

2.4 Normale verdeling

Inleiding

Je hebt nu naar frequentieverdelingen leren kijken en centrummaten en spreidingsmaten leren gebruiken. Maar er is nog een spreidingsmaat die vooral belangrijk is bij een veel voorkomende frequentieverdeling: de normale verdeling. Deze verdeling heeft de typerende klokvorm.



Figuur 1

Je leert in dit onderwerp

- de standaardafwijking te berekenen en te interpreteren;
- de normale verdeling te herkennen en te typeren.

Voorkennis

- de begrippen data, populatie, steekproef, aselekt en representatief, kwantitatief en kwalitatief, absolute en relatieve frequentie, discrete en continue variabele, klassenbreedte, klassenmiddelen en klassengrenzen;
- de verschillende centrummaten en spreidingsmaten berekenen;
- allerlei diagrammen interpreteren.

Verkennen

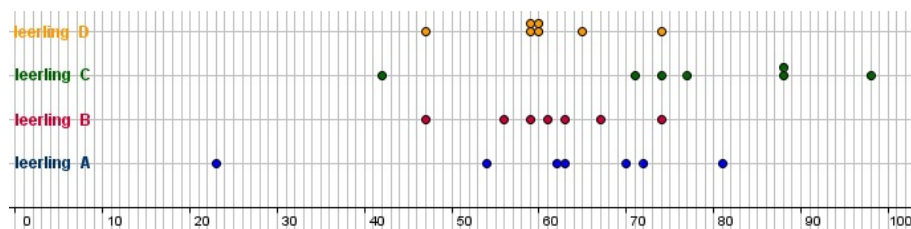
Opgave V1

Je ziet de SE-cijfers (schoolexamencijfers) van enkele leerlingen aan het eind van havo 5. Het eindcijfer is het gemiddelde van deze cijfers.

	se1	se2	se3	se4	se5	se6	se7
A	7,2	6,3	7,0	2,3	6,2	8,1	5,4
B	5,9	7,4	5,6	6,7	6,1	6,3	4,7
C	8,8	9,8	7,4	8,8	5,7	7,3	4,2
D	7,4	6,0	6,0	6,5	5,9	5,9	4,7

Tabel 1

- Welke eindcijfers krijgen deze leerlingen?
- In de figuur is voor iedere leerling elk SE-cijfer aangegeven door een bolletje op een getallenlijn (de komma in het cijfer is weggelaten). Geef in deze figuur per leerling de gemiddelde SE-cijfers aan.



Figuur 2

- De leerlingen A en B hebben hetzelfde gemiddelde. Toch is hun cijferbeeld nogal verschillend. In welk opzicht?
- De spreiding van de cijfers van A en C is vrijwel hetzelfde. Waarin verschilt hun cijferbeeld vooral?
- Bepaal bij elke leerling de spreidingsbreedte van hun cijfers.

- f De leerlingen B en D hebben vrijwel dezelfde spreidingsbreedte. Zou je de spreiding van hun cijfers ook hetzelfde willen noemen?

Opgave V2

Een andere maat voor de spreiding kun je vinden door te kijken hoe ver elk cijfer van het gemiddelde af ligt. Je doet dit door van elk cijfer het verschil met het gemiddelde te berekenen. Bekijk die verschillen voor leerling A met het gemiddelde 6,1.

	se1	se2	se3	se4	se5	se6	se7	
A	7,2	6,3	7,0	2,3	6,2	8,1	5,4	cijfer
	1,1	0,2	0,9	-3,8	0,1	2,0	-0,7	verschil

