Statistisk behandling av kalibreringsresultatene – Del 1.

v/ Rune Øverland, Trainor Elsikkerhet AS

Denne artikkelserien handler om statistisk behandling av kalibreringsresultatene.

I de fleste tilfeller er både forventningsverdien μ og standardavviket σ ukjent.

Forventningsverdien forstås som den samme verdien, for eksempel gjennomsnitthøyden på alle mennesker i Norge. Standardavviket forteller noe om variasjonen omkring gjennomsnittshøyden. Hvem vel har kapasitet til å teste hele populasjonen?

I stedet for å ta arbeidet med å måle hele populasjonen, kan vi ta stikkprøver; teste et lite, begrenset utvalg. Gjennomsnittsverdien til utvalget er det beste estimatoren for gjennomsnittsverdien til populasjonen. Et utvalg har en empirisk variasjon, som er den beste estimatoren for variasjonen til populasjonen.

Estimere verdi på utvalgsgjennomsnittet $\bar{X}$

Når verdien av en målestørrelse bestemmes ved å ta en serie enkeltmålinger, er beste estimat av målestørrelsen tatt for å være det aritmetiske gjennomsnittet (utvalgsgjennomsnittet) av enkeltmålingene.

Dermed, hvis målingene av målestørrelsen X, har observasjonsverdiene x_1, x_2, x_3, ..., x_n er det beste estimatet av verdien på utvalgsgjennomsnittet

$$\bar{X} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad [1]$$

Når vi regner ut utvalgsgjennomsnittet av n enkeltobservasjoner, vil hver av dem bidra med 1/n av sin verdi. Har vi en ekstrem observasjonen grunnet feilmåling, vil denne gi et mindre bidrag til utvalgsgjennomsnittet hvis vi samtidig har mange n. Hver enkelt feil vil derfor få mindre innflytelse på utvalgsgjennomsnittet når vi foretar flere målinger.

Senere i artikkelserien skal vi se på hvorledes vi kan forkaste et måleresultat ved å påstå at observasjonen er en ekstrem verdi som forekommer sjelden.

Eksempel 1: Temperaturen i et testkammer blir målt hver andre time over ett døgn.

Verdiene ble:

Kl. 00:00	Kl. 02:00	Kl. 04:00	Kl.06:00	Kl.08:00	Kl.10:00	Kl.12:00	Kl.14:00	Kl.16:00	Kl.18:00	Kl.20:00	Kl.22:00
Verdi 1	Verdi 2	Verdi 3	Verdi 4	Verdi 5	Verdi 6	Verdi 7	Verdi 8	Verdi 9	Verdi 10	Verdi 11	Verdi 12
20,05 °C	20,12°C	20,00 °C	19,81 °C	19,61 °C	19,74 C	19,91 °C	20,07 °C	20,38 °C	20,14 °C	20,09 °C	20,08 °C

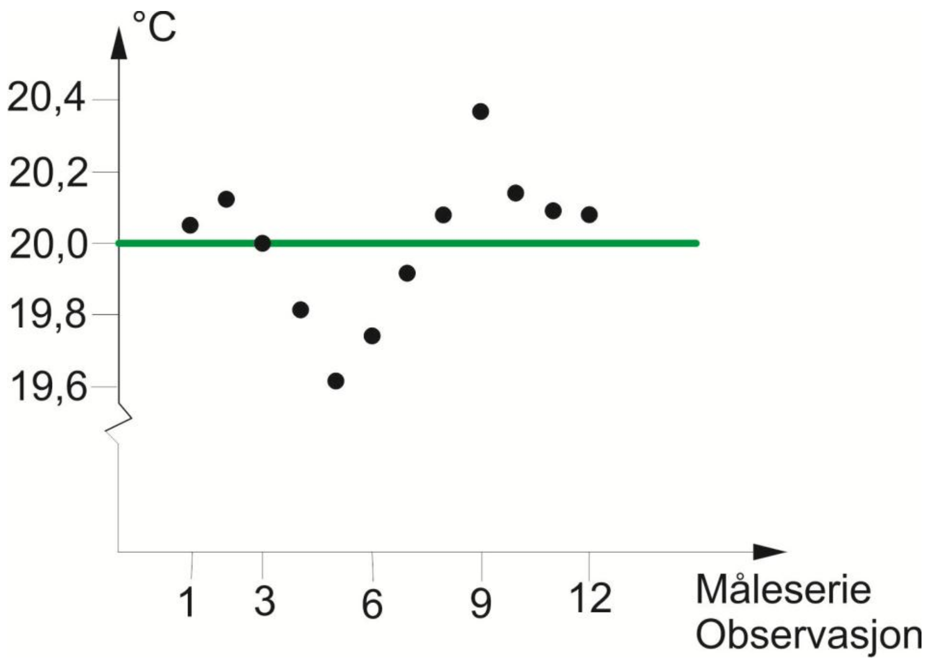
Tabell 1: Observasjoner tatt hver andre time av temperaturen i et testkammer

