Wiskunde A

6 VWO

Katern 1

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

Periodieke functies

1.1 Periodiciteit

Inleiding

Er zijn veel verschijnselen die zich herhalen in tijd of ruimte. Bijvoorbeeld een muzieknummer dat je afspeelt met herhaling, een voetbalfragment dat telkens wordt herhaald, zomer en winter, de dagen van de week, rondjes in een reuzenrad. Je noemt dat: periodieke verschijnselen. Als het verschijnsel ook nog met een functie te beschrijven is, spreek je van een periodieke functie. Als je één geschikt stukje kent (de periode) kun je het vervolg helemaal voorspellen. In dit onderdeel zul je dat voor verschillende soorten van periodieke functies doen.

Figuur 1.1

Je leert in dit onderwerp

- de periode vaststellen van een periodiek verschijnsel;
- berekeningen maken bij periodieke verschijnselen.

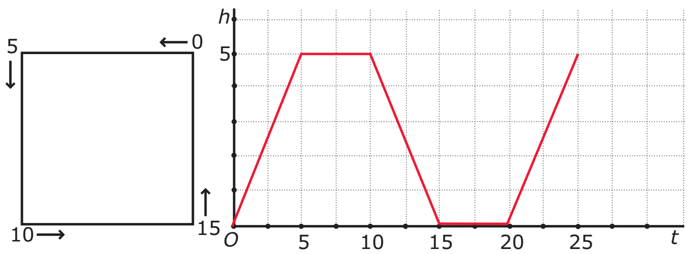
Voorkennis

- grafieken tekenen met grote getallen op de assen;
- lineaire vergelijkingen oplossen;
- vergelijkingen oplossen met twee of meer oplossingen.

Verkennen

Opgave V1

Een mier loopt op een verticaal vlak vierkantjes: 5 cm omhoog, 5 cm horizontaal, 5 cm recht naar beneden, 5 cm horizontaal, enz. De snelheid is constant 1 cm/s. Hier zie je de grafiek van de hoogte h boven de horizontale as afhankelijk van de tijd t.



Figuur 1.2

a Leg uit waarom de grafiek er zo uitziet.

b Welke periode heeft de grafiek?

c Hoe hoog zit de mier na 614 seconden?

Uitleg 1

Stel het is vandaag maandag. Hoe bepaal je welke dag het over 100 dagen is?

Bij de dagen van de week is er sprake van een herhaling per 7 dagen. Als er zo'n herhaling optreedt, noem je dit een periodiek verschijnsel: na een bepaalde periode begin je als het ware weer van voor af aan. In dit geval heeft het periodiek verschijnsel een periode van 7 dagen.

Figuur 1.3

Geef de eerste maandag dagnummer 0. Het is opnieuw maandag op dagnummer 7 en 14 en op elk veelvoud van 7. Omdat je weet dat het over 98 (14×7) dagen ook maandag is, weet je dat het over 99 dagen dinsdag is en over 100 dagen woensdag.

Kies in dit soort situaties een basispatroon dat zich telkens herhaalt. Dat kan zijn maandag tot en met zondag (7 dagen), maar ook zondag tot en met zaterdag (ook 7 dagen) of donderdag tot en met woensdag.

Opgave 1

Bekijk **Uitleg 1**.

a Waarom is het 1071 dagen na een maandag weer een maandag?

b Stel dat het vandaag donderdag is. Wat voor dag is het 198 dagen later?

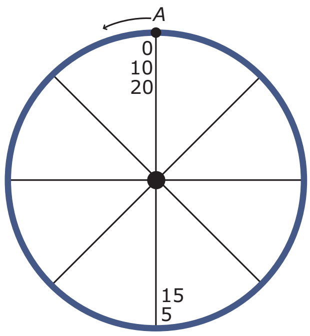
c Stel dat het vandaag zaterdag is. Wat voor dag was het 298 dagen geleden?

d Stel dat 12 maart op een maandag valt, op wat voor dag valt dan 2 mei van hetzelfde jaar?

Uitleg 2

Bekijk de applet

Een wiel draait met steeds dezelfde snelheid rond (dit wordt een eenparige cirkelbeweging genoemd) en maakt één omwenteling in 10 seconden. Op tijdstip $t = 0$ staat punt A precies bovenaan. Het wiel draait linksom. Op welke tijdstippen staat punt A weer bovenaan?



Figuur 1.4

Omdat het wiel in 10 seconden ronddraait, bevindt punt A zich ook bovenaan op $t = 0,10,20,...$ (elk veelvoud van 10).
Als het wiel al aan het draaien was, bevindt punt A zich ook bovenaan op $t = -10, -20,...$
Dit kun je schrijven als: $t = 0 + k \cdot 10$
met k een geheel getal, of korter: $t = k \cdot 10$. k is een soort teller.
Als $k = 1$, dan is $t = 1 \cdot 10 = 10$,
als $k = 2$, dan is $t = 2 \cdot 10 = 20$.
als $k = -2$, dan is $t = -2 \cdot 10 = -20$.

Op welke tijdstippen staat punt A helemaal rechts?

Dit gebeurt op drie vierde van een omwenteling, dus op 7,5 seconden en elke 10 seconden vroeger of later weer: $t = ...; -12{,}5; -2{,}5; 7{,}5; 17{,}5; ...$
Dus op: $t = 7{,}5 + k \cdot 10$ met k een geheel getal.

Omdat de periode van draaiing 10 seconden is, draait het wiel per seconde $\frac{1}{10}$ deel.
Dit heet de frequentie van de draaiing.

Opgave 2

Bekijk in **Uitleg 2** nog eens het verhaal van het ronddraaiende wiel. Ga er van uit dat het punt A in 10 seconden rond draait en op $t = 0$ bovenaan zit.

a Op welke tijdstippen zit punt A helemaal links?

b Beschrijf waar punt A zit op $t = 7 + k \cdot 10$ (teken eventueel zelf zo'n wiel, de grootte is onbelangrijk).

c De frequentie van draaiing is $\frac{1}{10}$ per seconde.
Hoe groot is de frequentie van dit wiel per minuut?

Theorie en voorbeelden

Om te onthouden

De **periode** van een periodiek verband is de lengte van het kortste interval waarin iets zich herhaalt.

In de grafiek is een herhalende golfbeweging zichtbaar. Er zijn verschillende intervallen voor de periode mogelijk:

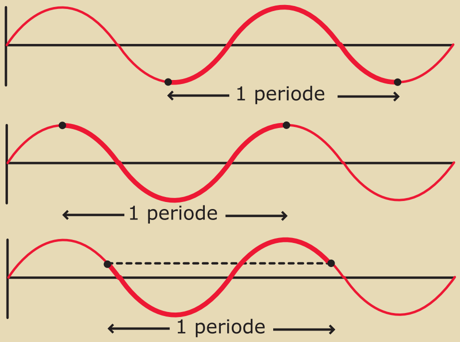
- van een maximum tot het eerstvolgend maximum;
- van een minimum tot het eerstvolgend minimum;
- een willekeurig punt tot het volgende punt waarop de grafiek dezelfde hoogte én dezelfde daling of stijging heeft als in het andere punt.

Voor de lengte van het interval waarop de grafiek zich herhaalt is het niet belangrijk uit welk deel van de grafiek het interval komt. Die lengte is altijd gelijk als de grafiek zuiver periodiek is. Soms wordt de periode ook wel **golflengte** genoemd.

Als op tijdstip t een bepaalde uitkomst voorkomt, komt diezelfde uitkomst ook voor op $t + k \cdot$ periode, waarin k een geheel getal is.

Het aantal periodes per tijdseenheid heet de **frequentie**:

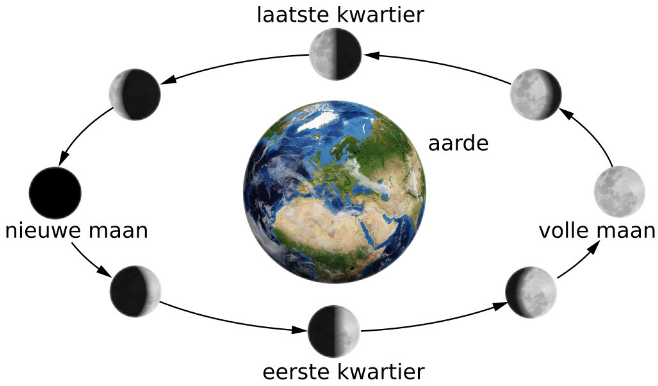
frequentie = $\frac{1}{\text{periode}}$.



Figuur 1.5

Voorbeeld 1

De schijngestalten van de maan staan bekend als nieuwe maan, eerste kwartier (halve maan), volle maan en laatste kwartier (weer halve maan). Dit is een periodiek verschijnsel met een periode van gemiddeld ongeveer 30 dagen (in de loop van het jaar wisselt de lengte een beetje).



Figuur 1.6

Op 12 januari 2017 was het volle maan.

- Bereken op welke datum in mei 2017 het ook volle maan is. Ga hierbij uit van een periode van gemiddeld 30 dagen.
- Bereken de jaarlijkse frequentie van nieuwe maan.

Antwoord

Tel 30 dagen op bij 12 januari, dat geeft 11 februari. Dit proces herhalend geeft 13 maart, 12 april en 12 mei. Op 12 mei 2017 was het ook volle maan. De volgende datum is in juni.

De jaarlijkse frequentie van nieuwe maan is $\frac{1}{\frac{30}{365}} = \frac{365}{30} \approx 12{,}2$ per jaar.

Opgave 3

Bekijk **Voorbeeld 1** over de schijngestalten van de maan.

a Hoe groot is de periode van dit verschijnsel?

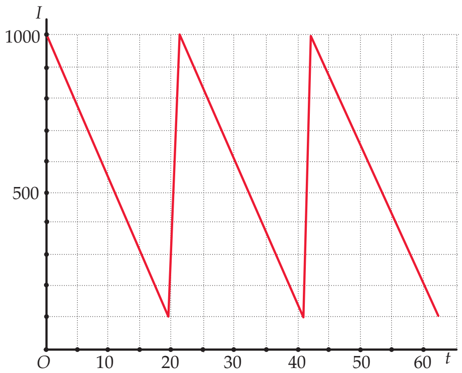
b Met welke frequentie is het volle maan? Bereken het aantal herhalingen per half jaar.

c Op welke datum is het volle maan in januari 2018, uitgaande van volle maan op 12 mei 2017?

Voorbeeld 2

Een opslagtank bevat 1000 liter brandstof op dag $t = 0$. In 20 dagen neemt die hoeveelheid gelijkmatig af tot 100 liter. Dan wordt de tank in een dag bijgevuld tot 1000 liter, enzovoort.

Hoeveel liter brandstof bevat de tank na 75 dagen?



Figuur 1.7

Antwoord

De periode van de inhoud is 21 dagen.
Op de tijdstippen $t = 75 + k \cdot 21$ is er een gelijke inhoud als op dag 75.
Op dag 12 heeft de tank dezelfde inhoud als op dag 75, want hier zitten precies drie periodes tussen ($75 - 3 \cdot 21 = 12$).
In 20 dagen gaat er 900 liter uit de tank, dat is 45 liter per dag.
Van $t = 0$ naar $t = 12$ gaat er $45 \cdot 12 = 540$ liter uit.
Er was 1000 liter. Er is $1000 - 540 = 460$ liter over op $t = 12$, en derhalve ook op $t = 75$.

Opgave 4

Bekijk **Voorbeeld 2**. De hoogte van de brandstof in de tank is een periodiek verschijnsel.

a Hoeveel bedraagt de periode?

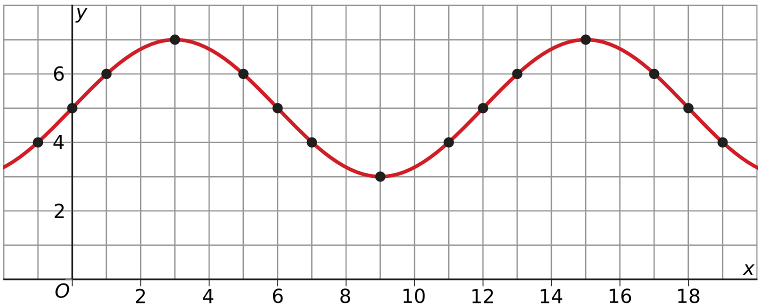
b Leg uit waarom er 550 liter in de tank zit op:
$t = 10 + k \cdot 21 \vee t = 20{,}5 + k \cdot 21$.

c Voor welke waarden van t zit er 100 liter in de tank?

d Hoeveel zit er in de brandstoftank na 500 dagen?

Opgave 5

Bekijk de grafiek van de periodieke functie f. De grafiek loopt naar beide kanten oneindig ver door.



Figuur 1.8

a Bepaal de periode van deze functie.

b Bepaal $f(81)$ en $f(91)$.

c Los op: $f(x) = 6$ met $75 \leq x \leq 85$.

d Bepaal $f(-5)$.

e Los op: $f(x) = 4$ met $-100 \leq x \leq -90$.

Voorbeeld 3

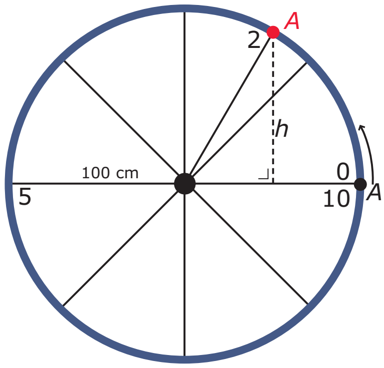
Bekijk de applet: waterrad

Hier is schematisch een waterrad weergegeven dat in 10 seconden ronddraait.

Punt A bevindt zich helemaal rechts op de cirkel op het tijdstip $t = 0$.
Gegeven is dat de straal MA van het waterrad 100 centimeter is.
Het waterrad draait tegen de wijzers van de klok in.

h is de hoogte van punt A ten opzichte van de as van het rad, zodat op tijdstip $t = 0$ de hoogte ook 0 centimeter is.

Hoe hoog is het punt op tijdstip $t = 42$?



Figuur 1.9

Antwoord

De periode van het waterrad is 10 seconden. De hoogte op $t = 42$ is hetzelfde als de hoogte op $t = 2$.

Punt A heeft dan $\frac{2}{10}$ van de cirkel doorlopen en is $\frac{2}{10} \cdot 360° = 72°$ gedraaid.
Voor het berekenen van de hoogte heb je de sinus nodig:

$\sin(72°) = \frac{h}{100}$ en hieruit volgt $h = 100 \cdot \sin(72°) \approx 95{,}1$ cm.
Op $t = 42$ is de hoogte ongeveer 95 cm.

Opgave 6

Bestudeer **Voorbeeld 3** over het ronddraaiende punt A.

a Bereken de hoogte h als $t = 1$. Rond af op gehele centimeters.

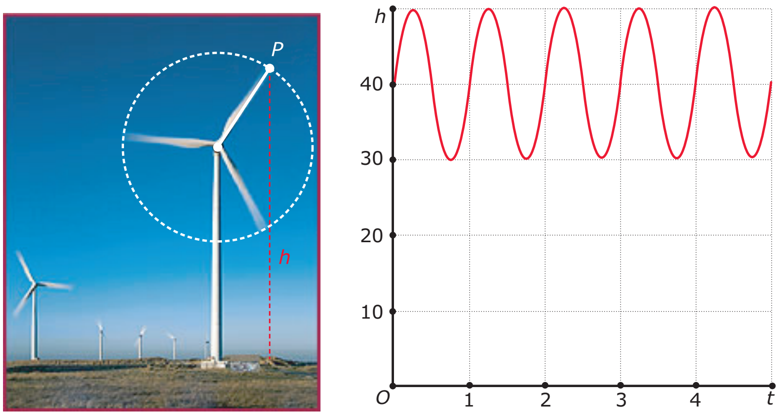
b Hoe groot is de hoogte h als $t = 31$?

c Bereken de hoogtes voor de gehele waarden van t vanaf 0 tot en met 10.
Rond indien nodig af op gehele centimeters.

d Met welke frequentie draait dit waterrad?
Ga uit van een tijdseenheid van een uur.

Opgave 7

Bekijk het punt P op de tip van een rotorblad van een ronddraaiende windmolen. De grafiek van de functie $h(t)$ is getekend, waarin h de hoogte van punt P boven de grond in meter voorstelt en t de tijd in seconden.



Figuur 1.10

a Met welke periode draait het rotorblad van de windmolen?
Met welke frequentie (omwentelingen per minuut) draait het rotorblad?

b Hoe hoog zit de as van de windmolen boven de grond?
En hoe lang is het rotorblad?

c Teken de grafiek van de hoogte van de tip van één van de twee andere rotorbladen.

d De wind neemt af, de windmolen gaat een half keer zo snel draaien.
Teken de bijbehorende grafiek.

Opgave 8

Een punt P beweegt linksom over een cirkel met straal 1 om de oorsprong O van een Oxy-assenstelsel. De afstand a die het punt heeft afgelegd hangt af van de hoek α waarover OP is gedraaid. Neem aan dat $a = 0$ als $\alpha = 0$.

a Hoeveel is $a(90°)$? En $a(180°)$? (Geef exacte waarden.)

b Leg uit waarom je nu te maken krijgt met hoeken die groter zijn dan $180°$. Leg ook uit waarom de draaihoek zelfs groter kan zijn dan $3690°$.

c Wat zou een draaihoek van $-60°$ betekenen?

d Bepaal nu $a(360°)$, $a(450°)$, $a(60°)$ en $a(-30°)$.

e Hoeveel is $a(1°)$?

f Is $a(\alpha)$ een periodieke functie? Licht je antwoord toe.

Verwerken

Opgave 9

Een draaimolen draait steeds met dezelfde snelheid in 20 seconden rond.

a Hoeveel bedraagt de frequentie van de draaiing per minuut?

b Hoeveel bedraagt de frequentie van de draaiing per uur?

c Hoeveel bedraagt de frequentie van de draaiing per seconde?

Opgave 10

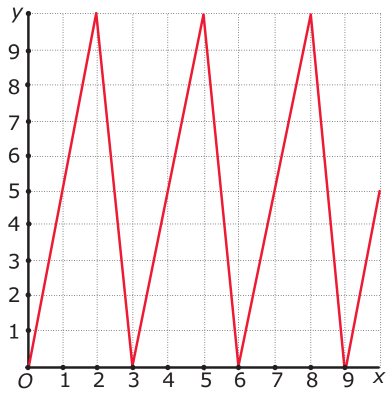
Bekijk de grafiek van de periodieke functie f.

a Hoeveel bedraagt de periode van de grafiek?

b Bereken $f(25)$.

c Voor welke waarden van x is $f(x) = 10$?

d Los op: $f(x) = 5$ met $12 \leq x \leq 15$.



Figuur 1.11

Opgave 11

Bekijk de figuur. Punt A ligt op een wiel op afstand 1 van het middelpunt. Noem de hoogte van punt A ten opzichte van de horizontale as door het middelpunt $h(t)$. Punt A begint rechts, zodat $h(0) = 0$. Het wiel draait in 6 seconden linksom rond.

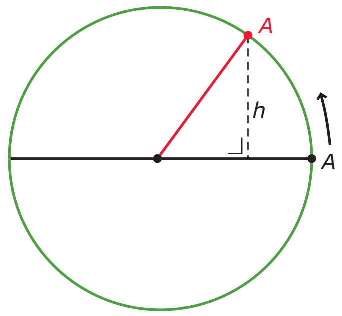
a Hoeveel bedraagt de frequentie van de draaiing per minuut?

b Bepaal $h(1{,}5)$.

c Bepaal $h(4{,}5)$ en $h(10{,}5)$.

d Bereken $h(0{,}5)$.

e Geef alle waarden van t die voldaan aan de vergelijking $h(t) = h(0{,}75)$.



Figuur 1.12

Opgave 12

Een torenklok heeft een grote wijzer met een lengte van $1{,}5$ m. De beide wijzers zitten bevestigd op de as van de klok op 45 m boven de grond. Punt T stelt de tip van deze grote wijzer voor. De hoogte h in m van T boven de grond hangt af van de draaihoek α. Neem aan dat $\alpha = 0$ om 12:00 uur.

a Hoe hoog zit T boven de grond op 2:10 uur?

b Schets een grafiek van $h(\alpha)$.

Opgave 13

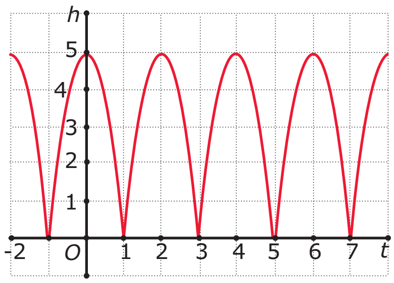
Bekijk de periodieke grafiek. Voor $-1 \leq x \leq 1$ geldt de formule: $h(t) = 5 - 5t^2$

a Bereken $h(0)$ en $h(0{,}5)$.

b Bepaal de periode van de grafiek.

c Bereken $h(6)$ en $h(6{,}5)$.

d Bereken $h(15)$ en $h(15{,}5)$.

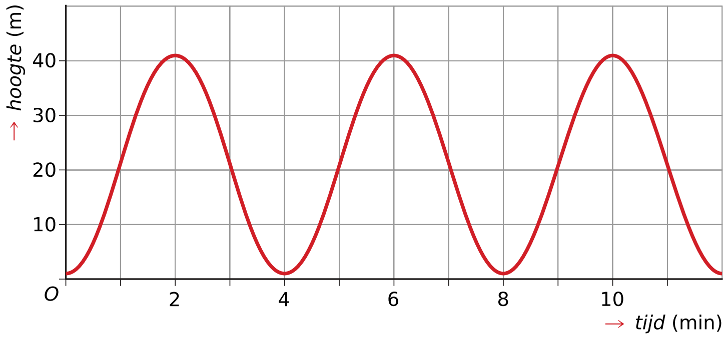


Figuur 1.13

Toepassen

Opgave 14: Reuzenrad

Peter stapt in een reuzenrad. Bekijk de grafiek waarin de hoogte van Peter ten opzichte van de grond is af te lezen na t minuten.



Figuur 1.14

a In hoeveel minuten draait het reuzenrad rond?

b Hoeveel rondjes heeft Peter gemaakt?

c Hoe groot is de straal van het reuzenrad?

Opgave 15: Hartslag

Op de intensive care van een ziekenhuis wordt met hartbewakingsapparatuur de hartfunctie van patiënten bewaakt. Bekijk de grafische weergave van de hartslag van een patiënt. De hartslagfrequentie wordt uitgedrukt in het aantal slagen per minuut.

Figuur 1.15

a Hoe groot is de hartslagfrequentie van deze patiënt?

b Wat gebeurt er met de periode van de grafiek als de hartslagfrequentie omhoog gaat?

Testen

Opgave 16

Een grafiek bestaat in een Oxy-assenstelsel uit rechte lijnstukjes tussen de punten $(100,1000)$, $(110,600)$, $(140,1000)$, $(150,600)$, $(180,1000)$, enz. Het patroon gaat naar links en rechts oneindig ver door.

a Welke periode heeft deze grafiek?

b Bereken de waarde van y bij $x = 250$.

c Teken de grafiek met $0 \leq x \leq 100$.

d Bereken de waarde van y bij $x = -250$.

e Hoeveel getallen x met $0 \leq x \leq 100$ bestaan er bij $y = 900$?

Opgave 17

Een wiel met een straal van 30 cm draait linksom rond met constante snelheid. De omlooptijd is 20 seconden. De hoogte in centimeter van punt A aan de buitenkant van het wiel, gemeten ten opzichte van het middelpunt, noem je $h(t)$ met t in seconden. Het punt A begint bovenaan, dus $h(0) = 30$.

a Met welke frequentie draait punt A?

b Bereken $h(35)$.

c Bereken $h(18)$ in één decimaal nauwkeurig.

d Bereken $h(76)$ in één decimaal nauwkeurig.

e Geef alle tijdstippen t met $-40 \leq t \leq 40$ waarvoor geldt $h = 0$.

1.2 Sinusfunctie

Inleiding

Een regelmatige cirkelbeweging is een belangrijk periodiek verschijnsel. De hoogte van een punt op een draaiende cirkel, uitgezet tegen de tijd, levert een heel regelmatige periodieke grafiek: de sinusoïde. Daarbij komt vanzelf een nieuwe manier tevoorschijn om de grootte van een hoek aan te geven: in radialen, in plaats van in graden.

Je leert in dit onderwerp

- de grafiek van $y = \sin(x)$ tekenen met x in graden of in radialen;
- graden omrekenen in radialen en omgekeerd.

Voorkennis

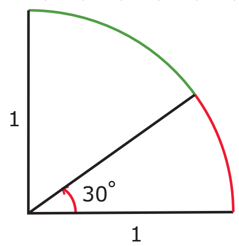
- grafieken tekenen met de grafische rekenmachine;
- werken met sinus in rechthoekige driehoeken.

Verkennen

Opgave V1

Bekijk de applet

Hoeken druk je al heel lang in graden uit. Toch hoeft dat niet, bekijk deze kwartcirkel maar eens. Hij heeft een straal van 1. Er staat een hoek van 30° in getekend.



Figuur 2.1

a Hoe lang is de getekende cirkelboog? Leg uit waarom 30° overeenkomt met een booglengte van $\frac{1}{6}\pi$.

Als je de grootte van een hoek door zijn booglengte in een cirkel met straal 1 beschrijft, krijg je hoeken in radialen. Dus 30° komt overeen met $\frac{1}{6}\pi$ radialen.

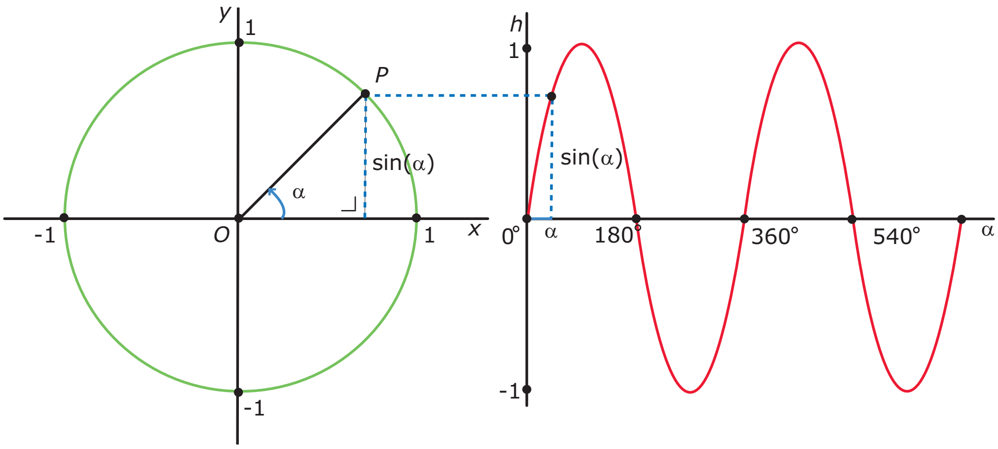
b Waarom is het van belang dat de cirkel waarin je de booglengte uitrekent een straal van 1 heeft?

c Reken maar eens een paar andere hoeken om van graden naar radialen.

