

## Effektiv tavleundervisning om SD og SE (ca 1 times undervisning).

Av Kathrine Frey Frøslie, førsteamanuensis i biostatistikk, NMBU. Blogg: [www.statistikk.no](http://www.statistikk.no)

Det er en vanlig misforståelse å blande sammen de to størrelsene **standardavvik** (SD) og **standardfeil** (standard error, SE). Dette dokumentet skisserer en undervisningsøkt for å rydde opp i dette. Den kan gjennomføres på fysisk eller digital tavle (nettbrett). Hvis tavle ikke er tilgjengelig\*, kan disse notatene også fungere som utgangspunkt for en diskusjonsoppgave, helst i smågrupper.

Undervisningen ble i utgangspunktet utviklet som en tavleundervisningsøkt for forskere i det medisinske fagfeltet. Jeg har gjennomført opplegget utallige ganger i undervisning på Oslo universitetssykehus og andre steder, og tilbakemeldingene har vært svært gode.

### Kort introduksjon for både lærere og studenter:

Gjennomsnittet (mean) og det empiriske standardavviket (SD) hører til under overskriften *deskriptiv statistikk*. Disse tallene brukes til å beskrive observasjoner i et utvalg. Gjennomsnittet beskriver «det typiske», «det vanlige» for kontinuerlige data som er symmetrisk fordelt, og SD beskriver spredningen (variasjonen) i de observerte verdiene.

For å forstå forskjellen på SD og SE, er det viktig å være klar over at gjennomsnittet kan brukes til to ting:

- 1) Til å beskrive et utvalg, og
- 2) Til å generalisere fra utvalget til populasjonen.

Det siste kaller vi å estimere.

Gjennomsnittsverdien kan altså brukes både som et oppsummeringstall og som et estimat.

Mens SD beskriver variasjonen i observerte verdier, er standardfeilen SE et tall som beskriver estimeringsusikkerheten til et gjennomsnitt (eller et annet estimat).

### Bruksanvisning for lærere:

Sammenligning av det som til slutt står på høyre og venstre side av tavla er essensielt. Start derfor økta med å dele tavla i to med en vertikal strek, eller ha tegnebrettet i liggende format og gjør det samme. Oppfordre studentene til å notere over en dobbeltside i boka si, eller i liggende format på nettbrett slik du gjør. Du fyller ut venstre side av tavla først, deretter høyre side, slik neste side viser. Siste side viser eksempel på hvordan dette ble løst på en ganske liten white-board.

Overskrifter (skrives på tavla) indikeres med at de er skrevet på grønn bakgrunn

Den gule teksten er muntlig informasjon eller (lette) plenumsspørsmål som kan brukes som utgangspunkt for diskusjon i hele gruppa eller i mindre grupper (2-4 personer), avhengig av gruppedynamikken og fysiske eller digitale undervisningsrom. (Vær obs på at det vil ta litt lenger tid å gjennomføre med digitale breakoutrooms.)

Resten av teksten (inkludert håndtegnede skisser av histogrammene) er ment å skrives på tavla.

Legg merke til at det er mer tekst i del 2 (estimering) enn i del 1 (deskriptiv statistikk). Grunnen er at del 2 er konseptuelt vanskeligere enn del 1, og at dette er det studentene trenger tid på å fordøye.

Disse notatene er mindre presise enn en lærebok om temaet ville vært. Men nettopp derfor kan de brukes som et supplement til lærebøkene. Studentene oppfordres til å renskrive egne notater med sine egne ord.

### Bruksanvisning for studenter som ikke får dette som tavleundervisning\*

- Skriv ut neste side i A3-format og med fargetrykk.
- Gå gjennom venstre side av tavla, deretter høyre side. Alle de gule kommentarene må diskuteres, og hver av dere noterer med egne ord supplement til den trykte teksten.

Anta at du jobber med rehabilitering av pasienter og er interessert i gangfunksjonen deres. Du utvikler en ny test for å måle det, som du kaller «6 Minute Walking Distance, 6MWD». Testen utføres sånn: Testpersonen går vanlig i 6 minutter, og distansen personen går, måles i meter. **Diskuter hvor langt du tror en person med normal funksjon vanligvis går på 6 minutter.**

### Studie 1: Pilotstudie (prøveprosjekt)

Mål: Å undersøke 6MWD blant menn i den generelle befolkningen.

Utvalg:  $n=25$ , fysioterapeuter ved et rehabiliteringssykehus. **Tilfeldig utvalg? Representativt?**

#### Oppgave 1: Beskrive observasjonene/utvalget (Deskriptiv statistikk)

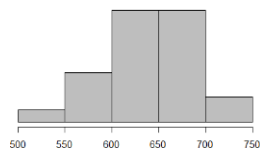
Hvilken type data?

