Repeterbarhetskrav vs antall Trails

v/ Rune Øverland, Trainor Automation AS

Artikkelserie

Dette er første artikkel i en serie av fire som tar for seg repeterbarhetskrav og antall ”trials”.

Formålet med artikkelserien er å bestemme sammenhengen mellom antall Trials og repeterbarhetskravet som må stilles for dette antall Trials.

Artikkel 1:
Her definerer vi måleusikkerhet og sentralverdi. Videre kommer vi innom at selv om vi gjentar en måling flere ganger vil vi som regel få forskjellige resultater. Artikkel avsluttes med å vise deg en spredning (variasjon) basert på en normalfordelingskurve.

Artikkel 2:
Sentralgrenseteoremet vises bred omtale i denne artikkelen. Vi kommer innom primær-populasjonens fordeling, og viktigheten av gjennomsnittverdien av et antall sekundær-populasjoner. Artikkelen avsluttes med hvordan man kan beregne standard usikkerhet for en populasjon.

Artikkel 3:
Her diskuterer vi påliteligheten av standardavvik, og vi introduserer begrepet utvidet usikkerhet av et gjennomsnitt. Du lærer mer om én-side og to-sidet avvik. Vi avslutter med å vise deg Student’s t-faktor.

Artikkel 4:
Vi viser deg hvorledes vi kan estimere standardavvik når dette ikke kan beregnes. Artikkelserien avsluttes med hvorledes vi kan kalkulere kombinert usikkerhet ved et visst antall ”Trials” – oppgitt ved 95 % Confidence Level” – til å være 0,03 % eller bedre.

Definisjon av usikkerhet

Et måleresultat består både av måleverdi og et intervall rundt denne der den sanne verdien med stor sannsynlighet ligger.

Estimert sann verdi = Måleverdi ± tilfeldige feil

Ut fra et statistisk synspunkt, vil tilfeldige feil være like mye på den positive siden som den negative.

Selv om vi har innrettet oss slik at verken systematiske, grove eller ensidige virkende feil kan få nevneverdige innvirkning på resultatet, vil vi som regel få forskjellige resultater når vi gjentar målingen flere ganger.

Måles en prøve gjentatte ganger oppnår vi derfor ikke en serie med identiske resultater, men de ligger mer eller mindre spredd innenfor et bestemt område. Denne spredningen skyldes tilfeldige feil, det vil si at tilfeldige feil påvirker presisjonen i målingen. Årsakene til at gjentatte målinger gir forskjellige resultat kan være mange.

Feilverdien varierer helt tilfeldig, og vi kan ikke på forhånd si i hvilken retning (pluss/minus) den vil forskyve resultatet. For å få et bilde av størrelsen av tilfeldige feil må vi ta i bruk statistiske metoder.

Sentralverdi

Sentralverdi kan i dagligtale forstås som en gjennomsnittsverdi. Vi skal bestemme tyngdepunktet i gruppe med tall. I statistikk har vi ofte tre ulike definisjoner på en sentralverdi:

- Aritmetisk middelverdi (som oftest denne vi bruker i det daglige)
- Median (det tall som ligger i midten i en tallserie)
- Mode (det tall som forekommer hyppisk i en tallserie)

La oss tenke oss tallsettet:

7	3	3	5	2	100

Her er det seks tall, så det er overkommelig å kommunisere disse tallene. Men, hadde vi hatt seks millioner tall i dette tallsettet, ville det straks vært mer uoversiktlig. Vi ønsker derfor å finne et nytt tall som representerer en sentralverdi (tyngdepunkt) for tallsettet.

Aritmetisk middelverdi

Her adderer vi individuelle tallverdier og deler dette med antall tall.

Sentralverdi = Aritmetisk middelverdi = $\frac{7+3+3+5+2+100}{6} = \frac{120}{6} = 20$

Hvert av de seks tallene har lik vekting på sentralverdien, det vil si 1/6 hver. Tallet 20 er dette tilfellet sentralverdien for tallsettet.