Figuur 4

leerling A		
A	A - gem	(A - gem) ²
7,2	1,1	1,27
6,3	0,2	0,05
7,0	0,9	0,86
2,3	-3,8	14,22
6,2	0,1	0,02
8,1	2,0	4,12
5,4	-0,7	0,45
	0,0	20,99

gem = 6,1
var = 3,00
sd = 1,73

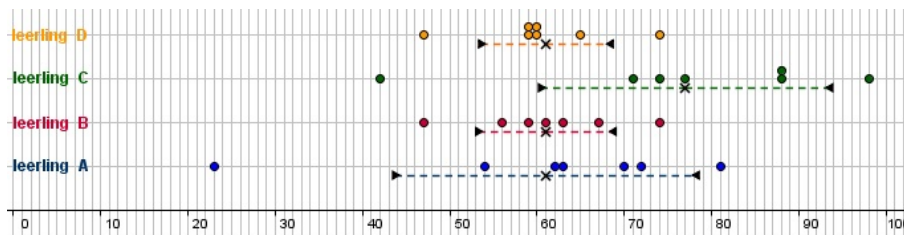
Figuur 3

- a Bereken het gemiddelde van deze verschillen. Wat valt je daarbij op?
- b Het gemiddelde van deze verschillen is geen goede spreidingsmaat. Dat komt doordat de negatieve afwijkingen wegvallen tegen de positieve. Door te kwadrateren verdwijnen die mintekens. Bekijk hoe je dat bij leerling A aan kunt pakken.
- Je ziet dat eerst een kolom met de verschillen met het gemiddelde wordt gemaakt. Daarna een kolom met de kwadraten van deze verschillen. De variantie var is het gemiddelde van de kwadraten van de verschillen. Reken dat getal na.
- c Welke verschillen leveren een grote bijdrage aan de variantie en welke niet?
- De wortel uit de variantie van A heet de 'standaardafwijking' van de set cijfers van A; hier afgekort met sd.
- d Maak voor leerling B ook zo'n tabel en bereken op dezelfde manier de standaardafwijking.
- e Bereken de standaardafwijking van de cijfers van de leerlingen C en D.

Uitleg 1

Je hebt naar frequentieverdelingen leren kijken en centrummaten en spreidingsmaten leren gebruiken. Maar er is nog een belangrijke spreidingsmaat, namelijk de standaardafwijking. De standaardafwijking houdt rekening met de afwijking vanaf het gemiddelde van elke waarde. Het is een soort gemiddelde afwijking ten opzichte van het gemiddelde.

Je ziet in één figuur vier sets met cijfers van vier leerlingen. De gemiddelden zijn aangegeven met x en je ziet de standaardafwijkingen links en rechts van het gemiddelde.



Figuur 5

Van leerling A zijn de cijfers (zie figuur): 2,3, 5,4, 6,2, 6,3, 7,0, 7,2 en 8,1.

cijfer	cijfer – gemiddelde	(cijfer – gemiddelde) ²
2,3	-3,8	14,44
5,4	-0,7	0,49
6,2	0,1	0,01
6,3	0,2	0,04
7,0	0,9	0,81
7,2	1,1	1,21
8,1	2,0	4,00
som van de kwadraten		21,00

Tabel 2

Je berekent zo'n standaardafwijking van bijvoorbeeld leerling A als volgt:

- Je bepaalt eerst het gemiddelde van zijn cijfers, dat is 6,1.
- Dan bepaal je van elk cijfer het verschil met het gemiddelde, de zogenoemde 'deviatie'.
- Elke deviatie kwadrateer je (om geen negatieve afwijkingen te krijgen).
- Van die kwadraten van de deviaties bereken je het gemiddelde. Daarmee heb je de variantie van de verzameling cijfers gevonden: $\text{variantie} = \frac{\text{som van de kwadraten van de deviaties}}{\text{totale aantal}}$
- De standaardafwijking is de wortel uit de variantie.

De variantie is dus $\frac{21,00}{7} \approx 3,5$ en de standaardafwijking is $\sqrt{\frac{21,00}{7}} \approx 1,87$.

Je kunt de standaardafwijking ook met de grafische rekenmachine berekenen, zie het **Practicum**.