Det beste estimat på gjennomsnittlig temperatur (utvalgsgjennomsnittet) er hentet fra enkeltresultatene registrert som følger:

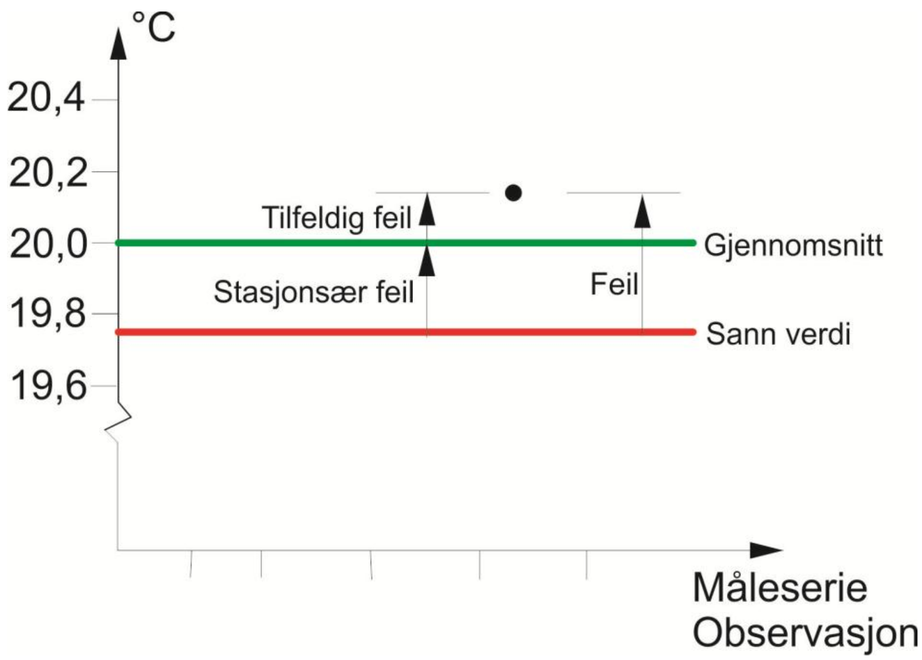
$$\bar{X} = \frac{1}{12}(20{,}05 + 20{,}12 + \ldots 20{,}09 + 20{,}08) \qquad [2]$$

$$= 20{,}0\ °C \qquad [3]$$

Vi plotter observasjonene i et diagram (figur 1).



Figur 1: 12 observasjoner (sorte prikker) av temperaturen i et testkammer, hvor utvalgsgjennomsnittet ble kalkulert til 20,0 °C (grønn heltrukken strek).



Figur 2: Definisjon av målefeil, og dennes bidrag

Hypotetisk tenker vi oss at den sanne temperaturen (rød heltrukken strek) var 19,75 °C gjennom hele testen. Tar vi utgangspunkt i observasjon nummer 10 hvor temperaturen ble målt til 20,14 °C, vil den totale målefeilen være 0,39 kelvin.

Denne målefeilen består imidlertid av to komponenter: Stasjonær feil og tilfeldig feil.

Den stasjonære feilen er differansen mellom utvalgsgjennomsnittet for måleserien og den sanneverdien. I vårt tilfelle er den stasjonære feilen 0,25 Kelvin.

Og den tilfeldige feilen er differansen mellom måleverdien og utvalgsgjennomsnittet. I vårt tilfellet er den tilfeldige feilen 0,14 Kelvin.

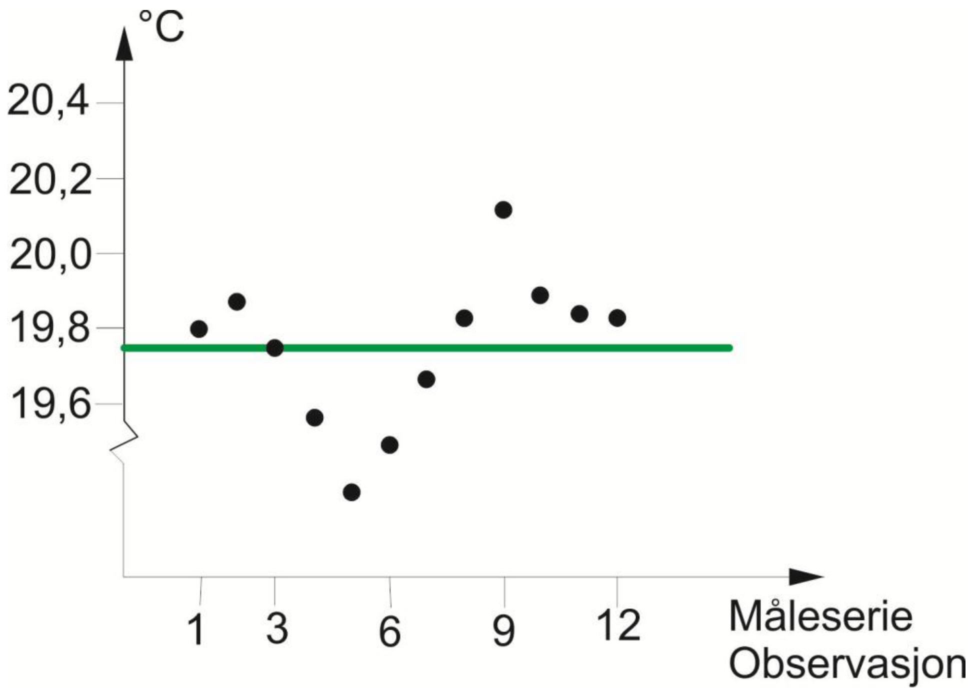
Den stasjonære feilen kan vi eliminere med kalibrering og justering.

