Wiskunde A

5 HAVO

Katern 2

ConTeXt College

Inhoud

Voorwoord

Het lesmateriaal in dit katern is gebaseerd op het materiaal dat je kunt vinden op de Math4All website `www.math4all.nl`. In de tekst staan dan ook regelmatig verwijzingen naar die website. Waar je precies moet zijn op die website kun je zien in de kopregel van iedere pagina.

Ieder hoofdstuk bestaat uit een aantal paragrafen en wordt steeds afgesloten met een paragraaf *Totaalbeeld* waar de leerstof wordt samengevat en/of herhaald. Iedere paragraaf is ingedeeld in vaste rubrieken die houvast geven bij de bestudering van het lesmateriaal.

- Verkennen
- Uitleg
- Theorie en Voorbeelden
- Verwerken
- Toepassen

Indien er in het lesmateriaal wordt verwezen naar werkbladen dan kun je deze terugvinden op de website en achterin je katern.

1

Allerlei verbanden

1.1 Werken met variabelen

Inleiding

Je hebt al leren werken met formules. Alleen formules van de vorm $y = ...$, waarbij op de stippeltjes een uitdrukking staat die één variabele bevat, kun je in je grafische rekenmachine invoeren om de grafiek te bekijken. Maar soms heb je met meerdere variabelen te maken. En hoe kun je dan toch grafieken maken met je GR?

Je leert in dit onderwerp

- formules herleiden tot de vorm $y = ...$;
- bij formules met meerdere variabelen een grafiekenbundel maken;
- formules met meerdere variabelen gebruiken en daarin eenvoudige uitdrukkingen invullen

Voorkennis

- werken met formules en grafieken, ook met de grafische rekenmachine;
- vergelijkingen en ongelijkheden oplossen.

Verkennen

Opgave V1

Een fabrikant wil een nieuw product op de markt brengen. Hij is de enige aanbieder van dit product en hij weet daarom dat de hoeveelheid (q) die hij dagelijks kan verkopen alleen zal afhangen van de prijs (p in euro) die hij ervoor vraagt. Er geldt $q = 400 - 25p$. Voor de opbrengst R van de verkoop van dit product geldt $R = p \cdot q$.

a Met hoeveel variabelen heb je hier te maken?

b Maak een formule voor de opbrengst R uitgedrukt in q.

c Bereken bij welke hoeveelheid q hij een zo groot mogelijke opbrengst R heeft.

Uitleg 1

Een fabrikant brengt een nieuw product op de markt. De hoeveelheid q die hij verkoopt, hangt af van de prijs p (euro). Er geldt: $q = 400 - 25p$.
Voor de opbrengst R (euro) van de verkoop van dit product geldt: $R = p \cdot q$.

Dit zijn twee formules waarin drie verschillende variabelen voorkomen.

De fabrikant wil weten bij welke hoeveelheid q hij een zo groot mogelijke opbrengst R heeft. Om dit uit te zoeken moet er een formule gemaakt worden waarbij de opbrengst R wordt uitgedrukt in q.

Druk eerst p uit in q:

$$\begin{aligned} q &= 400 - 25p \\ q + 25p &= 400 \\ 25p &= 400 - q \\ p &= 16 - 0{,}04q \end{aligned}$$

beide zijden $+25p$

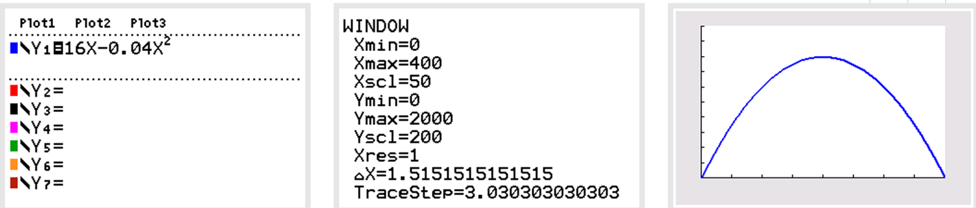
beide zijden $-q$

beide zijden $:\ 25$

Vervang in de opbrengstformule de p door de uitdrukking $16 - 0{,}04q$ en herleid:
$R = (16 - 0{,}04q) \cdot q = 16q - 0{,}04q^2$

Dit is een formule met twee variabelen van de vorm $R = ...$
Je kunt hierbij een grafiek plotten met de grafische rekenmachine.
Voer in: $y_1 = 16x - 0{,}04x^2$.

Zorg ervoor dat de grafiek goed in beeld is. Bedenk dat de hoeveelheid verkochte producten een positief getal moet zijn. Bekijk met de tabel op de grafische rekenmachine wat er met q gebeurt als p oploopt.



Figuur 1.1

Als de grafiek goed in beeld is, kan het maximum bepaald worden. Bekijk indien nodig in het **Practicum** hoe dat moet. Het maximum is $(200{,}1600)$. Bij een verkoop van 200 stuks per dag is de opbrengst van de fabrikant maximaal.

Opgave 1

Gebruik de gegevens uit **Uitleg 1**.

a Welke formule in de uitleg heeft meer dan twee variabelen?

b Waarom moet er een formule gemaakt worden waarin R wordt uitgedrukt in q?

c Waarom wordt daarna eerst p uitgedrukt in q?

d Hoe ontstaat vervolgens de formule $R = 16q - 0{,}04q^2$?

e Hoe worden de vensterinstellingen van de grafische rekenmachine bepaald voor het plotten van de grafiek bij deze formule?

f Hoeveel bedraagt de maximale opbrengst per dag?

Opgave 2

Er kan ook een formule gemaakt worden waarin R is uitgedrukt in p. De formule $q = 400 - 25p$ hoeft dan niet te worden herleid.

a Welke formule voor R uitgedrukt in p ontstaat dan?

b Bij de formule van a moet een grafiek geplot worden. Beredeneer de juiste vensterinstellingen voor de horizontale as.

c De instellingen van de verticale as kun je met de tabel op je grafische rekenmachine bepalen.
Plot de grafiek bij het verband tussen R en p.

d Voor welke prijzen is de opbrengst per dag meer dan € 1000,00?

Opgave 3

Een fabrikant wil een nieuw product op de markt brengen. Hij is de enige aanbieder van dit product en weet dat de hoeveelheid q die hij dagelijks kan verkopen daarom alleen zal afhangen van de prijs p (euro) die hij ervoor vraagt.
Er geldt: $q = 1200 - 30p$.
Voor de opbrengst R (euro) van de verkoop van dit product geldt: $R = p \cdot q$.

a Herleid de formule $q = 1200 - 30p$ naar de vorm $p = ...$

b Stel een formule op waarin R is uitgedrukt in q

Naast inkomsten (de opbrengst) zijn er ook kosten. Voor de productie van dit product is de fabrikant aan grondstoffen per stuk € 0,50 kwijt. Ook heeft hij er een machine voor aan moeten schaffen. Deze kost per dag € 1000,00.

c Stel een formule op voor de totale kosten per dag afhankelijk van de productie van q stuks.

De winst wordt berekend door van de opbrengst de kosten af te halen: $W = R - K$.

d Stel een formule op waarin W is uitgedrukt in q en herleid.

e Bereken bij welke hoeveelheid q de winst maximaal is.

Uitleg 2

Het subsidiebedrag B (euro) dat een sportclub jaarlijks ontvangt, hangt af van het aantal senioren s en het aantal junioren j. Er geldt: $B = 1000 + 10j + 5s$.

Ken je de waarden van s en j, dan kun je het subsidiebedrag berekenen.
Een sportclub met 60 junioren en 115 senioren ontvangt een subsidie van:
$B = 1000 + 10 \cdot 60 + 5 \cdot 115 = 2175{,}00$ euro.

De formule bevat drie variabelen. Door bijvoorbeeld voor B een rijtje vaste waarden te kiezen kun je er een grafiekenbundel bij tekenen.

Neem bijvoorbeeld $B = 1500$, $B = 2000$ en $B = 2500$.

Voor $B = 1500$ krijg je:

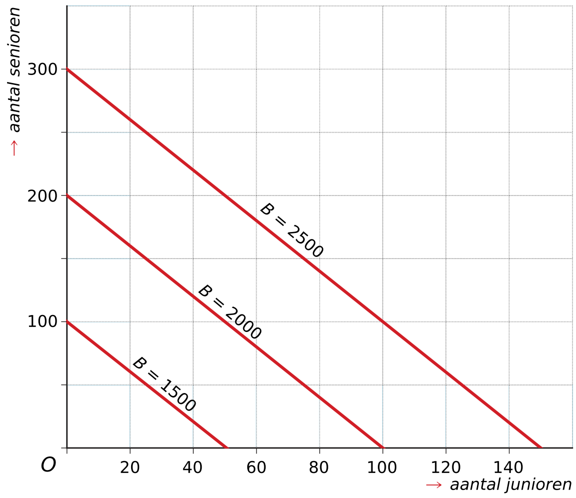
$$\begin{aligned} 1000 + 10j + 5s &= 1500 && \text{beide zijden } -1000 \\ 10j + 5s &= 500 && \text{beide zijden } -10j \\ 5s &= 500 - 10j && \text{beide zijden } : 5 \\ s &= 100 - 2j \end{aligned}$$

Je vindt zo:

- $B = 1500$ geeft $s = 100 - 2j$
- $B = 2000$ geeft $s = 200 - 2j$
- $B = 2500$ geeft $s = 300 - 2j$

Teken met behulp van een tabel of met de grafische rekenmachine de grafiekenbundel.

In een bepaald jaar ontvangt de club een subsidie van € 2500,00. Er zijn dat jaar 100 senioren. Met de grafiek kun je bepalen hoeveel junioren de club heeft. Je leest af dat de club 100 junioren heeft.
Het aantal junioren kun je ook berekenen door de vergelijking $2500 = 1000 + 10j + 5 \cdot 100$ op te lossen.



Figuur 1.2

Opgave 4

Gebruik de gegevens uit **Uitleg 2**.

a Waarom kan bij de formule voor het subsidiebedrag geen grafiek in een xy-assenstelsel getekend worden?

b Plot zelf de grafiekenbundel die hoort bij $B = 1500$, $B = 2000$ en $B = 2500$ op de grafische rekenmachine. Geef aan welke formules je invoert en welke vensterinstellingen je gebruikt.

c Stel, in een bepaald jaar is $B = 1600$. Teken de grafiek die hoort bij die situatie.

d Er zijn dat jaar 80 senioren. Hoeveel junioren zijn er dan?

Opgave 5

Gebruik de gegevens uit **Uitleg 2**.

a Hoeveel subsidie krijgt een sportclub met 45 junioren en 93 senioren?

b In een bepaald jaar was het aantal junioren precies drie keer zo groot als het aantal senioren. Daarbij hoort de vergelijking $j = 3s$. Gebruik dit om een formule te maken waarin B is uitgedrukt in s.

Opgave 6

Het subsidiebedrag B dat een sportclub jaarlijks ontvangt, hangt af van het aantal senioren s en het aantal junioren j.
Er geldt: $B = 800 + 20j + 10s$

a Hoeveel subsidie krijgt een sportclub met 60 junioren en 115 senioren?

b Maak een grafiekenbundel met $B = 1000$, $B = 1500$, $B = 2000$ en $B = 2500$.

c In een bepaald jaar ontvangt de club een subsidie van € 2600,00. Er zijn dat jaar 80 senioren. Bepaal met behulp van de grafiek hoeveel junioren de sportclub dat jaar heeft.

d Welke vergelijking kan er opgesteld worden bij c? Los deze vergelijking algebraïsch op.

e In een bepaald jaar was het aantal junioren precies twee keer zo groot als het aantal senioren. Daarbij hoort de vergelijking $j = 2s$. Gebruik dit om een formule te maken waarin B is uitgedrukt in s.

Theorie en voorbeelden

Om te onthouden

Soms heb je te maken met een of meer **formules met meerdere variabelen**. Alleen bij formules met twee variabelen van de vorm $y = ...$ (met op de puntjes een uitdrukking waarin alleen de variabele x voorkomt) kun je een grafiek maken in een xy-assenstelsel.

Van een formule met meer dan twee variabelen kun je grafieken bekijken door:

- meerdere formules te combineren tot één formule waarbij y is uitgedrukt in x. Daarbij moet je meestal een **uitdrukking invullen** in een formule en soms nog eerst een formule **herleiden**.
- een **grafiekenbundel** te maken door voor één of meer variabelen een vast aantal waarden te kiezen.

Voorbeeld 1

Hoeveel brandstof een personenauto verbruikt, hangt onder andere af van de af te leggen afstand, het aantal stops en het wachten voor verkeerslichten. Het brandstofverbruik B (mL) van een auto kan berekend worden met de formule: $B = a \cdot L + b \cdot S + c \cdot D$.
Hierin is:

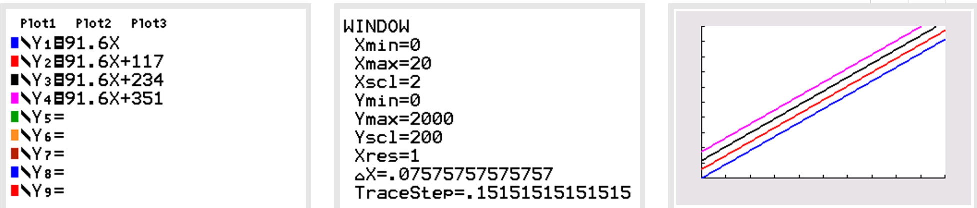
- L = ritlengte in km
- S = aantal stops onderweg
- D = totale wachttijd voor verkeerslichten in seconden

a en b zijn getallen die van de gemiddelde snelheid V (km/h) afhangen en c is een constante.
Voor a geldt: $a = 170 - 4{,}55V + 0{,}049V^2$.
Voor b geldt: $b = 0{,}0077V^2$.
Voor c geldt: $c = 0{,}39$.
Optrekken en afremmen worden buiten beschouwing gelaten, zodat in de uitdrukkingen voor a en b steeds een constante waarde voor V ingevuld kan worden.

- Neem een rit van één kilometer met een snelheid van 50 km/h, twee stops onderweg en een totale wachttijd van 40 seconden. Bereken het totale brandstofverbruik.
- Een vertegenwoordiger is voor zijn werk veel onderweg met de auto om klanten te bezoeken. Hij legt daarvoor afstanden af tussen de 0 en 20 km. Hij maakt geen stops onderweg en hij rijdt met een gemiddelde snelheid van 70 km/h. De totale wachttijd voor verkeerslichten varieert. Maak hiervoor een grafiekenbundel met daarin het brandstofverbruik B uitgezet tegen de ritlengte L. Neem voor de totale wachttijd voor verkeerslichten D een rijtje vaste waarden: $D = 0$, $D = 300$, $D = 600$ en $D = 900$.

Antwoord

- Voor deze rit geldt: $V = 50$. Invullen geeft:
  $a = 170 - 4{,}55 \cdot 50^2 + 0{,}049 \cdot 50^2 = 65$
  $b = 0{,}0077 \cdot 50^2 = 19{,}25$
  $c = 0{,}39$
  De waarden van L, S en D zijn gegeven. Invullen geeft:
  $B = 65 \cdot 1 + 19{,}25 \cdot 2 + 0{,}39 \cdot 40 = 119{,}1$ mL
- Voor de vertegenwoordiger geldt: $S = 0$ en $V = 70$. Invullen geeft:
  $a = 170 - 4{,}55 \cdot 70 + 0{,}049 \cdot 70^2 = 91{,}6$ en $b = 0{,}0077 \cdot 70^2 = 37{,}73$.
  Invullen in de formule geeft:
  $B = 91{,}6L + 37{,}73 \cdot 0 + 0{,}39D = 91{,}6L + 0{,}39D$
  Vul voor D het rijtje vaste waarden in:
  $D = 0$ geeft $B = 91{,}6L$
  $D = 300$ geeft $B = 91{,}6L + 117$
  $D = 600$ geeft $B = 91{,}6L + 234$
  $D = 900$ geeft $B = 91{,}6L + 351$
  Maak de grafiekenbundel met de grafische rekenmachine.



Figuur 1.3

Opgave 7

Gegeven is de formule uit **Voorbeeld 1**.

a Neem een rit van 10 kilometer met een snelheid van 80 km/h, twee stops onderweg en een totale wachttijd van 40 seconden.
Bereken het totale brandstofverbruik.

b Een andere vertegenwoordiger legt met de auto afstanden af tussen de 0 en 30 km. Zij maakt geen stops onderweg en rijdt met een gemiddelde snelheid van 90 km/h. De totale wachttijd voor verkeerslichten varieert. Maak hierbij een grafiekenbundel met daarin het brandstofverbruik B uitgezet tegen de ritlengte L. Neem voor de totale wachttijd voor verkeerslichten D een rijtje vaste waarden: $D = 0$, $D = 300$, $D = 600$ en $D = 900$.

Opgave 8

Bekijk **Voorbeeld 1** over de formule voor het brandstofverbruik.

a Iemand rijdt regelmatig een stuk over de snelweg met een snelheid van 130 km/h. Daar komt hij geen verkeerslichten tegen. Maak een grafiekenbundel met daarin het brandstofverbruik B uitgezet tegen het aantal stops S. Neem voor de ritlengte L een rijtje van vijf vaste waarden van 0 tot 40 km.

b De grafieken in de grafiekenbundel lijken bijna horizontaal te lopen. Leg uit wat dat betekent voor de invloed van het aantal stops op het brandstofverbruik.

Voorbeeld 2

Bij de verkoop van een bepaald artikel gelden de formules $TO = p \cdot q$ en $q = 450 - p$, waarin TO de totale maandelijkse opbrengst bij de verkoop van dat artikel is. p is de prijs (euro) en q is de verkochte hoeveelheid per maand.

- Combineer deze twee formules tot een formule voor TO afhankelijk van q.
- Voor de maandelijkse winst TW geldt: $TW = TO - TK$.
  $TK = 30q + 7500$ zijn de totale maandelijkse kosten voor dit artikel.
  Stel een formule op voor TW.
- Bij welke verkoopcijfers wordt winst gemaakt?

Antwoord

- $TO = p \cdot q = (\text{-}q + 450) \cdot q = \text{-}q^2 + 450q$
- $TW = TO - TK = (\text{-}q^2 + 450q) - (30q + 7500) = \text{-}q^2 + 420q - 7500$
- Voer in: $y_1 = \text{-}x^2 + 420x - 7500$.
  Venster bijvoorbeeld: $0 \leq x \leq 450$ en $\text{-}100 \leq y \leq 50000$.
  Nulpunten bij $x \approx 18{,}7$ of $x \approx 401{,}3$.
  Dus voor q van 19 tot en met 401 wordt winst gemaakt.

Opgave 9

Bij de verkoop van een bepaald artikel gelden de formules $TO = p \cdot q$ en $q = 300 - p$, waarin TO de totale maandelijkse opbrengst bij de verkoop van dat artikel is. p is de prijs (euro) en q is de verkochte hoeveelheid per maand.

a Combineer deze twee formules tot een formule voor TO afhankelijk van q.

b Voor de maandelijkse winst TW geldt: $TW = TO - TK$.
$TK = 40q + 6900$ zijn de totale maandelijkse kosten voor dit artikel.
Stel een formule op voor TW.

c Bij welke verkoopcijfers wordt winst gemaakt?

Opgave 10

Een fabriek produceert steps. Het bedrijf heeft als enige producent een monopoliepositie. Het aantal verkochte producten q, in duizendtallen, hangt uitsluitend af van de prijs p (euro): $q = 12 - 0{,}1p$. De bedrijfswiskundige heeft een model opgesteld voor de kosten van de productie van deze steps: $TK = 1{,}5q^3 - 22{,}5q^2 + 120q$. Hierin is TK gegeven in duizenden euro.

a Toon aan dat geldt: $p = 120 - 10q$.

b De prijs p is minimaal 0 euro.
Welke waarden kan q aannemen?

c Voor de totale opbrengst TO (euro) geldt de formule: $TO = p \cdot q$.
Stel een formule op voor de opbrengst TO uitgedrukt in q.

d Voor de totale winst TW (euro) geldt de formule: $TW = TO - TK$.
Stel een formule op voor de winst TW uitgedrukt in q.

e Bepaal de prijs van één step bij maximale winst.

f q is het aantal verkochte producten in duizendtallen. Geef een formule voor de gemiddelde totale kosten GTK per 1000 steps. Bepaal bij welk aantal verkochte producten de GTK minimaal is.

Verwerken

Opgave 11

Herleid tot de vorm $y = ...$

a $5x + 4y = 25$

b $x = 15 - 3y$

c $x \cdot y = 16$

d $\frac{2y}{5x} = 3$

Opgave 12

Gegeven zijn de formules: $r = 2p + q$ en $p = q + 3$
Vul de uitdrukking van p uit de tweede formule in de eerste formule in en herleid tot de vorm $r = aq + b$.

Opgave 13

De prijzen van een pretpark zijn € 30,00 per volwassene en € 7,50 per kind.

a Een gezin met vader, moeder en drie kinderen gaat naar het pretpark. Wat zijn de kosten voor de entree?

b Geef een formule voor de prijs P afhankelijk van het aantal volwassenen v en het aantal kinderen k.

c Teken een grafiekenbundel bij het aantal volwassenen: $v = 2,4,6,8,10$.

d Een basisschool gaat op schoolreis naar het pretpark. Per vier leerlingen is er één volwassene mee. Leg uit dat geldt $k = 4v$ en druk de prijs P uit in aantal volwassenen v.

e Leg aan de hand van de grafiek uit dat je kunt zien dat de prijs voor twee volwassenen en acht kinderen hetzelfde is als de prijs voor vier volwassenen zonder kinderen.

Opgave 14

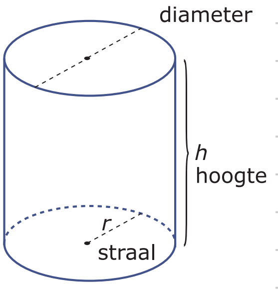
De oppervlakte A van een cilinder is $A = 2\pi r^2 + 2\pi r \cdot h$ met r en h in centimeter.

a Een blik met de vorm van een cilinder heeft een hoogte van 12 cm. Welke formule geldt voor de oppervlakte van dit blik?

De fabrikant van deze blikken wil dat de oppervlakte van het materiaal maximaal 10 dm^2 is.

b Plot de grafiek van A afhankelijk van r. Welke instellingen kies je?

c Welke straal heeft een blik met een oppervlakte van 10 dm^2? Geef je antwoord in millimeter nauwkeurig.



Figuur 1.4

Opgave 15

Een softwareontwikkelaar verkoopt softwarepakketten aan kleinere bedrijven. Deze softwareproducent rekent met de formule $p = 1200 - 3q$ om zijn prijs p (euro) te bepalen afhankelijk van het aantal pakketten q dat hij verkoopt. De kosten voor het versturen van dit pakket naar een klant bedragen € 10,00 per stuk.

a Voor de opbrengst R (euro) geldt de formule: $R = q \cdot p$.
Welke formule geldt voor de opbrengst R uitgedrukt in q?

b Welke formule geldt voor de kosten K uitgedrukt in q?

c Voor de winst W (euro) geldt de formule: $W = R - K$.
Stel een formule op voor de winst W. Schrijf de formule zonder haakjes.

d Breng de formule voor W volledig in beeld op de grafische rekenmachine en bereken de maximaal haalbare winst.

Opgave 16

ChemTech produceert een bepaald onkruidbestrijdingsmiddel. Voor de productiekosten per maand geldt:

q (duizend kg per maand)	1	2	3	4	5	6
TK (euro per maand)	775	1000	1220	2000	4000	8000

Tabel 1.1

Hierin is q de geproduceerde hoeveelheid per maand in duizenden kg en TK de totale kosten in euro. Verder verkoopt ChemTech dit middel voor € 2,25 per kg.

a De bedrijfsleiding heeft voor de kosten deze formule bedacht:
$TK = 100q^3 - 600q^2 + 1300q$
Laat zien dat deze formule redelijk goed bij de gegeven tabel past.

b Voor de totale winst TW (euro) geldt de formule: $TW = TO - TK$.
Hierin is TO de totale opbrengst (euro).
Stel een formule op voor de totale winst TW afhankelijk van q. Ga ervan uit dat de geproduceerde hoeveelheid elke maand ook wordt verkocht.

c Bepaal met de grafische rekenmachine bij welke productie per maand de winst maximaal is.

Toepassen

Opgave 17: Ontslagvergoeding

Bij grote bedrijven wordt er regelmatig gereorganiseerd. Hierbij vallen soms gedwongen ontslagen. Een ontslagen werknemer krijgt dan vaak eenmalig een ontslagvergoeding. Er zijn verschillende rekenmethodes om deze vergoeding te bepalen.
De eerste rekenmethode gebruikt de formule $V_1 = 0{,}5 \cdot m \cdot g$.
Hierin is V_1 de ontslagvergoeding (euro), m het bruto maandsalaris (euro) en g het aantal gewogen dienstjaren. De dienstjaren vóórdat iemand 40 jaar is, tellen elk voor 1, de volgende tien dienstjaren tellen elk voor $1{,}5$ en elk dienstjaar vanaf het moment dat iemand 50 jaar is geworden, telt voor 2.

a Bob Jansen is onlangs 52 jaar geworden. Hij werkt vanaf zijn dertigste verjaardag bij zijn huidige werkgever. Deze werkgever gaat reorganiseren met als gevolg dat Bob ontslagen wordt. Bob had een bruto maandsalaris van € 4300,00.
Toon aan dat Bob volgens de formule een ontslagvergoeding krijgt van meer dan € 60000,00.

De tweede rekenmethode gebruikt de formule $V_2 = 6 \cdot m + 2{,}4 \cdot m \cdot d$. Hierin is V_2 de ontslagvergoeding (euro), m het bruto maandsalaris (euro) en d het aantal dienstjaren, geteld vanaf de 36^e^ verjaardag. Er geldt dat $d = 0$ totdat de werknemer 37 jaar wordt. Op de dag dat de werknemer 37 jaar wordt, wordt $d = 1$. Op de 38^e^ verjaardag wordt $d = 2$, enzovoort. Het aantal dienstjaren wordt altijd in gehele jaren berekend.
Bij de tweede formule geldt een maximale ontslagvergoeding van 54 keer het bruto maandsalaris.

b Henk Klaassen werkt vanaf zijn 36^e^ verjaardag bij zijn huidige werkgever. Hij is bang dat hij bij een reorganisatie ontslagen wordt. Neem aan dat zijn maandsalaris door de jaren heen steeds gelijk blijft. Bereken na hoeveel dienstjaren hij voor het eerst zijn maximale ontslagvergoeding zou krijgen volgens de tweede formule.

Vakbonden hebben liever dat de formule voor V_2 gebruikt wordt dan de formule voor V_1. Toch is de formule voor V_2 niet altijd gunstiger. Er zijn situaties waarbij de eerste formule gunstiger is voor een ontslagen werknemer.

c Geef een rekenvoorbeeld van zo'n situatie en geef daarbij aan op welke leeftijd de werknemer in dienst is getreden en op welke leeftijd hij ontslagen wordt.

Met de tweede formule kan de ontslagvergoeding V_2 worden berekend als het bruto maandsalaris m en het aantal dienstjaren d, geteld vanaf de 36^e^ verjaardag, bekend zijn. De formule kan worden herschreven tot een formule die uitgaat van het bruto jaarsalaris in plaats van het bruto maandsalaris. Een jaarsalaris is meer dan 12 maandsalarissen, omdat een werknemer ook vakantiegeld en een eindejaarsuitkering uitbetaald krijgt. Hierdoor geldt dat het bruto jaarsalaris j gelijk is aan $13{,}5$ keer het bruto maandsalaris m.

d De tweede formule is daarmee te schrijven in de vorm: $V_2 = \ldots \cdot j + \ldots \cdot j \cdot d$.
Bereken de getallen die op de puntjes moeten staan. Rond af op twee decimalen. Licht je werkwijze toe.

(naar: examen havo wiskunde A in 2013, eerste tijdvak)

Opgave 18: Verf

Voordat je met verven begint, is het handig om te weten hoeveel (blikken) verf je nodig hebt. Omgekeerd kun je je ook afvragen hoeveel vierkante meter je kunt verven met één blik verf. Afhankelijk van het soort kwast dat wordt gebruikt, verlies je tussen de 5 en 10 procent van de verf. Het verband tussen deze zaken staat in de volgende formule, waarin ook rekening is gehouden met verlies van verf door gebruik van de kwast: $H = \frac{10 \cdot A \cdot d}{V \cdot (100 - p)}$.

Hierin is:

- H de hoeveelheid verf (liter)
- A de oppervlakte (m^2)
- d de dikte van de verflaag (micrometer)
- V het percentage vaste stof
- p het verliespercentage bij kwasten; dit varieert van 5 tot 10

a Stel je wilt verf gebruiken die wordt verkocht in blikken van $2{,}5$ liter. Op de blikken staat dat het percentage vaste stof 35 is. Met een kwast wil je een verflaag van 70 micrometer dikte aanbrengen. Bereken hoeveel vierkante meter je met zo'n blik verf maximaal kunt schilderen.

b Voor een andere kamer koop je verf met een percentage vaste stof van 55. De oppervlakte die je wilt verven is 90 m^2. Maak een grafiekenbundel met daarin de hoeveelheid verf H uitgezet tegen de dikte van de verflaag d. Neem voor het verliespercentage p het rijtje vaste waarden: $p = 5$, $p = 6$, ..., $p = 10$.

c De lijnen van de grafiekenbundel uit b lopen heel dicht bij elkaar. Wat zegt dat over de invloed van het verliespercentage van kwasten p op de oppervlakte A die geverfd kan worden?

Iemand heeft 15 liter verf gekocht met een percentage vaste stof van 67. Hij gaat een verflaag van 60 micrometer dikte aanbrengen. Met deze gegevens ingevuld, is de formule: $15 = \frac{10 \cdot A \cdot 60}{67 \cdot (100 - p)}$.

In deze formule is te zien dat de oppervlakte A die hij met deze hoeveelheid kan verven nu alleen nog afhangt van het verliespercentage p. Het verband tussen A en p is lineair. De formule is dus te herschrijven tot een formule van de vorm $A = a \cdot p + b$.

d Bereken a en b.

(naar: examen havo wiskunde A in 2009, tweede tijdvak)

Testen

Opgave 19

Een ondernemer maakt een bepaald product waarop hij het monopolie heeft. Voor zijn productiekosten (in honderden euro) geldt de formule $K = 0{,}5q^3 - 4q^2 + 11q + 4$ waarin q de geproduceerde hoeveelheid in honderden kilogram is.

De hoeveelheid product die hij aanbiedt aan zijn afnemers heeft invloed op de prijs. Er geldt: $p = 11 - q$ waarin p de prijs in honderden euro is. Ga er van uit dat deze ondernemer zijn totale productie ook verkoopt.

a De winst W kan worden uitgedrukt in q. Geef de bijbehorende formule.

b Maak de grafiek van de functie W op je grafische rekenmachine en bereken bij welke verkochte hoeveelheid zijn winst zo groot mogelijk is.

Opgave 20

Gegeven zijn de formules $P = 3q + 5r + 22$ en $r = 4q - 2$.

Combineer de formules, en stel formules op in de vorm $P = ar + b$ en $r = cP + d$.
Welke getallen zijn a, b, c, en d?

Practicum

Als je met functies werkt, wil je alle karakteristieken (nulpunten, toppen en asymptoten) in beeld. In het volgende practicum kun je nalezen hoe dat gaat. En hoe je dan maxima en minima kunt berekenen, nulpunten kunt berekenen, etc.

- **Functies en de TI84**
- **Functies en de TInspire**
- **Functies en de Casio fx-CG50**
- **Functies en de HPprime**
- **Functies en de NumWorks**

Inleiding

Je maakt een heg van thuja's en jeneverbessen. Thuja's kosten € 15,75 per stuk, jeneverbessen kost € 27,50 per stuk, want je wilt al meteen flinke struiken hebben. Je wilt maximaal 50 struiken planten en niet meer dan € 1000,00 uitgeven.
Zo'n probleem kun je beschrijven met behulp van ongelijkheden als je bijvoorbeeld het aantal thuja's x en het aantal jeneverbessen y noemt.

Figuur 2.1 jeneverbes

Je leert in dit onderwerp

- lineaire ongelijkheden zoals $px + qy \leq r$ oplossen (dus het bijpassende gebied tekenen);
- gebieden ingesloten door lineaire ongelijkheden interpreteren;
- gebieden binnen een grafiekenbundel interpreteren.

Voorkennis

- formules herleiden naar een vorm waarin y is uitgedrukt in x;
- getallen en uitdrukkingen in formules invullen;
- bij een formule met drie of meer variabelen een grafiekenbundel tekenen.

Verkennen

Opgave V1

Je maakt een heg van thuja's en jeneverbessen. Thuja's kosten € 15,75 per stuk, jeneverbessen kost € 27,50 per stuk, want je wilt al meteen flinke struiken hebben. Je wilt maximaal 50 struiken planten en niet meer dan € 1000,00 uitgeven.
Zo'n probleem kun je beschrijven met behulp van ongelijkheden als je bijvoorbeeld het aantal thuja's x en het aantal jeneverbessen y noemt.

a Stel twee ongelijkheden op die bij dit probleem passen.

b Hoeveel van beide planten ga je kopen?

Uitleg 1

Je maakt een heg van thuja's en jeneverbessen. Thuja's kosten € 15,75 per stuk, jeneverbessen kosten € 27,50 per stuk. Je wilt maximaal 50 struiken planten en niet meer dan € 1000,00 uitgeven. Hoeveel struiken van iedere soort kun je dan kopen?

Zo'n probleem beschrijf je met behulp van ongelijkheden. Noem bijvoorbeeld het aantal thuja's x en het aantal jeneverbessen y. Nu kunnen er twee ongelijkheden worden opgesteld:

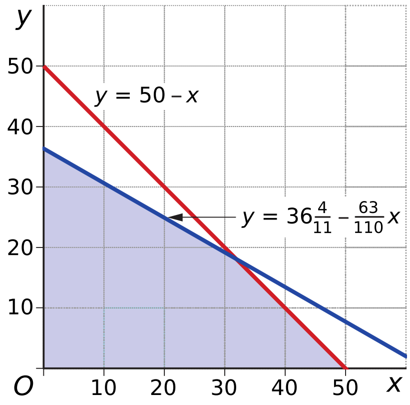
- $x + y \leq 50$
- $15{,}75x + 27{,}5y \leq 1000$

Bij de grenzen van deze ongelijkheden kun je vergelijkingen van de vorm $y = ...$ maken, zodat ze kunnen worden ingevoerd op de grafische rekenmachine.

