

2 Statistiek

Voorkennis Procenten en frequenties

Bladzijde 54

- 1 a $121 - 57 = 64$, dus $\frac{64}{121} \times 100\% \approx 52,9\%$.
 b $0,003 \times 8700 = 26,1$
 c $0,913 \times 3,2 \approx 2,9$ miljoen
 d totaal = $\frac{8,3}{0,126} \approx 65,9$ miljoen
- 2 a De procentuele toename is $\frac{7550 - 2521}{2521} \times 100\% \approx 199,5\%$.
 b Het aantal inwoners in Afrika in 2017 was $5,68 \times 221 \approx 1255$ miljoen.
 c De procentuele toename is $\frac{4504 - 1402}{1402} \times 100\% \approx 221,3\%$.
 In de periode 1950-2017 is de procentuele toename in Azië inderdaad minder dan die in Afrika.
 d $\frac{547}{2521} \times 100\% \approx 21,7\%$, dus in 1950 woonde 21,7% van de wereldbevolking in Europa.
 $\frac{742}{7550} \times 100\% \approx 9,8\%$, dus in 2017 woonde 9,8% van de wereldbevolking in Europa.

Bladzijde 55

- 3 a De totale frequentie $11 + 3 + 3 + 2 + 5 = 24$.
 De relatieve frequentie bij Schalkhaar is $\frac{3}{24} \times 100\% = 12,5\%$.
 b De relatieve frequentie van leerlingen uit Diepenveen is 20%.
 totaal = $\frac{6}{0,2} = 30$
 Er zitten 30 leerlingen in klas 4VB.
- 4 a totaal = $\frac{1310}{0,195} \approx 6718$. Er waren 6718 deelnemers in het voortgezet onderwijs.
 b Het verschil is $21,5\% - 18,3\% = 3,2\%$.
 Dus 3,2% van het totaal is 51.
 totaal = $\frac{51}{0,032} \approx 1594$
 Er waren 1594 deelnemers in het basisonderwijs.

2.1 De statistische cyclus

Bladzijde 56

- 1 De logische volgorde is 2, 4, 5, 1, 3.

Bladzijde 57

- 2 De juiste volgorde:
 activiteit C: fase 1
 activiteit D: fase 1
 activiteit A: fase 2
 activiteit G: fase 3
 activiteit E: fase 3
 activiteit B: fase 4
 activiteit F: fase 4

Bladzijde 58

- 3 fragment A: fase 3
fragment B: fase 4
fragment C: fase 1
fragment D: fase 2
- 4 a 'Hoeveel jongeren' is niet specifiek genoeg en 'gezond eten' is vaag geformuleerd.
b Dit zijn twee vragen ineen.
c Er wordt verondersteld dat jongeren die veel gamen minder sociaal zijn, maar dat is nog de vraag.

Bladzijde 60

- 5 a Causaal verband. Het effect wordt onderzocht en niet alleen of er een relatie is.
b Geen causaal verband. Als er een verband wordt gevonden is niet duidelijk welke van de twee het gevolg is van de andere.
c Causaal verband. Men wil nagaan of het verbieden van roken op het schoolplein tot gevolg heeft dat het roken wordt teruggedrongen.
d Geen causaal verband. Er wordt niet onderzocht of het ene het andere tot gevolg heeft.
- 6 a Nee, waarschijnlijk weet ze al genoeg van het onderwerp.
b *
c *
d *

Bladzijde 61

- 7 a Deelvraag 1: Welk percentage van de scholieren eet voldoende fruit volgens de richtlijnen van het voedingscentrum?
Deelvraag 2: Welk percentage van de scholieren krijgt dagelijks te veel calorieën binnen volgens de richtlijnen van het voedingscentrum?
b Deelvraag 1: Treedt er gewichtsverlies op bij jongeren die minstens vijf keer per week sporten?
Deelvraag 2: Welk percentage van de jongeren doet vijf keer in de week aan sport?
c Deelvraag 1: Hoeveel procent van de kinderen van ouders met een minimuminkomen doet aan paardrijden?
Deelvraag 2: Hoeveel kost paardrijden op jaarbasis?
- 8 Vraag 5a:
Deelvraag 1: Welk percentage van de jongeren gamet overmatig?
Deelvraag 2: Is er een verschil in het percentage jongeren dat overmatig gamet op de verschillende schoolniveaus?
Vraag 5b:
Deelvraag 1: Hoeveel uur per dag besteden jongeren gemiddeld aan sociale media?
Deelvraag 2: Hoeveel uur per dag besteden jongeren gemiddeld aan gamen?
Vraag 5c:
Deelvraag 1: Welk percentage van de jongeren rookt?
Deelvraag 2: Welk percentage van de rokende jongeren doet dat ook op het schoolplein?
Vraag 5d:
Deelvraag 1: Welk percentage van de scholieren in de vierde klas rookt?
Deelvraag 2: Welk percentage van de leerlingen in 4 vwo rookt?

- 9 a Onderzoeksvraag 1: B
Onderzoeksvraag 2: D
Onderzoeksvraag 3: A
Onderzoeksvraag 4: C

b Onderzoeksvraag 1:

Deelvraag 1: Wat is de ontwikkeling van de huizenverkoop in de afgelopen vijf jaar?

Deelvraag 2: Is er een verschil in de prijsontwikkeling tussen verschillende woningtypen?

Onderzoeksvraag 2:

Deelvraag 1: Heeft het beperken van de hypotheekrenteaftrek invloed op de woningprijzen?

Deelvraag 2: Zijn er op de lange duur voldoende woningen om in de totale woonwens te voorzien?

Onderzoeksvraag 3:

Deelvraag 1: Hoe was de woningmarkt in 2000?

Deelvraag 2: Hoeveel huizen zijn er bijgekomen sinds 2000?

Onderzoeksvraag 4:

Deelvraag 1: Hoe heeft de woningmarkt zich ontwikkeld in die regeerperiode?

Deelvraag 2: Welke onvoorziene omstandigheden zijn er geweest?

Bladzijde 62

- 10** **a** Dat is veel te veel werk.
b Dan pas geeft het onderzoek informatie over de hele bevolking.

Bladzijde 63

- 11** **a** De examenleerlingen van je school.
b Alle weekbladen.
c Alle inwoners van Appelscha.
d Hooikoortspatiënten.
- 12** **a** De leeftijdsopbouw in de figuur is niet gelijkmatig. Je kunt dus niet van elke leeftijd evenveel mensen ondervragen om een representatieve steekproef te krijgen.
b Ja, dat kan als de populatie ook alleen maar uit mannen van pensioengerechtigde leeftijd bestaat.
c Ja, dat kan als het aantal onderzochte aspecten niet groot is en de populatie niet zo divers.
- 13** **a** Er zijn grote groepen mensen die zelden of nooit op zaterdagmiddag in de stad (kunnen) rondlopen, dus de steekproef is niet aselekt.
b Wie geen auto heeft of zelden in de ochtendspits in de file zit, heeft weinig kans om in de steekproef voor te komen, dus de steekproef is niet aselekt.
c De steekproef van 12 personen is te klein en niet aselekt: iemand in Drenthe heeft meer kans dan iemand in Zuid-Holland om in de steekproef voor te komen.
d De bezoekers van een natuurgebied vormen – zeker als het gaat over het milieuprobleem – geen goede afspiegeling van de hele bevolking, dus de steekproef is niet aselekt.
e Mensen zonder smartphone kunnen niet in de steekproef voorkomen, dus de steekproef is niet aselekt.