Ulempen med denne metoden for å angi sentralverdien er at unormale verdier, i dette tilfellet tallet 100, raskt trekker opp sentralverdien.

Median

Her ordner vi først tallene i stigende rekkefølge, og finner det tallet som er i midten av tallserien.

Ordnet tallserie:

2	3	**3**	**5**	7	100

I dette tilfellet har vi ikke ett tall som ligger i midten, men to. Pr definisjon tar vi derfor den aritmetiske middelverdien av disse:

Altså:

$$\text{Sentralverdi} = \text{Median} = \frac{3+5}{2} = \frac{8}{2} = 4 \qquad [1.1]$$

Tallet 4 er i dette tilfellet sentralverdien for tallsettet.

Det fine med denne metoden er at unormale tall, i dette tilfellet 100, ikke ”ødelegger” den folkelige oppfatningen av hva som er sentralverdien for tallsettet.

Mode

Denne definisjonen sier at sentralverdien er det tall som forekommer hyppigst (har høyest frekvens) i tallsettet.

Altså:

Tall	Antall (frekvens)
2	1
3	**2**
5	1
7	1
100	1

Tallet 3 er i dette tilfellet sentralverdien for tallsettet.

Hvilken metode er best for å bestemme tyngdepunktet?

Nå som vi har vært inne på tre ulike definisjoner av sentralverdier, kan vi spørre oss:

Hvilken av de tre sentralverdiene i eksemplet over er best?

20, 4 eller 3?

Vel, i utgangspunktet er alle like riktige. Men, det kan være nyanser i hva som er ”best”. Basert på skjønn vil sikkert noen si at tallet 100 i vårt tallsett ”ødelegger” dersom vi bruker definisjonen aritmetisk middelverdi. Hadde vi utelatt dette tallet fra tallsettet, ville den aritmetiske middelverdien vært:

$$\text{Modifisert Aritmetisk middelverdi} = \frac{7+3+3+5+2}{5} = \frac{20}{5} = 4$$

Dersom vi har mistanke om at et tall ikke er representativ for tallsettet, det vil si at tallet kan ha oppstått grunnet en grov feil, kaller vi dette tallet for en ”uteligger”.

Det er flere metoder for å fjerne ”uteligggere”.

Variasjon rundt sentralverdien (tyngdepunktet)

La oss vurdere tre tallsett:

Tallsett	Utvalg
Nr. 1	2, 3, 3, 5, 7
Nr. 2	3, 4, 4, 5, 4
Nr. 3	4, 4, 4, 4, 4

For de tre tallsettene er den aritmetiske gjennomsnittsverdien (sentralverdien = $\bar{x}$) for hvert tallsett verdien 4.

Og, vi ser lett at utvalgene likevel er ganske forskjellige. Tallsett 1 har en nokså stor variasjon i enkeltverdiene rundt sentralverdien, mens tallsett 3 ikke har noen variasjon i det hele tatt.

I tillegg til sentralverdien vil det ofte være ønskelig at en angir en størrelse som forteller hvordan måleresultatene er spredt i måleområdet.

Metoder som kan brukes for å angi spredningen rundt sentralverdien er:

- Grenseverdi
  $$G = \frac{x_{maksimum} - x_{minimum}}{2}$$
- Midlere avvik:
  $$M = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \ \dots\ |x_n - \bar{x}|}{n}$$
- Varians:
  $$Var = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ \dots\ + (x_n - \bar{x})^2}{n}$$

  eller

  standardavvik:

  $$s = \sqrt{Var^2} = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ \dots\ + (x_n - \bar{x})^2}{n}}$$

Jo nærmere datapunktene ligger middelverdien, jo bedre vil en si datasettet er.