Verdiene etter korreksjon av stasjonær feil ble:

Verdi 1	Verdi 2	Verdi 3	Verdi 4	Verdi 5	Verdi 6	Verdi 7	Verdi 8	Verdi 9	Verdi 10	Verdi 11	Verdi 12
19,80 °C	19,87 °C	19,75 °C	19,56 °C	19,36 °C	19,49 C	19,66 °C	19,82 °C	20,13 °C	19,89 °C	19,84 °C	19,83 °C

Tabell 2: Korrigerte verdier

Etter slikt arbeid, ville måleresultatserien se slik ut



Figur 3: Måleserie etter justering av stasjonær feil. Alle enkeltverdier er korrigert tilsvarende som stasjonær feil.

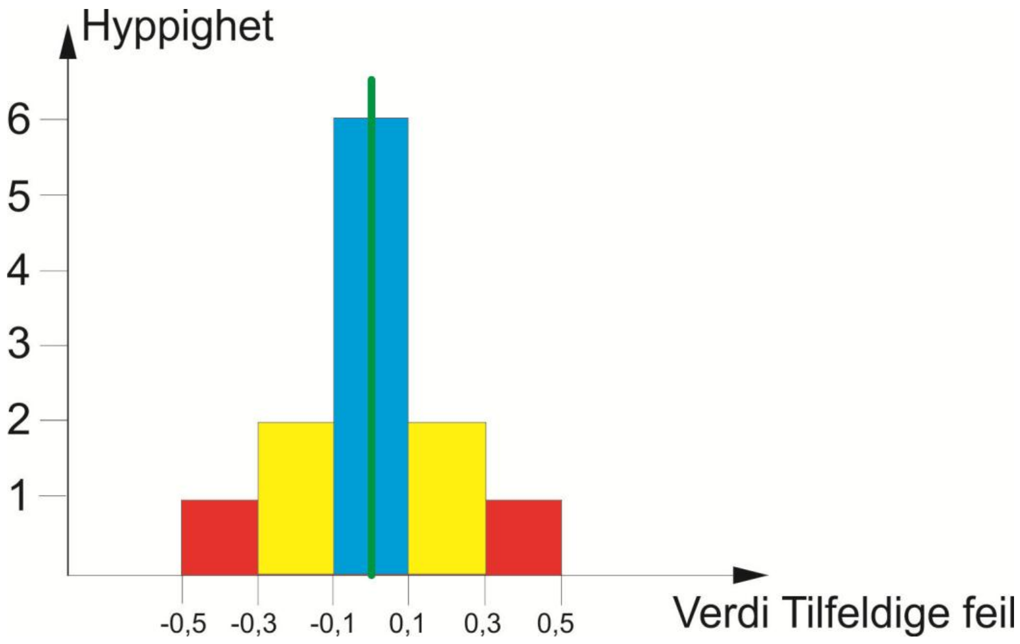
Vi tar en nærmere kikk på de tilfeldige feilene.

De tilfeldige feilene (variasjonene) ble:

Verdi 1	Verdi 2	Verdi 3	Verdi 4	Verdi 5	Verdi 6	Verdi 7	Verdi 8	Verdi 9	Verdi 10	Verdi 11	Verdi 12
0,05 K	0,12 K	0,00 K	-0,19 K	-0,39 K	-0,26 K	-0,09 K	0,07 K	0,38 K	0,14 K	0,09 K	0,08 K

Tabell 3: Tilfeldige feil i antall kelvin.

Vi tegner opp et histogram for grafisk se på hvorledes de tilfeldige feilene distribueres.



Figur 4: Histogram av tilfeldige feil fra vår måleserie. Stolpene har 0,2 Kelvin bredde.