2.2 Data verzamelen

Bladzijde 65

- 14** **a** leeftijd, nationaliteit, geslacht, lengte, tak van sport
b leeftijd, lengte
c *

Bladzijde 66

- 15** **a** 1 Suggestieve vraagstelling.
2 Vraagstelling met dubbele ontkenning. Bovendien suggestief.
3 Onduidelijke vraag. Wat is 'veel'?
4 Doet een beroep op de herinnering en niet alle antwoordmogelijkheden zijn opgenomen.

- b** 1 Ervaart u overlast van burens?
I nooit II soms III regelmatig IV vaak
- 2 Vindt u het verstandiger om een huis te kopen of te huren?
I huren II kopen
- 3 Hoeveel uren heeft u gisteren tv gekeken?
I 0 tot 1 uur II 1 tot 2 uur III 2 tot 3 uur IV 3 tot 5 uur V 5 uur of meer
- 4 Op hoeveel dagen heeft u vorige week de website NU.nl bezocht?
I op 0 dagen II op 1 of 2 dagen III op 3 of 4 dagen V op 5 of meer dagen

- 16** a kwalitatief
b kwantitatief
c kwalitatief
d kwalitatief

Bladzijde 67

- 17** a discreet
b continu
c continu
d discreet
e continu
f discreet

- 18** Als de variabele is weergegeven in gehele jaren dan is het een discrete variabele. Zijn alle leeftijden mogelijk, dus bijvoorbeeld ook 17,3434 jaar, dan is de variabele continu te noemen.

- 19** a $\frac{754}{1300} \times 100\% = 58\%$
b $\frac{12}{663} \times 100\% \approx 1,8\%$
 $\frac{14}{637} \times 100\% \approx 2,2\%$
c De variabele geslacht is kwalitatief, want het is een eigenschap. De variabele eenzaamheid is kwalitatief, want het verschil tussen twee opeenvolgende getallen heeft geen betekenis
d Daar is deze wel van belang, want het gaat van niet eenzaam tot zeer ernstig eenzaam.

Bladzijde 69

- 20** woonplaats: nominaal
reisafstand naar school in km: ratio
geboortejaar: interval
aantal broers/zussen: ratio
lid van een sportclub: nominaal
aantal uur per week bijbaan: ordinaal

- 21** a kwalitatief: nominaal en ordinaal
kwantitatief: interval en ratio
b nominaal: bestemming
ordinaal: aantal sterren van het hotel
interval: gemiddelde jaartemperatuur op plaats van bestemming
ratio: kosten van de reis

- 22** Op grond van de peiling zou je $\frac{205}{1200} \times 100\% \approx 17,1\%$ van de stemmen op de VVD verwachten, dus 21,3% is meer dan je op basis van de peiling zou mogen verwachten.

Bladzijde 70

populatieproportie $p = \frac{120}{10000} = 0,012$ en de steekproefproportie $\hat{p} = \frac{4}{250} = 0,016$.

- 24 Er geldt $p = 0,418$ en er zijn 627 mensen met bloedgroep A.

$0,418 = \frac{627}{\text{aantal patiënten in het ziekenhuis}}$, dus het aantal patiënten in het ziekenhuis

is $\frac{627}{0,418} = 1500$.

Er geldt $\hat{p} = 0,45$ en er zijn 27 patiënten met bloedgroep A in de steekproef.

$0,45 = \frac{27}{\text{aantal patiënten in de steekproef}}$, dus het aantal patiënten in de steekproef is $\frac{27}{0,45} = 60$.

- 25 Voor het kenmerk 'personeelslid is man en heeft kinderen' geldt $p = 0,3 \cdot 0,6 = 0,18$.
 Voor het kenmerk 'personeelslid is vrouw en heeft kinderen' geldt $p = 0,7 \cdot 0,5 = 0,35$.
 Voor het kenmerk 'personeelslid heeft kinderen' geldt $p = 0,18 + 0,35 = 0,53$.
 Dus voor kenmerk 'personeelslid met kinderen is man' geldt $p = \frac{0,18}{0,53} \approx 0,34$.

2.3 Centrum- en spreidingsmaten

Bladzijde 72

- 26 a nominaal
 b Nee, want je hebt geen kwantitatieve variabele.
 c De modus, want daarvoor hoef je alleen maar te weten wat het vaakst voorkomt.

Bladzijde 73

- 27 a De totale frequentie is $3 + 3 + 7 + 0 + 3 + 2 + 2 = 20$, dus het onderzoek duurde 20 dagen.
 b gemiddelde $= \frac{3 \cdot 1 + 7 \cdot 2 + 3 \cdot 4 + 2 \cdot 5 + 2 \cdot 6}{20} = 2,55$
 mediaan $= \frac{10^{\text{e}} \text{ getal} + 11^{\text{e}} \text{ getal}}{2} = \frac{2 + 2}{2} = 2$
 modus $= 2$, want 2 heeft de grootste frequentie.

Bladzijde 74

- 28 a gemiddelde $= \frac{5 \cdot 1 + 5 \cdot 2 + 3 \cdot 3 + 3 \cdot 4 + 2 \cdot 5 + 1 \cdot 32}{20} = 3,9$
 b mediaan $= \frac{10^{\text{e}} \text{ getal} + 11^{\text{e}} \text{ getal}}{2} = \frac{2 + 2}{2} = 2$
 Dus de klas kan op de mediaan wijzen, want die is in beide klassen gelijk.

- 29 a Op volgorde zetten geeft 5,9; 6,3; 6,6; 6,8; 7,0; 7,0; 7,2; 8,0; 8,4.
 9 cijfers, dus mediaan $= 5^{\text{e}} \text{ cijfer} = 7,0$.
 b Op volgorde zetten geeft 5,7; 5,9; 6,3; 6,6; 7,0; 7,1; 7,3; 7,4.
 8 cijfers, dus mediaan $= \frac{4^{\text{e}} \text{ cijfer} + 5^{\text{e}} \text{ cijfer}}{2} = \frac{6,6 + 7,0}{2} = 6,8$.
 c Nils zou bijvoorbeeld kunnen proberen om van het cijfer NE een 7,1 te maken.
 Dan zou gelden mediaan $= \frac{4^{\text{e}} \text{ cijfer} + 5^{\text{e}} \text{ cijfer}}{2} = \frac{7,0 + 7,1}{2} = 7,05$.
 Hij zou ook van SK, BIO of AK een 7,1 kunnen proberen te maken.

- 30 a Mediaan, vanwege de uitschieter.
 b Modus, kwalitatieve gegevens.
 c Gemiddelde of mediaan.
 d Modus of mediaan, vanwege de uitschieter (een meetfout).