De standaardafwijking is de meestgebruikte spreidingsmaat bij statistisch onderzoek. Deze afwijking van het gemiddelde wordt links en rechts vanaf het gemiddelde uitgezet.

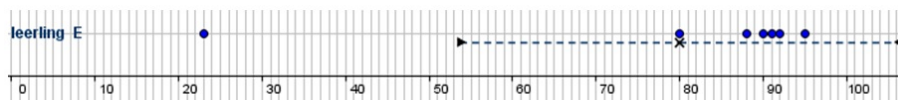
Opgave 1

In **Uitleg 1** zie je hoe de standaardafwijking wordt berekend van de cijfers van leerling A.

Doe dit zelf van leerling B.

Opgave 2

Lees **Uitleg 1**. Bekijk de gegevens van een vijfde leerling E.



Figuur 6

- In de verdeling van de cijfers van leerling E zie je een uitschieter: een 2,3. Daarom is de getekende standaardafwijking ongeschikt om de spreiding van de verdeling te beschrijven. Licht dit toe.
- Stel dat leerling E de 2,3 mag herkansen en dat hij dan een 7,9 haalt. Welke richting zal het gemiddelde dan opgaan en wat betekent dit voor de standaardafwijking, denk je? Geeft de standaardafwijking de spreiding nu beter weer?

Opgave 3

Stel je voor dat de laatste drie behaalde cijfers twee keer zo zwaar meetellen als de eerste vier cijfers. Bij leerling E zijn die laatste drie cijfers 9,0 en 8,8 en 2,3. Je noemt het eindcijfer in zo'n geval een 'gewogen gemiddelde'.

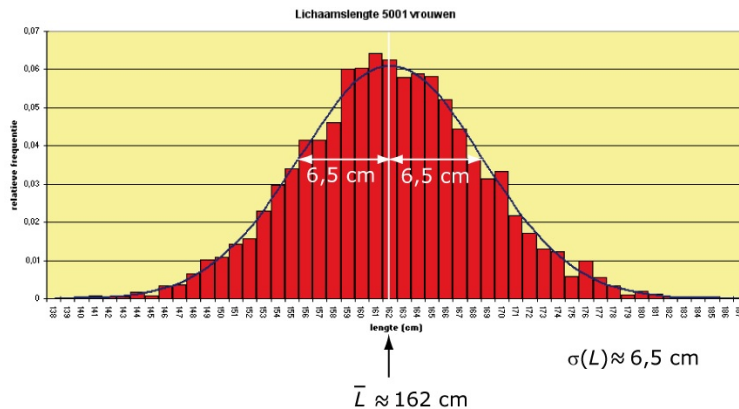
- Waarom moet je de laatste cijfers nu twee keer in de berekeningen van het gemiddelde en de standaardafwijking opnemen?
- Bereken het gewogen gemiddelde van deze leerling E met de bijbehorende standaardafwijking. Maak een tabel met cijfer, frequentie, frequentie min gemiddelde en $(\text{frequentie} - \text{gemiddelde})^2$. Bereken vervolgens de variantie en de standaardafwijking.

Opgave 4

Op welke van de drie spreidingsmaten heeft een uitschieter geen invloed?

Uitleg 2

Je ziet een staafdiagram van de lengtes van 5001 vrouwen uit de dataset **Statistiek Bijenkorf 1947**. Die verdeling heeft een vrijwel zuivere 'klokvorm'. Je noemt dit een normale verdeling, elke normale verdeling heeft zo'n klokvorm. De gemiddelde lengte $\bar{L} \approx 162$ cm en de standaardafwijking $\sigma(L) \approx 6,5$ cm zijn in de figuur aangegeven.



Figuur 7

Beide getallen leggen de normale verdeling volledig vast, ze zijn karakteristiek voor deze normale verdeling. De standaardafwijking geef je aan met een Griekse letter s, de 'sigma' in $\sigma(L)$. Voor het gemiddelde gebruik je ook wel een Griekse letter m, de 'mu' in $\mu(L)$. Op grond van alleen het gemiddelde en de standaardafwijking kan elke normale verdeling worden getekend.

