

STK1000: Løsningsforslag Uke 43

2023-10-23

Oppgave 6.62

Dette er faktisk et veldig vanlig problem inne forskning og poenget her er at vi skal tenke gjennom nedsidene. Vi gir her noen eksempler på hva vi kan diskutere:

Når negative resultater ikke publiseres vil det gi et inntrykk av at alle studier foretatt på antidepressive stoffer gir positive utfall. Dette er farlig siden man ikke får et balansert bilde.

Vider kan vi også lære av de negative studiene hva som ikke fungerer, og det kan hindre andre forskere i å bruke tid på å repetere forsøk (som ikke har blitt publisert).

Å vri negative resultater for å få en positiv vinkling er selvfølgelig ikke bra. Det er ikke spesifisert hvordan fremstillingen er vridd, men man forstår at dette kan være problematisk siden leseren kan trekke gale konklusjoner fra artiklene.

Oppgave 6.76

Vi bruker altså $\frac{\alpha}{k} = \frac{0.05}{6} = 0.008333$. Da er bare p-verdiene på 0.002 og < 0.001 signifikante.

Oppgave 7.6

- Vi har frihetsgrader $df = n - 1 = 21$.
- I Tabell D ser vi at de to nærmeste t-verdiene er 2.189 og 2.518.
- P-verdiene som svarer til t-verdiene i b) er 0.02 og 0.01.
- Fra c) ser vi at p-verdien som svarer til $t = 2.24$ er mellom 0.01 og 0.02. Vi har derfor signifikante resultater på 5% nivå, men ikke på 1% nivå.
- Her bruker vi R og ser at den sanne p-verdien stemmer bra med svarene over.

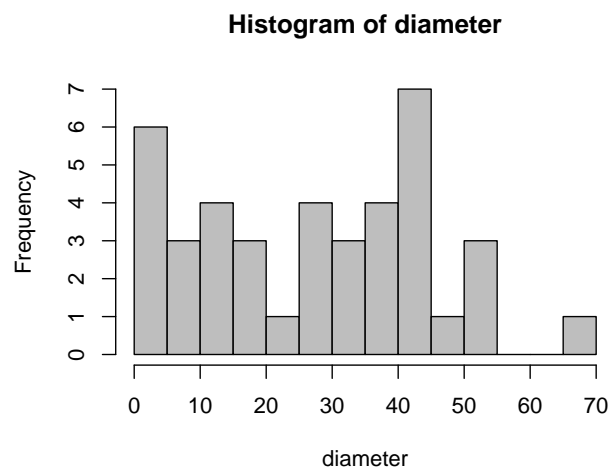
```
t = 2.24
df = 21
1 - pt(t, df)
```

```
[1] 0.01801848
```

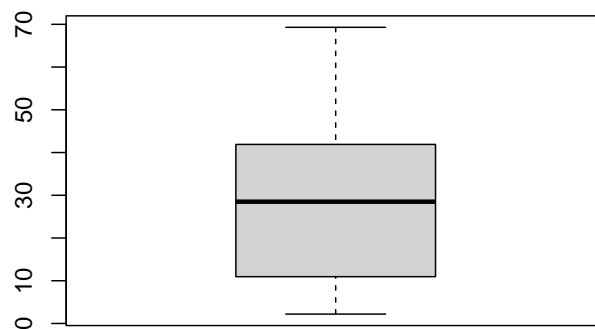
Oppgave 7.19(R)

a)

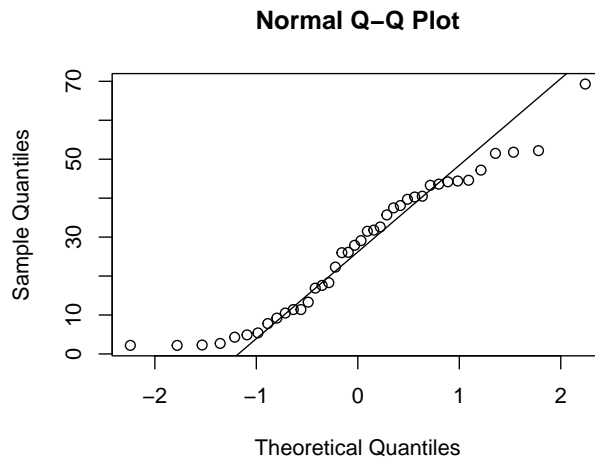
```
data = read.csv('../..//ips10e_csv_data_sets/ips10e_ch7_csv_data_sets/EX07-019PINES.csv')
diameter = data$Diameter
hist(diameter, col = 'gray', breaks = 10)
```



```
boxplot(diameter)
```



```
qqnorm(diameter)
qqline(diameter)
```



Fra histogrammet kan det se ut som fordelingen har to topper. Videre ser vi at boksplottet ikke er symmetrisk og at punktene i QQ Plottet ikke følger linjen veldig bra. Det ser derfor ikke ut til at diametrene er normalfordelt. Videre kan det se ut som vi har en outlier til høyre i histogrammet.

b)

Kanskje. Vi konkluderte i a) med at fordelingen ikke er normal, men vi har trolig tilstrekkelig med observasjoner til at gjennomsnittet er noenlunde normalfordelt (sentralgrenseteoremet).

c)

Hvis vi bare er interessert i et konfidensintervall kan vi gjøre det med

```
t.test(diameter, conf.level = 0.95)$conf.int
```

```
[1] 21.6274 32.9526
attr(,"conf.level")
[1] 0.95
```

Men siden vi også blir bedt om å finne “margin of error” kan vi like gjerne bruke den til å lage intervallet

```
x_bar = mean(diameter)
x_bar # gjennomsnitt
```

[1] 27.29

```
sigma = sd(diameter)
n = length(diameter)
df = n - 1
alpha = 0.05
t = qt(1 - alpha/2, df)
m = t * sigma / sqrt(n)
m # margin of error
```

[1] 5.662603

```
c(x_bar - m, x_bar + m) # konfidensinterval
```

[1] 21.6274 32.9526

d)

Ulike resonnementer vil legge vekt på ulike faktorer, og det viktigste er ikke konklusjonen men et velbegrunnet resonnement som en leser kan ta stilling til. Resultatene kan i prinsippet sies å gjelde en populasjon som vi kan se utvalget som et enkelt tilfeldig utvalg fra. I hvilken grad vi kan strekke dette til 'lignende trær' og 'samme område', spørs selvfølgelig på hvordan vi avgrensar disse begrepene. Jordsmonn, solforhold og andre faktorer som kan påvirke treets utvikling bør være en del av evalueringa.

Oppgave 2 fra avsluttende eksamen 2015 OG oppgave 1 fra avsluttende eksamen 2014

Løsningsforslag til eksamensoppgavene er på emnets semesterside

<https://www.uio.no/studier/emner/matnat/math/STK1000/oppgaver/losningsforslag/>