Uitleg

Bekijk de applet: sinusgrafiek

Een punt P draait linksom (tegen de klok in) over een cirkel met straal 1, een eenheidscirkel.
Straal OP maakt een hoek α met de horizontale as vanuit O.
De hoogte h van punt P is: $h = 1 \cdot \sin(\alpha) = \sin(\alpha)$.
Dit is precies de y-coördinaat van punt P: $y_P = h = \sin(\alpha)$.



Figuur 2.2

Als punt P zich onder de x-as bevindt, is $y_P = \sin(\alpha)$ negatief.
P kan ook rechtsom (met de klok mee) draaien. Op deze manier is de hoek α negatief. De sinus van zo'n negatieve hoek kan weer positief of negatief zijn.

Naast de eenheidscirkel zie je de grafiek van $y_P = \sin(x)$. De grafiek ontstaat door voor elke draaihoek de sinus van die hoek uit te zetten op de verticale as. In plaats van α in graden wordt de hoek weergegeven door de bijbehorende booglengte x. De eenheid voor deze hoek heet radiaal, afgekort rad.

In een eenheidscirkel is de straal 1 en de omtrek dus $2\pi \cdot 1 = 2\pi$.
Dus bij 360° hoort 2π rad.
En bij 180° hoort π rad.

Bij een booglengte van 1 op de eenheidscirkel hoort een hoek van 1 rad.

Hoeken worden vanaf nu, tenzij anders vermeld, gegeven in radialen. Je ziet:

$\sin(\pi) = 0$, $\sin\left(1\frac{1}{2}\pi\right) = -1$ en $\sin\left(2\frac{1}{2}\pi\right) = \sin\left(2\pi + \frac{1}{2}\pi\right) = 1$.

Om graden om te rekenen naar radialen gebruik je dat $180° = \pi$ rad.

En zo is $40° = \frac{40}{180}\pi = \frac{2}{9}\pi$ rad.

Opgave 1

Bekijk de **Uitleg**. Teken een eenheidscirkel (een cirkel met een straal 1).

a Teken P als de draaihoek $\alpha = 30°$. Bereken de bijbehorende waarde van h. Hoeveel radialen is α?

b Teken Q als de draaihoek $\alpha = 150°$. Bereken h. Hoeveel radialen is α?

c Teken R als de draaihoek $\alpha = 210°$. Bereken h. Hoeveel radialen is α?

d Teken S als de draaihoek $\alpha = 270°$. Bereken h. Hoeveel radialen is α?

e Hoeveel radialen hoort er bij $360°$?

f Bij welke draaihoeken is $h = 1$?
Geef je antwoord in graden en daarna in radialen.

Opgave 2

Teken met je grafische rekenmachine de grafiek van $y = \sin(x)$. Neem x in graden en stel het venster zo in dat $-360 \leq x \leq 720$ en $-1{,}5 \leq y \leq 1{,}5$.

a Hoeveel periodes van de sinusgrafiek krijg je zo in beeld?

b Bereken $\sin(30°)$ en $\sin(390°)$. Leg uit waarom beide uitkomsten gelijk zijn.

c Bereken $\sin(30°)$ en $\sin(150°)$. Leg uit waarom beide uitkomsten gelijk zijn.

d Bij welke waarden van x vind je dezelfde uitkomst als $\sin(30000°)$?

e Bij welke waarden van x vind je dezelfde uitkomst als $\sin(-10000°)$?

Opgave 3

Teken met je grafische rekenmachine de grafiek van $y = \sin(x)$. Neem x in radialen en stel het venster zo in dat $-2\pi \leq x \leq 6\pi$ en $-1{,}5 \leq y \leq 1{,}5$.

a Hoeveel periodes van de sinusgrafiek krijg je zo in beeld?

b Bereken $\sin(1)$ en $\sin(1 + 2\pi)$. Leg uit waarom beide uitkomsten gelijk zijn.

c Bereken $\sin(1)$ en $\sin(\pi - 1)$. Leg uit waarom beide uitkomsten gelijk zijn.

d Bij welke waarden van x vind je dezelfde uitkomst als $\sin(211{,}5\pi)$?

e Bij welke waarden van x vind je dezelfde uitkomst als $\sin(-1500\pi)$?

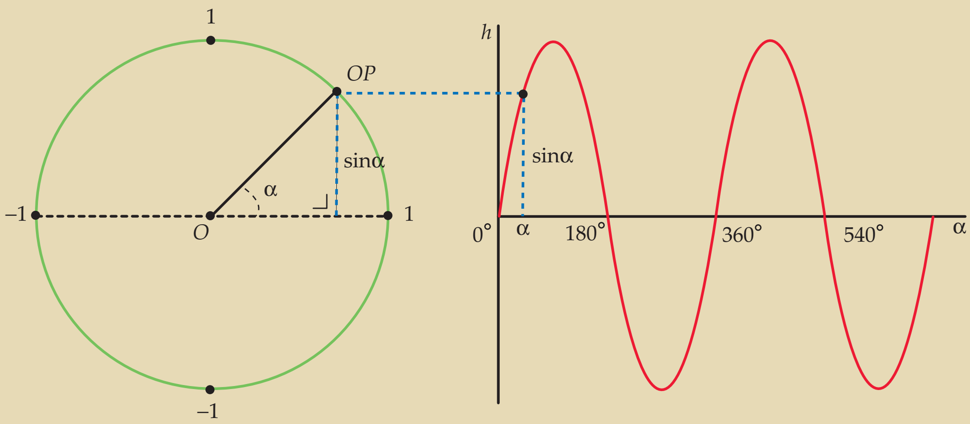
Theorie en voorbeelden

Om te onthouden

Bekijk de **eenheidscirkel** (cirkel met straal 1). Punt P beweegt over de eenheidscirkel en begint in het punt uiterst rechts te draaien. De bijbehorende **draaihoek** x is positief als je linksom draait, negatief als je rechtsom draait en kan alle waarden aannemen.

De draaihoek wordt in **radialen** uitgedrukt, dus door de lengte van de bijbehorende boog op de eenheidscirkel.

Bekijk de applet: sinusgrafiek 2



Figuur 2.3

De hoogte h van P ten opzichte van een horizontale lijn door het middelpunt van de cirkel is $h = \sin(x)$. De grafiek van h heet de **sinusgrafiek** en is periodiek met **periode** 2π.

$180°$ komt overeen met π radialen of π rad, ofwel de halve omtrek van de eenheidscirkel. Tenzij anders aangegeven, wordt de hoekeenheid rad gebruikt. Let op de instelling van de rekenmachine.

Voorbeeld 1

Reken de hoeken in graden om naar radialen tussen 0 en 2π, en omgekeerd.

- $1°$
- $90°$
- 1 rad
- $\frac{1}{6}\pi$ rad

Antwoord

Bij het omrekenen van graden naar radialen gebruik je $180° = \pi$ rad.

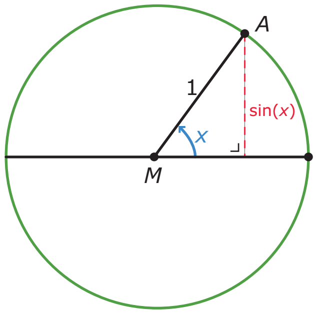
- $1°$ wordt $\frac{\pi}{180} = \frac{1}{180}\pi$ rad.
- $90°$ wordt $90 \cdot \frac{\pi}{180} = \frac{1}{2}\pi$ rad.

En omgekeerd:

- 1 rad komt overeen met $\left(\frac{180}{\pi}\right)^\circ \approx 57°$.
- $\frac{1}{6}\pi$ rad komt overeen met $\left(\frac{1}{6}\pi \cdot \frac{180}{\pi}\right)^\circ = 30°$.

Opgave 4

Punt A beweegt tegen de klok in over een eenheidscirkel met middelpunt M.
α is de draaihoek in graden en x is de lengte van de cirkelboog die bij die draaihoek hoort.



Figuur 2.4

a Hoeveel bedraagt x als $\alpha = 360°$?

b Vul de tabel in.

α	$0°$	$30°$	$45°$	$60°$	$90°$	$120°$	$225°$	$270°$	$330°$
x									

Tabel 2.1

c Hoeveel radialen is $10°$?

d Hoeveel graden is 10 radialen? Rond af op een geheel getal.

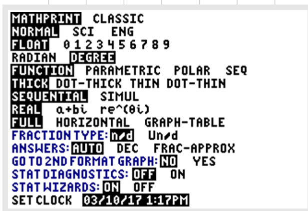
Voorbeeld 2

Bereken met de rekenmachine: $\sin(10°)$, $\sin(100°)$, $\sin(1000°)$ en $\sin(10000°)$

Waarom zijn de laatste twee uitkomsten hetzelfde?

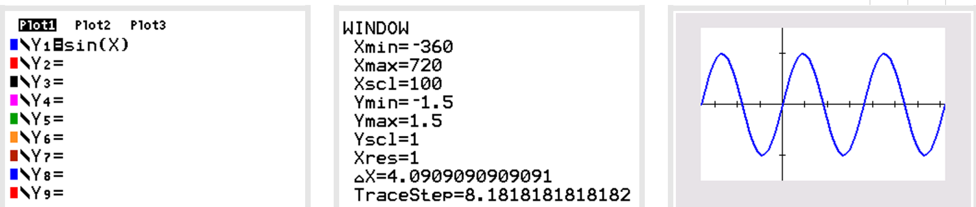
Antwoord

De draaihoek is gegeven in graden. Zorg ervoor dat de rekenmachine met graden rekent.
Ga na dat $\sin(10°) \approx 0{,}174$; $\sin(100°) \approx 0{,}985$; $\sin(1000°) \approx -0{,}985$ en $\sin(10000°) \approx -0{,}985$.



Figuur 2.5

De grafiek van $y = \sin(x)$ met x in graden heeft een periode van $360°$. De laatste twee uitkomsten zijn gelijk omdat het verschil tussen $1000°$ en $10000°$ gelijk is aan $9000° = 25 \cdot 360°$.
Dus dat zijn precies 25 periodes.



Figuur 2.6

Opgave 5

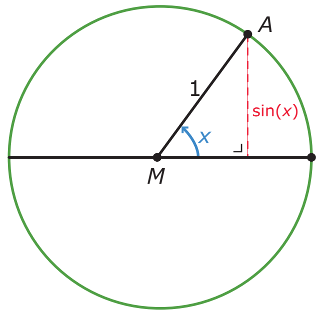
In **Voorbeeld 2** bekijk je de waarden van $\sin(x)$ als x in graden is. Hier zie je nog eens waar je de waarden van $\sin(x)$ in de eenheidscirkel vindt.

a Hoeveel bedraagt de periode van deze sinusfunctie?

b Leg uit waarom $\sin(x) = \sin(x + k \cdot 360°)$.

c Leg uit waarom $\sin(x) = \sin(180° - x)$.

d Bepaal $\sin(45°)$? Voor welke andere hoek x tussen 0° en 360° geldt: $\sin(x) = \sin(45°)$?



Figuur 2.7

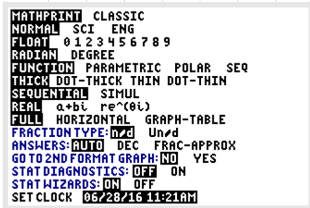
Voorbeeld 3

Bepaal met de rekenmachine $\sin(1)$, $\sin(10)$, $\sin\left(\frac{1}{6}\pi\right)$ en $\sin(360)$.

Bepaal ook $\sin(\pi - 1)$ en licht toe dat $\sin(\pi - 1) = \sin(1)$.

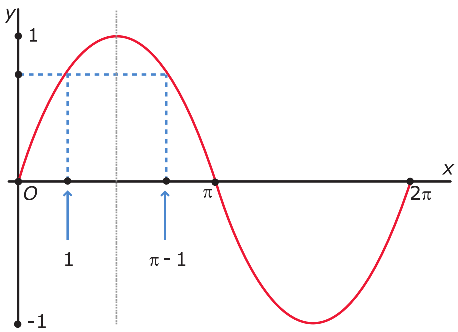
Antwoord

Er wordt nu in radialen gerekend, want er zijn geen gradentekens. Stel de rekenmachine in op radialen.
Ga na dat $\sin(1) = \sin(\pi - 1) \approx 0{,}841$; $\sin(10) \approx -0{,}544$; $\sin\left(\frac{1}{6}\pi\right) = 0{,}5$ en $\sin(360) \approx 0{,}959$.



Figuur 2.8

Bekijk de grafiek van $y = \sin(x)$.
Je ziet dat $\sin(1)$ en $\sin(\pi - 1)$ symmetrisch liggen ten opzicht van de lijn $x = 0{,}5\pi$.
Hieruit volgt $\sin(1) = \sin(\pi - 1)$.
In het algemeen geldt: $\sin(x) = \sin(\pi - x)$.



Figuur 2.9

Opgave 6

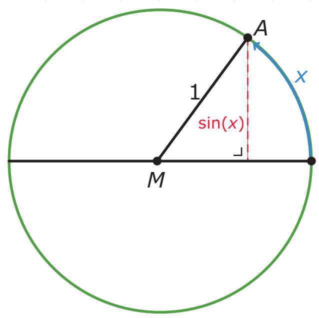
In **Voorbeeld 3** bekijk je de waarden van $\sin(x)$ als x in radialen is. Hier zie je nog eens waar je de waarden van $\sin(x)$ in de eenheidscirkel vindt.

a Hoeveel bedraagt de periode van deze sinusfunctie?

b Leg uit waarom $\sin(x) = \sin(x + k \cdot 2\pi)$.

c Leg uit waarom $\sin(x) = \sin(\pi - x)$.

d Welke waarden kan $\sin(x)$ aannemen?



Figuur 2.10

Opgave 7

Gegeven is $\sin(x) = 0{,}1$ met x in radialen.

a Geef alle mogelijke hoeken x die hieraan voldoen en waarvoor geldt $0 \leq x < 2\pi$ aan in een eenheidscirkel.

b Hoe groot is $\sin(x + 2\pi)$?

c Hoe groot is $\sin(x + \pi)$?

d Hoe groot is $\sin(\pi - x)$?

Verwerken

Opgave 8

De hoeken zijn gegeven in graden. Bereken de bijbehorende booglengtes in de eenheidscirkel in radialen.

a $30°$, $10°$, $270°$, $630°$

De booglengtes in de eenheidscirkel zijn gegeven. Bereken de bijbehorende hoeken in graden nauwkeurig.

b $\frac{1}{2}\pi$; $\frac{1}{3}\pi$; 2; π; $4{,}5$; 10π

Opgave 9

Plot de grafiek van $f(x) = \sin(x)$ op het interval $-2\pi \leq x \leq 4\pi$.

a Bereken $f\left(1\frac{5}{6}\pi\right)$ en $f\left(-\frac{1}{6}\pi\right)$. Verklaar de overeenkomst.

b Bereken $f\left(\frac{1}{4}\pi\right)$ en $f\left(-\frac{1}{4}\pi\right)$. Verklaar de overeenkomst.

Opgave 10

Plot de grafiek van $y = \sin(x)$ op het interval $-3\pi \leq x \leq 8\pi$.

a Hoeveel periodes zijn in beeld?

b Reken na dat $\sin\left(\frac{1}{6}\pi\right) = 0{,}5$.

c Geef een exacte waarde voor x die voldoet aan $\sin(x) = -0{,}5$.

d Is er een waarde voor x zo, dat $\sin(x) = 1{,}5$? Leg uit.

Opgave 11

Gegeven is dat $\sin(x) = 0{,}2$ met x in radialen.

a Welke waarde heeft $\sin(x + 2\pi)$?

b Welke waarde heeft $\sin(x + \pi)$?

c Welke waarde heeft $\sin(\pi - x)$?

d Welke waarde heeft $\sin(x + 11\pi)$?

e Welke waarde heeft $\sin(x + 12\pi)$?

Toepassen

Naast graden en radialen wordt ook de **decimale graad** gebruikt. Deze graad is gedefinieerd als $\frac{1}{400}$ deel van een cirkel en wordt aangegeven met grad (of gra of gon).
Een volledige cirkel is 400 grad.

Opgave 12

Bekijk wat onder decimale graden wordt verstaan.

a Hoeveel grad is $90°$?
En hoeveel grad is $\frac{\pi}{3}$ rad?

b Toon aan dat $x° = \frac{10x}{9}$ grad.

c Geef een omzettingsformule om van radialen naar grad te gaan.

Testen

Opgave 13

Geef de antwoorden exact indien mogelijk, anders in drie decimalen benaderd.

a Deze hoeken zijn gegeven in graden: $60°$, $45°$, $180°$, $300°$, $330°$, $350°$, $-350°$.
Reken om naar radialen.

b Deze booglengtes van een eenheidscirkel zijn gegeven in radialen: π, $\frac{1}{3}\pi$, $-\frac{1}{4}\pi$, 2π, $\frac{5}{6}\pi$, $\frac{13}{12}\pi$, 2, $\frac{5}{3}\pi$.
Bereken de bijbehorende hoeken in graden.

Opgave 14

Bekijk de grafiek van $f(x) = \sin(x)$ op $[-2\pi, 4\pi]$.

a Bereken $f\left(\frac{1}{4}\pi + \frac{1}{3}\pi\right)$ en $f\left(\frac{1}{4}\pi\right) + f\left(\frac{1}{3}\pi\right)$. Verklaar het verschil.

b Bereken $f\left(\frac{1}{4}\pi\right)$ en $f\left(-\frac{3}{4}\pi\right)$. Verklaar de overeenkomst.

c Laat in deze grafiek zien dat $\sin(-x) = \sin(\pi + x)$.

Practicum: Eenheidscirkel

In deze eenheidscirkel vind je de waarden van sinus en cosinus van een hoek in graden of in radialen in twee decimalen nauwkeurig.

Bekijk de applet

1.3 Vergelijkingen met sinus

Inleiding

Met periodieke functies kun je, door de herhaling, gemakkelijk voorspellingen doen. Dezelfde uitkomsten komen met een vaste regelmaat weer terug. Dit bemoeilijkt echter het terugrekenen. Zelfs een eenvoudige vergelijking als $\sin(x) = 0{,}5$ heeft in principe oneindig veel oplossingen. Hoe vind je die allemaal en hoe schrijf je ze allemaal netjes op? Daarover gaat dit onderdeel. Bij de grafiek van $y = \sin(x)$ neem je x altijd in radialen.

Je leert in dit onderwerp

- de vergelijking $\sin(x) = c$ oplossen als c een constante is.

Voorkennis

- een grafiek tekenen op de grafische rekenmachine en dan uit de grafiek x vinden, als y gegeven is;
- symmetrie in grafieken gebruiken.

Verkennen

Opgave V1

Gebruik deze grafiek van $f(x) = \sin(x)$ en de symmetrie ervan.

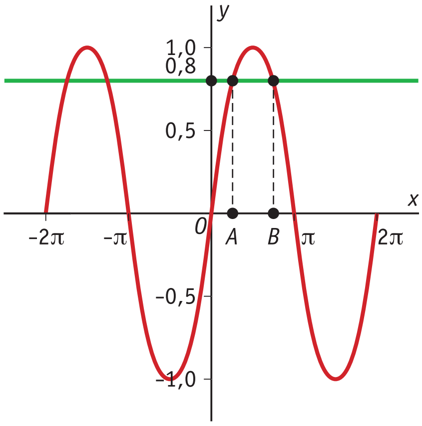
Figuur 3.1

a Los op: $\sin(x) = 0{,}8$ met $0 \leq x \leq 4\pi$. Rond je antwoord af op drie decimalen.

b Los op: $\sin(x) = 0{,}8$ voor elke mogelijke waarde van x. Rond je antwoord af op drie decimalen.

Uitleg

Bekijk de applet.



Figuur 3.2

Bekijk de grafiek van $y = \sin(x)$ en de lijn $y = 0{,}8$.

De lijn $y = 0{,}8$ snijdt de grafiek van $y = \sin(x)$ meerdere malen, er zijn dus meerdere oplossingen voor de vergelijking $\sin(x) = 0{,}8$. Per periode zijn er twee oplossingen. Als je voor x alle getallen toelaat zijn er zelfs oneindig veel oplossingen.

Je wilt $\sin(x) = 0{,}8$ oplossen:

- Zoek eerst de oplossing die zo dicht mogelijk bij de y-as ligt. Deze oplossing heet arcsinus van 0,8: $x = \arcsin(0{,}8) \approx 0{,}927$. Je (grafische) rekenmachine heeft een knop die de arcsinus berekent, meestal aangeduid door $\sin^{-1}$.
- Bepaal daarna de andere oplossing in dezelfde periode. Je gebruikt de symmetrie van de sinusgrafiek.
  Die oplossing is: $x = \pi - \arcsin(0{,}8) \approx 2{,}214$.
- Gebruik de periode om alle oplossingen op te schrijven.
  De periode is 2π, de oplossingen zijn:
  $x \approx 0{,}927 + k \cdot 2\pi \vee x \approx 2{,}214 + k \cdot 2\pi$ met k een geheel getal.
  Op het interval $-2\pi \leq x \leq 2\pi$ zijn de oplossingen:
  $x \approx -5{,}356 \vee x \approx -4{,}069 \vee x \approx 0{,}927 \vee x \approx 2{,}214$.

Bijzondere gevallen zijn:

$\sin(x) = 1$ geeft $x = \frac{1}{2}\pi + k \cdot 2\pi$

$\sin(x) = -1$ geeft $x = -\frac{1}{2}\pi + k \cdot 2\pi$

$\sin(x) = 0$ geeft $x = 0 + k \cdot 2\pi \vee x = \pi + k \cdot 2\pi$
Voeg dit samen tot: $x = k \cdot \pi$

Als in $\sin(x) = c$, de c groter is dan 1 of kleiner is dan -1, dan zijn er geen oplossingen.

Opgave 1

Bekijk de **Uitleg**.

a Welke oplossingen heeft $\sin(x) = 0{,}8$ als $0 \leq x \leq 4\pi$?

b Los op: $\sin(x) = 0{,}2$.

c Los op: $\sin(x) = -0{,}2$

Opgave 2

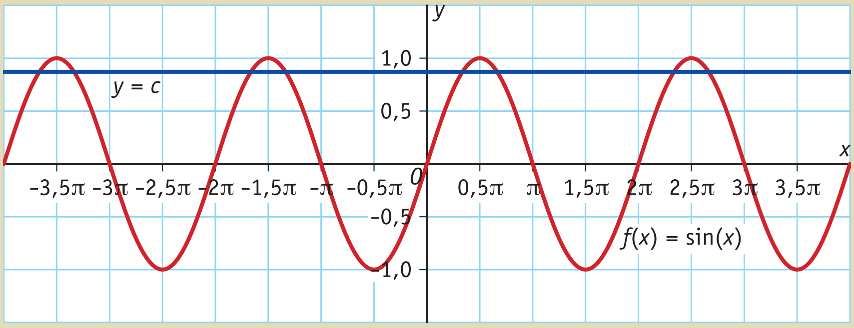
Waarom heeft $\sin(x) = 1{,}2$ geen oplossingen?

Theorie en voorbeelden

Om te onthouden

Bekijk de applet

Bekijk de grafiek van $f(x) = \sin(x)$. De lijn $y = c$ snijdt de grafiek van f meerdere malen. Er zijn dus meerdere oplossingen voor de vergelijking $\sin(x) = c$. Per periode zijn er twee oplossingen. Als het domein $\mathbb{R}$ is, dan zijn er zelfs oneindig veel oplossingen.



Figuur 3.3

De oplossing van $\sin(x) = c$ die het dichtst bij de y-as ligt, heet de **arcsinus** van c: $x = \arcsin(c)$. Binnen één periode is er (vaak) nog een oplossing, alleen de vergelijkingen $\sin(x) = 1$ of $\sin(x) = -1$ hebben slechts één oplossing per periode.

De periode van $y = \sin(x)$ is 2π. Gebruik dit om alle oplossingen te vinden.

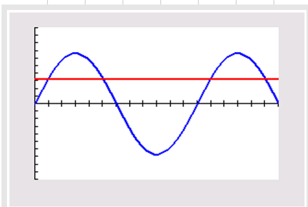
De vergelijking $\sin(x) = c$ heeft alleen oplossingen als $-1 \leq c \leq 1$.

Voorbeeld 1

Los op: $\sin(x) = 0{,}5$ met $0 \leq x \leq 3\pi$.

Antwoord

Plot eerst met de grafische rekenmachine de grafieken van $y_1 = \sin(x)$ en $y_2 = 0{,}5$ op het gegeven interval.



Figuur 3.4

Een oplossing van de vergelijking is: $x = \arcsin(0{,}5) \approx 0{,}524$.
In de grafiek zie ja dat er binnen de eerste periode nog een oplossing is, namelijk $x = \pi - \arcsin(0{,}5) \approx 2{,}618$.

Je ziet ook dat de lijn $y = 0{,}5$ de grafiek van $y = \sin(x)$ op $0 \leq x \leq 3\pi$ nog twee keer snijdt. Er zijn dus nog twee andere oplossingen.
De periode van $y = \sin(x)$ is 2π, zodat de andere twee oplossingen ongeveer $0{,}524 + 2\pi$ en $2{,}618 + 2\pi$ zijn.
De vier oplossingen: $x \approx 0{,}524 \vee x \approx 2{,}618 \vee x \approx 6{,}807 \vee x \approx 8{,}901$.

Opgave 3

Bekijk **Voorbeeld 1**.

a Bereken alle oplossingen van $\sin(x) = 0{,}5$.

b Los nu op $\sin(x) = -0{,}5$.

Opgave 4

Gegeven is $f(x) = \sin(x)$ met $0 \leq x \leq 3\pi$.

a Los op $\sin(x) = 0{,}6$. Geef benaderingen in twee decimalen nauwkeurig.

b Los op $\sin(x) < 0{,}6$. Geef benaderingen in twee decimalen nauwkeurig.

c Los op $\sin(x) < -0{,}6$. Geef benaderingen in twee decimalen nauwkeurig.

Voorbeeld 2

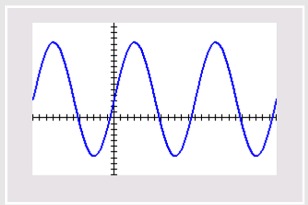
Los op: $3 \cdot \sin(x) + 1 = 0$.