Opgave 5

Bekijk de dataset **Gegevens 154 havo 4-leerlingen**. Je vindt daar de lengtes van 154 leerlingen. Bekijk de lengtes van jongens en meisjes afzonderlijk.

- a Laat met behulp van de grafische rekenmachine of met Excel zien dat de jongens een gemiddelde lengte van 180 cm hebben met een standaardafwijking van 8 cm en dat de meisjes een gemiddelde lengte van 169 cm hebben met een standaardafwijking van 7 cm.

De lengtes van de jongens en de meisjes van deze groep leerlingen vormen geen normale verdeling omdat de groep daarvoor te klein is. Maar stel je voor dat de gevonden gemiddelden en de gevonden standaardafwijkingen wel bij een normale verdeling zouden horen.

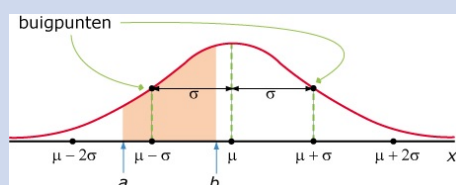
- b Maak dan een schets van deze normale verdelingen. Let vooral goed op de correcte schaalverdeling op de lengte-as.

Theorie en voorbeelden

Om te onthouden

Symmetrische verdelingen met de vorm van een klok laten zich goed beschrijven door de centrummaat het **gemiddelde** en de spreidingsmaat de **standaardafwijking** of **standaarddeviatie**.

Die symmetrische verdeling noem je **normale verdeling**.



Figuur 8

In deze normale verdeling zie je het gemiddelde μ en de standaardafwijking σ aangegeven.

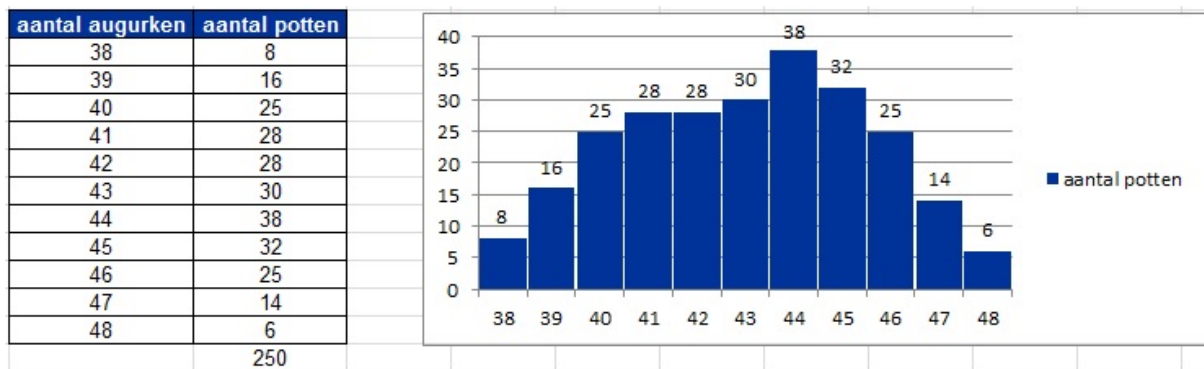
Je berekent de standaardafwijking σ zo:

- Je bepaalt het gemiddelde.
- Dan bepaal je van elke waarde van de variabele het verschil met het gemiddelde, dat is de **deviatie** van die waarde.
- Al die deviaties kwadrateer je om negatieve afwijkingen te voorkomen.
- Van die kwadraten van de deviaties bereken je het gemiddelde. Het getal dat je dan krijgt, heet de **variantie** van de verzameling cijfers.
- De standaardafwijking is de wortel uit de variantie.