- $x + y = 50$ geeft $y = 50 - x$
- $15{,}75x + 27{,}5y = 1000$ geeft $y = 36\frac{4}{11} - \frac{63}{110}x$

De rode grafiek hoort bij $y = 50 - x$ en het gebied daaronder voldoet aan de ongelijkheid $x + y \leq 50$.

De blauwe grafiek hoort bij $y = 36\frac{4}{11} - \frac{63}{110}x$ en het gebied daaronder voldoet aan de ongelijkheid $15{,}75x + 27{,}5y \leq 1000$.
Het paarse gebied voldoet dus aan beide ongelijkheden. Daarom zijn alle punten binnen het paarse gebied oplossingen van het probleem. Dit gebied heeft als grenslijnen de vergelijkingen $x + y = 50$ en $15{,}75x + 27{,}5y = 1000$. In het **Practicum** zie je hoe dit met een grafische rekenmachine kan.
Zo is bijvoorbeeld het punt $(20,10)$ een juiste oplossing. Er worden dan in totaal $20 + 10 = 30$ struiken gekocht, en de kosten daarvoor zijn: $15{,}75 \cdot 20 + 27{,}50 \cdot 10 = 590{,}00$ euro.



Figuur 2.2

Opgave 1

Gebruik de gegevens uit **Uitleg 1**.

a Leg uit waarom de ongelijkheden $x + y \leq 50$ en $15{,}75x + 27{,}5y \leq 1000$ bij het probleem horen.

b De rode grafiek hoort bij $y = 50 - x$. Aan welke ongelijkheid voldoet het gebied onder de rode grafiek? En onder de blauwe grafiek?

c Waarom bevinden zich binnen het paarse gebied de oplossingen van beide ongelijkheden?

d Leg uit hoe het gebied bij twee ongelijkheden getekend kan worden.

e Laat met een berekening zien dat het punt $(10,30)$ een juiste oplossing is.

f Laat met een berekening zien dat het punt $(20,30)$ geen juiste oplossing is.

Opgave 2

Iemand wil een heg van liguster en laurier maken. Ligusterstruiken kosten € 6,00 per stuk, laurierstruiken kosten € 13,50 per stuk. Ze wil maximaal 30 struiken planten en niet meer dan € 300,00 uitgeven.

a Stel twee ongelijkheden op bij dit probleem.

b Schrijf de vergelijkingen van de grenslijnen van de ongelijkheden in de vorm $y = ...$

c Plot de grafieken van beide grenslijnen. Kleur het gebied dat voldoet aan beide ongelijkheden.

d Ga na welk van deze punten juiste oplossingen zijn van het probleem: $(5{,}10)$, $(30{,}0)$, $(25{,}10)$, $(15{,}15)$.

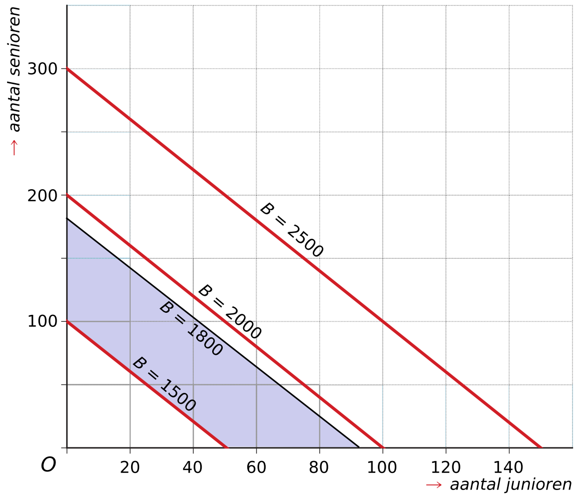
Uitleg 2

Het subsidiebedrag B (euro) dat een sportclub jaarlijks ontvangt, hangt af van het aantal senioren s en het aantal junioren j. Er geldt:
$B = 1000 + 10j + 5s$
Je kunt een grafiekenbundel bij deze formule maken.
Al jarenlang ontvangt men een subsidie tussen de € 1500,00 en de € 1800,00. In de grafiekenbundel kun je het bijbehorende gebied aanduiden.

Kleur het gebied waarvoor geldt: B ligt tussen de 1500 en 1800.
$B = 1500$ staat al in de grafiekenbundel.
$B = 1800$ levert de formule:

$$\begin{aligned} 1800 &= 1000 + 10j + 5s \\ 10j + 5s &= 800 \\ 2j + s &= 160 \\ s &= 160 - 2j \end{aligned}$$

Teken de lijn die voldoet aan de formule in de grafiekenbundel erbij en kleur het gebied tussen $B = 1500$ en $B = 1800$.



Figuur 2.3

Opgave 3

Gebruik de gegevens uit **Uitleg 2**.

a Wat is de betekenis van het gekleurde gebied in de grafiekenbundel?

b Het aantal senioren blijft al jaren constant, met tachtig personen. Geef deze situatie in de figuur aan door een lijn te trekken.

c Geef de snijpunten van de lijn uit b met de grafieken die horen bij $B = 1500$ en $B = 1800$. Wat is de betekenis van deze snijpunten?

d Stel dat het subsidiebedrag al jaren tussen € 2000,00 en € 2500,00 schommelt, welk gebied hoort daar dan bij?

e Stel dat het subsidiebedrag al jaren tussen € 2000,00 en € 2500,00 schommelt, en dat het aantal senioren al jaren constant rond de 80 zit. Tussen welke aantallen heeft het aantal junioren dan gevarieerd? Hoe is dat af te lezen uit de grafiekenbundel?

Opgave 4

Het subsidiebedrag B dat een sportclub jaarlijks ontvangt, hangt af van het aantal senioren s en het aantal junioren j. Er geldt: $B = 800 + 20j + 10s$.

a Maak de bijbehorende grafiekenbundel met $B = 1000$, $B = 1500$, $B = 2000$ en $B = 2500$.

b Al jarenlang ontvangt men een subsidie tussen de € 2500,00 en de € 2800,00. Geef in de grafiek het bijbehorende gebied aan.

c Het aantal senioren blijft ook al jaren constant, ongeveer tachtig personen. Tussen welke aantallen heeft het aantal junioren dan gevarieerd? Geef dit in de figuur aan.

Theorie en voorbeelden

Om te onthouden

Ongelijkheden met twee variabelen hebben niet één oplossing, maar een **gebied** met alle mogelijke oplossingen van het bijbehorende probleem. Wanneer bij zo'n probleem twee (of meer) ongelijkheden horen, bepalen de bijbehorende vergelijkingen de **grenslijnen** van het gebied.

Ook in een grafiekenbundel kun je gebieden aangeven. Zo'n gebied bevindt zich tussen twee grafieken van de bundel en kun je met ongelijkheden beschrijven. In het **Practicum** zie je hoe een grafische rekenmachine zo'n gebied kan arceren.

Voorbeeld 1

Teken het gebied bij de lineaire ongelijkheid $4x + 5y \leq 40$ in een xy-assenstelsel.

Antwoord

Om een lineaire ongelijkheid te tekenen schrijf je de bijbehorende vergelijking in de vorm van $y = ...$
Daarna teken je de grafiek van de vergelijking. Bepaal door middel van twee controlepunten aan welke kant van de lijn het juiste gebied zit. Kleur of arceer het gebied dat voldoet aan de ongelijkheid. In het **Practicum** zie je hoe dit met een grafische rekenmachine kan.

LET OP: Bij $<$ en $>$ doet de grenslijn niet mee met het gebied, de grenslijn is dan een stippellijn. Bij $\leq$ en $\geq$ doet de grenslijn wel mee met het gebied, de grenslijn is dan een doorgetrokken lijn.

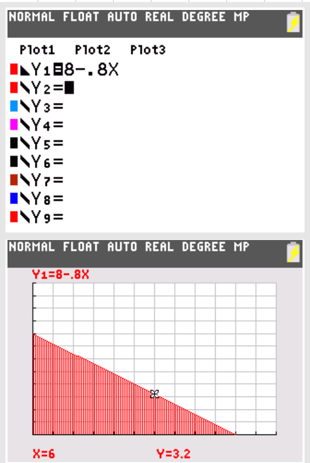
De bijbehorende vergelijking is: $y_1 = 8 - \frac{4}{5}x$.

Voer de vergelijking in op de grafische rekenmachine.
Venster bijvoorbeeld: $-4 \leq x \leq 12$ en $-4 \leq y \leq 12$.

Om het gebied te bepalen kies je aan beide kanten van de lijn een controlepunt.
Bijvoorbeeld: $(6,6)$ en $(4,2)$.
Onderzoek welk van de controlepunten aan de ongelijkheid voldoet.

$4{,}6 + 5{,}6 = 40$: klopt niet.
$4{,}4 + 5{,}2 = 40$: klopt.

Het juiste gebied bevindt zich dus op en onder de grenslijn $y = 8 - \frac{4}{5}x$.



Figuur 2.4

Opgave 5

Bekijk **Voorbeeld 1** over het tekenen van de lineaire ongelijkheid $4x + 5y \leq 40$.

a Welk gebied hoort bij de ongelijkheid $4x + 5y \geq 40$?

b Welk gebied hoort bij de ongelijkheid $4x + 5y < 40$?

Opgave 6

Teken de lineaire ongelijkheid $3x + 6y > 22$ in een assenstelsel.

Voorbeeld 2

Iemand wil sla en prei in haar moestuin zaaien. Een zakje zaadjes voor sla kost € 0,55 en een zakje zaadjes voor prei kost € 0,25. Ze wil maximaal 40 zakjes zaad kopen, en maximaal € 18,00 uitgeven.
Hoeveel zakjes slazaad en hoeveel zakjes preizaad kan ze kopen? Maak een schets van het gebied waarin de oplossingen van dit probleem zich bevinden. Bepaal het snijpunt van de grenslijnen en leg uit wat de betekenis daarvan is.

Antwoord

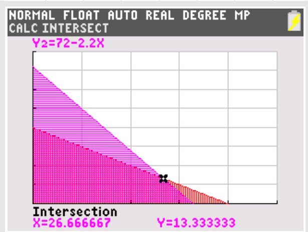
Neem x voor het aantal zakjes slazaad en y voor het aantal zakjes preizaad. Dan kunnen er twee ongelijkheden opgesteld worden:

- $x + y \leq 40$
- $0{,}55x + 0{,}25y \leq 18$

Schrijf beide bijbehorende vergelijkingen in de vorm $y = ...$ zodat ze kunnen worden ingevoerd op de grafische rekenmachine:

- $y = 40 - x$
- $y = 72 - 2{,}2x$

Voer in: $y_1 = 40 - x$ en $y_2 = 72 - 2{,}2x$.
Venster bijvoorbeeld: $0 \leq x \leq 50$ en $0 \leq y \leq 80$.



Figuur 2.5

Het dubbel gearceerde gebied voldoet aan beide ongelijkheden. Daarom zijn alle punten binnen dit gebied oplossingen van het probleem. De grenslijnen hebben de vergelijkingen $x + y = 40$ en $0{,}55x + 0{,}25y = 18$.

Bij het snijpunt van de grenslijnen is de verhouding tussen het aantal zakjes slazaad en het aantal zakjes preizaad precies zo dat het totaalbedrag gelijk is aan het maximum van € 18,00 en het aantal zakjes gelijk is aan het maximum van 40.
Sijpunten bij: $x \approx 26{,}7$ en $y \approx 13{,}3$. In dit geval moet x naar beneden worden afgerond omdat je anders over het maximum bedrag heen gaat. Dus $x = 26$ en $y = 14$.

Opgave 7

Bekijk **Voorbeeld 2**.

a Laat zien hoe je aan $y_1 = 40 - x$ en $y_2 = 72 - 2{,}2x$ komt.

b Bereken het snijpunt van de grenslijnen zonder de grafische rekenmachine te gebruiken.

c In het punt $(10{,}30)$ heb je ook 40 plantjes. Ben je dan goedkoper of duurder uit.

Opgave 8

Een paar kinderen verkopen limonade en zelfgebakken cakejes op de vrijmarkt. Een bekertje limonade verkopen ze voor € 1,50 en een cakeje voor € 2,00. Ze hebben 50 cakejes bij zich en 60 bekertjes limonade. Ze willen hiermee geld inzamelen voor een goed doel, minimaal € 150,00. De kosten voor de ingrediënten en het materiaal krijgen ze van hun ouders, dus al het geld dat ze inzamelen is winst.

a Stel drie ongelijkheden op bij dit probleem.

b Plot het gebied waarin de oplossingen van dit probleem zich bevinden.

c De cakejes blijken erg goed te verkopen. Al snel hebben ze alle cakejes verkocht. Wat is het minimale aantal bekertjes limonade dat ze moeten verkopen om hun doel te behalen?

d Hoe hoog is de maximale winst die ze kunnen behalen?

Voorbeeld 3

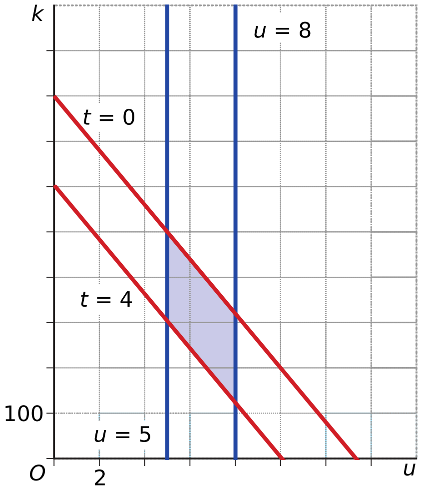
Een gemeente wil het water in haar buitenzwembad op 20 °C houden. De gemeente legt hiervoor een verwarmingsinstallatie aan. Omdat je in de zomermaanden ook van de warmte van de zon kunt profiteren, voorspelt een verwarmingsdeskundige dat de verwarmingskosten k zullen voldoen aan de formule: $k = 800 - 60u - 50t$
Hierin is u het gemiddeld aantal zonuren per dag en t het gemiddeld aantal graden Celsius dat de buitentemperatuur afwijkt van de 20 °C.
k wordt gerekend in euro per dag.

De hele maand juli is bijgehouden wat de gemiddelde buitentemperatuur was. Deze schommelde tussen 20 °C en 24 °C. Het aantal zonuren schommelde tussen de 5 en 8 uur per dag.

- Teken het gebied dat bij de maand juli hoort.
- Welke ongelijkheden beschrijven dit gebied?
- Tussen welke bedragen varieerden de verwarmingskosten in de maand juli?

Antwoord

- De gemiddelde buitentemperatuur schommelde tussen 20 °C en 24 °C, dus t schommelde tussen 0 en 4. Het gebied zit dus tussen $t = 0$ en $t = 4$.
  $t = 0$ geeft $k = 800 - 60u$.
  $t = 4$ geeft $k = 800 - 60u - 200$ en dus $k = 600 - 60u$.
  Het aantal zonuren schommelde tussen de 5 en 8 uur per dag, dus tussen $u = 5$ en $u = 8$.
  Teken nu het juiste gebied.
- Het gebied wordt omschreven door de ongelijkheden:
  $k \leq 800 - 60u$, $k \geq 600 - 60u$, $u \geq 5$ en $u \leq 8$.
- De verwarmingskosten in juli zijn maximaal in de linkerbovenhoek van het gebied en minimaal in de hoek rechtsonder. Vul $u = 5$ in de formule van $t = 0$ in, en vul $u = 8$ in de formule van $t = 4$ in. De maximale kosten waren: $800 - 60 \cdot 5 = 500{,}00$ euro. De minimale kosten waren: $600 - 60 \cdot 8 = 120{,}00$ euro.



Figuur 2.6

Opgave 9

Een andere gemeente wil het water in haar buitenzwembad ook op 20 °C houden. Voor hen geldt de formule: $k = 1000 - 80u - 60t$, waarin u het gemiddeld aantal zonuren per dag is en t het gemiddeld aantal graden Celsius dat de buitentemperatuur afwijkt van de 20 °C. k wordt gerekend in euro per dag.

a De hele maand augustus is bijgehouden wat de gemiddelde buitentemperatuur iedere dag was. Deze schommelde tussen 19 °C en 23 °C. Het aantal zonuren schommelde tussen de 5 en 7 uur per dag.
Teken het gebied dat bij de maand augustus hoort.

b Welke ongelijkheden beschrijven dit gebied?

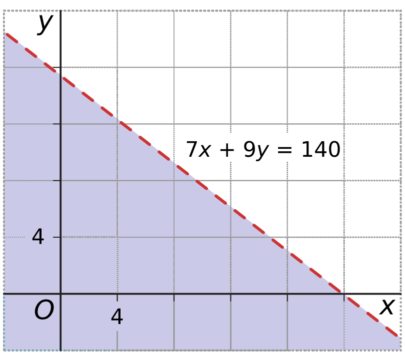
c Tussen welke bedragen varieerden de verwarmingskosten in de maand augustus?

Verwerken

Opgave 10

Aan welke ongelijkheid voldoet het gekleurde gebied?

A. $7x + 9y < 140$
B. $7x + 9y > 140$
C. $7x + 9y \leq 140$
D. $7x + 9y \geq 140$



Figuur 2.7

Opgave 11

Teken het gebied dat voldoet aan de gegeven ongelijkheid.

a $5x + 8y \leq 30$

b $10x - 7y > 45$

Opgave 12

Een stel dat gaat trouwen heeft aan de gasten geld gevraagd voor de huwelijksreis. Vrienden van het stel besluiten er iets grappigs van te maken door het hele bedrag in losse munten van € 0,50 en € 0,05 te geven. Ze stoppen deze munten in een lange koker waarin alle munten op elkaar gestapeld worden. Deze koker is 1 meter lang. Een munt van € 0,50 is 2 mm dik, een munt van € 0,05 is 1,5 mm dik. In totaal willen ze een bedrag van maximaal € 175,00 in de koker stoppen.

a Stel ongelijkheden op en breng het gebied waarin de oplossingen van dit probleem zich bevinden in beeld.

b Bepaal het snijpunt van de grenslijnen en leg uit wat de betekenis daarvan is.

Opgave 13

Het energieverbruik bij volwassen mannen kan berekend worden met de Harris- en Benedict-formule die in 1984 herzien is door Roza en Shizgal. De formule is: $E = 88{,}362 + 13{,}397G + 4{,}799H - 5{,}677L$. Hierin is E het energieverbruik per 24 uur in rust in kcal, G het gewicht in kg, H de lichaamslengte in centimeter en L de leeftijd in jaar.

a Meneer Jackson is een bejaarde man met een lengte van 1,83 m. Zijn gewicht schommelde gedurende zijn leven regelmatig. Maak voor meneer Jackson een grafiekenbundel met daarin zijn energieverbruik in rust uitgezet tegen zijn gewicht, en neem daarbij voor zijn leeftijd de vaste waarden $L = 20$, $L = 40$, $L = 60$ en $L = 80$.

b Tussen zijn 20ste en zijn 40ste schommelde het gewicht van meneer Jackson tussen de 70 en 100 kg.
Schets de grafiekenbundel en kleur daarin het gebied dat hierbij hoort.

c Welke ongelijkheden beschrijven dit gebied?

d Tussen welke waarden kan het energieverbruik in rust van meneer Jackson tussen zijn 20ste en 40ste levensjaar gevarieerd hebben?

Opgave 14

Willemijntje heeft € 2,50 gekregen om iets lekkers te kopen in een snoepwinkel. Toverballen kosten € 0,40 per stuk, kauwgomballen kosten € 0,15 per stuk. Willemijntje wil maximaal 10 stuks kopen. Stel ongelijkheden op bij dit probleem. Maak een schets en kleur het gebied dat voldoet aan de ongelijkheden.

Opgave 15

Een bedrijf produceert en verkoopt twee producten A en B. Het bedrijf heeft per week 2100 arbeidsuren beschikbaar en 3200 machine-uren. Per 1000 producten van soort A zijn er 12 arbeidsuren en 23 machine-uren nodig. Per 1000 producten van soort B zijn er 18 arbeidsuren en 15 machine-uren nodig. Het bedrijf moet van zowel product A als product B minstens 25000 stuks per week leveren.

a Stel voor dit bedrijf vier ongelijkheden op. Neem x voor het aantal te produceren producten A in duizendtallen. Neem y voor het aantal te produceren producten B in duizendtallen.

b Teken de bijbehorende grafieken. Maak een schets en kleur het gebied dat aan alle vier de ongelijkheden voldoet.

c De optimale combinatie is het grootste aantal te produceren producten A en B dat aan alle ongelijkheden voldoet. Waar in het gebied bevindt zich de optimale combinatie? Hoeveel producten A en B worden er dan geproduceerd? Rond af op duizendtallen.

d Het bedrijf wil zo veel mogelijk winst maken. De totale constante kosten per week zijn € 175000,00. Voor product A geldt een omzet per 1000 verkochte producten van € 2300,00 en voor product B is dat € 2800,00. Hoeveel bedraagt de maximale winst per week voor dit bedrijf?

Toepassen

Opgave 16: Formule van Strouhal

Uit biologisch onderzoek blijkt dat vogels, vleermuizen en insecten op een vergelijkbare manier met hun vleugels bewegen als vissen met hun staartvin. Onderzoekers hebben een verband ontdekt tussen de slagfrequentie (het aantal slagen per seconde van de vleugels of staartvin), de slaggrootte (de afstand tussen de uiterste staartvin- of vleugelstanden tijdens een slag) en de kruissnelheid (de gemiddelde snelheid).

Voor dieren als vissen, dolfijnen, vogels en insecten is het verband hetzelfde. Er geldt namelijk: $\frac{f \cdot d}{v} = 0{,}3$.

Dit wordt wel de formule van Strouhal genoemd.

In deze formule is:

- f de slagfrequentie (het aantal slagen per seconde van de vleugels of staartvin)
- d de slaggrootte (meter)
- v de kruissnelheid (meter per seconde)

a De kolibrie is een klein vogeltje dat vliegt met een hoge slagfrequentie. Een kolibrie heeft een slaggrootte van 8 cm en een kruissnelheid van $13{,}5$ meter per seconde.
Toon aan dat een kolibrie een slagfrequentie van ruim 50 heeft.

b Teken de grafiekenbundel met de slaggrootte d uitgezet tegen de slagfrequentie f. Neem voor de kruissnelheid v een rijtje van vier vaste waarden tussen 8 en 20 m/s. Zorg dat in de grafiek een slaggrootte tot 25 cm af te lezen is.

c Kleine vogels hebben een slaggrootte tussen 9 en 12 cm en een slagfrequentie tussen 40 en 60 slagen per seconde. Geef in de grafiekenbundel het gebied aan dat bij kleine vogels hoort.

d Wat kun je zeggen over de kruissnelheid v van kleine vogels?

(naar: examen havo wiskunde A in 2007, tweede tijdvak)

Opgave 17: Verf (2)

Voordat je met verven begint, is het handig om te weten hoeveel (blikken) verf je nodig hebt. Omgekeerd kun je je ook afvragen hoeveel vierkante meter je kunt verven met één blik verf. Afhankelijk van het soort kwast dat wordt gebruikt, verlies je tussen de 5 en 10 procent van de verf. Het verband tussen deze zaken staat in de volgende formule, waarin ook rekening is gehouden met verlies van verf door gebruik van de kwast: $H = \frac{10 \cdot A \cdot d}{V \cdot (100-p)}$.

Hierin is:

- H de hoeveelheid verf (liter)
- A de oppervlakte (m^2)
- d de dikte van de verflaag (micrometer)
- V het percentage vaste stof
- p het verliespercentage bij kwasten

a Je koopt verf met een percentage vaste stof van 45 en een kwast met een verliespercentage van 8.
Teken voor die situatie een grafiekenbundel met de hoeveelheid verf H uitgezet tegen de oppervlakte A. Neem voor de dikte van de verflaag d een rijtje vaste waarden tussen 0 en 100 micrometer. Neem de vensterinstellingen zo dat de grafiek af te lezen is tot 10 liter verf.

b Je verwacht dat de dikte van de verflaag tussen de 40 en 60 micrometer zal zijn. Kleur het gebied in de grafiekenbundel dat daarbij hoort.

c Je koopt 6 liter verf. Geef deze situatie in de figuur aan door een lijn te trekken.

d Hoe groot is de oppervlakte die je kunt verven?

Testen

Opgave 18

Gegeven is de ongelijkheid: $y > 3x - 9$.
Teken het gebied dat eraan voldoet. Geef duidelijk aan of de grenslijn wel of niet bij het gebied hoort.

Opgave 19

Evi heeft in een winkel leuke armbandjes gezien en wil er een aantal kopen. Er zijn twee soorten. Soort 1 kost € 1,50 per stuk en soort 2 kost € 2,75 per stuk. Evi wil maximaal € 10,00 aan de armbandjes uitgeven.
Stel een lineaire ongelijkheid op bij dit probleem en teken het bijbehorende gebied in een assenstelsel.

Opgave 20

Een wielrenner moet tijdens een circa vier uur durende etappe van de Tour de France (afhankelijk van de etappe en het weer) minimaal 3000 kcal binnen krijgen. Hiervoor kan hij flesjes sportdrank drinken, die bevatten 150 kcal per fles van 0,5 L. Verder kan hij energierepen eten, die bevatten 220 kcal per reep. De coach adviseert de wielrenner om tijdens de etappe 3 tot 7 energierepen te eten. Tijdens het fietsen verliest een wielrenner 1 tot 2 liter vocht per uur, dat moet aangevuld worden met sportdrank.

a Stel ongelijkheden op en teken het gebied dat voldoet aan de caloriebehoeften van de wielrenner.

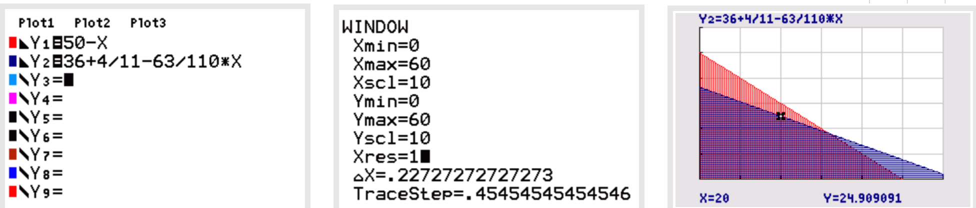
b De wielrenner krijgt onderweg van zijn coach niet meer sportdrank en energierepen dan volgens dit gebied nodig is. Bereken het aantal kcal dat de wielrenner tijdens een etappe maximaal binnen kan krijgen.

Practicum

Als je met functies werkt, wil je alle karakteristieken (nulpunten, toppen en asymptoten) in beeld. In het volgende practicum kun je nalezen hoe dat gaat. En hoe je dan maxima en minima kunt berekenen, nulpunten kunt berekenen, etc.

- **Functies en de TI-84**
- **Functies en de TInspire**
- **Functies en de Casio fx-CG50**
- **Functies en de HPprime**
- **Functies en de NumWorks**

Op sommige grafische rekenmachines kun je het **gebied arceren** dat onder of boven een ingevoerde formule ligt. Hier zie je hoe dat met TI-84 gaat voor de figuur in de **Uitleg 1**.
Door op het tekentje voor Y1= te gaan staan en [ENTER] te drukken kun je de vorm van de grafiek aanpassen en kiezen voor een arcering onder of een arcering boven de grafiek. Ook kun je de kleur instellen.



Figuur 2.8

1.3 Omgekeerd evenredig

Inleiding

Ga je met de auto van het centrum van Apeldoorn naar dat van Deventer, dan geeft de ANWB-routeplanner aan dat je een stuk van 16 km op de snelweg moet rijden. Hoe sneller je rijdt, hoe korter je over die 16 km doet. De tijd die je nodig hebt is omgekeerd evenredig met de snelheid.
Maar als je onderweg moet tanken, ligt dit allemaal weer anders...

Je leert in dit onderwerp

- met recht evenredige, omgekeerd evenredige en hyperbolische verbanden werken.

Voorkennis

- werken met grafieken en formules, ook met de grafische rekenmachine;
- formules herleiden, vergelijkingen en ongelijkheden oplossen, ook algebraïsch;
- werken met breuken.

Verkennen

Opgave V1

Ga je met de auto van het centrum van Apeldoorn naar dat van Deventer, dan geeft de ANWB-routeplanner aan dat je een stuk van 16 km op de snelweg moet rijden. Hoe sneller je rijdt, hoe korter je over die 16 km doet. De tijd die je nodig hebt is omgekeerd evenredig met de snelheid.

a Je mag maximaal 120 km/h rijden op de snelweg. Hoeveel minuten ben je dan onderweg?

b Het is druk dus je rijdt (gemiddeld) 80 km/h. Hoeveel minuten ben je onderweg?

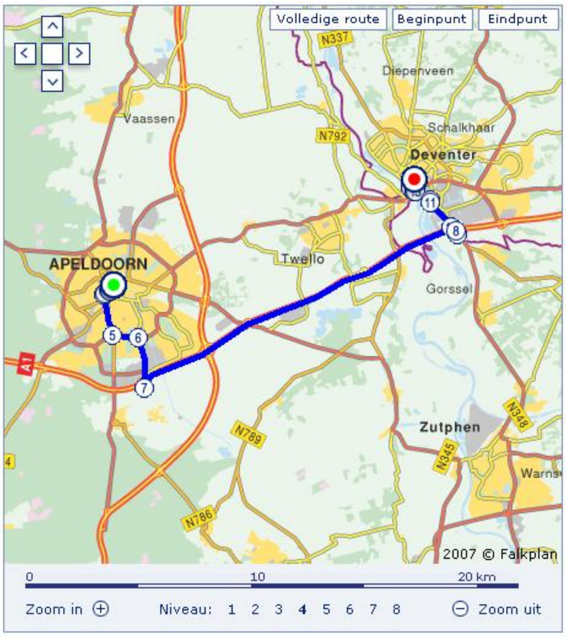
c Welke formule kun je opstellen voor de reistijd t (in min.) als functie van de snelheid v (in km/h)?

Uitleg 1

Van Apeldoorn naar Deventer is het met de auto 16 km over de snelweg. Met een snelheid van 120 km/h wordt er iedere minuut $\frac{120}{60} = 2$ km afgelegd. De afgelegde afstand is bij een vaste snelheid recht evenredig met de tijd: als de reistijd twee keer zo lang wordt, dan wordt ook de afgelegde afstand twee keer zo groot.

Hoe hoger de snelheid, hoe korter de reistijd over die 16 km. De reistijd is omgekeerd evenredig met de snelheid: wordt er twee keer zo snel gereden, dan is de helft van de reistijd nodig.

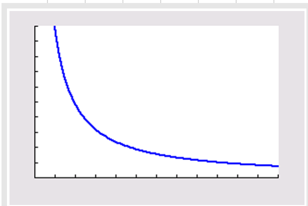
- Bij een snelheid van 120 km/h is de reistijd $\frac{16}{120} \cdot 60 = 8$ minuten.
- Door de drukte kan er (gemiddeld) maar 60 km/h gereden worden. Dan is de reistijd $\frac{16}{60} \cdot 60 = 16$ minuten.



Figuur 3.1

De reistijd t in minuten kan berekend worden door de afstand van 16 km te delen door de snelheid v (km/h) en met 60 te vermenigvuldigen: $t = \frac{16}{v} \cdot 60 = \frac{960}{v}$.

Bekijk de grafiek van zo'n omgekeerd evenredig verband. Voor snelheden dicht bij 0 wordt de reistijd heel erg groot. Voor hele grote snelheden wordt de reistijd vrijwel 0.



Figuur 3.2

Opgave 1

Gebruik de gegevens uit **Uitleg 1**.

a Welk soort verband bestaat er tussen de afgelegde weg en de tijd bij een vaste snelheid?

b Vul in: als bij een vaste snelheid de reistijd drie keer zo lang wordt, dan wordt de afgelegde afstand ... keer zo groot.

c Welk soort verband bestaat er bij een vaste afstand tussen de reistijd en de snelheid?

A. recht evenredig verband
B. omgekeerd evenredig verband

d Vul in: wordt er vier keer zo snel gereden, dan wordt de reistijd ... keer zo groot.

e De reistijd t in minuten kan berekend worden door de afstand van 16 km te delen door de snelheid v (km/h) en met 60 te vermenigvuldigen.
Waarom moet er met 60 vermenigvuldigd worden?

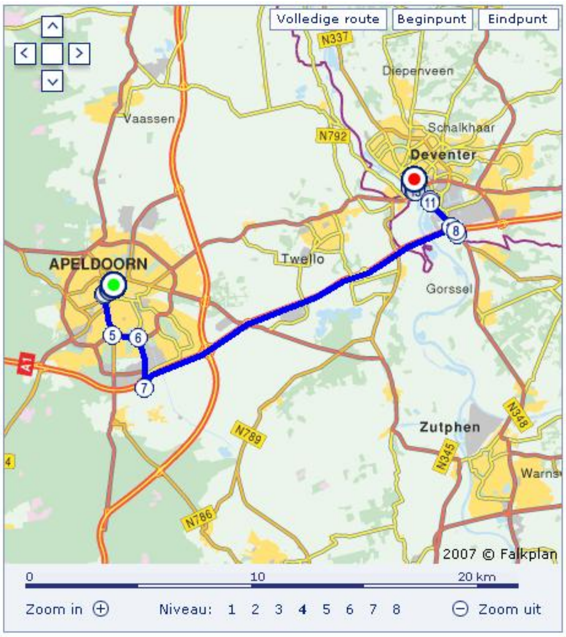
f Leg uit waarom bij snelheden dicht bij 0 de reistijd heel erg groot wordt.

Uitleg 2

Van Apeldoorn naar Deventer is met de auto 16 km over de snelweg. Hoe hoger de snelheid, hoe korter de reistijd over die 16 km. Onderweg wordt 5 minuten gestopt voor het tanken van brandstof.

- Bij een snelheid van 120 km/h is de reistijd $\frac{16}{120} \cdot 60 + 5 = 13$ minuten.
- Door de drukte kan er (gemiddeld) maar 60 km/h gereden worden. Dan is de reistijd $\frac{16}{60} \cdot 60 + 5 = 21$ minuten.

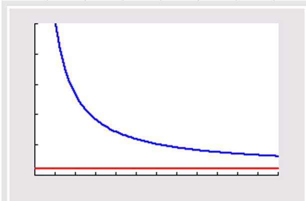
Nu betekent een verdubbeling van de snelheid niet een halvering van de reistijd. Snelheid en reistijd zijn niet omgekeerd evenredig.



