

# STK1000: Løsningsforslag Uke 36

2019

## Oppgave 1.120 (8. utgave)

a) Vi standardiserer normalfordelte variabler  $x$  med formelen

$$z = \frac{x - \mu}{\sigma},$$

hvor  $\mu$  og  $\sigma$  er forventning og standardavvik for fordelingen. Vi får da

$$(-1.0, 2.1, -1.8, 0.4, 0.1, 2.6, -0.8, -1.7, 0.8, -0.1)$$

b)  $P(Z \geq z) = 0.15$  gir at  $P(Z < z) = 1 - 0.15 = 0.85$ . Fra Tabell A i boken finner vi at dette gir  $z = 1.04$ . Så grensen for A i form av en standardisert poengsum er 1.04.

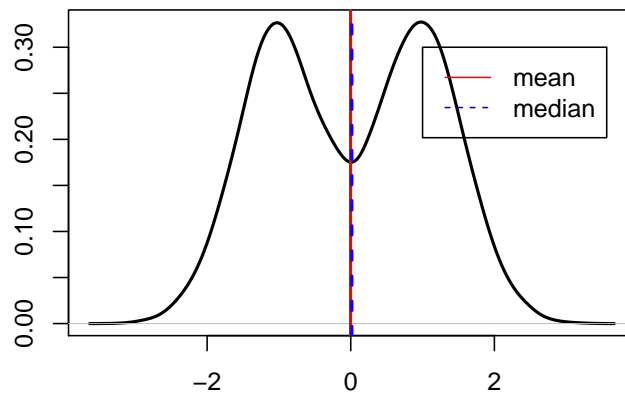
c) Vi må nå reversere standardiseringen for å få det tilbake på poeng-skalaen. Dette kan vi gjøre med å snu om på uttrykket i a)

$$x = z \cdot \sigma + \mu = 82.4.$$

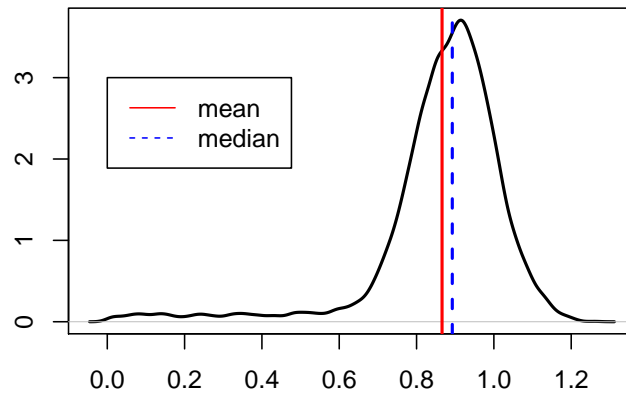
Så grensen for A er 82.4 poeng. Derfor er det to av studentene (93 og 98) som fikk A.

## Oppgave 1.101

a)

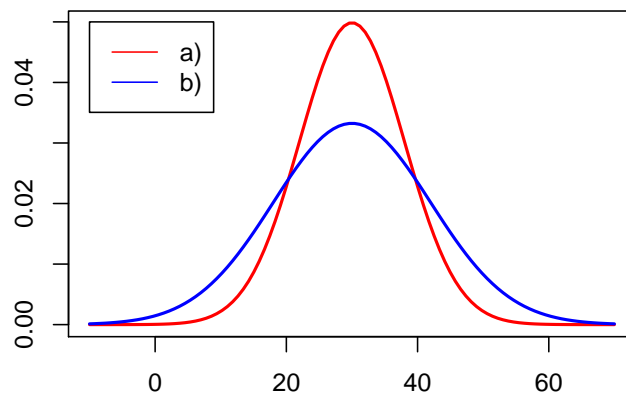


b)



### Oppgave 1.102

a) og b)

