ISTx1002 Usikkerhet og støy i målinger

Feilforplantning i målesysemer; Kalibrering

Stefanie Muff, Institutt for matematiske fag, NTNU Trondheim

Oktober 24 og 30, 2023

Plan for i dag og mandag 14:15-15:00

Tema 1: Feilforplantning i målesysemer

Tema 2: Kalibrering

Pensum og læringsressurser

Husk lenken til den eksterne modulsiden:

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

Pensum del 2:

Korte videoer: (by Charles H. A. Curry)

Målefunksjoner og kombinert standardavvik (7:08)
Kalibrering (5:06)

Denne forelesningen

Disse slides med alle notater og beregninger som er vist fram

Målesystemer

Et **målesystem** beskriver en sammenhang mellom ‘en eller flere inngangsstørrelser og den størrelsen vi egentlig vi måle.

Eksempel:

Vi skal måle volumen til en gjenstand som er en sylinder med høyde h og diameter d. Volumen er gitt ved

$$V(d, h) = d^2 h \ .$$

Målefunksjonen defineres som $V(d, h)$ og ses som funksjon av variablene h og d.

Feilforplantning i målesystemer

Vi har sett på standard usikkerhet i en eneste variabel. Nå skal vi gjøre ting litt mer kompliserte:

1) Usikkerhet i en linear kombinasjon av variabler
2) Usikkerhet i en ikke-linear *funksjon* av en variabel $f(X)$
3) Usikkerhet i en ikke-linear *kombinasjon* av flere variabler $f(X,Y)$ for ukorrelerte X, Y
4) Usikkerhet i en ikke-linear *kombinasjon* av flere variabler $f(X,Y)$ for korrelerte X, Y

Kombinert standard usikkerhet (eller “feilforplantning”)

Eksempler:

Mål summen av flere steg (X_1, X_2,..), slik at du tar $X_1 + X_2 + \; + X_n$.

Mål en størrelse X, men så er du interessert i en transformert version, for eksempel X^2 (tenk at du vil måle arealeat av en firkant).

Mål lengde (L) og bredde (B) til en rektangel og beregn arealet $f(L, B) = L \;\; B$.

Mål masse m og hastighet v av et objekt og beregn bevegelsesenergi $f(m, v) = \frac{1}{2}mv^2$.

Og alle størrelser vi måler har forskjelle grader av usikkerhet som “forplanter” seg videre til den endelige størrelsen vi egentlig vil måle.

1) Lineare kombinasjoner av variabler

Problemstilling: Vi har to variabler, X_1 og X_2 som måler steglengden til to personer. Vi måler steglenger $x_1 = 75$cm og $x_2 = 82$cm med usikkerhet $u(X_1) = 6$ og $u(X_2) = 9$cm. Vi antar steglengdene er uavhengige ($\mathrm{Cov}(X_1, X_2) = 0$).

a) Hva er usikkerheten i summen av 5 uavhengige steg gått av person 1?

b) Hva er usikkerheten i gjennomsnittig steglengde til de to personene?

Generell regel:

For en kombinert størrelse

$$Y = f(X_1, \; , X_n) = a_1 X_1 + a_2 X_2 + \; + a_n X_n$$

kan vi beregne variansen som

$$\begin{aligned}\mathrm{Var}(Y) =& a_1^2 \mathrm{Var}(X_1) + a_2^2 \mathrm{Var}(X_2) + \; + a_n^2 \mathrm{Var}(X_n) \\ &+ 2 \mathop{}_{i=1}^{n1} \mathop{}_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j) \; .\end{aligned}$$

Husk: $\mathrm{Var}(X) = u^2(X)$, fordi standard usikkerheten er bare standardavviket av størrelsen.

Nå kan vi beregne løsning for problemstillingen ovenfor:

2) Transformert versjon av en variabel

Problemstilling: Du måler lengden X til en firkant med 55cm (nøyaktigheten er bare 1cm), og så vil du beregne arealet til firkanten. Hva er usikkerheten i arealmålingen?

55cm med en usikkerhet som er firkantfordelt, og derfor…

Men hvordan bruker vi det når vi må gange X med seg selv…?

Usikkerheten i en transformert variabel $f(X)$

Generell regel:
Når vi har en variabel X som vi kan måle og vet usikkerheten $u(X)$, men vi er interessert i en transformert version $f(X)$ (for eksempel $f(X) = X^2$, $\log(X)$,...), kan vi bruke det følgende:

Trick: Taylor approximasjon
For en gitt måleverdi a har vi omtrent

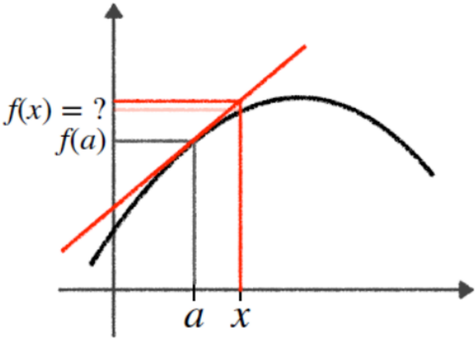
$$
\begin{aligned}
f(X) \quad & f() + f()(X \quad) \\
& = [\underset{a}{f() \quad f()}] + \underset{b}{f()}\, X
\end{aligned}
$$

Derfor har vi omtrent

$$\mathrm{Var}(f(X)) = \underset{b^2}{f()^2}\ \underset{u^2(X)}{\mathrm{Var}(X)}\ .$$

$$\text{Eller:} \quad u(f(X)) = \ f()^2\ \ u^2(X)\ .$$

Visualisering: Taylor approximasjon i 1 dimension:



$$f(x) \approx f(a) + \boxed{f'(a)} \cdot (x - a)$$

$f(a)$ heter følsomhet.