Figuur 3.3

De reistijd t in minuten kan berekend worden door de afstand van 16 km te delen door de snelheid v (km/h), dan met 60 te vermenigvuldigen en tenslotte nog 5 bij de uitkomst op te tellen:

$t = \frac{16}{v} \cdot 60 + 5 = \frac{960}{v} + 5$



Figuur 3.4

Bekijk de grafiek van zo'n hyperbolisch verband. Voor snelheden dicht bij 0 wordt de reistijd heel erg groot. Voor hele grote snelheden komt de reistijd in de buurt van de 5 minuten.

Opgave 2

Bekijk **Uitleg 2** over hyperbolische verbanden.

a Waarom is er in de situatie in de uitleg geen omgekeerd evenredig verband tussen snelheid en reistijd?

b De reistijd t in minuten kan berekend worden door de afstand van 16 km te delen door de snelheid v (in km/h), dan met 60 te vermenigvuldigen en tenslotte nog 5 bij de uitkomst op te tellen. Waarom moet er op het laatst nog 5 bij opgeteld worden?

c Leg uit waarom bij hele grote snelheden de reistijd in de buurt van de 5 minuten komt.

Opgave 3

Iemand uit Apeldoorn bezoekt regelmatig familie in Rotterdam. De afstand tussen Apeldoorn en Rotterdam is ongeveer 150 kilometer. Onderweg naar de familie in Rotterdam wordt altijd een half uur gepauzeerd.

a Hoe lang duurt de rit als er gemiddeld 100 kilometer per uur wordt gereden?

b Bereken de gemiddelde snelheid als de rit 2,5 uur duurt.

c Geef een formule met t uitgedrukt in v.

d Geef ook een formule die v uitdrukt in t.

e Voor welke waarden van t en v is deze formule bruikbaar?

Theorie en voorbeelden

Om te onthouden

Twee variabelen x en y zijn **recht evenredig** wanneer geldt: als een x-waarde k keer zo groot wordt, wordt de bijbehorende y-waarde ook k keer zo groot. Bij een recht evenredig verband hoort een formule van de vorm $y = cx$, waarin c een constante is. De grafiek van een recht evenredig verband is een rechte lijn die door de oorsprong gaat en die c als richtingscoëfficiënt heeft.

Bekijk de applet.

Twee variabelen x en y zijn **omgekeerd evenredig** wanneer geldt: als een x-waarde k keer zo groot wordt, wordt de bijbehorende y-waarde k keer zo klein. Bijvoorbeeld: wordt x twee keer zo groot, dan wordt y een half keer zo groot (ofwel twee keer zo klein). Bij een omgekeerd evenredig verband hoort een formule van de vorm $x \cdot y = c$. Deze formule is ook te schrijven als $y = \frac{c}{x}$.

Bij een formule van de vorm $y = \frac{c}{x} + a$ spreek je van een **hyperbolisch verband**. Hoewel dan x en y niet omgekeerd evenredig zijn, lijkt de grafiek wel op die van een omgekeerd evenredig verband. Hij is alleen a verschoven in de y-richting.

Naarmate in een hyperbolisch verband de x-waarde hogere waarden aanneemt, komt de grafiek heel dicht bij de y-waarde a. De grafiek komt dus steeds dichter bij de lijn $y = a$.
En zo neemt de grafiek hele grote waarden aan als die dichter bij de lijn $x = 0$ komt. Die grote waarden kunnen positief of negatief zijn. Dat hangt af aan welke kant van de y-as de grafiek loopt.

Voorbeeld 1

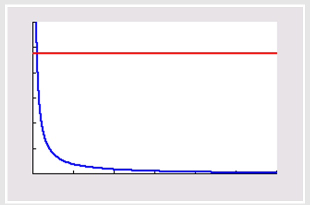
Bekijk de applet.

Als een rechthoekig tafelblad een oppervlakte van 1 m^2 heeft, kunnen lengte l en breedte b variëren.

- Stel een passende formule op met l en b in centimeter.
- Laat zien dat l en b omgekeerd evenredig zijn.
- Welke breedte heeft een rechthoekige tafel met een oppervlakte van 1 m^2 en een lengte van minimaal 240 cm? Stel hierbij een ongelijkheid op en los deze algebraïsch op.

Antwoord

- $1\text{ m}^2 = 10000\text{ cm}^2$.
  Voor deze rechthoek geldt: *oppervlakte* $= 10000 = l \cdot b$.
  Dus: $l = \frac{10000}{b}$.
- Als b verdubbelt, halveert l, bijvoorbeeld:
  $b = 50$ cm geeft $l = \frac{10000}{50} = 200$ cm;
  $b = 100$ cm geeft $l = \frac{10000}{100} = 100$ cm.
  Dit geldt ook voor andere waarden, dus l en b zijn inderdaad omgekeerd evenredig.
- Uit $\frac{10000}{x} = 240$ volgt $x = 41\frac{2}{3}$.
  Voer in: $y_1 = \frac{10000}{x}$ en $y_2 = 240$.
  Venster bijvoorbeeld: $0 \leq x \leq 300$ en $0 \leq y \leq 300$.
  Lees uit de grafieken de juiste oplossing af: $b \leq 41\frac{2}{3}$.
  De tafel wordt $41\frac{2}{3}$ cm breed of minder.



Figuur 3.5

Opgave 4

Bas wil een rechthoekige tafel met een oppervlakte van 2 m^2 en een lengte van minimaal 440 cm. Welke breedte hoort hierbij? Stel een ongelijkheid op en los deze algebraïsch op.

Opgave 5

Lengte l en breedte b van een rechthoek met een oppervlakte van 1 m^2 zijn omgekeerd evenredig.

a Welke drie formules passen bij het verband tussen l en b als beide in centimeter zijn?

b Plot de grafiek met l uitgedrukt in b.

c Laat met behulp van de formule uit b zien dat l wordt gehalveerd als b wordt verdubbeld. Vervang daarvoor b in $2b$.

d Als b tien keer zo groot wordt, hoeveel keer zo groot wordt l dan?

e Als b met $\frac{1}{10}$ wordt vermenigvuldigd, hoeveel keer zo groot wordt l dan?

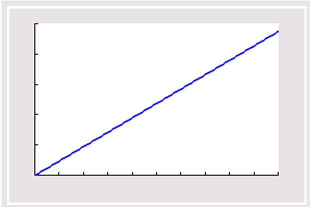
Voorbeeld 2

Voor een vakantiebaantje krijg je een vast bedrag per uur. Als je 8 uur op een dag werkt, verdien je € 38,00.

- Waarom is het salaris dat je dan verdient recht evenredig met het aantal uur dat je hebt gewerkt?
- Stel een formule op waarmee het salaris s (euro) berekend kan worden als het aantal gewerkte uren u bekend is en plot de grafiek.
- Je wilt minimaal € 200,00 verdienen aan dit baantje. Hoeveel uur moet je daarvoor werken?

Antwoord

- Omdat je een vast bedrag per uur krijgt en verder niets.
- De formule is: $s = 4{,}75u$.
  Voer in: $y_1 = 4{,}75x$.
  Venster bijvoorbeeld: $0 \leq x \leq 50$ en $0 \leq y \leq 250$.
  De grafiek is een rechte lijn met richtingscoëfficiënt $4{,}75$ en hij gaat door de oorsprong.
  Als u twee keer zo groot wordt, dan wordt s twee keer zo groot.
- Je moet nu $4{,}75s \geq 200$ oplossen.
  Uit $4{,}75s = 200$ volgt $s \approx 42{,}1$.
  Je moet dus minimaal 43 uur werken.



Figuur 3.6

Opgave 6

Gegeven is de formule: $y = 21x$

a Plot de grafiek die hoort bij deze formule.

b Welke eigenschappen heeft de grafiek?

c Vul in: als x vijf keer zo groot, dan wordt y ... keer zo groot.

d Wat voor soort verband beschrijft de formule?

e Los de ongelijkheid $21x \leq 15$ exact op.

Opgave 7

Stel een formule op voor iedere situatie.

a Het jaarinkomen I van een sportclub als ieder lid l een jaarlijkse contributie van € 65,00 betaalt.

b De kosten K voor de flessen cola c die ik koop als iedere fles € 1,10 kost.

c Voor een toets kun je 40 punten halen en je cijfer C kun je berekenen door een kwart te nemen van het behaalde aantal punten p.

d Je totale opbrengst O bij een sponsorloop als ieder rondje r € 2,50 oplevert.

Voorbeeld 3

Voor het laten drukken van folders moet een vast bedrag van € 10,00 plus € 0,04 per folder betaald worden. De kosten per folder zijn daarom hoog als er maar weinig afgedrukt moeten worden. Stel een formule op voor de kosten per folder en bereken met behulp daarvan bij welk aantal folders de drukkosten niet hoger zijn dat € 0,05.

Antwoord

Voor de totale kosten TK (in euro) geldt: $TK = 10 + 0{,}04a$. Hierin is a het aantal folders.

Voor de totale kosten per folder TKF (in euro) geldt: $TKF = \frac{10+0{,}04a}{a} = \frac{10}{a} + 0{,}04$.

$\frac{10}{a} + 0{,}04 = 0{,}05$ oplossen geeft $a = 1000$.

Dus vanaf 1000 folders zijn de drukkosten per folder lager dan € 0,05.

Opgave 8

Voor het laten drukken van luxe folders moet een vast bedrag van € 25,00 en daar bovenop € 0,12 per folder betaald worden. De totale kosten per folder zijn hoog als er maar weinig afgedrukt moeten worden. De totale kosten per folder TKF (in euro) hangen af van het aantal folders a dat gedrukt moet worden.

a Stel een formule op voor de totale kosten per folder TKF.

b Waarom is TKF niet omgekeerd evenredig met a?

c Teken de grafiek met TKF uitgedrukt in a op de grafische rekenmachine.

d Bereken met behulp van de formule bij welk aantal folders de drukkosten niet hoger zijn dan € 0,15 per folder.

Opgave 9

Los de vergelijkingen algebraïsch op:

a $\frac{400}{v} + 120 = 200$

b $\frac{20}{a-10} - 6 = 10$

Verwerken

Opgave 10

Iemand uit Apeldoorn bezoekt regelmatig familie in Rotterdam. De afstand tussen Apeldoorn en Rotterdam is ongeveer 150 kilometer. Hij rijdt gemiddeld 100 kilometer per uur.

a Hoeveel kilometer wordt er iedere minuut afgelegd?

b Met welke formule kan voor deze persoon de afgelegde afstand berekend worden als de tijd bekend is? Neem de tijd t in minuten en de afgelegde afstand a in kilometer.

c Plot de grafiek bij deze formule. Noem twee opvallende eigenschappen van de grafiek.

Opgave 11

In een groot winkelbedrijf wordt onderzocht hoe de tomatenverkoop afhangt van de prijs. Iemand beweert dat de volgende formule geldt: $a = \frac{500}{p}$. Hierin is a de verkoop per dag in kg en p de prijs per kg in euro.

a Plot een grafiek waaruit je de verkoop kunt aflezen voor prijzen tussen de € 1,00 en € 5,00 per kilogram.

b Iemand zegt: “Een verdubbeling van de prijs zorgt voor een halvering van de verkoop.” Klopt dat?

A. De bewering is waar.
B. De bewering is niet waar.

c Klopt deze bewering met de formule: ‘Als de prijs vijf keer zo hoog wordt, wordt de verkoop vijf keer zo klein.’?

A. De bewering is waar.
B. De bewering is niet waar.

d Geef twee andere formules voor hetzelfde verband tussen a en p.

e In het bedrijf heeft men een voorraad van 300 kg tomaten. Deze tomaten zijn niet lang meer houdbaar en men wil er binnen een dag vanaf. Bereken de maximale prijs volgens de formule.

Een formule zoals $a = \frac{500}{p}$ is meestal slechts op een beperkt gebied bruikbaar. Dat kun je zien als je voor p extreme gevallen neemt.

f Hoe groot is de verkoop bij een prijs van € 0,01? En bij een prijs van € 100,00? Zal dit in werkelijkheid ook zo zijn?

Opgave 12

Een kaasboer houdt bij hoeveel kilo geraspte kaas hij per week verkoopt. Het blijkt dat de hoeveelheid k (kg) die hij verkoopt omgekeerd evenredig is met de prijs p per kilo. Bij een prijs van € 13,00 per kilo verkoopt hij 15 kg geraspte kaas.

a Bereken $p \cdot k$ met behulp van de gegevens en geef vervolgens een formule die k uitdrukt in p.

b Bereken het aantal verkochte kilo's als de prijs € 10,00 per kilogram is.

Er is een nieuw restaurant in het pand naast hem geopend. De eigenaar van het restaurant neemt iedere week 15 kg kaas af. (De prijs die hij betaalt, wordt in onderling overleg met de kaasboer bepaald.)

c Bereken in de nieuwe situatie het aantal verkochte kilo's als de prijs € 10,00 per kilo is.

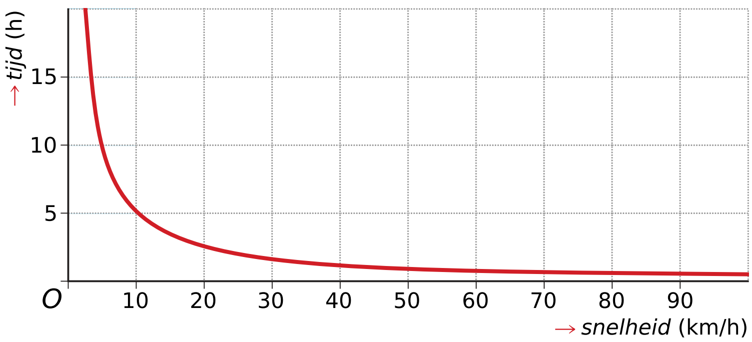
d Leg uit dat de nieuwe formule voor k de vorm $k = \frac{195}{p} + c$ heeft en bepaal de waarde van de constante c.

e Plot de grafiek van k en bepaal bij welke prijs per kilogram de verkoop per week 40 kg is.

f Bereken ook algebraïsch bij welke prijs de verkoop per week 40 kg is.

Opgave 13

Bekijk de grafiek die het verband tussen de snelheid en de tijd weergeeft voor iemand die van Utrecht naar Den Bosch reist. De (gemiddelde) snelheid v (km/h) is omgekeerd evenredig met de tijd t (h).



Figuur 3.7

a Een wandelaar heeft een gemiddelde snelheid van vijf kilometer per uur.
Hoelang doet hij over de afstand Utrecht - Den Bosch?

b Een fietser doet $2{,}5$ uur over deze afstand. Bereken zijn gemiddelde snelheid.

c Bereken uit je antwoorden bij a en b het product van de tijd en de snelheid. Stel een formule op die bij de grafiek past.

d De trein van Utrecht naar Den Bosch doet er ongeveer 25 minuten over.
Hoeveel kilometer per uur rijdt de trein?

Opgave 14

Los de vergelijkingen algebraïsch op.

a $\frac{2{,}25}{p} = 0{,}45$

b $4{,}50 + \frac{300}{k} = 4{,}70$

c $\frac{1200}{k+12} - 42 = 6$

Opgave 15

De overheid besteedt veel geld aan campagnes die waarschuwen voor de gevolgen van roken en drinken. Als deze campagnes effect hebben, dan zou binnen redelijke grenzen moeten gelden: hoe meer geld de overheid eraan besteedt, hoe minder mensen er roken. Stel dat de volgende formule geldt voor het percentage rokers van de Nederlandse bevolking:

$p = \frac{50}{b} + 15$

Hierin is p het percentage rokers en b het bedrag dat de overheid aan antirookcampagnes besteedt in miljoenen euro.

a Geldt voor deze formule inderdaad dat het percentage rokers afneemt naarmate de overheid meer geld aan campagnes besteedt?

A. ja
B. nee

b Bereken het percentage rokers als de overheid 200 miljoen euro aan campagnes besteedt.

c Er is een percentage van de bevolking dat ondanks alle campagnes hardnekkig blijft roken. Hoe groot is dat percentage?

d Het percentage rokers is nog nooit meer dan 90 geweest en zal dat waarschijnlijk ook nooit worden. Bereken welke waarde van b bij $p = 90$ hoort. Denk je dat de formule voor deze waarde van b nog geldig is?

Toepassen

Opgave 16: Geld lenen kost geld

Geld lenen kost geld. Soms kost het heel veel geld. Vooral als je direct een paar honderd euro nodig hebt. In dit soort situaties kun je een flitslening nemen. Je leent een niet al te groot geldbedrag en betaalt dit na een korte periode terug. Er bestaan verschillende websites waar je geld kunt lenen. Op de website flitsmoney.nl staat dat er geen rente wordt berekend. Je hoeft alleen behandelingskosten te betalen. In de tabel staat hoe hoog deze kosten zijn.

te lenen bedrag (euro)	behandelingskosten (euro)
100,00	25,00
250,00	62,50
300,00	75,00
375,00	93,75

Tabel 3.1

Als je bijvoorbeeld € 100,00 wilt lenen, krijg je dit geld binnen 10 minuten op je bankrekening. Dit bedrag moet samen met de € 25,00 behandelingskosten na 30 dagen worden terugbetaald.
Er is bij Flitsmoney een recht evenredig verband tussen het totaal terug te betalen bedrag en het te lenen bedrag. Laat dit met berekeningen zien. Controleer hiervoor alle waarden in de tabel.

(naar: examen havo wiskunde A in 2015, tweede tijdvak)

Testen

Opgave 17

Dat wijnglazen groter zijn dan portglazen is niet toevallig. Voor de inhoud I (in ml) van een glas en het alcoholpercentage p van de drank die erin hoort geldt namelijk ongeveer deze formule: $p \cdot I = 1200$.

a Leid uit het bovenstaande af of wijn een hoger of lager alcoholpercentage heeft dan port.

b Gewoon bier heeft een alcoholpercentage van ongeveer 5%. Bereken de inhoud van een bierglas.

c Een jeneverglas heeft een inhoud van 35 ml. Hoe groot is het alcoholpercentage van jenever?

d Geef een formule voor I als functie van p.

e Alcoholarm bier heeft een alcoholpercentage van 2%. Is de formule nog bruikbaar voor alcoholarm bier?

Opgave 18

Los de volgende vergelijkingen algebraïsch op.

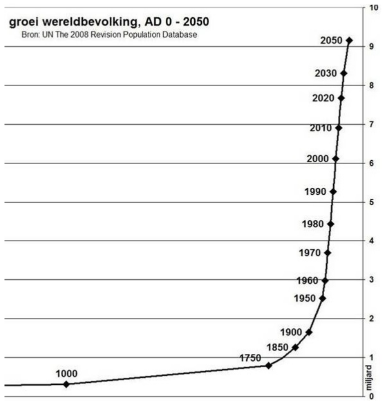
a $\frac{300}{T-2} = 50$

b $\frac{300}{T} - 2 = 48$

1.4 Groei en verval

Inleiding

Vooral in ontwikkelingsgebieden is vaak sprake van een sterke groei van de bevolking. Het maakt daarbij nogal veel verschil of er jaarlijks een bepaalde (ongeveer) vaste hoeveelheid bij komt, of dat er jaarlijks een (ongeveer) vast percentage van de hoeveelheid van het jaar ervoor bijkomt. In dit tweede geval groeit de totale hoeveelheid steeds sterker, de groei is exponentieel. Maar niet altijd is dit verschil in het begin goed te zien...



Figuur 4.1

Je leert in dit onderwerp

- lineaire en exponentiële groei en afname met elkaar vergelijken.

Voorkennis

- werken met lineaire en exponentiële verbanden, de bijbehorende formules opstellen, grafieken maken, etc.;
- vergelijkingen en ongelijkheden oplossen, zowel met de GR als (waar dat kan) met de hand.

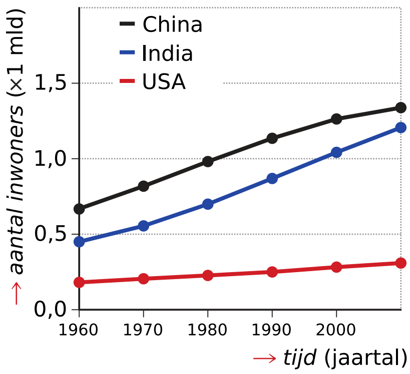
Verkennen

Opgave V1

Je ziet hier een tabel met de aantallen inwoners van China, India en de USA, afgerond op miljoenen inwoners.

jaar	1960	1970	1980	1990	2000	2010
China	667	818	981	1135	1263	1338
India	450	555	699	869	1042	1206
USA	181	205	227	250	282	309

Tabel 4.1



Figuur 4.2

De figuur laat dit ook nog eens zien.

a In welke van deze drie landen is de bevolking het minst gestegen?

b Ga voor de USA uit van lineaire groei en stel een formule op voor het aantal mensen in miljoenen N_{USA} afhankelijk van de tijd t in jaren na 1960.

c In India lijkt de groei (zeker vlak na 1960) steeds sterker te worden. Ga voor dit land uit van exponentiële groei en stel een formule op voor het aantal mensen in miljoenen N_{India} afhankelijk van de tijd t in jaren na 1960.

d China heeft jarenlang een één kind politiek gevoerd. Wat wordt hiermee bedoeld en waar zie je dit aan?

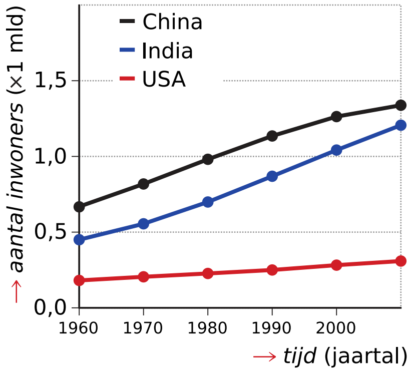
e Ga voor China uit van lineaire groei en voor India van exponentiële groei. Bepaal daarmee het jaar waarin India voor het eerst meer inwoners zal hebben dan China.

Uitleg

Bekijk de tabel en grafiek.

jaar	1960	1970	1980	1990	2000	2010
China	667	818	981	1135	1263	1338
India	450	555	699	869	1042	1206
USA	181	205	227	250	282	309

Tabel 4.2



Figuur 4.3

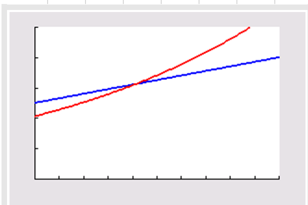
Onderzoekers voorspellen dat de bevolkingsomvang van China vanaf het jaar 2000 lineair zal toenemen. Voor India is de verwachting dat de bevolkingsomvang vanaf het jaar 2000 exponentieel toeneemt. In welk jaar zal India naar verwachting het land China inhalen qua bevolkingsomvang?

Om antwoord te geven op deze vraag, kun je voor beide landen een formule opstellen voor de bevolkingsgroei. Neem hierin N de bevolkingsomvang, t de tijd in jaar na het jaar 2000, b de beginwaarde bij $t = 0$. Vervolgens kan met behulp van de grafische rekenmachine het juiste jaar gevonden worden.

- De formule voor de bevolkingstoename van China is van de vorm: $N = at + b$.
  Hierin is a de richtingscoëfficiënt, ofwel de toename per jaar.
  Er geldt: $N_C = 7{,}5t + 1263$.
- De formule voor de bevolkingstoename van India is van de vorm: $N = b \cdot g^t$.
  Hierin is g de groeifactor per jaar.
  Er geldt: $N_I = 1042 \cdot 1{,}01^t$.
  Hierbij hoort een groeipercentage van 1% per jaar.

Voor het vinden van het jaartal waarin de bevolking van India die van China inhaalt moet de ongelijkheid $N_I > N_C$ worden opgelost.

Hier wordt gebruikgemaakt van extrapoleren: op basis van de meest recente cijfers over de groei van de bevolkingsomvang zijn formules opgesteld en die zijn gebruikt om de bevolkingsomvang in de jaren daarna te schatten. De formules kunnen ook gebruikt worden voor interpoleren: de bevolkingsomvang kan daarmee in een jaar tussen 2000 en 2010 geschat worden.



Figuur 4.4

Opgave 1

Gebruik de gegevens uit de **Uitleg**.

a Welke algemene formule geldt voor lineaire groei?

b Stel de lineaire formule op die hoort bij de bevolkingsgroei van China vanaf 2000.

c Welke algemene formule geldt voor exponentiële groei?

d Laat zien hoe je de groeifactor per jaar berekent en stel daarmee de exponentiële formule op die hoort bij de bevolkingsgroei van India vanaf 2000.

e Laat zien hoe je het groeipercentage van de bevolkingsgroei in India berekent.

f Bereken in welk jaar de bevolking van India die van China naar verwachting inhaalt.

Opgave 2

Bekijk in de **Uitleg** de tabel met gegevens over de bevolkingsomvang van de Verenigde Staten. Onderzoekers voorspellen dat de bevolkingsomvang van de VS vanaf het jaar 2000 lineair zal toenemen. Van een groot ontwikkelingsgebied is bekend dat in het jaar 2000 de bevolkingsomvang zo'n 170 miljoen was. Dit aantal groeide exponentieel tot 210 miljoen in 2010. De voorspelling is dat de groei zich op deze manier zal voortzetten.

a Stel een formule op voor de bevolkingsomvang N van de VS met $t = 0$ in 2000.

b Stel een formule op voor de bevolkingsomvang N van het ontwikkelingsgebied met $t = 0$ in 2000.

c De vraag is in welk jaar het ontwikkelingsgebied naar verwachting de VS in zal halen qua bevolkingsomvang.
Stel bij deze vraag een ongelijkheid op en los deze op met de grafische rekenmachine.

d Bepaal de verdubbelingstijd van de bevolkingsomvang voor de VS en voor het ontwikkelingsgebied en vergelijk de uitkomsten met elkaar.

Theorie en voorbeelden

Om te onthouden

Als er een **lineair verband** tussen y en x is, heeft de bijbehorende formule de vorm $y = ax + b$, hierin is:

- a de **richtingscoëfficiënt**, dus de toe- of afname van y als x met stappen van 1 toeneemt;
- b de **beginwaarde**, de uitkomst bij $x = 0$.

Als er een **exponentieel verband** tussen y en x is, heeft de bijbehorende formule de vorm $y = b \cdot g^x$, hierin is:

- b de **beginwaarde**, de uitkomst bij $x = 0$.
- g de **groeifactor** per tijdseenheid, dus het getal waarmee y wordt vermenigvuldigd als x met stappen van 1 toeneemt.
  Als bijvoorbeeld de groeifactor per 5 jaar g is, is de groeifactor per jaar $g^{\frac{1}{5}}$.

Het **groeipercentage** geeft aan met hoeveel procent een hoeveelheid per tijdseenheid toe- of afneemt.

Bij exponentiële verbanden spelen de **verdubbelingstijd** en de **halveringstijd** soms een belangrijke rol. Dat is achtereenvolgens de tijd die het kost om de beginwaarde te verdubbelen (als $g > 1$) of te halveren (als $0 < g < 1$). Die kun je met de grafische rekenmachine bepalen.

Voorbeeld 1

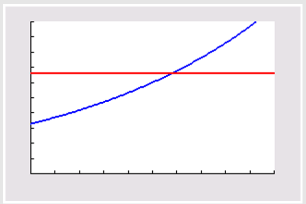
Bekijk de groeitabel van plaats A. Zowel exponentiële groei als lineaire groei zijn denkbaar. Bepaal door middel van zowel lineair extrapoleren als exponentieel extrapoleren de bevolkingsomvang in 2030 en vergelijk de resultaten met elkaar. Bereken ook in beide situaties de verdubbelingstijd.

jaar	1985	1995	2005	2015
aantal inwoners plaats A (× 1000)	662	701	743	784

Tabel 4.3

Antwoord

- Bij een lineair verband geldt $N = 4{,}07t + 662$.
  In 2030 is $t = 45$. $N = 4{,}07 \cdot 45 + 662 = 845{,}15$, dus 845150 inwoners.
  Bij een exponentieel verband geldt $N = 662 \cdot 1{,}006^t$.
  Hier is in 2030 $N = 662 \cdot 1{,}006^{45} \approx 866{,}497$, dus 866497 inwoners.
  De resultaten liggen redelijk dicht bij elkaar.
- De verdubbelingstijd bij het lineaire verband is te vinden door een vergelijking algebraïsch op te lossen: dit geeft afgerond 163 jaar.
  De verdubbelingstijd bij het exponentiële verband is te vinden door de vergelijking $662 \cdot 1{,}006^t = 1324$ op te lossen met de grafische rekenmachine. Dit geeft afgerond 116 jaar.



Figuur 4.5

Opgave 3

Bekijk **Voorbeeld 1** over het aantal inwoners in plaats A.

a Zowel exponentiële groei als lineaire groei zijn denkbaar. Laat dit met berekeningen zien.

b Stel een lineaire formule op voor de bevolkingsomvang van plaats A, met de tijd t in jaar.

c Bepaal door middel van lineair extrapoleren het aantal inwoners van plaats A in 2043.

d Stel een exponentiële formule op voor de bevolkingsomvang van plaats A, met de tijd t in jaar.

e Bepaal door middel van exponentieel extrapoleren het aantal inwoners van plaats A in 2043.

f Bereken in beide situaties de verdubbelingstijd in jaar.

Opgave 4

Bekijk de groeitabel van het aantal herten in een bosgebied. Eerst groeide het aantal herten exponentieel, maar sinds 2012 wordt er gecontroleerd op de herten gejaagd omdat te veel herten een gevaar opleveren voor het verkeer.

jaar	2006	2008	2010	2012	2014	2016
aantal herten	32	51	82	131	140	150

Tabel 4.4

a Vanaf 2012 zijn zowel exponentiële groei als lineaire groei denkbaar. Laat dit met berekeningen zien.

b Bepaal door middel van zowel lineair extrapoleren als exponentieel extrapoleren het aantal herten in 2025 en vergelijk de resultaten met elkaar.

c Bereken ook in beide situaties de verdubbelingstijd.

Voorbeeld 2

Sommige stoffen in de natuur vallen spontaan uit elkaar en zenden daarbij straling uit. Dit proces heet radioactief verval. Bekijk de vervaltabel met daarin de hoeveelheid van stof M in microgram (µg).

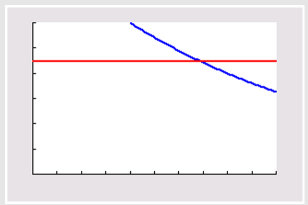
tijd (dag)	0	2	4	6	8
hoeveelheid stof M (µg)	450	432	415	398	381

Tabel 4.5

Zowel lineair als exponentieel verval zijn denkbaar. Bepaal door middel van zowel lineair extrapoleren als exponentieel extrapoleren de hoeveelheid van stof M na 13 dagen en vergelijk de resultaten met elkaar. Bereken in beide situaties de halveringstijd.

Antwoord

- Bij een lineair verband geldt $M = \text{-}8{,}5t + 450$. Na 13 dagen is $t = 13$. $M = \text{-}8{,}5 \cdot 13 + 450 = 339{,}5$ µg.
  Bij een exponentieel verband geldt $M = 450 \cdot 0{,}98^t$. Hier is na 13 dagen $M = 450 \cdot 0{,}98^{13} \approx 346{,}1$ µg.
  De resultaten liggen redelijk dicht bij elkaar.
- De halveringstijd bij het lineaire verband is te vinden door een vergelijking algebraïsch op te lossen: dit geeft afgerond 26 dagen.
  De halveringstijd bij het exponentiële verband is te vinden door de vergelijking $450 \cdot 0{,}98^t = 225$ op te lossen met de grafische rekenmachine. Dit geeft afgerond 34 dagen.



Figuur 4.6

Opgave 5

Bekijk **Voorbeeld 2** over het verval van de hoeveelheid van stof M.

a Zowel lineair als exponentieel verval zijn denkbaar. Laat dit met berekeningen zien.

b Stel een lineaire formule op voor de hoeveelheid stof M, met de tijd t in dagen.

c Bepaal door middel van lineair extrapoleren de hoeveelheid van stof M na 360 uur.

d Stel een exponentiële formule op voor de hoeveelheid stof M, met de tijd t in dagen.

e Bepaal door middel van exponentieel extrapoleren de hoeveelheid van stof M na 360 uur.

f Bereken in beide situaties de halveringstijd in dagen.

Opgave 6

Kernenergie levert weinig afval op, maar het is wel afval dat speciale aandacht vereist. Het is namelijk radioactief en het blijft nog tientallen jaren warmte afgeven. In 2003 is in Zeeland een gebouw geopend waar de komende honderd jaar kernafval zal worden opgeslagen. Het gebouw heet HABOG, Hoogradioactief Afval Behandelings- en Opslag Gebouw. In het HABOG wordt het afval van de kerncentrale van Borssele opgeslagen. Over honderd jaar zijn de radioactiviteit en de warmte van het afval zo veel afgenomen dat het afval op een andere plaats kan worden opgeslagen. Het afval uit Borssele bestaat jaarlijks uit zes glasblokken met hoogradioactief afval. In het begin geeft zo’n blok evenveel warmte af als een kachel van 1800 Watt. Na 100 jaar is de warmteafgifte verminderd tot die van drie gloeilampen, ofwel 180 Watt. De warmteafgifte neemt exponentieel af.

a Bereken het percentage waarmee de warmteafgifte per jaar afneemt. Rond af op twee decimalen.

b Het gebouw is knaloranje geverfd. In grote groene letters zijn er beroemde formules van Einstein en Planck op aangebracht. Elke tien jaar wordt het gebouw opnieuw geverfd, telkens in een iets lichtere tint om de afname van de warmteafgifte aan te geven.
Je mag er in de rest van de opgave van uitgaan dat de warmteafgifte met $2,3$% per jaar afneemt.
Bereken het percentage waarmee de warmteafgifte in een periode van tien jaar afneemt. Rond af op één decimaal.

c Bereken na hoeveel jaar de warmteafgifte nog maar de helft is van de oorspronkelijke hoeveelheid. Rond af op één decimaal.

(naar: examen wiskunde A1,2 in 2005, tweede tijdvak)

Voorbeeld 3

De verkoop van SmartWatches daalt na het uitbrengen ervan. Ditzelfde geldt voor de verkoop van een gewoon horloge.

