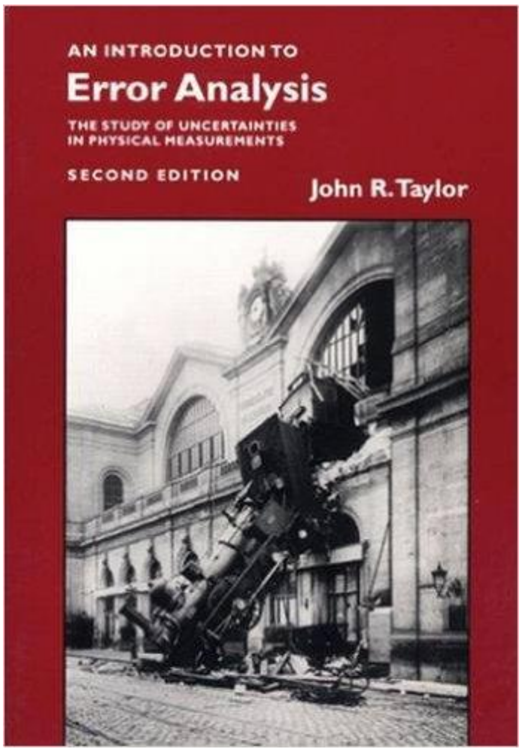
«Uncertainty of the Uncertainty»
Del 5 av 6

v/Rune Øverland, Trainor Elsikkerhet AS

Dette er femte del i artikkelserien om «Uncertainty of the Uncertainty».

Jeg skal vise deg utledning av «Uncertainty of the Uncertainty»-formelen:



$\frac{1}{\sqrt{2v}} = \frac{1}{\sqrt{2(n-1)}}$ [1]

Denne finner du i ISO «Guide to the Expression of Uncertainty in Measurement» seksjon E.4.3 formel (E.7).

Min presentasjon av formel [1] bygger på boken til John R. Taylor: Error Analysis.

«Uncertainty of the Uncertainty»-formelen uttrykker usikkerheten i estimeringen av bredden av standardavviket. Det vil si hvor mange signifikante sifre skal vi bruke for å angi standardavviket. Så, hvorledes kommer man frem til denne $\frac{1}{\sqrt{2(n-1)}}$ formelen?

Hvordan varierer andelen av usikkerheten?

Variansen til et måleoppsett (med n < 30 observasjoner) kan kalkuleres slik:

$s^2 = \frac{1}{(n-1)} \sum (x_i - \bar{x})^2$ [2]

Da ($\frac{1}{(n-1)}$) vil ha en konstant verdi når man har valgt antall [n] observasjoner, vil ikke denne brøken variere. I den videre utledningen om hvorledes andelen av usikkerheten varierer, kan vi derfor se bort fra ($\frac{1}{(n-1)}$).

Andelen av usikkerhet av s^2 [i likning 2] vil derfor ligge i usikkerheten i deluttrykket $\sum (x_i - \bar{x})^2$.

Vi oppsummerer dette slik:

Usikkerhet i estimatet av [s^2] = Usikkerhet i estimatet av [$\sum (x_i - \bar{x})^2$]. [3]

Generelt uttrykk for usikkerhet

Generelt, så kan usikkerheten i estimatet av en parameter [q] – «quantity» - uttrykkes slik:

usikkerhet i estimatet av q = $\sqrt{\overline{q^2} - \bar{q}^2}$ [4]

, hvor

$\overline{q^2}$ gjennomsnittlig verdi av kvadratet av parameteren

$\bar{q}^2$ kvadratet av gjennomsnittlig verdi av parameteren.

Uttrykk for usikkerhet i estimatet

Fra [3] har vi at:

$\bar{q} = \sum(x_i - \bar{x})^2$ [5]

$\bar{q}$ finner vi ved å kvadrere hver residual $(x_i - \bar{x})$ og deretter finner summen [Σ] av disse.

