ISTx1002 Usikkerhet og støy i målinger

Grunnlager om usikkerhet og støy i målinger

Stefanie Muff, Institutt for matematiske fag, NTNU Trondheim

Oktober 23 og 24, 2023

Hva dreier seg denne modulen om?

- Kan vi stole på resultatet av en måling?
- Hvor stor er usikkerheten eller feilen i en måling?
- Kan statistikk hjelpe oss til å handtere det?

Tema 1: Grunnlager om usikkerhet og støy i målinger → 2 timer

Tema 2: Feilforplantning i målesysemer, kalibrering → 2 timer

Tema 3: Usikkerhet og feil i variabler i lineær regresjon → 2 timer

Læringsmål (av modulen)

Etter du har gjennomført denne modulen skal du:

- ha forståelse for usikkerhet og feil i målinger av enkle variabler av interesse, og hvordan man kvantifiserer og rapporterer denne usikkerheten på en meningsfull måte.
- kjenne til de mest grunnleggende feilfordelingene og hvordan man beregner deres varianser.
- vite hvordan man håndterer og beregner usikkerhet for transformerte variabler.
- kjenne til hvordan man håndterer og beregner usikkerhet ved lineære og ikke-lineære kombinasjoner av feilmålte variabler (feilpropagering).
- være i stand til å utføre de nødvendige beregningene.
- forstå grunnlager om kalibrering og kunne gjennomføre en enkel kalibrering selv.
- forstå effekten av usikkerhet/målefeil i kovariater i lineære regresjonsmodeller.
- være i stand til å korrigere parameterestimatene i lineær regresjon når kovariater med målefeil brukes.

Organisasjon

- Forelesninger på zoom mandag 15.15-16.00 og tirsdag 14.15-16.00 (første to uker) → 6 timer totalt, 2 for hvert tema
- Temavideoer
- Øvingstimer – se på kurssiden
- Gruppearbeid (3-6)
- Innleveringsfrist prosjektoppgave: **mandag 20. november 2023 kl 12.00**

Den eksterne kurssiden:

https://wiki.math.ntnu.no/istx1002/2023h/start

Prosjektoppgaven

- Vi ser hvor informasjonen ligger på Blackboard og på den eksterne nettsiden:

https://wiki.math.ntnu.no/istx1002/2023h/start

- Vi ser på prosjektoppgaven på https://github.com/imfdrift/istx100y2023.

- Karakteren teller 30% til den endelige karakteren.

- Vi bruker prosentvurderingsmetoden: Konverterer poengene i en % (heltall, avrundet) og så bruker vi følgende skala:

Karakterskala for prosentvurderingsmetoden

A: 89–100 poeng

B: 77–88 poeng

C: 65–76 poeng

D: 53–64 poeng

E: 41–52 poeng

F: 0–40 poeng

Hvem er vi?

- **Studentene**: 8 studieprogram, omtrent 600 studenter totalt.
- **Faglig ansvarlig** for innholdet i modulen er Stefanie Muff (stefanie.muff@ntnu.no).
- I **veilederteamet** (for prosjektet) inngår i tillegg
  - Trondheim: øvingslærer Kenneth Aase og studentassistent Oscar Ovanger
  - Gjøvik: Stine Marie Berge
  - Ålesund: Siebe B. van Albada

Pensum og læringsressurser

Pensum og læringsressurser del 1:

- **Korte videoer**: (by Charles H. A. Curry)
  - Standardusikkerhet (4:35 min)
  - Fordelinger (3:28 min)
  - Relative usikkerheter (3:42 min)
- Denne forelesningen
- **Disse slides** med alle notatene og beregningene som er vist fram

Alt ligger på den eksterne kurssiden:

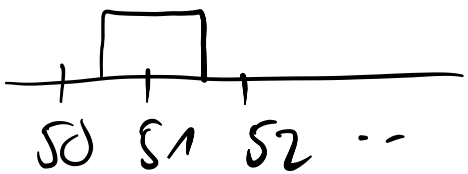
https://wiki.math.ntnu.no/istx1002/2023h/start

I tillegg vil jeg bruke litt materialier fra denne boken:

- Dere trenger ikke å kjøpe boken.
- Alt dere trenger blir diskutert i forelesningen.

