Hoofdstuk 2

De normale verdeling

Opdracht 1 bladzijde 44

1 Gegeven de volgende vijf meetwaarden: 7, 14, 17, 19 en 23. Hun gemiddelde $\overline{x}$ is 16. Het verschil tussen een waarde en het gemiddelde noemen we de **afwijking** van die waarde. Zo is $7 - 16 = -9$ de afwijking van 7. Dit betekent: 7 ligt 9 eenheden onder het gemiddelde.

Bereken het gemiddelde van de vijf afwijkingen. We noemen dat de **gemiddelde afwijking** van deze vijf waarden.

$$\frac{-9+(-2)+1+3+7}{5}=0$$

2 We veralgemenen het voorbeeld hierboven. Beschouw de n waarden $x_1, x_2, \ldots, x_n$ met gemiddelde $\overline{x}$.

Bereken $\dfrac{(x_1-\overline{x})+(x_2-\overline{x})+\ldots+(x_n-\overline{x})}{n}$.

Er zijn verschillende manieren om dit gemiddelde te berekenen. Hieronder een van de mogelijkheden:

$$\frac{(x_1-\overline{x})+(x_2-\overline{x})+\ldots+(x_n-\overline{x})}{n}$$

$$=\frac{x_1+x_2+\ldots+x_n-n\cdot\overline{x}}{n}$$

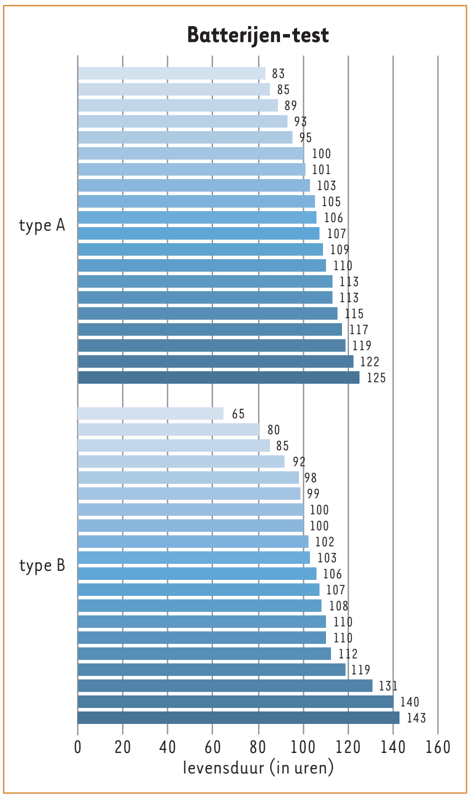
$$=\frac{x_1+x_2+\ldots+x_n}{n}-\frac{\cancel{n}\cdot\overline{x}}{\cancel{n}}$$

$$=\overline{x}-\overline{x}$$

$$=0$$

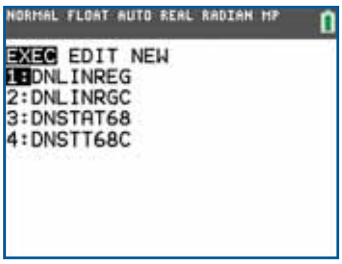
Opdracht 2 bladzijde 48

Een consumentenorganisatie onderzocht de levensduur van twee soorten batterijen. Ze kozen lukraak twintig batterijen van elke soort en gingen na hoeveel uren een zaklamp met elke batterij bleef branden. De resultaten zie je hiernaast.



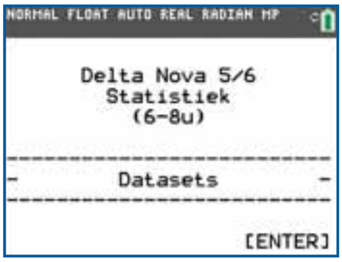
Via Scoodle zijn enkele programma's te vinden voor grafische rekentoestellen van Texas Instruments. Met deze programma's kunnen de datasets bij de opdrachten snel ingeladen worden.

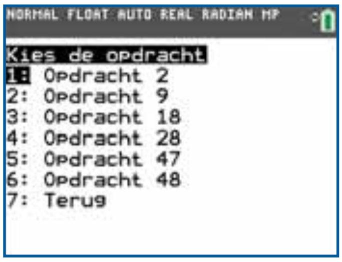
Op de schermafdruk hieronder zijn de programma's DN68STAT en DN68STTC zichtbaar, die het hoofdmenu oproepen. Het eerste is geschreven voor de TI-83/84 Plus met monochroom scherm, het laatste is aangepast aan het grotere scherm van de TI-84 Plus C en andere varianten met een kleurenscherm. Op elk toestel is maar één van beide programma's nodig.

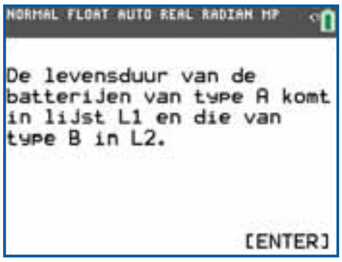


Het programma DNLINREG of DNLINRGC is enkel nodig indien de datasets van hoofdstuk 4 ingeladen moeten worden. Het eerste wordt gebruikt met DN68STAT, het tweede met DN68STTC. Ze zijn nogal groot en kunnen best weggelaten of gearchiveerd worden indien lineaire regressie niet wordt behandeld.

Via een eenvoudige menustructuur worden de datasets geladen.







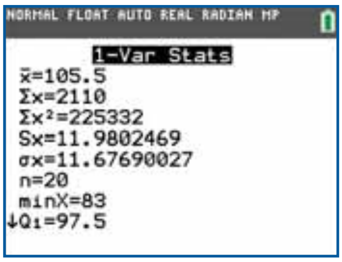
Telkens het programma wordt uitgevoerd, worden de lijsten L_1 en L_2 overschreven.

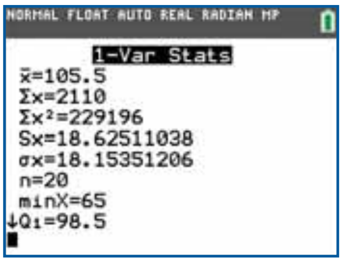
1 Bepaal op basis van de grafische voorstelling welk type batterij wellicht de grootste standaardafwijking heeft.

