

# SIGNIFIKANS, STATISTIKK OG EKSPERIMENTELLE DATA.

## TIPS OG TRIKS

October 26, 2017

## Contents

<b>1</b>	<b>Kort om teksten</b>	<b>2</b>
<b>2</b>	<b>Oversikt</b>	<b>3</b>
<b>3</b>	<b>Desimaler og signifikante siffer</b>	<b>4</b>
3.1	Antall desimaler . . . . .	4
3.2	Summasjon av reelle tall . . . . .	4
3.3	Antall signifikante siffer . . . . .	4
3.4	Noen unntak ved telling av signifikante siffer . . . . .	5
3.5	Multiplikasjon og divisjon med signifikante siffer . . . . .	5
<b>4</b>	<b>Statistikk</b>	<b>6</b>
4.1	Noen begreper . . . . .	6
4.2	Gjennomsnitt . . . . .	6
4.3	Standardfeil . . . . .	7
4.4	Standardavvik . . . . .	8
<b>5</b>	<b>Behandling av eksperimentelle data</b>	<b>9</b>
5.1	Motivasjon: enkel feilforplantning . . . . .	9
5.2	Signifikans og målepresisjon . . . . .	10
5.3	Noen konvensjoner . . . . .	10
<b>6</b>	<b>Mål</b>	<b>10</b>

# 1 Kort om teksten

Dette dokumentet er ment å gi en kortfattet og superenkel guide til bruk av:

- Signifikans (*antall signifikante siffer, antall desimaler, regneoperasjoner med korrekt antall signifikante siffer og desimaler*).
- Eksperimentelle data (*resultater fra laboratorieforsøk, undersøkelser eller observasjoner*).
- Statistikk (*matematiske regler for å behandle, tolke og presentere data*).

Hensikten er at du på noen timer skal kunne lære deg alt du trenger for å mestre disse verktøyene. Det krever selvfølgelig litt innsats fra din side også. Hver seksjon er på noen få linjer, og jeg har forsøkt så godt som mulig å sikre at disse få linjene inneholder alt du trenger å vite om saken (for nå, i det minste). Det som er din oppgave er å *forstå* hva som står der. Det innebærer å *lese* teksten, men også å *tenke* over det du leser. Om noe virker uklart, rart eller uforståelig kan du lene deg tilbake, se litt opp i taket og tenke nøye gjennom over hva du nettopp har lest. Kanskje klarer det, kanskje ikke. I verste fall er det jeg som har skrevet noe feil, og i så fall håper jeg du kan sende meg en e-post på [audunsh4@gmail.com](mailto:audunsh4@gmail.com), så skal jeg fikse det så fort som mulig.

Å studere vitenskap er litt som å drive med mental kampsport. Jeg er helt sikker på at du vil få dette til, men ta i mot et råd fra en 900 år gammel Jedi-mester: *"Do, or do not. There is no try."* Lykke til! :)

Med vennlig hilsen Audun

## 2 Oversikt

Her er en oversikt over hva du bør kunne etter å ha lest dokumentet:

- Signifikans. Du bør kunne
  - Telle antall signifikante siffer og antall desimaler.
  - Utføre aritmetiske operasjoner som multiplikasjon, divisjon, addisjon og subtraksjon med riktig antall gjeldende siffer og desimaler.
- Statistikk. Du bør
  - Kunne beregne gjennomsnitt, standardavvik og standardfeil.
  - Kjenne til begrepet *populasjonens gjennomsnitt* og kunne skille mellom dette og *utvalgets gjennomsnitt*.
  - Kjenne til forskjellen mellom standardfeil og standardavvik, og kunne redegjøre for når det er hensiktsmessig å bruke det ene eller andre.
- Behandling av eksperimentelle data. Du bør
  - Kunne benytte reglene for signifikans for å unngå falsk presisjon i gjengivelse av eksperimentelle data.
  - Kunne redegjøre for sammenhengen mellom standardavvik, standardfeil og måleusikkerhet.
  - Kunne redegjøre for begrepene presisjon og nøyaktighet, og relatere disse til statistiske størrelser.
  - Kjenne til enkel feilforplantning.
  - Kjenne til noen vanlige konvensjoner for gjengivelse av eksperimentelle data.