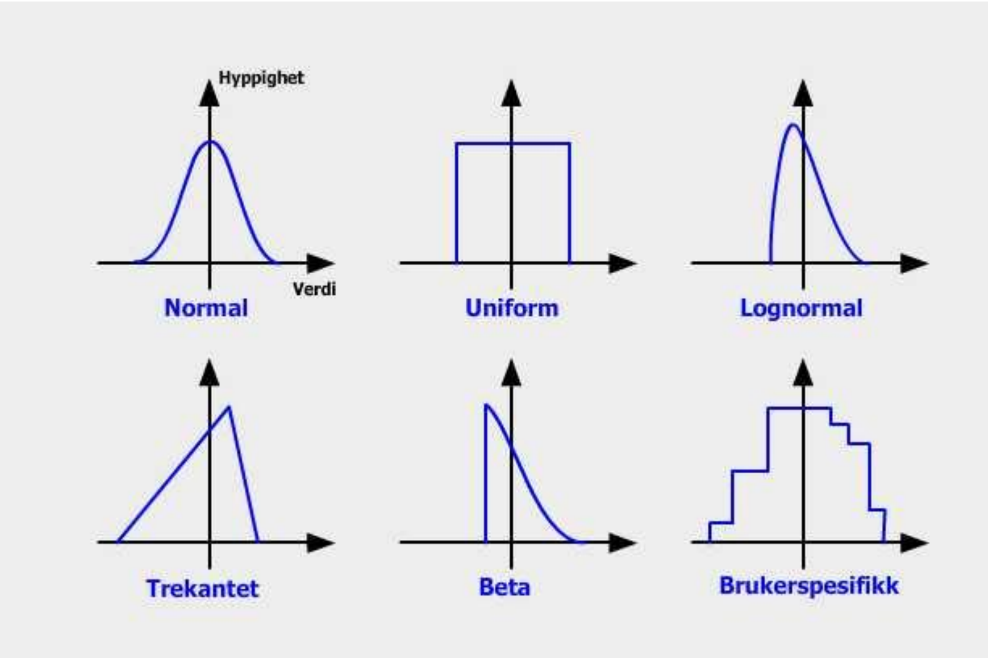
Seks av observasjonene (verdiene 0,05/0,00/-0,09/0,07/0,09 og 0,08)) lå innenfor vårt blå intervall -0,1 til 0,1.

To av observasjonene (-0,19/-0,26) lå i det gule intervallet -0,3 til -0,1. To observasjoner (0,12/0,14) lå i det gule intervallet 0,1 til 0,3.

En observasjon (-0,39 og 0,38) lå i hver av de ytterste røde intervallene.

Ulike typer distribusjon

Generelt, vi har mange ulike typer distribusjon av tilfeldige feil.



Figur 5: Ulike distribusjoner av tilfeldige feil.

Figur 5 viser eksempler på ulike former for distribusjon av tilfeldige feil. I de aller fleste tilfeller benyttes normalfordelingskurven, som vist øverst til venstre.

Normalfordelingskurven er også omtalt som Gauss-kurven.

Frekvensen (hyppigheten) i en tilfeldig målefeil vil ha følgende sammenheng:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\left(-\frac{[x-\bar{x}]^2}{2\sigma^2}\right)} \qquad [4]$$

Hvor:

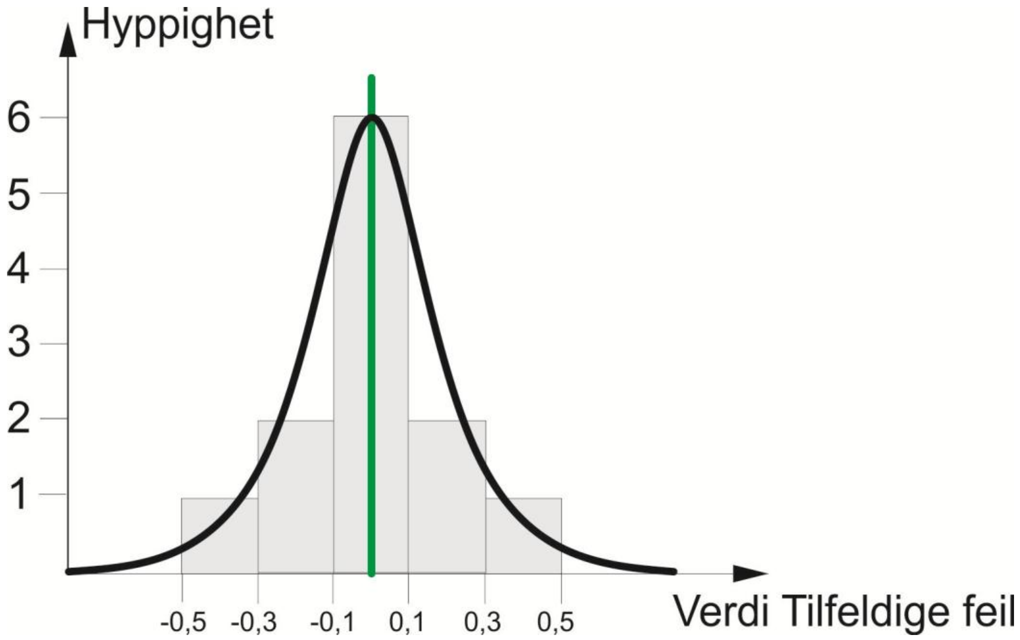
σ ett standardavvik

π 3,14 . . .