Bladzijde 75

- 31 $7 \cdot 5,5 - 7 \cdot 5,06 = 3,08$, dus Daan moet 3,1 hoger halen.
 $8 \cdot 5,5 - 8 \cdot 5,08 = 3,36$, dus Tim moet 3,4 hoger halen.
 Dus Tim moet zich het meeste verbeteren.

- 32** Als p het percentage meisjes is, dan is $100 - p$ het percentage jongens.

Er geldt $(100 - p) \cdot 3 = 100 \cdot 1,2$

$$300 - 3p = 120$$

$$-3p = -180$$

$$p = 60.$$

Dus 60% van de klas is meisje.

Alternatieve uitwerking

$\frac{1,2}{3} \times 100\% = 40\%$, dus de jongens bepalen voor 40% het cijfer.

Als 40% jongen is, dan is 60% meisje.

- 33**
- a Ratio, want er is een natuurlijk nulpunt.
 - b 25% van 60 is 15.
Bij 14, 12 en 10 nachten horen de frequenties 2, 4 en 9.
 $2 + 4 + 9 = 15$, dus de 25% langst verblijvende gasten bleven 10, 12 of 14 nachten.
 - c De middelste 50%, dus de laagste en de hoogste 25% niet.
Bij de laagste 25% horen 15 gasten.
Bij 1, 2, 3 en 4 nachten horen de frequenties 3, 1, 6 en 5.
 $3 + 1 + 6 + 5 = 15$, dus de 25% kortst blijvende gasten bleven 1, 2, 3 of 4 nachten.
De gasten die de middelste 50% vormen bleven 5, 6 of 7 nachten.

Bladzijde 76

- 34**
- a De spreidingsbreedte is $3216 - 1648 = 1568$ euro.
De interkwartielafstand is $2304 - 1952 = 352$ euro.
 - b 50% van de werknemers, dus $0,50 \cdot 250 = 125$ werknemers.
 - c Die hebben een salaris van tussen 2304 en 3216 euro per maand.

Bladzijde 77

- 35**
- a Onjuist, het zijn er evenveel, beide 25% van de leerlingen.
 - b Onjuist, 25% van de leerlingen is lichter dan 46,9 kg. Dat zijn $0,25 \cdot 2185 \approx 546$ leerlingen. Dat zijn er minder dan 550.
 - c Onjuist, zowel binnen als buiten de box zit 50% van de leerlingen.
 - d Juist. Het gemiddelde van het kleinste en grootste gewicht is $\frac{34,7 + 92,8}{2} = 63,75$ kg.
Dat valt in de categorie met de zwaarste 25% leerlingen. Dus meer dan driekwart van de leerlingen is lichter.
- 36**
- a 50% van de bedrijven, dus $0,50 \cdot 18\,056 = 9028$ bedrijven.
 - b Het gemiddelde is hoger dan de mediaan omdat de boxplot rechts van de mediaan veel verder doorloopt dan links van de mediaan.
 - c Er zijn evenveel bedrijven met meer dan 117 koeien als bedrijven met minder dan 55 koeien.
 - d In ieder kwartiel zitten $0,25 \cdot 18\,056 = 4514$ bedrijven.
De bedrijven met minder dan 117 koeien hebben dus samen maximaal $4514 \cdot 55 + 4514 \cdot 83 + 4514 \cdot 117 = 1\,115\,070$ koeien.
De bedrijven met meer dan 117 koeien moeten dan per bedrijf gemiddeld meer dan $\frac{1\,115\,070}{4514} = 255$ koeien hebben om meer koeien te hebben. Dat is gezien de boxplot zeker mogelijk.
De bewerking kan dus kloppen.
 - e 100 is precies het midden van de rechterhelft van de box. Naar schatting zal dus 62,5% van de bedrijven minder dan 117 koeien hebben.
Dat zijn $0,625 \cdot 18\,056 = 11\,285$ bedrijven.

Bladzijde 78

- 37** In klas 4VC liggen veel cijfers dicht bij het gemiddelde, dus hier is de spreiding het kleinste. Bij 4VA is de spreiding groter en bij 4VB het grootst, omdat veel cijfers vrij ver van het gemiddelde liggen.

Bladzijde 79

- 38 a** 1 Het gemiddelde is $\frac{0+2+3+4+6}{5} = 3$.
 2 Waarneming – gemiddelde geeft $-3, -1, 0, 1, 3$.
 3 Kwadrateren en optellen geeft $(-3)^2 + (-1)^2 + 0^2 + 1^2 + 3^2 = 20$.
 4 Delen door aantal waarnemingsgetallen geeft $\frac{20}{5} = 4$.
 5 De wortel nemen geeft $\sqrt{4} = 2$. Dus $\sigma = 2$.
- b** Alle getallen zijn 10 hoger, dus de afwijkingen van het gemiddelde zijn hetzelfde als bij A. De standaardafwijking is dus gelijk.
- c** Alle waarnemingsgetallen zijn hetzelfde, dus $\sigma = 0$.
- 39 a** Omdat zelfs het laagste en het hoogste cijfer niet zover afwijken.
- b** De standaardafwijking is kleiner geworden, want een van de waarnemingen is dichterbij het gemiddelde gekomen.
- 40 a** Bij groep B, want daar is de spreiding van de lengtes het grootst.
- b** Figuur b, want daar is de spreiding, dus de standaardafwijking, kleiner dan bij figuur a.

2.4 Klassenindeling en polygonen**Bladzijde 81**

- 41** Het geeft geen beter overzicht, want elk waarnemingsgetal heeft frequentie 1.

Bladzijde 82

- 42 a** Discrete variabele, omdat prijzen in supermarkten op centen worden afgerond.
- b** In de klasse $(200, 300]$.
- c** De totale frequentie is $26 + 17 + 21 + 8 + 13 + 2 = 87$.
 gemiddelde = $\frac{26 \cdot 50 + 17 \cdot 150 + 21 \cdot 250 + 8 \cdot 350 + 13 \cdot 450 + 2 \cdot 550}{87} \approx 217$ cent.
- d** De mediaan is het 44^e waarnemingsgetal.
 $26 + 17 = 43$, dus de mediaan ligt in de klasse $(200, 300]$.
- e** De modale klasse is de klasse $(0, 100]$, want die heeft de grootste frequentie.

43 a

oppervlakte in m ²	frequentie
[50, 150)	8
[150, 250)	5
[250, 350)	3
[350, 450)	2
[450, 550)	2

- b** De modale klasse is de klasse $[50, 150)$, want die heeft de grootste frequentie.
- c** Het aantal werknemers is $6 + 8 + \dots + 19 = 266$.

$$\text{gemiddelde} = \frac{8 \cdot 100 + 5 \cdot 200 + 3 \cdot 300 + 2 \cdot 400 + 2 \cdot 500}{266} \approx 16,9 \text{ m}^2$$

De berekening met de klassenmiddens is een benadering, want hierbij ga je ervan uit dat de gegevens gelijkmatig verdeeld zijn over de klasse en dat is niet zo. Hierdoor ontstaat het verschil in uitkomsten.