### 3 Desimaler og signifikante siffer

Læringsressurs: Kahn Academy - Significant Figures (video, lenke)

Du bør kunne

- Telle antall signifikante siffer og antall desimaler.
- Utføre aritmetiske operasjoner som multiplikasjon, divisjon, addisjon og subtraksjon med riktig antall gjeldende siffer og desimaler.

For å få til dette må du først lære deg å telle desimaler og signifikante siffer, og deretter lære deg de to enkle reglene for aritmetiske operasjoner på tallene.

#### 3.1 Antall desimaler

Du finner antallet desimaler i et tall ved å **telle alle siffer etter komma**. Her er noen eksempler:

- 1 har ingen desimaler.
- 1.0 har 1 desimal.
- 1.0000 har 4 desimaler.

#### 3.2 Summasjon av reelle tall

Når du **summerer** reelle tall skal resultatet rundes av til like mange desimaler som **leddet med færrest** desimaler. Her er noen eksempler:

- $1 + 1.0 = 2.$
- $1.23 + (-1.1) = 0.1.$
- $1.000 + 1.0 = 2.0.$

#### 3.3 Antall signifikante siffer

Du finner antallet signifikante siffer ved å telle fra venstre mot høyre alle oppgitte siffer fra og med første siffer ulikt 0. Her er noen eksempler (pass på at du er enig i dem alle):

- **1** har 1 signifikant siffer.
- **1.0** har 2 signifikante siffer.
- **0.12345** har 5 signifikante siffer.
- **0.00012345** har 5 signifikante siffer.
- **1.02** har 3 signifikante siffer.
- **$2.0 \cdot 10^{-2}$**  har 2 signifikante siffer.

Ikke bli forvirret av nullene. Alle disse måtene å skrive tallet på har 5 signifikante siffer:

$$2.0000 \cdot 10^{-2} = 0.020000 = \mathbf{20000} \cdot 10^{-6} \quad (1)$$

### 3.4 Noen unntak ved telling av signifikante siffer

Det finnes noen få unntak når du teller signifikante siffer. Disse er:

- Enkelte størrelser, slik som heltall (1,2,3...) og noen konstanter ( $\pi$ ,  $e$ ), kan antas å ha uendelig mange signifikante siffer, slik at de ikke påvirker endelig resultat. Et eksempel er konverteringen fra Torr til atomsfære i lab 2. Et annet er når man beregner volum av en sylinder basert på måling av diameter og høyde.

### 3.5 Multiplikasjon og divisjon med signifikante siffer

Når du **multipliserer (eller dividerer)** reelle tall skal resultatet rundes av til samme antall signifikante siffer som **faktoren**<sup>1</sup> **med færrest** signifikante siffer. Her er noen eksempler:

$$\frac{1.23}{2.0} = 1.23 \cdot 0.50 = 0.62 \quad (2)$$

$$\frac{1.234}{2.0} = 0.62 \quad (3)$$

$$\frac{1.23}{2.00} = 0.615 \quad (4)$$

$$\frac{1.23}{2.000} = 0.615 \quad (5)$$

---

<sup>1</sup>eller dividend, divisor.

## 4 Statistikk

Læringsressurs: Kahn Academy - Sampling distribution of a sample mean (video, lenke)

Etter å ha gjennomgått denne seksjonen bør du:

- Kunne beregne gjennomsnitt, standardavvik og standardfeil.
- Kjenne til begrepet *populasjonens gjennomsnitt* og kunne skille mellom dette og *utvalgets gjennomsnitt*.
- Kjenne til forskjellen mellom standardfeil og standardavvik, og kunne redegjøre for når det er hensiktsmessig å bruke det ene eller andre.