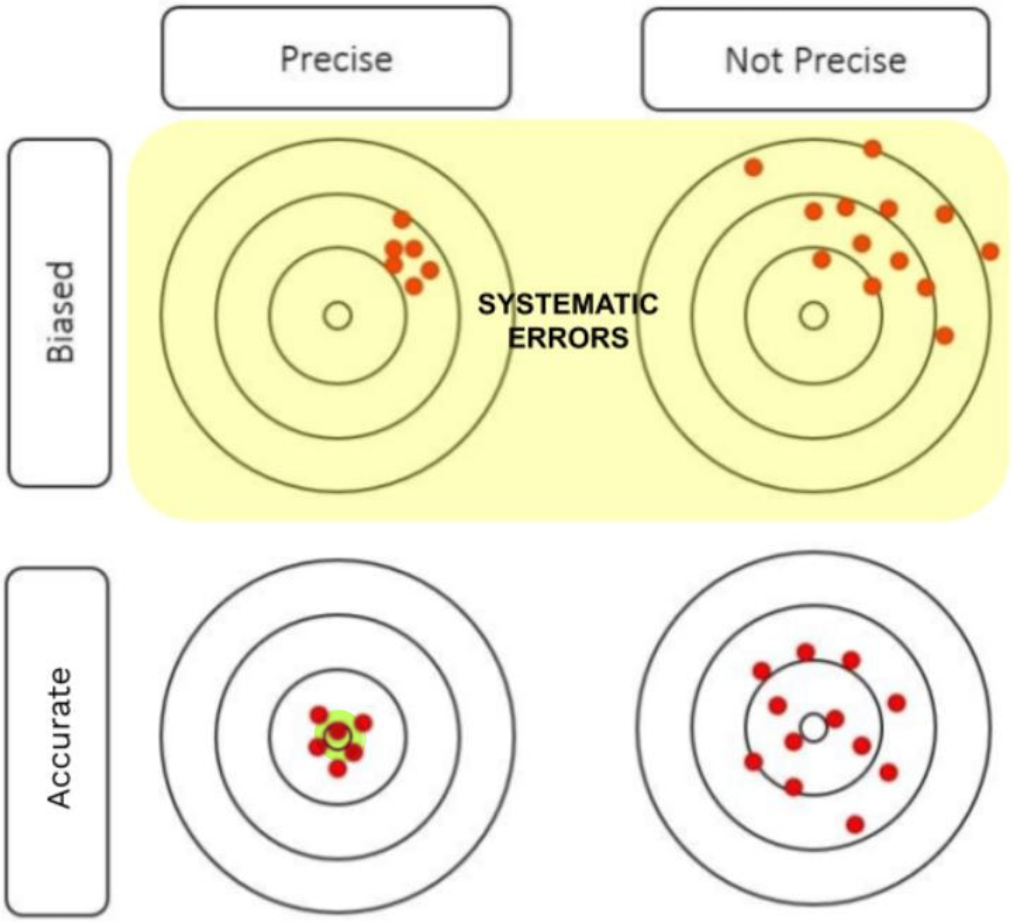
Kilder til måleusikkerhet / målefeil

- **Upresise målinger** i feltarbeid eller labor (lengde, vekt, blodtrykk...)
- Upresisjon i apparater/instrumenter.
- Feil grunnet **ufullstendige eller skjeve observasjoner** (for eksempel selvrapporterte kostvaner, helsehistorie).
- Skjeve observasjoner grunnet **preferansebasert utvalg** eller gjentatte observasjoner.
- Avrundingsfeil, sifferpreferanse.
- ...



Hvilke av de er *systematiske* feil, og hva er *tilfeldig variasjon*?

Systematisk vs tilfeldig variasjon



https://www.biologyforlife.com/error-analysis.html

Målinger

Hvis vi har eliminert alle systematiske påvirkninger, står vi igjen med *tilfeldig variasjon.*

En eller flere målinger

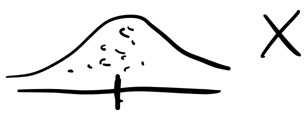
Vi kan måle

- **én gang**, eller
- **flere ganger**.

Hva er fordelen med de to strategiene?

Gjentatte målinger former et punktsverm rundt sentrum (se forrige folie).

- Det vi måler er alltid en *tilfeldig variabel*, vanligvis betegnet som X.



- Det betyr at X har en sannsynlighetsfordeling som ikke alltid er kjent (men vi ser på noen av de vanligste her).

- Hvis vi har n uavhengige målinger $x_1, x_2, \ldots, x_n$, kan vi bruke gjennomsnittet som estimator for μ og den empiriske variansen som estimator for σ^2:

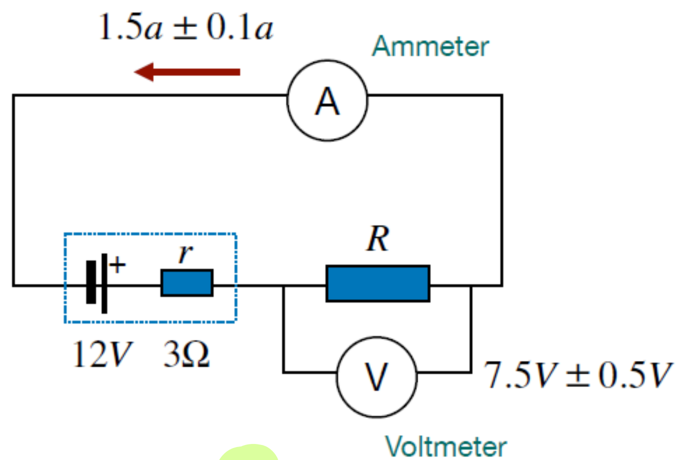
estimater til μ $\hat{\sigma}^2$

$$\begin{cases} \hat{\mu} = \overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} \\ \hat{\sigma}^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x}) \end{cases}$$

Husk! Det er alltid lurt å lagre verdiene til gjentatte målinger!!
(Hvorfor?)

