STK1000: Løsningsforslag Uke 42

2019

Oppgave 6.15

a) Hun glemte å dele standardavviket på $\sqrt{500}$.

b) Vi bruker konfidensintervall til å si noe om *populasjonsgjennomsnittet* (population mean), mens hun sier *utvalgsgjennomsnittet* (sample mean).

c) Et (korrekt) 95 %-konfidensintervall vil i snitt inneholde den sanne parameteren i 95 % av tilfellene det er konstruert. Hun påstår derimot at det er 95 % sannsynlig at den sanne parameteren er i intervallet hennes (**dette er en vesentlig forskjell som er VIKTIG å forstå**).

d) Størrelsen på utvalget påvirker ikke fordelingen til alumni-ratingene. I stedet har vi at gjennomsnittet i utvalget vil gå mot en normalfordeling når størrelsen på utvalget øker.

Oppgave 6.19

For 95% konfidens er feilmarginen $m = 1.96\frac{6.5}{\sqrt{31}} = 2.29$ og et 95%-konfidensintervall for μ er 13.2 ± 2.29 som gir intervallet [10.91, 15.49].

Oppgave 6.20

Her er er 95%-konfidensintervall for μ gitt av $33.4 \pm 1.96\frac{19.6}{\sqrt{31}}$ som gir intervallet [26.50, 40.30].

Oppgave 6.30*

a)

Vi vil finne $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, som kan gjøre i R med

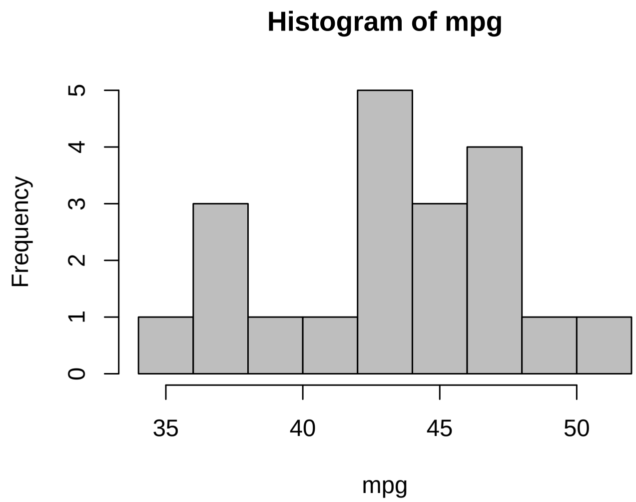
```
sigma = 3.5
n = 20
sigma_xb = sigma / sqrt(n)
sigma_xb
```

```
## [1] 0.7826238
```

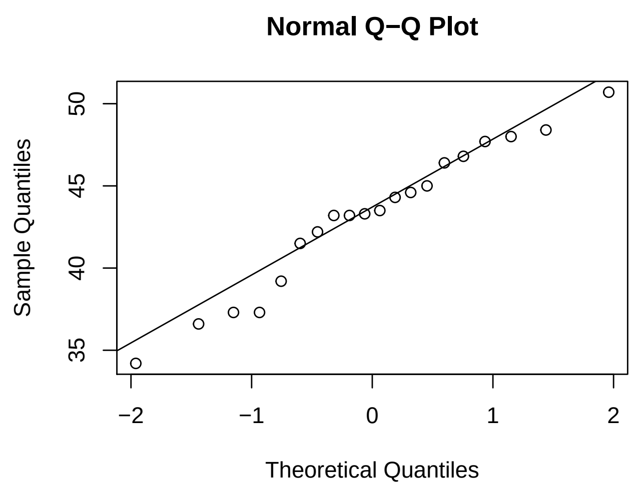
b)

Vi lager et histogram, et QQ Plot og skriver ut kvantiler og gjennomsnitt. Vi finner ikke noen tydelige tegn på skjevheter eller ikke-normalitet. Derfor virker det fornuftig å videre anta en normalfordeling.

```
mpg = c(41.5,50.7,36.6,37.3,34.2,45,48,43.2,47.7,42.2,43.2,44.6,48.4,46.4,46.8,39.2,37.3,
        43.5,44.3,43.3)
hist(mpg, col = 'gray')
```




```
qqnorm(mpg)
qqline(mpg)
```




```
summary(mpg)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   34.20   40.92   43.40   43.17   46.50   50.70 
```

c)

Her bruker vi funksjonen som er oppgitt på fagsiden og lager både et 95%- og et 99%-konfidensintervall for μ.

```
konfidens_intervall = function(x, sigma, level) {
  n=length(x)
  z=qnorm(1-(1-level)/2)
  return(c(mean(x)-(z*sigma/sqrt(n)), mean(x)+(z*sigma/sqrt(n))))
}

konfidens_intervall(mpg, sigma, 0.95) # 95% intervall
```

```
## [1] 41.63609 44.70391
```

```
konfidens_intervall(mpg, sigma, 0.99) # 97% intervall
```

```
## [1] 41.15409 45.18591
```

Vi ser at 99%-intervallet er bredere enn det med 95%. Dette er fordi intervallene skal i 95% og 99% av tilfellene dekke den sanne populasjonsparameteren. For å være sikrere på å dekke parameteren trenger man et større intervall. Derfor vil et 90%-konfidensintervall være smalere enn disse to. Det kan vi få bekreftet med

```
konfidens_intervall(mpg, sigma, 0.90)
```

```
## [1] 41.8827 44.4573
```

Oppgave 6.33

$n \geq \left(\frac{1.96 \cdot 6.5}{1.5}\right)^2 = 72.14$. Derfor trenger vi et utvalg av størrelse $n = 73$.

Oppgave 6.36

a) $10.0023 \pm 2.57 \cdot \frac{0.0002}{\sqrt{6}} = (10.00209, 10.00251)$.

b) Siden intervallet i a) ikke dekker den sanne verdien 10, virker det ikke som vekten er nøyaktig.

c) Vi snur om på uttrykket i a) og får $n \geq \left(\frac{2.57 \cdot 0.0002}{0.0001}\right)^2 = 26.4$, så $n = 27$.