Bekijk de twee vervaltabellen met daarin het aantal verkochte SmartWatches en gewone horloges per dag na het uitbrengen.

tijd (t)	7	14	21
verkochte aantal SmartWatches	200000	20000	2000

tijd (t)	10	100	1000
verkochte aantal gewone horloges	5000	500	50

Tabel 4.6

De eerste dagen na het uitbrengen worden er altijd heel veel SmartWatches verkocht, hetzelfde geldt voor gewone horloges. Vanaf de tweede week kunnen zowel exponentiële groei als omgekeerd evenredige groei passen bij de verkoopdalingen. Welke soort groei hoort bij welk apparaat?

Antwoord

Bij beide tabellen wordt het verkochte aantal steeds gedeeld door 10 ofwel vermenigvuldigd met $\frac{1}{10}$. Het verschil zit hem in de tijdstippen waarop dat gebeurt.

In de tabel van de gewone horloges worden de tijdstippen steeds met 10 vermenigvuldigd. Hier geldt: als de tijd wordt vermenigvuldigd met 10, dan wordt het verkochte aantal gedeeld door 10. Dat is een omgekeerd evenredig verband.

In de tabel van de SmartWatches worden voor de tijd steeds gelijke stappen genomen, iedere 7 dagen wordt het verkochte aantal vermenigvuldigd met $\frac{1}{10}$. Dat is een exponentieel verband.

Opgave 7

Gebruik de gegevens uit **Voorbeeld 3**.

a Stel voor beide apparaten een formule op die het verkochte aantal A na t dagen beschrijft.

b Bereken met behulp van extrapoleren de hoeveelheid verkochte apparaten van beide soorten na 30 dagen en vergelijk de resultaten met elkaar.

Opgave 8

Bekijk de twee vervaltabellen met daarin de hoeveelheid van stof M en stof P in microgram (µg).

tijd (h)	1	4	16	64
hoeveelheid stof M (µg)	1000	250	62,5	15,6

tijd (h)	0	24	48	72
hoeveelheid stof P (µg)	1000	250	62,5	15,6

Tabel 4.7

Zowel exponentiële groei als omgekeerd evenredige groei kunnen passen bij het verval van een stof.

a Welke soort groei hoort bij welke stof?

b Stel van beide stoffen een formule op die de hoeveelheid stof na t uur beschrijft.

c Bereken met behulp van extrapoleren de hoeveelheid van beide stoffen na 100 uur en vergelijk de resultaten met elkaar.

d Bereken voor beide stoffen de halveringstijd.

Verwerken

Opgave 9

Geef voor iedere tabel aan van welke soort groei er (bij benadering) sprake is.

a

tijd (h)	2	4	6	8
aantal	150	180	216	259

Tabel 4.8

A. lineaire groei
B. exponentiële groei
C. omgekeerd evenredige groei

b

tijd (h)	2	4	6	8
aantal	150	180	210	240

Tabel 4.9

A. lineaire groei
B. exponentiële groei
C. omgekeerd evenredige groei

c

tijd (h)	2	4	6	8
aantal	150	75	50	37,5

Tabel 4.10

A. lineaire groei
B. exponentiële groei
C. omgekeerd evenredige groei

Opgave 10

Gegeven zijn twee punten van een grafiek in een Oxy-assenstelsel. Stel een formule op voor het bijbehorende verband tussen x en y. Doe dit zowel voor een lineair als voor een exponentieel verband.

a $(0,12)$ en $(20,42)$

b $(0,55)$ en $(40,10)$

Opgave 11

Lisette en Elma huren allebei een appartement sinds het jaar 2012. Het eerste jaar betaalden ze allebei evenveel huur per maand. In de tabel is te zien dat de huur ieder jaar verhoogd wordt en dat dit bij Lisette en Elma niet om hetzelfde bedrag gaat.

tijd (jaar)	2012	2013	2014	2015	2016
maandelijkse huur Lisette (euro)	600	660	726	798,60	878,46
maandelijkse huur Elma (euro)	600	660	720	780	840

Tabel 4.11

a Van welke soort groei is er bij Lisette sprake?

A. exponentiële groei
B. lineaire groei

b Van welke soort groei is er bij Elma sprake?

A. exponentiële groei
B. lineaire groei

c Stel voor beide dames een formule op waarmee de huurprijs H in het jaar t kan worden berekend met $t = 0$ in 2012.

d Bepaal door middel van extrapoleren de huurprijs van beide dames in 2020.

e Bepaal voor beide dames de verdubbelingstijd van de huurprijs.

Opgave 12

In twee verschillende bosgebieden heerst een konijnenplaag. In het ene gebied probeert men hier iets aan te doen door jagers een bepaald aantal konijnen per week te laten schieten. In het andere gebied worden wolven losgelaten die naar verwachting heel wat konijnen zullen opeten.
In de tabel staan de aantallen konijnen in beide gebieden. In week 0 zijn de jagers begonnen en de wolven losgelaten.

tijd (week)	0	4	8	12	16
aantal konijnen in gebied met jagers	1500	1400	1300	1200	1100
aantal konijnen in gebied met wolven	1500	1385	1280	1180	1085

Tabel 4.12

a Van welke soort groei is er bij het gebied met de jagers en het gebied met de wolven sprake?

b Stel voor beide gebieden een formule op waarmee het aantal konijnen K in week t kan worden berekend.

c Bepaal door middel van extrapoleren het aantal konijnen in beide gebieden na 19 weken.

d In week 0 zijn er in beide gebieden evenveel konijnen. De wolven eten in het begin meer konijnen op dan dat de jagers er neerschieten. De wolven eten naarmate de tijd vordert echter steeds minder konijnen, en op een gegeven moment zijn er in beide gebieden weer evenveel konijnen. Zoek uit wanneer dat is.

e Het is de bedoeling dat het aantal konijnen halveert. Daarna moeten de jagers stoppen met het jagen op konijnen en worden de wolven gevangen en uit het gebied verwijderd. Zoek uit wanneer dat het geval is.

Opgave 13

Bekijk de twee vervaltabellen met daarin de hoeveelheid van stof K en L in microgram (µg).

tijd (h)	1	3	9	27
hoeveelheid stof K (µg)	600	200	66,7	22,2

tijd (h)	0	12	24	36
hoeveelheid stof L (µg)	600	200	66,7	22,2

Tabel 4.13

Zowel exponentiële groei als omgekeerd evenredige groei kunnen passen bij het verval van een stof.

a Welke soort groei hoort bij welke stof?

b Bepaal door middel van zowel exponentieel extrapoleren als omgekeerd evenredig extrapoleren de hoeveelheid van beide stoffen na 50 uur en vergelijk de resultaten met elkaar.

c Plot beide grafieken.

d Bereken voor beide stoffen de halveringstijd.

Opgave 14

Er wordt een stroomcircuit gemaakt met een 12 Volt accu en een weerstand R (ohm) ertussen. Door dit circuit loopt een stroom I (ampère).
Bekijk in de tabel de verschillende waarden van R en I.

R (ohm)	0,5	1	2	3	4	5	10
I (ampère)	24	12	6	4	3	2,4	1,2

Tabel 4.14

Uit de tabel is bijvoorbeeld af te lezen dat wanneer op een accu van 12 Volt een weerstand van $0{,}5$ ohm wordt aangesloten, een stroom van 24 ampère door het stroomcircuit loopt.

a Welke soort groei hoort bij deze tabel?

b Bepaal door middel van extrapoleren de hoeveelheid stroom bij een weerstand van 12 ohm.

c Plot de grafiek die hoort bij het verband tussen I en R.

d Wat gebeurt er met de stroomsterkte als de weerstand heel groot wordt? Licht je antwoord toe.

Toepassen

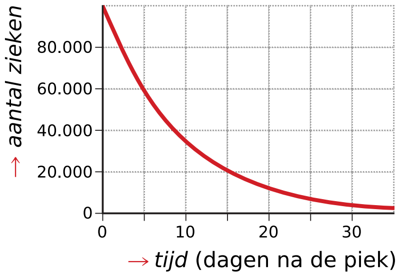
Opgave 15: Griepepidemie

Aan het begin van een griepepidemie neemt het aantal ziektegevallen exponentieel toe. Na een piek bereikt te hebben, neemt het aantal zieken ook weer af.

a Bekijk de grafiek. De eerste vijf dagen na de piek zou er zowel van lineair als exponentieel verval sprake kunnen zijn.
Stel voor beide gevallen een formule op voor het aantal zieken met als tijdseenheid het aantal dagen na de piek van de griepepidemie.

b Na de eerste vijf dagen is duidelijk te zien dat de grafiek niet lineair maar exponentieel verloopt.
Bepaal in uren nauwkeurig de halveringstijd van het aantal zieken.

c Hoeveel zou de halveringstijd bedragen als ook na de vijfde dag het aantal zieken lineair zou blijven dalen?



Figuur 4.7

Opgave 16: Aardappelteelt

Het lijkt goed te gaan met het terugdringen van het gifgebruik in de aardappelteelt. Nederlandse aardappelboeren gebruikten in 1998 gemiddeld 32 kg chemische bestrijdingsmiddelen (gif) per hectare (ha). In 2007 was dat gedaald tot $24{,}5$ kg per ha. En het gebruik daalt nog steeds.

a Neem aan dat dit gebruik lineair afnam en ook na 2007 op dezelfde wijze lineair blijft afnemen.
Bereken hoeveel kg gif per ha er dan in 2015 gebruikt wordt.

Steeds meer mensen willen biologisch geteelde aardappelen kopen. Hierdoor neemt in Flevoland het aantal hectaren waarop aardappelen biologisch geteeld worden vanaf 2007 toen het 680 ha was, exponentieel toe. Hierdoor zal dit aantal hectaren iedere 12 jaar verdubbelen. Neem aan dat de totale oppervlakte voor aardappelen in Flevoland vanaf 2007 gelijk blijft aan 20700 ha.

b Bereken in welk jaar in Flevoland het aantal hectaren biologisch geteelde aardappelen voor het eerst meer dan 10% van de totale oppervlakte voor aardappelen zal zijn.

(naar: examen havo wiskunde A in 2015, eerste tijdvak)

Opgave 17: Fukushima

De zeebeving van 11 maart 2011 met de daaropvolgende tsunami zorgde voor grote problemen bij de kerncentrale Fukushima I. Om de reactoren te koelen, werd zeewater in de reactoren gepompt. Dit water lekte, radioactief geworden, weer terug in zee. Hierdoor raakte vis besmet met radioactief jodium en moest de visvangst tijdelijk worden stopgezet.
Radioactief jodium verdwijnt volgens een exponentieel proces. De halveringstijd van radioactief jodium is 8 dagen. Op 6 april 2011 gaven metingen aan dat er 4800 keer de maximaal toegestane hoeveelheid radioactief jodium in het zeewater aanwezig was. De maximaal toegestane hoeveelheid radioactief jodium is 5 becquerel/liter. Op het moment dat de maximaal toegestane hoeveelheid werd bereikt, mocht er weer gevist worden. We gaan ervan uit dat

er na 6 april 2011 geen nieuw radioactief jodium meer in zee lekte. Bereken na hoeveel dagen er weer gevist mocht worden.

(naar: examen vwo wiskunde C in 2015, eerste tijdvak)

Testen

Opgave 18

In een klein dorpje in de provincie Zeeland werden in 2012 zeven nieuwe huizen gebouwd. Er waren toen ook precies zeven gezinnen die een huis in het dorp wilden laten bouwen.

De gemeente heeft sinds dat jaar bepaald, dat er ieder jaar maximaal drie nieuwe huizen meer in het dorp gebouwd mogen worden dan in het jaar ervoor gebouwd werden. Dus in 2013 mochten er tien nieuwe huizen gebouwd worden, in 2014 dertien, enzovoort.

De vraag naar nieuwbouwhuizen in het dorp stijgt echter vanaf 2012 exponentieel met 30% per jaar.

a Stel een formule op voor het aantal nieuwe huizen H dat in het jaar t in het dorp gebouwd mag worden. Stel ook een formule op voor het aantal gezinnen G dat in het jaar t een huis wil laten bouwen in het dorp. Neem $t = 0$ in 2012.

b In welk jaar zijn er voor het eerst $2{,}5$ keer zo veel gezinnen die een nieuwbouwhuis willen kopen in het dorp dan dat er nieuwbouwhuizen beschikbaar zijn?

Opgave 19

Jodium-131 is een onstabiele radioactieve en zeer gevaarlijke stof. Jodium-131 komt van nature niet op aarde voor, maar is vrijgekomen tijdens de kernrampen in Tsjernobyl en Fukushima. Ook wordt Jodium-131 geproduceerd in een kernreactor om het te kunnen gebruiken in de nucleaire geneeskunde ter behandeling van onder andere schildklierkanker.
Bekijk de vervaltabel van Jodium-131.

tijd (dag)	0	2	4	6	8	10	12	14	16
hoeveelheid (µg)	400	336	282	238	202	166	142	116	98

Tabel 4.15

a Welke soort groei hoort bij deze tabel?

b Bepaal door middel van extrapoleren de hoeveelheid Jodium-131 na 23 dagen.

c Plot de grafiek die hoort bij het verval van Jodium-131.

d Hoeveel dagen is de halveringstijd van Jodium-131 ongeveer?

1.5 Totaalbeeld

Samenvatten

Je hebt nu het onderwerp **Allerlei verbanden** doorgewerkt. Er moet een totaalbeeld van deze leerstof ontstaan. Ga na, of je alle bij dit onderwerp behorende begrippen kent en weet wat je er mee kunt doen. Ga ook na of je de activiteiten die staan genoemd kunt uitvoeren. Maak een eigen samenvatting!

Begrippenlijst

- formule met meerdere variabelen — grafiekenbundel
- gebied — grenslijn
- recht evenredig — omgekeerd evenredig — hyperbolisch verband
- lineair en exponentieel verband — extrapoleren

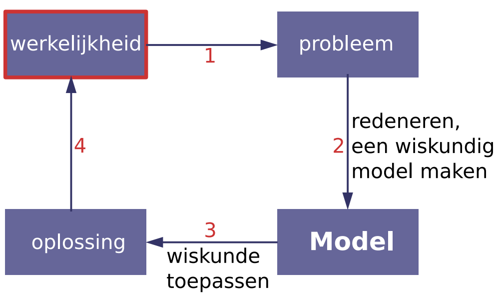
Activiteitenlijst

- werken met formules van meerdere variabelen, met name waarden invullen en formules combineren — grafiekenbundel tekenen
- een gebied aangeven met ongelijkheden — bij een stelsel ongelijkheden een gebied tekenen met behulp van grenslijnen
- een recht evenredig, een omgekeerd evenredig en een hyperbolisch verband herkennen — hierbij vergelijkingen en ongelijkheden oplossen — asymptotisch gedrag benoemen
- lineaire en exponentiële verbanden herkennen en vergelijken — bijbehorende vergelijkingen en ongelijkheden oplossen

Achtergronden

Tegenwoordig heeft in veel vakgebieden de wiskunde zijn intrede gedaan. Vaak gaat het dan om het opstellen van een **wiskundig model** aan de hand waarvan voorspellingen gedaan kunnen worden. Zo'n model bestaat uit variabelen en bijbehorende formules.

Een model is een vereenvoudiging van de werkelijkheid waarin nog alle eigenschappen zijn terug te vinden die belangrijk zijn voor de beschrijving van het verschijnsel dat je wilt verklaren. Het bewust opstellen van zo'n model heet **modelleren**.
Bij het modelleren volg je een viertal vaste stappen.



Figuur 5.1

1. Je kijkt naar de werkelijkheid en stelt jezelf een vraag: de probleemstelling. Je bedenkt welke grootheden en variabelen een rol spelen.
2. Je vereenvoudigt de werkelijkheid door aannames te doen en ontwerpt een wiskundig model dat zo goed mogelijk bij de probleemstelling past. Je geeft duidelijke definities van de grootheden waartussen je verbanden gaat zoeken. Je moet ook goed bijhouden waarom je bepaalde dingen weglaat.
3. Je zoekt het antwoord op je vraag door in je model wiskundige berekeningen toe te passen. Het antwoord kan de oplossing van

het probleem zijn, maar ook een beschrijving van de bepaalde situatie.

4. Je kijkt of je antwoord wel bij de werkelijkheid past. Je moet je antwoord 'terugvertalen'. Als dat kan, ontwerp je ook een test. Daarmee onderzoek je of je model goed genoeg was of moet worden bijgesteld en doorloop je de cyclus opnieuw.

Testen

Opgave 1

In de micro-economie wordt het volgende rekenmodel voor de winst van de verkoop van een bepaald product gehanteerd als het bedrijf de enige aanbieder is.

Het aantal verkochte producten hangt alleen af van de prijs p in euro per stuk. Hoe hoger de prijs, hoe lager de hoeveelheid q die van dit product wordt verkocht per tijdseenheid. Bijvoorbeeld: per week kan dan gelden $q = 500 - 2p$.

De inkoopkosten hangen weer af van de inkoopprijs per eenheid en de voorraadkosten. De inkoopprijs van een eenheid product kan bijvoorbeeld € 5,00 zijn en de voorraadkosten kunnen € 2000,00 per week zijn.

Voor de opbrengst als wekelijks de hele voorraad wordt verkocht, geldt $TO = p \cdot q$, de wekelijkse kosten noem je TK en de winst is $TW = TO - TK$.

a Waarom is $TO = p \cdot q$?

b Stel een formule op met TK uitgedrukt in q.

c Laat zien dat $TW = pq - 2000 - 5q$.

d Toon aan dat je bij invullen van $q = 500 - 2p$, $TW = -2p^2 + 510p - 4500$ krijgt.

e Bepaal de maximale winst.

Opgave 2

Een olympische triathlon bestaat uit de onderdelen zwemmen, fietsen en hardlopen. Het energieverbruik in kcal tijdens een triathlon kan berekend worden met de formule $E = 6kZ + 12kF + 11{,}5kH$. Hierin is E het energieverbruik in kcal, k het gewicht in kg, Z het aantal uur zwemmen, F het aantal uur fietsen, en H het aantal uur hardlopen.

a Willem weegt 70 kg en traint voor een triathlon. Hij gaat eerst een kwartier zwemmen, dan een half uur fietsen, en dan nog drie kwartier hardlopen.
Bereken het energieverbruik van Willem in kcal.

b Een olympische triathlon bestaat uit 1,5 km zwemmen, 40 km fietsen, en 10 km hardlopen. Willem doet gemiddeld een half uur over het onderdeel zwemmen. Zijn hardlooptijd varieert nog heel erg. Zijn fietstijd varieert tussen de 1,2 uur en 1,6 uur.
Maak met de grafische rekenmachine bij deze formule een grafiekenbundel met $F = 1{,}2$, $F = 1{,}4$, $F = 1{,}6$.

c Bij de laatste keren dat Willem trainde voor de olympische triathlon had hij een energieverbruik van tussen de 2200 en 2400 kcal en een fietstijd van tussen de $1{,}2$ en $1{,}4$ uur.
Geef in de grafiekenbundel het gebied aan dat hierbij hoort.

d Welke ongelijkheden beschrijven dit gebied?

e Tussen welke tijden zat de hardlooptijd van Willem tijdens deze laatste trainingen? Geef je antwoord in minuten nauwkeurig.

Opgave 3

Los algebraïsch de volgende vergelijkingen op. Rond indien nodig af op twee decimalen.

a $\frac{500}{v-10} = 20$

b $\frac{500}{v} - 10 = 20$

Opgave 4

De gymnastiekdocenten van een school organiseren een prestatieloop voor de vierde klassen. Er moet een afstand van vijftien kilometer worden afgelegd. De gemiddelde snelheid voor een loper in kilometer per uur is v, de totale tijd t.

a Wat voor soort verband bestaat er tussen de snelheid en de tijd?

b Geef een formule die de looptijd t uitdrukt in de gemiddelde snelheid v.

c Hoe groot is de snelheid bij een looptijd van 100 minuten?

Alle lopers zijn onderweg ongeveer vijf minuten tijd kwijt met het wachten bij een aantal stempelposten.

d Maak met dit gegeven een formule voor t van de vorm: $t = \frac{a}{v} + c$.

e Bereken met de tweede formule de gemiddelde snelheid van een loper die in het totaal een uur en twintig minuten nodig heeft.

Opgave 5

De afgelopen jaren is het aantal leerlingen op veel basisscholen sterk afgenomen. In het jaar 2002 zaten er op basisschool De Regenboog nog 280 leerlingen, en in 2012 waren dat er nog maar 210. De afname op basisschool De Regenboog verloopt lineair. Op basisschool De Margriet zaten in 2002 nog 250 leerlingen, en in 2014 waren dat er nog maar 188. Hier is sprake van exponentieel verval. Voor beide basisscholen is de verwachting dat de afname van leerlingen zich zo zal blijven voortzetten.

a Stel voor beide basisscholen een formule op waarmee het aantal leerlingen berekend kan worden.

b De gemeente waarin beide basisscholen staan, heeft de scholen medegedeeld dat ze moeten fuseren (samenvoegen) als ze aan het begin van het jaar 2025 gezamenlijk minder dan 250 leerlingen hebben. Ga na of de twee scholen in 2025 zullen moeten fuseren.

c Bepaal voor beide scholen de halveringstijd.

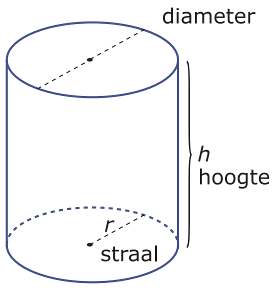
Toepassen

Opgave 6: Literblik

Stel je voor dat een fabrikant zuiver cilindervormige blikken nodig heeft met een inhoud van 1 L. Om de kosten te drukken wil hij zo min mogelijk blik gebruiken. Welke afmetingen moet zo'n literblik krijgen?

Voor een cilindervormig literblik (afmetingen in cm) geldt:

- de inhoud is: $I = \pi \cdot r^2 \cdot h = 1000$ cm^3;
- de oppervlakte is: $A = 2\pi r h + 2\pi r^2$ cm^2.



Figuur 5.2

Je kunt hierbij een formule opstellen voor A als functie van r. Zo kun je bepalen voor welke waarde van r de oppervlakte van het literblik zo klein mogelijk is. En zo bepaal je de afmetingen van het literblik met de kleinste materiaalkosten.

a Wordt bij de formule voor de oppervlakte A ook het deksel en de bodem van het blik meegerekend? Waaraan zie je dat?

b Laat zien hoe je de formule voor A als functie van r kunt afleiden.

c Maak zelf de bijpassende grafiek van A en bepaal voor welke r de materiaalkosten zo klein mogelijk zijn.

d Voor bepaalde stoffen (bijvoorbeeld verschillende soorten olie) worden zuiver balkvormige blikken gebruikt. In verband met het makkelijk stapelen worden voor een bepaald merk olie 10-liter-blikken gebruikt waarvan lengte en breedte gelijk zijn. Bereken ook voor deze blikken de afmetingen bij een zo klein mogelijke oppervlakte aan blik.

Opgave 7: Paracetamol

Paracetamol is een veelgebruikte pijnstiller, die in tabletvorm te koop is. Voor volwassenen zijn er tabletten die 500 mg paracetamol bevatten.

Na het innemen van een tablet wordt de 500 mg paracetamol via maag en darmen bijna volledig in het bloed opgenomen. Gebruik het volgende wiskundige model voor de opname van paracetamol in het bloed:

Tien minuten na het innemen van een tablet is de helft van de paracetamol opgenomen in het bloed. De andere helft zit dan nog in maag en darmen. Van de achtergebleven paracetamol in maag en darmen wordt in de volgende tien minuten weer de helft opgenomen in het bloed. Ook daarna wordt iedere tien minuten de helft van de paracetamol die nog in maag en darmen zit, opgenomen in het bloed.

Een volwassene neemt om 9:00 uur één tablet van 500 mg in.

a Laat met een berekening zien dat na één uur ongeveer 492 mg paracetamol in het bloed is opgenomen.

b Stel een formule op voor het aantal mg paracetamol P in de maag na t uur.

Het pijnstillend effect is merkbaar zolang de hoeveelheid paracetamol in het bloed meer is dan 200 mg. Als de hoeveelheid paracetamol onder de 200 mg zakt, is het pijnstillend effect niet meer merkbaar: de tablet is uitgewerkt. Een volwassene die om 9.00 uur een tablet heeft ingenomen, zal merken dat deze tablet in de loop van de middag is uitgewerkt.

c Bereken op welk moment de tablet is uitgewerkt. Geef je antwoord in uren en minuten nauwkeurig.

Examen

Opgave 8: Park ’N Fly

In de Verenigde Staten komen veel mensen met de auto naar het vliegveld. Ze parkeren hun auto op een parkeerterrein in de buurt. Eén van de parkeerterreinen bij het vliegveld van Minneapolis wordt beheerd door het bedrijf Park ’N Fly. In 2010 had dit terrein 2100 parkeerplaatsen. Het normale parkeertarief in 2010 was $ 10,00 (10 dollar) per dag. Online gekochte parkeerkaarten waren $ 1,00 per dag goedkoper.

Ga in deze opgave uit van de situatie in 2010 en neem aan dat alle klanten die hun parkeerkaart online kopen, komen opdagen. Reken alleen met de parkeerprijs per dag.

a Op een dag zijn 2065 parkeerplaatsen bezet. De totale inkomsten voor het bedrijf zijn die dag $ 20214,00.
Bereken hoeveel klanten die dag hun parkeerkaart online gekocht hebben.

b Park ’N Fly wil dat de opbrengst per dag minimaal $ 20250,00 is.
Stel bij deze situatie twee ongelijkheden op en kleur het gebied waarbinnen de oplossingen zich bevinden.

c Lees uit het gekleurde gebied een oplossing af. En laat met berekeningen zien dat deze oplossing voldoet aan beide ongelijkheden.

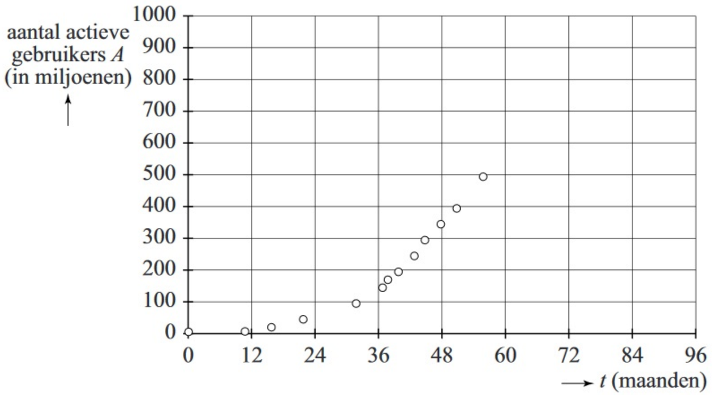
Van maandag tot en met donderdag is het parkeerterrein goed gevuld. Maar op vrijdag en in het weekend zijn er nogal wat lege plaatsen. Het bedrijf wil graag dat deze plaatsen benut worden, desnoods tegen een lager tarief. Er wordt voor vrijdag en het weekend een nieuw tarief geïntroduceerd, het actietarief. De hoogte van het actietarief wordt slechts een paar dagen van tevoren bepaald en parkeerkaarten tegen dit tarief kunnen alleen online gekocht worden. Uit onderzoek blijkt dat bij een actietarief van $ 6,00 er 1500 klanten hun auto tegen dit tarief zullen parkeren bij Park ’N Fly. Bij een actietarief van $ 5,00 zijn dat er zelfs 1700.
Stel dat het actietarief wordt bepaald op $ 4,20.

d Bereken met lineair extrapoleren, uitgaande van de gegeven waarden, hoeveel klanten hun auto dan tegen dit tarief bij Park ’N Fly zullen parkeren.

(naar: examen havo wiskunde A in 2014, tweede tijdvak)

Opgave 9: Sociaal netwerk

Facebook is een sociaalnetwerksite, opgericht door Mark Zuckerberg in februari 2004. In het begin konden alleen studenten van Harvard College lid worden, later werden ook studenten van andere universiteiten toegelaten. In september 2006 werd Facebook geheel openbaar. Iedereen vanaf 13 jaar, waar ook ter wereld, kreeg de mogelijkheid om zich te registreren en actief gebruik te gaan maken van de site.
Het aantal actieve gebruikers steeg de eerste jaren spectaculair. Zie figuur 1, waarin het aantal actieve gebruikers op verschillende momenten is aangegeven.



Figuur 5.3

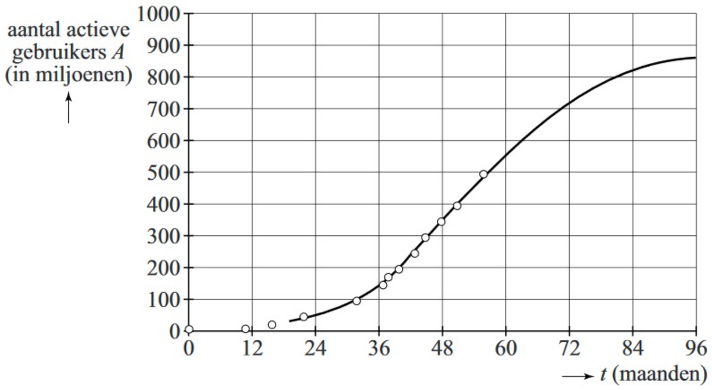
Op 1 december 2005, dat is bij $t = 0$, waren er $5{,}5$ miljoen actieve gebruikers, 43 maanden later, op 1 juli 2009, waren het er al 244 miljoen. Neem aan dat er in deze periode bij benadering sprake was van exponentiële groei.

a Bereken voor deze periode het groeipercentage per maand.

In de maanden na 1 juli 2009 groeide het aantal actieve gebruikers niet meer exponentieel maar bij benadering lineair, van 244 miljoen op 1 juli 2009 tot 493 miljoen op 1 augustus 2010. Er werd in 2011 voorspeld dat de groei zich op deze manier zou voortzetten.

b Bereken hiermee het aantal actieve gebruikers op 1 december 2013.

Het bleek erg optimistisch om aan te nemen dat de groei zich lineair voortzet. Al in 2011 voorspelden sommigen dat de groei verder zou afnemen. In figuur 2 zie je een grafiek die bij deze voorspelling past.



Figuur 5.4

Bij deze grafiek hoort de formule:

$$A = \frac{4500}{5+310 \cdot 0{,}926^t}$$

Hierin is A het aantal actieve gebruikers in miljoenen en t de tijd in maanden met $t = 0$ op 1 december 2005.

c Bereken voor welke gehele waarde van t er volgens de formule voor het eerst meer dan 730 miljoen actieve gebruikers zijn.

Volgens de formule zal het aantal actieve gebruikers uiteindelijk nauwelijks meer toenemen en een grenswaarde benaderen.

d Bepaal deze grenswaarde met behulp van de formule.

(bron: examen havo wiskunde A in 2015, eerste tijdvak)

2

Conclusies trekken

2.1 Soorten variabelen

Inleiding

Je hebt al gezien wat het verschil is tussen kwalitatieve en kwantitatieve statistische variabelen. Ook weet je dat je kwantitatieve variabelen kunt verdelen in discrete en continue variabelen. Maar ook kwalitatieve variabelen kun je nog onderverdelen. Er is bijvoorbeeld verschil tussen de variabele *man of vrouw* en de variabele *maand*. Daarbij wordt de term meetniveaus gebruikt.

Figuur 1.1

Je leert in dit onderwerp

- of een kwalitatieve variabele ordinaal of nominaal is;
- of een kwantitatieve variabele discreet of continu is.

Voorkennis

- soorten statistische variabelen herkennen;
- de begrippen onderzoek, steekproef, populatie en representatief, simulatie;
- het begrip normale verdeling met gemiddelde en standaardafwijking;
- betrouwbaarheidsintervallen en foutenmarges bepalen bij het schatten van populatieproporties en populatiegemiddelden.

Verkennen

Opgave V1

Er zijn verschillende soorten statistische variabelen, het verschil tussen kwalitatieve en kwantitatieve variabelen en dat tussen discrete en continue variabelen ken je al.

a Geef een paar voorbeelden van een kwalitatieve variabele.

b Geef een paar voorbeelden van kwantitatieve variabelen.

c Wat is het verschil tussen een discrete en een continue variabele? Waarom kan dit alleen bij kwantitatieve variabelen?

Uitleg 1

Bij statistische vragen naar kenmerken worden kwalitatieve variabelen gebruikt, bijvoorbeeld geslacht, lievelingskleur, tevredenheid of beroep.

Bij sommige kwalitatieve vragen heeft het zin om de antwoorden te ordenen, bij andere niet.

Een voorbeeld waarbij een ordening zin heeft:
"Hoe tevreden bent u?"
Antwoordmogelijkheden kunnen bijvoorbeeld worden geordend als:
zeer tevreden / tevreden / neutraal / ontevreden / zeer ontevreden
De kwalitatieve variabele is ordinaal.

Een voorbeeld waarbij een ordening geen zin heeft:
"Wat is uw favoriete automerk?"
Antwoordmogelijkheden zijn dan bijvoorbeeld Audi, BMW, Dacia, Renault, enzovoort. Daarin zit geen ordening. De kwalitatieve variabele is nominaal.

Figuur 1.2

Opgave 1

Geef van de statistische kwalitatieve variabelen aan of het nuttig is om de antwoorden te sorteren of te vergelijken.

a *soort boom*

b *gevoel van geluk*

c *e-mailadres*

d *woonplaats*

Uitleg 2

Bij statistische vragen naar hoeveelheden worden kwantitatieve variabelen gebruikt. Kwantitatief wil zeggen dat het antwoord is uit te drukken in een getal en dat daarmee gerekend kan worden. Voorbeelden zijn snelheid, lengte en leeftijd.

Soms zijn alle tussenliggende waarden als antwoord mogelijk:
"Hoe lang bent u?"
Antwoord: $182{,}3$ cm.
De kwantitatieve variabele is continu.

Soms zijn tussenliggende waarden niet mogelijk:
"Hoeveel kinderen heeft u?"
Antwoord $4{,}33$ is dan onzin.
De kwantitatieve variabele is discreet.

Opgave 2

Geef van de statistische kwantitatieve variabelen aan of alle tussenliggende waarden als antwoord voor kunnen komen.

a *schoenmaat*

b *aantal leerlingen*

c *gemiddeld aantal leerlingen in een klas*

Theorie en voorbeelden

Om te onthouden

Een **kwantitatieve variabele** kan in een getal worden uitgedrukt (bijvoorbeeld *lengte*, *hoogte van het inkomen*, *aantal* of *gemiddelde*).