Kategorisk eller kontinuerlig?

Kontinuerlige data ->

Hva gjør vi for å vurdere fordelingen?

Må lage histogram:



OBS: «Som et digitalt bilde med dårlig oppløsning: Grovt, upresist»

Symmetrisk eller skjev fordeling?

-> Symmetrisk fordeling:

Hvilke tall er gode oppsummeringstall for midten og for spredningen?

Gjennomsnittet (mean)=636m og SD=46m er gode oppsummeringstall

Hva betyr disse tallene?

### Study 2: The national research council is enthusiastic about the project, and gives you a grant to study this in a population-wide study.

**Objective?** Objective: Same.

Sample:  $n=10,000$  males from the general population. **Random sample? Representative?**

#### Task 1: Still (and always) to summarize the observed data (Descriptive statistics)

Fortsatt:

Hvilken type data?

Kontinuerlig ->

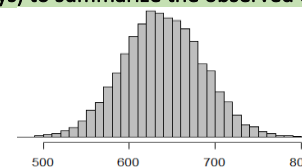
Fortsatt: Hva må vi?

Må lage histogram.

«Basert på piloten:

Hvilken fordeling

forventer du å se?»



OBS: «Som et bilde med supergod oppløsning: Presist og detaljert.

Symmetrisk fordeling, altså: Gj.snitt og SD er gode oppsummeringstall. **Hva tror du tallene er nå når n er større?**

Gj.snittet: **Større? Mindre? Ca likt?**

Mean=639m (nesten likt)

**Hva med SD? Større? Mindre? Ca likt?**

SD=47m (nesten likt)

OBS: Nesten samme tall som i piloten

Egenskapene til normalfordelingen (som histogrammet ligner på) forteller at:

Intervallet  $636 \pm 2 \cdot 46 = (543, 728)$ \* inneholder **hvor mange?** de fleste observasjonene (ca. 95%) av 6MWD

Intervallet  $636 \pm 3 \cdot 46 = (497, 774)$  inneholder **hvor mange?** nesten alle (ca. 100%) observasjonene

Egenskapene til normalfordelingen (som også dette histogrammet ligner på) forteller fortsatt at:

Intervallet  $639 \pm 2 \cdot 47 = (545, 732)$ \* inneholder **hvor mange?** de fleste observasjonene (ca. 95%)

Intervallet  $639 \pm 3 \cdot 47 = (499, 779)$  inneholder **hvor mange?** nesten alle (ca. 100%) observasjonene

OBS: Nesten samme som i pilotstudien, men disse tallene kan vi stole mer på. **Hvorfor?**

#### Oppgave 2: Si noe om befolkningens 6MWD, basert på observasjonene i utvalget (inferens)

Tenker: Hvis gj.snittet i utvalget er 636, er antakelig gj.snittet i befolkningen omtrent det samme, hvis utvalget er representativt.

Det ukjente befolkningsgjennomsnittet (også kalt forventningsverdien) noteres  $\mu$ .

M.a.o., vi antar at det observerte gj.snittet,  $\bar{x}$  ligner på  $\mu$  (og vi tenker at  $\bar{x} \approx \mu$ , men ikke skriv det!)

Vi sier at  $\bar{X}$  er en estimator for  $\mu$ , og at tallet  $\bar{x} = 636$  er et estimat for  $\mu$ .

Men hvor mye stoler vi på tallet 636? Hvor langt unna 636 er det rimelig å anta at det ukjente befolkningsgjennomsnittet  $\mu$  kan være? Kan det like gjerne være 640? 650? 680? 600? 700?

For å svare på det, må vi beregne estimeringsusikkerheten til  $\bar{X}$ : Standarfeilen, eller standard error (S.E.)

Estimeringsusikkerheten S.E. til et gjennomsnitt er gitt ved

$$SE(\bar{X}) = SD/\sqrt{n} = 46/\sqrt{25} = 9.2$$

Dette tallet har ikke en like meningsfull tolkning i seg selv, slik SD har, men er i stedet en mellomregning når vi skal beregne et (95%) konfidensintervall for befolkningsgjennomsnittet  $\mu$ .

I akkurat denne situasjonen er et 95% KI for  $\mu$  gitt ved

$$\bar{X} \pm 2.064 \cdot S.E.(\bar{X}) \rightarrow 636 \pm 2.064 \cdot 46/5 \rightarrow (617, 655)$$

OBS: Dette intervallet er ment å dekke den ukjente verdien til befolkningsgjennomsnittet  $\mu$ , og ikke enkeltobservasjoner av 6MWD, som i Oppgave 1.\*

Tolkning: Vi er 95% sikre på at intervallet (617, 654) inneholder den sanne verdien til  $\mu$ .