Vi har nå:

Usikkerhet i estimatet av q= $\sqrt{(\overline{q^2}) - (\bar{q})^2}$ [6]

Utlede uttrykket $(\bar{q})^2$ i [6]

Vi bestemmer først kjernen ($\bar{q}$) i [6]. Fra tidligere [2] har vi

$s^2 = \frac{1}{(n-1)} \sum(x_i - \bar{x})^2$ [7]

som gir:

$(n-1)\ s^2 = \sum(x_i - \bar{x})^2$ [8]

$\sum(x_i - \bar{x})^2 = (n-1)\ s^2$ [9]

$\bar{q} = (n-1)\ \sigma^2$ [10]

Når vi kvadrerer venstre og høyre side, får vi:

$(\bar{q})^2 = [(n-1)\ \sigma^2]^2 = (n-1)^2\sigma^4$ [11]

Vi har nå utledet det høyre leddet i [6] til å kunne skrives $(n-1)^2\sigma^4$.

Utlede [$\overline{(q^2)}$]-leddet i [6]

Å bestemme [$\overline{(q^2)}$]-leddet i [6] er svært komplisert.

Vi starter med å bestemme (q^2), for deretter å bestemme gjennomsnittsverdien: $\overline{(q^2)}$.

Det kan vises at (q^2) kan settes sammen av tre ledd

$$q^2 = \left(\frac{n-1}{n}\right)^2 \sum_i^n (x_i)^2 \ \sum_j^n \left(x_j\right)^2 - 2\left(\frac{n-1}{n^2}\right) \sum_i^n (x_i)^2 \sum_j^n \sum_{k\neq j}^n x_j x_k + \frac{1}{n^2} \sum_j^n \sum_{k\neq j}^n x_j x_k \sum_m^n \sum_{n\neq m}^n x_m x_n \quad [12]$$

$$q^2 = A - B + C \quad [13]$$

For å bestemme gjennomsnittsverdien av q^2 må vi addere gjennomsnittsverdiene av $\bar{A}$, $\bar{B}$ og $\bar{C}$.

$\overline{(q^2)} = \bar{A} + \bar{B} + \bar{C}$

Her må man analysere leddene $\bar{A}$, $\bar{B}$ og $\bar{C}$ hver for seg, og vi starter med leddet $\bar{A}$.

$\bar{A}$-leddet i q^2

$$\bar{A} = \left(\frac{n-1}{n}\right)^2 \sum_i^n (x_i)^2 \ \sum_j^n \left(x_j\right)^2 \quad [14]$$

Vi studerer først høyresiden av uttrykket [14] hvor vi har en dobbelsum [$\sum_i^n (x_i)^2 \ \sum_j^n \left(x_j\right)^2$].

Dobbelsummen [$\Sigma\Sigma$] i A inneholder først [n(n-1)]-termer av i $\neq$ j, som gir $\left(\overline{x^2}\right)^2$. Fordi x er normalt distribuert rundt $\mu = 0$, har vi at $\overline{x^2} = \sigma^2$. Dette gir: [n(n-1)] σ^2.

Dobbelsummen [$\Sigma\Sigma$] i A inneholder deretter n-termer av i = j, som danner $()^2()^2 = ()^4$ som gir ($\overline{x^4}$). En integrering viser at $\overline{x^4} = 3\sigma^4$. Dette gir: [n]$3\sigma^4$.

Således, sammenfatter vi disse to deluttrykkene, får vi dobbelsummen [$\Sigma\Sigma$] i A med et gjennomsnitt:

$$\overline{\Sigma\Sigma} = n(n-1)\,\sigma^4 + 3n\sigma^4 \quad [15]$$

$$= \sigma^4\,(3n + n^2 - n) \quad [16]$$

$$= \sigma^4\,(n^2 + 2n) \quad [17]$$

$$\overline{\Sigma\Sigma} = n(n+2)\,\sigma^4 \quad [18]$$

Vi kan nå kan skrive [14]:

$$\bar{A} = \left(\frac{n-1}{n}\right)^2 \overline{\Sigma\Sigma} \quad [19]$$

$$= \left(\frac{n-1}{n}\right)^2 n(n+2)\,\sigma^4 \quad [20]$$

$$= \frac{(n-1)^2}{n^2}\,\overline{n(n+2)}\,\sigma^4 \quad [21]$$

[22]

$\bar{A} = \frac{(n-1)^2(n+2)}{n}\sigma^4$

$\bar{B}$-leddet i q^2

$$\bar{B} = 2\left(\frac{n-1}{n^2}\right)\sum_i^n (x_i)^2 \sum_j^n \sum_{k\neq j}^n x_j x_k \qquad [23]$$

Hver term i summen av B inneholder oddetall i eksponenten av x_i (enten x_i^1 eller x_i^3). Da x er normalt distribuert rundt 0, vil gjennomsnittet av enhver oddetallsekspoent være lik 0.