### 4.1 Noen begreper

En *populasjon* er for vårt anliggende gitt ved alle mulige utfall av en måling. Som regel kan vi ikke gjøre alle mulige målinger, men må heller ta til takke med et begrenset antall målinger som vi antar er representative for populasjonen. En slik begrenset mengde målepunkter kalles et *utvalg*. Fra et slikt utvalg kan vi beregne en rekke størrelser<sup>2</sup> slik som *gjennomsnitt*, *standarfeil* og *standardavvik*, og ønsker å benytte disse for å si noe om populasjonen som helhet.

En kortfattet oppsummering av nye ord:

- *Populasjon* - mengden av alle mulige målinger.
- *Utvalg* - en begrenset mengde målinger tatt fra populasjonen.
- *Populasjonens gjennomsnitt* - Det "sanne" gjennomsnittet, beregnet fra alle mulige målinger.
- *Utvalgets gjennomsnitt* - Gjennomsnittet beregnet på mengden.
- *Parameter* - Størrelser beregnet fra populasjonen.
- *Statistic* - Størrelser beregnet fra utvalget.

### 4.2 Gjennomsnitt

Gjennomsnittet  $\bar{x}$  i et utvalg  $x_i$  er gitt ved

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{1}{N} (x_1 + x_2 + \dots + x_N), \quad (6)$$

hvor  $N$  er antall målinger<sup>3</sup>.

Vi skiller mellom populasjonens gjennomsnitt og utvalgets gjennomsnitt, da det som regel er det siste vi måler.

<sup>2</sup>I statistikk kalles disse størrelsene for "statistics", for å skille de fra "parameters" som er de tilsvarende størrelsene for populasjonen.

<sup>3</sup>samples

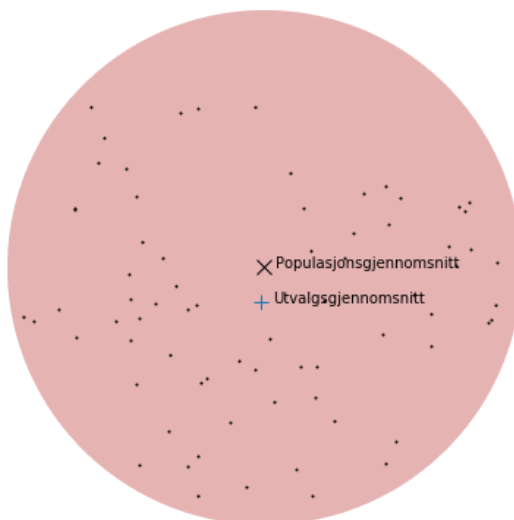


Figure 1: Populasjon angitt med rødt, utvalget representert ved sorte punkter. Populasjonens gjennomsnitt  $\bar{\mu}$  er gjennomsnittet av alle mulige målinger. Utvalgets gjennomsnitt  $\bar{x}$  beregnes derimot fra et begrenset utvalg fra populasjonen. Vi antar at det er mulig å nærme seg populasjonens gjennomsnitt ved å øke antallet målinger. Avstanden mellom  $\bar{\mu}$  og  $\bar{x}$  antas å være relatert til utvalgets standardfeil. Presisjonen i målingene relateres derimot ofte til målingenes standardavvik.

### 4.3 Standardfeil

Standardfeilen  $SE$  regnes ut ved

$$SE = \frac{1}{N} \sqrt{\sum_{i=1}^N (x_i - \bar{x})^2}. \quad (7)$$

Det er mulig å vise<sup>4</sup> at standardfeilen er relatert til standardavviket i utvalgets gjennomsnitt, med andre ord at avstanden mellom populasjonen og utvalgets gjennomsnitt er relatert til standardfeilen, som vist i figur 1.