Antwoord

Plot de grafiek van de functie $y = 3 \cdot \sin(x) + 1$ met de grafische rekenmachine. Het gaat bij de vergelijking $3 \cdot \sin(x) + 1 = 0$ om de nulpunten van deze functie, dat zijn er oneindig veel.

De vergelijking $3 \cdot \sin(x) + 1 = 0$ kun je herleiden tot $\sin(x) = -\frac{1}{3}$.

De oplossingen hiervan zijn:
$x \approx -0{,}340 + k \cdot 2\pi \vee x = 3{,}481 + k \cdot 2\pi$



Figuur 3.5

Opgave 5

Bekijk **Voorbeeld 2**.

a Reken de oplossingen van de vergelijking $3 \cdot \sin(x) + 1 = 0$ na.

b Los op: $3\sin(x) + 1 = 2$

c Waarom kun je de vergelijking $3 \cdot \sin(x) + 1 = 5$ niet oplossen?

Opgave 6

Los op: $2 \cdot \sin(x) \leq -1{,}5$ met $-2\pi \leq x \leq 2\pi$.

Verwerken

Opgave 7

Los op.

a $\sin(x) = 0{,}35$

b $\sin(x) = -0{,}35$

Opgave 8

Geef alle oplossingen van:

a $\sin(x) = 1$

b $\sin(x) = \sin(1)$

c $\sin(1) = x$

Opgave 9

Gegeven is de functie f met $f(x) = 2 \cdot \sin(x) - 1$ met $-2\pi \leq x \leq 2\pi$.

a Bereken alle nulpunten van deze functie in twee decimalen.

b Los op: $f(x) = 0{,}5$ in twee decimalen nauwkeurig.

c Wat is de kleinste waarde voor c zodanig dat de vergelijking $f(x) = c$ een oplossing heeft?

Opgave 10

Los op.

a $4 \cdot \sin(x) - 2 = -1$

b $10 + 20 \cdot \sin(x) = 8$

Opgave 11

Gegeven is de functie g met $g(x) = \sin(x - 2)$ met $0 \leq x \leq 2\pi$.

a Welke periode heeft g?

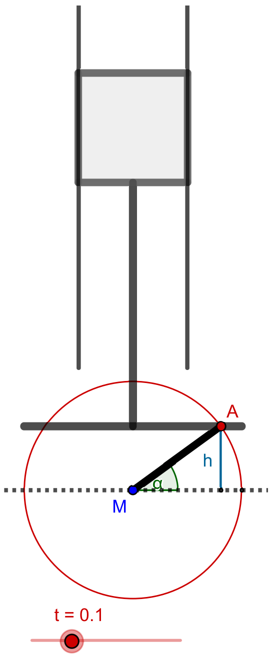
b Los op: $g(x) = 0{,}5$

Toepassen

Bekijk de applet: krukstang.

In de figuur zie je een schematische weergave van een **krukstang** MA die aan een zuiger is bevestigd. Als de zuiger op en neer beweegt, draait de krukstang rond.
Punt A zit helemaal rechts op de cirkel op $t = 0$.
Gegeven is $MA = 1$ decimeter.
De krukstang draait tegen de wijzers van de klok in, $x = \alpha$ is de draaihoek.

De hoogte van het punt A ten opzichte van de horizontale stippellijn is $h(x) = \sin(x)$.



Figuur 3.6

Opgave 12

Bekijk de formule voor de hoogte h van punt A boven de horizontale stippellijn.

a In welke eenheid is h uitgedrukt?

b Welke periode heeft h als x in graden wordt uitgedrukt?
En als x in radialen wordt uitgedrukt?

c Kun je een voordeel noemen van het werken met radialen ten opzichte van het werken met graden?

Het werken met decimeters als eenheid is niet gebruikelijk, liever werk je met meter, centimeter, millimeter.

d Hoe wordt het functievoorschrift voor de hoogte als je in mm werkt?
En wat verandert er dan aan de grafiek?

Opgave 13

De formule voor h in cm als functie van x in radialen is $h = 10 \cdot \sin(x)$.

a Maak de grafiek van h.

b Bij welke waarden voor x is $h(x) = 5$ cm?

c Bij welke waarden voor x is $h(x) = -5$ cm?

Testen

Opgave 14

Bekijk de grafiek van $f(x) = \sin(x)$. Los de volgende vergelijkingen op. Geef benaderingen in drie decimalen nauwkeurig.

a $\sin(x) = 0{,}95$

b $\sin(x) = -0{,}95$

Opgave 15

Gegeven is de functie $f(x) = 4\sin(x) + 1$ op $[-2\pi, 2\pi]$.

a Bereken alle nulpunten van de grafiek van f in twee decimalen nauwkeurig.

b Los op $f(x) < 0$.

1.4 Sinusoïden

Inleiding

Je hebt leren werken met de functie $y = \sin(x)$ (met x in radialen). Als je op deze functie transformaties toepast, krijg je andere periodes en kunnen de grafieken om een andere lijn dan de x-as gaan slingeren met een andere uitwijking. Dat is belangrijk omdat de sinusfunctie dan kan worden gebruikt om periodieke verschijnselen meer in het algemeen te beschrijven. Functies die door transformatie ontstaan uit $y = \sin(x)$ noem je sinusoïden.

Je leert in dit onderwerp

- het begrip sinusoïde en de bijbehorende karakteristieken kennen en de grafieken ervan tekenen;
- de vergelijkingen $a \cdot \sin(b(x+c)) + d = p$ oplossen.

Voorkennis

- de grafieken van $y = \sin(x)$ tekenen met x in radialen;
- de vergelijkingen $\sin(x) = c$ oplossen als c een constante is.

Verkennen

Opgave V1

Gegeven is de functie $f(x) = 2 \cdot \sin(4x) + 3$.

a Maak met de grafische rekenmachine de grafiek van f op $[0{,}2\pi]$.

b Bepaal de periode van deze periodieke functie.

c Bereken de coördinaten van alle toppen van de grafiek op $[0{,}2\pi]$.

Gegeven is de functie $g(x) = 4 \cdot \sin(0{,}5(x - \pi)) - 1$.

d Maak met de grafische rekenmachine de grafiek van g op $[0{,}4\pi]$.

e Bepaal de periode van deze periodieke functie.

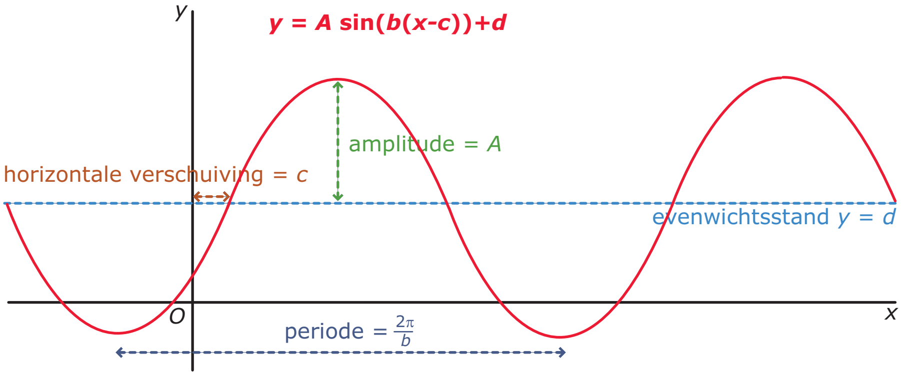
f Bereken de coördinaten van alle toppen van de grafiek op $[0{,}4\pi]$.

Uitleg

Bekijk de applet: sinusoïden

Door vervormingen van de grafiek van $f(x) = \sin(x)$ kun je grafieken met functies van de vorm $g(x) = a \cdot \sin(b(x - c)) + d$ maken. Zulke grafieken heten sinusoïden.

Een sinusoïde is periodiek en schommelt om de evenwichtsstand. De evenwichtsstand is het gemiddelde van het maximum en het minimum. Het verschil tussen het maximum en de evenwichtsstand heet de amplitude.



Figuur 4.1

Bekijk wat er gebeurt als je a, b, c en/of d verandert.

- a is de maximale uitwijking, de amplitude.
- b bepaalt de periode. De periode is $\frac{2\pi}{b}$.
- c is de horizontale verschuiving.
- d is de evenwichtsstand, de horizontale lijn $y = d$.

Wil je de grafiek van de sinusoïde $g(x) = 1{,}5 \sin(\pi(x-1)) + 0{,}5$ maken, dan lees je eerst uit de formule af:

- de amplitude is $1{,}5$
- de evenwichtsstand is $0{,}5$
- de periode is $\frac{2\pi}{\pi} = 2$
- de horizontale verschuiving is 1

Dit gebruik je om op een grafische rekenmachine geschikte vensterinstellingen te kiezen.

Het bereik van de functie is: $\text{B}_g = [0{,}5 - 1{,}5; 0{,}5 + 1{,}5] = [-1{,}2]$.

Opgave 1

In de **Uitleg** zie je de functie $g(x) = 1{,}5 \sin(\pi(x-1)) + 0{,}5$.

a Maak de grafiek van g met de applet en controleer zo de afgelezen waarden voor de periode, de amplitude, de evenwichtsstand en de periode.

b Plot de grafiek van g op je grafische rekenmachine. Welke vensterinstellingen kies je?

c Het punt $(0{,}0)$ ligt op de grafiek van $y = \sin(x)$. Welk punt op de grafiek van f ontstaat uit het punt $(0{,}0)$ als je de grafiek van $y = \sin(x)$ vervormt?

d Welke toppen heeft de grafiek van g op het interval $0 \leq x \leq 2\pi$?

Opgave 2

Gegeven is de functie f met $f(x) = 1 + 2\sin(3(x+2))$.

a Lees uit het functievoorschrift de periode, de amplitude, de evenwichtsstand en de horizontale verschuiving af.

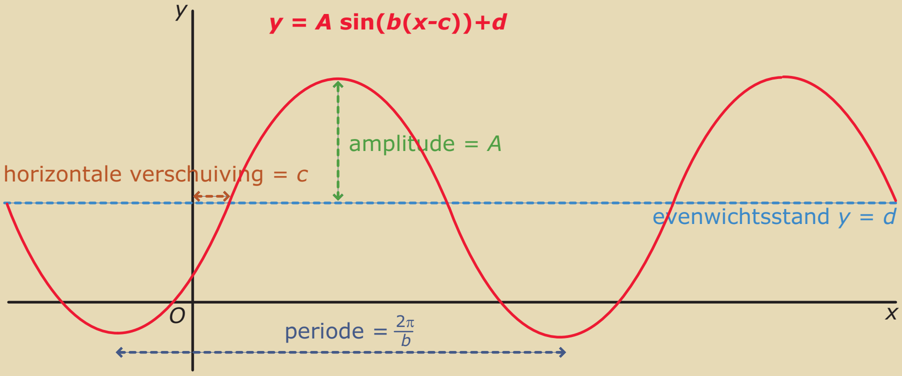
b Controleer met de applet in de **Uitleg** of de grafiek juist is.

c Oefen dit een aantal keer met andere sinusoïden.

Theorie en voorbeelden

Om te onthouden

Bekijk de applet: sinusoïden



Figuur 4.2

De grafiek van de functie $g(x) = a \cdot \sin(b(x-c)) + d$ is een **sinusoïde**. Dat is een functie die door vervormingen ontstaat uit $f(x) = \sin(x)$. Voor de grafiek van g geldt:

- de **amplitude** (maximale uitwijking uit de evenwichtsstand) is a (of - a als a negatief is);
- de **periode** is $\frac{2\pi}{b}$ en dus is $b = \frac{2\pi}{\text{periode}}$;
- de **horizontale verschuiving** is c, dit is de verschuiving in de x-richting;
- de **evenwichtsstand** is $y = d$.

Voorbeeld 1

Plot de grafiek van $y = 2\sin(\pi(x-1)) + 5$ met geschikte vensterinstellingen.

Bepaal daarvoor eerst de periode, de evenwichtsstand en de amplitude van deze sinusoïde.

Antwoord

De periode is $\frac{2\pi}{\pi} = 2$, de evenwichtsstand is $y = 5$ en de amplitude is 2.

Verder zie je in de formule dat er sprake is van een horizontale verschuiving van 1.

De grafiek heeft dus het punt $(1,5)$ op de evenwichtsstand.

Omdat de periode 2 is en je zeker één hele periode in beeld wilt hebben, kies je in ieder geval $0 \leq x \leq 3$.

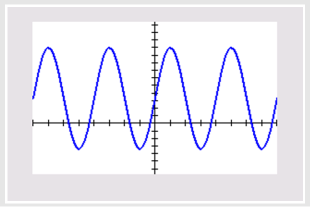
Verder zijn de maximale waarden $5+2 = 7$ en de minimale waarden $5-2 = 3$.
Dus kies je $0 \leq y \leq 8$ om ook in verticale richting de grafiek en de assen in beeld te hebben.

Ga na, dat je met deze instellingen de grafiek goed in beeld hebt.

Opgave 3

Bekijk deze grafiek van $y = 10\sin(4x) + 5$.

Welke vensterinstellingen moet je gebruiken om de grafiek zo in beeld te krijgen?



Figuur 4.3

Voorbeeld 2

Gegeven is de functie: $f(x) = 6\sin\left(\frac{\pi}{12}(x - 8)\right) + 10$.

- Bepaal de periode en de coördinaten van alle toppen van de grafiek van f.
- Los op: $f(x) = 13$.

Antwoord

De periode is $\frac{2\pi}{\frac{\pi}{12}} = 24$.

Je kunt nu met de grafische rekenmachine binnen één periode de toppen bepalen. Je kunt ook bedenken dat het maximum een kwart periode voorbij $x = 8$ zit, en dus bij $x = 8 + \frac{24}{4} = 14$ moet zitten. En de grootte is $10 + 6 = 16$. Zo kun je ook de top bepalen die bij het minimum hoort.
Dit geeft als toppen: $(14{,}16)$ en $(26{,}4)$.
Omdat je de periode weet, kun je alle coördinaten van de toppen geven: $(14 + k \cdot 24{,}16)$ en $(26 + k \cdot 24{,}4)$.

Oplossing van de vergelijking $f(x) = 13$:
Bereken met de grafische rekenmachine de x-coördinaten van de snijpunten van de grafiek van f en $y = 13$ binnen één periode.
Dit geeft bijvoorbeeld: $x = 10$ en $x = 18$.
De oplossing is: $x = 10 + k \cdot 24 \vee x = 18 + k \cdot 24$.

Opgave 4

Bekijk de functie in **Voorbeeld 2**.

a Beredeneer de coördinaten van een top met een minimum.

b Los op (benaderingen in drie decimalen nauwkeurig): $f(x) = 11$.

Opgave 5

Gegeven is de functie f met $f(x) = 3\sin(\pi(x - 1)) + 10$.

a Bepaal de periode en de coördinaten van alle toppen.

b Los op: $f(x) = 11{,}5$.
Rond af op twee decimalen.

Verwerken

Opgave 6

De grafieken van de functies zijn sinusoïden. Geef van iedere sinusoïde de periode, de evenwichtsstand en de amplitude. Breng daarna de grafiek in beeld zodat je op de grafische rekenmachine twee periodes ziet.

a $y = 12\sin(x)$

b $y = 50\sin(2\pi x) + 10$

c $y = 2 + 3\sin(4(x-2))$

d $y = -25\sin(0{,}5\pi x) - 5$

Opgave 7

Los de vergelijkingen op.

a $10\sin(\pi(x-1)) + 2 = 9$

b $5\sin\left(\frac{1}{2}x + 4\right) = 1$

c $50 - 30\sin\left(\frac{2\pi}{15}x\right) = 45$

Opgave 8

Gegeven is de functie $f(x) = 12\sin(0{,}5(x-2)) + 1$ met $0 \leq x \leq 8\pi$.

a Welk bereik heeft f?

b Welke periode heeft f?

c Bereken de nulpunten van f. Rond af op twee decimalen.

d Bereken de toppen van de grafiek van f.
Rond indien nodig af op drie decimalen.

Opgave 9

De hoogte boven de grond van iemand die zich in een reuzenrad bevindt, kan worden beschreven door:

$h = 11 + 10 \cdot \sin\left(\frac{\pi}{10} \cdot t\right)$

Hierin is h uitgedrukt in meter en t in seconden.

a Plot de grafiek van h.

b De getallen 11 en 10 uit de formule hebben een betekenis voor het reuzenrad. Welke?

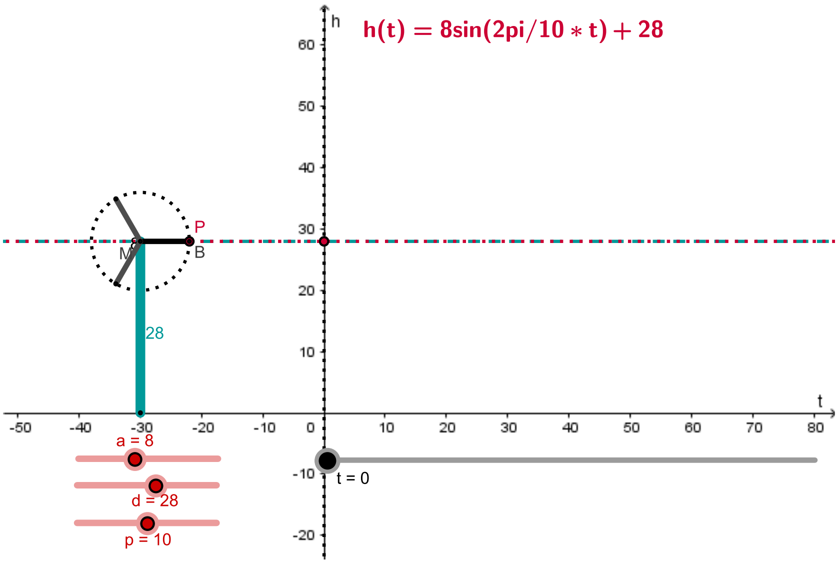
c Na één periode is het reuzenrad precies één keer rondgedraaid. Bepaal de periode in seconden.

d Bereken hoelang een bakje van het reuzenrad binnen een periode hoger dan 18 meter boven de grond zit. Geef je antwoord in seconden afgerond op één decimaal.

Toepassen

Hier zie je een schematische weergave van een **windmolen** waarvan je de lengte van de wieken a (in m), de hoogte van het draaipunt d (in m) en de omwentelingstijd, de periode p (in s) kunt aanpassen.
De grafiek gaat over de hoogte van de uiterste punt P van de wiek, afhankelijk van de tijd t in seconden.

Bekijk de applet.



Figuur 4.4

Opgave 10

Bekijk de windmolen in **Toepassen**.

Van een zekere windmolen is de hoogte van het draaipunt 30 m, de lengte van de wieken 15 m en de omwentelingstijd (bij een zekere windsnelheid) 5 s.

a Stel een bijpassende formule op voordat je deze instellingen in de applet doet.
Controleer je antwoord met de applet.

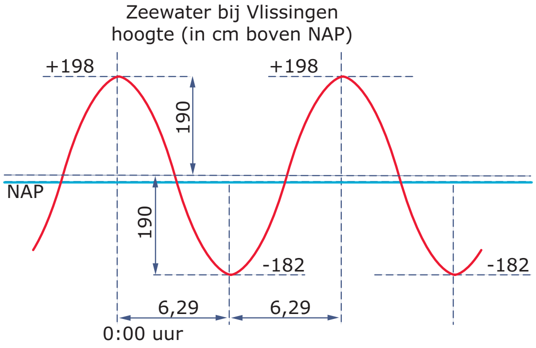
Deze windmolen staat achter een boerderij die een hoogte heeft van 20 m.

b Hoe lang is elke omwenteling de hoogte van de top van de wiek groter dan de hoogte van de boerderij?

c Je kunt de opdrachten bij a en b variëren door andere getallen te kiezen. Oefen met een medeleerling.

Opgave 11: Getijden

De grafiek in de volgende figuur geeft globaal de getijdenbeweging van het zeewater voor de haven van Vlissingen weer. Er wordt geen rekening gehouden met de invloed van de wind, met springtij, en dergelijke.



Figuur 4.5

a Hoe hoog is de gemiddelde waterstand volgens deze grafiek?

b Hoe groot is de maximale afwijking van de waterstand ten opzichte van het gemiddelde?

c Hoe groot is de periode van de getijdenbeweging?

Een benadering van de getijdenbeweging wordt gegeven door de volgende formule:

$$y = 8 + 190\sin\left(\frac{2\pi}{12{,}25}\cdot(t + 3{,}0625)\right)$$

met t in uren t.o.v. middernacht op 21 juni 2008 en y in cm ten opzichte van het NAP.

d Vergelijk de grafiek van deze functie met de grafiek in de figuur hierboven. Vind je dat de formule een goed beeld geeft van de getijdenbeweging?

e Hoe groot is volgens de formule de periode en de amplitude?

f Hoeveel uur per periode is de waterstand hoger dan 180 cm?

Testen

Opgave 12

Bepaal van de volgende functies de periode, de amplitude, de evenwichtslijn en de horizontale verschuiving ten opzichte van $y = \sin(x)$.

a $y = 4\sin(4\pi x)$

b $y = 6 + 2\cos(x + 8)$

c $y = 0{,}5\sin(0{,}5\pi x)$

Opgave 13

Gegeven is de functie $h(t) = 20 + 12\sin(0{,}5\pi(t - 2))$.

a Met welke vensterinstellingen krijg je de grafiek van h op je grafische rekenmachine goed in beeld?

b Bereken de coördinaten van de toppen van de grafiek van h

c Los op: $h(t) \geq 30$. Benaderingen in drie decimalen nauwkeurig.

1.5 Periodieke modellen

Inleiding

Je hebt tot nu toe berekeningen gemaakt en grafieken getekend bij gegeven sinusoïden. Het omgekeerde kan ook: bij een gegeven grafiek van een sinusoïde de formule opstellen. Met die formule kun je snel nieuwe punten van de grafiek vinden.

Verder kun je periodieke verschijnselen waarvan de grafiek golfvormig is, vaak goed benaderen met een sinusoïde. Die sinusoïde is dan een model voor het verschijnsel.

Je leert in dit onderwerp

- bij een getekende sinusoïde de formule opstellen;
- sinusoïden gebruiken als model voor een periodiek verschijnsel.

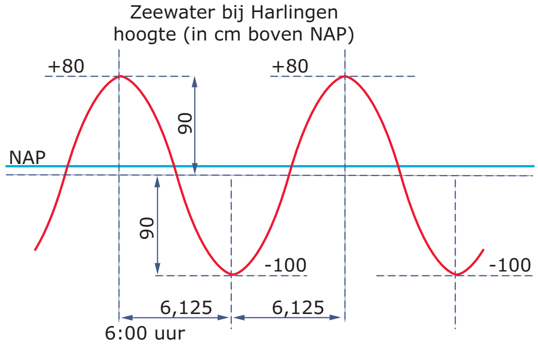
Voorkennis

- de grafiek van een sinusoïde tekenen;
- de periode, de amplitude, de evenwichtsstand en de horizontale verschuiving van een sinusoïde aflezen uit de formule, dan wel uit de grafiek.

Verkennen

Opgave V1

In de getijdeninformatie van Harlingen kun je aflezen dat bij hoogwater de waterstand h ongeveer 80 cm boven NAP (Normaal Amsterdams Peil) zit en dat bij laagwater de waterstand ongeveer 100 cm onder NAP zit. Verder liggen de opeenvolgende tijdstippen van hoogwater (net als die van laagwater) ongeveer 12 uur en 15 minuten uit elkaar. Dat betekent een periode van $12{,}25$ uur. Op een zekere dag is het hoogwater om 6:00 uur.



Figuur 5.1

De bijbehorende grafiek lijkt op een sinusoïde.

a Bepaal de periode, de amplitude en de evenwichtslijn van die sinusoïde.

b Stel een passende formule op.

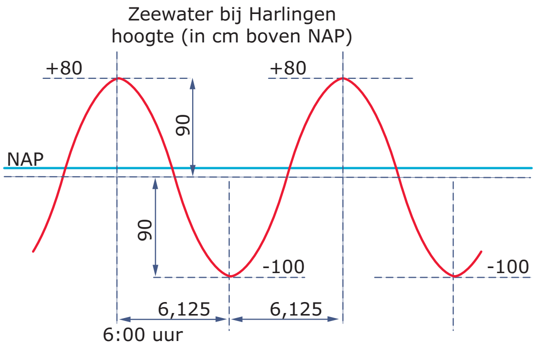
c Wat is het nut van zo'n formule?

Uitleg

Bekijk de applet: waterstanden Harlingen

Periodieke verschijnselen waarvan de grafiek golfvormig is, kun je vaak goed benaderen met een sinusoïde. Die sinusoïde is dan een model voor het verschijnsel.

In de getijdeninformatie van Harlingen kun je aflezen dat bij hoogwater de waterstand h ongeveer 80 cm boven NAP (Normaal Amsterdams Peil) zit en dat bij laagwater de waterstand ongeveer 100 cm onder NAP zit. Verder liggen de opeenvolgende tijdstippen van hoogwater (net als die van laagwater) ongeveer 12 uur en 15 minuten uit elkaar. Dat betekent een periode van 12,25 uur. Op een zekere dag is het hoogwater om 6:00 uur.



Figuur 5.2

Bekijk de schets van een grafiek die past bij de getijdeninformatie van Harlingen.

De bijbehorende formule bij de grafiek heeft de vorm: $h(t) = a \cdot \sin(b(t+c)) + d$.

In de grafiek kun je de volgende gegevens aflezen.

- De periode is 12,25 uur: $b = \frac{2\pi}{12{,}25} \approx 0{,}52$.
- De waterstand ligt tussen 0,8 m en -1,0 m. De amplitude is $a = 0{,}9$ m.
- De evenwichtsstand ligt 0,9 m onder hoogwater: $d = -0{,}1$.
- Hoogwater moet bij $t = 6$ zitten. Het direct ervoor liggende punt op de evenwichtsstand zit daar een kwart periode voor. Dit is bij $t = 6 - 3{,}0625 \approx 2{,}94$. Dit betekent dat $c \approx -2{,}94$.

De bijpassende sinusoïde wordt: $h(t) \approx 0{,}9 \sin(0{,}52(t - 2{,}94)) - 0{,}1$.

Opgave 1

Gegeven is de opgestelde sinusoïde als model voor de waterstand bij Harlingen in de **Uitleg**. Stel nu dat alle punten 20 centimeter hoger liggen. De hele grafiek verschuift dan 20 centimeter omhoog.

a Leg uit hoe uit de gegevens de periode, de evenwichtsstand en de amplitude worden gevonden.

b Leg uit dat $c \approx 2{,}94$.

Opgave 2

Ga uit van de functie $y = \sin(x)$. Schrijf het voorschrift op van de periodieke functies die ontstaan bij de gegeven wijzigingen.

a De amplitude wordt 4.