De grafische voorstelling laat vermoeden dat de batterijen van type B een grotere standaardafwijking zullen hebben: de aanwezigheid van enkele zeer kleine en grote waarden zal wellicht een relatief groot effect hebben.

2 Bereken nu ook van beide de standaardafwijking.

Via $\boxed{\text{STAT}}$, menu **CALC**, vind je het commando **1: 1-Var Stats** om de standaardafwijkingen van beide gegevensreeksen te laten berekenen.

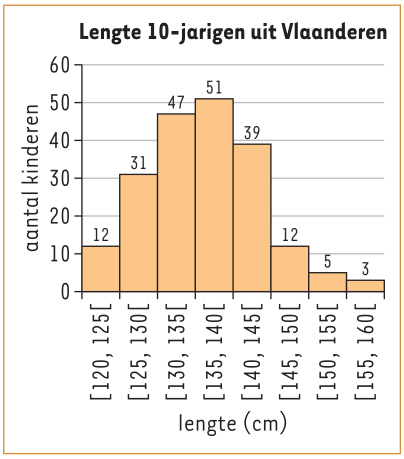




De steekproefstandaardafwijking voor de batterijen van type A is gelijk aan 11,98 en voor type B is dit 18,63, wat een bevestiging is van wat op basis van de grafische voorstelling werd vermoed.

Opdracht 3 bladzijde 49

De lengtes (in cm) van 200 tienjarige kinderen uit Vlaanderen zijn weergegeven in het histogram.



1 Bereken het percentage kinderen dat minstens 140 cm groot is met behulp van de aantallen op het histogram.

$$\frac{39+12+5+3}{200}=\frac{59}{200}=29{,}5\ \%$$

2 Bereken de verhouding van de oppervlakte van de rechthoekjes die overeenkomen met de kinderen die groter zijn dan of gelijk aan 140 cm, en de oppervlakte van het volledige histogram.

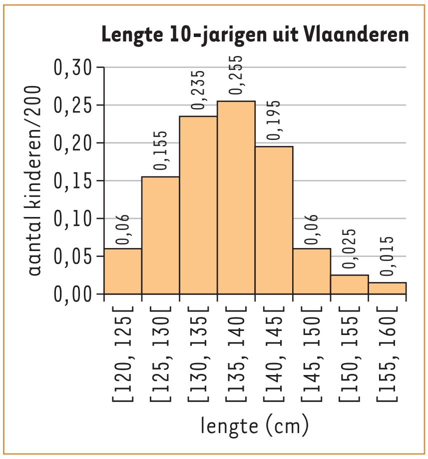
Wat stelt dit getal voor?

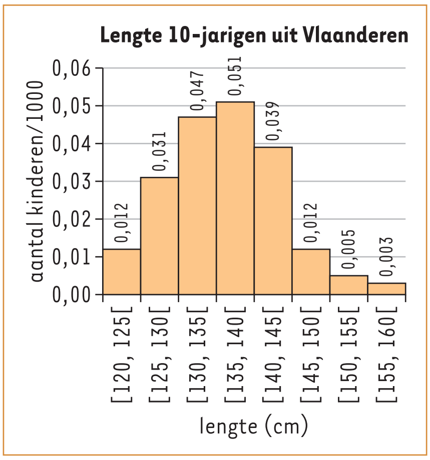
$$\frac{5\cdot(39+12+5+3)}{5\cdot(12+31+47+51+39+12+5+3)}=\frac{5\cdot 59}{5\cdot 200}=\frac{295}{1000}=0{,}295$$

Het is de relatieve frequentie van de kinderen die minstens 140 cm groot zijn.

3 Bij de onderstaande twee histogrammen is de schaal op de verticale as gewijzigd ten opzichte van het bovenstaande histogram: links zijn de waarden gedeeld door 200 (het aantal gegevens), rechts door 1000 (de oppervlakte van het histogram hierboven). De algemene vorm van het histogram wijzigt niet door de verandering van de schaal.

Bereken in beide gevallen de oppervlakte van de rechthoekjes die overeenkomen met kinderen die groter zijn dan of gelijk aan 140 cm en vergelijk met je antwoord in vraag **2**.





Oppervlakte van het linkse histogram: $5\cdot(0{,}195+0{,}06+0{,}025+0{,}015)=1{,}475$.

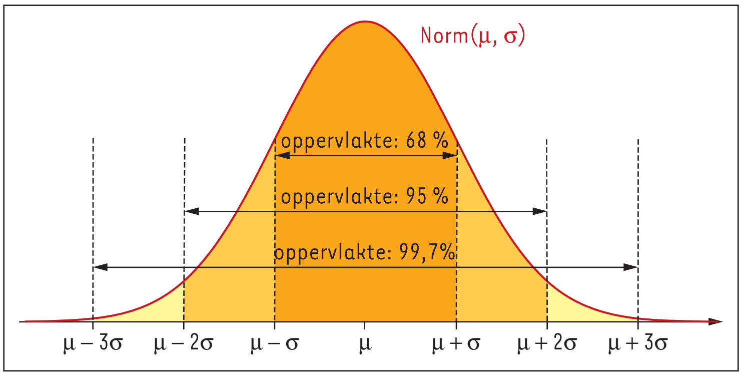
Van het rechtse histogram: $5\cdot(0{,}039+0{,}012+0{,}005+0{,}003)=0{,}295$. Deze oppervlakte komt precies overeen met het antwoord uit vraag 2, de relatieve frequentie van het aantal kinderen die minstens 140 cm groot zijn.

4 Het linkse histogram hierboven geeft op de verticale as de relatieve frequentie van elke klasse weer. Bepaal, zonder te rekenen maar enkel door te redeneren, de totale oppervlakte van het histogram. Wat zou de totale oppervlakte zijn indien de klassenbreedte 7 zou zijn?

Uit de berekeningen hierboven kun je inzien dat de oppervlakte gelijk is aan de klassenbreedte, vermenigvuldigd met de som van de hoogtes van de rechthoekjes. Die som is 1 wanneer de hoogtes relatieve frequenties voorstellen. De oppervlakte van het linkse histogram zal dus 5 zijn. Was de klassenbreedte 7, dan zou de totale oppervlakte van het histogram ook 7 zijn.