e eksponentiell funksjon

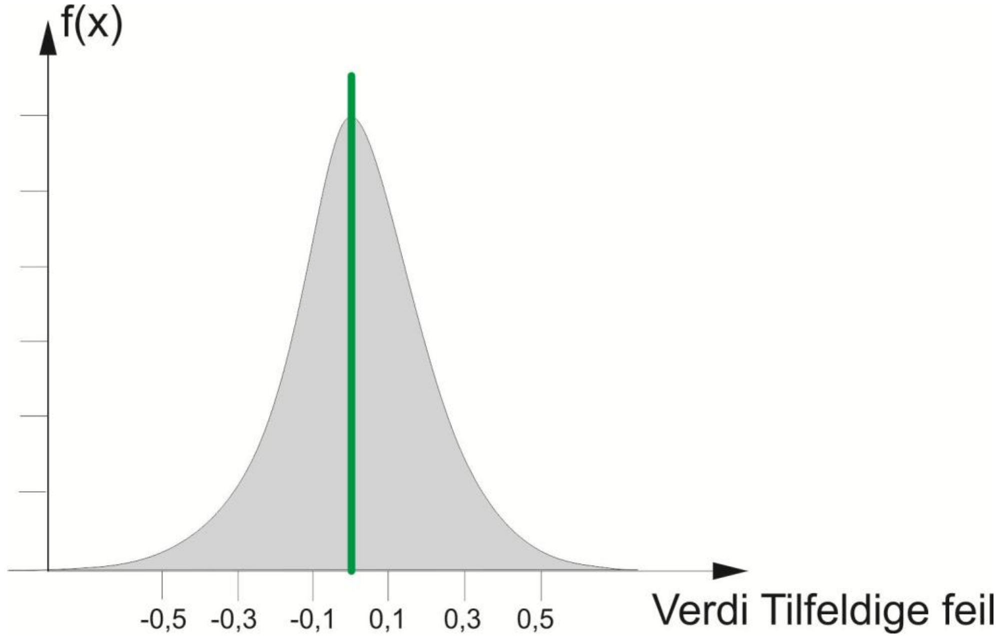
x Målestørrelse

$\bar{x}$ aritmetisk gjennomsnittsverdi



Figur 6: Ved å finne passende verdi på σ kan vi tilpasse Gauss-kurven (svart heltrukken strek) slik at den så godt som mulig beskriver histogrammet.

De grå søylene teller totalt 12 verdier, eller sagt på en annen måte; 100 prosent av antall observasjoner.



Figur 7: Arealet under Gauss-kurven er alltid lik 100 prosent, og representerer således alle observasjoner i måleserien.

Gauss-kurven er utformet slik at vi har hyppigst verdier rundt utvalgsgjennomsnittet. Desto lengre vekk fra utvalgsgjennomsnittet vi beveger oss, desto sjeldnere finner vi slike ekstreme observasjoner.

Variasjoner i observasjonene, og deres standardavvik

La oss gå tilbake til tabell hvor vi hadde loggført de tilfeldige feilene:

Verdi 1	Verdi 2	Verdi 3	Verdi 4	Verdi 5	Verdi 6	Verdi 7	Verdi 8	Verdi 9	Verdi 10	Verdi 11	Verdi 12
0,05 K	0,12 K	0,00 K	-0,19 K	-0,39 K	-0,26 K	-0,09 K	0,07 K	0,38 K	0,14 K	0,09 K	0,08 K

Tabell 4: Tilfeldige feil

Det fins ulike måter å angi hvor stor variasjonen er for måleserien. Den mest anerkjente metoden er å beregne variansen til måleserien ved å kvadrere den tilfeldige feilen, addere disse og dividere med «Degrees of Freedom».

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n-1} \qquad [5]$$

$$s^2 = \frac{(0{,}05)^2 + (0{,}12)^2 + \ldots + (0{,}08)^2}{12-1} \qquad [6]$$

$$s^2 = \frac{0{,}4642}{11} \qquad [7]$$

$$s^2 = 0{,}04\ (\mathrm{K})^2 \qquad [8]$$

Ulempen med varians er at vi ikke har samme enhet som for målestørrelsen. Dette ordner vi elegant ved å trekke kvadratroten av variansen!

Vi har dermed et nytt uttrykk vi kaller for standardavvik.

Standardavvik = $\sqrt{Varians} = \sqrt{s^2}$ [9]

For vår måleserie:

s = $\sqrt{0{,}04\ (K^2)} = 0{,}205$ K [10]

Estimere standard usikkerhet for utvalgsgjennomsnittet $\bar{X}$ (CI = Confidence Interval)

Vår første måleserie, som bestod av 12 enkeltobservasjoner, har gitt oss muligheter for å tallfeste et usikkerhetsintervall for måleserienes utvalgsgjennomsnittet.