Elke normale verdeling wordt bepaald door twee getallen: het gemiddelde μ en de standaardafwijking σ . De buigpunten van de kromme zitten precies één standaardafwijking ($1 \cdot \sigma$) van de symmetrieas af.

Voorbeeld 1

Van 250 potten met augurken (uitlekgewicht 370 gram) is geteld hoeveel augurken de potten bevatten. Van het resultaat zie je een tabel en een staafdiagram.



Figuur 9

Bereken het gemiddelde en de standaardafwijking van het aantal augurken per pot.

Antwoord

In de praktijk bereken je dit met Excel of met de grafische rekenmachine. Ook de grafische rekenmachine kan met zo'n frequentietabel het gemiddelde en de standaardafwijking berekenen. Deze uitwerking is bedoeld om je nog een keer te laten zien welke stappen je handmatig uitvoert om de standaardafwijking σ te krijgen.

Bekijk de tabel. Let daarbij goed op het meewegen van de frequenties.

aantal augurken	aantal potten			
A	f	A * f	(A - μ) ²	(A - μ) ² * f
38	8	304	25	200
39	16	624	16	256
40	25	1000	9	225
41	28	1148	4	112
42	28	1176	1	28
43	30	1290	0	0
44	38	1672	1	38
45	32	1440	4	128
46	25	1150	9	225
47	14	658	16	224
48	6	288	25	150
	250	10750		1586
	$\mu = 10750 / 250 =$	43		
	var = $1586 / 250 =$	6,344		
	$\sigma =$	2,52		

Tabel 3

Opgave 6

Bekijk [Voorbeeld 1](#).

- a Laat met een berekening zien hoe het gemiddelde μ wordt berekend.
- b Bij het berekenen van de variantie wordt van elk aantal augurken het kwadraat van het verschil met het gemiddelde (de deviatie) berekend. Laat zien dat de waarde 39 inderdaad 16 wordt.
- c Waarom is de laatste kolom in de tabel $(A - \mu)^2 \cdot f$ nodig?
- d Hoe wordt de standaardafwijking σ berekend? Controleer die berekening.

Opgave 7

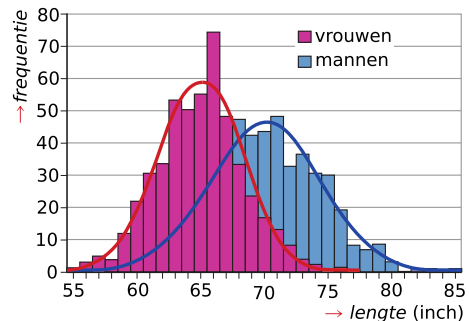
In [Uitleg 2](#) kwam je de verdeling tegen van de lengtes van de vrouwen uit het onderzoek in 1947 van Freudenthal en Sittig in opdracht van De Bijenkorf. Bekijk de dataset [Statistiek Bijenkorf 1947](#) en zoek daarin de lengtes van deze vrouwen op.

Bereken μ en σ met behulp van Excel.

Voorbeeld 2

Je ziet een verdeling van de lengtes van een groep mannen en ook een verdeling van de lengtes van een groep vrouwen. De lengtes zijn gegeven in inches.

Schat zowel van de mannen als van de vrouwen de gemiddelde lengte en de bijbehorende standaardafwijking. Geef je antwoorden in gehele inches.



Figuur 10

Antwoord

Bekijk de twee geschetste normale verdelingen en schat hun buigpunten.

Voor de mannen geldt:

- De gemiddelde lengte is ongeveer 70 inches.
- De bijbehorende standaardafwijking is ongeveer $70 - 65 = 5$ inches.