Bladzijde 83

- 44** a In de klasse $[35, 45)$.
 b De totale frequentie is $18 + 3 + 40 + 9 + 2 = 72$.
 Het kleinst mogelijke gemiddelde krijg je door te rekenen met de laagst mogelijke leeftijd per klasse, dat geeft

$$\text{gemiddelde} = \frac{18 \cdot 25 + 3 \cdot 35 + 40 \cdot 45 + 9 \cdot 55 + 2 \cdot 65}{72} \approx 41,4 \text{ jaar.}$$

 Het kan dus niet kloppen.
 c De mediaan is het gemiddelde van het 36^{e} en het 37^{e} getal.
 $18 + 3 = 21$ en $18 + 3 + 40 = 61$, dus de mediaan ligt in de klasse $[45, 55)$.
 De mediaan kan dus 47 jaar zijn.
 d In de klasse $[45, 55)$ vallen 10 leeftijden. Doordat deze klasse frequentie 40 heeft komt in elk geval één leeftijd uit deze klasse minstens 4 keer voor.
 De leeftijd 36 jaar valt in de klasse $[35, 45)$ met frequentie 3 en komt dus hoogstens 3 keer voor.
- 45** a Discreet, want een leerling kan alleen een geheel aantal broers of zussen hebben.
 Het meetniveau is ratio, want er is een natuurlijk nulpunt.
 b Er zijn 6 leerlingen met 2 broers of zussen.
 c $5 + 12 + 6 + 2 + 2 + 1 = 28$ leerlingen.

Bladzijde 84

- 46** a De variabele gewicht ei in gram. Deze variabele is kwantitatief.
 b De modale klasse is $[66, 69)$, want die heeft de grootste frequentie.
 c Het kleinst mogelijke totale gewicht krijg je door te rekenen met het laagste gewicht per klasse, dat geeft totale gewicht $= 8 \cdot 60 + 11 \cdot 63 + 13 \cdot 66 + 12 \cdot 69 + 6 \cdot 72 = 3291$ gram.
 Het hoogst mogelijke totale gewicht krijg je door te rekenen met het grootste gewicht per klasse, dat geeft totale gewicht $= 8 \cdot 63 + 11 \cdot 66 + 13 \cdot 69 + 12 \cdot 72 + 6 \cdot 75 = 3441$ gram.
 Dus Matthijs heeft inderdaad gelijk.
 d Bij het kleinst mogelijke totale gewicht is het gemiddelde gewicht $\frac{3291}{50} = 65,82$ gram.
 Dus het is mogelijk.
- 47** a De variabele klasse is kwalitatief. Je kunt bijvoorbeeld niet een gemiddelde klasse bepalen.
 b Nee, dat kan niet. Omdat de klassen S en XL aan één kant onbegrensd zijn, kan er sprake zijn van grote uitschieters en dus is een goede schatting niet mogelijk.
 c Een histogram gebruik je bij een kwantitatieve variabele, maar klasse is kwalitatief.
- 48** a De variabele is gemiddeld bonbedrag in euro's. Deze variabele is continu, want alle tussenliggende bedragen zijn mogelijk.
 b 10 euro, want daar begint de eerste klasse.
 c $6 + 7 + 4 + 3 + 5 + 3 = 28$ winkels.
 d De mediaan is het gemiddelde van het 14^{e} en 15^{e} waarnemingsgetal.
 $6 + 7 = 13$, dus de mediaan ligt in de klasse $[30, 40)$.
 e De relatieve frequentie is $\frac{5}{28} \times 100\% = 17,9\%$.

Bladzijde 85

- 49** a Ja, want het betreft kwantitatieve gegevens.
 b $\text{gemiddelde} = 0,10 \cdot 1 + 0,12 \cdot 2 + 0,14 \cdot 3 + 0,20 \cdot 4 + 0,16 \cdot 5 + 0,14 \cdot 6 + 0,08 \cdot 7 = 3,76$
 c Op 33 dagen meer dan 5 leerlingen en dat is op $14\% + 8\% = 22\%$ van de dagen.
 Dus in totaal duurde het onderzoek $\frac{33}{0,22} = 150$ dagen.
 d Gedurende 150 dagen kwamen er gemiddeld 3,76 leerlingen te laat, dus zij heeft $150 \cdot 3,76 = 564$ keer genoteerd dat een leerling te laat kwam.

- 50**
- a** $0,3\% + 0,7\% + 2,4\% + 5,9\% + 12,2\% = 21,5\%$ van de sporters is kleiner dan Diego.
 - b** $3,0\% + 1,1\% + 0,4\% + 0,2\% = 4,7\%$ van de sporters is langer dan 2 meter, dus het klopt.
 - c** $0,3\% + 0,7\% + 2,4\% + 5,9\% + 12,2\% + 16,3\% + 20,4\% = 58,2\%$.
Dus zijn lengte is minstens 185 cm en minder dan 190 cm.
 - d** Een percentielscore van 50.

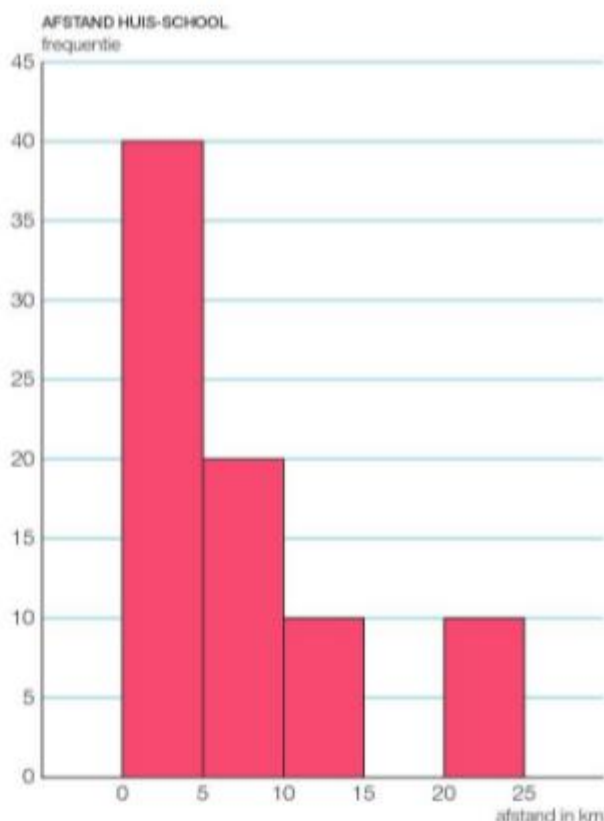
Bladzijde 86

- 51** De frequentie die je bij de rechtergrens afleest heeft dan de betekenis 'aantal met deze waarde of lager'. Zie bijvoorbeeld de polygoon in theorie C. In de klasse $[10, 20)$ zitten 6 bonbedragen. De waarnemingen in deze klasse liggen verspreid tussen 10 en 20 euro. Als het punt bij het klassenmidden 15 uitgezet zou zijn, dan zou het net lijken of alle 6 bonbedragen minder dan 15 euro zouden zijn.