Opdracht 4 bladzijde 59

Beschouw de verdeling Norm(μ, σ). Beantwoord aan de hand van een schets: wat is de oppervlakte onder die normale verdeling in het interval



1 $]-\infty, \mu - 2\sigma]$?

Tussen $\mu - 2\sigma$ en $\mu + 2\sigma$ is de oppervlakte 95 %. De oppervlakte erbuiten is dus 5 %.

De gevraagde oppervlakte is de helft hiervan, dus 2,5 % of 0,025.

2 $[\mu - 2\sigma, \mu - \sigma]$?

Tussen $\mu - 2\sigma$ en $\mu + 2\sigma$ is de oppervlakte 95 % en tussen $\mu - \sigma$ en $\mu + \sigma$ is de oppervlakte 68 %.

Het verschil is 27 %.

De gevraagde oppervlakte is de helft hiervan, dus 13,5 % of 0,135.

3 $[\mu, \mu + \sigma]$?

Tussen μ en $\mu + \sigma$ is de oppervlakte de helft van 68 %, dus 34 % of 0,34.

4 $[\mu + 2\sigma, \mu + 3\sigma]$?

Tussen $\mu - 3\sigma$ en $\mu + 3\sigma$ is de oppervlakte 99,7 % en tussen $\mu - 2\sigma$ en $\mu + 2\sigma$ is de oppervlakte 95 %. Het verschil is 4,7 %.

De gevraagde oppervlakte is de helft hiervan, dus 2,35 % of 0,0235.

5 $[\mu + 3\sigma, +\infty[$?

Tussen $\mu - 3\sigma$ en $\mu + 3\sigma$ is de oppervlakte 99,7 %. De oppervlakte erbuiten is dus 0,3 %.

De gevraagde oppervlakte is de helft hiervan, dus 0,15 % of 0,0015.

Opdracht 5 bladzijde 59

De geboortegewichten van 100 meisjes zijn Norm(3,325; 0,503) verdeeld.
De geboortegewichten van 100 meisjes zijn Norm(3,325; 0,503) verdeeld.

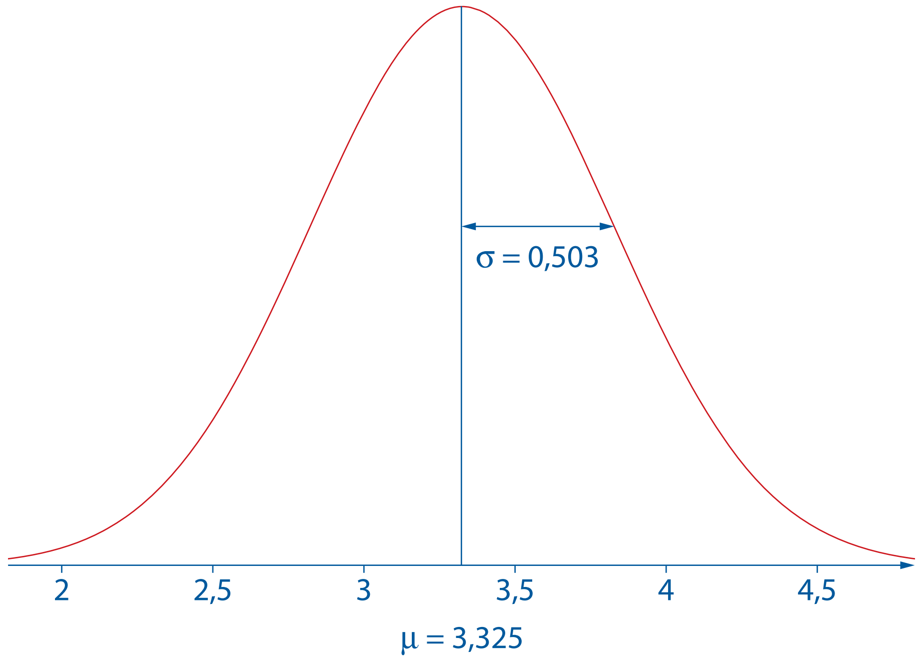
1 Tussen welke waarden op de *x*-as kun je het best de normale kromme tekenen?

De kromme wordt meestal drie standaardafwijkingen links en rechts van het gemiddelde getekend, hier dus voor x-waarden tussen $3{,}325 - 3 \cdot 0{,}503$ en $3{,}325 + 3 \cdot 0{,}503$, dus tussen 1,816 en 4,834.

2 Hoeveel bedraagt de maximale hoogte van deze kromme?

De maximale hoogte bedraagt $\frac{0{,}4}{\sigma} = \frac{0{,}4}{0{,}503} \approx 0{,}795$.

3 Maak een schets van de kromme en duid μ en σ aan.



Opdracht 6 bladzijde 59

De *Wechsler Intelligence Scale for Children* is een intelligentietest die wordt gebruikt voor kinderen tussen 6 en 16 jaar. Deze test wordt zo gestandaardiseerd dat voor een bepaalde leeftijd het gemiddelde kind een IQ heeft van 100 met een standaardafwijking van 15. IQ-scores zijn normaal verdeeld.

1 Schets de normale kromme die bij deze gegevens hoort.

$\mu - 3 \cdot \sigma = 100 - 3 \cdot 15 = 55$

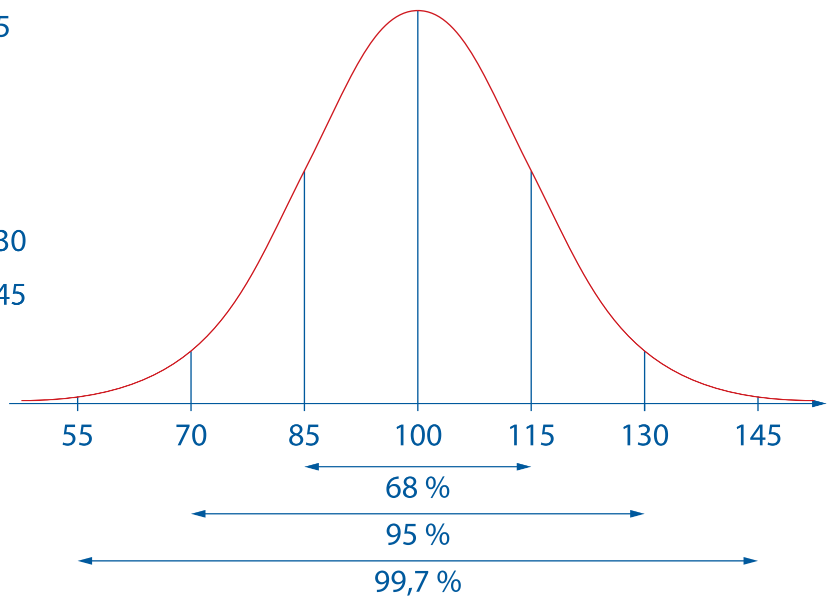
$\mu - 2 \cdot \sigma = 100 - 2 \cdot 15 = 70$