Direkte og indirekte målinger

- **Direkte målinger:** Vi måler størrelsen vi ønsker direkte, for eksempel en vekt, temperatur, elektrisk motstand osv.
- **Indirekte målinger:** Vi måler flere størrelser og kombinerer de til en størrelse vi egentlig vil måle (vi an målefunksion). Eksempel:



$$R = \frac{U}{I}$$

Andre eksempler?

Med indirekte målinger snakker vi om **målesystemer** (ikke i dag).

Måleusikkerhet

$M = 52\,\mathrm{kg} \quad u(M) = 0.5\,\mathrm{kg}$

Litt terminologi...

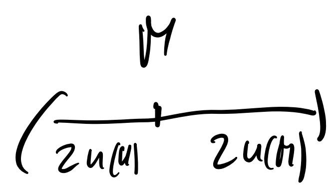
- **Standard usikkerhet:** Et annet ord for “standardavvik”, nemlig i fordelingen til en *målestørrelse M*. Vi skriver

$$u(M)\,, \rightarrow \sigma(M)$$

men det er også greit (og jeg foretrekker egentlig) å bruke $\sigma(M)$.

- **Dekningsfaktor k:** Ofte gir vi et intervall omkring måleresultatet som et multiplum k av standardusikkerhet. Usikkerhet gis da som $k \cdot u(M)$.

$$M \pm \underbrace{2}_{k} \cdot u(M)$$



- **Relativ standard usikkerhet:**

$$\frac{0.5}{52} \approx 1\% \qquad \Bigg\} \; \frac{u(M)}{M}$$

Vi går til **www.menti.com**

- Vi har målt en spenning på $U = 100V$ og det angis et intervall mellom 90V og 110V med en deknigsfaktor $k = 2$. Hva er da standard usikkerheten u i målingen?
- Vi måler lengden til et objekt med $l = 1m$ og vet at $u(l) = 2cm$. Hva er den *relative usikkerheten* i denne målingen?

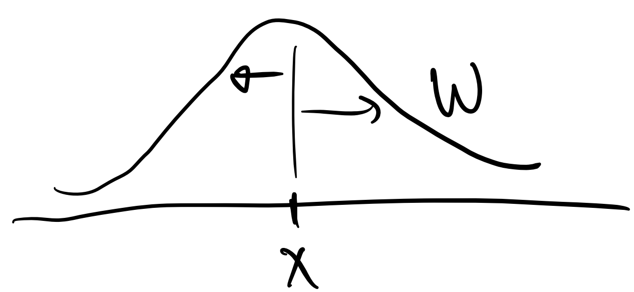
Hvordan (be)skriver vi usikkerheten?

- En typisk måte å skrive usikkerheten er jo som sagt ved bruk av $u(M) = \sigma(M)$, eller som et intervall $M \pm k \cdot u(M)$ med dekningsfaktor k.

- Men en annen måte som vi kanskje bruker mer i statistikk er å skrive en observasjon W som en summe av den egentlige størrelsen (X) plus støy/feil (U):

$$W = \boxed{X} + U \, , \qquad U \sim \mathcal{N}(0, \sigma_u^2)$$

hvor vi kan anta en egnet fordeling til U (se tema "fordelinger" senere).



To typer usikkerheter

I noen situasjoner er usikkerheten til et måleinstrument kjent, men i andre situasjoner må vi selv vurdere og kvantifisere usikkerheten i målingene vi gjør.

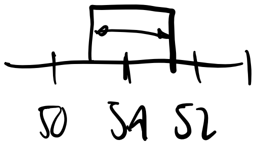
Type A: Usikkerheten finnes ved gjentatte målinger og det ukjente standardavviket kan da estimeres som

$$\hat{u}(x) = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})} \ ,$$

som kalles for *standard usikkerhet* i måleteknikk.

Type B: Vi kjenner usikkerheten $u(x)$ presist, enten via

- antakelser
- kunnskap om variasjon (for exsempel når det gjelder avlesningsusikkerhet)
- informasjon i bruksanvisning



$\rightarrow$ Ingen gjentatte målinger nødvendig.

Når vi har en type A usikkerhet, så må vi ikke bare måle verdien, men også usikkerheten selv.

Hvordan kan vi gjøre det?

Når vi har en type A usikkerhet, så må vi ikke bare måle verdien, men også usikkerheten selv.

Hvordan kan vi gjøre det?

Idé: Standard usikkerheten finnes ved å ta gjentatte målinger og å bruke det empiriske standardavviket som estimator.