- 52**
- a** Aflezen geeft 80 studenten.
 - b** Aflezen geeft 60 studenten.
 - c** $80 - 40 = 40$ studenten.
 - d** $70 - 60 = 10$ studenten.

e

afstand in km	frequentie
$[0, 5)$	40
$[5, 10)$	20
$[10, 15)$	10
$[15, 20)$	0
$[20, 25)$	10



Bladzijde 87

- 53** a Aflezen geeft 60%.
Dat zijn $0,60 \cdot 300 = 180$ klanten.
- b Aflezen geeft 85%.
Dat zijn $0,85 \cdot 300 = 255$ klanten.
- c 2,5 minuten = 150 seconden
Aflezen geeft $100\% - 70\% = 30\%$.
Dat zijn $0,30 \cdot 300 = 90$ klanten.
- d Aftrekken van de antwoorden van a en b geeft $255 - 180 = 75$ klanten.
- 54** a Van 10:00 tot 20:00 is $10 \cdot 4 = 40$ kwartieren in totaal.
Aflezen bij 20 klanten geeft 80%.
Dat zijn $0,80 \cdot 40 = 32$ kwartieren.
- b Aflezen bij 15 klanten geeft 40%.
Dus $100\% - 40\% = 60\%$ van de kwartieren minstens 15 klanten.
Dat zijn $0,60 \cdot 40 = 24$ kwartieren.
- c Aflezen bij 10 klanten geeft 20%.
Aflezen bij 25 klanten geeft 90%.
Dus $90\% - 20\% = 70\%$ van de kwartieren minstens 10, maar minder dan 25 klanten.
Dat zijn $0,70 \cdot 40 = 28$ kwartieren.

d	aantal per kwartier	relatieve frequentie	frequentie
	[0, 5)	10%	4
	[5, 10)	10%	4
	[10, 15)	20%	8
	[15, 20)	40%	16
	[20, 25)	10%	4
	[25, 30)	10%	4

2.5 Soorten verdelingen**Bladzijde 89**

- 55** a Er waren 7 wedstrijden waarin 2 keer werd gescoord.
- b $8 + 4 + 7 + 2 + 3 + 6 + 2 + 1 + 1 = 34$ wedstrijden gespeeld.
- c $8 \cdot 0 + 4 \cdot 1 + 7 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + 6 \cdot 5 + 2 \cdot 6 + 1 \cdot 7 + 1 \cdot 8 = 93$ doelpunten gescoord.
- d Het kunnen alleen wedstrijden met 4 doelpunten zijn. Dus maximaal 3 wedstrijden zijn in 2-2 geëindigd.
- e De mediaan is het gemiddelde van het 17^e en 18^e waarnemingsgetal.
 $8 + 4 = 12$ en $8 + 4 + 7 = 19$, dus de mediaan is 2.
Het gemiddeld aantal doelpunten per wedstrijd is $\frac{93}{34} \approx 2,7$.
Dus het gemiddeld aantal doelpunten is groter dan de mediaan.

Bladzijde 90

- 56** a Het gemiddelde is lager bij figuur 2.30 dan bij figuur 2.28, door de uitschieter.
- b De mediaan is hetzelfde, want daar heeft de uitschieter geen invloed op.
- c De uitschieter zorgt ervoor dat de standaardafwijking groter is bij figuur 2.30 dan bij figuur 2.28.

Bladzijde 91

- 57** **a** Mogelijk betrof het een eerste en een zesde klas, zodat de linkertop bij de eerste klas en de rechtershoek bij de zesde klas hoorde.
- b** De modale klasse is de klasse $[182, 184)$, want die heeft de grootste frequentie.
- c** Het gemiddelde is ongeveer 176 cm. De lengten wijken hier gemiddeld ongeveer 8 cm van af. Een schatting van de standaardafwijking is dus 8 cm.
- d** Als boxplot I bij het histogram zou passen dan zou 25% van de waarnemingen tussen 160 en 164 liggen, dat is in figuur 2.31 duidelijk niet het geval. Er zitten rechts van 176 cm een stuk meer waarnemingsgetallen dan links van 176 cm. De mediaan zal dus groter zijn dan 176 cm. Boxplot III past dus niet bij het histogram. Boxplot II past het beste bij figuur 2.31, want ongeveer een kwart van de waarnemingen zit tussen 179 cm en 183 cm.

Bladzijde 92

- 58** **a** De modale klasse is $[12\,000, 18\,000)$, de mediaan ligt rechts hiervan, want de cumulatieve frequentie bij deze klasse is ongeveer $10,2\% + 13,6\% + 15,8\% = 39,6\%$, en dat is minder dan 50%.
- b** Nee, als iedereen een gelijk inkomen heeft, dan is dat het enige inkomen dat voorkomt en bestaat de verdeling dus uit één staaf.
- c** Boxplot I past het beste, want 50% van de inkomens zit ongeveer tussen 12 000 euro en 40 000 euro.

Bladzijde 93

- 59** **a** symmetrisch
b rechts-scheef
c rechts-scheef
d symmetrisch
e rechts-scheef
- 60** Bij I hoort D.
Bij II hoort A.
Bij III hoort B.
Bij IV hoort C.

Bladzijde 94

- 61** **a** Veel geboorten vinden plaats in het ziekenhuis, dus nuljarigen maken veel kosten.
- b** De staaf bij de klasse 55-59 is de eerste die boven 1593 euro uitkomt. Dus vanaf de leeftijdsgroep van 55-59 jaar zijn de kosten meer dan gemiddeld.
- c** De kosten zijn per persoon en niet elke leeftijdsgroep bevat evenveel personen.
- d** De klassenbreedtes zijn niet gelijk. De klasse van de nuljarigen is 1 jaar breed. De meeste andere klassen zijn 5 jaar breed.
- e** $45-49 \quad 1200 \cdot 1\,300\,000 \approx 1,6$ miljard euro
 $75-79 \quad 4100 \cdot 550\,000 \approx 2,3$ miljard euro
De leeftijdsgroep van 75-79 jaar heeft dus de hoogste totale kosten.

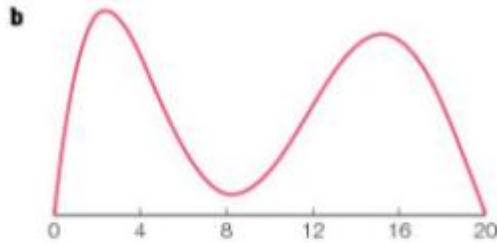
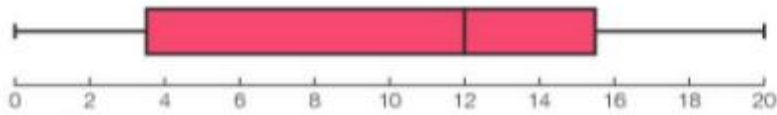
Bladzijde 95

- 62** De modus zal kleiner zijn dan de mediaan, want de modus zit vrijwel helemaal links en dan heb je nog niet de helft van de waarnemingsgetallen.

Bladzijde 96

- 63** **a** Bij de top van de verdelingskromme hoort de grootste frequentie, dus de cumulatieve verdelingskromme neemt daar het meest toe, oftewel daar is het steilste stuk van de verdelingskromme.
- b** De cumulatieve verdelingskromme is overal even steil, de verdelingskromme is dus overal even hoog, dus er is sprake van een uniforme verdeling.