Således har vi:

$$\bar{B} = 0 \qquad [24]$$

$\bar{C}$-leddet i q^2

$$\bar{C} = \frac{1}{n^2}\sum_j^n \sum_{k\neq j}^n x_j x_k \sum_m^n \sum_{n\neq m}^n x_m x_n \qquad [25]$$

Den høyre siden av uttrykket inneholder kvadruppelsummen ΣΣΣΣ. Denne inneholder n(n-1)-termer hvor j = m og k = n, hver som har gjennomsnittet $\left(\overline{x^2}\right)^2 = \sigma^4$. Den inneholder også n(n-1)-termer hvor j = n og k = m, hver som har gjennomsnittet σ^4. Dette gir: n(n-1) σ^4 + n(n-1) σ^4 = 2n(n-1) σ^4

Alle de gjenværende termer inneholder oddetallsekspoent av x, og har gjennomsnittet 0.

Således, kvadruppelsummen ΣΣΣΣ i C har gjennomsnittet 2n(n-1) σ^4 + 0 = 2n(n-1) σ

$$\overline{\Sigma\Sigma\Sigma\Sigma} = 2n(n-1)\,\sigma^4 \qquad [26]$$

Dette gir oss:

$$\bar{C} = \frac{1}{n^2}\overline{\Sigma\Sigma\Sigma\Sigma} \qquad [27]$$

$$= \frac{1}{n^2} 2n(n-1)\,\sigma 4 \qquad [28]$$

$$\bar{C} = \frac{2(n-1)}{n}\sigma^4 \qquad [29]$$

Bestemme $\overline{(q^2)}$

Vi kan nå sette sammen de tre leddene [22], [24] og [29].

$$\overline{(q^2)} = \bar{A} + \bar{B} + \bar{C} \qquad [30]$$

$$= \frac{(n-1)^2(n+2)}{n}\sigma^4 + 0 + \frac{2(n-1)}{n}\sigma^4 \qquad [31]$$

$$= \frac{(n-1)^2(n+2)+2(n-1)}{n}\sigma^4 \qquad [32]$$

$$\overline{(q^2)} = (n^2 - 1)\,\sigma^4 \qquad [33]$$

Bestemme usikkerhet i q [6]

Vi har nå bestemt verdiene på leddene $(\bar{q})^2$ [fra 11] og $\overline{(q^2)}$ [fra 33].

Vi får nå usikkerhet i estimatet fra [4]

$$q \quad = \sqrt{\overline{(q^2)} \qquad - \{(\bar{q})^2\}} \qquad [34]$$

$$= \sqrt{(n^2 - 1)\,\sigma^4 - \{(n-1)^2\sigma^4\}} \qquad [35]$$

$$= \sqrt{[(n^2 - 1) - (n-1)^2]\,\sigma^4} \qquad [36]$$

$$= \sqrt{[(n^2 - 1) - (n^2 - 2n + 1)]\,\sigma^4} \qquad [37]$$

$$= \sqrt{[n^2 - 1 - n^2 + 2n - 1]\,\sigma^4} \qquad [38]$$

$$= \sqrt{[2n - 2]\,\sigma^4} \qquad [39]$$

$$= \sqrt{[2n - 2]}\ \sqrt{\sigma^4} \qquad [40]$$

$$= \sqrt{2(n-1)}\ \cdot \sigma^2 \qquad [41]$$

Relativ usikkerhet i q

For å bestemme relativ andel av usikkerheten, får vi:

$$\text{Relativ usikkerhet} \quad = \frac{\text{Usikkerhet i estimatet at } q}{Usikkerhet} \qquad [42]$$

$$= \frac{[fra\ 41]}{[fra\ 10]} \qquad [43]$$

$$= \frac{\sqrt{2(n-1)}\cdot\sigma^2}{(n-1)\cdot\sigma^2} \qquad [44]$$

$$= \frac{\sqrt{2(n-1)}\cdot\sigma^2}{\sqrt{(n-1)(n-1)}\cdot\sigma^2} \qquad [45]$$

$$= \sqrt{\frac{2}{(n-1)}} \qquad [46]$$

Vi nærmer oss målet . . .