$\mu - 1 \cdot \sigma = 100 - 15 = 85$

$\mu + 1 \cdot \sigma = 100 + 15 = 115$

$\mu + 2 \cdot \sigma = 100 + 2 \cdot 15 = 130$

$\mu + 3 \cdot \sigma = 100 + 3 \cdot 15 = 145$



2 Beantwoord de volgende vragen met behulp van de 68-95-99,7-regel.

a Tussen welke IQ-grenzen bevindt zich de middelste 95 % van alle kinderen?

De middelste 95 % van alle kinderen hebben een IQ tussen $\mu - 2\sigma$ en $\mu + 2\sigma$, dus tussen 70 en 130.

b Hoeveel procent van de kinderen heeft een IQ van minder dan 70?

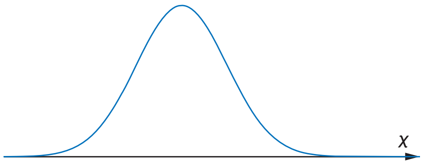
70 is twee standaardafwijkingen links van 100. De oppervlakte buiten het interval $[\mu - 2\sigma, \mu + 2\sigma]$ is 5 %. Wegens de symmetrie heeft dus 2,5 % van de kinderen een IQ minder dan 70.

c Welk is het minimale IQ van de 16 % intelligentste kinderen?

Tussen $\mu - \sigma$ en $\mu + \sigma$ is de oppervlakte 68 %. De oppervlakte erbuiten is dus 32 %. Wegens symmetrie is de oppervlakte rechts van $\mu + \sigma$ gelijk aan 16 %. Het minimale IQ van de 16 % intelligentste kinderen is dus 115.

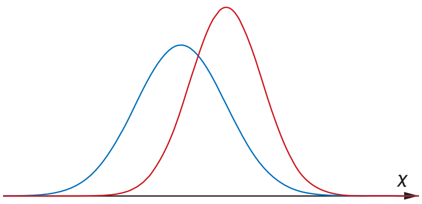
Opdracht 7 bladzijde 59

De onderstaande figuur toont een grafiek van een normale verdeling.

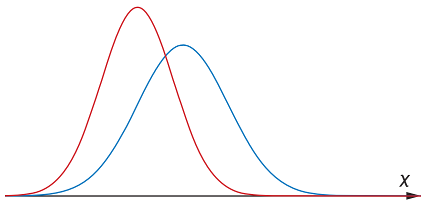


Als het gemiddelde kleiner wordt en de standaardafwijking groter wordt, welke figuur verkrijg je dan?

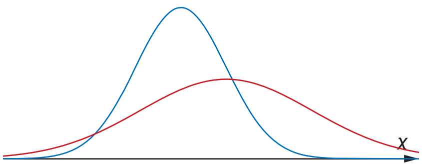
A



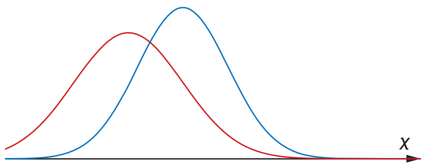
C



B



D



Het gemiddelde wordt kleiner, de grafiek wordt dus naar links verschoven. De standaardafwijking wordt groter, de grafiek wordt dus breder en minder hoog, want de totale oppervlakte moet 1 zijn. Je krijgt dan figuur D.

Opdracht 8 bladzijde 61

Plot de volgende normale krommen telkens samen in een gepast assenstelsel en verklaar gelijkenissen en verschillen.

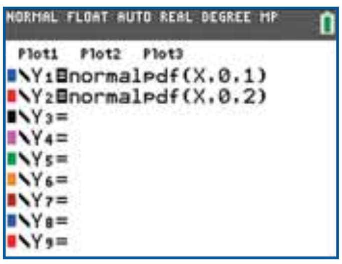
1 Norm(0, 1) en Norm(0, 2)

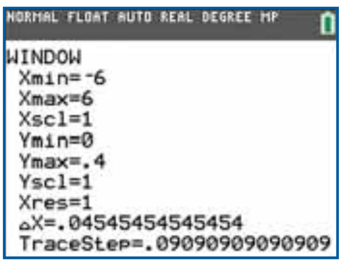
Vensterinstelling: voor Norm(0, 1): $x_{min} = -3$ en $x_{max} = 3$, $y_{min} = 0$ en $y_{max} = 0{,}4$

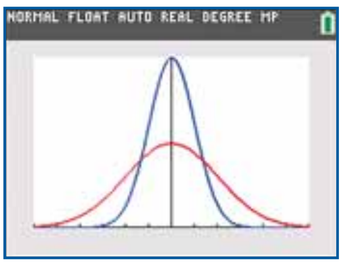
Vensterinstelling: voor Norm(0, 2): $x_{min} = -6$ en $x_{max} = 6$, $y_{min} = 0$ en $y_{max} = 0{,}2$

Om beide krommen voldoende zichtbaar te plotten, kiezen we als vensterinstelling:

$x_{min} = -6$ en $x_{max} = 6$, $y_{min} = 0$ en $y_{max} = 0{,}4$.







Beide grafieken hebben dezelfde symmetrieas. De tweede grafiek is twee keer zo breed en ook maar half zo hoog.