64 a Aflezen geeft $\min = 0$, $Q_1 = 3,5$, mediaan = 12, $Q_3 = 15,4$ en $\max = 20$.



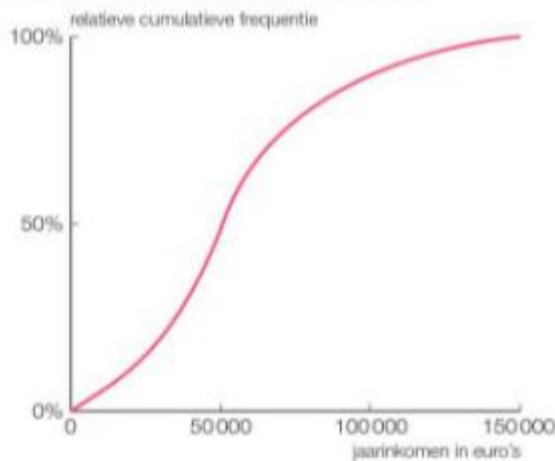
c Een tweetoppige verdeling.

65 a Weinig waarnemingsgetallen in het begin en veel op het eind.



- b** Het is een links-scheve verdeling.
c Het gemiddelde ligt links van de mediaan.

- 66** a De mediaan ligt rechts van de modus.
b Het steilste stuk zit bij ongeveer 25 000 euro. De mediaan is ongeveer 50 000 euro.

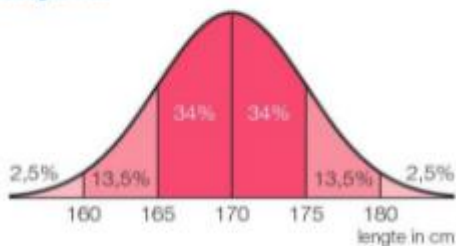


Bladzijde 97

- 67** $100\% - 68\% = 32\%$ van de waarnemingsgetallen is kleiner dan 165 of groter dan 185. De helft daarvan, dus 16% is groter dan 185.

Bladzijde 98

68 a



- b** $34\% + 34\% + 13,5\% = 81,5\%$

c 2,5%

$13,5\% + 2,5\% = 16\%$

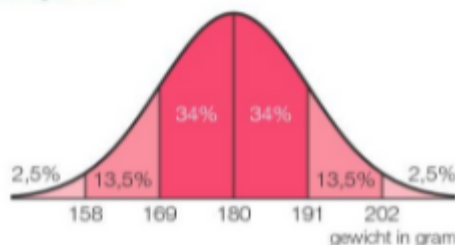
$13,5\% + 34\% = 47,5\%$



- 69 a nee
b ja
c nee
d ja
e nee
f ja

Bladzijde 99

70 a



$$13,5\% + 34\% + 34\% = 81,5\%$$

Dus $0,815 \cdot 5000 = 4075$ appels.

b 84%, dus $0,84 \cdot 5000 = 4200$ appels.

c 125 is $\frac{125}{5000} \times 100\% = 2,5\%$.

Dus zwaarder dan 202 gram.

- 71 a Met stijgende leeftijd neemt iemands reactietijd toe. Bij de 18-jarigen hoort kromme A (kleinste gemiddelde), bij de 60-jarigen hoort kromme C (grootste gemiddelde).
b Bij kromme C hoort de grootste standaardafwijking, dus bij 60-jarigen is de kans het grootst.

- 72 Kromme A: $\mu = 65$ en van toenemend stijgend naar afnemend stijgend bij 63,5 geeft $\sigma = 65 - 63,5 = 1,5$.
Kromme B: $\mu = 66,5$ en van toenemend stijgend naar afnemend stijgend bij 65,3 geeft $\sigma = 66,5 - 65,3 = 1,2$.
Kromme C: $\mu = 67,5$ en van toenemend stijgend naar afnemend stijgend bij 65,5 geeft $\sigma = 67,5 - 65,5 = 2,0$.
Kromme D: $\mu = 70$ en van toenemend stijgend naar afnemend stijgend bij 69,5 geeft $\sigma = 70 - 69,5 = 0,5$.

2.6 Betrouwbaarheidsintervallen

Bladzijde 101

- 73 a *
Vind je bijvoorbeeld 14, dan is de steekproefproportie 0,14.
b *
c *
d Bij een grotere steekproeflengte liggen de grenzen waartussen de steekproefproportie ligt dichterbij elkaar.
e Tussen 0,09 en 0,15.
f Dat zijn de grenzen 0,10 en 0,14.

Bladzijde 105

- 74 a $\hat{p} = \frac{378}{600} = 0,63$ en $\sigma = \sqrt{\frac{0,63 \cdot 0,37}{600}} = 0,01971... \approx 0,0197$
b $\hat{p} - 2\sigma = 0,63 - 2 \cdot 0,0197... \approx 0,591$
 $\hat{p} + 2\sigma = 0,63 + 2 \cdot 0,0197... \approx 0,669$
Het 95%-betrouwbaarheidsinterval is $[0,591; 0,669]$.
c Bij de steekproeflengte 1200 hoort een kleiner 95%-betrouwbaarheidsinterval dan bij steekproeflengte 600.
Dus bij steekproeflengte 1200 hoort een grotere nauwkeurigheid dan bij steekproeflengte 600.

75 68% van de waarnemingen ligt tussen $\hat{p} - \sigma$ en $\hat{p} + \sigma$.

$$\hat{p} - \sigma = 0,63 - 0,014 = 0,616$$

$$\hat{p} + \sigma = 0,63 + 0,014 = 0,644$$

Het 68%-betrouwbaarheidsinterval is $[0,616; 0,644]$.

76 a $\hat{p} = \frac{117}{150} = 0,78$

b $\hat{p} = 0,78$ en $\sigma = \sqrt{\frac{0,78 \cdot 0,22}{150}} = 0,0338\dots$

$$\hat{p} - 2\sigma = 0,78 - 2 \cdot 0,0338\dots \approx 0,712$$

$$\hat{p} + 2\sigma = 0,78 + 2 \cdot 0,0338\dots \approx 0,848$$

Het 95%-betrouwbaarheidsinterval is $[0,712; 0,848]$.

77 a $\hat{p} = \frac{47}{575} = 0,0817\dots$ en $\sigma = \sqrt{\frac{0,0817\dots \cdot 0,9182\dots}{575}} = 0,0114\dots$

$$\hat{p} - 2\sigma = 0,0817\dots - 2 \cdot 0,0114\dots \approx 0,059$$

$$\hat{p} + 2\sigma = 0,0817\dots + 2 \cdot 0,0114\dots \approx 0,105$$

Het 95%-betrouwbaarheidsinterval is $[0,059; 0,105]$.

b $4\sigma = 0,296 - 0,224 = 0,072$

Dus de standaardafwijking is $\sigma = \frac{0,072}{4} = 0,018$.

78 De steekproefproportie van de mensen die zeggen op A te gaan stemmen is $\hat{p} = \frac{487}{935} = 0,5208\dots$

De bijbehorende standaardafwijking is $\sigma = \sqrt{\frac{0,5208\dots \cdot 0,4791\dots}{935}} = 0,0163\dots$

$$\hat{p} - 2\sigma = 0,5208\dots - 2 \cdot 0,0163\dots \approx 0,488$$

Omdat $0,488 < 0,5$ kun je niet met een betrouwbaarheid van 95% zeggen dat kandidaat A de verkiezingen zal gaan winnen.