**Kan befolkningsgjennomsnittet  $\mu$  like gjerne være 640? 650? 680? 600? 700?**

#### Oppgave 2: Si noe om befolkningens 6MWD, basert på observasjonene i utvalget (inferens)

Tenker: Hvis gj.snittet i utvalget er 639, er antakelig gj.snittet i befolkningen omtrent det samme, hvis utvalget er representativt.

Det ukjente befolkningsgjennomsnittet (forventningsverdien) noteres fortsatt  $\mu$ .

(tenker fortsatt « $\bar{x} \approx \mu$ », men ikke skriv det på eksamen!)

Vi sier at  $\bar{X}$  er en estimator for  $\mu$ , og at tallet  $\bar{x} = 639$  er et nytt estimat for  $\mu$ .

Men hvor mye stoler vi på tallet 639? Hvor langt unna 639 er det rimelig å anta at det ukjente befolkningsgjennomsnittet  $\mu$  kan være? Kan det like gjerne være 640? 650? 630?

Estimeringsusikkerheten til dette nye estimatet kan beregnes med samme formel:  $S.E.(\bar{X}) = SD/\sqrt{n}$

**Forventer du nå at verdien til S.E. skal være større, mindre eller ca lik som S.E. i pilotstudien?**

$$SE(\bar{X}) = SD/\sqrt{n} = 47/\sqrt{10,000} = 0.47 \text{ (MYE mindre!)}$$

Akkurat som i pilotstudien betrakter vi dette som en mellomregning på vei mot konfidensintervallet for  $\mu$ , i stedet for å gi det en egen tolkning, slik vi gjør med SD.

Et 95% KI for  $\mu$  er nå gitt ved

$$\bar{X} \pm 1.96 \cdot S.E.(\bar{X}) \rightarrow 639 \pm 1.96 \cdot 0.47 \rightarrow (637.7, 639.5)$$

OBS: Dette intervallet er ment å dekke den ukjente verdien til befolkningsgjennomsnittet  $\mu$ , og ikke enkeltobservasjoner av 6MWD, som i Oppgave 1.\*

Tolkning: Vi er 95% sikre på at intervallet (637.7, 639.5) inneholder den sanne verdien til  $\mu$ .

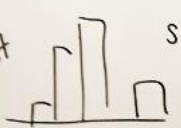
**Kan befolkningsgjennomsnittet  $\mu$  like gjerne være 640? 650? 630?**

Her er et eksempel på hvordan en slik tavleøkt ble løst med ganske dårlig plass på en liten whiteboard. Selv det går an:

Interessert i gangfunksjon : 6 minutt walking test 6MWT

Pilot : n=11 Fysio SS menn 30-60

HUNT : n = 10.000 menn 30-60

hist  Symmetri eller Skjevt? (Table I)

① Oppsummer data : Deskriptiv stat, beskrivelse av utvalg

$\bar{X} = 750\text{m}$   
 $SD = 85\text{m}$

---

② "Si noe om pop" basert på utvalget: generalisere

Estimerer gjennomsnittlig 6MWT i pop:  $\mu$  (ukjent)  
formått (forventningsverdien)

Jordt gj.sn. i utvalget ligner på pop. gj.sn.  
Bruker vi  $\bar{X}$  som estimat for  $\mu$ :  $\bar{X} \approx \mu$


Et estimat for  $\mu$  er  $\bar{X} = 750\text{m}$

Må vite estimeringsusikkerheten (standard error)  
 $S.E. = SE(\bar{X}) = SE(\text{mean}) = \frac{SD}{\sqrt{n}} = \frac{85}{\sqrt{11}} \approx 26$

Beregn et 95% konfidensintervall for  $\mu$

[698, 802]

---

hist  Symmetri eller Skjevt?

① oppsummer data: Beskrive utvalget

$\bar{X} = 725$   
 $SD = 90$

---

② "Si noe om pop" basert på utvalget: generalisere

Estimerer gjennomsnittlig 6MWT i pop:  $\mu$  (ukjent)

Nytt estimat for  $\mu$  er  $\bar{X}$  i dette utvalget:  
 $\mu \approx \bar{X} = 725$  (punktestimert for  $\mu$ )

Beregn estimeringsusikkerheten (standard error)  
til dette estimatet:  
 $SE(\bar{X}) = \frac{SD}{\sqrt{n}} = \frac{90}{\sqrt{10.000}} = 0.9$

95% konfidensintervall for  $\mu$ :  
95% KI : [723, 727]  
95% CI : [723, 727]