- **Discrete kwantitatieve variabelen** zijn variabelen die een beperkt aantal waarden kunnen aannemen. Een voorbeeld is *aantal*.
- **Continue kwantitatieve variabelen** zijn variabelen die ook tussenliggende waarden kunnen aannemen. Een voorbeeld is *lengte*.

Een **kwalitatieve variabele** beschrijft een bepaald kenmerk (*geslacht*, *kleur ogen*, *godsdienst*, *naam*, *tevredenheid*, *rijkdom*).

- **Ordinale kwalitatieve variabelen** zijn variabelen waarvan de waarden kunnen worden geordend. Een voorbeeld is *tevredenheid*.
- **Nominale kwalitatieve variabelen** zijn variabelen waarvan het ordenen van de antwoorden geen zin heeft. Een voorbeeld is *kleur van de ogen*.

Je noemt dit wel de **meetniveaus** van de kwalitatieve variabelen. Ook kwantitatieve variabelen hebben meetniveaus, maar die vallen buiten het bestek van havo wiskunde A.

Verder kun je deze indeling niet altijd heel strikt hanteren: bijvoorbeeld rapportcijfers zijn op de eerste blik kwantitatief, maar als je 10 = uitmuntend, 9 = zeer goed, 8 = goed, en dergelijke gebruikt, zijn ze eerder kwalitatief, ordinaal.

Voorbeeld 1

Geef van de variabelen aan of ze kwalitatief, kwantitatief, nominaal, ordinaal, discreet en/of continu zijn.

1. *politieke voorkeur*
2. *inkomen*
3. *telefoonnummer*
4. *kwaliteitsklasse*
5. *aantal talen*
6. *gemiddeld aantal*

Antwoord

1. *politieke voorkeur*: kwalitatief, nominaal
2. *inkomen*: kwantitatief, continu
3. *telefoonnummer*: kwalitatief, nominaal
4. *kwaliteitsklasse*: kwalitatief, ordinaal
5. *aantal talen*: kwantitatief, discreet
6. *gemiddeld aantal*: kwantitatief, meestal continu

Opgave 3

Geef van de variabelen aan of ze kwalitatief, kwantitatief, nominaal, ordinaal, discreet en/of continu zijn.

a *favoriete band*

b *aanschafprijs smartphone*

c *bloedgroep*

d *maximale windsnelheid*

e *huisnummer*

f *kwaliteit docent*

Voorbeeld 2

Drie centrummaten zijn: gemiddelde, mediaan en modus. Spreidingsmaten zijn bijvoorbeeld: interkwartielafstand en standaardafwijking.

Geef van de variabelen aan welke centrummaten voor die variabele bruikbaar zijn, en geef aan of voor deze variabele spreidingsmaten bruikbaar zijn of niet.

1. *politieke voorkeur*
2. *inkomen*
3. *telefoonnummer*
4. *kwaliteitsklasse*
5. *aantal talen*
6. *gemiddeld aantal*

Antwoord

1. *politieke voorkeur*
   Centrummaten: modus. Spreidingsmaten: niet bruikbaar.
2. *inkomen*
   Centrummaten: gemiddelde, mediaan, modus. Spreidingsmaten: wel bruikbaar.
3. *telefoonnummer*
   Centrummaten: geen. Spreidingsmaten: niet bruikbaar.
4. *kwaliteitsklasse*
   Centrummaten: modus, mediaan kan ook maar is vaak niet erg nuttig. Spreidingsmaten: niet bruikbaar.
5. *aantal talen*
   Centrummaten: gemiddelde, mediaan, modus. Spreidingsmaten: wel bruikbaar.
6. *gemiddeld aantal*
   Centrummaten: gemiddelde, mediaan, modus. Spreidingsmaten: wel bruikbaar.

Opgave 4

Geef van de variabelen aan welke soorten diagrammen er bij gemaakt kunnen worden.

a *favoriete band*

b *aanschafprijs smartphone* (met klassenindeling)

c *bloedgroep*

d *maximale windsnelheid* (zonder klassenindeling)

e *huisnummer*

f *kwaliteit docent*

Voorbeeld 3

Geef van de variabelen aan welke soorten diagrammen er bij gemaakt kunnen worden.

	soort variabele				
soort diagram	kwalitatief ordinaal (*kwaliteitsklasse*)	kwalitatief nominaal (*favoriet merk*)	kwantitatief discreet (*aantal kinderen*)	kwantitatief continu zonder klassenindeling (*lichaamslengte*)	kwantitatief continu met klassenindeling (*lichaamslengte*)
staafdiagram					
lijndiagram					
dotplot					
cirkeldiagram					
boxplot					
steelbladdiagram					
cumulatief frequentiepolygoon					

Tabel 1.1

Antwoord

Een staafdiagram, lijndiagram, dotplot en cirkeldiagram kunnen alleen worden gebruikt bij een klein aantal antwoorden of bij een indeling in klassen. Een lijndiagram wekt de indruk dat er sprake is van stijging of daling, maar dat is bij een nominale of ordinale variabele niet zo.

Een boxplot kan alleen worden gebruikt bij kwantitatieve variabelen. Een klassenindeling is al een samenvatting van gegevens, vervolgens een boxplot toepassen zou een volgende samenvatting zijn. Het is dan beter om geen klassenindeling te gebruiken.

Een steelbladdiagram kan alleen bij kwantitatieve variabelen gebruikt worden.

Een cumulatief frequentiepolygoon kan alleen bij kwantitatieve of bij ordinale variabelen gebruikt worden.

	soort variabele				
soort diagram	kwalitatief ordinaal (*kwaliteits-klasse*)	kwalitatief nominaal (*favoriet merk*)	kwantitatief discreet (*aantal kinderen*)	kwantitatief continu zonder klassen-indeling (*lichaams-lengte*)	kwantitatief continu met klassen-indeling (*lichaams-lengte*)
staafdiagram	x	x	x		x
lijndiagram			x		x
dotplot	x	x	x		x
cirkeldiagram	x	x	x		x
boxplot			x	x	
steelbladdiagram			x	x	x
cumulatief frequentiepolygoon	x		x	x	x

Tabel 1.2

Opgave 5

Geef van de variabelen aan welke centrummaten voor die variabele bruikbaar zijn, en geef aan of voor deze variabele spreidingsmaten bruikbaar zijn of niet.

a *favoriete band*

b *aanschafprijs smartphone*

c *bloedgroep*

d *maximale windsnelheid*

e *huisnummer*

f *kwaliteit docent*

Verwerken

Opgave 6

Geef van de variabelen aan of ze kwalitatief of kwantitatief, nominaal of ordinaal en discreet of continu zijn.

a *geluksgevoel*

b *het gewicht van sinaasappels*, gemeten in grammen

c *voorkeur voor een automerk*

d *de sterkte van een aardbeving*, gemeten volgens de schaal van Richter

Opgave 7

Leg uit waarom ...

a de variabele *politieke partij waarop iemand stemt* een kwalitatieve nominale variabele is.

b de variabele *schoenmaat* een kwantitatieve discrete variabele is.

Opgave 8

De kwaliteit van de schoonmaak van een school kan worden beschreven met variabelen.
Geef vier variabelen, elk van een andere soort, om schoonmaakwerk te beschrijven.

Opgave 9

Drie veelgebruikte centrummaten zijn: gemiddelde, mediaan en modus. Geef van de volgende variabelen aan welke centrummaten voor die variabele bruikbaar zijn.

a *telefoonnummers in een adressenlijst*

b *mate van tevredenheid*

c *aantal kinderen in een gezin*

d *buitentemperatuur*

e Hoe zit het bij *buitentemperatuur*, als de metingen ingedeeld zijn in klassen?

Opgave 10

Spreidingsmaten zijn bijvoorbeeld: interkwartielafstand en standaardafwijking.

a Zijn deze maten bruikbaar bij nominale kwalitatieve variabelen, ja of nee?

b Zijn deze maten bruikbaar bij discrete kwantitatieve variabelen, ja of nee?

Opgave 11

Er zijn verschillende (soorten) diagrammen, zoals:

- dotplot
- cirkeldiagram
- steelbladdiagram
- cumulatief frequentiepolygoon

Geef bij de volgende soorten variabelen aan in welke van deze diagrammen de variabele kan worden weergegeven. Geef ook aan waarom de andere diagrammen niet kunnen.

a continue kwantitatieve variabele

b continue kwantitatieve variabele (met klassenindeling)

c nominale kwalitatieve variabele

d ordinale kwalitatieve variabele

Toepassen

In de **Theorie** werd opgemerkt, dat ook kwantitatieve variabelen meetniveaus hebben.
Dat zijn het **ratio-meetniveau** en het **interval-meetniveau**.
Bij een ratio-meetniveau geldt: een twee keer zo grote waarde van de variabele betekent altijd een twee keer zo grote hoeveelheid (soms van een andere variabele). Ook hebben kwantitatieve variabelen met een ratio-meetniveau een absoluut (natuurlijk) nulpunt en geen negatieve waarden.

Voorbeelden van variabelen met ratio-meetniveau zijn *gewicht*, *leeftijd*, *lengte* en *afstand*.

Voorbeelden van variabelen met interval-meetniveau zijn:

- *Intelligentiequotiënt* (IQ), want er geldt niet: iemand met IQ 110 is $1{,}1$ keer zo intelligent als iemand met IQ 100.
- *Temperatuur in graden Celcius*, want deze kan negatief zijn en bovendien betekent een 2 keer zo hoge temperatuur niet dat het ook 2 keer zo warm is.

Opgave 12: Meetniveaus kwantitatieve variabelen

Bekijk in **Toepassen** welke meetniveaus een kwantitatieve variabele kan hebben.

a Geef van de volgende variabelen aan welk meetniveau ze hebben en waarom.

- *temperatuur in Kelvin*;
- *geluksgevoel* op een schaal van 1 tot 10;
- *score voor een toets*;
- *percentage deelnemers*.

b *Voorkeur voor een politieke partij* is een nominale variabele: je kunt de waarden niet zinnig sorteren.

Mate van teleurstelling is een ordinale variabele: de waarden kunnen worden gesorteerd.

IQ is een variabele met een interval-meetniveau: verschillen tussen waarden van die variabele kunnen worden berekend.

Welke extra mogelijkheid hebben variabelen met een ratio-meetniveau om ermee te rekenen?

c Hebben de meetniveaus van kwantitatieve variabelen ook iets te maken met het discreet of continu zijn van zo'n variabele? Geef voorbeelden.

Testen

Opgave 13

Bekijk het bestand **Gegevens 154 leerlingen**. Je ziet daarin onder andere de statistische variabelen *geslacht*, *lengte*, *huiswerk*, *profiel* en *plezier*.

a Bekijk deze vijf variabelen. Beschrijf bij elke kwalitatieve variabele welk meetniveau hij heeft.

b Beschrijf bij elk van deze vijf variabelen welk van de volgende diagrammen kan worden gebruikt om de gegevens te presenteren: dotplot, boxplot, staafdiagram, lijndiagram, cirkeldiagram.

c Beschrijf bij elk van deze vijf variabelen welk van de volgende centrummaten kan worden gebruikt om de gegevens te presenteren: gemiddelde, modus, mediaan.

d Beschrijf bij elk van deze vijf variabelen welk van de volgende spreidingsmaten kan worden gebruikt om de gegevens te presenteren: spreidingsbreedte, kwartielafstand, standaardafwijking.

2.2 Verschil kwalitatieve variabelen

Inleiding

Kun je zonder meer zeggen dat mannen langer zijn dan vrouwen? Dat er meer meisjes wiskunde A kiezen dan jongens? Dat Belgen meer vreemde talen spreken dan Nederlanders?

Bij de verschillende soorten statistische variabelen horen verschillende meetniveaus, die ken je al. Als je verschillen tussen statistische variabelen in kaart wil brengen, hangt de manier waarop je dit kunt doen af van het meetniveau van deze variabelen.

Figuur 2.1

Je leert in dit onderwerp

- drie manieren om het verschil tussen twee kwalitatieve variabelen te beschrijven en te interpreteren;
- werken met een formulekaart om deze verschillen te beschrijven.

Voorkennis

- soorten statistische variabelen herkennen;
- de begrippen onderzoek, steekproef, populatie en representatief, simulatie;
- meetniveaus onderscheiden bij antwoordmogelijkheden op vragen.

Verkennen

Opgave V1

Langzamerhand kom je in de buurt van echt statistisch onderzoek.

Probeer te beschrijven hoe je zou onderzoeken of meisjes naar verhouding structureel vaker wiskunde A kiezen en jongens naar verhouding wiskunde B.

Uitleg 1

Er is een sportdag op school. De leerlingen uit klas 1 en klas 2 mogen kiezen: binnensport of buitensport. De sportkeuze is een nominale kwalitatieve variabele. Van de leerlingen uit klas 1 kiest 21% binnensport. Van de leerlingen uit klas 2 kiest 43% binnensport.

Hoe verschillend kiezen de leerlingen uit klas 1 en klas 2?

Het verschil tussen deze percentages is $43 - 21 = 22\%$.
Om antwoord te geven op de vraag gebruik je vuistregels zoals deze:

- Als het verschil groter of gelijk is aan 40% is de conclusie:
  Het verschil is groot.
- Als het verschil tussen 20% en 40% is, is de conclusie:
  Het verschil is middelmatig.

- Als het verschil kleiner of gelijk is aan 20% is de conclusie:
  Het verschil is gering.

Volgens deze vuistregels is de conclusie:
Het verschil tussen klas 1 en klas 2 is ‘middelmatig’ als je kijkt naar de (percentages) sportkeuze.

Leer dergelijke vuistregels niet uit het hoofd, je krijgt ze gegeven als je ze moet gebruiken.

Opgave 1

Gebruik de gegevens uit **Uitleg 1**. In klas 3 kiest 48% voor binnensport.

a Als je op de percentages binnensportkiezers let, wat kun je dan zeggen over het verschil tussen klas 2 en klas 3?

b Als je op de percentages binnensportkiezers let, wat kun je dan zeggen over het verschil tussen klas 1 en klas 3?

Opgave 2

Om te onderzoeken of vitamine C helpt tegen verkoudheid heeft een onderzoeker 141 personen vitamine C toegediend en 138 personen een placebo (fopmiddel). De personen in kwestie wisten niet of ze vitamine C of de placebo toegediend kregen.

De onderzoeker telde het aantal personen dat verkouden werd en maakte deze tabel.

	verkouden	niet verkouden
vitamine C	19	122
placebo	29	109
totaal	48	231

Tabel 2.1

a Van welke soort is de statistische variabele?

b Hoeveel procent van de verkouden personen slikte vitamine C?

c Hoeveel procent van de personen die niet verkouden zijn, slikte vitamine C?

d Wat is het verschil in percentage verkouden personen van de vitamine C-slikkers en de placebo-slikkers?

e Is het verschil tussen wel of niet verkouden worden bij wel of niet vitamine C slikken gering, middelmatig of groot?

Uitleg 2

Regelmatig heb je aantallen eerst omgerekend naar percentages. Er is ook een manier om direct uit aantallen conclusies te trekken.

Om te onderzoeken of vitamine C helpt tegen verkoudheid heeft een onderzoeker 141 personen vitamine C toegediend en 138 personen een placebo (fopmiddel). De personen in kwestie wisten niet of ze vitamine C of de placebo toegediend kregen.

	verkouden	niet verkouden
vitamine C	19 (a)	122 (b)
placebo	29 (c)	109 (d)

Tabel 2.2

De onderzoeker telde het aantal personen dat verkouden werd en maakte deze kruistabel.

Met deze aantallen kun je de groepen vergelijken met behulp van het getal *phi* ofwel φ (spreek uit: ‘fi’):

- Noem de waarden in de vier cellen van de kruistabel a, b, c en d op de manier die je in de tabel ziet.
- Vul deze waarden in de formule voor *phi* $= \varphi$ in:

  $$\varphi = \frac{ad-bc}{\sqrt{(a+b)\cdot(a+c)\cdot(b+d)\cdot(c+d)}}$$

- Je vindt: $\varphi \approx -0{,}100$.

Conclusies trek je met de volgende vuistregels:

- Als $\varphi < -0{,}4$ of $\varphi > 0{,}4$ is de conclusie: het verschil is groot.
- Als $-0{,}4 \leq \varphi < -0{,}2$ of $0{,}2 < \varphi \leq 0{,}4$ is de conclusie: het verschil is middelmatig.
- Als $-0{,}2 \leq \varphi \leq 0{,}2$, is de conclusie: het verschil is gering.

De vuistregels staan ook op de **Formulekaart**.

Op grond van deze vuistregels trek je een conclusie.

Opgave 3

Gebruik de gegevens uit **Uitleg 2**.

a Bereken φ nu zelf.

b Welke conclusie trek je volgens de vuistregels op de formulekaart?

Opgave 4

Gebruik de gegevens uit **Uitleg 2**.

a Bereken φ bij de volgende tabel.

aantal	jongen	meisje
houdt van Harry Potter	30	13
houdt niet van Harry Potter	12	40

Tabel 2.3

b Wat kun je zeggen over de grootte van het verschil tussen jongens en meisjes, als je kijkt of ze houden van Harry Potter?

Uitleg 3

De leerlingen in 4 havo zijn naar een toneelvoorstelling geweest. Daarna werd hun gevraagd of ze de voorstelling boeiend vonden, hun *voorstellingsbeleving*.

In de eerste tabel staan de resultaten uitgesplitst naar profiel.

Om goed te kunnen vergelijken bij groepen die niet even groot zijn, moeten eerst de absolute aantallen worden omgezet naar percentages. In de tweede tabel staan onder p de percentages en onder cp de cumulatieve percentages.

	CM	EM	NG	NT
1 = niet boeiend	5	8	6	17
2 = gaat wel	12	12	18	13
3 = boeiend	9	18	15	8
4 = erg boeiend	9	10	6	2

	EM			NG			
	aantal	p	cp	aantal	p	cp	Vcp
1 = niet boeiend	8	16,7	16,7	6	13,3	13,3	3,3
2 = gaat wel	12	25,0	41,7	18	40,0	53,3	11,7
3 = boeiend	18	37,5	79,2	15	33,3	86,7	7,5
4 = erg boeiend	10	20,8	100,0	6	13,3	100,0	0,0
	48			45			

Figuur 2.2

De percentages onder p van de EM-leerlingen en de NG-leerlingen verschillen: procentueel hebben veel meer NG-leerlingen ‘gaat wel’ geantwoord dan dat dit bij de EM-leerlingen het geval was. En procentueel hebben veel minder NG-leerlingen ‘erg boeiend’ geantwoord dan dat dit bij de EM-leerlingen het geval was. Toch hebben procentueel ook minder NG-leerlingen ‘niet-boeiend’ geantwoord. Een conclusie trekken is nog niet eenvoudig.

Daarom is voor deze situatie afgesproken dat het maximale verschil van de cumulatieve percentages $\max V_{cp}$ wordt gebruikt als maat voor het verschil tussen twee groepen. De vuistregels zijn nu:

- Als $\max V_{cp} \leq 20\%$ is het verschil ‘gering’.
- Als $20\% < \max V_{cp} \leq 40\%$ is het verschil ‘middelmatig’.
- Als $\max V_{cp} > 40\%$ is het verschil ‘groot’.

Je vindt deze afspraken op de **Formulekaart**.

Opgave 5

Gebruik de gegevens uit **Uitleg 3**.

a Waarom is de variabele *voorstellingsbeleving* een ordinale kwalitatieve variabele?

b Vergelijk nu de cumulatieve percentages. Wat betekent het dat de cumulatieve percentages van de EM-leerlingen zowel bij ‘gaat wel’ als bij ‘boeiend’ lager zijn dan die van de NG-leerlingen?

c In de kolom V_{cp} zie je het verschil van de cumulatieve percentages. Laat zien dat daarbij niet wordt gelet op welk cumulatieve percentage groter is.

d Waarom is de V_{cp} bij de hoogste waarde van de variabele altijd 0?

e Hoe groot is het maximale verschil $\max V_{cp}$ van de cumulatieve percentages?

f Schrijf de conclusie op.

Opgave 6

Vergelijk de *voorstellingsbeleving* van de NG-leerlingen en de NT-leerlingen.

a Bepaal bij elke waarde van deze variabele het verschil tussen de cumulatieve relatieve percentages. Neem het absolute verschil, laat dus mintekens weg.

b Bepaal nu $\max V_{cp}$.

c Als je dezelfde criteria hanteert als in het voorbeeld, is er dan een gering, middelmatig of groot verschil tussen NG- en NT-leerlingen?

Theorie en voorbeelden

Om te onthouden

In statistisch onderzoek worden vaak twee groepen onderzocht en met elkaar vergeleken.

Als er vragen worden gesteld met twee mogelijke antwoorden (kwalitatieve, nominale variabele), kunnen percentages of aantallen voor deze vergelijking worden gebruikt.

- **Percentages vergelijken**.
  Gebruik de vuistregels:
  - Als het verschil in percentage groter of gelijk is aan 40% is de conclusie:
    Het verschil is groot.
  - Als het verschil in percentage tussen 20% en 40% is, is de conclusie:
    Het verschil is middelmatig.
  - Als het verschil in percentage kleiner of gelijk is aan 20% is de conclusie:
    Het verschil is gering.
- **Aantallen vergelijken**:
  Maak een 2×2 **kruistabel**.
  Bereken *phi* $= \varphi$:
  $\varphi = \frac{ad-bc}{\sqrt{(a+b)\cdot(a+c)\cdot(b+d)\cdot(c+d)}}$
  Conclusies trek je met behulp van de vuistregels op de **Formulekaart**.

aantal	groep 1	groep 2	
wel	a	b	$a+b$
niet	c	d	$c+d$
	$a+c$	$b+d$	

Tabel 2.4

Als er vragen worden gesteld met meerdere mogelijke antwoorden die kunnen worden geordend (kwalitatieve, ordinale variabele), kunnen **cumulatieve percentages** worden gebruikt. Dan wordt het maximale verschil van de cumulatieve percentages, $\max V_{cp}$, berekend. De bijbehorende vuistregels staan op de **Formulekaart**.

Voorbeeld 1

Bekijk in de tabel de keuze voor wiskunde A voor jongens en meisjes in 4 havo.

	jongens	meisjes
wiskunde A	27	32
geen wiskunde A	33	25

Tabel 2.5

Is er een gering, een middelmatig of groot verschil tussen jongens en meisjes op deze school voor wat betreft de keuze voor wiskunde A? Gebruik percentages en de vuistregels:

- Als het verschil in percentage groter of gelijk is aan 40% is de conclusie:
  Het verschil is groot.
- Als het verschil in percentage tussen 20% en 40% is, is de conclusie:
  Het verschil is middelmatig.
- Als het verschil in percentage kleiner of gelijk is aan 20% is de conclusie:
  Het verschil is gering.

Antwoord

Er is sprake van een nominale variabele (variabele is wel of geen wiskunde A).
Je kijkt nu naar het verschil in procenten tussen jongens en meisjes.

	percentage jongens	percentage meisjes
wiskunde A	45%	56,1%

Tabel 2.6

Reken de aantallen om naar percentages. Het totaal aantal jongens is 60, het totaal aantal meisjes is 57.

Het verschil in procenten tussen jongens en meisjes die wiskunde A kiezen, is $56{,}1 - 45 = 11{,}1\%$.

Dit verschil in procenten is $\leq 20\%$. Je spreekt dus van een gering verschil.

Opgave 7

Gebruik de gegevens uit **Voorbeeld 1**.

a Bereken zelf de percentages die in de tabel staan.

b Bereken het procentuele verschil tussen jongens en meisjes die geen wiskunde A kiezen. Wat is dan je conclusie?

Opgave 8

Bekijk de tabel met de keuze voor Frans voor jongens en meisjes in 4 havo.

	jongens	meisjes
Frans	22	40
geen Frans	38	17

Tabel 2.7

Is er een gering, een middelmatig of groot verschil tussen jongens en meisjes op deze school voor wat betreft de keuze voor Frans? Gebruik de percentages en de vuistregels uit het voorbeeld.

Voorbeeld 2

Er wordt veel onderzoek gedaan naar bijwerkingen van medicijnen. Onderzoek onder mannen en vrouwen naar de bijwerkingen gaf de volgende resultaten:

	bijwerking	geen bijwerking
mannen	86	14
vrouwen	61	39

Tabel 2.8

Bereken *phi*= φ en trek een conclusie met behulp van de vuistregels op de **Formulekaart**.

Antwoord

phi berekenen geeft $\varphi \approx 0{,}28$.

Volgens de vuistregels voor vergelijken van groepen is dus sprake van een middelmatig verschil.

Opgave 9

Gebruik de gegevens uit **Voorbeeld 2**.

Bereken zelf φ en controleer de conclusie.

Opgave 10

Dit is het resultaat van een statistisch onderzoek.

	slagen	zakken
gemotiveerd	70	10
ongemotiveerd	10	30

Tabel 2.9

a Bereken φ.

b Wat is je conclusie over de mate van verschil tussen gemotiveerde en ongemotiveerde leerlingen?

c Trek ook een conclusie door percentages te gebruiken.

d Hoe groot zou φ zijn geweest als alle gemotiveerde leerlingen zouden zijn geslaagd en alle ongemotiveerde leerlingen gezakt?

Voorbeeld 3

Hotelgasten is gevraagd het dinerbuffet te beoordelen. In de tabel staan de resultaten uitgesplitst naar werelddeel van herkomst.

	Europa	Afrika	Azië	Noord-Amerika	Zuid-Amerika	Oceanië
1 = zeer goed	12	4	6	3	4	0
2 = goed	15	6	22	6	8	4
3 = matig	10	2	15	4	2	2
4 = slecht	3	0	4	4	0	2

Tabel 2.10

Op grond van deze tabel lijkt het alsof de Europeanen het dinerbuffet hoger hebben gewaardeerd dan de Aziaten, maar die groepen zijn niet even groot.

Is er een gering, een middelmatig of groot verschil tussen de Europeanen en de Aziaten en hun oordeel over het dinerbuffet? Gebruik percentages en de vuistregels op de **Formulekaart**.

Antwoord

Om goed te kunnen vergelijken bij groepen die niet even groot zijn, moeten eerst de absolute aantallen worden omgezet naar percentages.

Bekijk in de tabel onder p de percentages, onder cp de cumulatieve percentages en onder V_{cp} het verschil van cp.

	Europa			Azië			
	aantal	p	cp	aantal	p	cp	V_{cp}
1 = zeer goed	12	30	30	6	12,8	12,8	17,2
2 = goed	15	37,5	67,5	22	46,8	59,6	7,9
3 = matig	10	25	92,5	15	31,9	91,5	1
4 = slecht	3	7,5	100	4	8,5	100	0
	40			47			

Tabel 2.11

Er geldt $\max V_{cp} \leq 20\%$, dus het verschil is gering.

Opgave 11

Gebruik de gegevens uit **Voorbeeld 3**.
Is er een gering, een middelmatig of groot verschil tussen de Afrikanen en de Oceaniërs en hun oordeel over het dinerbuffet?

Verwerken

Opgave 12

Het verschil bij een onderzoek onder 200 patiënten naar het gebruik van twee medicijnen is 14%.
Is er een gering, een middelmatig of groot verschil?

A. gering verschil
B. middelmatig verschil
C. groot verschil

Opgave 13

Er is veel onderzoek gedaan naar de effecten van roken. Gekeken is bijvoorbeeld naar aandoeningen aan hart- en bloedvaten bij rokers en niet-rokers met het volgende resultaat:

	aandoening	geen aandoening
roker	86	14
niet-rokers	61	39

Tabel 2.12

a Gaat het hier om kwalitatieve of kwantitatieve variabelen?

b Gaat het om nominale of ordinale variabelen?

c Bereken het verschil in procenten tussen rokers en niet-rokers als er wordt gelet op wel of geen aandoeningen aan hart- en bloedvaten.

d Is het verschil tussen roken en niet roken gering, middelmatig of groot volgens de vuistregels?

e Welke kritische vragen heb je bij dit onderzoek?

f Zegt dit onderzoek iets over de oorzaak van de aandoeningen aan hart- en bloedvaten?

Opgave 14

Dit is het resultaat van een onderzoek door de Rijksoverheid naar de uitstroom van studenten uit het voltijd HBO per studierichting. De getallen zijn percentages.

studierichting	werk en/of uitkering	geen werk en/of uitkering
Onderwijs	87	13
Landbouw en natuur	76	24
Techniek	77	23
Gezondheidszorg	86	15
Economie	77	23
Gedrag en maatschappij	83	17
Taal en Cultuur	62	38

Tabel 2.13

a Welke studierichting heeft het laagste percentage uitstroom van studenten met werk en/of uitkering?

b Voor welke studierichtingen geldt: er is een middelmatig verschil met uitstroom van studenten van de richting Taal en Cultuur, als gelet wordt op het percentage werk en/of uitkering?

Opgave 15

Er is veel onderzoek naar bijwerkingen van medicijnen. Onderzoek onder gebruikers van aspirines gaf de volgende 2×2 tabel.

	hartfalen	geen hartfalen
aspirine gebruikers	139	10898
geen aspirine gebruikers	239	10795

Tabel 2.14

Je wilt weten welk verschil er is tussen gebruikers van aspirine en niet-gebruikers. De grootheid φ is een maat voor dit verschil.

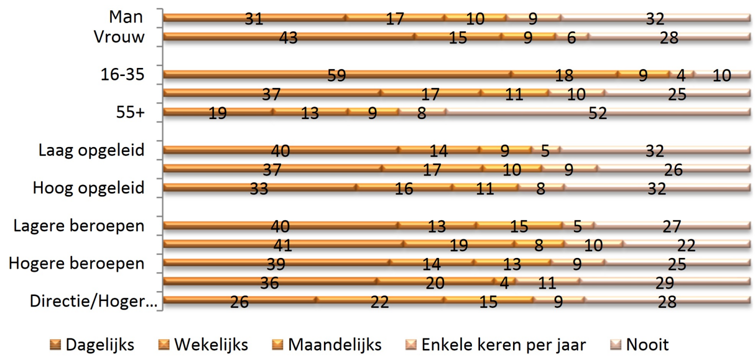
a Bereken φ.

b Welke conclusie trek je over de samenhang tussen gebruik van aspirine en hartfalen? Gebruik de **Formulekaart**.

c Reken de gegevens in de tabel om naar procenten.
Welke conclusie kun je nu trekken?

Opgave 16

In een rapport van Universiteit Twente staat het volgende overzicht over het bezoek aan sociale netwerksites.



Figuur 2.3

Bekijk of er een verschil is tussen mannen en vrouwen. Gebruik de cijfers uit de figuur.

a Vul de percentages voor mannen en vrouwen in de tabel in.

cp is de afkorting voor cumulatief percentage.

	mannen	cp mannen	vrouwen	cp vrouwen
nooit				
enkele keren per ja				
maandelijks				
wekelijks				
dagelijks				

Tabel 2.15

b Gebruik de vuistregels op de **Formulekaart**. Is er een gering, middelmatig of groot verschil tussen mannen en vrouwen? Licht je antwoord toe.

Opgave 17

Een marketingmanager heeft de volgende resultaten van een onderzoek naar de werking van een medicijn ontvangen. Welke conclusie zal hij eruit trekken?

aantal	positieve reactie	geen reactie
medicijn niet gebruikt	161	739
medicijn wel gebruikt	1796	1306

Tabel 2.16

Toepassen

Opgave 18: Intelligentie van hondenrassen

Voor een onderzoek naar de intelligentie van hondenrassen is aan hondenbezitters gevraagd hoelang het duurde voor de hond naar zijn naam luisterde vanaf het moment dat ze hem als pup kregen. De resultaten van labradors en beagles staan in de tabel.

aantal	labrador	beagle
luistert naar naam binnen een week	55	39
luistert naar naam na meer dan een week	32	27

Tabel 2.17

a Op welke manier zou je hier een conclusie uit kunnen trekken?

b Trek aan de hand van de **bijbehorende vuistregels** en de tabel een conclusie over het verschil in intelligentie tussen een labrador en een beagle.

Testen

Opgave 19

Dit is het resultaat van een statistisch onderzoek.

	huid verbrandt snel	huid verbrandt niet snel
rood haar	55	8
bruin haar	87	102

Tabel 2.18

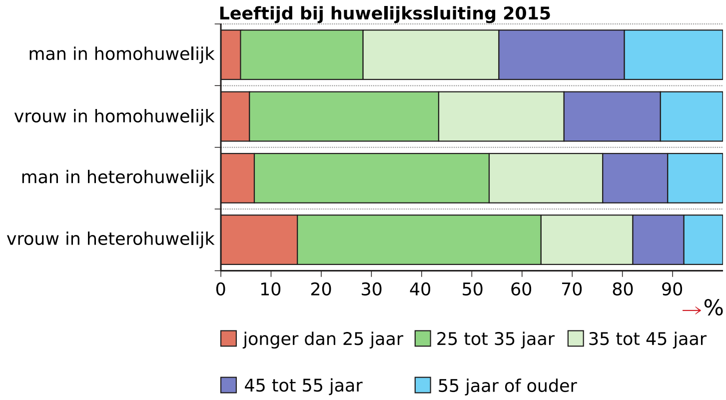
a Trek op grond van deze gegevens een conclusie over het verschil tussen roodharigen en bruinharigen met betrekking tot het verbranden van de huid door de zon. Gebruik de vuistregels van de **Formulekaart**.

b Trek op grond van percentages een conclusie over het verschil tussen roodharigen en bruinharigen met betrekking tot het verbranden van de huid door de zon. Gebruik de vuistregels:

- Als het verschil in percentage groter of gelijk is aan 40% is de conclusie:
  Het verschil is groot.
- Als het verschil in percentage tussen 20% en 40% is, is de conclusie:
  Het verschil is middelmatig.
- Als het verschil in percentage kleiner of gelijk is aan 20% is de conclusie:
  Het verschil is gering.

Opgave 20

Bekijk de leeftijd bij huwelijkssluiting in 2015, gegroepeerd naar soort huwelijk.



Figuur 2.4

Maak een cumulatieve tabel en trek op grond van deze tabel een conclusie over het verschil in huwelijksleeftijd tussen mannen in homohuwelijk en mannen in heterohuwelijk.

2.3 Verschil kwantitatieve variabelen

Inleiding

Hoe kun je bijvoorbeeld de levensduur van twee verschillende typen batterijen met elkaar vergelijken? Je neemt dan steekproeven. Maar hoe kun je die dan weer vergelijken?