Vi benytter oss av sentralgrenseteoremet. Mer om dette i en senere artikkel. I dag nøyer vi oss med å sette opp konklusjonen:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} \qquad [11]$$

Uttrykt med ord:

Standard usikkerhet for distribusjonen av måleserienes utvalgsgjennomsnittet er lik standardavviket til en måleserie dividert med kvadratroten av antall observasjoner i en måleserie.

$$s_{\bar{x}} = \frac{0{,}205}{\sqrt{12}}\ °\mathrm{C} = 0{,}059\ °\mathrm{C} \qquad [12]$$

Det er med andre ord ca 68 prosent sannsynlighet for at den neste måleserien har et utvalgsgjennomsnitt som ligger i intervallet

$$\bar{X} \pm s_{\bar{x}} \qquad [13]$$

<19,75 – 0,059 , 19,75 + 0,059> [14]

<19,69 , 19,81> [15]

I tilfellet et nytt utvalgsgjennomsnitt ($\bar{X}$), er det derfor cirka 1/3 sjanse (100 – 68 = 32 %) for at dette utvalgsgjennomsnittet vil ligge på utsiden av usikkerhetsintervallet definert av standardavviket. Dette er i mange måletilfeller en uakseptabel risiko.

Estimere utvidet standard usikkerhet for utvalgsgjennomsnittet $\bar{X}$

Likeledes kan det vises at i intervallet to standardavvik fra middelverdien, vil 95 prosent av alle enkeltobservasjoner ligge. Dette er kun riktig dersom antall observasjoner (n) i en måleserie er høy, det vil si fler enn 30 enkeltobservasjoner.

Har vi færre enn 30, må vi innføre en korreksjonsfaktor. Denne omtales som Student t-faktor.

For å øke troverdigheten av at den sanne verdien vil ligge i intervallet, øker vi intervallet ved å multiplisere standard usikkerhet med en faktor t. Resultatverdien kaller vi for utvidet usikkerhet, U.

For et utvalg med enkeltmålinger, har vi følgende:

$$U(\bar{x}) = t \cdot s_{\bar{x}} \qquad [16]$$

Verdien på t hentes fra Student's tabellen.

"Degrees of Freedom" (v = n – 1)	**Student's t-faktor**		
	Konfidensnivå ("Confidence Level")		
	90 %	95 %	99 %
1	6,31	12,71	63,66
2	2,92	4,30	9,92
3	2,35	3,18	5,84
4	2,13	2,78	4,60
5	2,02	2,57	4,03
6	1,94	2,45	3,71
7	1,89	2,36	3,50
8	1,86	2,31	3,36
9	1,83	2,26	3,25
10	1,81	2,23	3,17
11	1,80	2,20	3,11
12	1,78	2,18	3,05
13	1,77	2,16	3,01
14	1,76	2,14	2,98
15	1,75	2,13	2,95
20	1,72	2,09	2,85
25	1,71	2,06	2,79
60	1,67	2,00	2,70
100	1,66	1,98	2,63
∞	1,64	1,96	2,58

Tabell 5: Student t-faktor som tar hensyn til ulike størrelser på måleseriene

Vi vårt eksempel er «Degrees of Freedom» lik verdien 11 siden vi har 12 observasjoner i måleserien.

Og, skal vi ha 95 % konfidensnivå, avleser vi t-verdien 2,20 ifra tabellen.

$U(\bar{x}) = t \cdot s_{\bar{x}}$ [17]

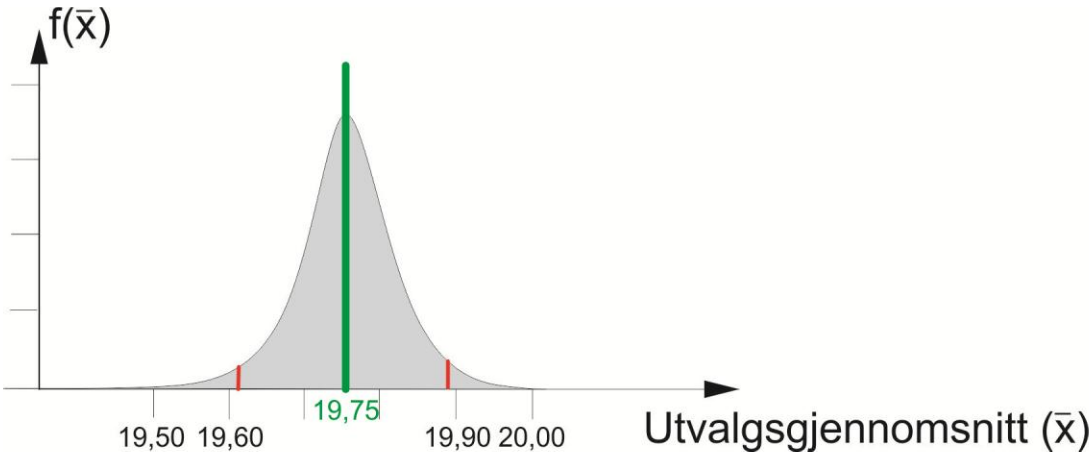
$U(\bar{x}) = 2{,}20 \cdot 0{,}0{,}59 = 0{,}13\ K$ [18]

Med 95 prosent sannsynlighet vil verdien på det neste utvalgsgjennomsnittet ligge innenfor intervallet

$\bar{X} \pm t \cdot s_{\bar{x}}$ [19]

<19,75 – 2,20·0,059 , 19,75 + 2,20·0,059 > [20]

<19,62 , 19,88> [21]



Figur 8: Distribusjon av utvalgsgjennomsnittene ($\bar{X}$), og dennes utvidede usikkerhetsintervall symbolisert med røde streker <19,62 , 19,88>.