<sup>4</sup>For en mer utfyllende forklaring se læringsressurs gitt innledningsvis i seksjonen.

## 4.4 Standardavvik

Standardavviket i et utvalg kan regnes ut fra standardfeilen ved

$$\sigma = \sqrt{N} \cdot SE = \frac{1}{\sqrt{N}} \sqrt{\sum_{i=1}^N (x_i - \bar{x})^2}. \quad (8)$$

Hvor  $N$  er antall målinger. Det er vanlig å regne ut standardavviket direkte fra det siste uttrykket over, og deretter standardfeilen ved å dele på  $\sqrt{N}$ .<sup>5</sup>

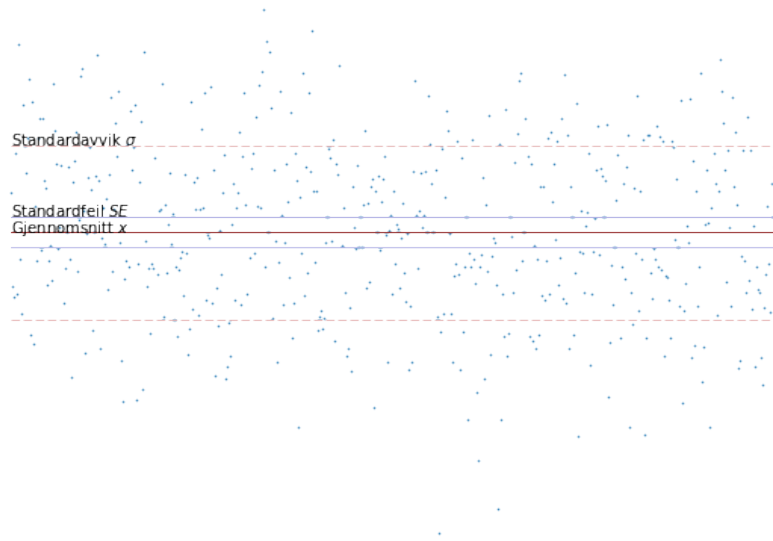


Figure 2: Standardavvik og standardfeil for et datasett.

---

<sup>5</sup>Legg gjerne merke til at vi her har brukt faktoren  $\frac{1}{\sqrt{N}}$ . Dette er for å holde ligningene rene og pene. Når standardavviket defineres med faktoren  $\frac{1}{\sqrt{N-1}}$  kalles det for "Bessel's korreksjon", og hensikten er å ikke underestimere standardavviket.



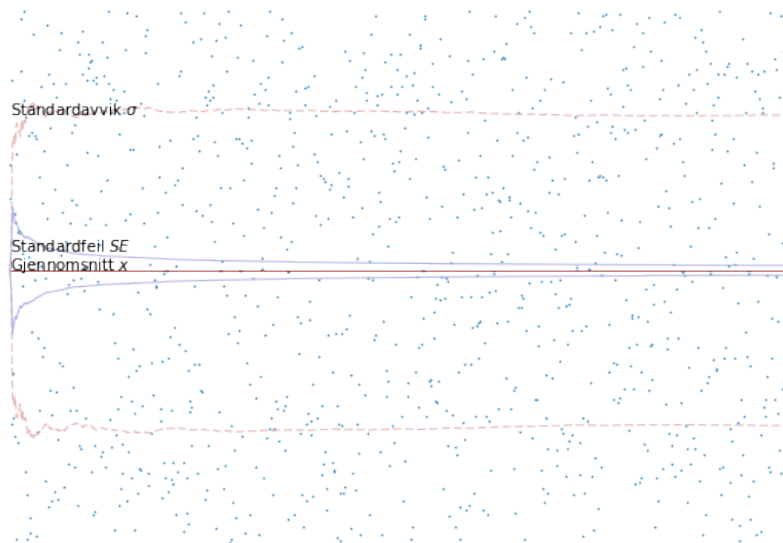


Figure 3: Standardavvik og standardfeil som funksjon av antall målinger (øker mot høyre).