Eksempel: Vi måler vår reaksjonstid!

- For en legeattest vil vi måle din typiske reaksjonstid, men det er jo mye støy i en sånn måling. Her får du litt tid for å gjennomføre fem målinger av din reaksjonstid. Test det én eller flere runder før du noterer de fem målte tidene (i millisekunder) i google sheetsen under.
- Til slutt ønsker vi at du noterer et anslag for hvor mange timer so sov i natt.

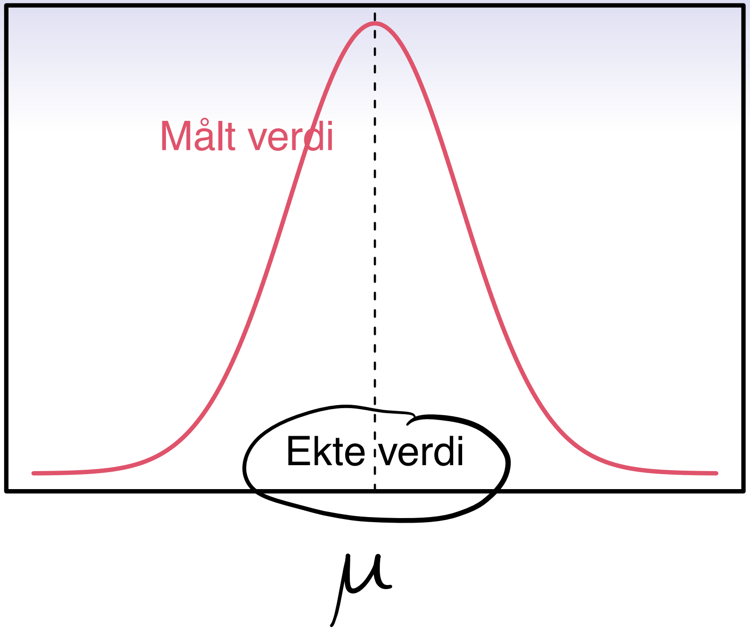
https://www.mathsisfun.com/games/reaction-time.html

Og så må de rapportere resultatene her:

https://docs.google.com/spreadsheets/d/1jswtNl8zmIdWwjuuRKSnKzaN-M5DLFISrnCIXl17w10/edit#gid=0

Tenk over det:

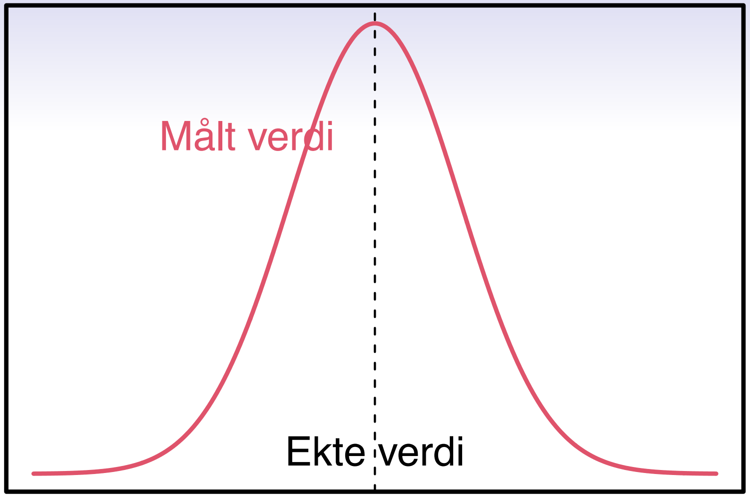
- Hva er et godt estimat for din typiske reaksjonstid?
- Har du en intuitiv forståelse av hvordan de målte verdiene sprer seg rundt den “ekte” verdien (μ)?



Alle må beregne

- gjennomsnitt $\longrightarrow$
- standardavvik $\longrightarrow$

til sine egne resultater.



Alle må beregne

- gjennomsnitt
- standardavvik

til sine egne resultater.

$$s = \frac{1}{n-1}\left((x_1-\bar{x})^2 + \ldots \ldots + (x_5-\bar{x})^2\right)$$

Her kan vi jobbe med mine tider: 256, 287, 257, 302, 299 ms.

$$\hat{\mu} = \bar{x} = \frac{256 + 287 + 257 + 302 + 299}{5} = 280.2$$

$$\hat{\sigma} = s = \sqrt{\frac{1}{4}\left((256-300.2)^2 + (287-300.2)^2 + \ldots\right)} = 22.35$$

Har vi blitt enige om at normalfordelingen er vel en realististisk approximasjon til feilfordelingen i måling av reaksjonstiden?

Oppgave: Finn flere situasjoner hvor en lik feilfordeling er realistisk.

Usikkerhet i én vs flere målinger

Tenk over hvordan standard usikkerheten hadde forandret seg om vi hadde tatt

- Bare én måling
- Ti målinger

Forskjellen mellom

- Usikkerheten i en eneste måling
- Usikkerheten i gjennomsnittet av flere målinger

Noen beregninger...