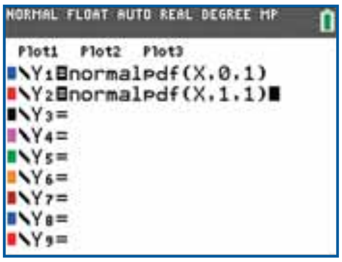
2 Norm(0, 1) en Norm(1, 1)

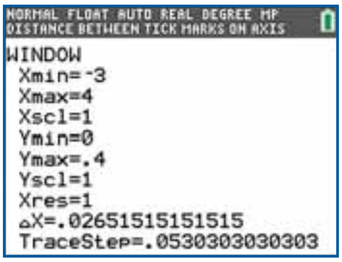
Vensterinstelling: voor Norm(0, 1): $x_{min} = -3$ en $x_{max} = 3$, $y_{min} = 0$ en $y_{max} = 0{,}4$

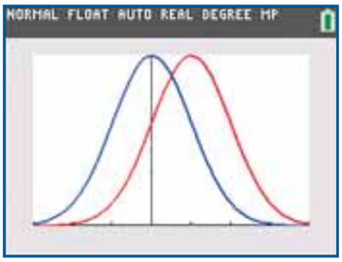
Vensterinstelling: voor Norm(1, 1): $x_{min} = -2$ en $x_{max} = 4$, $y_{min} = 0$ en $y_{max} = 0{,}4$

Om beide krommen voldoende zichtbaar te plotten, kiezen we als vensterinstelling:

$x_{min} = -3$ en $x_{max} = 4$, $y_{min} = 0$ en $y_{max} = 0{,}4$.







Beide grafieken zijn even hoog. De kromme Norm(1, 1) is naar rechts verschoven over een afstand 1 ten opzichte van de kromme Norm(0, 1).

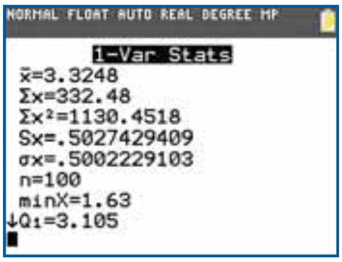
Opdracht 9 bladzijde 62

In de tabel zijn de geboortegewichten, in kg, van 100 meisjes weergegeven, gerangschikt van laag naar hoog.

1,63	2,77	2,99	3,13	3,22	3,31	3,45	3,54	3,76	3,95
1,86	2,77	3,04	3,13	3,22	3,31	3,45	3,54	3,76	3,95
2,31	2,77	3,04	3,13	3,27	3,31	3,45	3,58	3,76	4,08
2,40	2,81	3,08	3,18	3,27	3,31	3,45	3,58	3,76	4,08
2,45	2,81	3,08	3,22	3,27	3,36	3,49	3,63	3,76	4,13
2,45	2,86	3,13	3,22	3,27	3,36	3,54	3,67	3,81	4,13
2,54	2,90	3,13	3,22	3,27	3,36	3,54	3,67	3,90	4,22
2,63	2,90	3,13	3,22	3,27	3,36	3,54	3,67	3,90	4,31
2,68	2,99	3,13	3,22	3,27	3,40	3,54	3,67	3,90	4,45
2,72	2,99	3,13	3,22	3,31	3,40	3,54	3,72	3,90	4,58

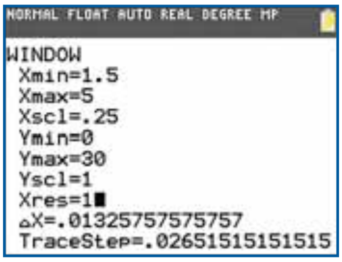
1 Controleer dat $\overline{x} = 3{,}325$ en $s = 0{,}503$.

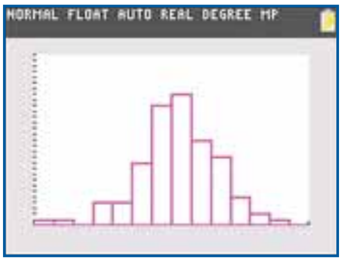
Voor deze opdracht kunnen we het programma gebruiken, zoals uitgelegd bij opdracht 2.



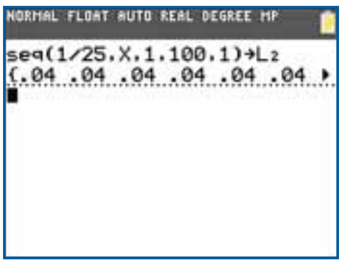
2 Teken een dichtheidshistogram en controleer dat de gegevens vrij goed gemodelleerd worden door een normale verdeling, door de normale kromme bovenop het histogram te tekenen.

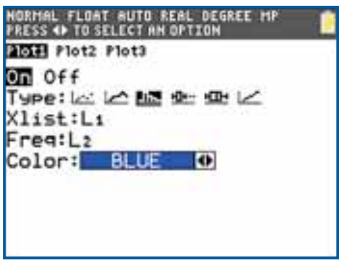
Een goede klassenbreedte voor een histogram is 0,25. De oppervlakte van het histogram is dan $0{,}25 \cdot 100 = 25$.



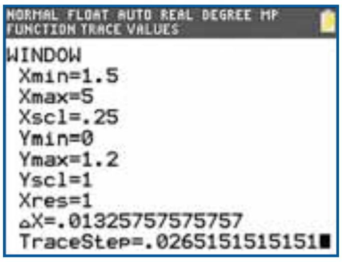


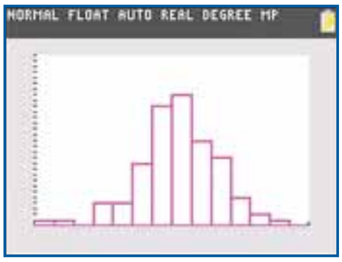
Omdat de oppervlakte van een dichtheidshistogram 1 is, zullen we werken met een frequentielijst L_2, zie schermafdrukken.



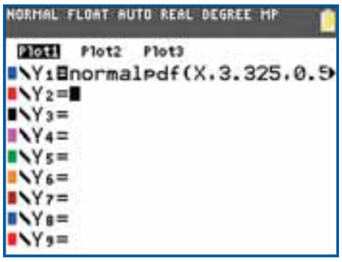


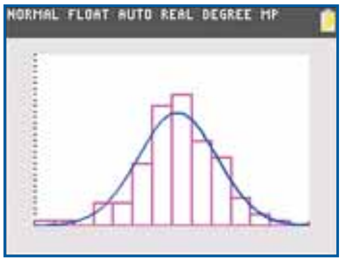
We bekijken de vorm van de verdeling. Het histogram ziet er tamelijk klokvormig uit.