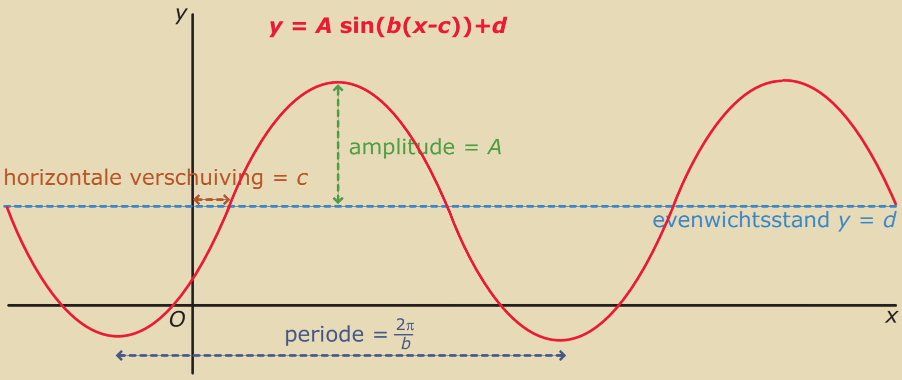
b De amplitude wordt 10 en de evenwichtsstand wordt 20.

c De periode wordt 4π en de amplitude wordt 4.

d De horizontale verschuiving is 2, de periode wordt 10, de amplitude wordt 5 en de evenwichtsstand wordt 10.

Theorie en voorbeelden

Om te onthouden



Figuur 5.3

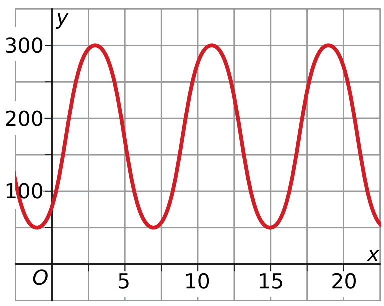
Wanneer je een periodiek verschijnsel kunt beschrijven met een sinusoïde, is het bijbehorende functievoorschrift: $y = a\sin(b(x-c)) + d$ met $b = \frac{2\pi}{p}$

Hierin is:

- d de **evenwichtsstand**, dit is de lijn $y = d$
- a de **amplitude**, de maximale uitwijking ten opzichte van de evenwichtsstand
- p de **periode**
- c de **horizontale verschuiving** ten opzichte van de standaardgrafiek $y = \sin(x)$

Voorbeeld 1

Bekijk de sinusoïde. Welk functievoorschrift kun je bij deze sinusoïde opstellen?



Figuur 5.4

Antwoord

De formule is van de vorm $y = a \cdot \sin(b(x-c)) + d$.

Het maximum van de functie is 300 en het minimum 50. Dit betekent dat:

- de amplitude $a = \frac{300-50}{2} = 125$ is;
- de evenwichtsstand $d = 300 - 125 = 175$ is.

Twee opvolgende maxima zitten bij $x = 3$ en $x = 11$, zodat de periode 8 is.

Dit betekent dat $b = \frac{2\pi}{8} = \frac{1}{4}\pi$.

De horizontale verschuiving is de x-waarde waarop de grafiek stijgend door de evenwichtsstand gaat.

Hier is dat $x = 1$.

Het functievoorschrift wordt: $y = 125\sin\left(\frac{1}{4}\pi(x-1)\right) + 175$.

Opgave 3

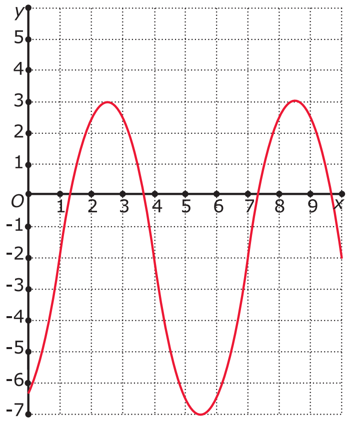
Bekijk **Voorbeeld 1**.

a Waarom is $y = 125\sin\left(\frac{\pi}{4}(x-9)\right) + 175$ ook een goed functievoorschrift?

b Geef nog een ander goed functievoorschrift.

Opgave 4

Je ziet hier een sinusoïde getekend.



Figuur 5.5

Maak er een functievoorschrift bij, uitgaande van $y = \sin(x)$.

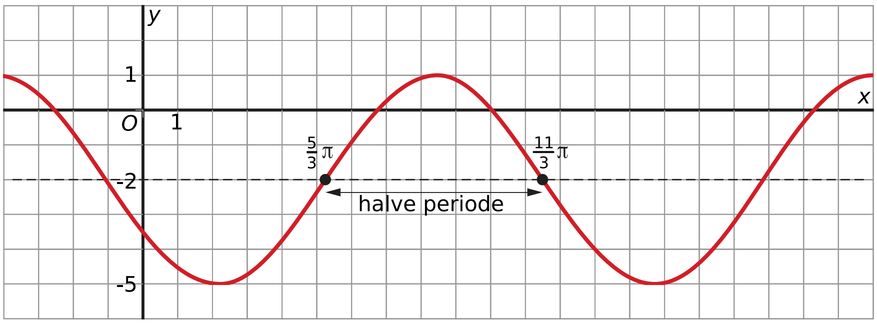
Voorbeeld 2

Een sinusoïde heeft een maximum van 1 en een minimum van -5. Het domein is $\mathbb{R}$. De evenwichtsstand $y = -2$ wordt onder andere bereikt als $x = \frac{5}{3}\pi$ en daarna als $x = \frac{11}{3}\pi$. Tussen deze beide x-waarden ligt de grafiek boven de evenwichtsstand.

Stel een formule op voor de beschreven sinusoïde.

Antwoord

De formule heeft de vorm $y = a\sin(b(x-c)) + d$.
Breng de situatie eerst in beeld.



Figuur 5.6

De twee punten op de evenwichtsstand liggen een halve periode uit elkaar.

- De periode is $2 \cdot \left(\frac{11}{3}\pi - \frac{5}{3}\pi\right) = 4\pi$, zodat $b = \frac{2\pi}{4\pi} = \frac{1}{2}$.
- De evenwichtsstand is -2.
- De amplitude $a = 3$.

De horizontale verschuiving is $\frac{5}{3}\pi$ ten opzichte van de standaardsinus, want bij die x-waarde gaat de grafiek stijgend door de evenwichtsstand.

De gevraagde formule is: $y = 3\sin\left(\frac{1}{2}\left(x - \frac{5}{3}\pi\right)\right) - 2$.

Opgave 5

De grafiek van een sinusoïde f heeft minimum 10 voor $x = 1$ en eerstvolgend maximum 26 voor $x = 13$.

a Bereken de periode, de evenwichtslijn en de amplitude. Bekijk eventueel eerst **Voorbeeld 2**.

b Geef een passende formule.

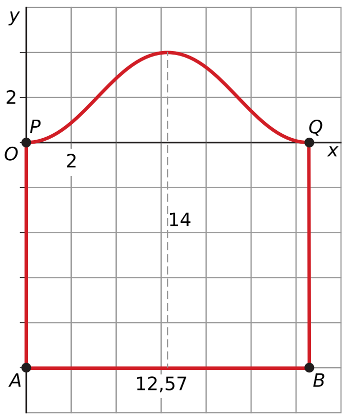
c Bereken in twee decimalen nauwkeurig: $f(12)$, $f(12{,}25)$, $f(12{,}5)$, $f(12{,}75)$ en $f(13)$.

d Los op: $f(x) > 20$.

Voorbeeld 3

Bekijk de figuur. Als je een cilinder met een diameter van 4 centimeter eerst schuin doorsnijdt zodat er twee cilinders ontstaan met een schuin vlak, en dan de ene cilinder in de lengterichting openknipt en plat neerlegt, krijg je de figuur. De bovenrand is een zuivere sinusoïde.

Stel voor deze rand een formule op. Gebruik daarvoor het assenstelsel in de figuur en neem aan dat punt P de coördinaten $(0{,}0)$ heeft.



Figuur 5.7

Antwoord

Bepaal vanuit de figuur dat:

- de evenwichtsstand is 2;
- de amplitude is 2;
- de periode is 4π.

Het maximum zit halverwege de bovenrand bij $x = 2\pi$.
Ten opzichte van de standaardsinus is de verschuiving π.
De formule wordt: $y = 2\sin(0{,}5(x - \pi)) + 2$.

Opgave 6

Bestudeer **Voorbeeld 3**.

a Stel voor de bovenrand een formule op uitgaande van $y = \sin(x)$.

b Waarom is de periode 4π?

c De lijn $y = 3$ snijdt de sinusoïde uit het voorbeeld in de punten A en B.
Bereken de lengte van lijnstuk AB.

d Een lijn evenwijdig aan PQ snijdt de bovenrand in A en B. Gegeven is $AB = 4$ cm. Bepaal de coördinaten van A en B.

Verwerken

Opgave 7

Het waterpeil bij een plaats in Nederland was om 3:25 uur hoog en om 15:28 uur weer hoog. Het water was toen 90 centimeter boven NAP. Dit verschilt 200 centimeter met laagwater.

Hoe hoog het water staat ten opzichte van NAP kun je goed beschrijven met een sinusoïde.
Stel een formule daarvoor op. Neem de tijd t in uur met $t = 0$ om 0:00 uur en de hoogte h in centimeter.

Opgave 8

Gegeven zijn karakteristieken van sinusoïden. Stel een passend functievoorschrift op van de vorm $y = a \cdot \sin(b(x - c)) + d$.

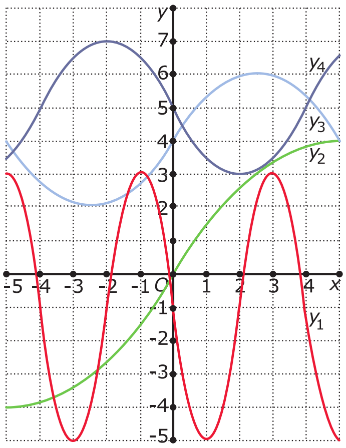
a De amplitude is 3, de periode is π, de evenwichtsstand is -1 en het maximum zit bij $x = \frac{\pi}{2}$.

b De amplitude is 5, de periode is 2, de evenwichtsstand is 2 en het maximum zit bij $x = 1{,}5$.

c De amplitude is 2, de periode is 6, de evenwichtsstand is 0 en het minimum zit bij $x = 3$.

Opgave 9

Stel bij de vier sinusoïden in de figuur een passend functievoorschrift op.



Figuur 5.8

Opgave 10

De grafiek van $f(x)$ is een sinusoïde. De evenwichtsstand is 1, de amplitude is 2, de periode is π en de grafiek gaat stijgend door het punt $\left(\frac{1}{6}\pi, 1\right)$.

a Stel een formule op voor f.

b Bereken $f(0)$. Rond af op twee decimalen.

c Bereken de nulpunten van f. Rond af op twee decimalen.

Opgave 11

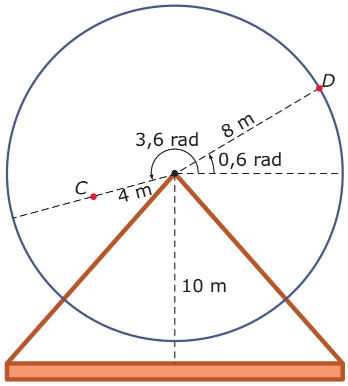
Een reuzenrad bevat de stoeltjes C en D. Stoeltje C draait op een afstand van 4 meter van de as in de rondte. De as van het reuzenrad bevindt zich op 10 meter boven de grond. Bekijk de getekende situatie. Het reuzenrad draait in 8 seconden één keer rond. Op $t = 2$ staat stoeltje C zo hoog mogelijk. Het reuzenrad draait tegen de wijzers van de klok in.

a Bereken bij elke stand de hoogte van de stoeltje C ten opzichte van de grond.

b Stel een passend functievoorschrift op voor de hoogte van stoeltje C.

c Hoe hoog staat stoeltje C op tijdstip $t = 65$?

d Hoe lang zit je in stoeltje C elk rondje hoger dan 12 m?



Figuur 5.9

Toepassen

Opgave 12: Temperatuur zeewater

De temperatuur T (°C) van het zeewater bij Zandvoort is goed te benaderen door een sinusoïde. De periode is 12 maanden. De laagste temperatuur is 6 °C op 1 maart en de hoogste 22 °C op 1 september.

a Stel een passende formule op voor T na t maanden met $t = 0$ op 1 januari.

b Hoe hoog is de zeewatertemperatuur op 1 oktober?

c Stel dat het zeewater door de opwarming van de aarde met 2 °C toeneemt. Welke formule kun je nu voor T opstellen?

Opgave 13: Ventiel fietsband

Als je bij een constante snelheid van een fiets de hoogte van een ventiel van een fietsband uitzet tegen de tijd krijg je een sinusoïde.

Tom fietst met een snelheid van 15 km/h. De diameter van zijn wielen is 90 centimeter. De hoogte van het ventiel van zijn banden varieert van 5 tot 85 centimeter.

a Hoe vaak draait het ventiel per seconde rond? Rond af op één decimaal.

b Stel een formule op van de hoogte van het ventiel h in centimeter na t seconden. Neem $t = 0$ als het ventiel op 45 centimeter hoogte zit en omhoog beweegt.

Testen

Opgave 14

Functie f met voorschrift $f(x)$ heeft een sinusvormige grafiek met een minimum in het punt $(20,300)$ en een eerstvolgend maximum in het punt $(32,400)$.

a Maak een schets van deze grafiek met x van 0 tot ten minste 40.

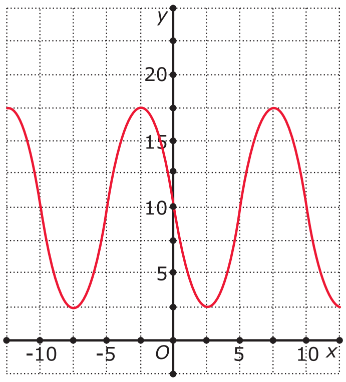
b Bereken de periode, de amplitude en de evenwichtslijn en stel een passend functievoorschrift op.

c Bereken $f(50)$, $f(51)$ en $f(52)$.

d Los op: $f(x) = 325$.

Opgave 15

Stel bij deze sinusoïde twee passende functievoorschriften op.



Figuur 5.10

Opgave 16

Onze ademhaling is bij benadering een periodiek verschijnsel. Een gezonde volwassen man ademt ongeveer 12 keer per minuut in en weer uit. De longinhoud $V(t)$ kan daarbij met zo'n halve liter toe- of afnemen, waarin t de tijd in seconden is. Het longvolume na inademen is 5,2 liter.

a Hoe groot is de ademhalingsfrequentie per minuut?

b Ga ervan uit dat $V(t)$ een sinusoïde is met op $t = 0$ een maximale longinhoud. Maak de grafiek van de longinhoud V uitgezet tegen de tijd t.

c Stel bij deze situatie een formule op voor $V(t)$.

1.6 Totaalbeeld

Samenvatten

Je hebt nu alle theorie van **Periodieke functies** doorgewerkt. Er moet een totaalbeeld van deze leerstof ontstaan...
Ga na, of je al de bij dit onderwerp horende 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

- periodiek verschijnsel — periode — golflengte — frequentie;
- eenheidscirkel — draaihoek — radialen — standaardsinus;
- arcsinus;
- sinusoïde — periode — amplitude — evenwichtslijn — horizontale verschuiving.
- periodiek model

Activiteitenlijst

- bij een periodiek verschijnsel de periode, de golflengte en de frequentie bepalen;
- draaihoeken omrekenen van graden naar radialen en omgekeerd — de standaard sinusgrafiek tekenen;
- vergelijkingen met de standaardsinus oplossen, exact (met arcsin) en met de GR;
- bij een sinusoïde de periode, de amplitude, de evenwichtslijn en de horizontale verschuiving bepalen, zowel vanuit de grafiek als vanuit de formule — toppen en nulpunten van sinusoïden berekenen — vergelijkingen bij sinusoïden oplossen;
- bij een gegeven periodiek verschijnsel een sinusoïde opstellen die dat verschijnsel zo goed mogelijk beschrijft.

Achtergronden

Bekijk de applet: sinus

In de Indische wiskunde heette de helft van de koorde van een cirkelboog de ardhâ-jyâ (ardha = half; jyâ = koorde) van die boog. Dit werd, afgekort tot jyâ of jîv en door de Arabieren als vgîb geschreven. Toen het wetenschappelijk centrum van de wereld verschoof, werden de Arabische werken in de 12e eeuw vertaald naar het Latijn. Bij de vertaling werd vgîb gelezen als het Arabische vgaib wat ‘plooi’ of ‘boezem’ betekent. Dit werd door **Gerard van Cremona (1114—1187)** letterlijk vertaald als **sinus**.

Sinus en cosinus werden al in de Griekse Oudheid bestudeerd en later o.a. door de Indische geleerden **Aryabhata (476–550)**, **Brahmagupta (598—670)** en **Bhaskara (1114—1185)** en de Perzische wetenschappers **Mohammad ibn Musa al-Khwarizmi (ongeveer 780—845)**, **Omar Khayyam (1048—1131)**, **Nasir al-Din al-Tusi (13e eeuw)**, **Ghiyath al-Kashi (14e/15e eeuw)**. In

de 12de eeuw werden deze begrippen in West-Europa bekend.
In de oorspronkelijke definitie waren sinus en cosinus verhoudingen van bepaalde zijden in een rechthoekige driehoek. De grootte van deze verhouding verandert niet zo lang de hoeken even groot blijven.
Maar deze definitie brengt het probleem met zich mee dat stompe hoeken geen sinus of cosinus hebben (want er bestaan geen rechthoekige driehoeken met een stompe hoek). Om dit probleem op te lossen werden de sinus en cosinus opnieuw gedefinieerd met behulp van de eenheidscirkel. Voordeel van deze definitie is dat de sinus en de cosinus van elke hoek kunnen worden bepaald.

Testen

Opgave 1

Reken de hoeken in graden om naar radialen tussen 0 en 2π en omgekeerd.
Rond indien nodig af op twee decimalen.

a $120°$

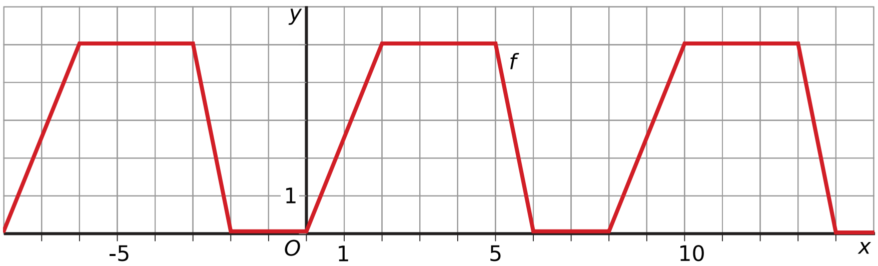
b $700°$

c $\frac{5}{6}\pi$ rad

d $\frac{5}{6}$ rad

Opgave 2

Bekijk de periodieke grafiek f. De grafiek loopt links en rechts oneindig ver door.



Figuur 6.1

a Hoe groot is de periode van f?

b Bereken $f(36)$ en $f(2449)$.

c Bereken $f(-250{,}5)$.

d Voor welke waarde van x is $f(x) = 3{,}75$?

e Los op: $f(x) = 0$ met $90 \leq x \leq 100$.

Opgave 3

Gegeven is de sinusoïde $y = 50\sin\left(\frac{1}{4}\pi(x-1)\right) + 200$ met $0 \leq x \leq 30$.

a Geef de periode, de evenwichtsstand en de amplitude van deze sinusoïde.

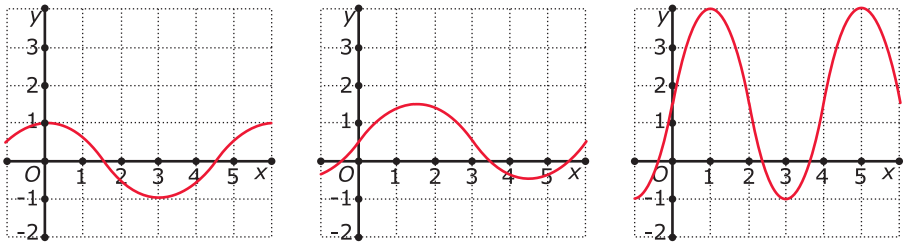
b Plot de grafiek. Welke vensterinstellingen kies je?

c Hoeveel periodes krijg je in beeld?

d Los in drie decimalen nauwkeurig op: $y \leq 180$.

Opgave 4

Bekijk de drie figuren met sinusoïden. Geef telkens een bijpassend functievoorschrift.



I II III

Figuur 6.2

Opgave 5

Bij het bepalen van de gewenste dijkhoogte langs de Nederlandse kust is het belangrijk dat de dijk hoger is dan de te verwachten maximale waterhoogte bij een stormvloed. De gemiddelde waterhoogte is daarbij niet van belang. Bij normale omstandigheden kan de getijdenbeweging van het zeewater bij de Hondsbosse zeewering bij Petten redelijk worden beschreven door de functie:

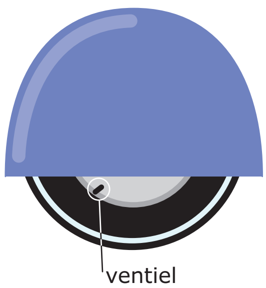
$$y = 0{,}4 + 1{,}5 \cdot \sin\left(\frac{2\pi}{12{,}25} \cdot t\right)$$

Hierbij is t in uur ten opzichte van middernacht op 21 juni 1998 en de waterhoogte y in meter ten opzichte van het NAP.
Onder invloed van de stand van de zon en de maan kan de amplitude van de getijdenbeweging variëren van 10% tot 140% van de amplitude van de gegeven functie. Afhankelijk van de windsterkte, kan de gemiddelde waterhoogte bij aanlandige wind $1{,}5$ tot $2{,}5$ meter hoger zijn dan normaal.

Hoe hoog moet de zeedijk van Petten volgens jou minimaal zijn? Licht je antwoord toe aan de hand van het gegeven functievoorschrift.

Opgave 6

Van het autowiel in de figuur is slechts het onderste deel zichtbaar. Van de wielhoogte is $\frac{3}{4}$ deel afgeschermd achter het spatbord.



Figuur 6.3

a Hoeveel procent van de tijd is het ventiel zichtbaar als de auto met een constante snelheid rijdt?

'Zichtbaar' kun je aangeven met een 1, 'onzichtbaar' met een 0. Je kunt dan de grafiek van de zichtbaarheid van het ventiel uitzetten tegen de tijd.

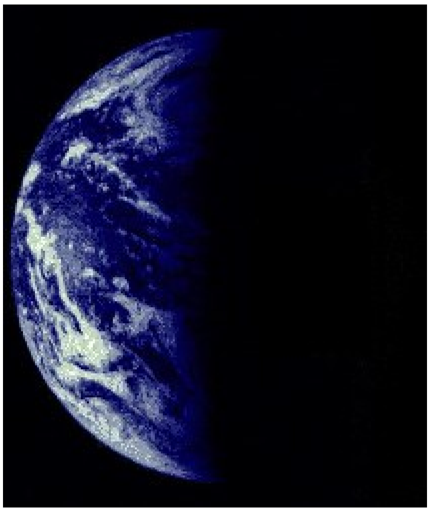
b Is dit een periodieke functie? Zo ja, teken een periode op schaal.

Toepassen

Opgave 7: Daglengte

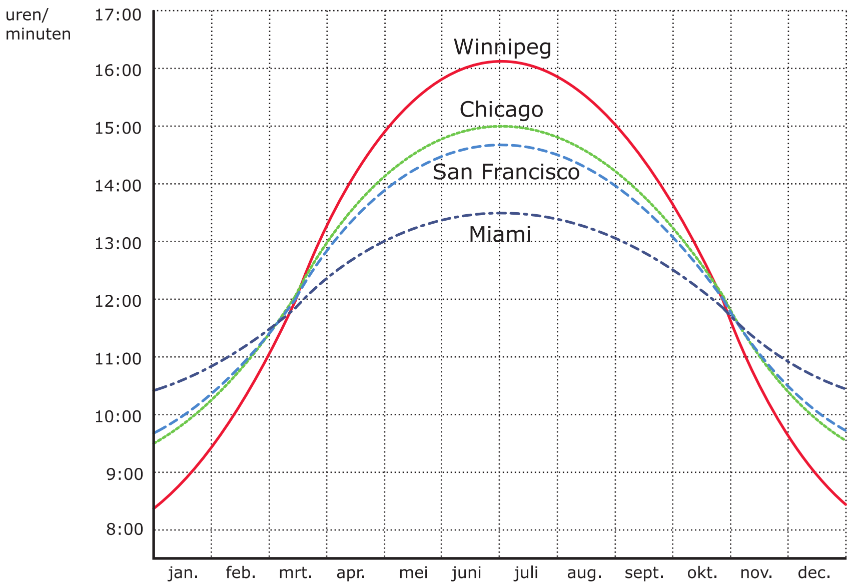
De daglengte varieert door het jaar heen. De daglengte is het verschil in tijd tussen zonsopkomst en zonsondergang. Dit is een heel mooi periodiek verschijnsel dat behoorlijk nauwkeurig is te beschrijven met behulp van een sinusoïde.

Via internet kun je een **tabel voor zonsopkomst en - ondergang in De Bilt** vinden.
Een dergelijke tabel kun je in een rekenblad invoeren en dan grafieken maken voor de tijdstippen van zonsopkomst en zonsondergang. **Hier zie je er een voorbeeld van**. Het zijn de vereenvoudigde gegevens van een bepaald jaar voor Amsterdam. De daglengte is het verschil van beide en ook daarvan is eenvoudig een grafiek te maken.
Je kunt de grafieken benaderen met sinusoïden en zo nauwkeurig de langste dag en de kortste dag berekenen...



Figuur 6.4

Het variëren van de daglengte hangt nogal af van de breedtegraad op aarde. Dat komt omdat de aardas niet precies loodrecht op de ecliptica (het vlak waarin de aardbaan om de zon ligt). Ook leuk om nader te onderzoeken...



Figuur 6.5

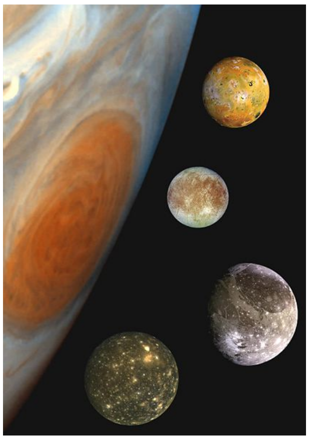
a Stel voor de vier steden een voorschrift op voor de daglengte als functie van de tijd t in dagen; $t = 0$ op 1 januari.

b Op welke datum is de langste dag van het jaar? En de kortste?

c Hoeveel dagen per jaar is de daglengte meer dan 14 uur?