Usikkerheten i en enkel måling vs gjennomsnitt

- Standardavvik i én måling $SD(X) = \sigma$

  i et utvalg $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}$

- $\bar{X} = \frac{X_1 + \ldots + X_n}{n}$ $SD(\bar{X}) = ?$

$$\mathrm{var}(\bar{X}) = \mathrm{var}\Big(\underbrace{\tfrac{1}{n}}_{a} \cdot \underbrace{\sum_{i=1}^{n} X_i}_{Y}\Big) = \underbrace{a^2}_{\frac{1}{n^2}} \cdot \underbrace{\mathrm{var}\Big(\sum X_i\Big)}_{= \underbrace{\sum_{i=1}^{n} \mathrm{var}(X_i)}_{n \cdot \mathrm{var}(X_i)}}$$

$$= \frac{1}{n^{\cancel{2}}} \cdot \cancel{n} \cdot \mathrm{var}(X) = \frac{\mathrm{var}(X)}{n} = \frac{\sigma^2}{n}$$

$$\mathrm{var}(\bar{X}) = \frac{\mathrm{var}(X)}{n} = \frac{\sigma^2}{n}$$

$$\Rightarrow SD(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

$n = 2$

$$\frac{\sigma}{\sqrt{2}}$$

$n = 4$

$$\frac{\sigma}{\sqrt{4}} = \frac{\sigma}{2}$$

Oppgave – mentimeter

$\mathrm{var}(M) = \sigma^2$
$SD(M) = \sigma$

En målevariabel er antatt $M \sim \mathsf{N}(0, \sigma^2)$. Hvor mange uavhengige målinger må tas for at gjenomsnittet av målingene blir normalfordelt med standardavvik $\sigma/2$?

www.menti.com

Standardavvik vs standardfeil

Hva er forskjellen?

- **Standardavvik** (σ[redacted]): et mål til *spredning i en populasjon.*
- **Standardfeil (SE)** = Standardavvik i en estimator = Standard usikkerhet. Høres det for teoretisk ut?

Eksempel: Vår estimator er $\hat{\mu} = \overline{x}$, det vil si at vi tar gjennomsnittet av flere målinger for a estimere μ.

- Estimer standardavviket (spredningen) ved å bruke alle målingene. $\rightarrow\ s = \hat{\sigma}$
- Estimer standardfeilen i gjennomsnittet av de fem målingene.

$$SD(\bar{x}) = \frac{\hat{\sigma}}{\sqrt{n}} = SE(\bar{x})$$

Spredningsintervaller

- Vi ønsker ofte å angi et intervall slik at det for eksempel dekker den ekte verdien i 95% av tilfellene.
- I den første delen av kurset var det gjennomsnittet i en populasjon dere var interessert i.
- Her bytter vi "populasjon" med en størrelse som har en "riktig verdi", men som er vanskelig å måle. Vi tar gjentatte målinger for å få informasjon om spredningen og hvordan målingene fordeler seg rundt den ekte verdien.

For en normalfordelt størrelse $X \sim \mathsf{N}(\overset{\mu}{\cancel{0}}, \sigma^2)$ kan vi beregne et $1 - \underset{0.05}{\alpha}$ konfidensintervall som:

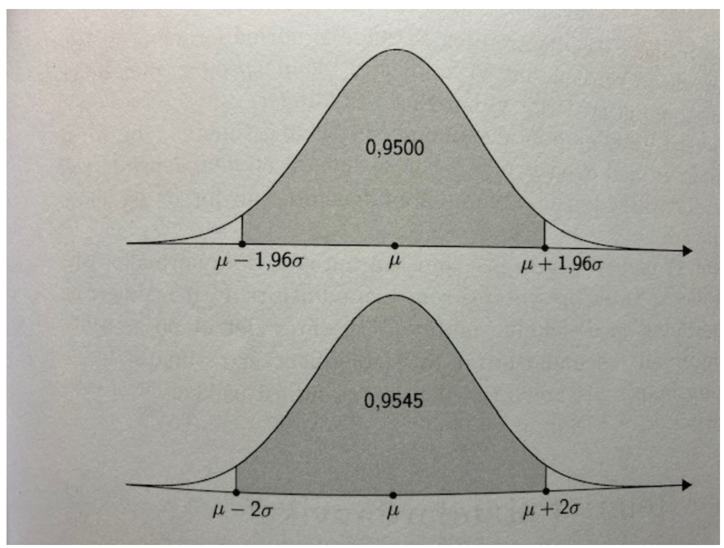
$$\boxed{[\mu - z_{\alpha/2} \cdot \sigma,\ \mu + z_{\alpha/2} \cdot \sigma]} \ . \quad \to 95\%\ \text{KI}$$