We plotten ook de normale kromme met parameters 3,325 en 0,503. Het normale model sluit tamelijk goed aan bij de verdeling van de gegevens.





Je kunt ook op het oorspronkelijke histogram, met oppervlakte 25, nagaan of de normale verdeling een goed model is, door de grafiek van de functie Norm(3,325; 0,503) verticaal uit te rekken met factor 25. Deze kromme is dan wel geen dichtheidsfunctie meer, maar voor een snelle inschatting van hoe goed een model is, is dit geen probleem.

3 Bij normaal verdeelde gegevens zou ongeveer 68 % van de meetwaarden tussen $\overline{x} - s$ en $\overline{x} + s$ moeten liggen. Ga na of aan deze eigenschap voldaan is.

$\overline{x} - s = 3{,}325 - 0{,}503 = 2{,}822$

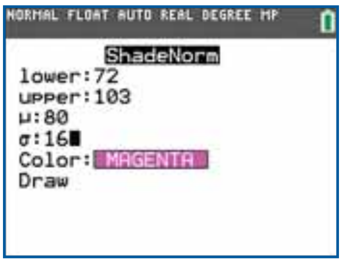
$\overline{x} + s = 3{,}325 + 0{,}503 = 3{,}828$

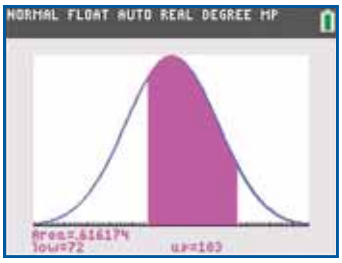
We tellen het aantal resultaten tussen 2,83 en 3,82, dat zijn er $5 + 6 \cdot 10 + 6 = 71$. Omdat er 100 gegevens zijn, is dit 71 % en dit ligt dus in de buurt van 68 %.

Opdracht 10 bladzijde 66

Gegeven de normale verdeling met $\mu = 80$ en $\sigma = 16$.

Bepaal grafisch de oppervlakte onder de kromme tussen $x = 72$ en $x = 103$.





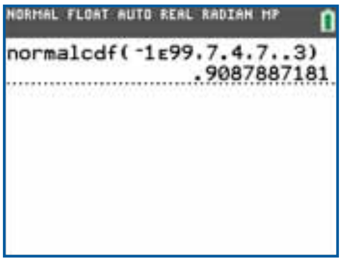
De oppervlakte is 0,616.

Opdracht 11 bladzijde 66

Neem aan dat de massa van koffiepads normaal verdeeld is met gemiddelde 7 g en standaardafwijking 0,3 g.

Hoeveel procent van de koffiepads weegt dan minder dan 7,4 g?

Ongeveer 90,9 % van de pads weegt minder dan 7,4 g.

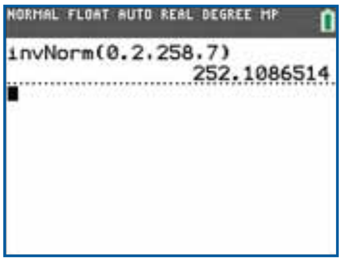


Opdracht 12 bladzijde 66

Een fabrikant produceert potten confituur. De massa (in gram) is Norm(258, 7) verdeeld.

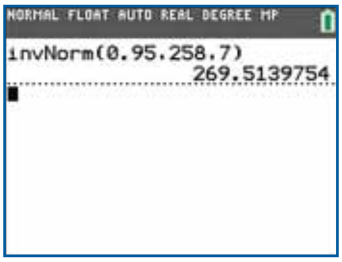
1 Wat is de maximale massa van de 20 % lichtste potten?

De maximale massa is ongeveer 252 g.



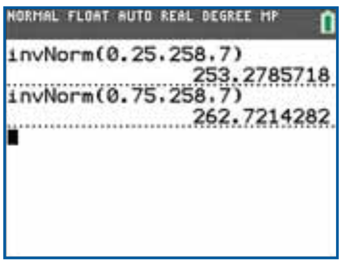
2 Wat is de minimale massa van de 5 % zwaarste potten?

De minimale massa is ongeveer 270 g.



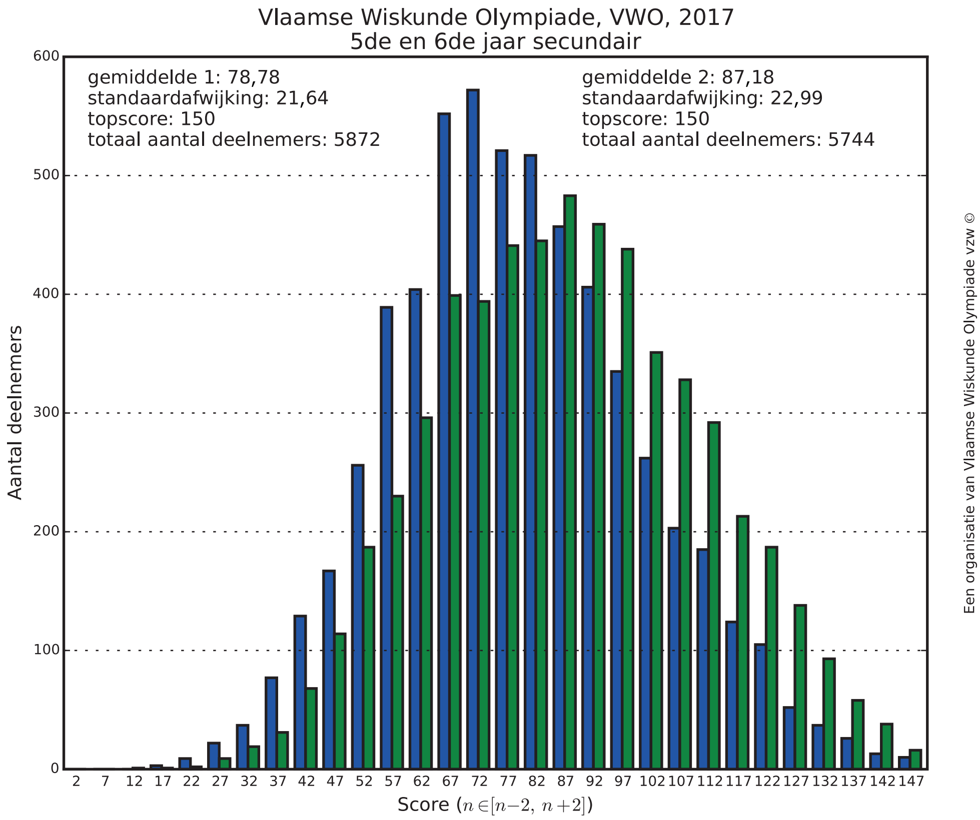
3 Tussen welke waarden bevindt zich de massa van de middelste 50 % van de confituurpotten?