Opgave 8: De manen van Jupiter

In 1610 werden de vier helderste **Jupitermanen** ontdekt door Galileï. De manen beschrijven bij benadering cirkelvormige banen om Jupiter, alle vier in dezelfde omlooprichting. Deze banen liggen (vrijwel) in één vlak met Jupiter en de Aarde. Daarom zie je Jupiter en de vier manen in een kijker altijd op één horizontale lijn liggen. De onderlinge posities van de manen in het kijkerbeeld veranderen voortdurend. Voor amateurastronomen worden maandelijks grafieken gepubliceerd waaruit ze op ieder moment de posities van de manen kunnen aflezen. Zie **hemel.waarnemen.com: Galileïsche manen van Jupiter, slingerdiagram september 2008**
Het diagram op de website geeft informatie over de maand september in 2008.
Deze slingerdiagrammen zijn vrijwel zuivere sinusoïden.

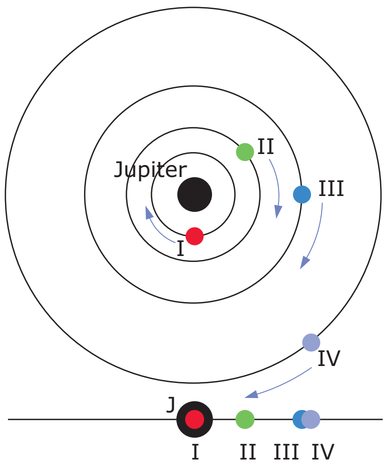


Figuur 6.6

Voor Ganymedes bijvoorbeeld wordt deze harmonische beweging goed beschreven door $u(t) = 15\sin\left(\frac{2\pi}{29{,}5}(t-17)\right)$ waarin t de tijd in dagen is met $t = 1$ op 1 september 2008 om 0:00 uur en u de uitwijking t.o.v. Jupiter gemeten in Jupiterstralen.

Zo kun je ook van de beweging van de drie andere Galileïsche manen een formule opstellen.
En verder kun je op elk moment tekenen hoe je deze manen t.o.v. Jupiter vanaf Aarde ziet. Nog een leuke puzzel...

Op 1 september 2008 om 0:00 uur waren dus van links (west) naar rechts (oost) in de kijker te zien: Io (I, voor Jupiter), Europa (II), Ganymedes (III) en Callisto (IV). Hier zie je van de vier manen de posities op hun cirkelbanen op 1 januari 1990 om 0:00 uur getekend.



Figuur 6.7

a Teken in de figuur voor deze vier manen het deel van de baan dat ze doorlopen van 1 september 0:00 uur tot 5 september 0:00 uur.

In de kijker zie je de beweging van elk van die manen als een in de tijd veranderende uitwijking $u(t)$ t.o.v. Jupiter op een horizontale as. Die uitwijking kan goed worden beschreven met een sinusoïde. u wordt uitgedrukt in veelvouden van de straal van Jupiter en t is in dagen.
Voor Callisto geldt bij goede benadering $u(t) = 26\sin(0{,}365(t-24))$.
(Hierbij is er van uit gegaan dat ‘West’ een positieve waarde van u betekent en ‘Oost’ een negatieve.)

b Laat zien dat deze formule redelijk overeenkomt met de gegeven grafiek. Bereken met de formule de omlooptijd van Callisto.

c Stel zelf zo'n formule op voor Ganymedes.

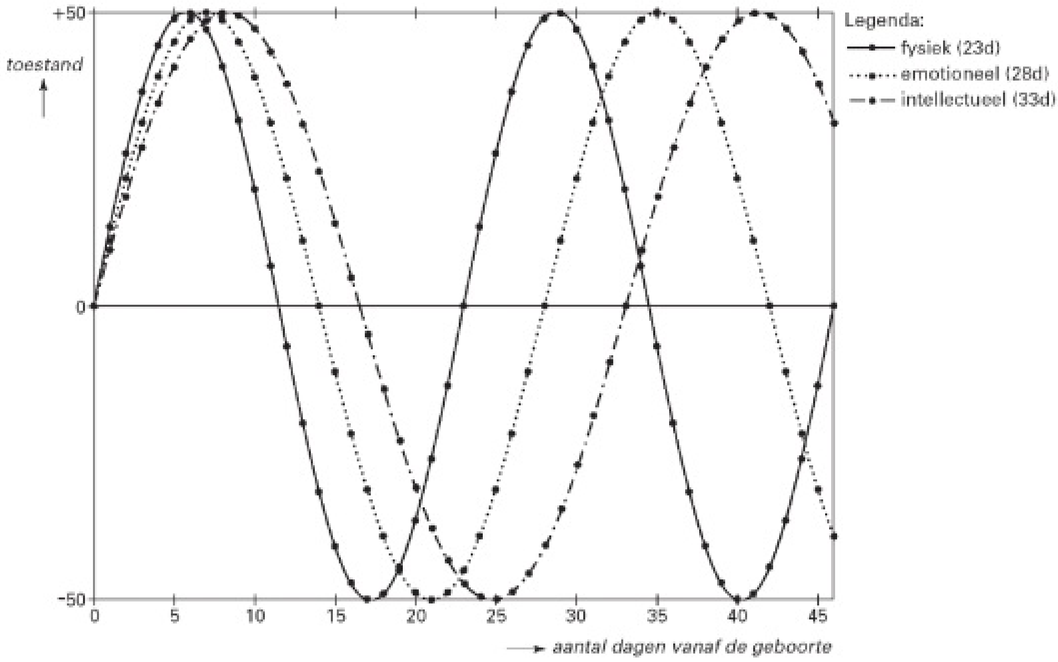
De manen zijn in de figuur naar verhouding veel te groot getekend. In werkelijkheid zijn het stipjes. Dus als $-1 \leq u(t) \leq 1$ dan kunnen de manen achter Jupiter zitten.

d Bereken met behulp van de formule voor Ganymedes hoe lang deze maan achter Jupiter zit.

Examen

Opgave 9: Bioritme

Op een pagina op Internet staat te lezen dat ons leven beheerst wordt door een drietal toestanden, namelijk door onze fysieke, onze emotionele en onze intellectuele toestand. Op de ene dag voel je je fysiek (lichamelijk) beter dan op een andere dag. Deze 'fysieke toestand' kunnen we weergeven op een schaal van -50 (fysiek op dieptepunt) tot $+50$ (fysiek opperbest). Deze fysieke toestand varieert in de tijd volgens een sinusoïde.
Ook de 'emotionele toestand' en de 'intellectuele toestand' variëren op een schaal van -50 tot $+50$ volgens een sinusoïde. Zie figuur.



Figuur 6.8

Bij de geboorte van een mens zou elke cyclus zich in dezelfde begintoestand bevinden, zoals is weergegeven in de figuur. Tezamen bepalen de drie cycli het zogenaamde bioritme van een mens. Sommigen beweren dat het bioritme volledig vastlegt tot welke prestaties een mens op een bepaald moment in staat is. Zo zou je bijvoorbeeld kunnen uitrekenen op welke dag je het best kunt solliciteren.
Voor de fysieke cyclus is de periode 23 dagen, voor de emotionele cyclus 28 dagen en voor de intellectuele cyclus is de periode 33 dagen.

Het bioritme in de figuur betreft een pasgeboren baby. E is de emotionele toestand van de baby t dagen na de geboorte. Hierbij hoort een formule van de vorm $E = a\sin(bt)$.

a Geef de waarden van a en b.

Zodra de emotionele toestand beneden -25 komt, zou het moeilijker worden om de emoties onder controle te houden.

b Hoeveel procent van een periode heeft de emotionele toestand een waarde die kleiner is dan -25? Licht je antwoord toe.

c F is de fysieke toestand van de baby. Onderzoek of F op de eerste verjaardag een dalend of een stijgend verloop heeft.

Annelies is op 1 januari 1983 geboren. Op 1 januari 2001 wordt ze dus 18 jaar. Vanaf die dag mag ze rijexamen doen. Ze wil dat doen op een dag waarop zowel haar fysieke als haar intellectuele toestand positief is. (De jaren 1984, 1988, 1992, 1996 en 2000 hebben een dag extra, dus 366 dagen.)

d Onderzoek welke de eerste drie dagen van januari 2001 zijn die voor het rijexamen in aanmerking komen.

(bron: examen wiskunde B1,2 havo 2000, eerste tijdvak)

Opgave 10: Eb en vloed

De staatsuitgeverij publiceert elk jaar de Getijdentabellen voor Nederland. Daarin worden voor een aantal kustplaatsen zowel de dagelijkse tijdstippen voor hoogwater en laagwater als de verwachte hoogten (in centimeters) ten opzicht van Normaal Amsterdams Peil (NAP) vermeld. De volgende tabel is ontleend aan zo'n Getijdentabel. Hierin kan bijvoorbeeld worden afgelezen dat hoogwater te Harlingen op 1 juli 1989 verwacht werd op zowel het tijdstip 7 uur en 24 minuten ('s ochtends) als het tijdstip 20 uur en 4 minuten ('s avonds).

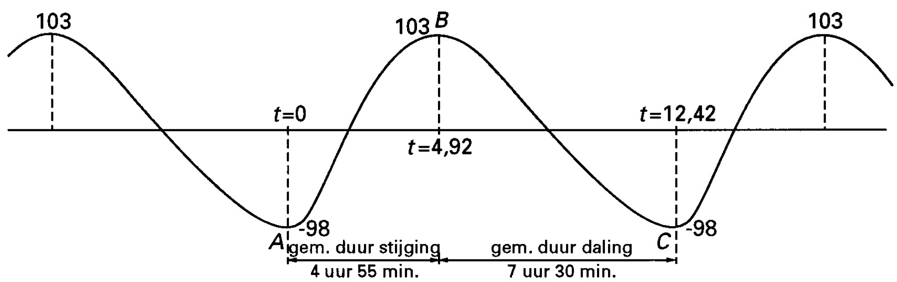
HARLINGEN

juli 1989

datum	laagwater		hoogwater	
	u.min	NAP -cm	u.min	NAP +cm
1 za	2.39	-95	7.24	78
	15.09	-105	20.04	91
2 zo	3.49	-97	8.45	93
	16.15	-109	21.05	88
3 ma	4.40	-99	9.40	105
	17.19	-109	22.15	84

Figuur 6.9

Door de gegevens over zeer lange tijd te middelen, krijgt men voor Harlingen de gemiddelde getijkromme die is weergegeven in de figuur.



Figuur 6.10

Uit de tabel kan voor zes gevallen de tijdsduur worden berekend die verstrijkt van laagwater tot het eerstvolgende hoogwater.

a Onderzoek met een berekening of het gemiddelde van die tijdsduren meer dan twee minuten afwijkt van de in de grafiek vermelde gemiddelde duur van 4 uur en 55 minuten.

De vorm van de grafiek laat duidelijk zien dat een model voor de gemiddelde getijdenbeweging dat uitgaat van één enkele sinusoïde niet erg realistisch is. Beter is het om het stijgende deel AB en het dalende deel BC elk met een afzonderlijke sinusoïde te beschrijven. Omdat de tijdsduur van 4 uur en 55 minuten ongeveer overeen komt met $4{,}92$ uur, geldt voor de waterhoogte (h) voor waarden van t tussen 0 en $4{,}92$ bij benadering:

$h = 100{,}5 \cdot \sin(0{,}64(t - 2{,}46)) + 2{,}5$

b Bereken uitgaande van dit model op welk tijdstip het water het snelst stijgt.

c Waarom gaat dit model niet op voor het stijgende deel van de grafiek na $t = 12{,}42$?

Voor het dalende deel BC geldt bij benadering
$h = 100{,}5 \cdot \sin(a(t - b)) + 2{,}5$.

d Bereken a en b in twee decimalen nauwkeurig.

In de praktijk gebruikt men in combinatie met de Getijdentabellen voor het benaderen van de waterstand soms de *twaalfdelenregel*. Bij opkomend getij let men op de stijging (S) van de waterhoogte gerekend vanaf de laagwaterstand tot de eerstvolgende hoogwaterstand. De periode van opkomend getij wordt verdeeld in zes even grote tijdvakken en men veronderstelt:

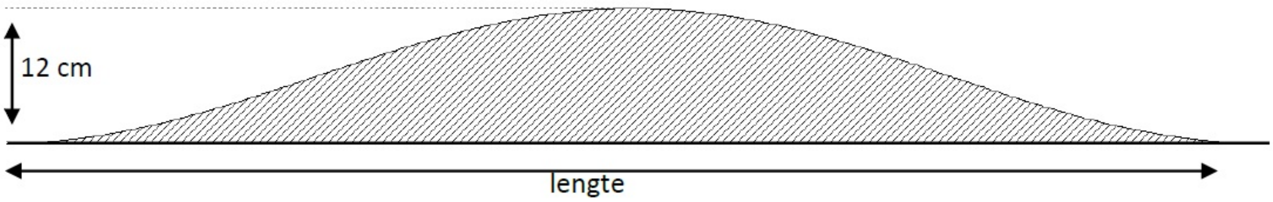
- in het eerste en het zesde tijdvak neemt de waterhoogte gelijkmatig met $\frac{1}{12}$ deel van S toe;
- in het tweede en het vijfde tijdvak neemt de waterhoogte gelijkmatig met $\frac{2}{12}$ deel van S toe;
- in het derde en het vierde tijdvak neemt de waterhoogte gelijkmatig met $\frac{3}{12}$ deel van S toe.

e Benader met de twaalfdenregel en de gegevens van de tabel de waterhoogte te Harlingen op 3 juli 1989 omstreeks 8:00 uur 's ochtends.

(bron: examen wiskunde A vwo 1992, tweede tijdvak)

Opgave 11: Verkeersdrempels

In België zijn vorm en afmetingen van verkeersdrempels sinds 1983 wettelijk vastgelegd. Het zijaanzicht van een verkeersdrempel heeft een sinusvorm. Zie de onderstaande figuur.



Figuur 6.11

Voor de verkeersdrempel van de figuur hierboven, die hoort bij een maximumsnelheid van 30 km/uur, is de volgende formule opgesteld: $h = 0{,}06 + 0{,}06 \sin\left(\frac{1}{2}\pi x - \frac{1}{2}\pi\right)$ Hierin is h de hoogte en x de horizontale afstand vanaf het (linker-)begin van de drempel, beide in meter.

a Bereken hoeveel meter de lengte van deze drempel is.

b Met de formule kun je berekenen over welke lengte deze drempel meer dan 10 cm hoog is. Bereken deze lengte in cm nauwkeurig.

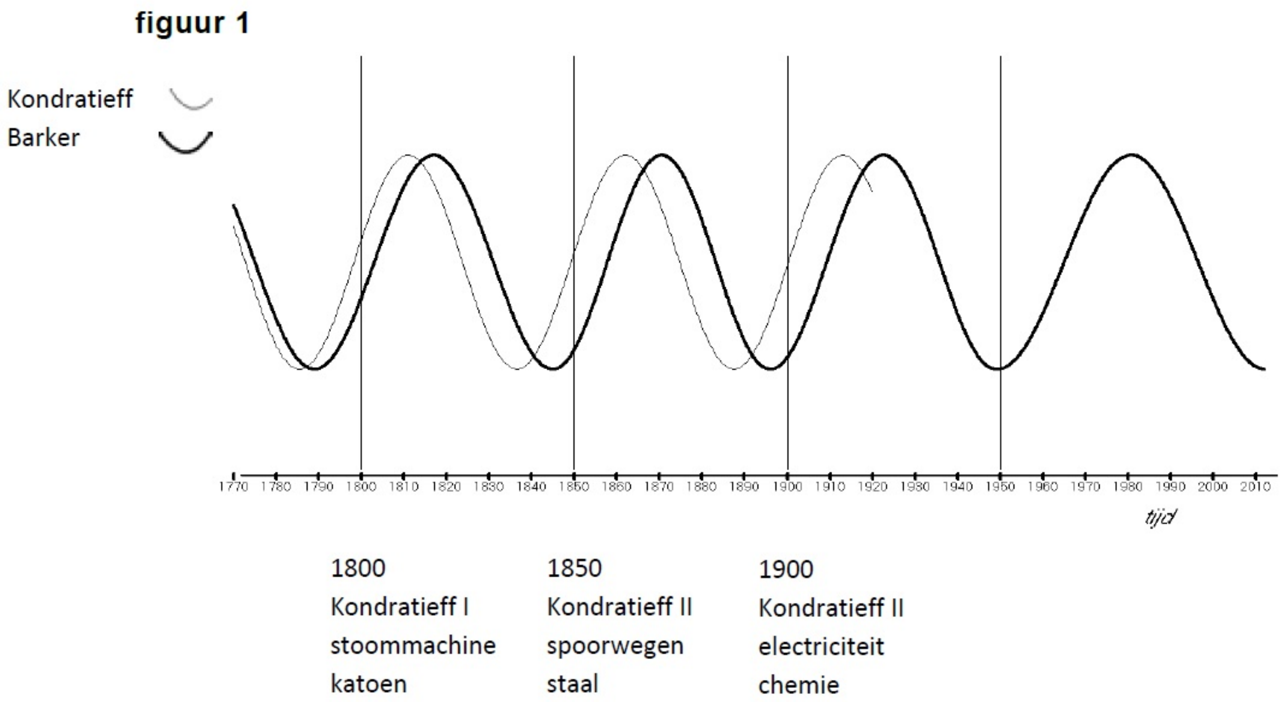
c De helling van de drempel is niet overal even groot. Hoeveel meter van het begin van de drempel ligt het eerste punt waar de drempel maximale helling heeft?

d Een verkeersdrempel die hoort bij een maximumsnelheid van 60 km/uur is 12 meter lang en 14 cm hoog. Daarbij hoort een formule van de vorm: $h = a + b \sin(c(x - d))$ Ook hier is h de hoogte en x de horizontale afstand vanaf het (linker-)begin van de drempel, beide in meter. Hoe groot zijn a, b, c en d?

(bron: voorbeeldexamenopgave vwo A)

Opgave 12: Economische cycli

Golfbewegingen volgens Kondratieff en Barker. In de economie komen vaak golfbewegingen voor: het gaat afwisselend beter en slechter met de economie. Economen proberen deze golfbewegingen te analyseren, onder andere om een volgende economische crisis te kunnen voorspellen. In november 2010 stond hierover een artikel in dagblad Trouw. In figuur 1, gebaseerd op dit artikel, zijn twee verschillende golfbewegingen te zien.



Figuur 6.12

De Russische econoom Kondratieff presenteerde rond 1920 de theorie dat er in de(kapitalistische) wereldeconomie golven of cycli voorkomen met een periode tussen de 50 en 60 jaar: na grote technische vernieuwingen leeft de economie steeds op, om een aantal jaren later weer in een crisis of slechte tijd te belanden. In figuur 1 is onder andere de golfbeweging volgens Kondratieff getekend tot 1920. Als je deze golfbeweging met dezelfde vaste periode ook na 1920 voortzet, wordt de crisis van 2009 hiermee niet goed voorspeld.

a Laat met een redenering gebaseerd op figuur 1 zien dat 2009 volgens Kondratieff niet in een periode van economische neergang zit.

De Amerikaanse beursanalist Barker gaat uit van een iets andere golfbeweging. Ook de golfbeweging volgens Barker is in figuur 1 getekend. Vanaf het dieptepunt in 1949 heeft de golfbeweging volgens Barker een periode die constant is. In figuur 1 is te zien dat de golfbewegingen volgens Kondratieff en Barker steeds meer van elkaar gaan verschillen. In bepaalde perioden laten de beide grafieken zelfs een tegengestelde beweging van de economie zien: de grafiek volgens Barker stijgt, terwijl die van Kondratieff daalt of andersom.

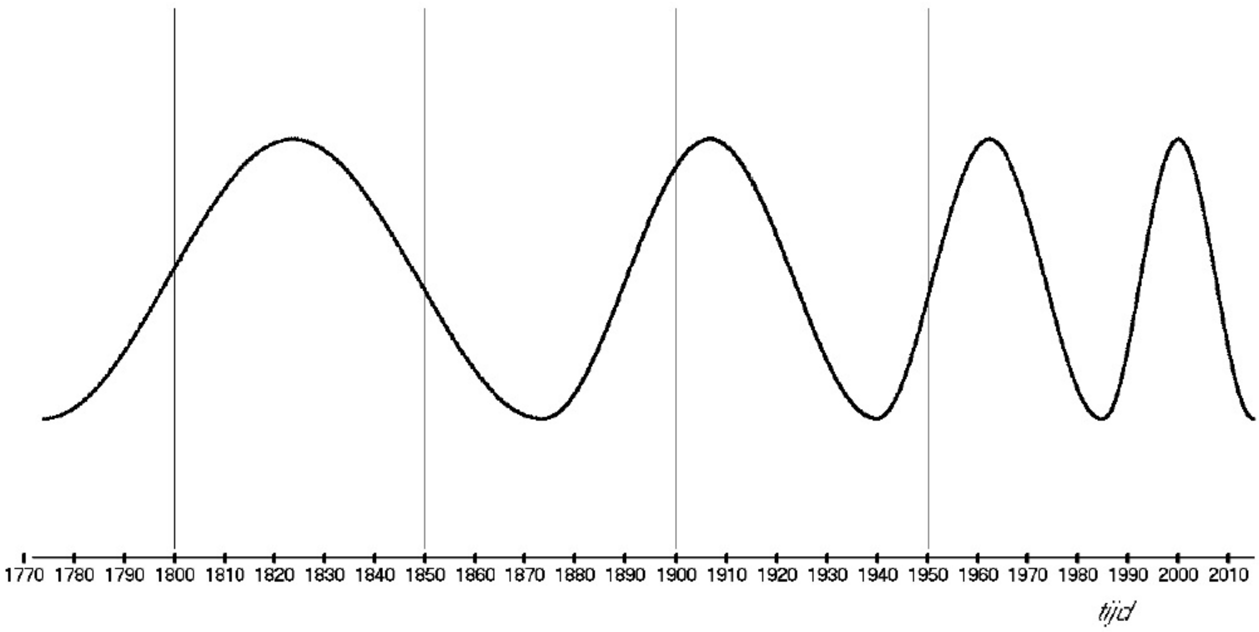
b Onderzoek met behulp van de figuur 1 in welke perioden tussen 1950 en 2050 de grafieken van Kondratieff en Barker een tegengestelde beweging van de economie laten zien.

De golfbeweging volgens Barker kan vanaf het dieptepunt in 1949 benaderd worden met de formule: $B = \sin\left(\frac{2\pi}{63}(t - 1965)\right)$ met t het jaartal. Omdat we hier alleen het stijgen en dalen van de golfbeweging bekijken, doet het er niet toe welke evenwichtsstand en welke amplitude we kiezen. In deze formule is gekozen voor evenwichtsstand 0 en amplitude 1. Voor de golfbeweging volgens Kondratieff kan een soortgelijke formule opgesteld worden.

c Stel een formule voor de golfbeweging volgens Kondratieff op.

In figuur 2 is een derde grafiek getekend: de Slowaakse onderzoeker Smihula ging ook uit van golfbewegingen in de economie, maar volgens hem wordt de periode van deze golven steeds korter. Volgens Smihula begint en eindigt een golf bij een dieptepunt.

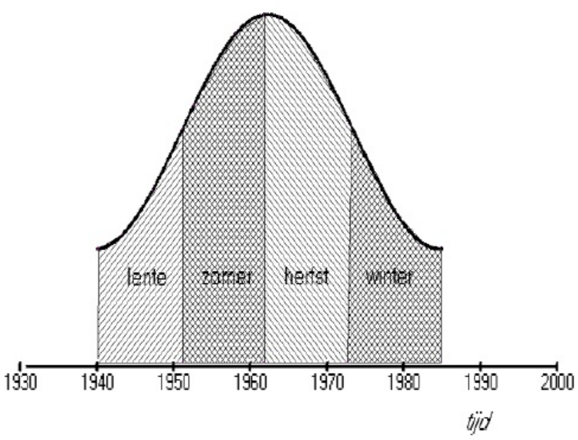
figuur 2



Figuur 6.13

In figuur 3 zie je de golf volgens Smihula tussen 1940 en 1985.De periode van deze golf is verdeeld in vier gelijke delen: deze delen worden respectievelijk lente, zomer, herfst en winter genoemd. De volgende golf volgens Smihula loopt van 1985 tot 2015. De periode van deze golf is tweederde van de periode van de vorige golf. Neem aan dat dit zich na 2015 zo voortzet, dus dat elke nieuwe golf een periode heeft die tweederde is van de vorige.

figuur 3



lente: invloedrijke innovatie trekt economie uit dal;
zomer: bloei van de economie tot oververhitting;
herfst: economie koelt af, koersen stijgen, inflatie daalt, zeepbel opgebouwd;
winter: zeepbel knapt, economische crisis

Figuur 6.14

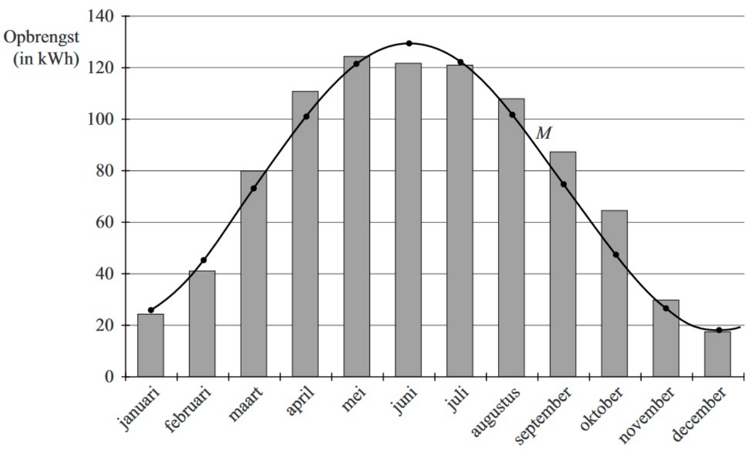
d Bereken in welk ‘seizoen’ (lente, zomer, herfst, winter) het jaar 2040 volgens Smihula zal vallen.

e Als elke nieuwe golf een periode heeft die tweederde is van de vorige, worden de perioden op den duur erg kort. Het is de vraag of dit realistisch is. Bereken in welk jaar er volgens deze regelmaat voor het eerst een periode begint die korter is dan één jaar.

(bron: voorbeeldexamenopgave vwo A)

Opgave 13: Zonnepanelen

Met zonnepanelen kan elektriciteit geproduceerd worden. De opbrengst van zonnepanelen varieert door het jaar heen: in de zomer is de opbrengst groter dan in de winter.
Bekijk het staafdiagram van de gemiddelde maandopbrengsten van een zonnepanelensysteem bij Leiden. Om de gemiddelde maandopbrengsten te bepalen, worden de maandopbrengsten van de laatste 10 jaar gebruikt. De opbrengst wordt gemeten in kilowattuur (kWh).