Ook bij kwantitatieve variabelen horen een aantal manieren waarop je verschillen tussen statistische variabelen in kaart kunt brengen.

Figuur 3.1

Je leert in dit onderwerp

- kwantitatieve variabelen vergelijken met behulp van boxplots;
- bij kwantitatieve variabelen de effectgrootte van een bepaalde handeling meten.

Voorkennis

- soorten statistische variabelen herkennen;
- de begrippen onderzoek, steekproef, populatie en representatief, simulatie;
- meetniveaus onderscheiden bij antwoordmogelijkheden op vragen.

Verkennen

Opgave V1

Van twee types batterijen wordt de levensduur (in uren) vergeleken. Van beide types worden 15 batterijen onderzocht. In de tabel zie je de resultaten.

Type I	560	625	580	605	598	602	602	613	650	583	588	595	601	623	589
Type II	630	620	595	590	635	660	610	654	632	680	624	590	643	625	671

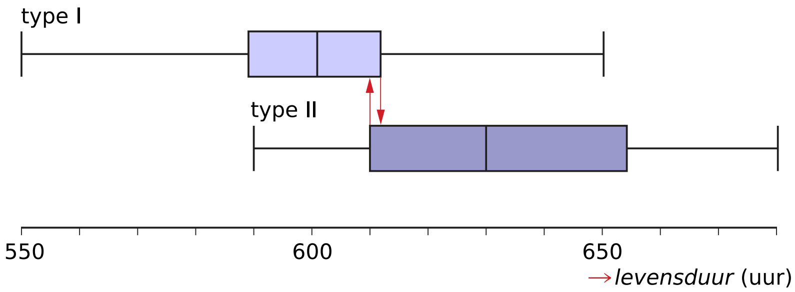
Tabel 3.1

Probeer een manier te verzinnen om deze twee steekproeven te vergelijken.

Uitleg 1

Statistisch onderzoek wordt vaak gebruikt om variabelen te vergelijken. Er worden dan uitspraken gedaan als: “Het verschil in levensduur van deze twee typen batterijen is gering.” Hoe kunnen dit soort uitspraken gedaan worden?

Dit kan bijvoorbeeld op de volgende manier. Van twee typen batterijen wordt de levensduur in uren onderzocht. Van beide typen worden 15 batterijen onderzocht. Bij het onderzoek horen de volgende boxplots:



Figuur 3.2

De ‘box’ is het interval vanaf het eerste kwartiel tot en met het derde kwartiel. Om uit deze boxplots conclusies te trekken, zijn vuistregels afgesproken:

- Als de boxen elkaar niet overlappen, dan is het verschil groot.
- Als de boxen elkaar wel overlappen en minstens één mediaan buiten de box van de andere boxplot ligt, dan is het verschil middelmatig.
- In alle andere gevallen is het verschil gering.

De vuistregels staan op de **Formulekaart**.

Opgave 1

Gegeven zijn de boxplots uit **Uitleg 1**.

a Schrijf de waarde van de mediaan van de levensduur van de batterijen van type I op.

b Schrijf het eerste en derde kwartiel van de levensduur van de batterijen van type II op.

c De boxen overlappen elkaar. Schrijf de conclusie van de vergelijking van de levensduur van de verschillende typen batterijen op.

Opgave 2

De volgende gegevens hebben betrekking op het aantal uur sport per week bij 73 jongens en 102 meisjes.

mediaan jongens $\approx 4{,}0$ uur	mediaan meisjes $\approx 2{,}0$ uur
interkwartielafstand $\approx 4{,}2$ uur	interkwartielafstand $\approx 4{,}0$ uur
eerste kwartiel $\approx 1{,}8$ uur	derde kwartiel $\approx 4{,}3$ uur

Tabel 3.2

a Bereken de ontbrekende kwartielen.

b Teken de boxplots voor zover mogelijk. Teken ze boven elkaar.

c Welke conclusie kun je trekken ten aanzien van het verschil in het aantal uur sport per week van jongens en meisjes?

Uitleg 2

Van een bepaald type batterijen wordt het productieproces aangepast om de levensduur (uur) te verlengen. Er worden 15 batterijen van het oude productieproces vergeleken met 15 batterijen die op de nieuwe manier zijn geproduceerd. In de tabel staan de resultaten. L_I stelt de levensduur voor van batterijen die volgens het oude productieproces zijn gemaakt, L_{II} is de levensduur van een batterij in het nieuwe proces.

L_I (uur)	560	625	580	605	598	602	602	613	650	583	588	595	601	623	589
L_{II} (uur)	630	620	595	590	635	660	610	654	632	680	624	590	643	625	671

Tabel 3.3

Als het verschil tussen de gemiddelden in beide steekproeven erg groot is, is het verschil in levensduur dan ook erg groot? Als de bijbehorende standaardafwijkingen groot zijn, hoeft dat niet zo te zijn.

Uit de gegevens volgt:

Het gemiddelde $\overline{L}_I = 600{,}9$.
De standaardafwijking $S_I = 20{,}7$.
Het gemiddelde $\overline{L}_{II} = 630{,}6$.
De standaardafwijking $S_{II} = 26{,}9$.

Bereken nu de zogenaamde effectgrootte:

$E = \frac{\text{grootste gemiddelde} - \text{kleinste gemiddelde}}{\text{gemiddelde van de standaardafwijkingen}}$

Dus $E = \frac{630{,}6-600{,}9}{\frac{1}{2}(20{,}7+26{,}9)}$.

Er bestaan vuistregels om een conclusie te trekken:

- Als $E > 0{,}8$ dan is het verschil tussen de variabelen groot.
- Als $0{,}4 < E \leq 0{,}8$ dan is het verschil middelmatig.
- Als $E \leq 0{,}4$ dan is het verschil gering.

De formule voor de effectgrootte en de vuistregels staan op de **Formulekaart**.

Opgave 3

Gebruik de gegevens uit **Uitleg 2**.

a Bereken de gemiddelde levensduur van de twee typen batterijen.

b Bereken de standaardafwijkingen van de twee typen batterijen.

c Bereken de effectgrootte met de formule uit de uitleg.

d Is het verschil in levensduur tussen de twee typen batterijen gering, middelmatig of groot?

Opgave 4

Bekijk de formule voor de effectgrootte in **Uitleg 2**.

a Wat verandert er aan E als de gemiddelden worden verwisseld?

b De gemiddelden zijn verwisseld. Geef een voorbeeld van een waarde van E waarbij een foute conclusie zou worden getrokken.

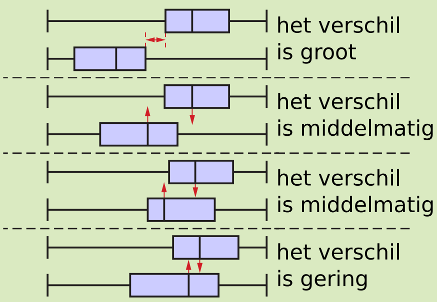
c Er worden twee variabelen vergeleken. Het komt regelmatig voor dat de standaardafwijkingen dan gelijk zijn. De formule voor E wordt dan eenvoudiger. Schrijf zo'n eenvoudige formule op.

Theorie en voorbeelden

Om te onthouden

Bij statistisch onderzoek wordt regelmatig het verschil tussen twee populaties onderzocht. Hier wordt dat met kwantitatieve variabelen gedaan. Dit kan op veel manieren.
Twee manieren zijn: met boxplots of met effectgrootte. Uiteraard moet er van elke populatie een steekproef worden genomen. Deze gegevens vormen de basis voor uitspraken over het verschil tussen de populaties.

Twee **boxplots vergelijken** gaat als volgt:



Figuur 3.3

Bekijk de applet.

De ‘box’ is het interval vanaf het eerste kwartiel tot en met het derde kwartiel. Conclusies trek je met de vuistregels op de **Formulekaart**.

Met twee gemiddelden en twee standaardafwijkingen gebruik je de **effectgrootte** $E = \frac{\text{grootste gemiddelde - kleinste gemiddelde}}{\text{gemiddelde van de standaardafwijkingen}} = \frac{\overline{X}_1 - \overline{X}_2}{\frac{1}{2}(S_1 + S_2)}$

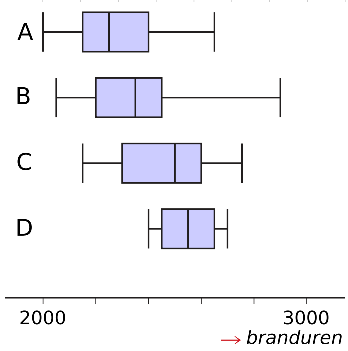
Let erop dat X_1 groter moet zijn dan X_2

Trek ook nu de conclusie weer met behulp van de vuistregels op de **Formulekaart**.

Voorbeeld 1

Bij lampen wordt onderzoek gedaan naar het aantal branduren. Bekijk de boxplots van de branduren van lampen. Deze laten het resultaat van steekproeven van vier typen lampen zien.

Vergelijk de branduren van type A met de andere typen. Trek de conclusies met behulp van de vuistregels op de **Formulekaart**.



Figuur 3.4

Antwoord

Voor het verschil in brandduur geldt:

- Het verschil tussen A en B is gering, want de boxen overlappen en er ligt geen mediaan buiten de andere box.
- Het verschil tussen A en C is middelmatig, want de boxen overlappen en een mediaan (zelfs beide) ligt buiten de andere boxen.
- Het verschil tussen A en D is groot, want de boxen overlappen niet.

Opgave 5

Gebruik de gegevens uit **Voorbeeld 1**. Trek alle overige mogelijke conclusies met de regels van de formulekaart.

Opgave 6

Gebruik de gegevens uit **Voorbeeld 1**.

a Naar een vijfde type, type E, is ook onderzoek gedaan. In de steekproef zaten alleen lampen met een levensduur van meer dan 3000 uur. Waarom weet je nog steeds niet 100% zeker dat elke lamp van type E langer brandt dan de lampen van type A?

b Als de steekproefomvang groter wordt gemaakt, welke invloed heeft dat dan op de conclusie?

c Hoeveel procent van de lampen van type A gaat langer mee dan de lamp van type D met de kortste brandtijd?

Voorbeeld 2

Om het effect van het taalonderwijs te onderzoeken is van twee even grote groepen Nederlanders en Belgen gekeken naar het aantal vreemde talen dat ze spreken. Bekijk de tabel met resultaten.

aantal gesproken vreemde talen	0	1	2	3	4	totaal
Belgen	122	168	184	103	34	611
Nederlanders	19	156	272	146	18	611
totaal	141	324	456	249	52	1222

Tabel 3.4

Welk verschil is er, statistisch gezien, tussen het aantal gesproken talen van Belgen en Nederlanders?

Antwoord

Het aantal talen is een kwantitatieve variabele. Er zijn twee bekende methodes om een vergelijking uit te voeren: boxplots vergelijken of effectgrootte berekenen en conclusies trekken. Boxplots zijn hier erg onnauwkeurig, dus effectgrootte blijft over.

De effectgrootte is: $E = \frac{\overline{M}_1 - \overline{M}_2}{\frac{1}{2} \cdot (S_1 + S_2)}$.

Voor de Belgen geldt: $\overline{M} \approx 1{,}606$ en $S \approx 1{,}144$.
Voor de Nederlanders geldt: $\overline{M} \approx 1{,}980$ en $S \approx 0{,}858$.

Omdat de *effectgrootte* een positief getal moet zijn, neem je: $\overline{M}_1 = 1{,}980$ en $\overline{M}_2 = 1{,}606$.

Dan geldt: $E = \frac{1{,}980 - 1{,}606}{\frac{1}{2} \cdot (0{,}858 + 1{,}144)} = \frac{0{,}374}{1{,}001} \approx 0{,}37$.

De effectgrootte is kleiner dan $0{,}4$. Dus het verschil is volgens de vuistregels op de **Formulekaart** gering.

Opgave 7

Gebruik de gegevens uit **Voorbeeld 2**.

a Welke gegevens uit de tabel worden niet gebruikt?

b Bereken zelf de gemiddelden en de standaardafwijkingen. Controleer of de berekende waarden in het voorbeeld juist zijn.

c Leg uit waarom het gebruik van een boxplot hier onnauwkeurig is.

Opgave 8

De volgende gegevens hebben betrekking op het aantal uur sport per week bij jongens en meisjes.

gemiddelde jongens $5{,}0$ uur	gemiddelde meisjes $3{,}4$ uur
standaardafwijking jongens $4{,}1$ uur	standaardafwijking meisjes $3{,}7$ uur
mediaan jongens ≈ 5 uur	mediaan meisjes ≈ 2 uur
interkwartielafstand ≈ 4 uur	interkwartielafstand ≈ 4 uur

Tabel 3.5

a Bereken met de formule de effectgrootte voor het aantal uur sport bij jongens en meisjes.

b Welke conclusie trek je over het verschil tussen jongens en meisjes?

c Waarom lukt het met deze gegevens niet om een statistische vergelijking met boxplots uit te voeren?

Verwerken

Opgave 9

De effectgrootte in de vergelijking van het aantal uur tv kijken per week van werkenden en niet-werkenden is $0{,}75$. Hoe beoordeel je dit verschil volgens de vuistregels op de **Formulekaart**?

A. gering verschil
B. middelmatig verschil
C. groot verschil

Opgave 10

Cito meet kennis en vaardigheden in de vorm van toetsen. Cito heeft onderzocht of er een verschil in vaardigheden is tussen vwo- en gymnasiumleerlingen. Cito heeft dit gedaan voor de onderwerpen Nederlands leesvaardigheid, Nederlands woordenschat, Engels leesvaardigheid, rekenen, wiskunde en taalverzorging. Bekijk de tabel.

leerling	vak	aantal leerlingen	aantal scholen	gemiddelde	standaard-afwijking	effect-grootte
gym	nlv	1654	87	284,24	15,77	0,55
gym	nws	1598	84	296,71	22,15	0,58
gym	elv	1605	84	297,53	29,65	0,48
gym	rek	1873	87	293,12	24,45	0,6
gym	wis	1872	87	234,93	13,77	0,62
gym	tv	1566	84	302,10	29,53	1,16
vwo	nlv	7400	170	275,68	15,62	
vwo	nws	6961	165	284,53	20,61	
vwo	elv	6955	160	284,32	26,80	
vwo	rek	7148	163	279,31	22,74	
vwo	wis	7145	163	227,15	12,08	
vwo	tv	6978	165	275,32	21,49	

Figuur 3.5

a Schrijf de statistische variabelen uit dit onderzoek op en geef aan of ze kwalitatief of kwantitatief zijn.

b In de laatste kolom staat de effectgrootte. Bereken zelf de effectgrootte voor Nederlands leesvaardigheid.

c Schrijf de conclusie op voor dit verschil.

d Controleer de effectgrootte voor taalverzorging. Wat valt je op? Verandert de conclusie?

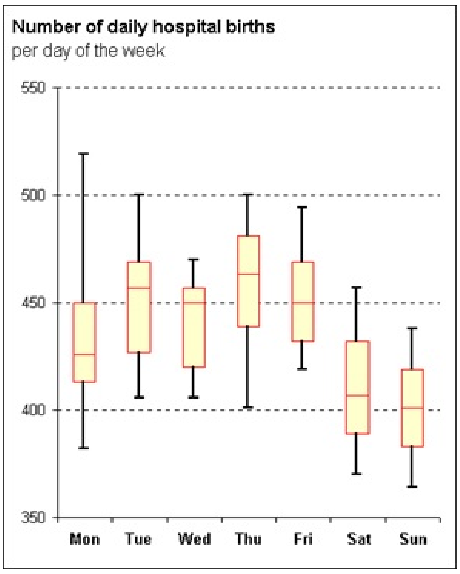
Opgave 11

Er worden veel statistieken bijgehouden over geboortes. In een land is een jaar lang het aantal geboortes per weekdag in alle ziekenhuizen bijgehouden. Bekijk de boxplots bij dit onderzoek.

a Leg uit dat je uit deze boxplots niet kunt concluderen dat er in dit land op zondag altijd minder baby's in een ziekenhuis worden geboren dan op elke andere dag.

b Het verschil tussen het aantal geboortes op zondag en op andere dagen is statistisch gezien voor sommige dagen groot. Voor welke dagen?

c Welke dagen verschillen middelmatig met donderdag?



Figuur 3.6

Opgave 12

Bij een statistisch onderzoek worden van twee groepen ondernemers drie variabelen onderzocht. Geef bij elk van deze variabelen aan welke van de volgende methodes kunnen worden gebruikt om de twee groepen te vergelijken:
A: verschil in percentages vergelijken
B: phi berekenen
C: boxplots vergelijken
D: effectgrootte berekenen

a bedrijfsomzet

b mate van geluk

c geslacht

Opgave 13

Een onderzoekster onderzoekt twee populaties. De ene populatie heeft geen behandeling ondergaan. Dit is de referentiepopulatie. De andere populatie heeft wel een behandeling ondergaan. Dit is de onderzoekspopulatie.
De onderzoekster heeft de volgende gegevens.

- Voor de referentiepopulatie geldt:
  Het gemiddelde van de variabele die ze onderzoekt is 210 gram, de standaardafwijking is 11 gram.
- Voor de onderzoekspopulatie geldt:
  De onderzoekster wil weten wat het gemiddelde van de variabele die ze onderzoekt moet zijn, zodat haar conclusie kan zijn: het verschil tussen de variabelen is groot. Ga ervan uit dat de standaardafwijking van deze variabele ook 11 gram is.

Voor welke gemiddelde waarden kan ze die conclusie trekken?

Opgave 14

Om twee groepen te vergelijken, kan soms de effectgrootte worden gebruikt.
De effectgrootte kan worden berekend met de formule: $E = \frac{\overline{X}_1 - \overline{X}_2}{\frac{1}{2} \cdot (S_1 + S_2)}$,
waarbij $\overline{X}_1$, S_1, $\overline{X}_2$, S_2 het gemiddelde en de standaardafwijking van de variabele in de ene groep respectievelijk de andere groep zijn. Geef aan wat er in de volgende situaties met de effectgrootte gebeurt. Kies uit: ‘wordt groter’, ‘verandert niet’, ‘wordt kleiner’ of ‘dat hangt ervan af’.

a Alleen $\overline{X}_1$ verandert en wordt groter.

b Alleen $\overline{X}_2$ verandert en wordt groter.

c Alleen S_2 verandert en wordt groter.

Toepassen

Opgave 15: Computertijd

aantal uur computer per week		
	jongen	meisje
Aantal waarnemingen	24472	25599
Gemiddelde	14,8	13,7
Mediaan	12,0	11
Modus	10	10
Minimum	0	0
Maximum	70	70
Standaardafwijking	10,60	10,24
VARn	112,42	104,89
Eerste kwartiel	7,0	7,0
Derde kwartiel	20,0	19,0
Kwartielafstand	13,0	12,0

Tabel 3.6

Deze tabel bevat informatie over het aantal uur dat jongens respectievelijk meisjes per week voor een computer zitten.

a Onderzoek met behulp van de effectgrootte hoe groot het verschil is tussen jongens en meisjes in het aantal uur dat zij per week voor een computer zitten.

b Onderzoek hetzelfde verschil met behulp van een globale vergelijking van de boxplots.

c Onderzoek dit verschil ook met behulp van de formule voor de grenzen van de intervallen: $med \pm 1{,}5 \cdot \frac{IQR}{\sqrt{n}}$.

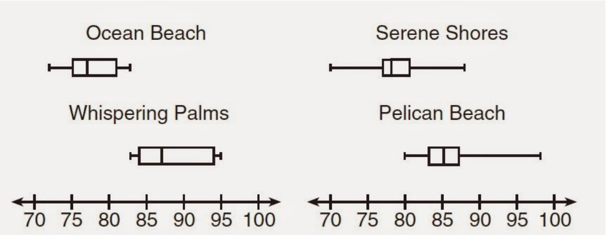
med is de mediaan, IQR is de interkwartielafstand, n is de steekproefomvang.

Vuistregel: als de intervallen niet overlappen dan is er verschil, bij overlap is er geen verschil.

Testen

Opgave 16

Vier stranden zijn vergeleken voor een strandvakantie in juli. De hoogste temperaturen op de stranden zijn dagelijks gemeten en in een boxplot verwerkt. De temperaturen staan in graden Fahrenheit. 95 graden Fahrenheit is 35 graden Celsius en 70 graden Fahrenheit is $21{,}11$ graden Celsius.



Figuur 3.7

Bron: **mrburkemath.blogspot.nl**

a Vergelijk de vier boxplots. Doe een uitspraak over elke van de vier bestemmingen en gebruik de begrippen mediaan, minimale en maximale temperatuur.

b Is er veel verschil tussen Ocean Beach en Serene Shores? En tussen Ocean Beach en Pelican Beach?

c Je zoekt een bestemming waar het vaak warm weer is en waar de kans op kil weer niet zo groot is. Welke van deze vier stranden voldoet daar het best aan? Motiveer je antwoord.

Opgave 17

Een bepaald type ledlamp blijkt volgens een onderzoek uit 2015 door de fabrikant gemiddeld 34300 uur mee te gaan met een bijbehorende standaardafwijking van 1200 uur.

Inmiddels heeft de fabrikant het productieproces van deze ledlampen verder verbeterd. Uit een nieuw onderzoek blijken deze ledlampen gemiddeld 35200 uur mee te gaan met een standaardafwijking van 1200 uur.

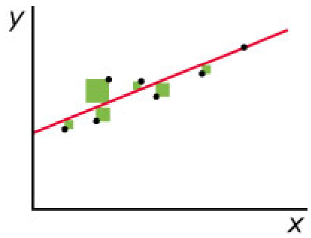
a Waarom kun je nu zeggen dat het effect op de levensduur van deze ledlampen door het verbeteren van het fabricageproces niet groot is geweest?

b Bij welk gemiddelde in het nieuwe onderzoek was het effect wel groot geweest?

2.4 Samenhang van variabelen

Inleiding

Behalve het onderzoeken van verschillen tussen statistische variabelen is het onderzoeken naar verbanden een belangrijke tak van sport: wanneer bestaat er een verband tussen twee statistische variabelen? Bestaat er bijvoorbeeld een verband tussen het aantal overvliegende ooievaars en het aantal geboorten in een bepaalde streek? Of bestaat er een verband tussen lengte en gewicht bij scholieren?



Figuur 4.1

Je leert in dit onderwerp

- de (statistische) samenhang tussen twee variabelen te beschrijven en te interpreteren;
- te werken met correlatiecoëfficiënt en regressielijn (trendlijn).

Voorkennis

- soorten statistische variabelen herkennen;
- de begrippen onderzoek, steekproef, populatie en representatief, simulatie.

Verkennen

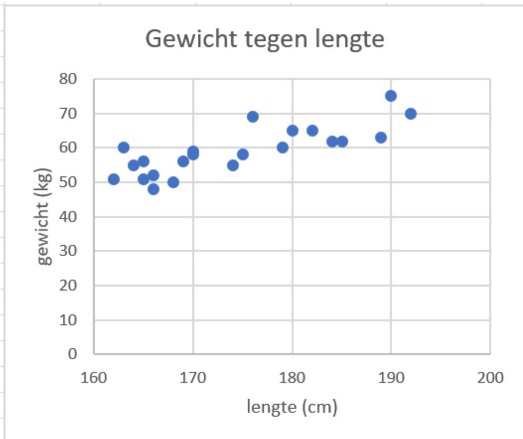
Opgave V1

Om te onderzoeken of er een verband bestaat tussen lengte en gewicht bij mensen van 15 tot 17 jaar oud heb je gegevens nodig. Op het werkblad **LengteGewicht22H4.xls** vind je de gegevens van een 4HAVO-klas van 22 leerlingen.

a Welke gegevens zijn er verzameld?

b Welke afspraken moet je maken bij het verzamelen van deze gegevens? Beschrijf er een paar. (Denk om de manier van meten!)

c Bekijk het getekende spreidingsdiagram. Trek je op grond van de gegevens op het werkblad de conclusie dat er zo'n verband bestaat? En is dat dan uitsluitend een statistisch verband of is het ook een oorzakelijk verband, m.a.w. wordt een groter gewicht veroorzaakt door een grotere lengte?

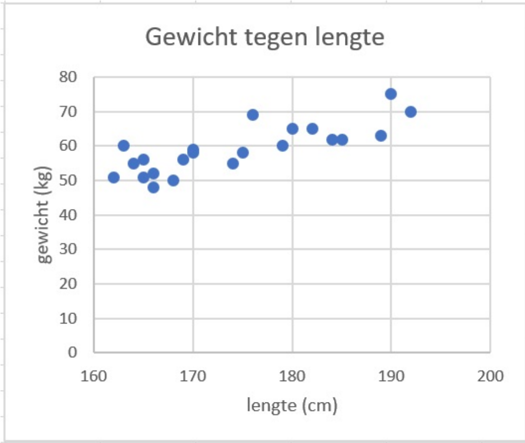


Figuur 4.2

Uitleg 1

Je wilt onderzoeken of er een verband bestaat tussen lengte en gewicht bij mensen van 15 tot 17 jaar oud. Op het werkblad **LengteGewicht22H4.xls** vind je de gegevens van een steekproef van 22 leerlingen. Hiernaast is een spreidingsdiagram van die gegevens getekend.
Is er binnen deze groep sprake van een verband tussen lengte en gewicht?

Als je de figuur bekijkt, zie je dat bij grotere lengtes vaak ook grotere gewichten horen.
Er lijkt dus een zeker verband te zijn.
Maar de getekende punten liggen zeker niet op één lijn, dus hoe sterk is het verband?



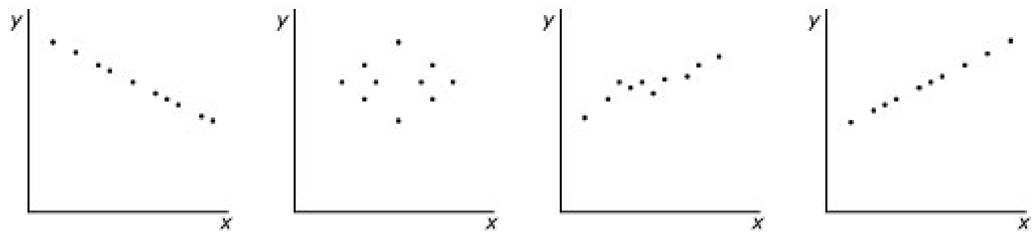
Figuur 4.3

De vorm van de puntenwolk, het spreidingsdiagram, zegt iets over de samenhang tussen twee variabelen. Zie daarvoor de afbeeldingen hieronder.

Een ander woord voor statistisch verband of samenhang is correlatie.

De mate van correlatie wordt uitgedrukt in de correlatiecoëfficiënt r.

Correlatie zegt iets over de statistische samenhang van twee variabelen en niets over een eventuele causale relatie (dus een oorzaak-gevolgrelatie) tussen de twee variabelen. Er is samenhang tussen de verkoop van handschoenen en de verkoop van winterbanden, maar de oorzaak is dan het mogelijke koude winterweer.



$r_{xy} = -1$ $r_{xy} = 0$ $r_{xy} = 0{,}6$ $r_{xy} = 1$

Figuur 4.4

Opgave 1

Gebruik het werkblad **LengteGewicht22H4.xls** in **Uitleg 1**.

a Maak zelf zo'n spreidingsdiagram.

b Hoe zie je in het spreidingsdiagram dat er een zeker verband tussen l en G bestaat?

c Bereken met behulp van Excel de correlatiecoëfficiënt r_{lG} in twee decimalen nauwkeurig, bekijk het **Practicum**.

Je hebt de waarde van r_{lG} berekend. Maar hoe weet je nu of deze correlatie genoeg is om vast te stellen dat er een verband is? Daarvoor zijn vuistregels zoals: Als $r \geq 0{,}7$ dan is er sprake van sterke positieve samenhang.

d Kun je vaststellen dat er een duidelijk verband is tussen lengte en gewicht van deze jongeren?

Opgave 2

Je hebt in de voorgaande opgave een verband aangetoond tussen lengte en gewicht bij mensen van 15 tot 17 jaar oud.
Daarbij gebruikte je een steekproef van 22 studenten.

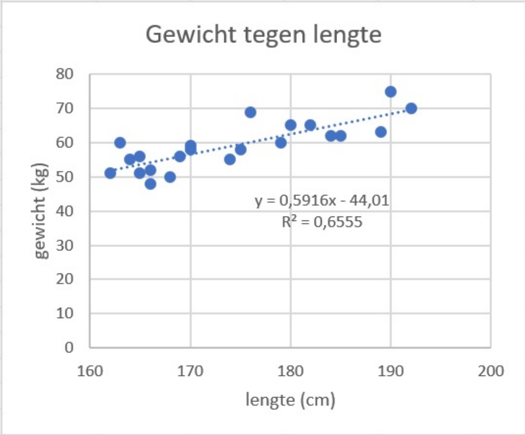
a Geef commentaar op deze steekproef.

Stel je voor dat je met een aselecte steekproef te maken hebt. En neem ook aan dat de groep representatief is voor studenten van 15 tot 17 jaar oud. Er is een duidelijk verband gevonden.

b Is er nu sprake van een oorzakelijk verband, dus is de lengte oorzaak van het gewicht?
Of is er alleen een statistisch verband? Licht het antwoord toe.

Uitleg 2

Je wilt onderzoeken of er een verband bestaat tussen lengte en gewicht bij mensen van 15 tot 17 jaar oud. Op het werkblad **LengteGewicht22H4.xlsx** vind je de gegevens van een steekproef van 22 studenten. Hiernaast is een spreidingsdiagram van die gegevens getekend.
Er is binnen deze groep studenten sprake van een verband tussen lengte en gewicht, maar hoe beschrijf je dit verband?



Figuur 4.5

Je ziet in de figuur hoe Excel automatisch het kwadraat van de correlatiecoëfficiënt r berekent en een best passende lijn door de puntenwolk trekt. Deze lijn heet de regressielijn of trendlijn. En r^2 heet de determinatiecoëfficiënt.

Bij de trendlijn kun je ook zelf een formule opstellen. Daarvoor lees je eerst twee geschikte punten op de trendlijn af, bijvoorbeeld $(160,50)$ en $(190,68)$.

Met behulp van deze punten kun je de formule opstellen. Je vindt:

$G = 0{,}6 \cdot l - 46$

Hiermee kun je voorspellingen doen over bijvoorbeeld het gewicht van iemand in de doelgroep van 2 meter.

Opgave 3

Je ziet in **Uitleg 2** hoe Excel een lijn door de puntenwolk trekt en de bijbehorende formule berekent.

a Je ziet in de uitleg hoe je zelf de formule voor de trendlijn kunt berekenen.
Voer die berekening uit.

Je kunt nu met behulp van de gevonden regressielijn (trendlijn) voorspellingen doen.

b Hoe zwaar zou iemand van 2,00 m volgens de regressielijn moeten zijn?

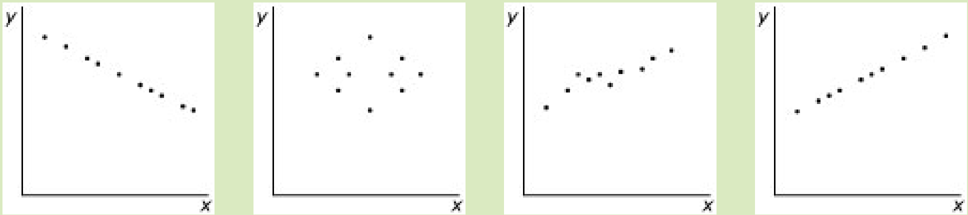
Theorie en voorbeelden

Om te onthouden

Wanneer je binnen een dataset zoekt naar een verband tussen twee statistische variabelen, gebruik je een **puntenwolk** of **spreidingsdiagram** (Engels: scatter plot).

Afhankelijk van de vorm van een puntenwolk kun je vastellen of er een verband tussen beide variabelen is en zo ja, of dat verband sterk is. Een maat daarvoor is de **correlatiecoëfficiënt** r, een getal met waarden vanaf - 1 tot en met 1. Hoe dichter r bij 1 of - 1 ligt, hoe sterker het verband. In een volgend voorbeeld staan vuistregels om aan te geven of de samenhang zwak, matig of sterk is.

In een spreidingsdiagram kun je met een **trendlijn** de (lineaire) samenhang weergeven. Hoe je deze trendlijn en de correlatiecoëfficiënt berekent, zie je in het **Practicum**.

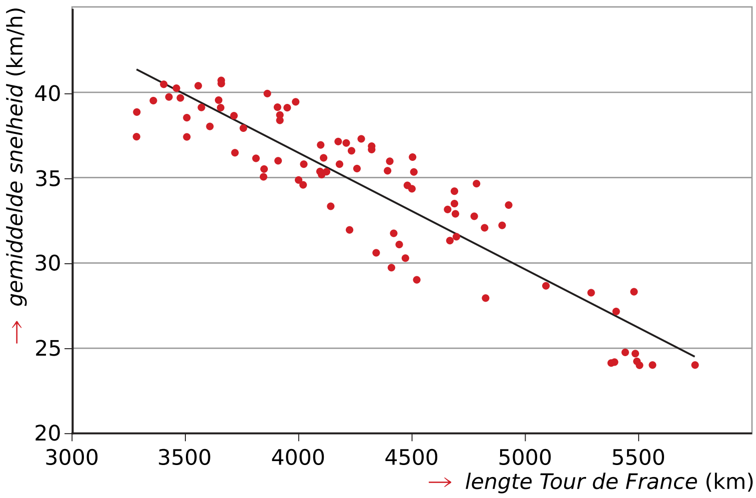


$r_{xy} = -1$ $r_{xy} = 0$ $r_{xy} = 0{,}6$ $r_{xy} = 1$

Figuur 4.6

Let op! Je stelt alleen vast dat er een **statistische samenhang** tussen beide variabelen is. Het is de vraag of dat verband ook causaal is. Je kunt dus niets zeggen over oorzaak en gevolg. Er kan best een derde variabele de oorzaak zijn van de samenhang.

Voorbeeld 1



Figuur 4.7

In deze figuur staan twee statistische variabelen. Op de horizontale as staat de totale lengte van de Tour de France in kilometer. Op de verticale as staat de gemiddelde snelheid van de winnaar.

Welke samenhang is er tussen de lengte van de Tour de France en de gemiddelde snelheid van de winnaar?