Nå kan vi beregne arealet og approximert usikkerhet til arealet for firkanten:

3) Ikke-linear *kombinasjon* av flere variabler $f(X_1, X_2)$, med ukorrelerte X_1, X_2

Problemstilling: Vi vil beregne bevegelsesenergi til et objekt og måler

masse $m_0 = 0.45, u(m) = 0.01kg$
hastighet $v_0 = 10.8, u(v) = 0.05m/s$

Hva er usikkerheten i den beregnete begevelsesenergien?

$$f(m, x) = \frac{1}{2}mv^2$$

Hvordan går vi frem i flere dimensioner?

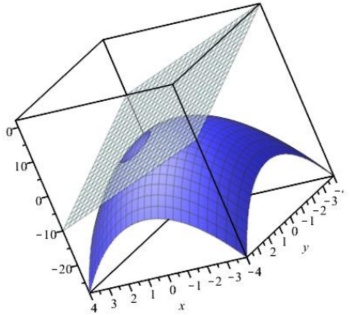
Generell regel:

Trick: Taylor approximasjon (igjen) for flere dimensioner

For gitte måleverdier $_1$ og $_2$ av en funksjon $f(X_1, X_2)$ har vi omtrent

$$
\begin{aligned}
f(X_1, X_2) \quad & f(_1, _2) \\
& + \underbrace{\frac{f}{X_1}(_1, _2)}_{a}(X_1 \quad _1) + \underbrace{\frac{f}{X_2}(_1, _2)}_{b}(X_2 \quad _2)
\end{aligned}
$$

Visualisering: Taylor approximasjon 2 dimensioner:



$$f(x,y)$$
$$\approx f(a,b) + \boxed{\frac{\partial f}{\partial x}(a,b)} \cdot (x-a)$$
$$+ \boxed{\frac{\partial f}{\partial y}(a,b)} \cdot (y-a)$$

$\frac{f}{x}(a,b)$ og $\frac{f}{y}(a,b)$ heter følsomhet.

Og derfor er

$$\mathrm{Var}(f(X_1, X_2)) \quad a^2\mathrm{Var}(X_1) + b^2\mathrm{Var}(X_2)$$

$$u(f(X_1, X_2)) \quad a^2\mathrm{Var}(X_1) + b^2\mathrm{Var}(X_2)$$

Nå kan vi beregne bevegelsesenergien og approximert usikkerhet:

4) Ikke-linear *kombinasjon* av flere variabler $f(X, Y)$, med korrelerte X, Y

Generell regel:

$$
\begin{aligned}
\mathrm{Var}(f(X_1, X_2)) \quad & (\frac{f}{X_1}(_1, _2))^2 \mathrm{Var}(X_1) \\
& + (\frac{f}{X_2}(_1, _2))^2 \mathrm{Var}(X_2) \\
& + 2 \ \frac{f}{X_1}\frac{f}{X_2} u(X_1, X_2) \ ,
\end{aligned}
$$

hvor $u(X_1, X_2) = u(X_1)u(X_2)\mathrm{Cor}(X_1, X_2)$.

Det må dere ikke lære utenat.

Men husk: sterkere korrelasjoner mellom X_1 og X_2 betyr større avvik fra tilfelle 3) hvor vi hadde uavhengige variabler.

Kalibrering – grunnlegende idé

Prosessen hvor vi presiserer forhold mellom målte og ekte verdier (eller nominell og ekte verdi)

Retting av **systematisk feil**, for eksempel i et måleminstrument

For å kalibrere, sammenligner vi resultatene av målinger fra et instrument med verdier fra et annet kilde som vi vet er mer nøyaktig (3-5 ganger nøyaktigere).

Eksempler: Måle en kjent kilde på 100V, en kjent lengte på 2m eller en kjent vekt på 1kg.

Kalibrering eksempel: Voltmeter

Vi måler en kjent kilde på 100V med to voltmetere, 10 ganger hver:

	A	B
1	100,5	90,0
2	95,5	92,0
3	101,5	91,5
4	104,0	90,5
5	100,5	89,5
6	103,0	90,0
7	99,5	91,0
8	101,0	89,5
9	98,5	88,5
10	103,0	92,0
gjennomsnitt	100,7	90,45
standardavvik	2,49	1,17

Hva vet vi nå om de to voltmetere? Hva ville du gjort hvis du må bruke en av disse to voltmetere?