I det utvidede intervallet fra middelverdien (verdier høyere enn 19,62 og lavere enn 19,88 slik visualisert med røde streker) vil utgjøre 95 % av tilfellene.

Hva med fremtidig enkeltobservasjon?

I mange applikasjoner ønsker vi å spå en temperaturverdi i stedet for å estimere hvor utvalgsgjennomsnittet ligger.

Vi skal estimere et intervall (PI – Prediction Interval) basert på at enkeltverdien kommer fra en normalfordelt populasjon. Intervallet skal være slik at det er 95 prosent sannsynlig at den fremtidige enkeltverdien ligger i intervallet.

I en senere artikkel vil jeg begrunne følge formel for å beregne PI

$$\bar{X} \pm \mathrm{t} \cdot s\sqrt{1+\frac{1}{n}} \qquad [22]$$

I vårt eksempel er utvalgsgjennomsnittet ($\bar{X}$) lik 19,75 °C, Student t-faktor (t) oppgitt ved 95 % konfidensnivå og «Degres of Freedom» 11 lik 2,20, standardavvik for måleserien (s) lik 0,205 K, og antall observasjoner (n) i måleserien lik 12.

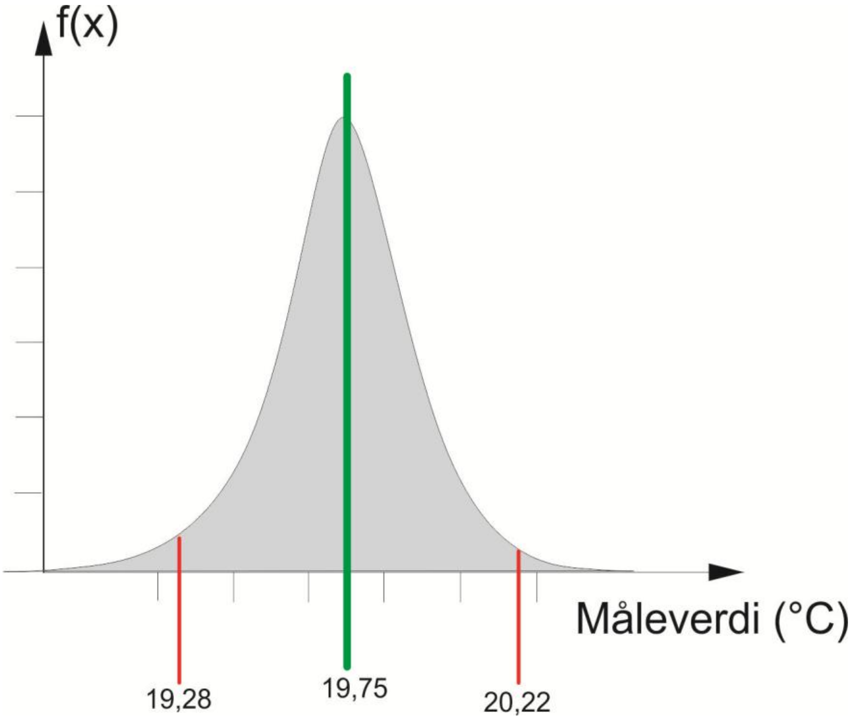
$$< \bar{X} - \mathrm{t} \cdot s\sqrt{1+\frac{1}{n}} \;,\; \bar{X} + \mathrm{t} \cdot s\sqrt{1+\frac{1}{n}} > \qquad [23]$$

$$< 19{,}75 - 2{,}20 \cdot 0{,}205\sqrt{1+\frac{1}{12}} \;,\; 19{,}75 + 2{,}20 \cdot 0{,}205\sqrt{1+\frac{1}{12}} > \qquad [24]$$

$$< 19{,}75 - 0{,}47 \;,\; 19{,}75 + 0{,}47 > \qquad [25]$$

$$< 19{,}28 \;,\; 20{,}22 > \qquad [26]$$

Med 95 prosent sannsynlighet vil en fremtidig enkeltobservasjon ligge innenfor intervallet 19,28 °C og 20,22 °C.



Figur 9: Utvidet usikkerhetsintervall for en fremtidig enkeltmåling (x_{13})

Dersom fiskal olje-/gassmålestasjon

Dersom vår temperaturmåler hadde vært benyttet på en fiskal olje- eller gassmålestasjon, ville vår temperaturmåler håndtert regelverkskravet?