## 5 Behandling av eksperimentelle data

Etter å ha lest denne seksjonen bør du:

- Kjenne til enkel feilforplantning.
- Kunne benytte reglene for signifikans for å unngå falsk presisjon i gjengivelse av eksperimentelle data.
- Kunne redegjøre for sammenhengen mellom standardavvik, standardfeil og måleusikkerhet.
- Kunne redegjøre for begrepene presisjon og nøyaktighet, og relatere disse til statistiske størrelser.
- Kjenne til noen vanlige konvensjoner for gjengivelse av eksperimentelle data.

### 5.1 Motivasjon: enkel feilforplantning

De fleste eksperimentelle målinger vil ha en feilmargin, hvilket betyr at utfallet av målingen kan skrives på formen

$$x_0 \pm \Delta x. \quad (9)$$

Her er  $x_0$  den målte verdien og  $\Delta x$  er oppgitt eller beregnet for måleinstrumentet som benyttes. Når man gjør aritmetiske operasjoner på slike størrelser er det formelt korrekt å ta med feil-estimatet i operasjonene. Dette kalles for å la feilen

”forplante seg” eller ”propagere”. For eksempel vil produktet av en måling og en eksakt kjent størrelse  $a$  kunne skrives som

$$a(x_0 \pm \Delta x) = ax_0 \pm |a|\Delta x. \quad (10)$$

Som det fremgår av uttrykket ovenfor blir feilen skalert mot størrelsen  $a$ . Om eksempelvis  $a = 10$  vil også feilen øke med en størrelsesorden. Om vi ikke hadde latt feilen forplante seg kunne vi risikert å fått det som kalles *falsk presisjon*, at vårt endelige svar har lavere feilestimat enn vi har belegg for.

Reglene for hvordan feilen propagerer er ganske omfattende og ikke noe vi bekymrer oss for i denne teksten, men nå som du forstår hvordan feil forplanter seg, spirer og gror gjennom utregningene dine, skal vi se at du kan bruke dine evner i signifikansaritmetikk til å unngå falsk presisjon.

## 5.2 Signifikans og målepresisjon

En annen måte å unngå falsk presisjon er altså å benytte reglene for signifikans på de størrelsene som inngår i beregningen. Fremgangsmåten sikrer at det endelige resultatet ikke blir angitt med siffer som det ikke er belegg for.

Om vi igjen ser på tilfellet fra (9), så er det mulig å unngå falsk presisjon ved å kun oppgi  $x_0$  med signifikante siffer, samt runde av til like mange signifikante siffer i svaret. To punkter det er viktig å merke seg:

- Standardavvik gir et mål på spredning av dataene omkring utvalgets gjennomsnitt, og sier dermed noe om presisjonen i målingene.
- Standardfeil er et mål på graden av nøyaktighet (dersom vi kan se bort fra systematiske feil) i målingen. Nøyaktigheten blir bedre med økende antall målinger  $N$ .

## 5.3 Noen konvensjoner

- I det endelige svaret bør du runde av standardfeilen til ett signifikant siffer, og gjennomsnittet til like mange desimaler som den avrundede standardfeilen.
- I tabeller er det derimot fint om svarene oppgis i henhold til de aritmetiske reglene.
- Avrundingen av signifikante siffer og desimaler skal gjøres i det endelige svaret, ikke underveis i beregningen.

## 6 Mål

Om du har lest notatet fra begynnelse til slutt og forstått alt som har blitt presentert, så har du nå kommet i mål. I såfall; Du har klart det! Gratulerer! Husk likevel på at det er mange detaljer som skjuler seg under overflaten her, og at det ikke skader å repetere. Ta gjerne en titt på oversikten innledningsvis for å forsikre deg om at du har fått med deg alt.