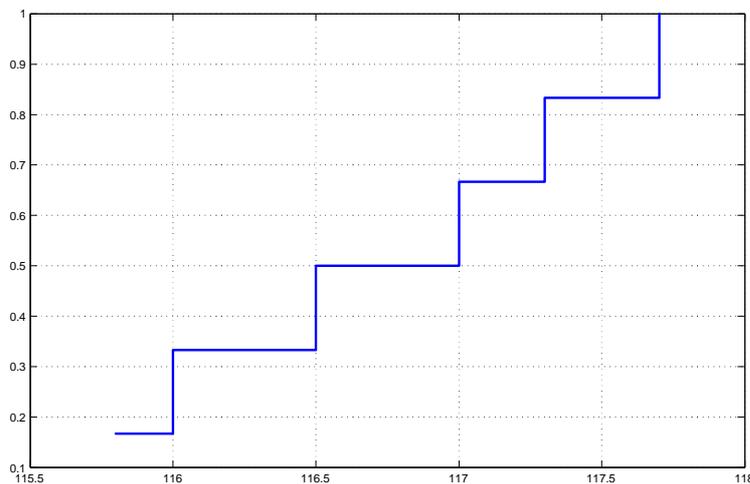
Problem 2.2.1

Medelvärde för operatör 1: 116.7

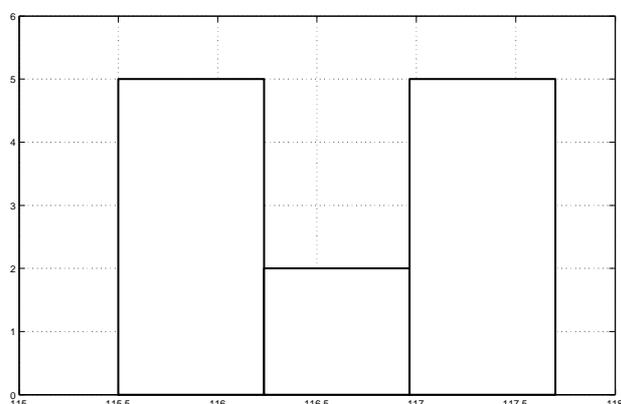
Standardavvikelse för operatör 1: 0.75

Den empiriska fördelningsfunktionen för operatör 1:

x	115.8	116.0	116.5	117.0	117.3	117.7
$\hat{F}(x)$	1/6	2/6	3/6	4/6	5/6	1



Histogram för hela datamängden (tre intervall):



Variationsvidden för hela datamängden: 2.2.

Problem 2.3.1

Konfidensintervall för differensen: 0.09 ± 0.14 (med 95% konfidens)
Rören har ej signifikant olika diametrar.

Problem 2.3.2

$$P(-2 \leq Z \leq 2) = 0.9545$$

Problem 2.3.3

Det krävs minst fyra mätningar.

Problem 2.3.4

Konfidensintervall för diametern: 37.22 ± 0.05 (med 95% konfidens)
Röret är godkänt på signifikansnivån 2.5%.

Problem 2.3.5

Konfidensintervall för differensen: 0.06 ± 0.28 (med 95% konfidens)
Rören har ej signifikant olika diametrar.

Problem 2.4.1

Sammanvägningen blir roten ur medelvärdet av variansskattningarna:

$$s_{SV} = \sqrt{\frac{1}{n} \sum_{i=1}^n s_i^2}$$

Problem 5.2.1

Metoden med negligerad samverkan:

$$GRR = \sqrt{0.003 + 0 + 0.00105} = 0.064$$

Metoden med inkluderad samverkan:

$$GRR = \sqrt{0.0013 + 0.0022 + 0.00095} = 0.067$$

Litteraturförteckning

- [1] ISO Guide, Guide to the expression of uncertainty in measurement, International Organisation for Standardization, 1995.
- [2] Semyon Rabinovich, Measurement Errors, Theory and Practise, AIP Press, 1992.
- [3] Hugh W. Coleman, W. Glenn Steele, Jr., Experimentation and uncertainty analysis for engineers, John Wiley & Sons, 1989.
- [4] Magnus Arnér, Mätosäkerhet, Kalibrering, felfortplantning och R&R-studier, Studentlitteratur, 2002.
- [5] I. Lira, Evaluation the mesurement uncertainty, Fundamentals and practical guidance, Institute of Physics Publishing, Bristol and Philadelphia, 2002.
- [6] Donald J. Wheeler, Richard W. Lyday, Evaluating the measurement process, SPC Press, 1989.
- [7] Douglas C. Montgomery, George C. Runger, Applied statistics and probability for engineers, John Wiley & Sons, 2003.
- [8] Leslie R. Pendrill, Meeting future needs for Metrological Traceability - A physicist's view, *Accreditation and Quality Assurance- Journal for Quality, Reliability and Comparability in Chemical Measurement*, 10, 133 - 9, 2005.
- [9] QS9000 Quality System Manuals, MSA-3, Measurement System Analysis, 2002.
- [10] ISO Standards, ISO 5725-2, Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method, International Organisation for Standardization, 1994.
- [11] B Bergman, B Klefsjö, Statistisk kvalitetsstyrning, Studentlitteratur, 1985.

Sakregister

- analysmetoden, 46, 68
- avrundning
 - fördelning, 72
 - regler, 52
- box-plot, 64
- empirisk fördelning, 19
- fördelningsfunktion, 17
- frihetsgrader, 21
 - beteckning, 17
- histogram, 19
- Inledning, 1
 - sammanfattning, 14
- justering, 85
- känslighet, 85
- känslighetskoefficient, 73
 - beteckning, 17
 - differensmetod, 74
 - partiell derivering, 76
- Kalibrering, 85
 - sammanfattning, 97
- konfidensintervall, 26, 31, 42, 47
- korrelation, 108
- kvalitetsstyrning, 103
- likformig fördelning, 23
- linjär regression, 35, 85
- Litteraturförteckning, 131
- mätförhållanden, 8
- mätfel, 5, 54
- mätmetod, 6
- mätmiljö, 6
- mätobjekt, 5
- mätosäkerhet, 7
 - utvidgad, 42
- mätsystemsanalys, 104
- mätutrustning, 5
- målvärde, 103
- medelvärde
 - beteckning, 17
- MSA, 104
- nomenklatur, 16
- normalfördelning, 21
- observationsmetoden, 46, 56
- operatör, 6
- osäkerhetskomponenter, 78
- referens, 56
- repeterbarhet, 3, 11
- reproducerbarhet, 3, 11
- shuntkalibrering, 85, 93
- signifikansnivå, 27
- slumpmässiga fel, 9
- sortering, 112
- spårbar, 85
- standardavvikelse, 9, 18, 42, 45
 - beteckning, 17
- standardosäkerhet, 9, 42, 45, 71
 - beteckning, 17
 - typ A, 68, 71
 - typ B, 68, 72
 - utvidgad, 9
- statistika, 26
- Statistiska grunder, 15
 - sammanfattning, 42

styrdiagram, 103
systematiska fel, 9, 51

t-fördelningen, 29, 49
 tabell, 30, 43
täckningsfaktor, 45, 49
 beteckning, 17
täthetsfunktion, 18
tolerans, 103, 105

utfall, 17

väntevärde, 18
varians, 18
variansanalys, 61
variationsvidd, 21
vinkelfel, 101

FCC
Chalmers Science Park
SE-412 88 Göteborg
Sweden

Telephone:+46 (0)31 7724000

Telefax:+46 (0)31 827035

E-mail:info@fcc.chalmers.se

Internet:www.fcc.chalmers.se