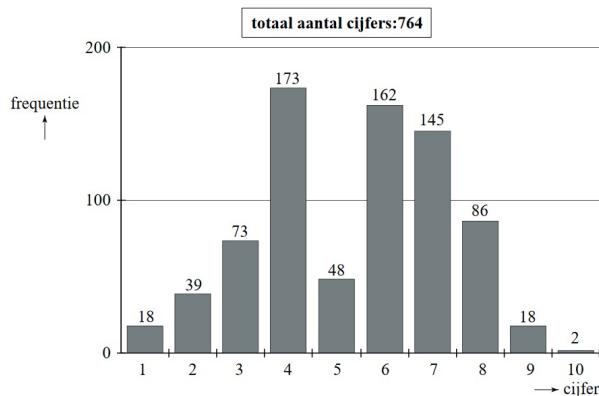
Voor de vrouwen geldt:

- De gemiddelde lengte is ongeveer 65 inches.
- De bijbehorende standaardafwijking is ongeveer $65 - 61 = 4$ inches.

Opgave 8

In januari 2008 verscheen er in de NRC een artikel over de becijfering van een tentamen Recht. In de figuur zie je de verdeling van de cijfers voor dat tentamen.

De schrijvers van het artikel waren erg kritisch.



Figuur 11

- a Wat is opvallend aan deze verdeling van cijfers?
- b Het gemiddelde is 5,37. Bereken de standaardafwijking in twee decimalen nauwkeurig.
- c De schrijvers van het artikel verwachtten een normale verdeling. Schets de verdeling die zij voor ogen hadden, met gemiddelde 5,4 en standaardafwijking 1,9.

(bron: examen vwo wiskunde A in 2012, eerste tijdvak)

Verwerken

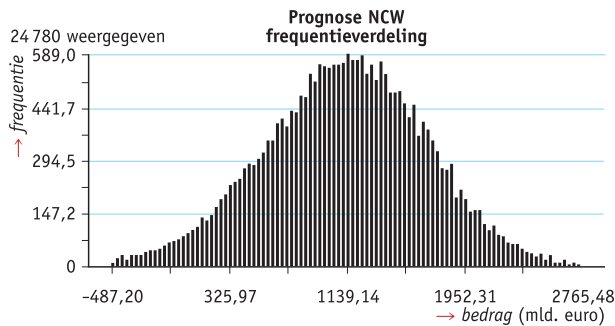
Opgave 9

In een grote dataset worden de twee grootste uitschieters naar boven weggestreept.

- a Bereken wat de gevolgen zijn voor de drie centrummaten.
- b Bereken wat de gevolgen zijn voor de drie spreidingsmaten.

Opgave 10

In een kamerstuk van de Tweede Kamer staat de volgende frequentieverdeling. Deze verdeling benadert de normale verdeling. Op de horizontale as staat het bedrag in miljarden euro's. Maak uit de grafiek een schatting van het gemiddelde en de standaarddeviatie.



Figuur 12

Opgave 11

De cijfers bij een toets wiskunde van dertig leerlingen worden voor alle leerlingen met een 0,5 verhoogd. De cijfers zijn (in de oude situatie) normaal verdeeld met gemiddelde 5,3 en standaardafwijking 1,2.

Bereken het gemiddelde en de standaardafwijking van de cijfers wiskunde na de verhoging.

Opgave 12

Voor deze opgave kun je het beste met Excel rekenen, maar met de grafische rekenmachine lukt het ook.

Je ziet de tabel met mouwlengtes van de 5001 vrouwen uit de dataset [Statistiek Bijenkorf 1947](#).

mouw- lengte	frequentie
49	3
50	11
51	22
52	53
53	89
54	163
55	250
56	405
57	519
58	660
59	578
60	653
61	560
62	421
63	260
64	159
65	106
66	52
67	18
68	15
69	3
70	0
71	1

Tabel 4

- Bereken de gemiddelde mouwlengte μ en de standaarddeviatie σ .
- Hoeveel procent van de mouwlengtes zit tussen $\mu - \sigma$ en $\mu + \sigma$?
- Hoeveel procent van de mouwlengtes zit tussen $\mu - 2\sigma$ en $\mu + 2\sigma$?