(for eksempel

Ofte bruker vi $\alpha = 0.05$ og får et 95% konfidensintervall med

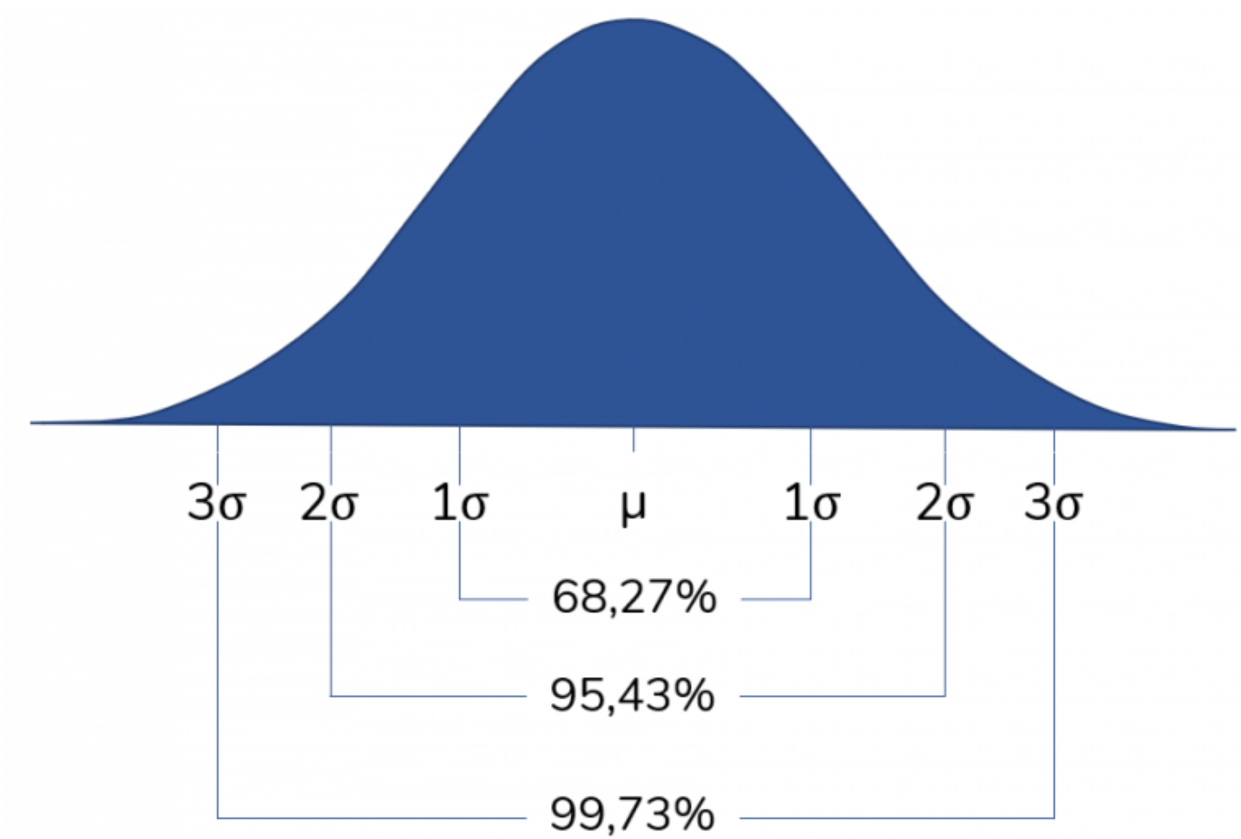
$$[\mu - 1.96 \cdot \sigma,\ \mu + 1.96 \cdot \sigma] \ ,$$

men vi bruker egentlig heller $1.96 \approx 2$.



Dekningsfaktor k

- k er tallet vi ganger standard usikkerheten med for å få forskjellige spredningsintervaller.
- Typisk $k = 1, 2, 3$:



https://eichhorn-coaching.de/optionstheorie-standardabweichung-normalverteilung-historische-volatilitaet/

Beregn 95% konfidensintervallet til din estimerte reaksjonstiden

- Husk at $\overline{X} \sim \mathsf{N}(\mu, \frac{\sigma^2}{n})$ $SD(\overline{X}) = \frac{\sigma}{\sqrt{n}}$
- Derfor er et 95% konfidensintervall gitt som

$$[\mu - \underbrace{z_{0.975}}_{=1.96} \frac{\sigma}{\sqrt{n}}; \quad \mu + z_{0.975} \frac{\sigma}{\sqrt{n}}]$$

- Men: vi vet jo ikke μ og σ!
- Vi har derfor estimert de to: $\hat{\mu}$ og $\hat{\sigma}$, og 95% konfidensitervallet er dermed gitt som

$$[\overline{x} - t_{0.975,n-1} \frac{s}{\sqrt{n}}; \quad \overline{x} + t_{0.975,n-1} \frac{s}{\sqrt{n}}] \ ,$$

hvor s er standardavviket i vårt utvalg av n gjentatte målinger.