Nedenfor oppsummerer vi grenseverdiene for spredningen som ligger rundt tallsettets sentralverdi:

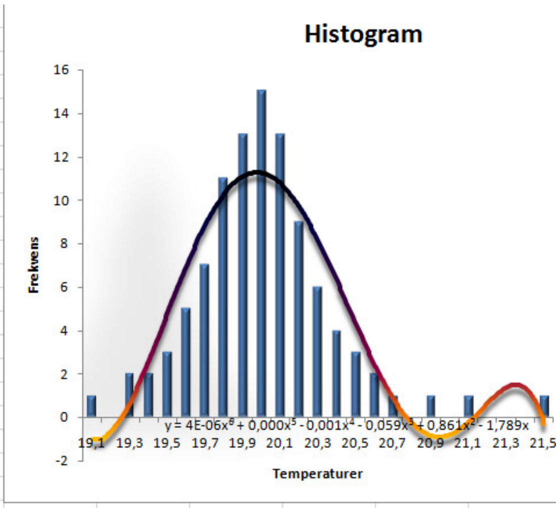
Tallsett	Utvalg	Grenseverdi	Midlere avvik	Varians	Standardavvik
Nr. 1	2, 3, 3, 5, 7	± 2,5	± 1,6	± 3,2	± 1,79
Nr. 2	3, 4, 4, 5, 4	± 1,0	± 0,4	± 0,4	± 0,63
Nr. 3	4, 4, 4, 4, 4	± 0,0	± 0,0	± 0,0	± 0,00

Som nevnt, har vi et tallsett som ovenfor er tallsettet oversiktlig. Og hadde vi et tallsett med mange millioner ”medlemmer”, ville det vært en grei måte å kommunisere dette tallsettet med å oppgi dennes sentralverdi og tallsettets spredningsgrenser.

Histogram

I et histogram tegner vi inn ”stolper” med en viss bredde.; for eksempel for hvert tiendedels grad Celsius, og en høyde. Diagrammet viser sammenhengen mellom måleverdi(er) i en gruppe, for eksempel temperaturer mellom 19,95 °C og 20,04°, og hvor ofte dette måleresultatet forekommer. Ved hjelp av dataverktøy kan man få tegnet opp en kurve som illustrerer en sammenhengen.

1	*Temperatur*	*Frekvens*
2	19,1	1
3	19,2	0
4	19,3	2
5	19,4	2
6	19,5	3
7	19,6	5
8	19,7	7
9	19,8	11
10	19,9	13
11	20,0	15
12	20,1	13
13	20,2	9
14	20,3	6
15	20,4	4
16	20,5	3
17	20,6	2
18	20,7	1
19	20,8	0
20	20,9	1
21	21,0	0
22	21,1	1
23	21,2	0
24	21,3	0
25	21,4	0
26	21,5	1



Figur: Temperaturen er målt 100 ganger, og det er tegnet et ”stolpe”-diagram (histogram). Programmet tegner opp en ”utjevnet” kurve.

Normalfordeling av spredning

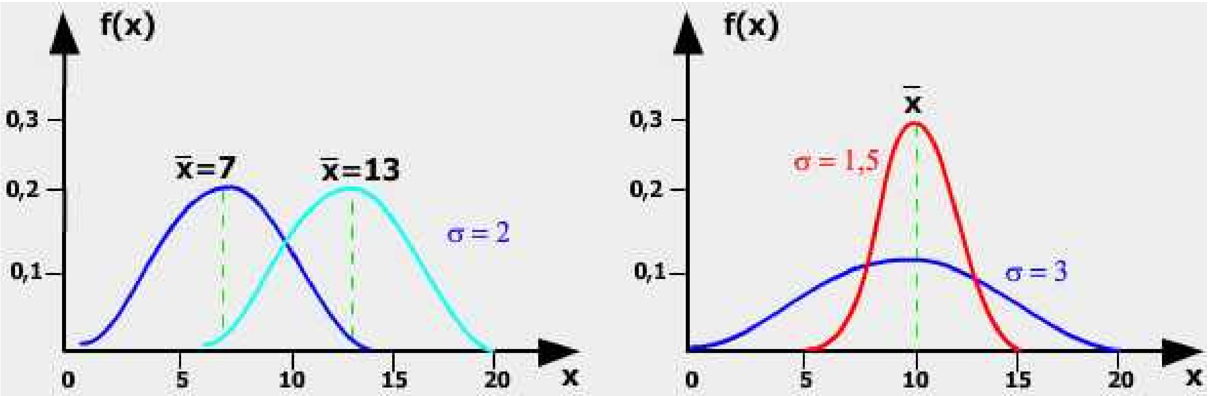
Årsaken til tilfeldig variasjon av tallverdier rundt tallsettets sentralverdi kan være mange. For de aller fleste tilfeller gjelder det at de fleste variasjoner er små, og tallverdien vil som oftest ligge nærmere sentralverdien. Desto større feil, desto sjeldnere vil denne opptre.