Opgave 13

Bekijk de dataset met de gegevens over de [Sportprestaties van 74 brugklassers](#). De gesprongen afstand is een variabele D . Werk met Excel.

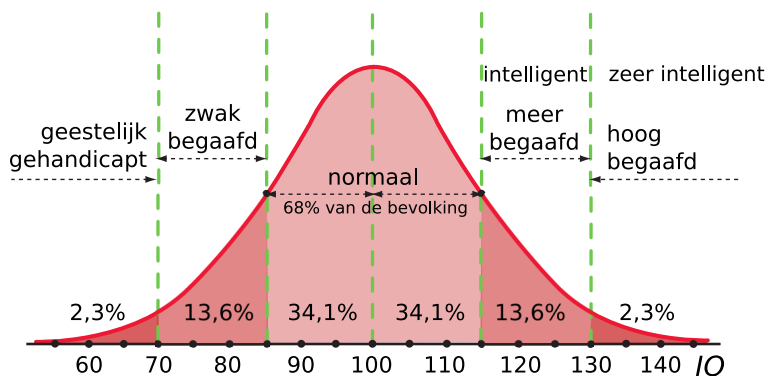
- Bereken het gemiddelde en de standaarddeviatie van D .
- Hoeveel procent van de prestaties zit onder het gemiddelde?
- Hoeveel procent van de resultaten zit tussen $\mu - \sigma$ en $\mu + \sigma$?

Toepassen

Opgave 14: Intelligentiequotiënt

In het begin van de vorige eeuw werd er veel waarde gehecht aan het zogenoemde ‘intelligentiequotiënt’ van met name kinderen. In 1904 werd de Franse psycholoog Alfred Binet door de Franse overheid gevraagd een test te ontwerpen om ‘slimme’ van ‘domme’ kinderen te onderscheiden.

Binet ontwierp een intelligentietest waarmee hij de intelligentieleeftijd van een kind vaststelde. Wanneer de intelligentieleeftijd wordt gedeeld door de werkelijke leeftijd, en met 100 vermenigvuldigd, dan krijg je het IQ. Dit IQ is normaal verdeeld met een gemiddelde dat op 100 is gesteld. Zo is de Stanford-Binet Intelligentieschaal ontwikkeld.