Figuur 6.15

De gemiddelde maandopbrengsten kunnen benaderd worden door een model: de kromme M in de figuur. De werkelijke gemiddelde maandopbrengst wijkt relatief het meest af in oktober van de door het model voorspelde waarde.

a Licht toe hoe je in de figuur kunt zien dat die relatieve afwijking inderdaad in oktober het grootst is en bereken deze relatieve afwijking.

b De kromme van de gemiddelde maandopbrengst M in de figuur is een sinusoïde. Stel een formule op voor M als functie van de tijd t in maanden. Neem hierbij voor januari $t = 1$.

c Bij een ander zonnepanelensysteem is voor elke dag in het jaar op basis van de gegevens van 10 jaar de gemiddelde dagopbrengst bepaald. De gemiddelde dagopbrengst kan benaderd worden met de formule:

$D = 6{,}34 + 4{,}19 \sin(0{,}0172(t - 74))$

Hierin is D de gemiddelde dagopbrengst in kWh en t de tijd in dagen met $t = 1$ op 1 januari.

Bereken op hoeveel dagen per jaar de gemiddelde dagopbrengst volgens deze formule groter is dan 10 kWh.

(bron: pilotexamen vwo wiskunde A in 2016, tweede tijdvak)

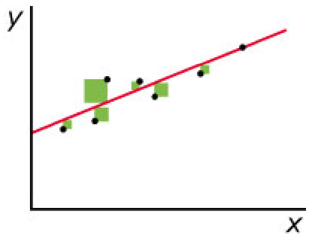
2

Verbanden en verschillen

2.1 Correlatie

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 1.1

Je leert in dit onderwerp

- de (statistische) samenhang tussen twee variabelen uitdrukken in de correlatiecoëfficiënt en deze zowel met de GR als in Excel berekenen;
- de formule van een regressielijn opstellen.

Voorkennis

- soorten statistische variabelen herkennen;
- de begrippen onderzoek, steekproef, populatie en representatief, simulatie.

Verkennen

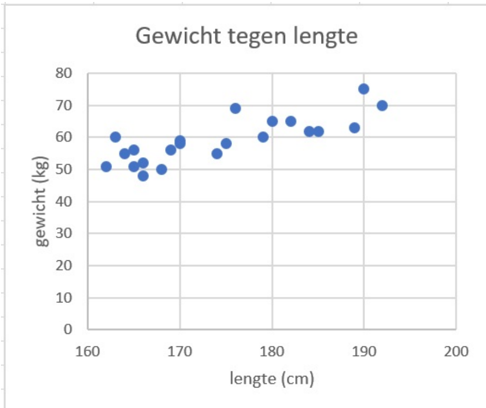
Opgave V1

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?

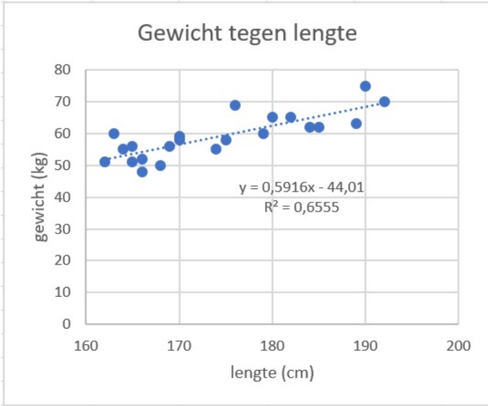


Figuur 1.2

Uitleg 1

Tot nu toe heb je meestal één kenmerk van een populatie afzonderlijk statistisch onderzocht. Maar met behulp van statistiek kun je ook betrouwbare uitspraken doen over een mogelijk statistisch verband tussen twee verschillende kenmerken.

Dit diagram, een puntenwolk of spreidingsdiagram, is de uitkomst van een dergelijk onderzoekje: in dit geval een onderzoek naar het mogelijke verband tussen gewicht en lengte van 22 leerlingen in 4Havo. Zie het werkblad **LengteGewicht22h4.xls**.



Figuur 1.3

De puntenwolk lijkt een richting te hebben: hij loopt grofweg van linksonder naar rechtsboven. Er lijkt dus een vorm van samenhang, van 'correlatie', te zijn tussen deze twee variabelen: grofweg geldt dat naarmate de lengte groter wordt, het gewicht ook groter wordt. Dit lijkt op een lineair verband.

Statistische samenhang betekent niet dat er ook een causaal verband bestaat. Met andere woorden dat een grotere lengte ook een groter gewicht veroorzaakt. Dit is een heel andere vraag en om daar een uitspraak over te doen moet je meer onderzoek doen.

De mate waarin een lineair verband tussen twee variabelen bestaat, wordt aangeduid met de correlatiecoëfficiënt r.

Als r een negatieve waarde heeft, is er sprake van negatieve correlatie en als r een positieve waarde heeft, is er sprake van positieve correlatie. Hoe dichter r bij 1 of -1 ligt, hoe sterker de correlatie tussen de twee variabelen is. Er kan ook sprake zijn van niet-lineaire vormen van samenhang, maar in dat geval heeft de correlatiecoëfficiënt geen zinvolle betekenis.

Je kunt de correlatiecoëfficiënt zowel met de grafische rekenmachine als in Excel berekenen. Ook zijn er vuistregels voor de mate van correlatie. Bekijk het **Practicum**.

Opgave 1

Bekijk de puntenwolk in **Uitleg 1**.

a Leg uit hoe je aan de puntenwolk kunt zien dat r ongelijk aan 0 is.

b Leg uit waaraan je kunt zien dat r positief zal zijn.

c Open het bestand **LengteGewicht22h4.xls** en bepaal de waarde van r^2 die je ook in de figuur ziet. Doe dit met behulp van Excel, zie het **Practicum**.

d Bereken de waarde van r en trek een conclusie op basis van de vuistregels.

Opgave 2

Er is altijd wel iemand nieuwsgierig naar de meest bijzondere weetjes. Zo zou je je kunnen afvragen: is er een verband tussen het soort 'smiley' dat mensen gebruiken in hun communicatie op sociale media en het weertype op het moment van communicatie?
Als je hier gegevens over zou verzamelen, is dit een voorbeeld van de gegevenstabel.

weertype	smiley
zonnig	🙂
stormachtig	☹
regenachtig	😐

Figuur 1.4

a Is een puntenwolk een bruikbare presentatiewijze voor deze gegevens?

Beargumenteer je antwoord, bijvoorbeeld met behulp van voorbeeldschetsjes van puntenwolken.

b Zou er een verband mogelijk kunnen zijn tussen deze twee variabelen?

Zo ja, wat zou dan ongeveer de waarde zijn van r?

c Zou hier ook van causaliteit sprake kunnen zijn? Zo ja: welke variabele is dan de veroorzaker van de waarde van de andere variabele?

Uitleg 2

Als er sprake is van lineaire correlatie tussen twee kwantitatieve variabelen, dan is het mogelijk om de formule op te stellen van een speciaal type trendlijn, de ‘regressielijn’.

De regressielijn is zo opgebouwd dat gemiddeld voor alle punten in de puntenwolk geldt dat de verticale afstand van het punt tot de regressielijn zo klein mogelijk is.

Dit is het geval als voor de richtingscoëfficiënt a van de regressielijn voor variabelen x en y geldt:

$$a = r_{xy} \cdot \frac{\sigma_y}{\sigma_x}$$

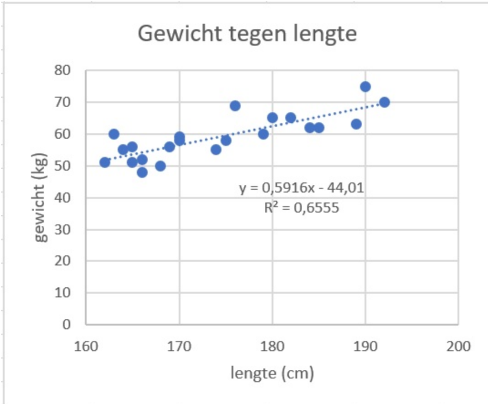
waarin r_{xy} de correlatiecoëfficiënt is van x en y en σ_x de standaardafwijking van variabele x is en σ_y de standaardafwijking van variabele y.

De regressielijn loopt altijd door het punt $(\overline{x}, \overline{y})$ en daarmee kun je de bijbehorende formule opstellen.

Het is heel verleidelijk om ervan uit te gaan dat y ook echt volledig afhankelijk is van x en dat er dus sprake is van én volledige correlatie én van causaliteit. Dit is echter meestal niet het geval!

De formule van een regressielijn wordt opgesteld om een schatting te maken. Bijvoorbeeld: welk gewicht zal naar schatting horen bij een lengte van $2{,}00$ m? De regressielijn van deze twee variabelen geeft op die vraag een betrouwbaar antwoord.

In het **Practicum** kun je zien hoe je de formule van een regressielijn opstelt met behulp van de grafische rekenmachine en Excel.



Figuur 1.5

Opgave 3

In **Uitleg 2** zie je de puntenwolk bij het werkblad **LengteGewicht22h4.xls**. Ook de regressielijn (of trendlijn) is getekend. Voer nu de gegevens van deze 22 leerlingen in je grafische rekenmachine in.

a Bereken de gemiddelde lengte $\overline{x}$ en de bijbehorende standaardafwijking σ_x en bereken ook het gemiddelde gewicht $\overline{y}$ en de bijbehorende standaardafwijking σ_y.

Ga uit van een lineaire correlatie tussen x en y. En gebruik de correlatiecoëfficiënt die je eerder hebt berekend of gebruik r^2 in de figuur.

b Stel met behulp van de gegevens bij a een formule op voor de lineaire regressielijn die je in de figuur ziet.

Je kunt ook je grafische rekenmachine zelf de correlatiecoëfficiënt en de formule van de regressielijn laten opstellen. Ga in het **Practicum** na hoe dat gaat.

c Maak met je grafische rekenmachine de puntenwolk die je in de figuur in de uitleg ziet en laat de formule voor de trendlijn bepalen.

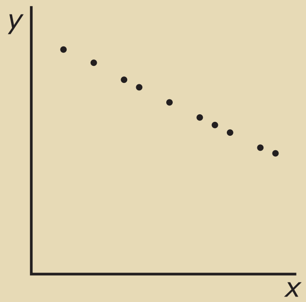
d Voorspel met behulp van de bij c gevonden regressielijn het gewicht van een scholier die $2{,}00$ m lang is.
Waarom zal deze voorspelling niet erg betrouwbaar zijn ondanks de hoge correlatie?

Theorie en voorbeelden

Om te onthouden

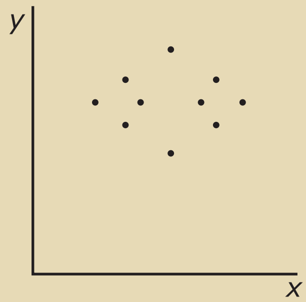
Met behulp van statistiek kunnen uitspraken worden gedaan over verbanden tussen twee verschillende variabelen. Als de twee variabelen kwantitatief zijn, kan hun mogelijke verband in beeld worden gebracht met een puntenwolk.

De **correlatiecoëfficiënt** r geeft de mate van **correlatie**, dat wil zeggen van samenhang, tussen de twee variabelen. Deze waarde is een getal tussen -1 en 1.



$r \approx -1$

goede correlatie



$r \approx 0$

geen correlatie

$r \approx 0{,}6$

matige correlatie

$r \approx 1$

goede correlatie

Figuur 1.6

De correlatie tussen twee variabelen wordt beter naarmate r dichter bij 1 of -1 ligt. Als r gelijk is aan 0 dan is er geen correlatie tussen de variabelen. In het **Practicum** zie je enkele vuistregels voor de mate van correlatie tussen twee kwantitatieve variabelen. Je vindt ze ook in dit **Formuleoverzicht**.

Als de correlatie redelijk is, kunnen met behulp van de **regressielijn** betrouwbare schattingen worden gemaakt. De richtingscoëfficiënt a van de regressielijn voor variabelen x en y bereken je met de formule:

$$a = r_{xy} \cdot \frac{\sigma_y}{\sigma_x}$$

waarin r_{xy} de correlatiecoëfficiënt is van x en y en σ_x de standaardafwijking van variabele x is en σ_y de standaardafwijking van variabele y.

De volledige formule vind je uit het gegeven dat de regressielijn door $(\overline{x}, \overline{y})$ gaat.

Zowel de correlatiecoëfficiënt als de formule voor de regressielijn bereken je in de praktijk met de grafische rekenmachine of met Excel. Bekijk het **Practicum**.

Het kan ook zo zijn dat de waarden van de ene variabele veroorzaakt worden (geheel of gedeeltelijk) door die van een andere variabele. Er is dan sprake van een **causaal verband**. Als er statistische samenhang is tussen twee variabelen betekent dat niet dat er ook een causaal verband bestaat. De correlatiecoëfficiënt en de regressielijn zeggen helemaal niets over een mogelijk causaal verband tussen twee variabelen.

Voorbeeld 1

De makers van een veelgebruikte weerapp en de uitbaters van een dierentuin doen een onderzoek naar een eventueel verband tussen de kans op regen die de weerapp voor een zaterdag voorspelt en het aantal dierentuinbezoekers op die zaterdag.

Na een aantal weken hebben ze de volgende gegevens verzameld:

kans op regen (%)	95	5	20	95	85	55	5	30
aantal bezoekers	876	3102	3263	1278	1491	2984	4312	2916

Tabel 1.1

Wat is de correlatiecoëfficiënt van de kans op regen en het aantal bezoekers in deze steekproef? Maak eventueel gebruik van de grafische rekenmachine of van Excel.

Antwoord

Bereken de correlatiecoëfficiënt r door voor beide variabelen een tabel op de grafische rekenmachine of in Excel (zie **Practicum**) te maken en vervolgens de berekening van de correlatiecoëfficiënt uit te laten voeren.

Je vindt: $r \approx -0{,}932$.

Opgave 4

Bekijk **Voorbeeld 1**.

a Wat betekent het voor de samenhang van de waarden in de steekproef dat de correlatiecoëfficiënt een negatief getal is?

b Wat gebeurt er met de correlatiecoëfficiënt als je de lijsten/kolommen omwisselt? Licht je antwoord toe.

c Beargumenteer wat er met de correlatiecoëfficiënt gebeurt als er nog een nieuwe meting bijkomt, namelijk 4000 bezoekers bij een kans op regen van 100%.

Opgave 5

Bekijk **Voorbeeld 1**.

Behalve de regenkans houdt men ook het voorspelde aantal uren zon bij. Bij de dierentuin zit ook een vennetje waarin kinderen kunnen zwemmen en dit aantal kinderen heeft men bijgehouden:

aantal uren zon	0,7	7,6	2,3	1,1	2,4	4,9	8,1	3,6
aantal zwemmende kinderen	6	802	121	6	48	123	964	32

Tabel 1.2

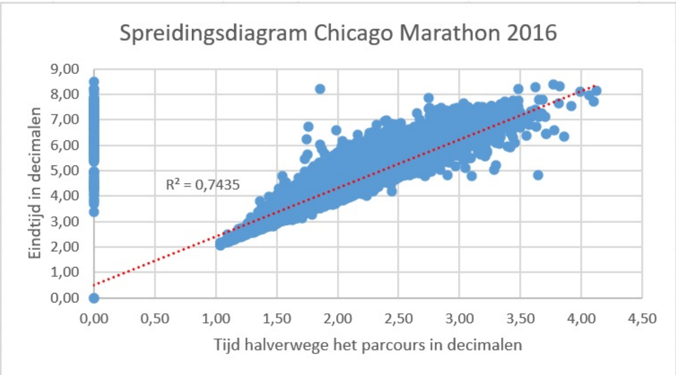
a Bereken de correlatiecoëfficiënt.

b Is dit een betekenisvol correlatieonderzoek? Zegt de correlatiecoëfficiënt in dit geval ook echt iets over een mogelijk causaal verband tussen het voorspelde aantal uren zon en het aantal kinderen dat gaat zwemmen?

Voorbeeld 2

Hier zie je een puntenwolk van alle deelnemers aan de Chicago Marathon van 2016. De eindtijd is uitgezet tegen de looptijd halverwege het parcours. Bekend is nog:

- looptijd halverwege: gemiddelde $\approx 2{,}13$ uur met een standaardafwijking van $\approx 0{,}40$ uur.
- totale looptijd: gemiddelde $\approx 4{,}55$ uur met een standaardafwijking van $\approx 0{,}92$ uur.



Figuur 1.7

Welke formule kun je opstellen voor de in de figuur getekende regressielijn?

Antwoord

Noem de looptijd halverwege h en de eindtijd E, beide in uur.

De correlatiecoëfficiënt is $r_{hE} = \sqrt{0{,}7435} \approx 0{,}86$.

De regressielijn gaat door $(2{,}13; 4{,}55)$.

De richtingscoëfficiënt ervan is $a = r_{hE} \cdot \frac{\sigma_E}{\sigma_h} \approx 0{,}86 \cdot \frac{0{,}92}{0{,}40} \approx 1{,}98$.

De regressielijn is dus $E \approx 1{,}98 \cdot h + 0{,}33$.

Hiermee kun je voorspellingen doen.

Opgave 6

Bekijk de figuur in **Voorbeeld 2**.

a Reken zelf de formule voor de regressielijn na.

b Als je de gegevens over gemiddelde tijden en de bijbehorende standaardafwijking niet hebt, kun je toch wel een vergelijking maken bij de getekende regressielijn. Hoe?

c Welke eindtijd zal een loper die aan deze marathon deelnam hebben als zijn tijd halverwege 3 uur was?

d Welke betekenis hebben de punten die op de verticale as liggen?

e Is hier sprake van een causaal verband of een statistisch verband?

f Hiermee kun je heel goed ‘voorspellingen achteraf’ doen.
Maar kan zoiets ook echt nut hebben?

Verwerken

Opgave 7

Bekijk de tabel met de correlatiecoëfficiënten die telkens de mate van samenhang aangeven tussen de lengte van een vrouw en een andere lichaamsmaat van dezelfde vrouw.

	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 1.3

Gebruik de vuistregels voor de mate van correlatie, zie het **Practicum**.

a Welke variabele heeft een sterke samenhang met lengte?

b Welke variabelen hebben een matige samenhang met lengte?

c Welke variabelen hebben een zwakke samenhang met lengte?

Opgave 8

Schets een mogelijke puntenwolk voor kwantitatieve variabelen X en Y als voor hun correlatiecoëfficiënt geldt:

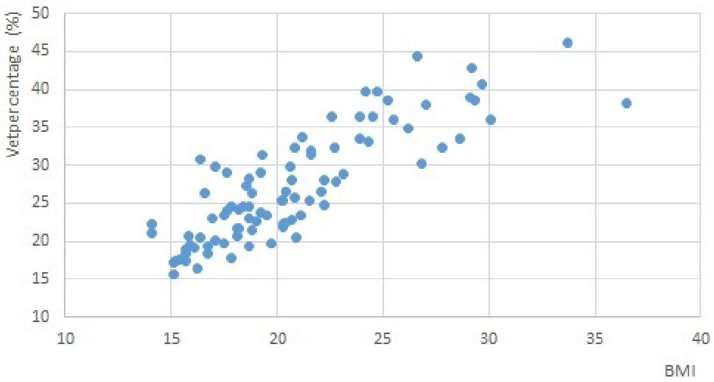
a $r \approx 0$

b $r = -1$

c $0{,}3 \leq r < 0{,}7$

Opgave 9

Bekijk de puntenwolk met de resultaten van een onderzoek naar het BMI en vetpercentage 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 1.8

a Is er een statistische samenhang?

b Is er een oorzakelijk verband?

c Welke statistische gegevens heb je van de twee variabelen gewicht en vetpercentage nodig om de regressielijn voor deze puntenwolk te kunnen maken?

Opgave 10

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 va der v in cm	173	168	178	170	180	165	185	175	180	178	183	188
lengte zo z in cm	180	175	180	173	183	175	180	173	188	178	180	185

Tabel 1.4

a Is er sprake van een positieve of een negatieve correlatie? Wat betekent dit in de praktijk?

b Stel de regressielijn op van z op v bij deze gegevens.

c Als een bepaalde vader 1,77 m lang is, hoe lang zou dan zijn oudste zoon moeten zijn?

Opgave 11

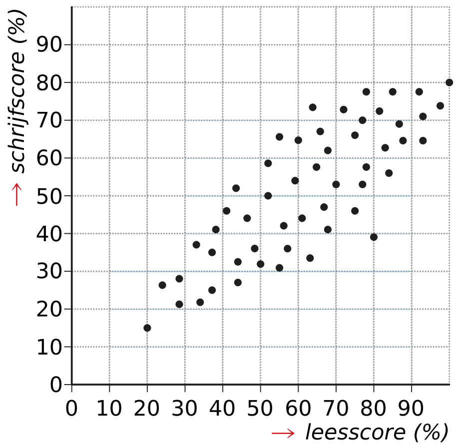
Een basisschool heeft een leestest en schrijftest Nederlands afgenomen bij de leerlingen in groep acht. De resultaten zijn verwerkt in een puntenwolk.

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 1.9

Opgave 12

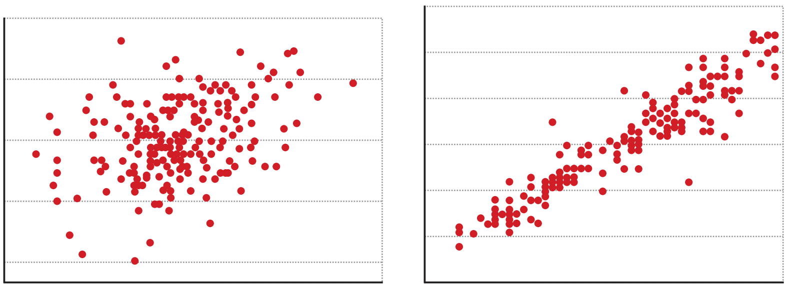
Om het verband tussen het gewicht G (in pounds) en de braadtijd voor kalkoenen te onderzoeken, werd onder gelijke omstandigheden nagegaan hoeveel minuten t het duurde tot het binnenste van een kalkoen de temperatuur van 85 °C bereikte. Er werden diverse kalkoenen aan dit onderzoek onderworpen. Ze hadden een gemiddeld gewicht van $15{,}24$ pounds met een standaardafwijking van $6{,}07$. Voor de waarden van t vonden de onderzoekers een gemiddelde van $205{,}4$ minuten met een standaardafwijking van $59{,}1$.
De regressielijn van t op G had de vergelijking: $t = 9{,}65G + 58{,}40$.

Hoeveel bedroeg de correlatiecoëfficiënt?

Toepassen

Opgave 13: 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:

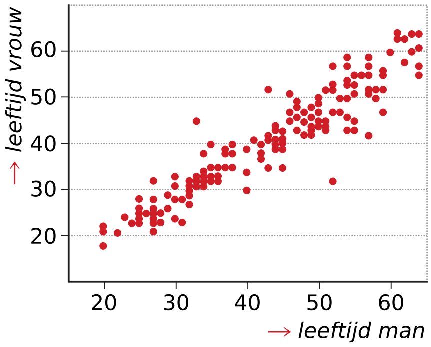


Figuur 1.10

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 1.11

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 1.5

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.

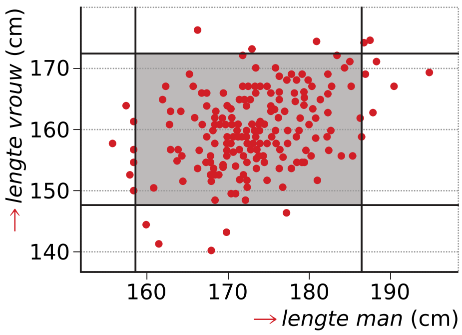
d Leg uit hoe je aan deze waarden komt.

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 1.12

Testen

Opgave 14

In een Amerikaans laboratorium heeft men proeven genomen waarbij gelet werd op het verband tussen de hoogte van de bewaartemperatuur F in graden Fahrenheit en de werkzaamheid W van een bepaald geneesmiddel. Bij temperaturen van 30°, 50°, 70° en 90° (Fahrenheit) werden drie porties van gelijk gewicht uit eenzelfde productie 20 dagen bewaard. Na deze periode werd op identieke wijze de werkzaamheid van de porties vastgesteld. De werkzaamheid werd uitgedrukt in percentages van de werkzaamheid zoals die was voor het bewaren.

Bewaartemperatuur F (°F)	30	50	70	90
Werkzaamheid W (%)	39, 42, 35	32, 26, 33	19, 27, 23	14, 19, 21

Tabel 1.6

a Verwerk deze gegevens in een spreidingsdiagram en bereken de correlatiecoëfficiënt. Is er sprake van een correlatie tussen W en F?

b Stel de vergelijking op van de regressielijn van W op F. Waarom ligt deze regressielijn meer voor de hand dan die van F op W?

Het verband tussen de temperatuur in graden Fahrenheit F en die in graden Celsius C wordt zoals bekend gegeven door: $F = 1{,}8C + 32$.

c Stel nu een vergelijking op van de regressielijn van W op C.

d Is de correlatiecoëfficiënt tussen W en C anders dan die tussen W en F? Verklaar je antwoord.

Uit andere experimenten is gebleken dat de werkzaamheid bij een vaste bewaartemperatuur exponentieel afhangt van de lengte van de bewaarperiode.

e Schat de gemiddelde werkzaamheid van porties die 40 dagen bij een temperatuur van 20 °C zijn bewaard.

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 HPprime**

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'.

- **Correlatie en regressie**

OPMERKING:
In de praktijk worden de volgende **vuistregels gebruikt voor de correlatie r tussen twee kwantitatieve variabelen**:

- 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.

Je vindt ze ook in dit **Formuleoverzicht**.

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 meetniveau's, 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.

Je leert in dit onderwerp

- het verschil tussen twee kwalitatieve variabelen te beschrijven en te interpreteren;
- te werken met vuistregels om deze verschillen te beschrijven.

Voorkennis

- soorten statistische variabelen herkennen;
- de begrippen onderzoek, steekproef, populatie en representatief.

Verkennen

Opgave V1

Langzamerhand kom je in de buurt van echt statistisch onderzoek. Hier een klein voorproefje.

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

Uitleg 1

Bekijk de tabel. In deze kruistabel is de waardering, de voorstellingsbeleving, gegeven van een groep leerlingen die een toneelvoorstelling heeft bezocht, uitgesplitst per examenprofiel.

	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

Tabel 2.1

Beide variabelen in deze kruistabel zijn kwalitatief. Variabele voorstellingsbeleving heeft oplopende waarden en heet daarom een ordinale variabele.