(Andel av usikkerhet i estimatet for σ^2) = (andel i usikkerhet i q)

Da σ er kvadratroten av σ^2, må andel av usikkerheten av σ være halvparten av q; slik

σ	σ	=	q
σ		=	$\frac{1}{2}q$

Andel av usikkerhet i estimatet av σ = $\frac{1}{2}$ (andel av usikkerhet i q) [47]

Andel av usikkerhet i estimatet av σ $= \frac{1}{2}\left(\sqrt{\frac{2}{(n-1)}}\right)$ [48]

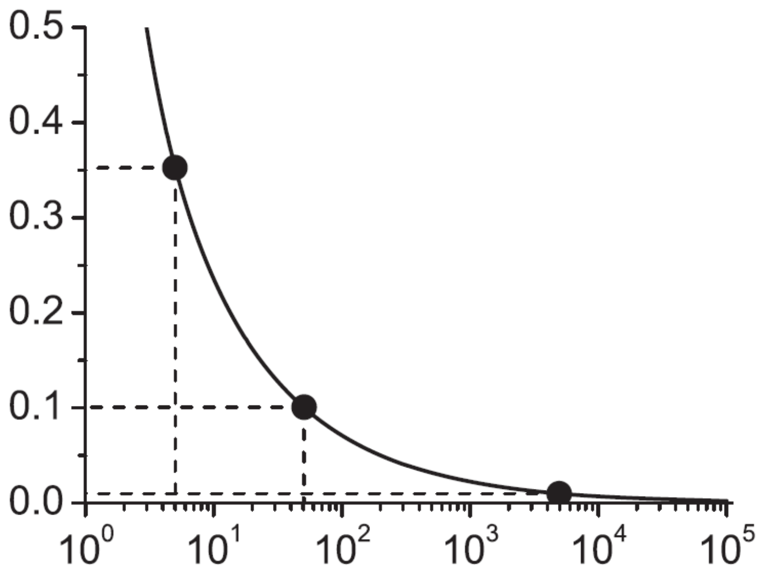
$= \sqrt{\frac{2}{(2\cdot 2)(n-1)}}$ [49]

$= \sqrt{\frac{1}{2(n-1)}}$ [50]

$= \sqrt{\frac{1}{2\upsilon}}$ [51]

Vi har nå kommet til veis ende i utledning av formelen som knytter sammen «Uncertainty of the Uncertainty», og antall [n] observasjoner som inngår i en måleserie.

Grafisk presentasjon av «Uncertainty of the Uncertainty» $\frac{1}{\sqrt{2\upsilon}}$



På den horisontale aksen har vi antall [n] observasjoner i måleserien. På den vertikale aksen har vi andelen av «Error in the Error».

La oss se på tre tilfeller på kurven markert med sorte prikker:

n = 5	Kurven har verdien 0,35. Det innebærer at for en måleserie med fem observasjoner, vil usikkerheten i estimeringen av standard usikkerhet utgjøre 35 prosent. Presisjonen er 65 prosent på standard usikkerhet.
n = 50	Kurven har verdien 0,1. Det innebærer at for en måleserie med femti observasjoner, vil usikkerheten i estimeringen av standard usikkerhet utgjøre 10 prosent. Presisjonen er 90 prosent på standard usikkerhet. Vi ser at etter hvert som antall observasjoner øker, øker også presisjonen på

	standard usikkerhet.
n = 5000	Kurven har verdien 0,01. Det innebærer at for en måleserie med femtusen observasjoner, vil usikkerheten i estimeringen av standard usikkerhet utgjøre 1 prosent. Presisjonen er 99 prosent på standard usikkerhet.

Denne tabellen viser $\frac{1}{\sqrt{2v}} = \frac{1}{\sqrt{2(n-1)}}$, det vil si sammenhengen mellom antall observasjoner [n] i en måleserie og presisjon på standard usikkerhet.