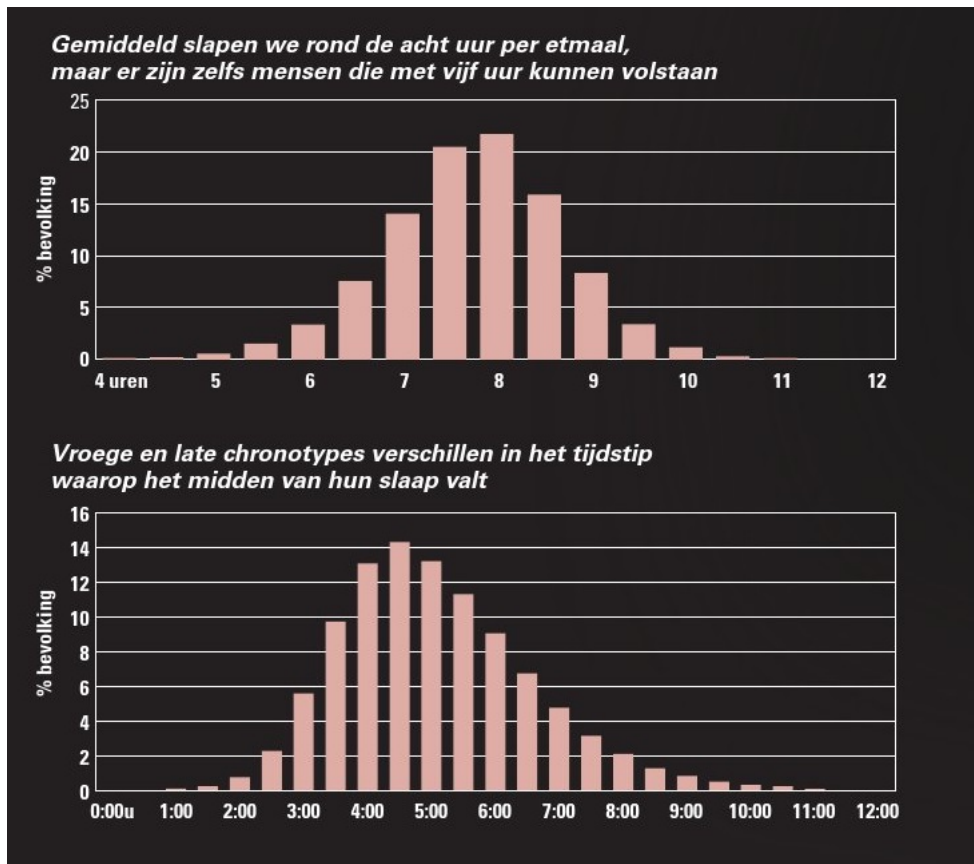
Figuur 13

- Welk gemiddelde en welke standaardafwijking heeft het IQ volgens de Stanford-Binet Intelligentieschaal?
- Hoeveel procent van de mensen heeft een minder dan normale intelligentie?
- Hoeveel procent van de mensen heeft een IQ van boven de 145?
- Welk IQ heeft de intelligentste 10% van de bevolking?

Testen

Opgave 15

Hier zie je twee staafdiagrammen die informatie geven over het aantal uren slaap en het tijdstip van in slaap vallen van de Nederlanders.



Figuur 14

- Welke van deze twee frequentieverdelingen kun je het beste benaderen met een normale verdeling?
- Schat gemiddelde μ en standaardafwijking σ van dit diagram.
- Hoeveel procent van de bevolking kan volstaan met minder slaap dan $\mu - \sigma$?
- Hoeveel procent van de bevolking moet meer dan $\mu + \sigma$ slapen?

(bron: slaapmlb.blogspot.nl)

Practicum

Als je met niet al te grote datasets te maken hebt, kun je met de **grafische rekenmachine** de gegevens verwerken, analyseren en presenteren. Hoe dat gaat zie je in de practica:

- [Statistiek en de TI84](#)
- [Statistiek en de TIinspire](#)
- [Statistiek en de Casio fx-CG50](#)
- [Statistiek en de HPprime](#)
- [Statistiek en de NumWorks](#)

Met **Excel** (een spreadsheetprogramma, een rekenblad) werken is bij statistiek eigenlijk onontbeerlijk. Je kunt er grote hoeveelheden gegevens in kwijt. Die gegevens kun je ordenen en presenteren. Bekijk de eerste vier delen van het practicum:

- **Data presenteren**


Je kunt ook data analyseren en presenteren met de app 'Data analyse' van **VUstat**. Daarin kun je eigen databestanden vanuit Google-Drive toevoegen, maar er zijn ook diverse datasets beschikbaar. Ga hiervoor naar:

- **Data analyse VUstat**



© 2024

Deze paragraaf is een onderdeel van het Math4All wiskundemateriaal.

Math4All stelt het op prijs als onvolkomenheden in het materiaal worden gemeld en ideeën voor verbeteringen in de content of dienstverlening kenbaar worden gemaakt. Klik op  in de marge bij de betreffende opgave. Uw mailprogramma wordt dan geopend waarbij het emailadres en onderwerp al zijn ingevuld. U hoeft alleen uw opmerkingen nog maar in te voeren.

Email: f.spijkers@math4all.nl

Met de Math4All Foliostaat kunnen complete readers worden samengesteld en toetsen worden gegenereerd. Docenten kunnen bij a.f.otten@math4all.nl een gratis inlog voor de maatwerkdienst aanvragen.