De massa bevindt zich tussen ongeveer 253 g en 263 g.



Opdracht 13 bladzijde 67

De resultaten van de eerste ronde van de Vlaamse Wiskunde Olympiade van 2017 zijn bij benadering normaal verdeeld. Bente en haar broer Rune hebben beiden meegedaan. Zij zit in het vijfde jaar en hij in het zesde jaar. Bente haalde 90 en Rune 98. Rune kan het niet laten om zijn zus daarmee te plagen. Volgens de website van de organisatoren was het gemiddelde in het vijfde jaar 78,78 met een spreiding van 21,64; voor het zesde jaar was dat 87,18 met spreiding 22,99.

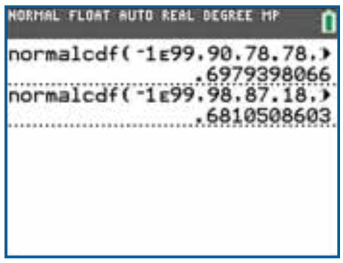


1 Wie scoorde beter ten opzichte van de eigen groep, Rune of Bente?

We berekenen het percentage leerlingen dat minder scoorde dan Rune, respectievelijk Bente, ten opzichte van hun groep.

Voor Bente is dit 69,79 % en voor haar broer Rune 68,11 %.

Bente scoorde beter ten opzichte van haar eigen groep.



2 Maarten haalde als zesdejaarsleerling 110. Schat, zonder rekentoestel te gebruiken, hoeveel procent van zijn jaargenoten het beter deed.

110 is ongeveer het gemiddelde plus 1 keer de standaardafwijking (87 + 23 = 110).

Ongeveer 16 % $\left(\frac{100-68}{2}=\frac{32}{2}=16\right)$ van zijn jaargenoten deed het beter.

3 Damian zit in het vijfde jaar, maar is ten opzichte van zijn jaargenoten even sterk als Maarten. Welk resultaat haalde Damian ongeveer?

Damian moet dus ook als resultaat ongeveer het gemiddelde plus de standaardafwijking als resultaat hebben: 78,78 + 21,64 = 100,42. Hij behaalde ongeveer 100 punten.

Opdracht 14 bladzijde 70

IQ's zijn normaal verdeeld. Als je weet dat een IQ-test een gemiddelde van 100 oplevert met een standaardafwijking van 15, welke *z*-score komt dan overeen met een IQ van 118?

Leg in woorden uit wat dit getal betekent.

$$z = \frac{118 - 100}{15} = 1{,}2$$

Een IQ van 118 bevindt zich 1,2 keer de standaardafwijking boven het gemiddelde IQ.

Opdracht 15 bladzijde 71

In Nederland hadden de dienstplichtigen in 1987 een gemiddelde lengte van 181,3 cm met een standaardafwijking van 7 cm.

1 Geef de *z*-score van een jongen die 170 cm lang is.

$$z = \frac{170 - 181{,}3}{7} = -1{,}61$$

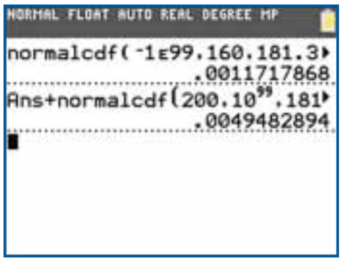
2 Michiels *z*-score is 2,5. Hoe lang is hij?

$x = \mu + z \cdot \sigma = 181{,}3 + 2{,}5 \cdot 7 = 198{,}8$

Michiel is 198,8 cm lang.

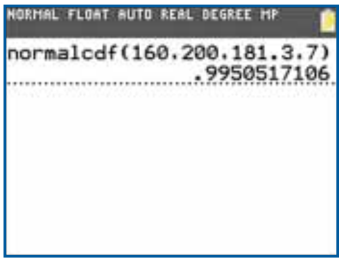
3 Jongens die kleiner waren dan 160 cm of groter dan 200 cm werden afgekeurd voor militaire dienst. Hoeveel procent van alle dienstplichtigen werd op die manier afgekeurd?

Ongeveer 0,5 % werd afgekeurd.



Je kunt dit resultaat ook verkrijgen door het percentage jongens te berekenen met een lengte tussen 160 cm en 200 cm.

Ongeveer 99,5 % werd aanvaard en dus werden er ongeveer 0,5 % jongens afgekeurd voor de militaire dienst.

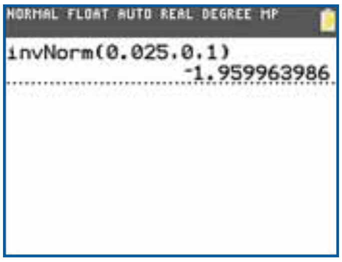


Opdracht 16 bladzijde 71

Een fabrikant koopt een vulmachine die literflessen met water vult. De machine geeft een normale inhoudsverdeling met een standaardafwijking $\sigma = 0{,}1$ cl.

Op welk gemiddelde moet de vulmachine ingesteld worden om ervoor te zorgen dat hoogstens 2,5 % van de flessen minder dan 1 liter inhoud heeft?

We berekenen de z-score die overeenkomt met een oppervlakte van 2,5 %.



De z-score is ongeveer –1,96. De corresponderende inhoud is 100 cl.

Uit $x = \mu + z \cdot \sigma$ volgt dan:

$100 = \mu - 1{,}96 \cdot 0{,}1$

$\Leftrightarrow \quad \mu = 100 + 1{,}96 \cdot 0{,}1 = 100{,}196$

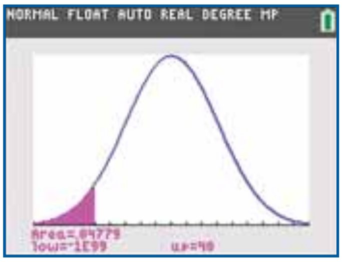
De vulmachine moet op een gemiddelde van 100,20 cl ingesteld worden.