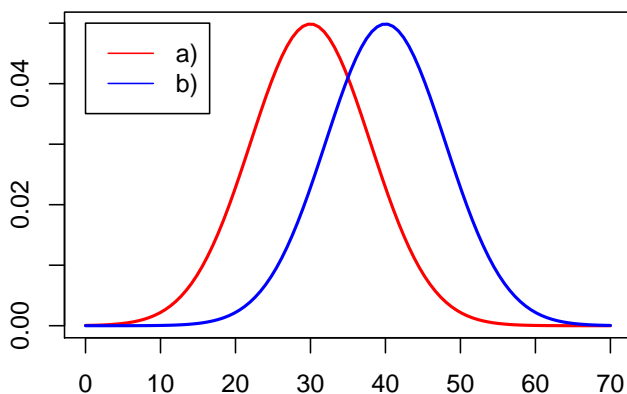


c)

Når vi forandrer standardavviket men holder forventningsverdien konstant ser vi at kurven blir bredere og smalere, men beholder samme midtpunkt.

## Oppgave 1.103

a) og b)



c)

Vi ser nå at hvis vi forandrer forventningsverdien mens vi holder standardavviket konstant, forflytter vi kurven langs x-aksen, mens kurven beholder samme form.

## Oppgave 1.110

a)

68-95-99.7 regelen tilsvarer avvik på 1, 2 og 3 standardavvik fra gjennomsnittet. Så for gjennomsnitt  $\mu$  og standardavvik  $\sigma$  kan vi regne ut  $\mu \pm \sigma \cdot \{1, 2, 3\}$ . Vi får da at

- 68 % av kvinner sier mellom 7856 og 20738 ord per dag.
- 95 % av kvinner sier mellom 1415 og 27179 ord per dag.
- 99.7 % av kvinner sier mellom -5026 og 33620 ord per dag (merk at nedre grense er negativ).

b)

Det er selvfølgelig ikke mulig å si et negativt antall ord, så fordelingen er ikke virkelig en normalfordeling.

c)

Tilsvarende får vi

- 68 % av menn sier mellom 5004 og 23116 ord per dag.
- 95 % av menn sier mellom -4052 og 32172 ord per dag.
- 99.7 % av menn sier mellom -13108 og 41228 ord per dag.

Vi får her også et negativt antall ord, som tilsier at fordelingen ikke er helt normal.

d)

Vi ser at den gjennomsnittlige kvinne sier litt flere ord en den gjennomsnittlige mann i løpet av en dag, men forskjellene er små. Siden menn har større standardavvik enn kvinner kan det bety at menn tilsier det at det

er større variasjon i antall ord en mann sier. Det er derfor både at menn sier flere ord per dag enn kvinner og vi kan ikke konkludere noe videre.

### Oppgave 1.112

- Figuren viser ikke noen y-akse, men vi vet at alle alle tetthetsfunksjoner skal ha areal lik 1. Derfor må høyden på funksjonen være 1.
- $1 - 0.44 = 0.56$ , altså ligger 56 % av observasjonene over 0.44.
- $0.70 - 0.44 = 0.26$ , altså ligger 26 % av observasjonene mellom 0.44 og 0.70.

### Oppgave 1.114

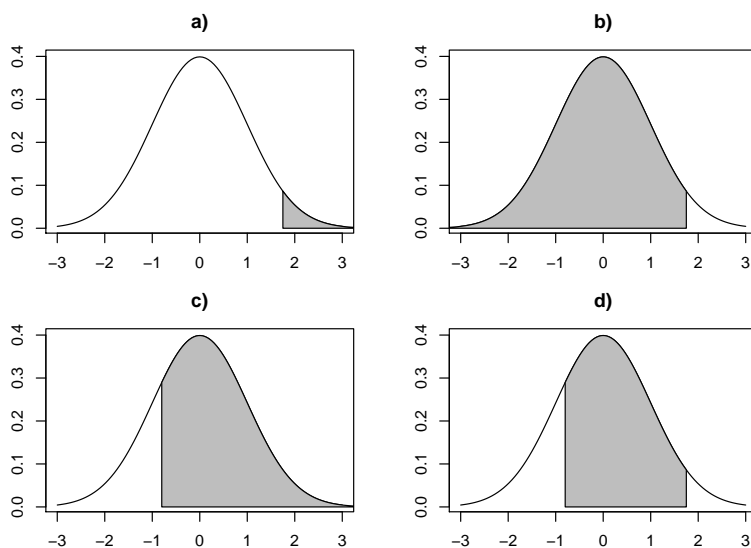
For en uniform fordeling mellom 0 og 1 er forventning og median begge 0.5. Kvartilene er  $Q_1 = 0.25$  og  $Q_2 = 0.75$ .

