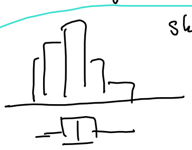
Problemstilling: Å finne referanseverdier for gangfunksjon for unge voksne, og beregne forventningsverdi for gangfunksjon i populasjonen.

Måler dette ved 6-minutters gangtest: 6MWD, som måler hvor langt du går på 6 minutter

Utvalg: $n = 25$ studenter. Observasjonelt design.

Kategoriske / kontinuerlige data?



skjevt / symmetrisk

BESKRIVE OBSERVASJONER

→ Oppsummeringstall: gjennomsnitt er et tall for det typiske i observasjonene

R sier at $\bar{x} = 623$ m

standardavviket er et tall for spredningen i observasjonene. R sier at sd = 83 m

ca 95% av studentene går mellom

$\bar{x} \pm 2 \cdot sd$ m : $(623 \pm \underbrace{2 \cdot 83}_{166})$ m : $\boxed{457\text{m og } 789\text{m}}$

omtrent alle går mellom

$\bar{x} \pm 3 \cdot sd$ m : $(623 \pm 3 \cdot 83)$ m : 374 m og 872 m

VIL BRUKE OBS TIL Å SI NOE OM POPULASJONEN : INFERENS

Populasjonens gjennomsnittlige 6MWD: ukjent. Kaller den for μ ← parameter

Fordi de $n = 25$ er representative for populasjonen, regner vi med at gjennomsnittet i utvalget ligner på gj.snitt i pop:

$\mu \approx \bar{x}$: $\bar{x}$ er et estimat for μ. $\hat{\mu} = \bar{x} = 623$ m

Vi vet at det er en viss usikkerhet her, for vi har bare undersøkt $n = 25$

Kan μ like gjerne være 600? ~~500~~? ~~580~~? 620? ~~780~~?

Estimeringsusikkerheten kalles også standardfeilen, eller standard error (S.E.) på engelsk tallfester dette.

Estimeringsusikkerheten til gjennomsnittet, $\bar{x}$: S.E.$(\bar{x})$

Vi skal vise at $SE(\bar{x}) = \frac{sd}{\sqrt{n}}$

Her: $SE(\bar{x}) = \frac{83}{\sqrt{25}} = 16.6$

Et konfidensintervall formidler estimeringsusikkerheten i form av et intervall med en viss "sikkerhet", kalt konfidensgrad, ofte 95%.

Her er et 95% KI for μ: $\bar{x} \pm 1.96 \cdot S.E(\bar{x})$

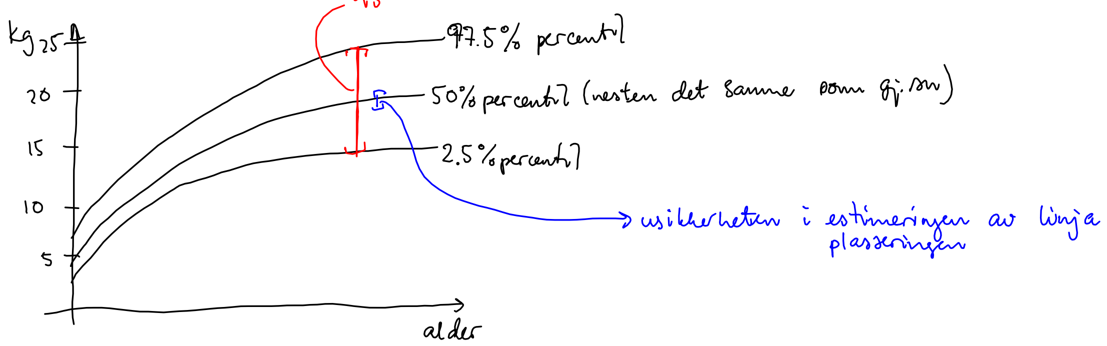
$623 \pm \underbrace{1.96 \cdot 16.6}_{32.5}$

[591, 656]

Konfidens Intervallet [591, 656] handler om μ, altså hvor vi forventer at populasjonsgjennomsn. er. Det forteller ikke noe om enkeltobservasjonene, slik det første intervallet gjorde

Husk: vekstkurver

95% av de enkelte barna har en vekt i dette intervallet.