Bladzijde 106

79 a $\hat{p} = \frac{225}{1250} = 0,18$ en $\sigma = \sqrt{\frac{0,18 \cdot 0,82}{1250}} = 0,0108\dots$

$$\hat{p} - 2\sigma = 0,18 - 2 \cdot 0,0108\dots \approx 0,158$$

$$\hat{p} + 2\sigma = 0,18 + 2 \cdot 0,0108\dots \approx 0,202$$

Het 95%-betrouwbaarheidsinterval is $[0,158; 0,202]$.

b $\hat{p} = \frac{143}{715} = 0,2$ en $\sigma = \sqrt{\frac{0,2 \cdot 0,8}{715}} = 0,0149\dots$

$$\hat{p} - 2\sigma = 0,2 - 2 \cdot 0,0149\dots \approx 0,170$$

$$\hat{p} + 2\sigma = 0,2 + 2 \cdot 0,0149\dots \approx 0,230$$

Het 95%-betrouwbaarheidsinterval is $[0,170; 0,230]$.

c Van de 225 personen uit Zuid-Holland waren er 143 man en dus $225 - 143 = 82$ vrouw.

$$\hat{p} = \frac{82}{535} = 0,153\dots$$
 en $\sigma = \sqrt{\frac{0,153\dots \cdot 0,846\dots}{535}} = 0,0155\dots$

$$\hat{p} - 2\sigma = 0,153\dots - 2 \cdot 0,0155\dots \approx 0,122$$

$$\hat{p} + 2\sigma = 0,153\dots + 2 \cdot 0,0155\dots \approx 0,184$$

Het 95%-betrouwbaarheidsinterval is $[0,122; 0,184]$.

d De breedte van $[0,158; 0,202]$ is $0,202 - 0,158 = 0,044$.

De breedte van $[0,170; 0,230]$ is $0,230 - 0,170 = 0,060$.

De breedte van $[0,122; 0,184]$ is $0,184 - 0,122 = 0,062$.

Van smal naar breed: $[0,158; 0,202]$, $[0,170; 0,230]$, $[0,122; 0,184]$.

Dat wordt veroorzaakt door het verschil in grootte van de steekproef.

e $4\sigma = 0,264 - 0,216 = 0,048$, dus $\sigma = \frac{0,048}{4} = 0,012$.

f $\hat{p} = \frac{0,216 + 0,264}{2} = 0,24$

Dus er waren $0,24 \cdot 1250 = 300$ bezoekers in de leeftijd 18-24 jaar.

Het 95%-betrouwbaarheidsinterval is $[0,216 \cdot 1250; 0,264 \cdot 1250] = [270; 330]$.

B0 a $4\sigma = 0,194 - 0,146 = 0,048$, dus $\sigma = \frac{0,048}{4} = 0,012$

$$\hat{p} = \frac{0,146 + 0,194}{2} = 0,17$$

$$\sigma = \sqrt{\frac{0,17 \cdot 0,83}{n}} = \sqrt{\frac{0,1411}{n}}$$

Je krijgt de vergelijking $\sqrt{\frac{0,1411}{n}} = 0,012$.

Voer in $y_1 = \sqrt{\frac{0,1411}{x}}$ en $y_2 = 0,012$.

De optie snijpunt geeft $x \approx 980$.

Het aantal stemgerechtigden in de steekproef was dus 980.

Alternatieve uitwerking

$$\sqrt{\frac{0,1411}{n}} = 0,012$$

$$\frac{0,1411}{n} = 0,012^2$$

$$n = \frac{0,1411}{0,012^2}$$

$$n \approx 980$$

Het aantal stemgerechtigden in de steekproef was dus 980.

- b** Mogelijk zijn mensen nog van mening veranderd of was de steekproef niet voldoende representatief. Ook kan het zijn dat er mensen niet zijn gaan stemmen.

Bladzijde 107

- B1** a $3750 + 2 \cdot 480 = 4710$. Dus 2,5% van de jongens was naar verwachting zwaarder dan 4710 gram. Dat zijn $0,025 \cdot 87\,159 \approx 2180$ jongens.

b $\frac{12\,332}{18\,138} \times 100\% \approx 68\%$

De gewichten wijken dus hoogstens σ af van het gemiddelde.

Die gewichten liggen dus tussen de $3750 - 480 = 3270$ gram en $3750 + 480 = 4230$ gram.

Bladzijde 108

- B2** 6 jaar en 9 maanden is $6 \cdot 12 + 9 = 81$ maanden.

Het 95%-betrouwbaarheidsinterval is $\left[81 - 2 \cdot \frac{10}{\sqrt{65}}, 81 + 2 \cdot \frac{10}{\sqrt{65}} \right] \approx [78,5; 83,5]$.

78,5 maanden is 6 jaar en 6,5 maanden

83,5 maanden is 6 jaar en 11,5 maanden

Met 95% zekerheid ligt de gemiddelde leeftijd waarop basisschoolleerlingen het zwemdiploma A halen tussen 6 jaar en 6,5 maanden en 6 jaar en 11,5 maanden.

Bladzijde 109

B3 $\bar{X} = \frac{170,2 + 179,8}{2} = 175$

$$4 \cdot \frac{S}{\sqrt{35}} = 179,8 - 170,2 \text{ oftewel } 4 \cdot \frac{S}{\sqrt{35}} = 9,6$$

Voer in $y_1 = 4 \cdot \frac{x}{\sqrt{35}}$ en $y_2 = 9,6$.

De optie snijpunt geeft $x \approx 14,2$.

Dus $S \approx 14,2$.

Alternatieve uitwerking

$$4 \cdot \frac{S}{\sqrt{35}} = 9,6$$

$$4S = 9,6 \cdot \sqrt{35}$$

$$S = 2,4 \cdot \sqrt{35} \approx 14,2$$

84 De breedte van een 95%-betrouwbaarheidsinterval voor een gemiddelde is $4 \cdot \frac{S}{\sqrt{n}}$.

De breedte van het 95%-betrouwbaarheidsinterval voor

meisjes van 6 maanden oud is $4 \cdot \frac{2,8}{\sqrt{38}} \approx 1,817$.

De breedte van het 95%-betrouwbaarheidsinterval voor

meisjes van 12 maanden oud is $4 \cdot \frac{3,2}{\sqrt{38}} \approx 2,076$.

De breedte van het 95%-betrouwbaarheidsinterval voor

jongens van 6 maanden oud is $4 \cdot \frac{3,1}{\sqrt{43}} \approx 1,891$.

De breedte van het 95%-betrouwbaarheidsinterval voor

jongens van 12 maanden oud is $4 \cdot \frac{3,4}{\sqrt{43}} \approx 2,074$.

Het interval voor meisjes van 6 maanden oud is het smalst.

85 $4 \cdot \frac{10}{\sqrt{n}} = 182 - 178 = 4$

Voer in $y_1 = 4 \cdot \frac{10}{\sqrt{x}}$ en $y_2 = 4$.

