6. Van steekproef naar populatie.

Opgave 6.1 **Steekproeven**

a Met slechts drie mannen zegt dit natuurlijk bijzonder weinig over de hele populatie.

b Waarschijnlijk is dat al een iets betere schatting, als de steekproef volgens de regels is uitgevoerd.

c Hoe groter de steekproef hoe betrouwbaarder meestal het resultaat.

d De gemiddeldes liggen dicht bij elkaar maar de standaarddeviatie wordt kleiner als het aantal samples per steekproef groter wordt.

e De nauwkeurigheid neemt dus toe met het aantal samples.

f De beste schatting van de gemiddelde lengte van de populatie mannen boven de 20 is 180,7 cm

g Het 99 % betrouwbaarheidsinterval hoort bij een kans van 0,995 (99% ligt tussen 0,005 en 0,995); dit levert een Z-waarde van 2,575.

$180{,}7 - 2{,}575 \times 1{,}5 < \mu < 180{,}7 + 2{,}575 \times 1{,}5$

176,8 cm $< \mu <$ 184,6 cm

Het 99 % betrouwbaarheidsinterval is dus *groter* dan het 95 % betrouwbaarheidsinterval

h $180{,}7 - 1{,}96 \times 2{,}5 < \mu < 180{,}7 + 1{,}96 \times 2{,}5$

dus bij 10 samples 175,8 cm $< \mu <$ 185,6 cm

en bij 50 samples 177,8 cm $< \mu <$ 183,6 cm

bij een grotere steekproef wordt de schatting nauwkeuriger

Opgave 6.2 **Standaardfout en populatie**

a $SE = \dfrac{\sigma}{\sqrt{n}} \Rightarrow \sigma = SE \times \sqrt{n}$

10 × 3 samples	$\sigma = SE \times \sqrt{n} = 4{,}2 \times \sqrt{3} = 7{,}3$
10 × 10 samples	$\sigma = SE \times \sqrt{n} = 2{,}5 \times \sqrt{10} = 7{,}9$
10 × 25 samples	$\sigma = SE \times \sqrt{n} = 1{,}5 \times \sqrt{25} = 7{,}5$

Ze verschillen heel weinig. De laatste zal wel het meest betrouwbaar zijn.

b Het klopt heel behoorlijk. Als je alles herhaalt komt er toch ook niet steeds weer hetzelfde uit.

Opgave 6.3 **Kan het niet met wat minder steekproeven?**

a Meting zoutgehalte $\bar{x}$ = 15,1 mg/L en s = 0,2 mg/L

σ = 2,5 % van 15,1 mg/L = 0,38 mg/L

b Bereken de standaardfout *SE*.

$SE = \dfrac{\sigma}{\sqrt{n}} \Rightarrow SE = \dfrac{0{,}38}{\sqrt{3}} = 0{,}22$

c Bereken het 95 % betrouwbaarheidsinterval voor de werkelijke waarde van het zoutgehalte.

$15{,}1 - 1{,}96 \times 0{,}22 < \mu < 15{,}1 + 1{,}96 \times 0{,}22$

$14{,}67 \text{ mg/L } < \mu < 15{,}53 \text{ mg/L}$

d Je kunt de steekproef groter maken dus meer metingen aan hetzelfde monster doen.

e Nee, want het 99 % betrouwbaarheidsinterval is groter dan het 95 % interval. Het aantal metingen verandert namelijk niet door een andere berekening. De steekproef blijft even (on)nauwkeurig.

f Reken uit hoe groot de steekproef minstens moet zijn om een afwijking van maximaal 2 % te krijgen.
2 % afwijking betekent 2 % van 15,1 = 0,302
dus 1,96 × *SE* = 0,302

$$SE = \frac{0{,}302}{1{,}96} = 0{,}15$$

$$SE = \frac{\sigma}{\sqrt{n}} \Rightarrow \sqrt{n} = \frac{\sigma}{SE} = \frac{0{,}38}{0{,}1541} = 2{,}47$$

$$n = 2{,}47^2 = 6{,}1$$

afgerond *n* =7

je moet dus nog 4 metingen extra doen.