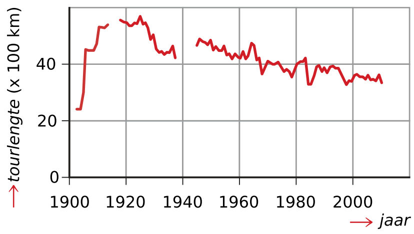
Antwoord

Als je kijkt naar de afbeelding, lijkt er een (sterke) negatieve samenhang te zijn. Hoe korter de Tour de France, hoe hoger de gemiddelde snelheid van de winnaar.

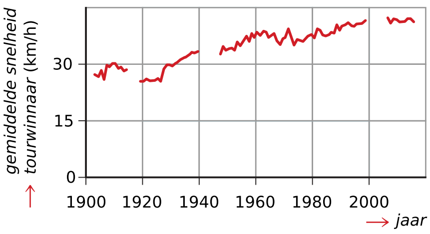
Maar pas op met oorzaak en gevolg. De oorzaak van een hogere gemiddelde snelheid van de winnaar hoeft niet de afnemende lengte van het parcours te zijn.

Opgave 4

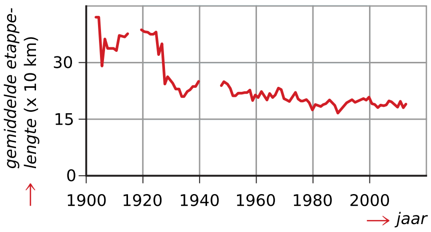
Bekijk de drie grafieken:



I



II



III

Figuur 4.8

Twee van de grafieken kun je combineren tot een spreidingsdiagram.

a Hoe komt het dat je deze grafieken kunt combineren tot een spreidingsdiagram?

b De drie grafieken hebben allemaal twee gaten. Wat betekent dit en wat is de oorzaak?

c Welke twee (van de drie) grafieken tonen gecombineerd het spreidingsdiagram uit het voorbeeld?

d Welke samenhang wordt zichtbaar, als je steeds twee andere grafieken combineert?

Opgave 5

Het is gevaarlijk om conclusies te trekken. Het gegeven spreidingsdiagram suggereert dat de gemiddelde snelheid van de winnaar omlaag gaat, als een Tour de France langer gemaakt wordt.

a Geef drie redenen waarom de gemiddelde snelheid van de winnaar los staat van de tourlengte.

b Noem nog een tweetal variabelen die vermoedelijk een sterke samenhang hebben met de gemiddelde snelheid van de winnaar.

c Noem een tweetal variabelen die geen oorzakelijk verband hebben met de gemiddelde snelheid van de winnaar.

Voorbeeld 2

De mate van samenhang is in een getal uit te drukken: de correlatiecoëfficiënt. Deze correlatiecoëfficiënt r kun je bijvoorbeeld met Excel of met de grafische rekenmachine laten uitrekenen. Zie het **Practicum**. Vervolgens kun je met vuistregels conclusies trekken.

- Als $r \leq -0{,}7$ dan is er sprake van sterke negatieve samenhang.
- Als $-0{,}7 < r \leq -0{,}3$ dan is er sprake van matige negatieve samenhang.
- Als $-0{,}3 < r < 0$ dan is er sprake van zwakke negatieve samenhang.
- Als $0 < r < 0{,}3$ dan is er sprake van zwakke positieve samenhang.
- Als $0{,}3 \leq r < 0{,}7$ dan is er sprake van matige positieve samenhang.
- Als $r \geq 0{,}7$ dan is er sprake van sterke positieve samenhang.

Opgave 6

Om te onderzoeken of er enig verband bestaat tussen de lengte van een vader en die van zijn zoon zijn de lengtes van 12 vaders en die van hun oudste zoons gemeten op het moment dat die zoons volwassen werden. De gegevens staan in deze tabel.

lengte vader v in cm	173	168	178	170	180	165	185	175	180	178	183	188
lengte zoon z in cm	180	175	180	173	183	175	180	173	188	178	180	185

Tabel 4.1

a Teken een spreidingsdiagram (een puntenwolk) bij deze gegevens.

b Bereken de correlatiecoëfficiënt in twee decimalen nauwkeurig.

c Bestaat er een lineair verband tussen v en z? Zo ja, stel dan een bijpassende formule op.

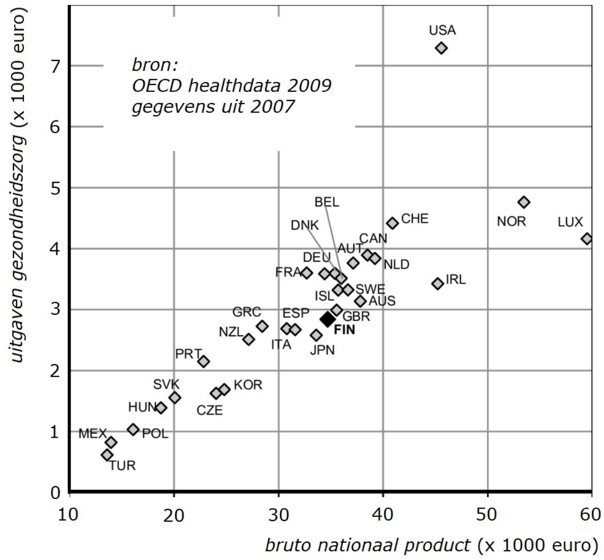
Voorbeeld 3

Bekijk het spreidingsdiagram. Daarin zijn de resultaten weergegeven van een onderzoek in dertig landen. Het bnp en de ug zijn gemiddelden per inwoner.

Je kunt door deze punten een trendlijn tekenen. De bijbehorende correlatiecoëfficiënt is een maat voor de sterkte van het verband. Is de lijn dalend, dan is de correlatiecoëfficiënt negatief en is ook de richtingscoëfficiënt van de trendlijn negatief.

Je kunt ook een trendlijn door deze gegevens tekenen. In de praktijk heb je de gegevens vaak in een database beschikbaar. De trendlijn bereken en teken je dan met de computer, zie het **Practicum**.

Welke vergelijking heeft deze trendlijn?



Figuur 4.9

Antwoord

Je zoekt twee punten. Deze punten moeten niet te dicht bij elkaar liggen, bijvoorbeeld de punten voor TUR en NLD: $(13; 0{,}5)$ en $(39; 3{,}9)$.

De vergelijking van de trendlijn is $ug \approx 0{,}13 \cdot bnp - 1{,}20$.

Opgave 7

Gebruik de gegevens uit **Voorbeeld 3**.

a Waarom moet je de punten niet te dicht bij elkaar kiezen om de vergelijking van een trendlijn op te stellen?

b Stel zelf de vergelijking van de trendlijn op.

Verwerken

Opgave 8

Bekijk het spreidingsdiagram. Op de horizontale as staat *kniehoogte* in cm en op de verticale as de *mouwlengte* in cm. De getallen stellen het aantal vrouwen voor die bij een bepaalde kniehoogte en mouwlengte horen. Je kunt hiervan een puntenwolk maken.

Statistisch onderzoek 1947
De Bijenkorf - Amsterdam
5001 vrouwen met leeftijd > 18 jr.

mouwlengte in cm \ kniehoogte in cm	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	
71																			1			1
70																						
69														1			1				1	3
68										2	2	1		3	3	2	2					15
67											3	1	2	4	2	2	3		1			18
66										4	5	5	4	6	10	7	7	3			1	52
65							1		2	7	3	15	13	26	16	11	6	3	1	1	1	106
64					1		2	2	5	5	13	22	26	31	22	18	4	4	4			159
63						3	1	5	5	20	36	44	48	40	29	16	6	4	2	1		260
62						2	6	15	21	61	61	84	62	59	27	14	6	1	2			421
61					3	2	6	29	43	65	99	111	91	61	33	9	7	1				560
60					3	5	17	40	68	112	117	121	84	49	31	2	2	2				653
59			1	1	2	11	20	60	110	108	88	66	72	22	12	3	1		1			578
58				3	8	21	50	72	110	141	103	90	34	22	5	1						660
57			1	2	9	28	60	80	92	87	72	53	20	12	2	1						519
56		1	2	4	13	30	50	79	82	60	40	32	8	3		1						405
55			3	6	11	27	38	46	46	39	14	16	1	3								250
54			4	7	11	17	32	37	23	14	8	4	1	1	3	1						163
53			3	6	6	16	12	17	15	8	6											89
52		1	2	3	5	9	13	10	5	4	1											53
51		1	2		4	4	5	2	2	1	1											22
50	1					3		1	3		3											11
49			1	2																		3
	1	3	19	34	76	178	313	495	632	738	675	665	466	343	195	88	45	18	12	2	3	5001

Figuur 4.10

Welke mate van samenhang verwacht je op grond van deze resultaten tussen *kniehoogte* (cm) en het *mouwlengte* (cm)?

Opgave 9

Bekijk de tabel met de correlatiecoëfficiënten van een aantal variabelen van de lengte van een vrouw. Gebruik nu de vuistregels uit **Voorbeeld 2**.

	gewicht	bovenwijdte	taille	heup	ruglengte	rugbreedte	vuistomvang	kniehoogte	voetlengte
lengte	0,2124	-0,0779	-0,1578	-0,0107	0,5933	0,0647	0,2668	0,8263	0,6737

Tabel 4.2

a Welke variabelen hebben een sterke samenhang met lengte?

b Welke variabelen hebben een matige samenhang met lengte?

c Welke variabelen hebben een zwakke samenhang met lengte?

Opgave 10

Soms is er wel een samenhang tussen twee statistische variabelen, maar kun je niet spreken van een causale (oorzaak-gevolg) relatie. Leg uit dat er een samenhang is tussen ijsverkoop en de verkoop van zonnebrillen, maar dat er geen oorzakelijk verband is tussen deze twee.

Opgave 11

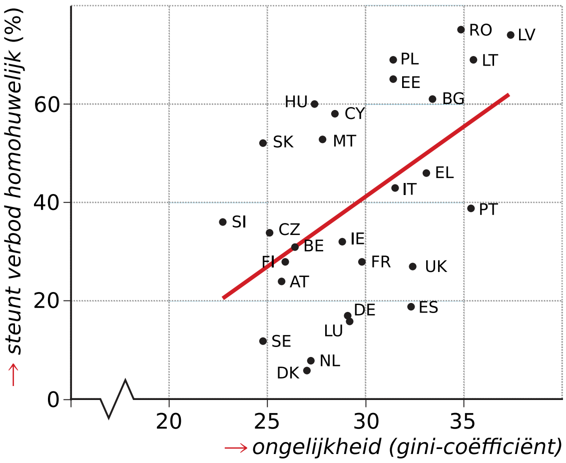
Bekijk deze puntenwolken.

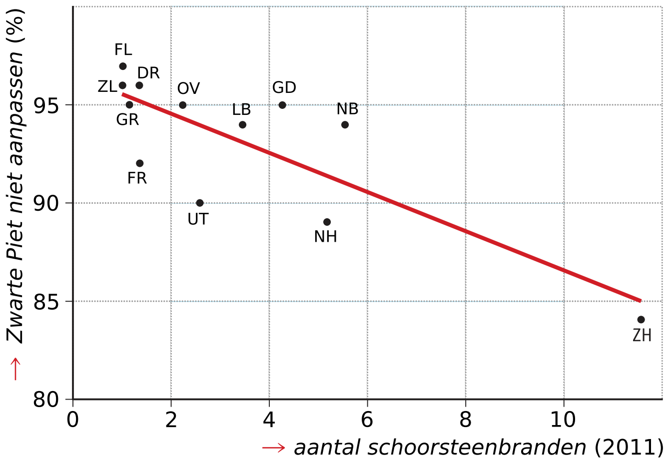
a Wat stelt een punt voor in de eerste figuur? En in de tweede figuur?

b Tussen welke statistische variabelen wordt in de figuren gezocht naar samenhang?

c Beschrijf de statistische samenhang tussen de variabelen op grond van de puntenwolken en de getekende trendlijn.

d Is er ook sprake van een causaal verband? Geef dan een mogelijke verklaring bij de puntenwolk.





Figuur 4.11

Opgave 12

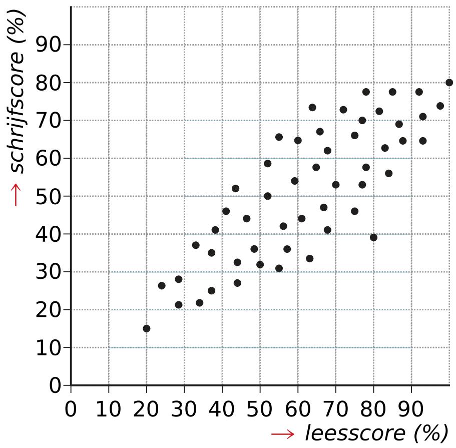
Een basisschool heeft een leestest en schrijftest Nederlands afgenomen bij de leerlingen in groep acht. De resultaten zijn verwerkt in een puntenwolk.

Er lijkt een verband te zijn tussen de schrijfscore S en de leesscore L.

a Stel een formule op voor de trendlijn die het verband tussen S en L weergeeft.

b Geef met behulp van de formule uit a een schatting van de schrijfscore bij een leesscore van 80%.

c Geef met behulp van de formule uit a een schatting van de leesscore bij een schrijfscore van 10%.

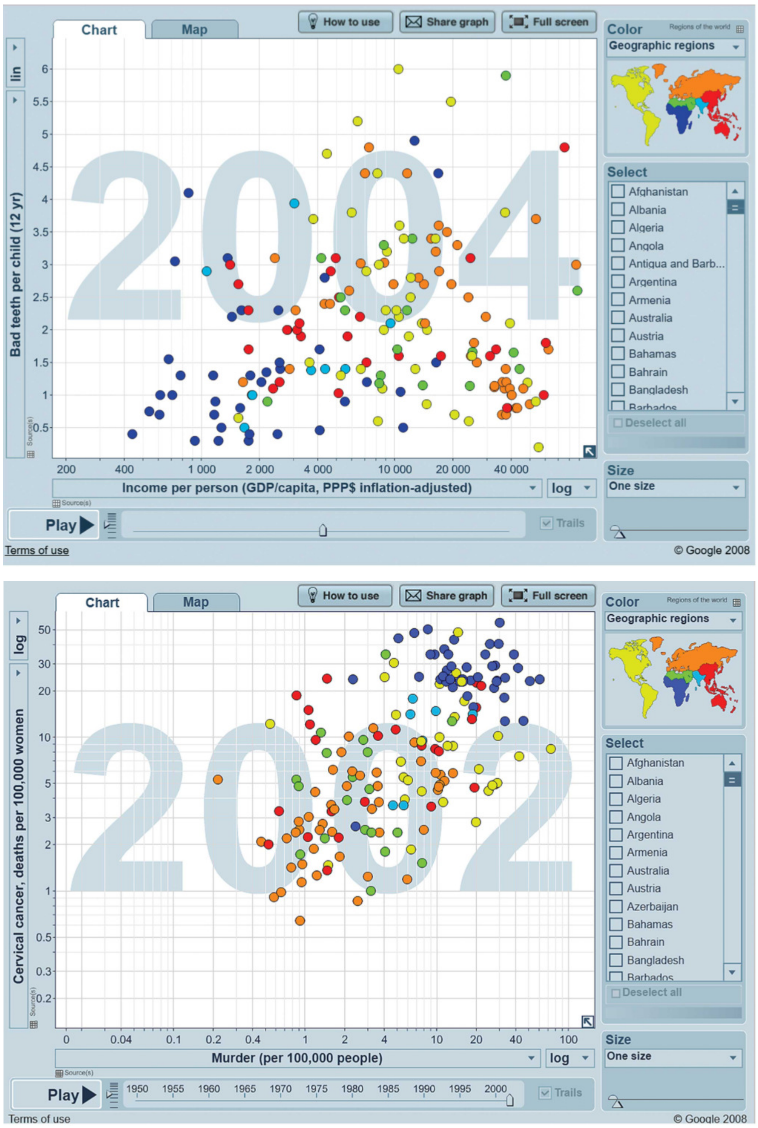


Figuur 4.12

Toepassen

Opgave 13: Gapminder

Met Gapminder (zie **http://www.gapminder.org/world**) kun je puntenwolken bekijken. Je kunt de statistische variabelen zelf kiezen en door op play te drukken kun je het verloop in de tijd zien. Let wel op de schaalverdelingen: lin = lineair; log = logaritmisch. Wil je de lineaire samenhang zien, dan moeten beide schalen op lin staan.



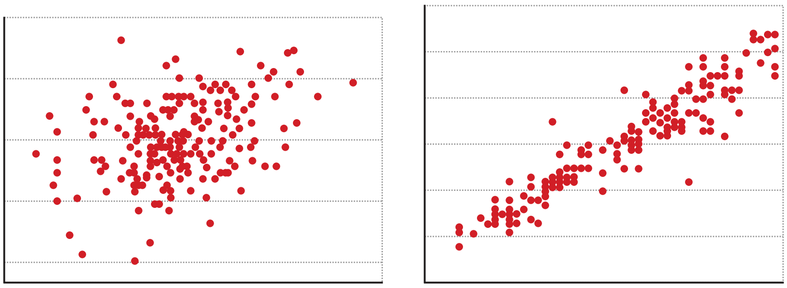
Figuur 4.13

a Door de vorm van de puntenwolk kun je een uitspraak doen over de lineaire samenhang. Bekijk de afbeeldingen en doe een uitspraak over de lineaire samenhang.

b Bij welke puntenwolk is er wel samenhang, maar geen lineaire samenhang?

Opgave 14: Huwelijken

In een onderzoek onder 199 echtparen is gevraagd naar de lengte en de leeftijd van de man en de vrouw. Onder andere werd onderzocht of er bij bepaalde eigenschappen van de gehuwden sprake was van een bepaalde statistische samenhang. Dit heeft geresulteerd in de volgende twee puntenwolken:

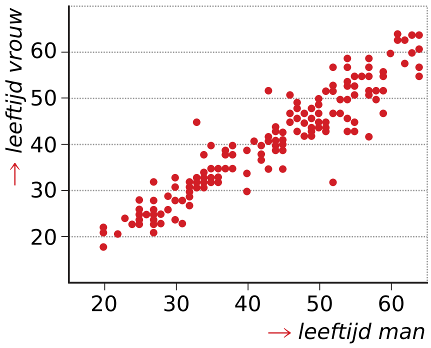


Figuur 4.14

Een van beide puntenwolken heeft betrekking op de leeftijden van de twee huwelijkspartners, waarbij de gegevens van de man op de horizontale as zijn uitgezet en die van de vrouw op de verticale as. De andere puntenwolk heeft betrekking op de lengte van beide partners. Ook hier zijn de gegevens van de man weer op de horizontale as uitgezet.

a Beredeneer dat, op basis van de vorm van de puntenwolk, de linker puntenwolk zeer waarschijnlijk betrekking heeft op de lengte en de rechter puntenwolk op de leeftijd.

b Bekijk de puntenwolk. Onderzoek met behulp van de puntenwolk of het in de betreffende 199 huwelijken vaker voorkomt dat de man ouder is dan de vrouw of dat het omgekeerde juist vaker voorkomt. Laat duidelijk zien hoe je tot je antwoord gekomen bent.



Figuur 4.15

Op basis van dergelijke puntenwolken wil men soms een schatting maken van de lengte of de leeftijd van een vrouw als men de lengte of de leeftijd van de man kent. Hoewel dit soort schattingen altijd een grote mate van onzekerheid hebben, is het toch mogelijk om aan te geven bij welk van de twee puntenwolken een dergelijke schatting het meest betrouwbaar zal zijn.

c Beredeneer bij welk van de twee puntenwolken, die met de leeftijden of die met de lengtes, een dergelijke schatting het meest betrouwbaar zal zijn.

In de tabel is een aantal kengetallen weergegeven uit het onderzoek.

	leeftijd man (jaar)	leeftijd vrouw (jaar)	lengte man (cm)	lengte vrouw (cm)
gemiddelde	42,6	40,7	173	160
minimum	20	18	156	141
maximum	64	64	195	176
standaardafwijking	11,6	11,4	6,9	6,2

Tabel 4.3

Ervan uitgaande dat de lengtes en de leeftijden van de huwelijkspartners nagenoeg normaal verdeeld zijn, is met behulp van deze gegevens uit te rekenen dat 95% van de lengtes van de mannen tussen de 159,2 cm en 186,8 cm zal liggen.

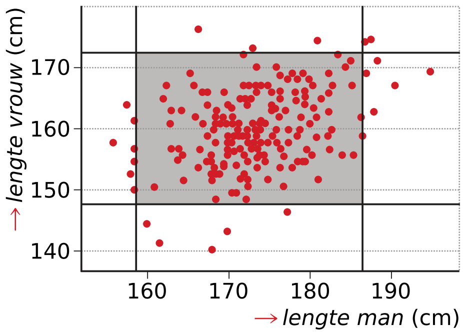
d Leg uit hoe je aan deze waarden komt.

e Bepaal tussen welke twee lengtes 95% van de vrouwen zit.

Omdat 5% van de mannen buiten de berekende grenzen zal vallen, evenals 5% van de vrouwen, concludeert de onderzoeker dat in totaal 10% van de punten uit de puntenwolk buiten de getekende rechthoek zullen vallen.

f Beargumenteer of je het met die conclusie eens bent of niet.

(bron: voorbeeldopgave Statistiek – syllabus havo A)

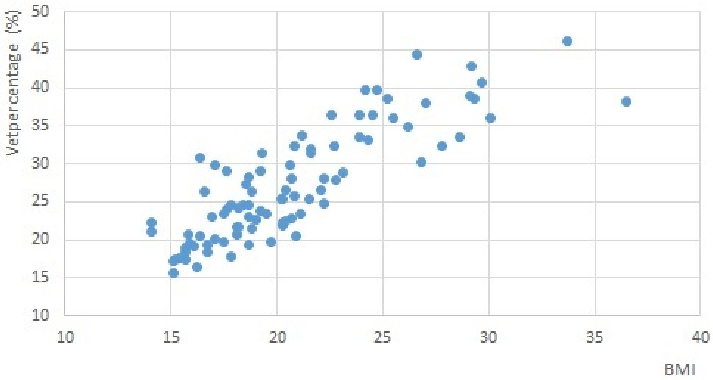


Figuur 4.16

Testen

Opgave 15

Bekijk de puntenwolk BMI-Vetpercentage. Daarin zijn de resultaten weergegeven van een onderzoek onder 90 jongeren. BMI is een getal dat samenhangt met lengte en gewicht, vetpercentage is het percentage van het lichaamsgewicht dat bestaat uit vet.



Figuur 4.17

a Is er een statistische samenhang? Geef een schatting van de correlatiecoëfficiënt.

b Is er een oorzakelijk verband?

c Bij deze puntenwolk kun je een trendlijn tekenen die door (15,17) en (30,42) gaat. Welke formule hoort bij deze trendlijn?

d Voorspel met behulp van de gevonden trendlijn het vetpercentage van iemand met een BMI van 21.

Practicum

Met deze practica leer je hoe je de **de trendlijn** met de grafische rekenmachine tekent en berekent.

- **Trendlijn, correlatie en de TI84**
- **Trendlijn, correlatie en de TInspire**
- **Trendlijn, correlatie en de Casio**
- **Trendlijn, correlatie en de HPprime**
- **Trendlijn, correlatie en de NumWorks**

Met het volgende practicum kun je zien hoe je **de trendlijn en de correlatiecoëfficiënt in Excel** berekent. Dat is handig als je een grote set gegevens hebt. Je treft er ook in aan hoe je de **regressielijn**, dat is de meest geschikte lijn door de puntenwolk, kunt tekenen en er door Excel de vergelijking van kunt laten opstellen. In Excel heet die lijn de 'trendlijn'. In het volgende onderdeel hoor je daar meer over.

- **Correlatie en regressie**

OPMERKING:
Natuurlijk is het veel mooier om een eigen dataset met gegevens van leerlingen in jouw jaargroep te gebruiken. Die moet je dan wel eerst zelf maken: statistisch onderzoek!

2.5 Totaalbeeld

Samenvatten

Je hebt nu alle theorie van het onderwerp ‘Conclusies trekken’ doorgewerkt. Het is tijd om een overzicht over het geheel te krijgen.

Begrippenlijst

- kwalitatief, kwantitatief — discreet, continu — nominaal, ordinaal
- verschil van kwalitatieve variabelen — procentueel verschil — cumulatief verschilpercentage — kruistabel en phi
- verschil van kwantitatieve variabelen — boxplot en vuistregels — effectgrootte
- verband tussen variabelen — trendlijn

Activiteitenlijst

- vuistregels normale verdeling toepassen
- drie methoden om het verschil tussen kwalitatieve variabelen in een getal uit te drukken toepassen — vuistregels via formulekaart
- drie methoden om het verschil tussen kwantitatieve variabelen in een getal uit te drukken toepassen — vuistregels via formulekaart
- statistisch verband tussen twee variabelen herkennen — verschil statistisch verband en oorzakelijk verband herkennen

Achtergronden

Je hebt diverse manieren gezien om statistische variabelen te vergelijken of juist naar een statistisch verband tussen twee variabelen te zoeken. Wie hebben al die methoden toch bedacht?

- De *phi* voor het bepalen of in een twee bij twee kruistabel de twee kwalitatieve variabelen weinig, een beetje of veel van elkaar verschillen, is bedacht door **Karl Pearson (1857—1936)**, één van de grondleggers van veel statistische methoden. In feite hangt dit getal nauw samen met de correlatiecoëfficiënt r, in dit geval voor twee punten in een assenstelsel.
- De *effectgrootte* is een statistische maat voor hoe sterk het effect van een handeling is op een populatie, waarbij vergeleken wordt met een andere populatie waarop die handeling niet wordt toegepast (‘controlegroep’). Deze maat is bedacht door de Amerikaanse statisticus en psycholoog **Jacob Cohen (1923—1998)**.
- De *boxplot* wordt gebruikt om twee kwantitatieve variabelen te vergelijken. Hij is voor het eerst geïntroduceerd door de Amerikaanse wiskundige **John Tukey (1915—2000)**. Tukey was één van de grondleggers van de data-analyse.

Figuur 5.1 Karl Pearson

- De *correlatiecoëfficiënt* r voor het bepalen of er tussen twee kwantitatieve variabelen een (statistisch) verband bestaat, is ook bedacht door **Karl Pearson (1857—1936)**, hoewel de trendlijn zelf al ouder was en door **Carl Friedrich Gauss (1777—1855)** met de kleinste-kwadraten-methode is afgeleid.

Testen

Opgave 1

Op de Universiteit van Leiden is een onderzoek gedaan onder een groep studenten. Hierbij werden onder andere de lichaamslengte en de armlengte van de studenten gemeten. Ook werden het geslacht, de kleur van de ogen en de voorkeurshand genoteerd. In totaal ging het om 34 mannelijke studenten en 32 vrouwelijke studenten. Enkele resultaten zie je in de tabel.

	voorkeurshand	
geslacht	links	rechts
man	5	29
vrouw	2	30

Tabel 5.1

Op basis van resultaten van dit onderzoek wil men uitspraken doen over alle mannelijke en vrouwelijke studenten in Leiden.

a Noem de variabelen uit de tabel en geef van elk van de variabelen aan of deze nominaal of ordinaal is. Licht je antwoord toe.

b Bepaal met behulp van de **Formulekaart** of het verschil in voorkeurshand tussen de mannelijke en de vrouwelijke studenten in het onderzoek groot, middelmatig of gering is.

In de tabel hieronder staan het gemiddelde en de standaardafwijking van de lichaamslengte van de studenten.

	lichaamslengte (cm)		
geslacht	aantal	gemiddelde	standaardafwijking
man	34	183,8	5,8
vrouw	32	170,8	8,1

Tabel 5.2

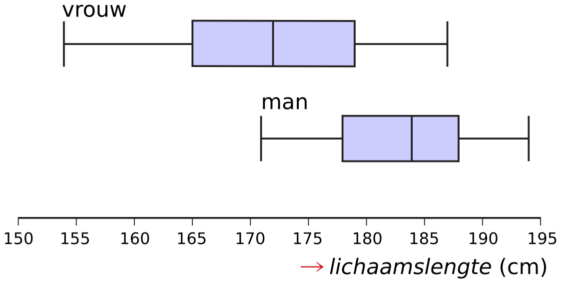
Op basis van deze tabel kunnen voor de lichaamslengte van mannen en vrouwen 95%-betrouwbaarheidsintervallen worden opgesteld.

c Onderzoek welk van deze twee intervallen het smalst is. Welke informatie geeft zo'n 95%-betrouwbaarheidsinterval in deze context?

Iemand die iets wil zeggen over de grootte van het verschil in lichaamslengte tussen de mannen en de vrouwen in het onderzoek, zou gebruik kunnen maken van de effectgrootte. Bij de berekening van de effectgrootte zijn de gemiddeldes en standaardafwijkingen van beide groepen van belang. Ga ervan uit dat het verschil van de gemiddeldes gelijk blijft.

d Leg uit op welke wijze grotere standaardafwijkingen de grootte van het verschil (groot, middelmatig, gering) kunnen beïnvloeden.

De resultaten van de lichaamslengtes werden ook in boxplots verwerkt. Zie figuur 1.



Figuur 1

Figuur 5.2

Je kunt met behulp van de gegevens over de lichaamslengtes in de tabel bij c onderzoeken hoe groot het verschil in lichaamslengte in het onderzoek is. Dit kun je ook doen met behulp van de gegevens in figuur 1.

e Toon met behulp van de **Formulekaart** aan dat je met deze twee gegevensbronnen een verschillende conclusie trekt over de grootte van het verschil in lichaamslengte tussen de mannelijke en de vrouwelijke studenten.

f Beargumenteer welk van beide conclusies het best te verdedigen is.

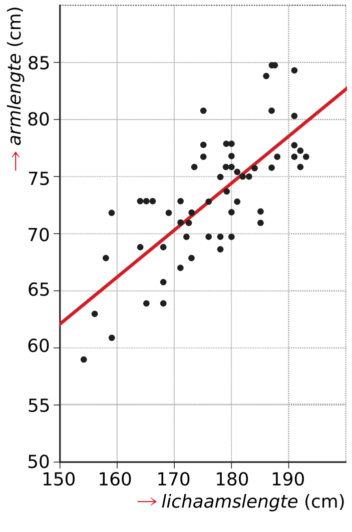
Om te onderzoeken of er een verband bestaat tussen de lichaamslengte en de armlengte van de studenten zijn de resultaten van het onderzoek in een spreidingsdiagram gezet. Zie figuur 2. Er blijkt in dit onderzoek bij benadering een lineair verband te zijn tussen de lichaamslengte L in centimeter en de armlengte A in centimeter. De getekende trendlijn geeft dit verband weer.

g Stel een formule op van de trendlijn.

De punten in het spreidingsdiagram liggen niet precies op de trendlijn. Ga ervan uit dat de armlengte normaal verdeeld is bij alle mogelijke lichaamslengtes tussen 150 cm en 200 cm en dat bij elke lichaamslengte de standaardafwijking van de bijbehorende armlengte gelijk is aan 3 cm. Als we de resultaten van dit onderzoek gebruiken om op basis van de lichaamslengte de armlengte van studenten te voorspellen, kunnen we hierbij figuur 2 gebruiken.

h Hoe kun je in figuur 2 het gebied waar de middelste 95% van de armlengten ligt voor elk van de lichaamslengten tussen 150 en 200 cm aangeven?

(bron: voorbeeldopgaven syllabus havo wiskunde A)



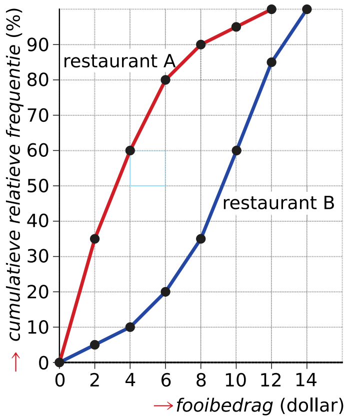
Figuur 2

Figuur 5.3

Opgave 2

In de Verenigde Staten is het gebruikelijk dat je in een restaurant een flinke fooi geeft aan degene die je bedient. Het basisloon is er zeer laag en daardoor is het bedienend personeel veel meer afhankelijk van fooien dan in Nederland. In New York bestudeerde een onderzoeker welke fooien er gegeven werden bij 500 rekeningen in restaurant A en 500 rekeningen in restaurant B.

In de figuur zijn de cumulatieve relatieve frequentiepolygonen van de fooien van de restaurants A en B getekend.



Figuur 5.4

a In welk restaurant werd in totaal meer fooi gegeven? Licht je antwoord toe.

b In welk restaurant werden er relatief meer fooien tussen de 6 en de 8 dollar gegeven? Licht je antwoord toe.

Ruim driekwart van de fooien in restaurant B is hoger dan de 75% laagste fooien van restaurant A.

c Leg uit hoe je dit in de figuur terug kunt vinden.

Op grond van de figuur kun je met behulp van de **Formulekaart** met twee verschillende vuistregels een uitspraak doen over het verschil tussen de fooien in restaurant A en die in restaurant B.

d Beschrijf hoe je met twee verschillende vuistregels dit verschil kunt bepalen en geef bij beide de conclusie: is het verschil groot, middelmatig of gering?

(bron: voorbeeldopgaven syllabus havo wiskunde A)

Opgave 3

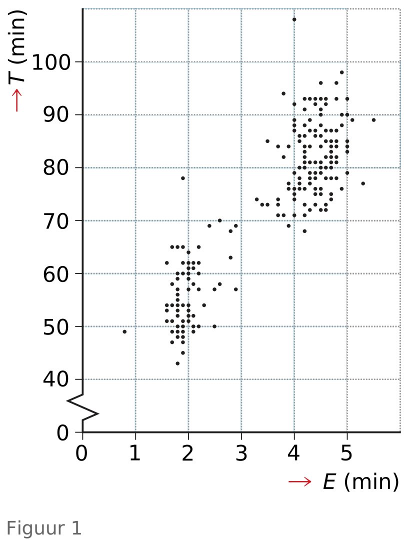
Vulkanen kunnen heel lang niet actief zijn. Dan zijn er geen erupties (uitbarstingen). Tijdens een actieve periode van een vulkaan zijn er wel erupties. Bij een eruptie komt er gesmolten steen, gas en as uit de vulkaan. Dat duurt een tijdje. Daarna is de vulkaan weer rustig, totdat de volgende eruptie begint.