De optie snijpunt geeft $x = 100$.

Dus de steekproefomvang was 100.

Alternatieve uitwerking

$$4 \cdot \frac{10}{\sqrt{n}} = 4$$

$$\frac{10}{\sqrt{n}} = 1$$

$$\sqrt{n} = 10$$

$$n = 100$$

De steekproefomvang was 100.

Diagnostische toets

Bladzijde 112

- 1**
 - a** Fase 1: De onderzoeksvraag.
Fase 2: Data verzamelen.
Fase 3: Data analyseren.
Fase 4: Conclusies trekken.
 - b** Een verband waarbij het één direct gevolg is van het ander.
Bijvoorbeeld je verslapen en te laat komen.
 - c** Een representatieve steekproef moet voldoende groot en aselekt zijn.
- 2**
 - a** Een kwantitatieve variabele hoort bij meetbare gegevens en wordt uitgedrukt in een getal, bijvoorbeeld snelheid.
Bij een kwalitatieve variabele gaat het om een eigenschap of kwaliteit, bijvoorbeeld beroep.
 - b** Een discrete variabele neemt alleen 'losse' waarden aan, bijvoorbeeld schoenmaat.
Een continue variabele kan alle tussenliggende waarden aannemen, bijvoorbeeld gewicht.
 - c** Nominaal, bijvoorbeeld beroep.
Ordinaal, bijvoorbeeld rugnummer bij voetbal.
Interval, bijvoorbeeld temperatuur in °C.
Ratio, bijvoorbeeld lichaamsgewicht.
 - d** De populatieproportie is het deel van een populatie dat een bepaalde eigenschap heeft.
De steekproefproportie is het deel van een steekproef dat een bepaalde eigenschap heeft.
In een representatieve steekproef zullen deze weinig verschillen.

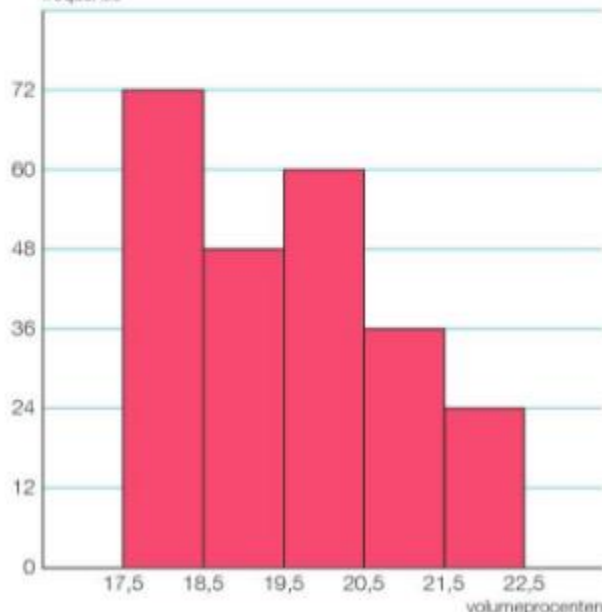
- 3 a** 25%, dus $0,25 \cdot 365 \approx 91$ dagen.
- b** De rechtergrenzen van de kwartielen zijn $10,8^\circ\text{C}$; $15,4^\circ\text{C}$; $21,8^\circ\text{C}$ en $31,9^\circ\text{C}$.
Het gemiddelde daarvan is $\frac{10,8 + 15,4 + 21,8 + 31,9}{4} = 19,975^\circ\text{C}$.
Dus de gemiddelde temperatuur kan onmogelijk meer dan 20°C zijn geweest.
- c** Het tweede kwartiel loopt van $10,8^\circ\text{C}$ tot $15,4^\circ\text{C}$.
In dit kwartiel zitten 47 mogelijke temperaturen.
Er liggen 91 dagen in dit kwartiel. Er zijn dus meerdere temperaturen die twee keer voorkomen, maar het is niet noodzakelijk dat er temperaturen drie keer voorkomen.
Als $15,2^\circ\text{C}$ de modus is moet deze temperatuur dus minstens drie keer zijn voorgekomen.
- d** De spreidingsbreedte is $31,9 - -1,9 = 33,8^\circ\text{C}$.
De interkwartielafstand is $21,8 - 10,8 = 11^\circ\text{C}$.
 $\frac{33,8}{11} > 3$. Lars heeft dus gelijk.
- e** De beste schatting voor de standaardafwijking is 9°C .

Bladzijde 113

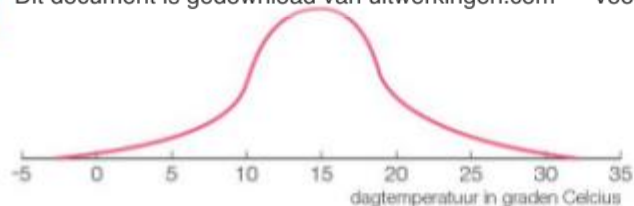
- 4 a** Het kleinst mogelijke totaal aantal klachten krijg je door te rekenen met het kleinst mogelijke aantal klachten per klasse. Dat geeft
 $2 \cdot 0 + 6 \cdot 5 + 55 \cdot 10 + 126 \cdot 15 + 135 \cdot 20 + 35 \cdot 25 + 6 \cdot 30 = 6225$ klachten.
 $6225 > 6215$, dus je kunt concluderen dat er dit jaar meer klachten waren.
- b** De mediaan is het 183^e waarnemingsgetal;
 $2 + 6 + 55 = 63$ en $2 + 6 + 55 + 126 = 189$, dus de mediaan ligt in de klasse $[15, 20)$.
- c** De modale klasse is de klasse $[20, 25)$, want die heeft de grootste frequentie.
- 5 a** 30%, dus $0,3 \cdot 240 = 72$ mannen met minder dan 18,5%.
25%, dus $0,25 \cdot 240 = 60$ mannen met minstens 20,5%.

zuurstofgehalte	relatieve frequentie	absolute frequentie
$[17,5; 18,5)$	30%	72
$[18,5; 19,5)$	20%	48
$[19,5; 20,5)$	25%	60
$[20,5; 21,5)$	15%	36
$[21,5; 22,5)$	10%	24

c ZUURSTOFGEHALTE
frequentie



6



7

a 97,5%, dus $0,975 \cdot 750 \approx 731$ potten.

b 13,5%, dus $0,135 \cdot 750 \approx 101$ potten.

8

a $\hat{p} = \frac{80}{500} = 0,16$ en $\sigma = \sqrt{\frac{0,16 \cdot 0,84}{500}} = 0,0163\dots$

$$\hat{p} - 2\sigma = 0,16 - 2 \cdot 0,0163\dots \approx 0,127$$

$$\hat{p} + 2\sigma = 0,16 + 2 \cdot 0,0163\dots \approx 0,193$$

Het 95%-betrouwbaarheidsinterval is $[0,127; 0,193]$.

b Het 95%-betrouwbaarheidsinterval is $\left[350 - 2 \cdot \frac{108}{\sqrt{80}}, 350 + 2 \cdot \frac{108}{\sqrt{80}} \right] \approx [326, 374]$.