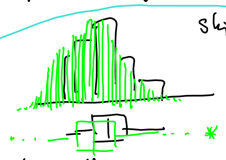


Problemstilling: Å finne referanseverdier for gangfunksjon for unge voksne, og beregne forventningsverdi for gangfunksjon i populasjonen.

Måler dette ved 6-minutters gangtest: 6MWD, som måler hvor langt du går på 6 minutter

Utvalg: n = 25 (10000) studenter. Observasjonelt design.

Kategoriske / kontinuerlige data? → skjevt / symmetrisk



BESKRIVE OBSERVASJONENE

└→ Oppsummeringstall:
gjennomsnitt er et tall for det typiske i observasjonene
R sier at $\bar{x} = 623$ (603) m

standardavviket er et tall for spredningen i observasjonene. R sier at sd = 83 m (89)

ca 95% av studentene går mellom
$\bar{x} \pm 2 \cdot sd$ m : $(623 \pm 2 \cdot 83)$ m : 457 m og 789 m (166)
425 – 781

omtrent alle går mellom
$\bar{x} \pm 3 \cdot sd$ m : $(623 \pm 3 \cdot 83)$ m : 374 m og 872 m
336 – 870

VIL BRUKE OBS TIL Å SI NOE OM POPULASJONEN : INFERENS

Populasjonens gjennomsnittlige 6MWD: ukjent. Kaller den for μ ← parameter

Fordi de n = 25 (10000) er representative for populasjonen, regner vi med at gjennomsnittet i utvalget ligner på gj.snitt i pop: ~~μ~~ $\bar{x}$: $\bar{x}$ er et estimat for μ. $\hat{\mu} = \bar{x} = 623$ (603) m

omtrent slik er det i populasjonen også

Vi vet at det er en viss usikkerhet her, for vi har bare undersøkt n=25 n=10000

Kan μ like gjerne være 600? ~~597~~? ~~500~~? ~~580~~? 620? ~~650~~? ~~750~~? ~~615~~?

Estimeringsusikkerheten kalles også standardfeilen, eller standard error (S.E.) på engelsk. Tallfester dette.

Estimeringsusikkerheten til gjennomsnittet, $\bar{x}$: $S.E.(\bar{x})$

Vi skal vise at $SE(\bar{x}) = \frac{sd}{\sqrt{n}}$

Her: $SE(\bar{x}) = \frac{83}{\sqrt{25}} = 16.6$

Et konfidensintervall formidler estimeringsusikkerheten i form av et intervall med en viss "sikkerhet", kalt konfidensgrad, ofte 95%.

Her er et 95% KI for μ: $\bar{x} \pm 1.96 \cdot S.E(\bar{x})$

$623 \pm 1.96 \cdot 16.6$ (32.5)

[591, 656]

Nå: $S.E.(\bar{x}) = \frac{89}{\sqrt{10000}} = 0.89$

95% KI for μ:
$603 \pm 1.96 \cdot 0.89 \Rightarrow$ [601.3, 604.7]

Problemstilling: Vil beregne Høyres oppslutning.
Meningsmåling: $n = 25$ 1001
5 av disse svarer H.
Andel i utvalg: $\frac{5}{25} = 0.2 = \hat{p}$ 25.4%

20% oppslutning for H.

UTVALG

Andelen i populasjonen som stemmer H: $p \leftarrow$ ukjent parameter
Fordi de $n=25$ er et representativt utvalg, vil andel i utvalg ligne på andel i pop.
Andel i utvalg er et estimat for andel i pop: $\hat{p} = \frac{5}{25} = 0.2$

~~Vet at~~ $n = 25$ er få, og ~~at dette~~ er usikkert.
Kan p like gjerne være 10%? 15%? 45%...

Estimeringsusikkerheten til $\hat{p}$: $S.E.(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.08$ 0.014

95% KI for p: $\hat{p} \pm 1.96 \cdot S.E.(\hat{p}) \rightarrow [0.04, 0.36]$

$[0.227, 0.281]$
22.7% – 28.1%

POP