Sannsynlighetsfordelingen er slik at vi forventer å finne flest resultater ved sentralverdien, og vi har en lavere forventning å finne tallverdier med større avvik. Sammenhengen kan beskrives slik:

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

hvor $\overline{x}$ er aritmetisk gjennomsnittsverdi.

Walter Shewhart arbeidet i 1920-årene i selskapet Bell Telephone Laboratories. Han omtales ofte som den moderne statistikkens far etter at i 1926 publiserte en av sine mange artikler om normalfordelinger. Normalfordelingskurven blir også omtalt som Bell-kurven.



I figuren, venstre halvdel, har vi to kurver med samme σ -verdi (spredningskarakteristikk) og ulike $\bar{x}$-verdier (sentralverdier; her 7 og 13).

I figuren; høyre halvdel; har vi to kurver med samme sentralverdi (her 10) og ulike σ -verdier (spredningskarakteristikker).

Det kan vises at ut i fra en normalfordelingskurve er det cirka 68 % sannsynlighet for et en enkeltverdi i et utvalg vil ligge innenfor intervallet pluss/minus ett standardavvik (s) av sentralverdien ($\bar{x} \pm s$).

Mange industri-normer bruker et krav på to ganger standardavvik som grenseverdier. Det vil her være cirka 95 prosent sannlighet for at en måleverdi vil ligge innenfor våre grenseverdier ($\bar{x} \pm 2s$).

Og likeledes, det vil være fem prosent sannsynlighet for at en måleverdi ligger utenfor grenseverdiene. I dette tilfellet kan måleresultatet være beheftet med så store feil (men også andre feiltyper enn bare tilfeldige feil) at vi forkaster måleresultatet. I all fall bør vi undersøke hvorfor vårt måleresultat ligger såpass langt unna sentralverdien.

Ulike estimater av variansen (Bessels's korreksjon)

Det er to regler for beregning av variansen basert på om sentralverdien er estimert eller om den er sann.

Estimert sentralverdi	Sann sentralverdi
$\frac{1}{n-1}$	$\frac{1}{n}$

$\frac{1}{n-1}$ brukes når s (standardavvik) baseres kun på et *estimat* av den sanne sentralverdien som tar utgangspunkt på en serie med enkeltverdier som innehar tilfeldige feil.

$\frac{1}{n}$ brukes dersom vi kjenner den sanne sentralverdien.

Korreksjonen (n-1 i stedet for n) kalles for Bessel's korreksjon.

Når man har flere målte enkelverdier i en større populasjon, bruker vi derfor $\frac{1}{n-1}$.

Formel for beregning av standardavvik	Brukes. . . :
$s = \sqrt{s^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}$	når vi skal bestemme standardavvik (σ) for hele populasjon, bruker vi ($\frac{1}{n}$).
$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}$	når vi skal bestemme standardavvik (s) for deler av en populasjon, det vil si at vi jobber med et utvalg av populasjonen, bruker vi ($\frac{1}{n-1}$).