Antall observasjoner	«Uncertainty of the Uncertainty»	Kommentarer
2	0,7071067812	
4	0,4082482905	For en måleserie med fire observasjoner, vil *usikkerheten* i estimatet av bredden standardavvik utgjøre 40 prosent av verdien av standardavviket.
8	0,2672612419	
16	0,1825741858	
32	0,1270001270	
64	0,0890870806	
512	0,0312805624	
1 024	0,0221078844	
2 048	0,0156288161	
4 096	0,0110498924	For en måleserie med 4096 observasjoner, vil *usikkerheten* i estimatet av bredden standardavvik utgjøre 1,1 prosent av verdien av standardavviket.
8 192	0,0078129769	
131 072	0,0019531325	
262 144	0,0013810706	For en måleserie med 262144 observasjoner, vil *usikkerheten* i estimatet av bredden standardavvik utgjøre 1,4 promille av verdien av standardavviket.
524 288	0,0009765634	
1 048 576	0,0006905343	

La oss ta et regneeksempel på «Uncertainty of the Uncertainty»:

Hvor mange [n] observasjoner trenger man for at andelen skal utgjøre en prosent?

$\frac{1}{\sqrt{(2n-2)}} = en\ prosent$	Vi setter opp likningen. [52]
$\left(\frac{1}{\sqrt{(2n-2)}}\right)^2 = 0{,}01^2$	Vi kvadrerer begge sider [53]
$\frac{1}{2n-2} = \frac{1}{10000}$	Vi rydder. [54]
2n – 2 = 10000	Vi kryssmultipliserer [55]
2n = 10002	Vi rydder [56]
n = 5001	Og, vi har kommet frem til løsningen

	[57]

Vi trenger altså 5001 observasjoner i måleserien før andelen av «Uncertainty of the Uncertainty» kommer ned i én prosent. Med andre ord vi trenger flere enn 5001 observasjoner før vi kan bruke to signifikante sifre!

Og, vi trenger flere enn femhundretusen observasjoner før vi kan bruke tre signifikante sifre.

Hva hvis signifikant siffer starter med 1?

Dersom vi skal rapportere med ett signifikant siffer, og dette starter med 1, må vi bruke normale avrundingsregler og sunn fornuft i henhold til GUM seksjon 7.2.6.

For eksempel; 1,4 avrundes til 1 og 1,5 avrundes til 2. Dette gir relativt store avrundingsfeil (rundt 25 %). Det anbefales derfor at når første signifikante siffer er 1, kan man bruke to signifikante sifre.

Eksempel:

σ_m er kalkulert til 0,12345. Denne er basert på 500 observasjoner. Standard usikkerhet rapporteres lik:

σ_m = 0,12.

Avsluttende talleksempel

Vi har følgende måleserie:

Observasjon 1	15,05 mA
Observasjon 2	15,11 mA
Observasjon 3	14,99 mA
Observasjon 4	15,13 mA
Observasjon 5	15,10 mA

Vi kalkulerer måleseriens gjennomsnitt $\bar{x}$ = $[\frac{\sum x_i}{n}]$ = 15,076 mA [58]

Vi kalkulerer måleseriens standard usikkerhet $s_{Bessel} = \sqrt{\frac{\sum(x_i - \bar{x})^2}{(n-1)}} = 0{,}05636149$ mA [59]

Vi kalkulerer utvalgets standard usikkerhet $\sigma_m = \frac{1}{\sqrt{n-1}} s_{Bessel}$ = 0,02819574 mA [60]

Vi rapporter verdiene fra måleserien:

$\bar{x}(\sigma_m) \rightarrow$ 15,08(0,03) mA

Oppsummering

Jeg viste deg utledningen av formelen for «Uncertainty of the Uncertainty»:

$$\frac{1}{\sqrt{2\upsilon}} \qquad [61]$$

Den viser at for måleserier, som inneholder 5000 eller færre observasjoner, vil andelen i «Uncertainty of the Uncertainty» være så høy, at vi kun kan bruke ett signifikant siffer.

Artikkelserien om «Uncertainty of the Uncertainty» avsluttes med Bessel's korreksjon.

På gjensyn!

Med vennlig hilsen
Trainor Elsikkerhet AS
Rune Øverland
Senioringeniør
Tønsberg juni 2016