Je bekijkt in deze opgave één actieve periode van één vulkaan. De actieve periode start bij de eerste eruptie. Wetenschappers hebben tijdens deze actieve periode gemeten:

- hoelang iedere eruptie duurt: de eruptieduur;
- hoelang de vulkaan rustig is tot de volgende eruptie begint: de tussentijd tot de eerstvolgende eruptie.

Tijdens deze actieve periode was de langste tijd tussen twee erupties 108 minuten. Na de allerlaatste eruptie was de vulkaan weer lange tijd rustig.

De eruptieduur is gemeten in tienden van een minuut nauwkeurig en de tussentijd in gehele minuten. De metingen zijn verwerkt in een puntenwolk. De eruptieduur E staat langs de horizontale as, de tussentijd tot de eerstvolgende eruptie T staat langs de verticale as.



Figuur 5.5

Het meest linkse punt in de figuur hoort bijvoorbeeld bij een eruptie die $0{,}8$ minuten duurde en waarna de vulkaan 49 minuten rustig was. In deze actieve periode zijn er 184 erupties geweest. De eerste 183 zijn weergeven in het figuur. De allerlaatste eruptie van deze actieve periode duurde $1{,}7$ minuten.

a De allerlaatste eruptie van een actieve periode kan niet in de puntenwolk worden weergegeven. Waarom niet?

In de figuur zijn de gegevens verwerkt van alle erupties, waarbij zowel de eruptieduur als de tussentijden tussen de erupties verwerkt zijn. Op basis van deze figuur is ook een schets te maken van de verdeling van de eruptieduur.

b Maak deze schets, met op de horizontale as de eruptieduur en op de verticale as het aantal erupties. Je hoeft hierbij niet het aantal erupties per eruptieduur exact te tellen, het gaat om een schets.

Iemand beweert op basis van deze schets van de eruptieduur te kunnen bepalen wat de standaardafwijking is van de eruptieduur. Hij doet dit door gebruik te maken van een van de vuistregels van de normale verdeling. Hij berekent het tijdsverschil tussen de langste eruptieduur en de kortste eruptieduur en deelt deze uitkomst door 6. De uitkomst van deze berekening beschouwt hij als een schatting van de standaardafwijking.

c Ben jij het met deze werkwijze eens? Geef duidelijk aan wat daarbij je argumenten zijn.

d Je kunt voor de 184 erupties de gemiddelde duur van de tussentijd T schatten. Met behulp van dit gemiddelde kun je een schatting maken van de lengte van de actieve periode van de vulkaan. Maak een schatting van de gemiddelde tussentijd tot de volgende eruptie en toon hiermee aan dat deze actieve periode van de vulkaan langer dan een week heeft geduurd.

In figuur 2 is een lijn getrokken die zo goed mogelijk bij de metingen past. Deze lijn gaat door de punten $(2{,}56)$ en $(5{,}90)$. Met behulp van deze lijn kun je bij een gegeven eruptieduur een grove schatting maken voor de tussentijd die je daarbij kunt verwachten.

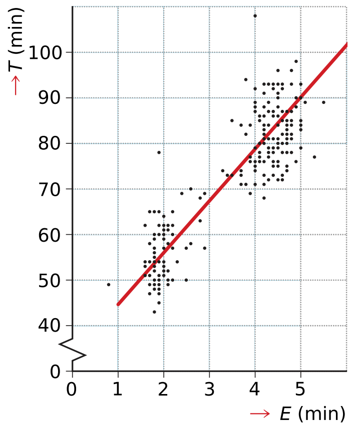
e Bereken met behulp van de getrokken lijn uit dit figuur wat de geschatte tussentijd is bij een eruptieduur van 6 minuten.

Ga ervan uit dat de tussentijden bij een eruptietijd van $4{,}5$ minuten normaal verdeeld zijn met een standaardafwijking van 4 minuten. Je kunt dan met een betrouwbaarheid van 95% aangeven tussen welke twee waarden de tussentijd tot de eerstvolgende eruptie zal liggen bij een eruptieduur van $4{,}5$ minuten.

f Geef deze twee waarden.

g Piet beweert op basis van de gegevens uit de figuur te kunnen concluderen dat de eruptieduur groter wordt naarmate de tussentijd tussen twee erupties groter is. Harm is het niet met Piet eens. Geef een argument dat Harm kan gebruiken om Piet te overtuigen van zijn ongelijk.

(bron: voorbeeldopgaven syllabus havo wiskunde A)

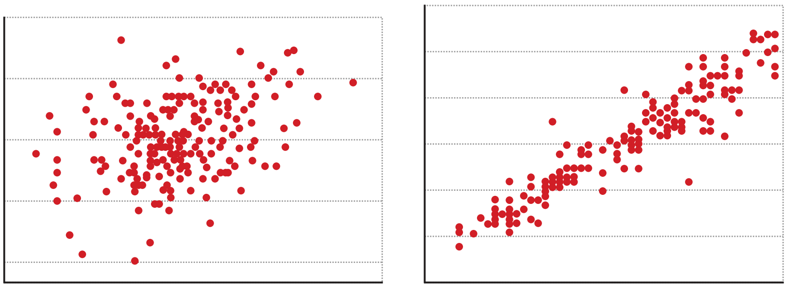


Figuur 2

Figuur 5.6

Opgave 4

In een onderzoek onder 199 echtparen is onder meer gevraagd naar de lengte en de leeftijd van de man en de vrouw. Er werd onder andere onderzocht of er bij bepaalde eigenschappen van de gehuwden sprake was van een bepaalde statistische samenhang. Dit heeft geresulteerd in de volgende twee puntenwolken:



Figuur 5.7

De rechter puntenwolk heeft betrekking op de leeftijden van de twee huwelijkspartners, waarbij de gegevens van de man op de horizontale as zijn uitgezet en die van de vrouw op de verticale as. De linker puntenwolk heeft betrekking op de lengte van beide partners. Ook hier zijn de gegevens van de man weer op de horizontale as uitgezet.

a Waarom kun je door de rechter puntenwolk het beste een trendlijn tekenen?

b Op het **werkblad** is de rechter figuur getekend met schaalverdelingen op beide assen.
Teken de trendlijn in de figuur.

In de tabel is een aantal kentallen weergegeven uit het onderzoek.

	leeftijd man (jaar)	leeftijd vrouw (jaar)	lengte man (cm)	lengte vrouw (cm)
gemiddelde	42,6	40,7	173	160
minimum	20	18	156	141
maximum	64	64	195	176
standaardafwijking	11,6	11,4	6,9	6,2

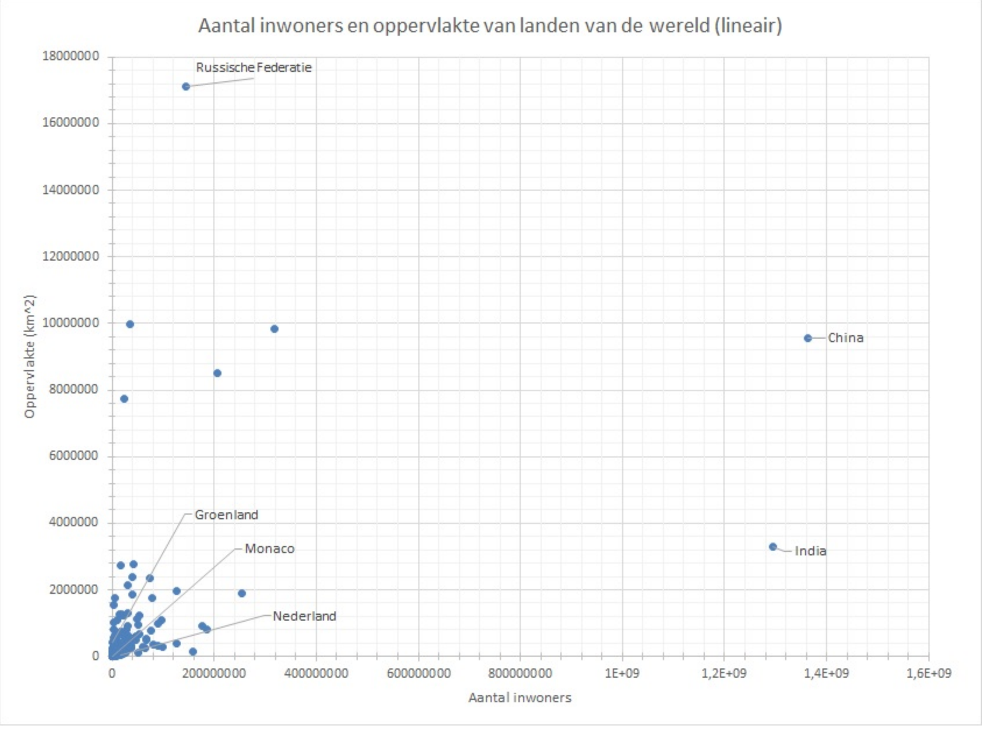
Tabel 5.3

c Bereken voor de leeftijden en de lengtes de effectgrootte om het verschil tussen mannen en vrouwen te vergelijken.
Gebruik de **Formulekaart** en rond af op twee decimalen.

(naar: voorbeeldopgaven syllabus havo wiskunde A)

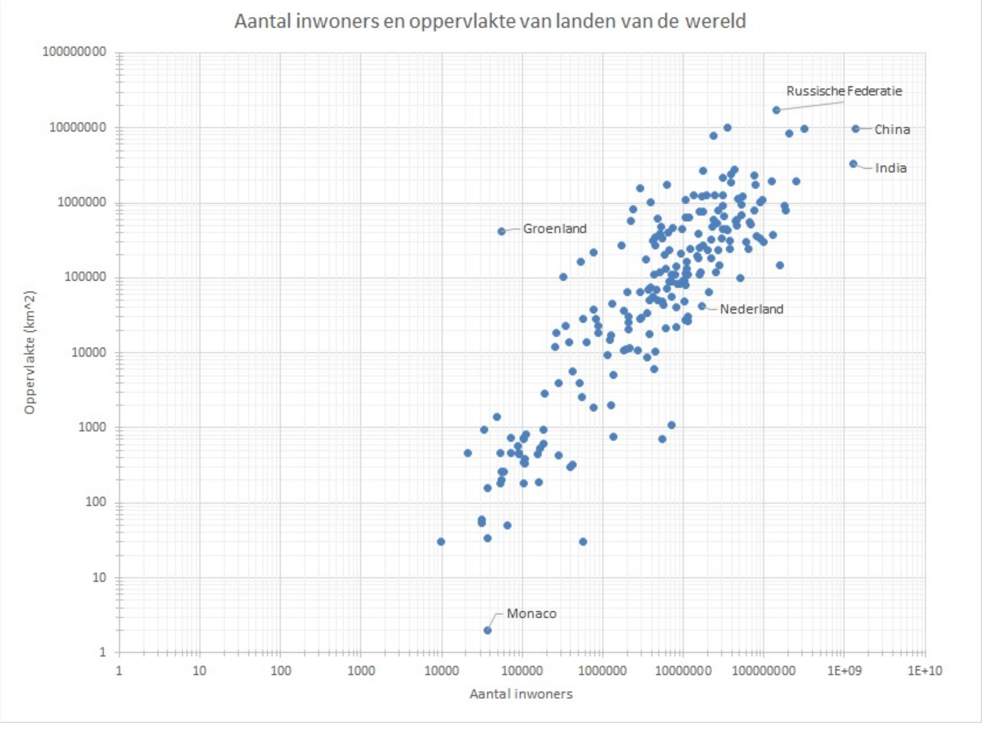
Toepassen

Opgave 5: Trendlijn op logaritmisch papier



Figuur 5.8

In de figuur hierboven is voor een groot aantal landen het aantal inwoners en de oppervlakte weergegeven. Het is duidelijk dat aflezen hier erg moeilijk is. Een manier om dit te verbeteren is het gebruiken van logaritmische schalen. Dat zie je in de volgende figuur.



Figuur 5.9

a Het land dat de meeste inwoners telt, is China.
Lees af hoeveel inwoners China heeft. Geef je antwoord in miljarden, afgerond op één decimaal.

b Lees af hoe groot het oppervlak van China is. Geef je antwoord in miljoenen, afgerond op één decimaal.

c Het land met de grootste oppervlakte is Rusland.
Bereken het aantal inwoners per km^2 voor China en Rusland.

In de figuur is ook de trendlijn getekend.

De vergelijking van de trendlijn is: $A = \frac{1}{10000} \cdot P^{\frac{4}{3}}$

waarin A de oppervlakte in km^2 en P de bevolkingsomvang is.

d Bereken met deze formule het aantal inwoners van een land van 200000 km^2.
Geef je antwoord in miljoenen, afgerond op één decimaal.

Examen

Opgave 6: Alutech

Bij het bedrijf Alutech, dat aluminium buizen produceert, werkt men aan één stuk door van maandagmorgen tot en met zaterdagmiddag. De werknemers van dit bedrijf werken in ploegendiensten van 8 uur. Daarbij zijn er twee dagdiensten, genaamd A en B, en één nachtdienst N. Zie tabel.

werktijden Alutech							
	ma	di	wo	do	vr	za	20
07.00 – 15.00 uur	A	A	A	A	A	A	-
15.00 – 23.00 uur	B	B	B	B	B	-	-
23.00 – 07.00 uur	N	N	N	N	N	-	-

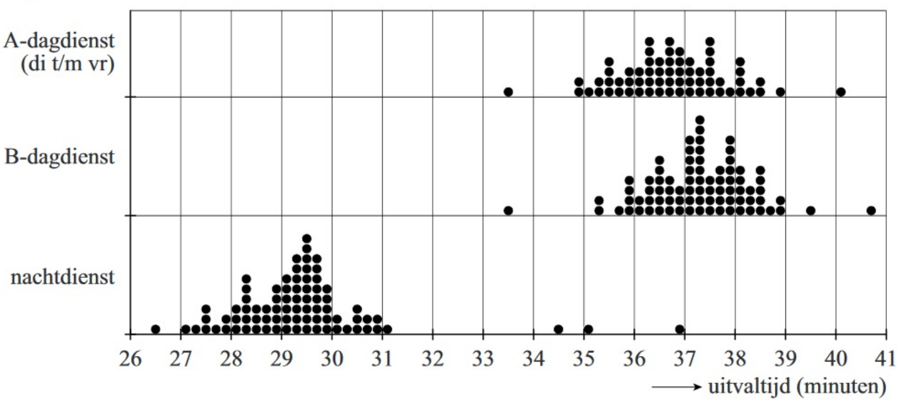
Tabel 5.4

Op maandagmorgen is men de eerste 3 uur van de A-dienst bezig om het productieproces op te starten en op zaterdagmiddag wordt het laatste uur van de A-dienst gebruikt om het proces gecontroleerd stil te zetten. Alle overige tijd wordt beschouwd als mogelijke productietijd.

a Bereken het aantal uren mogelijke productietijd per week.

Tijdens elke dienst komen er storingen voor. Het productieproces wordt dan een aantal minuten stilgelegd totdat de storing verholpen is. Telkens wordt bijgehouden hoe lang de storing duurt. Na afloop van de dienst wordt de totale tijd van alle storingen genoteerd. Deze tijd noemt men de uitvaltijd. De directie wil dat de uitvaltijd zo klein mogelijk is. Om te onderzoeken hoe groot de uitvaltijd is, heeft men van 16 werkweken de uitvaltijden van de dag- en nachtdiensten vergeleken. Daarbij heeft men bij de A-diensten alleen gekeken naar de diensten die op dinsdag tot en met vrijdag vallen. De resultaten staan in de dotplot in figuur 1.

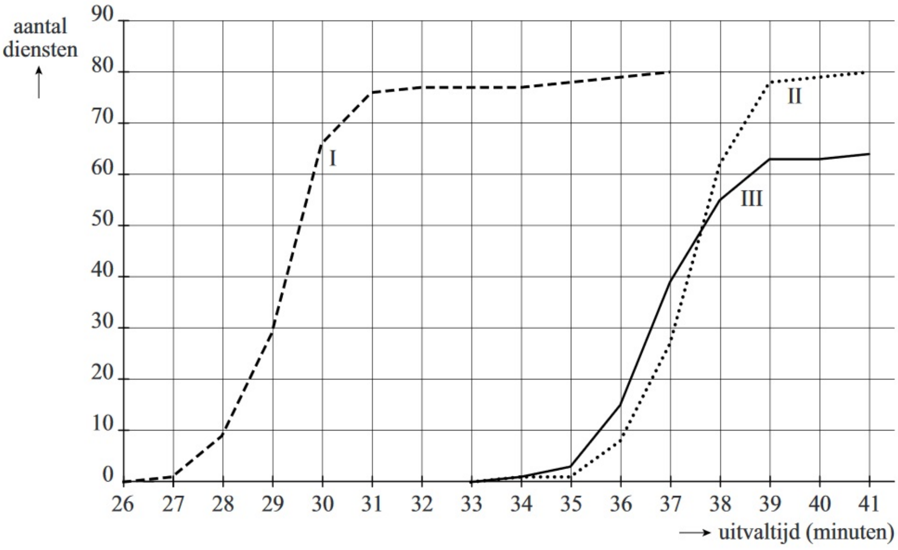
figuur 1 uitvaltijd per dienst



Figuur 5.10

Met de gegevens uit de dotplot zijn van alle diensten cumulatieve frequentiepolygonen gemaakt. Er is gebruikgemaakt van een klassenindeling met klassenbreedte 1 minuut. Zie figuur 2.

figuur 2



Figuur 5.11

b Beredeneer met behulp van figuur 1 welke cumulatieve frequentiepolygoon hoort bij dagdienst B.

Tot verbazing van de directie is de uitvaltijd tijdens de nachtdiensten in vrijwel alle gevallen lager dan tijdens de beide dagdiensten. Omdat elke minuut uitvaltijd het bedrijf geld kost, besluit de directie nader onderzoek te doen naar de uitvaltijden van de drie verschillende diensten. De directie berekent de mediaan, het eerste kwartiel en het derde kwartiel van de uitvaltijd voor elk van de drie diensten. Zie deze tabel.

uitvaltijd per dag- of nachtdienst in minuten			
	mediaan	eerste kwartiel	derde kwartiel
dagdienst A (di-vr)	36,7	36,1	37,5
dagdienst B	37,3	36,6	37,9
dagdienst N	29,3	28,5	29,7

Tabel 5.5

Eerst kijkt de directie alleen naar het verschil tussen dagdienst A en dagdienst B. Met behulp van boxplots kun je een uitspraak doen over het verschil tussen de uitvaltijden van de twee dagdiensten. Daarvoor hoeven de boxplots niet getekend te worden.

c Bepaal met behulp van de **Formulekaart** en de tabel of het verschil in uitvaltijd tussen dagdienst A en dagdienst B groot, middelmatig of gering is.

In figuur 1 is te zien dat er bij alle diensten waarnemingen zijn die opvallend afwijken van de rest. We leggen daarom vast wat we in deze opgave met een uitschieter bedoelen: “Een uitschieter is een waarneming die meer dan $1{,}5$ keer de interkwartielafstand onder het eerste of boven het derde kwartiel ligt.”

d Bepaal hoeveel waarnemingen bij dagdienst B uitschieters zijn.

Van elk van de drie verschillende diensten worden ook de gemiddelde uitvaltijd en de standaardafwijking berekend. Zie deze tabel.

uitvaltijd per dag- of nachtdienst in minuten		
	gemiddelde	standaardafwijking
dagdienst A (di-vr)	36,75	1,10
dagdienst B	37,29	1,04
dagdienst N	29,39	1,53

Tabel 5.6

Men vermoedt dat de lagere uitvaltijden tijdens de nachtdiensten te maken hebben met het feit dat de energietoevoer gedurende de nacht constanter is dan overdag. Daarom wordt de energietoevoer overdag verbeterd. De directie van Alutech hoopt daarmee de uitvaltijden van de dagdiensten terug te dringen.
Na verloop van tijd blijkt dat de gemiddelde uitvaltijd van de A-diensten van dinsdag tot en met vrijdag gelijk geworden is aan de gemiddelde uitvaltijd van de nachtdiensten. Ook de gemiddelde uitvaltijd van alle B-diensten is gelijk geworden aan de gemiddelde uitvaltijd van de nachtdiensten. De standaardafwijkingen van de A-diensten en B-diensten zijn niet veranderd.

Je kunt nu voor dagdienst B de oude met de nieuwe situatie vergelijken.

e Bereken met behulp van de **Formulekaart** voor dagdienst B of het verschil in uitvaltijd tussen de oude en de nieuwe situatie groot, middelmatig of gering is.

Doordat de gemiddelde uitvaltijd van de A-diensten van dinsdag tot en met vrijdag en van alle B-diensten is teruggedrongen tot het gemiddelde van de uitvaltijd van de nachtdiensten, zal de productietijd toenemen.

f Bereken met hoeveel uur de totale netto productietijd op jaarbasis toeneemt. Neem hierbij aan dat men 51 volledige werkweken per jaar werkt. Laat hierbij de A-diensten van maandag en zaterdag buiten beschouwing.

(bron: examen havo wiskunde A in 2017, tweede tijdvak)

2.6 Compleet onderzoek

Inleiding

Het is nu tijd voor een compleet statistisch onderzoek. Bij een statistisch onderzoek zoek of verzamel je data zodat je een uitspraak kunt doen over je onderzoeksvraag en een betrouwbare en begrijpelijke onderbouwing kunt geven. Deze data verwerk je tot een overzichtelijk geheel waar je conclusies uit kunt trekken.

Figuur 6.1

Je leert in dit onderwerp

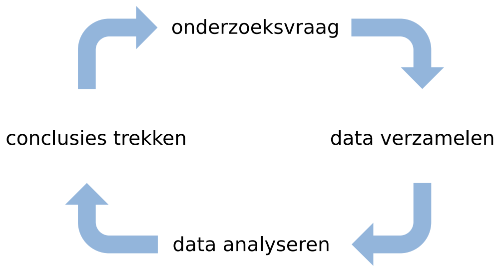
- een statistisch onderzoek opzetten.

Voorkennis

- alle kennis van het domein Statistiek.

Uitleg

Het is nu tijd voor een compleet statistisch onderzoek. Bij een statistisch onderzoek zoek of verzamel je data zodat je een uitspraak kunt doen over je onderzoeksvraag en een betrouwbare en begrijpelijke onderbouwing kunt geven. Bij de rapportage respecteer je altijd de privacy van deelnemers en verwerk je de gegevens altijd anoniem.



Figuur 6.2

Bij een onderzoek is het essentieel dat je een goede **onderzoeksvraag** hebt.

Er zijn globaal drie soorten onderzoeksvragen:

- de **frequentievraag**, hier gaat om 'hoeveel' en 'hoe vaak', het soort onderzoek dat er bij hoort is een beschrijvend onderzoek
- de **verschilvraag**, hier gaat het om het verschil tussen het ene en het andere, en het bijbehorende onderzoek is exploratief
- de **samenhangvraag**, hier gaat het om de samenhang tussen aspecten, het bijbehorende onderzoek is hier ook exploratief.

Er wordt onderscheid gemaakt tussen **veldonderzoek** en **bureauonderzoek**. Met veldonderzoek verwerf je eigen **data**, bijvoorbeeld met een digitale enquête, zie het practicum **Data verzamelen met een enquête**. Bij een bureauonderzoek gebruik je bestaande data. Veel data zijn te vinden op internet, bijvoorbeeld **http://statline.cbs.nl/Statweb**.

Vervolgens ga je met de data aan de slag:

- Tabellen en grafieken maken en statistische maten voor centrum en spreiding berekenen.
- Proporties en bijbehorende betrouwbaarheidsintervallen berekenen.

- Verschil tussen twee variabelen onderzoeken en conclusies trekken.
- Samenhang tussen twee variabelen onderzoeken en conclusies trekken.

Voor de verwerking van data ben je aangewezen op tools, zoals Excel, VUStat, DWO of GeoEnZo. Soms kun je ook een grafische rekenmachine gebruiken. Elke **tool** heeft voor- en nadelen; bovendien willen ze nog weleens veranderen.

Welke tool je ook gebruikt, je moet altijd zelf de uitspraken doen en onderbouwen. En bewaar altijd je (ruwe) data. Conclusies onderbouw je met tabellen, diagrammen, statistische maten en redeneringen. **Verwerk** alles in een mooi rapport, een heldere presentatie of een overzichtelijke poster. Beëindig het onderzoek met een kritische **reflectie**.

De tools vind je vaak op internet. Excel staat meestal al standaard op een pc of laptop.
VUStat vind je op **http://www.vusoft.be/vustat.html**
GeoEnZO vind je op **http://www.math4all.nl/pagina/geoenzo**
De DWO (nu Numworx) vind je op **http://www.fi.uu.nl/dwo/**
Het laden van grote databestanden kan soms best even duren. Vaak kun je alleen maar csv-bestanden inlezen. Sla je bestanden goed op, geef ze duidelijke namen en gooi niets weg.

Opgave 1

Wat is het verschil tussen een veldonderzoek en een bureauonderzoek?

Opgave 2

Wat voor werk kan een tool als Excel, DWO, VU-Stat of GeoEnZo voor je uit handen nemen?

Opgave 3

Conclusies onderbouw je met tabellen, diagrammen, en statistische maten.

a Wat is de toegevoegde waarde van tabellen?

b Waarom zou je diagrammen toevoegen?

c Waarom voeg je statistische maten toe?

Opgave 4

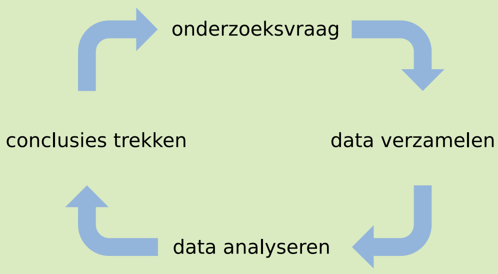
Wat is de bedoeling van een kritische reflectie na afloop van je onderzoek?

Theorie en voorbeelden

Om te onthouden

Bij een **statistisch onderzoek** zoek of verzamel je data zodat je een uitspraak kunt doen over je onderzoeksvraag en een betrouwbare en begrijpelijke onderbouwing kunt geven. Bij de rapportage respecteer je altijd de privacy van deelnemers en verwerk je de gegevens altijd anoniem.

Bij een onderzoek is het essentieel dat je een goede **onderzoeksvraag** hebt. Hierbij zoek je betrouwbare **data**. Deze data verwerk je tot een overzichtelijk geheel. Ze vormen de onderbouwing van het antwoord op je onderzoeksvraag.



Figuur 6.3

Je trekt **conclusies** die je onderbouwt met tabellen, grafieken, statistische maten en redeneringen.
Verwerk alles in een mooi rapport, een heldere presentatie of een overzichtelijke poster.

Het is gebruikelijk een goed onderzoek te beëindigen met een **kritische reflectie**.

Voorbeeld 1

Op 28 juli 2014 vielen grote hoeveelheden neerslag in Nederland, op sommige plaatsen ontstond wateroverlast.

De KNMI vroeg zich af: “Hoe vaak komt extreme neerslag zoals op 28 juli tegenwoordig voor, en is dat anders dan vroeger?”

Bekijk het artikel op de **KNMI** website.

Conclusie van KNMI na statistisch onderzoek:

“We hebben op vier verschillende manieren uitgerekend hoe veel vaker intense buien, zoals die op 28 juli 2014 waargenomen zijn, nu voorkomen ten opzichte van het midden van de twintigste eeuw. Het aantal dagen met 50 mm per dag of meer per zomer of per jaar is nu twee keer zo groot als rond 1950 volgens een eenvoudige trendanalyse. We beschouwen ook de kans op zo’n hoge waarde als op 28 juli binnen de verdeling van waarnemingen met de hoogste etmaalsom van het jaar op één van de ongeveer 325 neerslagstations. Een extreme-waarden analyse geeft dat die kans minstens 1,5 keer groter geworden is. Binnen de verdeling van de hoogste dagsommen van het jaar op alle stations is de kans op zo’n waarde nu een factor 2,0 tot 2,6 groter onder wat aannames. Tenslotte geeft het verband met de Clausius-Clapeyron relatie een factor 1,3 tot 2,4 meer kans. Alles wijst er dus op dat deze buien nu ongeveer twee keer vaker voorkomen dan rond 1950.”

Opgave 5

Bekijk **Voorbeeld 1** en lees het bijbehorende artikel.

a Welke soort onderzoeksvraag hoort bij dit voorbeeld?

b Wat is de conclusie van het KNMI naar aanleiding van het onderzoek?

c Naar welke statistische variabele hebben ze met name gekeken?

d Waarom hebben ze niet alleen maar gekeken naar de gemiddelde neerslag op een dag?

e Wat maakte het complex bij de verwerking van de meetgegevens?

f Tot welke periode van het jaar hebben ze zich beperkt en waarom?

Voorbeeld 2

Vlak voor het WK in 2014 in Brazilië kwam de vraag naar boven: “Wie is de beste keeper van de eredivisie?”

Een journalist onderzocht de topkeepers. Zijn onderzoeksvraag was: “Wat zijn de verschillen tussen de keepers in de eredivisie in het tegenhouden van ballen?”

Bekijk ook het artikel **De beste keeper...**.

Conclusie van de journalist na statistisch onderzoek:

“Jasper Cillessen als beste keeper van Nederland, dat strookt met onze intuïtie. Maar Kristoffer Nordfeldt van Heerenveen als een na beste keeper? Klopt dat ook? Het is in ieder geval contraintuïtief, maar wellicht tonen de cijfers iets wat we nog niet eerder wisten: dat Kristoffer Nordfeldt een ondergewaardeerd supertalent is.”

Interessant was de toegevoegde reflectie op het statistisch onderzoek naar ballen tegenhouden:

“Maar het kan ook zijn dat toeval hier een grotere rol speelt dan we denken. Als je een munt tien keer opgooit, heb je meestal vijf keer kop of munt, maar soms ook acht keer kop of zelfs negen keer munt. Precies zo kun je na een seizoen eredivisie uitkomen op Cillessen, Nordfeldt en Van Dijk als keepers 1, 2 en 3. De steekproef – 34 wedstrijden, ergens tussen de 300 en 600 schoten – is wellicht niet groot genoeg.”

En tot slot:

“... het grootste onderscheid tussen een professionele keeper en een amateur is dat de amateur veel minder ballen stopt. Maar een professionele keeper onderscheiden van andere professionele keepers, dat is een stuk moeilijker – en lukt in elk geval *niet* aan de hand van hun reddingen, want die blijken vrij willekeurig te zijn en per seizoen te verschillen.”

Opgave 6

Bekijk **Voorbeeld 2** en lees het artikel.

a Welke soort onderzoeksvraag hoort bij dit voorbeeld?

b Welke conclusie trekt de journalist?

c Naar welke statistische variabelen is gekeken in dit onderzoek?

d Wat zegt het spreidingsdiagram in dit artikel?

e Welke statistische variabelen zouden volgens de onderzoeker ook bekeken moet worden om het verschil tussen topkeepers te kunnen bepalen?

f Wat is er lastig aan een onderzoek naar de statistische variabelen die bij onderdeel e genoemd zijn?

Toepassen

Je gaat nu een eigen statistisch onderzoek uitvoeren. Belangrijk is dat je over een goede onderzoeksvraag en over genoeg data beschikt. Bespreek dit met je docent.

Hieronder staan voorbeeldonderwerpen voor onderzoek via ‘CBS in de klas’, naar hardlooptijden en een vergelijking van wielerrondes. Het kan natuurlijk goed zijn dat je een ander onderwerp veel geschikter vindt voor een onderzoek. De opgaven hieronder kun je dus als opstapje/voorbeeldvragen gebruiken voor je eigen onderzoek. Er staan geen antwoorden en uitwerkingen bij die vragen.

Figuur 6.4

Opgave 7: Statistisch onderzoek: Leren met het CBS

Via **Leren met het CBS** kun je een aantal grote databestanden vinden om statistisch onderzoek mee te doen.

a Kies één van de beschikbare databestanden en bedenk daarbij een goede onderzoeksvraag.

b Voer het onderzoek uit, maak geschikte tabellen en diagrammen, bereken passende centrum- en spreidingsmaten en trek conclusies. Formuleer een antwoord op je onderzoeksvraag.

Opgave 8: Chicago marathon 2016

Elke grote stad kent een marathon. Zo ook de stad Chicago. De resultaten van de Chicago Marathon 2016 (ruim 40000 atleten) staan in **dit Excel-document**. Je kunt nu aan de slag met grafieken maken, statistische maten berekenen, groepen vergelijken.

a Welke statistische variabelen zie je in dit document?

b Kijk naar de tijden op de finish. Bereken gemiddelde en mediaan van de finishtijden van de lopers.

c Bereken gemiddelde snelheden van alle lopers.

d Er is een kolom in het document die aangeeft in welke leeftijdsklasse de loper hoort. Geef een histogram van de frequentieverdeling van de leeftijden van de lopers. Verwacht je de vorm van een normale verdeling? Motiveer je antwoord.

e Welke vragen zou je allemaal nog kunnen beantwoorden met behulp van deze gegevens?

f Kies een eigen onderzoeksvraag en voer een statistisch onderzoek uit.

Opgave 9: Eindexamenresultaten

Je carrière in het voortgezet onderwijs wordt normaal gesproken afgesloten met een centraal schriftelijk eindexamen voor de meeste vakken. Maar er zijn ook een schoolexamencijfers die mede het eindcijfer bepalen.

In het bestand **HAVO-examenresultaten** zie je de schoolexamencijfers (SE) en de cijfers voor het centraal examen (CE) van alle kandidaten van een school voor HAVO in een bepaald jaar. Je kunt van alles onderzoeken.

a Is er een verband tussen de SE-cijfers en de CE-cijfers van een bepaald vak?

b Is er een significant verschil tussen de SE-cijfers en de CE-cijfers van een bepaald vak?

c Vergelijk de resultaten van twee vakken met elkaar en zoek naar een verband.

d Zijn er verschillen tussen mannelijke en vrouwelijke examenkandidaten voor een bepaald vak?

e En je kunt vast nog wel meer vragen verzinnen.
Vooral is het veel leuker om te kijken of je aan (geanonimiseerde) gegevens van jouw eigen school kunt komen.

Register

Het lesmateriaal in deze reader is gebaseerd op het materiaal dat ook op de Math4All website staat.

De reader is gegenereerd met de Math4All maatwerkdienst. De inhoud en de volgorde van de onderwerpen in deze reader zijn gekozen door docenten van het ConTeXt College.

Stichting Math4All

Inhoud Katern 2

www.math4all.nl

Werkblad bij Opgave 4 op pagina 109.