La oss først se litt på kravet. I seksjon 8 i Måleforskriften fremkommer det at sløyfe-usikkerhetsgrenser er 0,30 °C. Usikkerhetsgrensen er oppgitt ved 95 % konfidensnivå.

Vår temperaturmåler hadde en usikkerhetsgrenser på 0,47 °C, noe som er bredere enn 0,30. Vår måler ville ikke klart kravet hvis vår måler skulle stå på en fiskal olje- eller gassmålestasjon.

Kravet for usikkerhetsgrenser for brensels- og fakkelgasstasjoner er 0,50 °C. Og, i dette tilfellet ville vår temperaturmåler tilfredsstille kravet.

Hva med en annen størrelse på måleserien?

Vår måleserie bestod av 12 enkeltobservasjoner. Men, hva hvis måleserien ville ha bestått av tre enkeltobservasjoner.

Vi oppsummerer dette i tabellen nedenfor:

[27]

$<\bar{X} - t \cdot s\sqrt{1+\frac{1}{n}}\ ,\ \bar{X} + t \cdot s\sqrt{1+\frac{1}{n}}>$

Utvalg	x_1, x_2, x_3	x_4, x_5, x_6	x_7, x_8, x_9	x_{10}, x_{11}, x_{12}	Snitt
Enkeltobservasjoner	19,80 19,87 19,75	19,56 19,36 19,49	19,66 19,82 20,13	19,89 19,84 19,83	
Antall elementer (n)	3	3	3	3	
Utvalgsgjennomsnitt ($\bar{X}$)	19,81	19,47	19,87	19,85	19,75
Standardavvik (s)	0,06	0,10	0,24	0,03	
Student t-faktor (t)	4,30	4,30	4,30	4,30	
Utvidet usikkerhet $U_x = t \cdot s\sqrt{1+\frac{1}{n}}$	0,30	0,50	1,19	0,16	0,54
Intervall <>	<19,51 – 20,11>	<18,97 – 19,97>	<18,68 – 21,06>	<19,69 – 20,01>	

Tabell 6: Reduserte datasett (n = 3), og estimering av usikkerhetsgrenser for en fremtidig observasjon

Fra det første datasettet (x_1, x_2 og x_3) har usikkerhetsgrense på 0,30 kelvin. Temperatursløyfen er derfor i henhold til krav i Måleforskriften for både olje-/gass-, samt brensel- og fakkelgass

	Olje-/gassmålestasjoner Usikkerhetsgrense: 0,30 K	Brensel-/fakkelgass Usikkerhetsgrense: 0,50 K
Utvalg ($x_1 \rightarrow x_{12}$): 0,47 K	Ikke OK.	OK, men tvilsomt.
Utvalg ($x_1 \rightarrow x_3$): 0,30 K	OK, men tvilsomt	OK.
Utvalg ($x_4 \rightarrow x_6$): 0,50 K	Ikke OK.	OK, men tvilsomt.
Utvalg ($x_7 \rightarrow x_9$): 1,19 K	Ikke OK.	Ikke OK.
Utvalg ($x_{10} \rightarrow x_{12}$): 0,16 K	OK.	OK, med god margin.

Tabell 7: Ulike konklusjoner på hva vi skal gjøre med x_{13} = 20,05 °C

Selv om datasettene brukt i dette eksempel er konstruerte, viser dette likevel nødvendigheten av å ha et bredt antall observasjoner i en måleserie. Har vi kun tre observasjoner i en måleserie, må vi forvente at usikkerhetsintervallene vil variere.

Hvis vår trettende (x_{13}) temperaturmåling hadde verdien 20,05 °C, ville vi trekke følgende konklusjon:

Utvalg ($x_1 \rightarrow x_{12}$): < 19,28 , 20,22>	X_{13} ligger inne i intervallet, og er derfor ikke en «uteligger».
Utvalg ($x_1 \rightarrow x_3$): <19,51 , 20,11>	X_{13} ligger inne i intervallet, og er derfor ikke en «uteligger».
Utvalg ($x_4 \rightarrow x_6$): <18,97 , 19,97>	X_{13} ligger utenfor usikkerhetsgrensene, og er derfor en «uteligger», og det må undersøkes nærmere hvorfor dette betraktes som en ekstrem verdi.
Utvalg ($x_7 \rightarrow x_9$): <18,68 , 21,06>	X_{13} ligger inne i intervallet, og er derfor ikke en «uteligger».
Utvalg ($x_{10} \rightarrow x_{12}$): <19,69 , 20,01>	X_{13} ligger utenfor usikkerhetsgrensene, og er derfor en «uteligger», og det må undersøkes nærmere hvorfor dette betraktes som en ekstrem verdi.

Tabell 8: Ulike konklusjoner på hva vi skal gjøre med x_{13} = 20,05 °C