Op grond van deze tabel lijkt het alsof de EM-leerlingen de voorstelling hoger hebben gewaardeerd dan de NG-leerlingen, maar die groepen zijn niet even groot. Om goed te kunnen vergelijken bij groepen die niet even groot zijn, moet je de absolute frequenties eerst omzetten in relatieve frequenties.

Bekijk de tabel. In deze tabel staan onder p de percentages en onder cp de cumulatieve percentages.

	EM			NG			
	aantal	p	cp	aantal	p	cp	V_{cp}
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			

Tabel 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 ordinale kwalitatieve variabelen de volgende afspraak gemaakt:
het maximale verschil van de cumulatieve percentages $\max V_{cp}$ wordt gebruikt als maat voor het verschil tussen twee groepen.

Vuistregels voor de interpretatie van de $\max V_{cp}$ zijn:

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

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

Opgave 1

Bekijk **Uitleg 1**.

a Leg uit waarom de variabele voorstellingsbeleving een kwalitatieve variabele is en geen kwantitatieve variabele.

b Ordinale kwalitatieve variabelen zie je vaak terug in enquêtes.

Bedenk zelf een ordinale kwalitatieve variabele die in een enquête zou kunnen voorkomen en geef ook de mogelijke waarden en hun volgorde.

Opgave 2

Bekijk **Uitleg 1**.
Vergelijk de voorstellingsbeleving van de NG-leerlingen met de NT-leerlingen.
Bereken eerst de bijbehorende $\max V_{cp}$.

Uitleg 2

Om te onderzoeken of vitamine C helpt tegen verkoudheid kregen 139 personen vitamine C toegediend en 140 personen een placebo (fopmiddel). De proefpersonen wisten niet in welke categorie ze zaten.

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

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

Tabel 2.3

Met deze aantallen kun je de groepen vergelijken. Je gebruikt hierbij het getal *phi* ofwel φ:

- Noem de waarden in de vier cellen van de kruistabel a, b, c en d.
- 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: $\varphi \approx -0{,}100$.

Trek conclusies over de mate van verschil tussen beide groepen 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'.

Volgens deze vuistregels is in dit onderzoek sprake van een gering verschil.

Opgave 3

Over het algemeen maak je een kruistabel om een beeld te krijgen van het verband en/of verschil tussen twee variabelen. De kruistabel in **Uitleg 2** bevat absolute frequenties en dat is de minst inzichtelijke manier om een kruistabel samen te stellen.

Verander de absolute frequenties in de kruistabel van de uitleg in relatieve frequenties nadat je eerst hebt besloten of je horizontaal percenteert, verticaal of beide (wat vind je zinvol in deze situatie?) en geef aan of je daarmee ook al iets kunt zeggen over de mate van verschil tussen de twee variabelen.

Opgave 4

Bekijk de kruistabel in **Uitleg 2**. Er is een vaste plek in de kruistabel voor de frequenties a, b, c en d die gebruikt worden om verschilmaat φ te berekenen. Je kunt je afvragen of het dan ook van belang is hoe je de kruistabel samenstelt.

Maak rijen van de kolommen in de kruistabel en vice versa en bepaal voor deze nieuwe kruistabel de waarde van φ.
Is deze waarde anders dan de waarde die in de uitleg is berekend?

Theorie en voorbeelden

Om te onthouden

In statistisch onderzoek worden vaak twee groepen onderzocht en met elkaar vergeleken: in welke mate is er verschil? Er zijn legio methodes om inzicht in het verschil te krijgen, al naar gelang het soort kenmerk en dus het type variabele waarbij het verschil tussen de twee groepen wordt vergeleken.

Als de variabele kwalitatief is maar wel een ordening kent in de mogelijke waarden, dan spreek je van een **ordinale kwalitatieve variabele**. Bereken per groep de cumulatieve relatieve frequenties van de mogelijke waarden. Bepaal vervolgens per waarde het verschil in cumulatieve relatieve frequentie tussen de twee groepen. De waarde met het hoogste verschil heet officieel **max** V_{cp}.

Vuistregels voor de interpretatie van $\max V_{cp}$ zijn (zie ook dit **Formuleoverzicht**):

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

	groep 1	groep 2
waarde 1	a	b
waarde 2	c	d

Tabel 2.4

Als de variabele kwalitatief is en geen ordening heeft spreek je van een **nominale kwalitatieve variabele**. Als hij slechts twee waarden heeft maak je een 2-bij-2 **kruistabel** van de absolute frequenties.

Bereken daarmee de verschilmaat **phi** (φ) volgens de formule:

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

Vuistregels voor de interpretatie van φ zijn (zie ook dit **Formuleoverzicht**):

- Als $-0{,}2 \leq \varphi \leq 0{,}2$ dan is het verschil is ‘gering’.
- Als $-0{,}4 \leq \varphi < -0{,}2$ of $0{,}2 < \varphi \leq 0{,}4$ dan is het verschil ‘middelmatig’.
- Als $\varphi < -0{,}4$ of $\varphi > 0{,}4$ dan is het verschil ‘groot’.

Voorbeeld 1

Aan een aantal hotelgasten is gevraagd om 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 2.5

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.

Welke mate van verschil is er tussen de Europeanen en de Aziaten en hun oordeel over het dinerbuffet?

Antwoord

In deze situatie gaat het om twee groepen, Europeanen en Aziaten, en het verschil in hun oordeel over een dinerbuffet. Variabele *buffetbeoordeling* is een ordinale nominale variabele. Dat betekent dat je verschilmaat $\max V_{cp}$ kunt gebruiken om een conclusie over het verschil tussen beide groepen te trekken.

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 2.6

Er geldt $\max V_{cp} = 17{,}2\%$ en dus $\max V_{cp} \leq 20\%$.

Volgens de vuistregels is het verschil gering.

Opgave 5

Bekijk de tabel over de waardering van een dinerbuffet in **Voorbeeld 1**.
Deze opgave betreft het verschil in beoordeling van dit dinerbuffet door enerzijds Europeanen en anderzijds Noord-Amerikanen.

a Wat is je conclusie over dat verschil als je alleen kijkt naar de absolute frequenties die in de tabel staan?

b Wat is je conclusie over dat verschil als je alleen kijkt naar de relatieve frequenties die in de tabel staan?

c Welke mate van verschil is er op statistisch verantwoorde wijze te vinden tussen Europeanen en Noord-Amerikanen en hun oordeel over het dinerbuffet?

Opgave 6

Op een nieuwssite staat een onderzoek naar de beoordeling van een nieuwe app.
Bekijk de resultaten.

	12- tot 15-jarigen	16- tot 19-jarigen
1 = leuk	31	22
2 = niet leuk	27	38

Tabel 2.7

a Welke mate van verschil in beoordeling van de app zit er tussen de groep van 12- tot 15-jarigen en de groep van 16- tot 19-jarigen?

b Was je tot dezelfde conclusie gekomen als je alleen de absolute frequenties had bekeken?

c De mensen die het onderzoek uitgevoerd en/of gepubliceerd hebben, gaven beoordeling 'leuk' de waarde 1 en beoordeling 'niet leuk' waarde 2. Dit hadden ze ook andersom kunnen doen.

Zou dat hebben uitgemaakt voor de bepaling van de mate van verschil?

Licht je antwoord toe met berekeningen.

Voorbeeld 2

Onderzoek onder een groep mannen en vrouwen naar de bijwerkingen van een medicijn gaf deze resultaten.

Bepaal de mate van verschil tussen het ontstaan van bijwerkingen bij mannen en bij vrouwen.

	bijwerking	geen bijwerking
mannen	86	14
vrouwen	61	39

Tabel 2.8

Antwoord

Het gaat hier om het verschil tussen een groep mannen en vrouwen. *Bijwerking* is een nominale variabele met twee mogelijke waarden: er kan een 2-bij-2 tabel van worden gemaakt en om echt een uitspraak over de mate van verschil te kunnen doen, kun je verschilmaat φ berekenen met behulp van de gegeven formule.

Je vindt: $\varphi \approx 0{,}28$

Volgens de vuistregels voor de interpretatie van φ is er dus sprake van een middelmatig verschil tussen de bijwerkingen bij mannen en de bijwerkingen bij vrouwen.

Opgave 7

Op een nieuwssite staat een onderzoekje naar de beoordeling van een nieuwe app.
Bekijk de resultaten.

	12- tot 15-jarigen	16- tot 19-jarigen
1 = leuk	31	22
2 = niet leuk	27	38

Tabel 2.9

In **Opgave 6** heb je een verschilmaat berekend, namelijk $\max V_{cp}$. Hiermee bleek dat er sprake was van een gering verschil tussen de beide groepen omtrent hun beoordeling van de app.

a Kijk nu met behulp van verschilmaat φ in welke mate er verschil is tussen deze groepen en vergelijk de conclusie op basis van φ met de conclusie op basis van $\max V_{cp}$.

b Leg uit waarom je voor deze gegevens zowel $\max V_{cp}$ als φ kunt gebruiken als verschilmaat. Zou dit ook kunnen met de gegevens uit **Voorbeeld 2**?

Opgave 8

Bekijk de gegevens in de tabel.

a Bereken verschilmaat φ voor deze gegevens en trek daarmee een conclusie over het verschil tussen de twee groepen.

b Laat zien dat je bij deze gegevens helemaal geen verschilmaat hoeft te gebruiken: je kunt namelijk met behulp van de absolute frequenties (of met de relatieve frequenties, als je het niet zo snel ziet) al laten zien dat er helemaal geen verschil is tussen de twee groepen.

	groep 1	groep 2
waarde 1	45	62
waarde 2	135	186

Tabel 2.10

Verwerken

Opgave 9

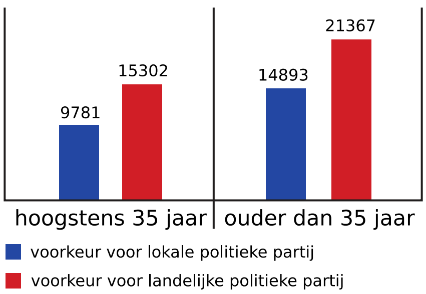
Bekijk de figuur. Je ziet de voorkeuren voor lokale en landelijke politieke partijen in een Nederlandse gemeente afhankelijk van de leeftijdsgroep.

a Maak een kruistabel voor deze gegevens.

b Leg uit waarom je de mate van verschil tussen het stemgedrag van stemgerechtigde inwoners die hoogstens 35 jaar zijn en degenen die ouder zijn dan 35 jaar in ieder geval met verschilmaat φ kunt bepalen.

De waarde van φ voor deze gegevens is -$0{,}02$.

c Bepaal op basis van deze waarde met de vuistregels de mate van verschil tussen het stemgedrag van stemgerechtigde inwoners die hoogstens 35 jaar zijn en degenen die ouder zijn dan 35 jaar.



Figuur 2.1

Opgave 10

Bekijk de drie 2-bij-2 kruistabellen met absolute frequenties die verzameld zijn voor een statistisch onderzoek.

	school A	school B
jonge docenten	83	78
oude docenten	115	47

kruistabel 1

	workshop tekenen	workshop muziek
klas A	20	10
klas B	14	7

kruistabel 2

uurloon voor vakkenvuller (€)	supermarkt A	supermarkt B
$2{,}5- < 3{,}0$	12	19
$3{,}0- < 3{,}5$	7	15

kruistabel 3

Tabel 2.11

Beantwoord de vragen voor ieder van de drie kruistabellen.

a Geef bij elke tabel aan welk verschil tussen de twee groepen er waarschijnlijk wordt bestudeerd. Geef ook het type variabele dat onderzocht wordt.

b Heeft de onderzoeker bij deze gegevens verschilmaat $\max V_{cp}$ en/of φ nodig voor een conclusie over het verschil tussen de twee groepen of volstaat hier het vergelijken van percentages? Licht je antwoord toe.

c Als de onderzoeker verschilmaat $\max V_{cp}$ of φ wil gebruiken om de mate van verschil tussen de groepen te bepalen, welke komt/komen dan in aanmerking voor deze gegevens? Licht je antwoord toe.

Opgave 11

Bekijk de tabel met de relatieve frequenties van het hoogst genoten onderwijs in twee regio's in een land.

	regio west	regio oost
basisschool	6	3
lager voortgezet onderwijs	11	19,5
hoger voortgezet onderwijs	5,5	8
MBO	40,5	37
HBO	28	20,5
WO	9	12

Tabel 2.12

Welke mate van verschil in hoogst genoten onderwijs bestaat er tussen de beide regio's?

Opgave 12

In een oud rapport van Kennisnet over ICT-gebruik in het onderwijs stonden deze gegevens over het gebruik van ICT-apparaten op scholen in het voortgezet onderwijs.

Leg uit in hoeverre je een betrouwbare uitspraak kunt doen omtrent het verschil in gebruik van ICT-apparaten tussen schooljaar 2012-2013 en schooljaar 2014-2015 op basis van de in a t/m d genoemde gegevens.

Geef aan welk verschil je ermee kunt onderzoeken (als het kan) en hoe betrouwbaar de conclusie is die je eruit kunt trekken.

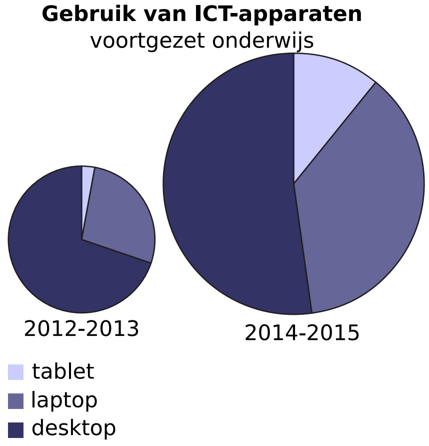
a De vorm van de cirkeldiagrammen.

b De percentages die bij iedere cirkelsector horen.

c $\max V_{cp}$

d φ

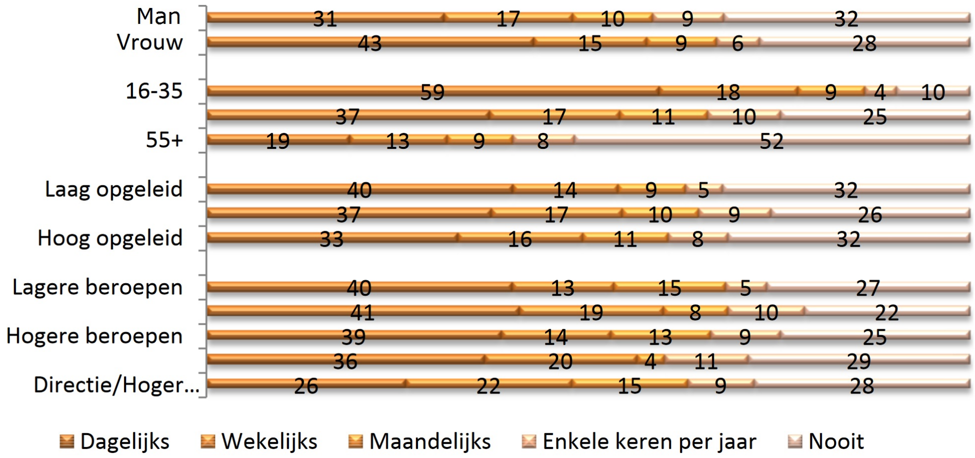
(bron: Kennisnet)



Figuur 2.2

Opgave 13

Bekijk de figuur. In een rapport van de Universiteit Twente staat een overzicht over het bezoek aan sociale netwerksites.



Figuur 2.3

Onderzoek de mate van verschil tussen mannen en vrouwen wat betreft de frequentie van hun bezoek aan sociale netwerksites.

Opgave 14

Een farmaceutisch bedrijf heeft een studie gedaan naar het ontstaan van hartfalen tijdens het slikken van hun nieuwe combinatietablet met aspirine. Een journalist is erg geïnteresseerd in de resultaten van deze studie, maar het bedrijf is niet erg scheutig met informatie. Hij weet ondertussen wel het volgende:

- de steekproefgrootte van de studie is 22071;
- het totale aantal personen in de steekproef dat géén last van hartfalen kreeg tijdens de studie is 21693;
- het aantal personen in de studie dat aspirine slikte en géén last van hartfalen kreeg tijdens de studie is 10898;
- het aantal personen in de studie dat géén aspirine slikte maar wel last van hartfalen kreeg tijdens de studie is 239.

Volgens de journalist heeft hij genoeg aan deze gegevens om zelf de mate van verschil te berekenen tussen het ontstaan van hartfalen tussen de groep die wel aspirine slikte en de groep die dat niet deed.

a Laat zien dat dat inderdaad kan en trek een conclusie omtrent de mate van het genoemde verschil.

Stel je de situatie voor dat de journalist alleen de absolute frequenties te horen kreeg zoals getoond in deze tabel:

	hartfalen	geen hartfalen
aspirinegebruikers		10898
geen aspirinegebruikers	239	10795

Tabel 2.13

Hij kan dan de minimale hoeveelheid personen berekenen die aspirine slikten en tijdens de studie hartfalen kregen zodanig dat het verschil tussen de groep die wel aspirine slikte en de groep die dat niet deed 'groot' te noemen is. Met behulp van deze kennis kan hij het farmaceutische bedrijf streng ondervragen.

b Stap nogmaals in de schoenen van de journalist en bereken dit minimumaantal.

Toepassen

Opgave 15: Intelligentie van honden

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.

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

Tabel 2.14

Bepaal op basis van deze tabel de mate van verschil in intelligentie tussen een labrador en een beagle.

Opgave 16: Geen verschil tussen beide groepen

Een onderzoek spitst zich toe op de vraag of er verschil in Y-waarde bestaat tussen de groepen met X_1 en met X_2.

Het resultaat halverwege het onderzoek naar eigenschappen X en Y in een populatie van 850 elementen zie je in de tabel.

Onderzoek hoe de kruistabel ingevuld moet zijn als er geen enkel verschil zit tussen beide groepen.

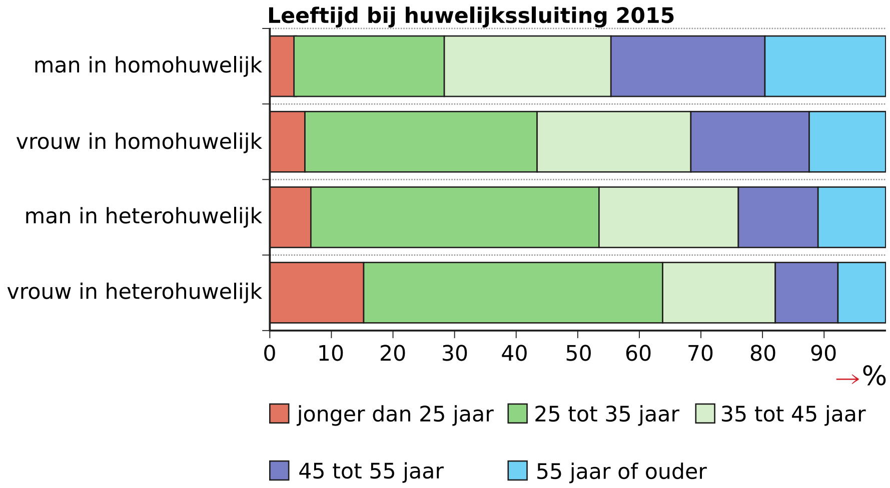
	X_1	X_2	
Y_1	...	...	102
Y_2	...	...	
		382,5	850

Tabel 2.15

Testen

Opgave 17

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 de conclusies van de soort “Er is een middelmatig verschil tussen het aantal ...” voor dezelfde leeftijdscategorieën.

Opgave 18

Dit is het resultaat van een statistisch onderzoek.

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

Tabel 2.16

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 daarbij het getal φ en de bijbehorende vuistregels.

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.

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 hoort een aantal manieren waarop je verschillen tussen statistische variabelen in kaart kunt brengen.

Figuur 3.1

Je leert in dit onderwerp

- twee populaties vergelijken met behulp van hun boxplots;
- twee gepaarde populaties vergelijken met behulp van de effectgrootte;
- twee normaal verdeelde populaties vergelijken met behulp van een verschiltoets op het gemiddelde.

Voorkennis

- soorten statistische variabelen herkennen;
- de begrippen onderzoek, steekproef, populatie en representatief, simulatie;
- meetniveau's te onderscheiden bij antwoordmogelijkheden op vragen;
- het begrip normale verdeling en de vuistregels;
- betrouwbaarheidsintervallen en foutenmarges bepalen bij het schatten van populatieproporties en populatiegemiddelden.

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

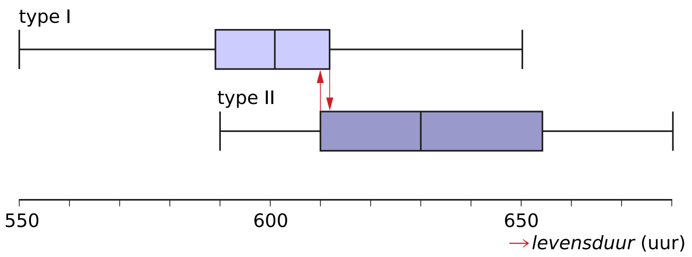
Probeer een manier te verzinnen om deze twee steekproeven te vergelijken.

Uitleg 1

Om een betrouwbare uitspraak te kunnen doen omtrent het verschil in levensduur van twee typen (twee populaties) batterijen, is van beide typen een steekproef van 15 stuks genomen.

Levensduur is een continue kwantitatieve variabele: je kunt statistische gegevens berekenen zoals steekproefgemiddelde, steekproefstandaardafwijking en de steekproefmediaan.

In dit geval is voor beide steekproeven de bijbehorende boxplot gemaakt.



Figuur 3.2

Door de boxplots boven elkaar boven dezelfde getallenlijn te presenteren, is het mogelijk een betrouwbare uitspraak over het verschil in levensduur van beide typen batterijen te doen.

Daarbij gebruik je de volgende vuistregels:

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

Bedenk: de ‘box’ is het interval vanaf het eerste kwartiel Q_1 tot en met het derde kwartiel Q_3.

Opgave 1

Bekijk **Uitleg 1**.
Ook zonder de afspraken waarmee je een maat kunt geven aan het verschil in levensduur tussen de type batterijen, kun je al met zekerheid het een en ander concluderen omtrent dat verschil.

a Noem minstens twee van dergelijke conclusies en geef aan hoe je tot die conclusies gekomen bent.

b Geef de conclusie over het verschil in levensduur van deze batterijen die je met behulp van de afspraken over verschillen in boxplots kunt trekken.

Opgave 2

Bekijk de tabel. De gegevens hebben betrekking op het aantal uur sport per week van 73 jongens en 102 meisjes.

mediaan jongens ≈ 4,0 uur	mediaan meisjes ≈ 2,0 uur
interkwartielafstand ≈ 4,2 uur	interkwartielafstand ≈ 4,0 uur
eerste kwartiel ≈ 1,8 uur	derde kwartiel ≈ 4,3 uur

Tabel 3.2

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

b 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 het **Formuleoverzicht**.

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.

Uitleg 3

Twee normaal verdeelde populaties kun je statistisch vergelijken met behulp van een hypothesetoets: de verschiltoets voor gemiddelden. Dat doe je door van beide populaties een steekproef te nemen en het verschil tussen beide steekproefgemiddelden als toetsvariabele te gebruiken.

Een fabrikant maakt oorbeschermers. Maar hij verkoopt ook tweedehands oorbeschermers. Hij wil de kwaliteit van beide populaties oorbeschermers vergelijken, want hij vermoedt dat de tweedehands versies van een lagere kwaliteit zijn. Hij neemt een steekproef van 100 nieuwe hoorbeschermers en 125 tweedehands oorbeschermers. Hij meet het aantal dB (decibel) dat de gehoorbeschermer dempt. De nieuwe hoorbeschermers dempen gemiddeld $\overline{X} = 30{,}1$ dB en de tweedehands oorbeschermers $\overline{Y} = 27{,}2$ dB.

De geluidsdemping van al deze oorbeschermers is normaal verdeeld met een standaardafwijking σ van 6 dB.

De te gebruiken toetsvariabele is het verschil $V = X - Y$.

Bij een normaal verdeelde verschiltoets voor gemiddelden geldt altijd als nulhypothese:

H$_0$: $\mu_V = 0$ (er is geen verschil tussen μ_X en μ_Y).

In dit geval geldt als alternatieve hypothese:

H$_1$: $\mu_V > 0$

De overschrijdingskans is:

$$\text{P}\left(\overline{V} > 2{,}9 | \mu_V = 0 \text{ en } \sigma_V = \sqrt{\frac{6^2}{100} + \frac{6^2}{125}}\right) \approx 0{,}0002$$

Conclusie:

Zelfs als hij een significantieniveau van 1% hanteert, wordt H$_0$ verworpen, want er is een significant verschil in gemiddelde demping tussen de twee populaties oorbeschermers.

Opgave 5

Bekijk **Uitleg 3**.

a Leg uit waarom de alternatieve hypothese in dit geval gelijk is aan $\mu_V > 0$.

b Leg uit waarom de overschrijdingskans gelijk is aan de normale kans $\text{P}\left(\overline{V} > 2{,}9\right)$.

c Bekijk de standaardafwijking die in de overschrijdingskans is gebruikt.

Laat zien uit welke onderdelen deze standaardafwijking is opgebouwd en leg uit waarom dat zo is.

d Laat zien hoe je de overschrijdingskans berekent als het verschil V gedefinieerd wordt als $Y - X$. Leidt dit tot dezelfde conclusie over het verschil in demping als in de uitleg?

Opgave 6

Een dameskapper houdt al jaren bij hoe lang het haar van zijn klanten met lang haar is. Daardoor weet hij dat deze haarlengte normaal verdeeld is met een standaardafwijking van $6{,}5$ cm.

Nu wil hij weten of er verschil is tussen de gemiddelde haarlengte van jonge vrouwen (tot 25 jaar) met lang haar en oudere vrouwen met lang haar. Als steekproef neemt hij de gegevens van 40 jonge en 40 oudere klanten uit de afgelopen 5 jaar.

De gemiddelde haarlengte van de groep jonge vrouwen met lang haar $\overline{J}$ is 56 cm.
De gemiddelde haarlengte van de groep oudere vrouwen met lang haar $\overline{O}$ is 54 cm.
De dameskapper kiest als significantieniveau 5%.

Is er significant verschil tussen de haarlengte van de jonge en de oudere vrouwen?

Theorie en voorbeelden

Om te onthouden

Twee populaties vergelijken op basis van een kwantitatieve variabele kan op meerdere manieren waarvan er hier drie benoemd worden.

Boxplots vergelijken
Dit ligt voor de hand als beide populaties klein zijn of als de steekproeven uit beide populaties klein zijn. De boxplots plaats je boven of naast één getallenlijn zodat de verschillen meteen te zien zijn.