Hva er *usikkerheten* i de to estimatene av systematiske feil?

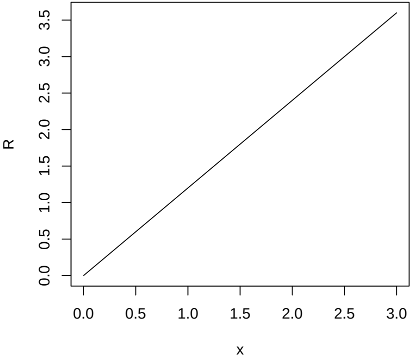
Hvilket voltmeter vil du heller bruke, og hvorfor?

Hvordan går du frem hvis du måler en spenning på omtrent 200V eller 25V med en av de to voltmetere?

Kalibreringskurver

Det er mulig at samme voltmeter som underestimerer når den skulle måle 100V, er mye nærmere den ekte verdien når den skulle måle 200V.

Mer generelt måler vi en størrelse R som har en linear sammenheng med en størrelse x vi er interessert i.



Kalibreringskurven kan beskrives som

$$R = k \ x \ ,$$

hvor k skal estimeres, eller

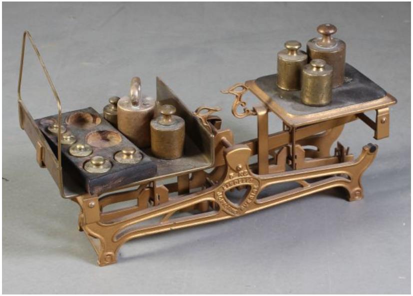
$$R = a + k \ x$$

hvis response ikke kan antas a være lik når $x = 0$.

Når vi kan anta at R er normalfordelt, kan vi bruke minste kvadratsumme (linear regresjon) for å finne k.

Eksempel 1: Kalibrering av et lodd

Her regner vi eksempelet for kalibrering av et lodd:



Et lodd med ukjent masse M_x på ca 10g skal kalibreres mot et 10g referanselodd.

1. Referanseloddet hadde ved siste kalibreringen massen $M_r = 10.005g$ og er antatt normalfordelt med standardavvik av $22.5mg$.
2. Mulig drift i referanseloddets masse siden kalibreringen: $M_r = 0$ med maksimal ś$15.5mg$
3. Målt differanse: Ukjent lodd - referanselodd: Ved 20 gjennomtatte målinger er gjennomsnitt $M_{diff} = 20mg$ med standardavvik $S = 64.6mg$, og derfor er standardavviket for middelverdien $S/\ 20 = 14.4mg$.

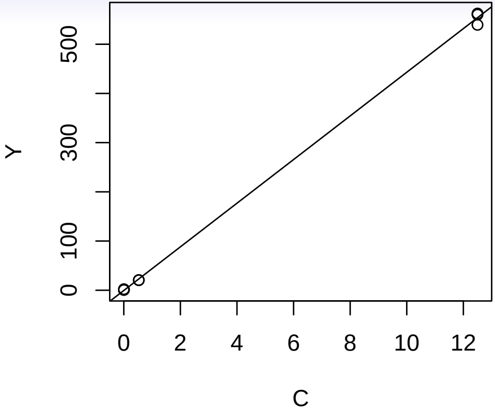
Oppgave: Angi en god estimat for verdien av det ukjente loddet og standard usikkerheten til den verdien.

Løsning:

Eksempel 2: Kalibreringskurve

Vi er interessert i konsentrajonen C av en radioaktivt stoff i en bergart. Siden det er vanskelig å måle den direkte, kan vi måle radon (en radioaktiv gass) som dannes av stoffen i bergarten. For å finne ut på relasjonen mellom konsentrasjon i steinen (C) og radon i luften (Y), måler vi åtte ganger strålingsintensiteten i tre kjente konsentrasjoner av uran:

prøve	Y_i	C_i
1	0.64	0.0036
2	0.67	0.0036
3	2.19	0.0036
4	20.35	0.53
5	20.80	0.53
6	539.4	12.5
7	560.2	12.5
8	562.4	12.5



Vi kan tilpasse en linear regresjonslinje:

$$Y = k\ C$$

med estimert $k = 44.36$ og $sd(k) = u(k) = 0.45$.

Anta vi målete en spesifisk stråleintensitet $y = 120$ i luften. Vi vil nå vite to ting:

1) Hva er den estimerte konsentrasjonen i bergarten?
2) Hva er usikkerheten i estimaten?

Løsning:

1)

$$c = \frac{y}{k} = \frac{120}{44.36} = 2.71$$

2) Usikkerheten er litt mer vanskelig å bestemme, og det finnes mer teoretiske og mer empiriske måter. Her gjør vi det empirisk og beregner forskjellen mellom de estimerte C_i verdiene med de observerte (det er et slags "residual", men for x-variablen istendenfor y)

$$W_i = \frac{Y_i}{k} \; C_i \; ,$$

og så tar vi standardavviket av alle W_i, $s = \frac{(W_i \overline{W})^2}{n1} = 0.16$