Opgave 6.4 Simulatie van steekproeven uit een populatie

a 5

b afwijking = 16 – 13,18 = 2,82
afwijking = $\frac{2{,}82}{16} \times 100\ \% = 17{,}6\ \%$

c 8 × 5 = 40

d afwijking = $\frac{0{,}10}{16} \times 100\ \% = 0{,}625\ \%$

het gemiddelde van de steekproef komt steeds dichter bij die van de populatie te liggen

d Die is 2,38

f $SE = \frac{\sigma}{\sqrt{n}} \Rightarrow \sigma = SE \cdot \sqrt{n} = 2{,}38 \times \sqrt{5} = 5{,}32$

Hoe meer metingen je doet, hoe beter dit gaat kloppen

Opgave 6.5 Betekenis van het betrouwbaarheidsinterval

a 10 steekproeven

b 9 van de 10 dus 90 %

c 15 van de 20 dus 75 %

d --

Opgave 6.6 **Schatting van het populatiegemiddelde bij een kleine steekproef**

a Tabel: 95%; tweezijdig; n =5 dus v = 4 → t = 2,78

b

$$\bar{x} - t \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t \cdot \frac{s}{\sqrt{n}}$$

$$0,91 - 2,78 \times \frac{0,06}{\sqrt{5}} < \mu < 0,91 + 2,78 \times \frac{0,06}{\sqrt{5}}$$

BI: 0,84 g/kg < μ < 0,98 g/kg

c De maximale waarde van 1,0 kg ligt boven het betrouwbaarheidsinterval, dus het gehalte is niet te hoog

Opgave 6.7 **Schatting van het populatiegemiddelde bij een grote steekproef**

a n =200, dus v = n – 1, tabel t = 1,64

$$\bar{x} - t \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t \cdot \frac{s}{\sqrt{n}}$$

$$\bar{x} - 1{,}64 \times \frac{118}{\sqrt{200}} < \mu < \bar{x} + 1{,}64 \times \frac{118}{\sqrt{200}}$$

$$1594 - 14 < \mu < 1594 + 14$$

€ 1580 < μ < € 1608

b

$$\bar{x} - 1{,}64 \times \frac{118}{\sqrt{190}} < \mu < \bar{x} + 1{,}64 \times \frac{118}{\sqrt{190}}$$

$$1594 - 14 < \mu < 1594 + 14$$

€ 1580 < μ < € 1608 en dat blijft door de afronding hetzelfde

Opgave 6.8 **BI bij controles - Koloniegetalbepaling**

a gemiddelde

$$\bar{x} - t \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t \cdot \frac{s}{\sqrt{n}}$$

$$75 - 2{,}78 \times \frac{12}{\sqrt{5}} < \mu < 75 + 2{,}78 \times \frac{12}{\sqrt{5}}$$

BI: 60 KVE/g < μ < 90 KVE/g

b Ja, het hele interval ligt onder de 100 dus het kiemgetal is niet te hoog

Opgave 6.9 **Controle - Zout in mineraalwater**

a Tabel: 95%; tweezijdig; n =25 dus v = 24 → t = 2,06
gemiddelde:

$$\bar{x} - t \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t \cdot \frac{s}{\sqrt{n}}$$

$$130-2{,}06\times\frac{12}{\sqrt{25}}<\mu<130+2{,}06\times\frac{12}{\sqrt{25}}$$

BI: 125 mg/L < μ < 135 mg/L

b zoals al eerder vermeld kan bij een redelijk grote steekproef de standaarddeviatie gebruikt worden als schatting van de standaarddeviatie van de populatie.

schatting $\sigma_n \approx \sigma_{n-1}$ = 12 mg/L

c

$$Z=\frac{x-\mu}{\sigma}=\frac{150-130}{12}=1{,}67$$

Z-tabel: $P(Z<1{,}67)$ = 0,9525

$P(Z>1{,}67)$ = 1 – 0,9525 = 0,0475

Dus 4,75 % van de flessen zal waarschijnlijk meer dan 150 mg/L zout bevatten

Opgave 6.10 **Statistisch significant of praktisch significant?**

a Middel B geeft de grootste gemiddelde gewichtsafname, maar ook de grootste onzekerheid, de afname kan zelfs negatief zijn, dat betekent dus sommige gebruikers een *gewichtstoename* kunnen verwachten. Van middel A wordt in ieder geval iedereen (95% betrouwbaar) lichter.

b Wie graag een gok waagt neemt middel B, als je op zekerheid speelt neem je A.