Merk at z-verdien har blitt byttet med t-verdien.

2.78

Beregn 95% konfidensintervallet til din estimerte reaksjonstiden

$\bar{x} = 280.2$ ms , $s = 22.35$

$$\left[\underbrace{280.2 - 2.78 \cdot \frac{22.35}{\sqrt{5}}}, \underbrace{280.2 + 2.78 \cdot \frac{22.35}{\sqrt{5}}}\right]$$

$$[252.4 , 308.0]$$

Eksempel

$$\left[10 - \underbrace{\overset{z_{0.995} \cdot \mu(R_s)}{0.000135}}, \; 10 + 0.000135\right]$$

En standard motstand viser måleverdien på $R_s = 10\Omega$ (ohm). Produsenten angir at verdien da ligger i et 99% konfidensintervall $\pm 135\mu\Omega$.

- Hvis vi antar at måleusikkerheten sprer seg normalfordelt, hva er standard usikkerheten i målingen da?
- Hva er den relative standardusikkerheten?

$$z_{0.005} = -z_{0.995} \qquad 1 - \frac{\alpha}{2}$$

$$135\,\mu\Omega = \underbrace{z_{0.995}}_{2.576} \cdot \mu(R_s) \Rightarrow \mu(R_s) = \frac{135}{2.576} = 52\,\mu\Omega$$

$$\frac{\mu(R_s)}{R_s} = \frac{52\,\mu\Omega}{10\,\Omega} = 0.00052\,\%$$

Fordelinger

Nå har vi sett eksempler for tilfeldig støy hvor den tilfeldige målevariablen X kan bli antatt normalfordelt:

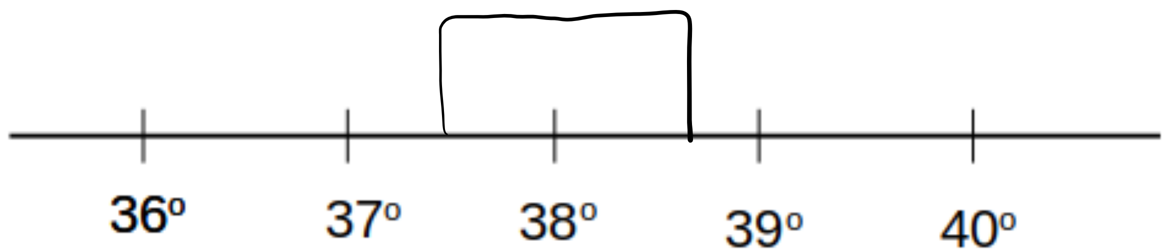
$$X \sim \mathsf{N}(\mu, \sigma^2) \ ,$$

med en sann, ukjent verdi for μ og σ^2.

Men hva med avlesningsusikkerhet?

Avlesningsusikkerhet

Eksempel: Vi måler temperatur med en måler som gir temperatur i hele grader.