Daarbij gebruik je de vuistregels op het **Formuleoverzicht**.

Effectgrootte
Als je van twee datasets de gemiddelden en de standaardafwijkingen weet, kun je de effectgrootte gebruiken:

$$E = \frac{\text{grootste gemiddelde} - \text{kleinste gemiddelde}}{\text{gemiddelde van de standaardafwijkingen}} = \frac{\overline{X}_1 - \overline{X}_2}{\frac{1}{2}(S_1 + S_2)}$$

Let erop dat X_1 groter moet zijn dan X_2

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

Verschiltoets voor gemiddelden uitvoeren

Deze toets kun je gebruiken bij twee normaal verdeelde populaties X en Y waarvan de populatiestandaardafwijkingen σ_X en σ_Y bekend zijn.

Van beide populaties trek je een grote steekproef. De beide steekproeflengtes n_X en n_Y mogen verschillend zijn. Het significantieniveau α van de toets moet vooraf vastgesteld zijn.

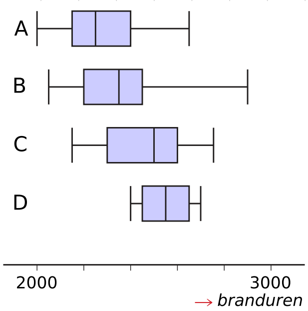
Als toetsvariabele gebruik je het verschil V, met $\mu_V = \mu_X - \mu_Y$ en $\sigma_V = \sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}$.

Je toetst altijd H_0: $\mu_V = 0$ ('er is geen verschil').

Voorbeeld 1

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

Vergelijk de brandduur van type A met de andere drie typen en doe een uitspraak over deze verschillen op basis van de vuistregels voor verschillen tussen boxplots.



Figuur 3.3

Antwoord

Voor het verschil in brandduur geldt (zie het **Formuleoverzicht**):

- 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 7

Gebruik de boxplots uit **Voorbeeld 1**. Trek alle overige mogelijke conclusies met de afspraken betreffende de boxplots.

Opgave 8

Gebruik de boxplots uit **Voorbeeld 1**.

a Naar een vijfde type lamp, type E, is ook onderzoek gedaan. In de steekproef zaten alleen lampen met een levensduur van meer dan 3000 uur.

Kun je met 100% zekerheid concluderen dat elke lamp van type E langer brandt dan de lampen van type A? Licht je antwoord toe.

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 het **Formuleoverzicht** gering.

Opgave 9

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 10

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?

Voorbeeld 3

Een datingapp, Vindn, wordt vaker bezocht door stadsbewoners dan door andere Nederlanders.

Het aantal minuten per week dat Vindn-bezoekers de app gebruiken is normaal verdeeld met een standaardafwijking van $4{,}5$ minuten.

De eigenaar van Vindn wil weten of er ook een significant verschil in bezoektijden zit tussen mensen uit de Randstad en mensen uit andere stedelijke gebieden. Ze laat een hypothesetoets, een verschiltoets voor gemiddelden, uitvoeren met een significantieniveau van 5%.

De onderzoekers trekken twee steekproeven:

- een steekproef van 55 Vindn-bezoekers uit de Randstad (R) heeft een steekproefgemiddelde van 1 uur en $24{,}5$ minuten per Vindn-bezoek per week
- een steekproef van 62 Vindn-bezoekers uit andere steden (A) heeft een steekproefgemiddelde van 1 uur en 26 minuten per Vindn-bezoek per week

Wat is de conclusie van de onderzoekers op basis van deze gegevens?

Antwoord

Als geldt $V = R - A$ dan is $\overline{V}$ het steekproefresultaat met $\overline{V} = 84{,}5 - 86 = \text{-}1{,}5$ minuten.

Verder geldt:

$\text{H}_0\text{: } \mu_V = 0$

$\text{H}_1\text{: } \mu_V \neq 0$

De bijbehorende overschrijdingskans is:

$$\text{P}\left(\overline{V} < \text{-}1{,}5 | \mu_V = 0 \text{ en } \sigma_V = \sqrt{\frac{4{,}5^2}{55} + \frac{4{,}5^2}{62}}\right) \approx 0{,}036$$

Dus H_0 wordt niet verworpen; er is geen sprake van een significant verschil tussen de Vindn-bezoekers uit de Randstad en de Vindn-bezoekers uit andere steden.

Opgave 11

Bekijk **Voorbeeld 3**.

a Leg uit waarom H_0 niet verworpen wordt ondanks het feit dat de overschrijdingskans kleiner is dan het genoemde significantieniveau.

b Beargumenteer waarom de overschrijdingskans exact gelijk is aan die uit het voorbeeld als de onderzoekers hadden gekozen voor $V = A - R$.

Opgave 12

Twee machines vullen pakken met suiker. Het gewicht van een pak suiker is normaal verdeeld met een standaardafwijking van $4{,}5$ gram.

De machines horen de pakken met een gelijke hoeveelheid suiker te vullen. Om te controleren of dit zo is, worden van beide machines steekproeven van 35 pakken genomen.

- Bij machine A is het gemiddelde gewicht van de pakken suiker in de steekproef $999{,}5$ gram.
- Bij machine B is het gemiddelde gewicht van de pakken suiker in de steekproef $1001{,}5$ gram.

Verschillen de gemiddelde vulgewichten van machine A en machine B van elkaar? Neem een significantie van 5%.

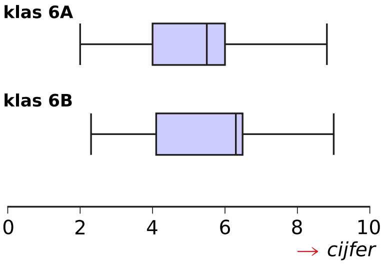
Verwerken

Opgave 13

Bekijk de twee boxplots van de cijfers voor een schoolexamen dat door twee zesde klassen gemaakt is.

a Welke mate van verschil bestaat er tussen de cijfers van klas 6A en die van klas 6B?

b Kun je deze twee groepen ook vergelijken met behulp van een verschiltoets voor gemiddelden als je daarvoor geschikte gegevens zou hebben?



Figuur 3.4

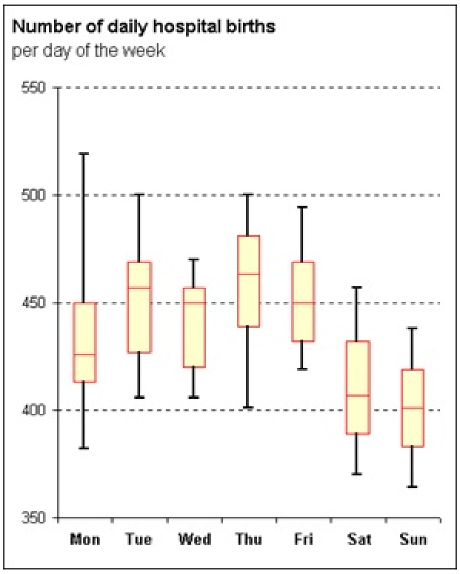
Opgave 14

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 3.5

Opgave 15

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

- Voor de referentiepopulatie geldt:
  Het gemiddelde van de variabele die hij onderzoekt is 220 gram, de standaardafwijking is 11 gram.
- Voor de onderzoekspopulatie geldt:
  Het gemiddelde van de variabele die hij onderzoekt is 210 gram, de standaardafwijking is ook 11 gram.

Welke conclusie kan hij trekken over het effect van de behandeling?

Opgave 16

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 17

Een pizzeria biedt pizza's aan met een gemiddelde diameter van 60 cm. Deze diameter is normaal verdeeld met een standaardafwijking van 4 cm.

Klanten van de pizzeria geven aan dat zij het idee hebben dat de diameter van de pizza's uit het zuidelijke filiaal groter is dan van de pizza's uit het noordelijke filiaal.

Uit een steekproef blijkt:

- de gemiddelde pizzadiameter $\overline{Z}$ van 75 pizza's uit het zuidelijke filiaal is $60{,}5$ cm;
- de gemiddelde pizzadiameter $\overline{N}$ van 60 pizza's uit het noordelijke filiaal is $59{,}2$ cm.

Hebben de klanten, met een significantieniveau van $2{,}5\%$, gelijk?

Opgave 18

Een zekere populatie bestaat uit mannen en vrouwen en men vraagt zich af of er een significant verschil voor normaal verdeelde variabele X is tussen mannen ($\sigma_m = 5$) en vrouwen ($\sigma_v = 3$). Van beide deelpopulaties is een steekproef getrokken:

	steekproef vrouwen	steekproef mannen
n	40	60
$\overline{X}$	830	829

Tabel 3.6

Er zijn meerdere mogelijkheden om dit verschil te onderzoeken en aan iedere mogelijkheid hangt een eigen kostenplaatje. Bovendien kan de ene methode een ander beeld geven over het verschil dan de andere methode.

Onderzoek het verschil met behulp van steekproefboxplots en met behulp van een hypothesetoets en vergelijk beide conclusies met elkaar.

Toepassen

Opgave 19: 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 3.7

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.

Opgave 20: Tennissers

Van professionele tennisspelers worden gemiddelde percentages bijgehouden van succesvolle servicebeurten. Van twee jonge tennissers die al op een hoog niveau spelen zijn deze gemiddelden ook bekend:

- K heeft een gemiddelde van 61% met een standaardafwijking van $3{,}5$%.
- R heeft een gemiddelde van $64{,}5$% met een standaardafwijking van $3{,}8$%.

De trainer wil weten of er een significant verschil is tussen het servicesucces van deze tennissers.

De genoemde percentages zijn op te vatten als populatiegemiddelden en -standaardafwijkingen: er zijn niet apart nog eens steekproeven getrokken. Dat betekent dat hier geen verschiltoets van gemiddelden op basis van steekproeven kan worden uitgevoerd.

a Neem aan dat beide standaardafwijkingen het resultaat zijn van een heel groot aantal metingen.
Bepaal met een hypothesetoets het verschil tussen tennisser K en tennisser R met een significantieniveau van 10%.

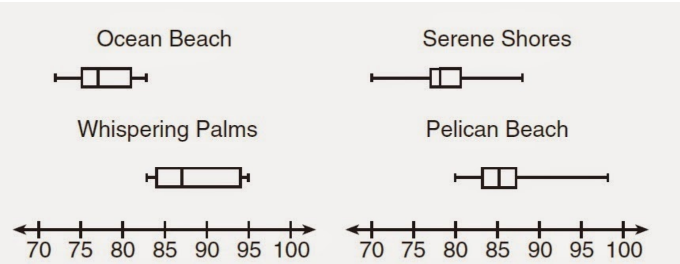
Stel dat de genoemde percentages afkomstig zijn van steekproeven van ieder 40 stuks.

b Toon met een berekening aan dat de conclusie dan anders zou zijn.

Testen

Opgave 21

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 3.6

Bron: http://www.mrburkemath.blogspot.nl/2015/05/january-2015-common-core-algebra.html

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 22

De diameters van machinaal geproduceerde bouten en de bijbehorende moeren zijn normaal verdeeld: de diameter van de moer is normaal verdeeld met een gemiddelde van $8{,}10$ mm en een standaarddeviatie van $0{,}05$ mm. De diameter van de bout is normaal verdeeld met een gemiddelde van $8{,}05$ mm en een standaardafwijking van $0{,}03$ mm. De bouten passen in de moeren als het verschil in diameter van de moer en de bout minder dan $0{,}02$ mm is. Er wordt regelmatig gecontroleerd of de machines die deze bouten en moeren maken niet moeten worden bijgesteld, omdat te veel moeren niet op de bouten passen. Wekelijks wordt een steekproef van 100 bouten en moeren getest.

a Waarom is hier sprake van een tweezijdige toets?

b Stel de nulhypothese en de alternatieve hypothese op.

c Welke standaardafwijking moet er worden gehanteerd? Waarom speelt nu ook de wortel-n-wet ($\sqrt{n}$-wet) een rol?

d Voer de toets uit met een significantieniveau van 5%. Bij welk gemiddelde verschil in de steekproef worden de machines bijgesteld?

Opgave 23

Een bepaald type ledlamp blijkt volgens een onderzoek uit 2015 door de fabrikant gemiddeld 34.300 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 35.200 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 Totaalbeeld

Samenvatten

Je hebt nu alle theorie van het onderwerp 'Verbanden en verschillen' doorgewerkt. Het is nu tijd om een overzicht over het geheel te krijgen.

Begrippenlijst

- verband tussen variabelen — correlatiecoëfficiënt — trendlijn/regressielijn
- verschil van kwalitatieve variabelen — cumulatief verschilpercentage — kruistabel en phi
- verschil van kwantitatieve variabelen — boxplots vergelijken — effectgrootte — verschiltoets van het gemiddelde

Activiteitenlijst

- statistisch verband tussen twee variabelen herkennen — verschil statistisch verband en oorzakelijk verband herkennen
- twee methoden om het verschil tussen kwalitatieve variabelen in een getal uit te drukken toepassen — vuistregels gebruiken
- boxplots vergelijken met behulp van vuistregels — een tekentoets uitvoeren — een verschiltoets van het gemiddelde uitvoeren

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 *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.
- 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**, 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 *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.
- Het *hypothese toetsen* wordt ook wel gebruikt om twee kwantitatieve variabelen te vergelijken (o.a. tekentoets en verschiltoets van het gemiddelde). De begrippen nulhypothese en significantieniveau kregen vooral betekenis door de Britse bioloog en statisticus **Ronald Aylmer Fisher (1890—1962)**. Hij schreef het invloedrijke boek 'Statistical methods for research workers' in 1925.

Figuur 4.1 **Karl Pearson**

Testen

Opgave 1

Van een groep bovenbouwleerlingen met wiskunde A/C in hun profiel is vastgelegd in hoeveel vreemde talen zij eindexamen doen. Dit is ook gebeurd voor een groep bovenbouwleerlingen met wiskunde B in hun profiel. Bekijk de tabel met de gegevens.

aantal vreemde talen	wiskunde A/C	wiskunde B
één	7	16
twee	26	58
drie	10	30
vier	0	7
	43	111

Tabel 4.1

a Leg uit waarom je het verschil in aantal gekozen vreemde talen tussen de groepen met wiskunde A/C en wiskunde B het best kunt onderzoeken met behulp van $\max V_{cp}$.

b Laat zien dat voor de gegevens in deze tabel geldt: $\max V_{cp} = 10\%$.

c Van welke mate van verschil is er sprake volgens de vuistregels voor $\max V_{cp}$?

(bron: cTwo)

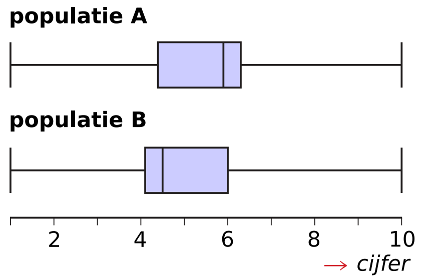
Opgave 2

Bekijk de boxplots van de gegevens van twee populaties.

a Om welk type gegevens gaat het hier?

b Geef van elk van de volgende beweringen aan of de bewering, de conclusie of beide wel of niet juist zijn en beargumenteer waarom dat volgens jou zo is.

1. De boxplots zijn even breed dus de beide populaties zijn gelijk.
2. De boxen van beide boxplots zijn even breed dus beide populaties zijn gelijk.
3. De boxen van beide boxplots verschillen van opbouw: beide populaties verschillen dus veel van elkaar.



Figuur 4.2

Opgave 3

In 1947 hielden de wiskundigen Freudenthal en Sittig een statistisch onderzoek ten behoeve van een nieuw maatsysteem voor vrouwenkleding in opdracht van het warenhuis De Bijenkorf. Zij lieten daarbij een grote verscheidenheid aan lichaamsmaten opmeten van 5001 vrouwen. Zij vonden onder andere een sterke correlatie tussen de taille (de omtrek van het lichaam gemeten ter hoogte van de navel) en de bovenwijdte (de omtrek van het lichaam gemeten ter hoogte van de borst). De gevonden correlatiecoëfficiënt bedroeg ongeveer $0{,}9058$.

In hun rapport ‘De juiste maat’ geven zij voor de bovenwijdte b een gemiddelde van $97{,}99$ cm met een standaardafwijking van $10{,}12$ cm en voor de taille t een gemiddelde van $80{,}45$ cm met een standaardafwijking van $10{,}80$ cm.

a Welke betekenis heeft deze hoge correlatie tussen b en t voor een spreidingsdiagram van deze twee variabelen?

b Stel een vergelijking op van de regressielijn voor b.

c Geef een statistisch verantwoorde schatting van de bovenwijdte van een vrouw met een taille van 90 cm.

Opgave 4

Bekijk de tabel met de (geschatte) woordenschat die eentalige Nederlandse kinderen hebben.

Wat mag je, op basis van de gegevens, uit deze tabel concluderen omtrent het verband tussen de leeftijd en de woordenschat van een eentalig Nederlands kind?

(bron: www.cbsdeakker.nl)

leeftijd	woordenschat
4	3000
5	3800
6	4500
7	5200
8	6000
9	8500
10	11000
11	14000
12	17000

Tabel 4.2

Opgave 5

Cito meet kennis en vaardigheden in door middel van 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 4.3

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 6

Patiënten die voor een behandeling enige tijd in een ziekenhuis worden opgenomen, lopen tijdens dit verblijf het risico een infectie te krijgen. Zo'n infectie wordt een zorginfectie genoemd. Een deel van de zorginfecties ontstaat na een operatie.

In de periode 2007 tot en met 2012 is een steekproef gehouden onder een deel van de Nederlandse ziekenhuizen. Bekijk de tabel met enkele resultaten hiervan.

	aantal
patiënten	95299
patiënten die een zorginfectie hebben opgelopen	4694
geopereerde patiënten	32664
geopereerde patiënten die een zorginfectie hebben opgelopen	1286

Tabel 4.3

Je mag aannemen dat de patiënten in deze ziekenhuizen representatief zijn voor alle patiënten die in een Nederlands ziekenhuis worden opgenomen.

Onderzoek in hoeverre er verschil is in het krijgen van een zorginfectie tussen geopereerde en niet-geopereerde patiënten.

(bron: pilotexamen 2016 - II havo A)

Opgave 7

Bekijk de resultaten van 19 leerlingen voor het SE en het CE.

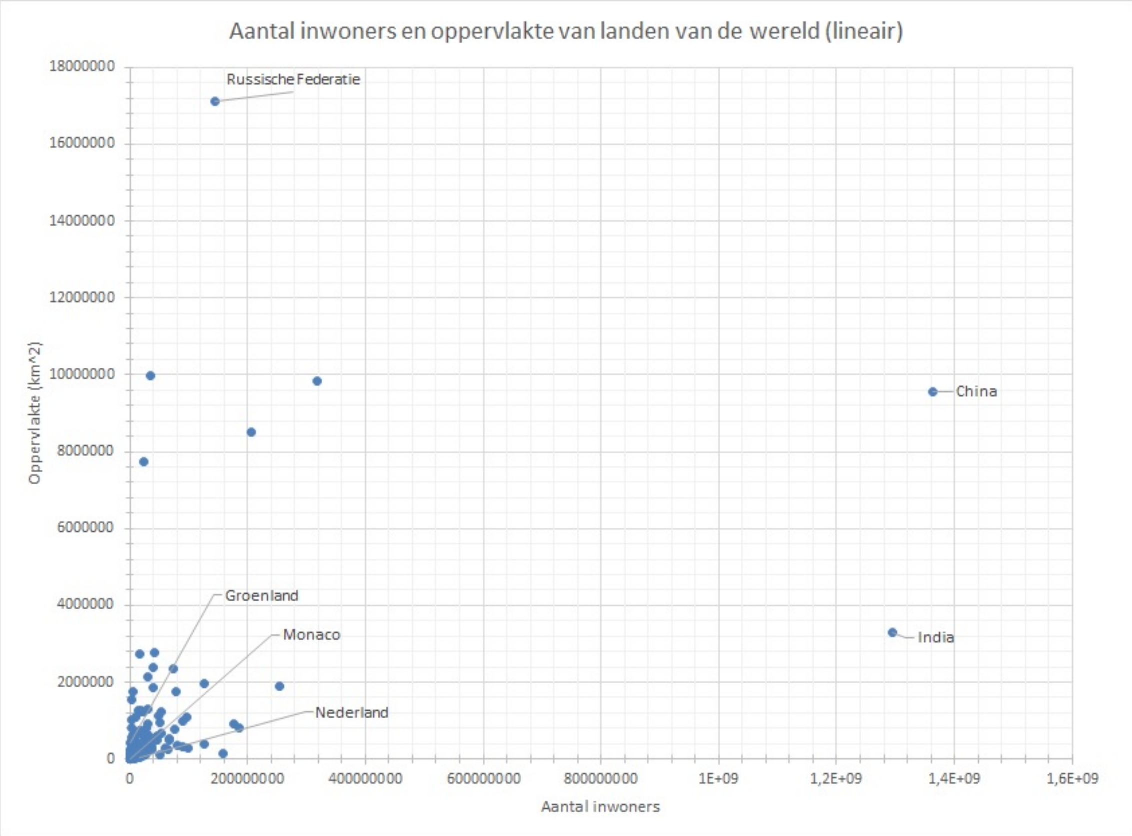
leerling	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
SE-cijfer	6,0	6,7	5,8	7,1	5,4	6,5	8,8	6,9	7,9	5,1	6,1	6,1	6,4	7,4	5,9	6,2	7,1	6,8	6,3
CE-cijfer	6,4	6,3	5,2	6,5	5,4	6,1	9,0	6,8	7,5	5,6	6,0	6,5	6,0	6,5	6,0	6,6	7,0	6,6	6,4

Tabel 4.4

Neem aan dat de gemiddelden van de steekproeven normaal verdeeld zijn. Ga na of uit een verschiltoets met een significantieniveau van 5% volgt dat het SE beter is gemaakt dan het CE.

Toepassen

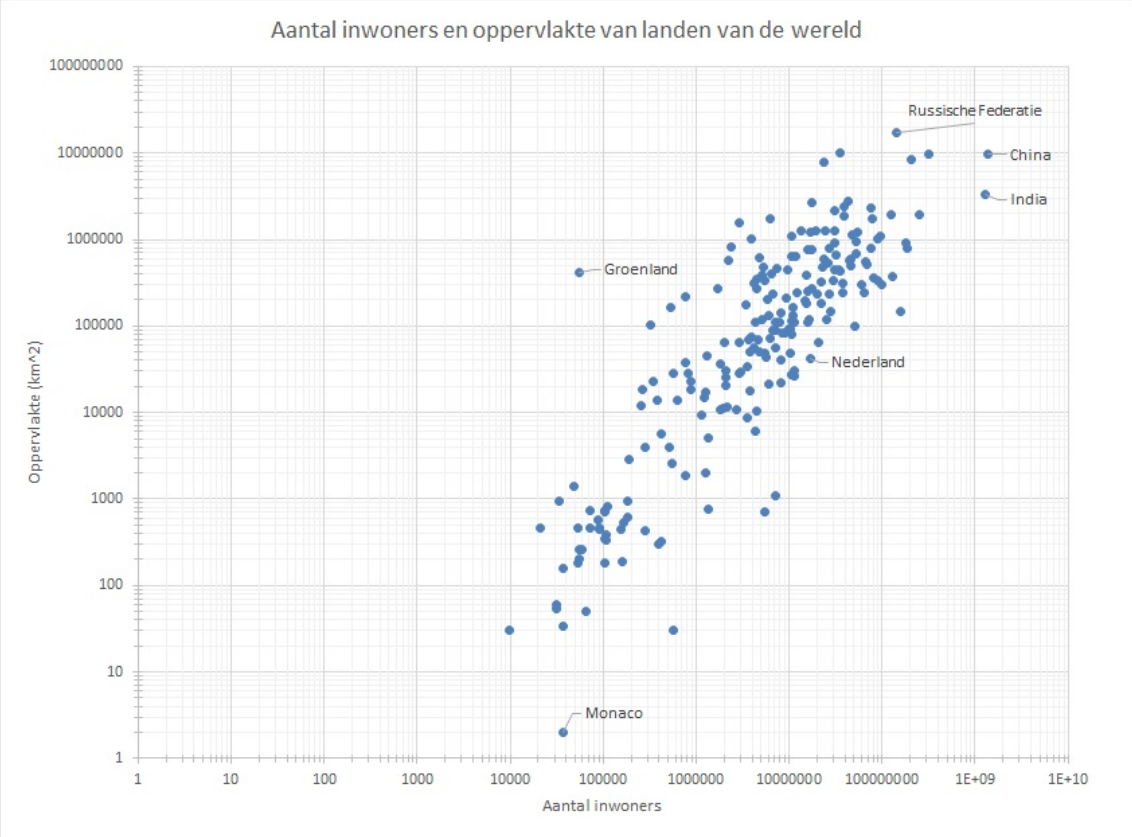
Opgave 8: Trendlijn op logaritmisch papier



Figuur 4.4

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 4.5

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.

2.5 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 5.1

Je leert in dit onderwerp

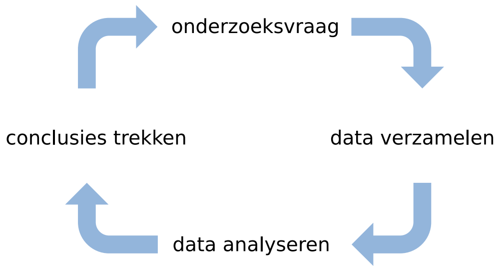
- een statistisch onderzoek opzetten.

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 5.2

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.

- UStat 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

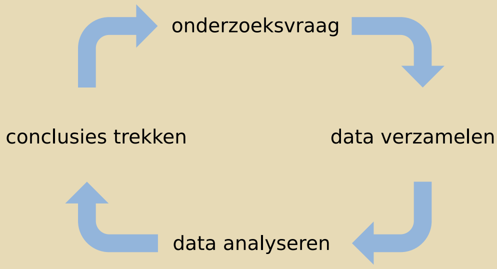
Om te onthouden

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.

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**.



Figuur 5.3

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?”

Lees **hier** het artikel.

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 n.a.v. 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?”

Lees **hier** het artikel.

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 Wat is de conclusie van 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 moeten worden om het verschil tussen topkeepers te kunnen bepalen?

f Wat is er lastig aan een onderzoek naar de statistische variabelen die bij 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. Deze opgaven kun je als opstapje/voorbeeldvragen gebruiken voor een eigen onderzoek. Er staan geen antwoorden bij. Maar waarschijnlijk vind je een ander onderwerp veel geschikter voor een onderzoek. Dan moet je wel zelf op zoek gaan naar geschikte data!

Figuur 5.4

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 m.b.v. 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 .

Stichting Math4All

Inhoud Katern 1

www.math4all.nl