### Oppgave 1.118

Vi kan enten bruke Tabell A fra boken eller kommandoen `pnorm` i R for å regne ut  $P(Z < z)$ .

- $P(Z > 1.75) = 1 - P(Z < 1.75) = 0.0401$ .
- $P(Z < 1.75) = 0.9599$ .
- $P(Z > -0.80) = 1 - P(Z < -0.80) = 0.7881$ .
- $P(-0.80 < Z < 1.75) = P(Z < 1.75) - P(Z < -0.80) = 0.7481$ .

Vi har skissert arealene i figuren under.



### Oppgave 1.121

- Vi bruker Tabell A i boken og ser at  $z = 1.175$  passer.
- $P(Z > z) = 0.12$  gir oss at  $P(Z < z) = 1 - 0.12 = 0.88$ , så vi får også her  $z = 1.175$ .

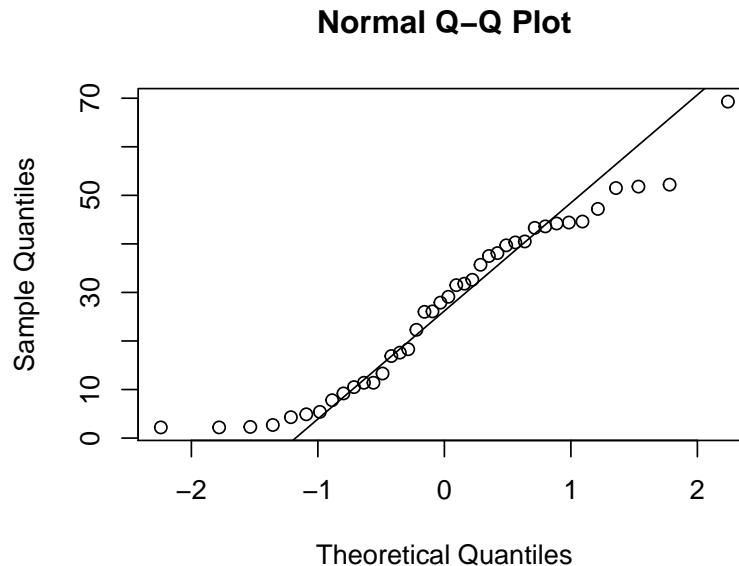
## Oppgave 1.134

- a) Vi må først standardisere verdien  $z = \frac{40-55}{15.5} = -0.97$ . Fra tabell A får vi da at  $P(Z < -0.97) = 0.1660$ , så 16.6 % av kvinner har lave HDL-verdier.
- b)  $z = \frac{60-55}{15.5} = 0.32$ . Fra tabell A har vi da at  $P(Z \geq 0.32) = 1 - P(Z < 0.32) = 0.3745$ . Altså har 37.45 % av kvinner har beskyttende verdier av HDL.
- c) Her bruker vi resultatene fra a) og b).  $P(-0.97 < Z < 0.32) = P(Z < 0.32) - P(Z < -0.97) = (1 - 0.3745) - 0.1660 = 0.4595$ . Altså har 45.95 % av kvinner har HDL-verdier som ligger mellom 40 og 60 mg/dl.

## Oppgave 1.142(R)

Vi laster inn diameterdataene og lager et Q-Q plott med følgende kode

```
data = read.csv('CSV/Chapter 1/EX01-142PINES.csv')
x = data$Diameter
qqnorm(x)
qqline(x)
```



Vi ser at punktene ikke følger linjen så bra, så fordelingen er nok ikke helt normal. Det er indikasjoner på høyreforskyvning, men det er ikke veldig tydelig.