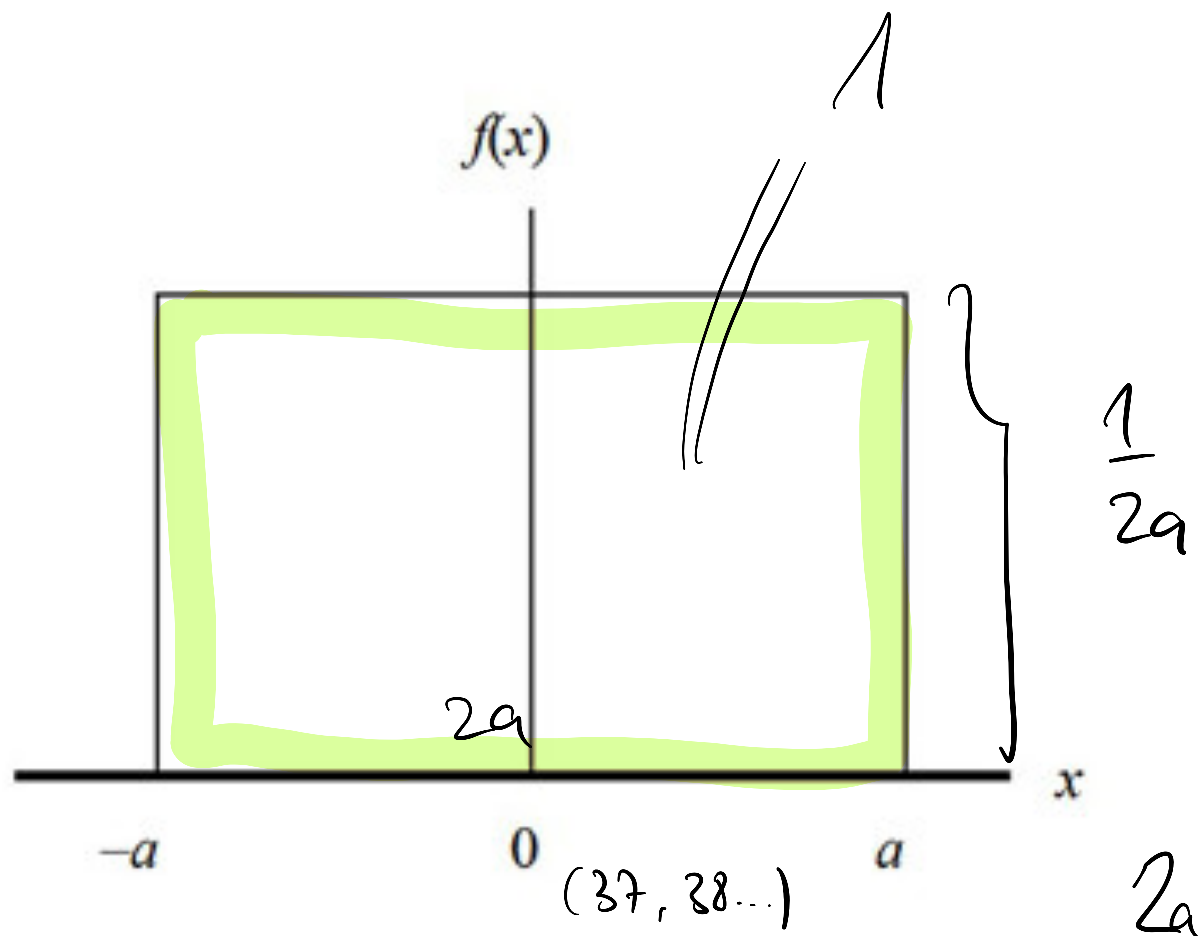


- Hva vet du når du, for eksempel, måler 38 grad?
- Hva er oppløsningen?
- Hva er standard usikkerhet (standardavvik)?

$$u(T) = ?$$

Firkantfordeling

For å forstå hvordan man får standard usikkerheten, må vi se på firkantfordelingen:

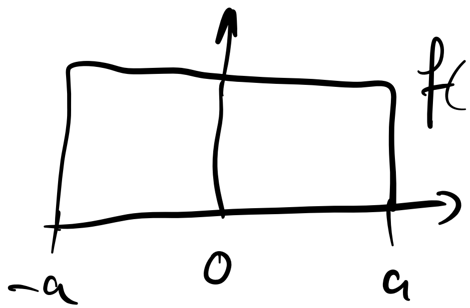


$2a \cdot \frac{1}{2a} = \underline{\underline{1}}$

$$f(x) = \begin{cases} \frac{1}{2a} & -a \leq x \leq a \\ 0 & \text{ellers} \end{cases}$$

Firkantfordeling

Forventningsverdi og varians (σ^2); standardavvik (σ):



$$f(x) = \begin{cases} \frac{1}{2a}, & -a \leq x \leq a \\ 0 \end{cases}$$

$$\underline{\underline{E[f(x)]}} = \int_{-a}^{a} x \cdot f(x)\, dx = \int_{-a}^{a} x \cdot \frac{1}{2a}\, dx = \frac{x^2}{2}\frac{1}{2a}\Bigg|_{-a}^{a}$$

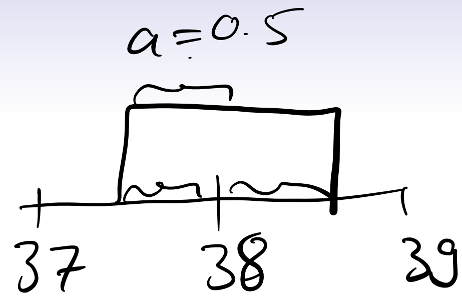
$$= \frac{a}{4} - \frac{a}{4} = \underline{\underline{0}}$$

$$\underline{\underline{\mathrm{var}(f(x))}} = \int_{-a}^{a} x^2 \underbrace{f(x)}_{\frac{1}{2a}}\, dx = \frac{1}{2a}\left[\frac{x^3}{3}\right]_{-a}^{a} = \frac{1}{\cancel{2a}}\underbrace{\left(\frac{a^3}{3} - \frac{(-a)^3}{3}\right)}_{\frac{\cancel{2}a^{\cancel{3}2}}{3}}$$

$$= \underline{\underline{\frac{a^2}{3}}}$$

$$\Rightarrow u(f(x)) = \sigma = \underline{\underline{\frac{a}{\sqrt{3}}}}$$

Temperatureksempelet:



$$f(t) = \begin{cases} 1\ , & 37.5 \le t \le 38.5\ , \\ 0\ , & \text{ellers.} \end{cases}$$

$$\mathrm{var}(f(t)) = \frac{a^2}{3} = \frac{0.5^2}{3}$$

$$\Big\downarrow \sqrt{}$$

$$\underline{\underline{u(f(t))}} = \frac{0.5}{\sqrt{3}} = \underline{\underline{0.289}}$$