Opdracht 17 bladzijde 73

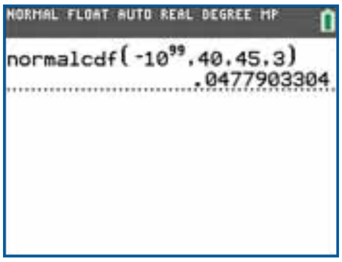
Een bepaald type kaars heeft een brandduur van gemiddeld 45 u, met een standaardafwijking van 3 u. De brandduur is normaal verdeeld.

1 Wat is de kans dat zo'n kaars minder dan 40 u zal branden?

$P(x < 40) = 0{,}048$

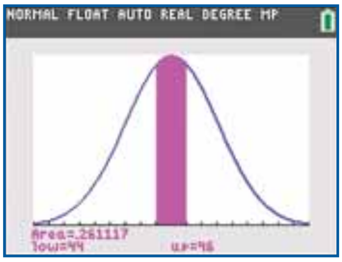


of

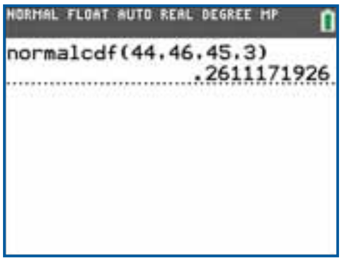


2 Wat is de kans dat ze tussen 44 en 46 u zal branden?

$P(44 < x < 46) = 0{,}261$

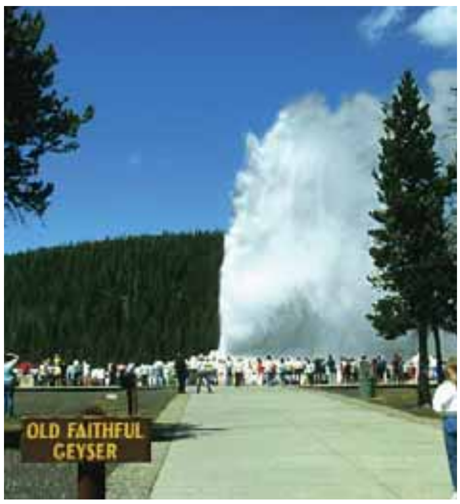


of



Opdracht 18 bladzijde 85

Een geiser is een natuurfenomeen waarbij zich ondergronds water ophoopt, tot de druk zo hoog is dat er zich een eruptie voordoet. Op dat ogenblik valt de druk op het water weg en wordt het oververhitte water in stoom omgezet. Dat geeft een nogal explosief effect. Het mengsel van stoom en water kan een hoogte van 60 m bereiken. Tussen twee erupties in kun je gerust dichterbij komen, maar dat houdt risico's in …



In het Yellowstone Park in de Verenigde Staten is een geiser die bekendstond om zijn regelmaat. Tussen twee erupties zat meestal ongeveer dezelfde tijd: in 1950 was dat zo'n 62 minuten. Daarom werd hij Old Faithful gedoopt. Om de een of andere (ondergrondse) reden is hij echter niet alleen trager geworden, die regelmaat is bovendien wat verloren gegaan. Vandaar dat in de loop van 1978-1979 een aantal metingen zijn uitgevoerd om te onderzoeken of er nog wel een regelmaat te bespeuren was.

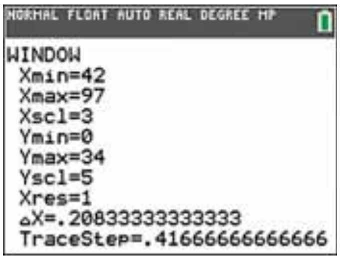
Hieronder vind je de 222 opgemeten tijden tussen twee erupties (in minuten).

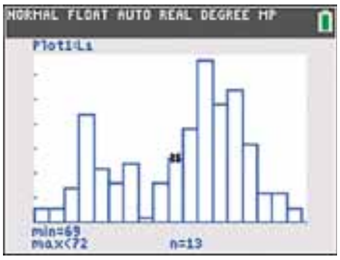
78	69	85	84	71	78	70	47	70	51	81	75	84	83
74	57	45	70	72	81	88	76	83	81	59	67	72	76
68	90	88	79	77	53	75	61	53	52	82	69	54	51
76	42	51	60	55	89	83	75	82	76	80	84	75	90
80	91	80	86	75	44	61	72	62	73	54	58	74	71
84	51	49	71	73	78	78	74	73	84	75	90	51	49
50	79	82	67	70	61	61	69	84	72	73	82	91	88
93	53	75	81	83	73	81	78	58	89	57	71	60	52
55	82	73	76	50	75	51	52	82	75	80	80	80	79
76	51	67	83	95	73	80	91	77	57	51	51	54	61
58	76	68	76	51	76	79	66	75	81	77	80	80	81
74	82	86	55	82	55	82	71	77	49	66	62	70	48
75	84	72	73	54	86	80	75	77	87	77	84	60	84
80	53	75	56	83	48	76	81	53	43	60	51	86	63
56	86	75	83	51	77	56	77	75	94	86	81	78	
80	51	66	57	80	73	82	74	78	45	62	83	51	

Via het programma met datasets (zie oplossing opdracht 2) werden de gegevens ingelezen.

1 Maak een histogram dat de gegevens zo duidelijk mogelijk weergeeft. Welke startwaarde kies je voor de tijd en welke klassenbreedte?

Een te grote klassenbreedte verbergt dat er eigenlijk twee pieken zijn bij deze verdeling (men spreekt ook van een 'bimodale verdeling').





2 Bereken het gemiddelde. Waar bevindt zich dat gemiddelde op je histogram? Is dit een zinvolle centrummaat voor deze reeks gegevens? Waarom (niet)?

Het gemiddelde is 71 (minuten). De klasse die deze waarde bevat is via TRACE geselecteerd op het histogram hierboven. Het is meteen duidelijk dat dit geen betrouwbare centrummaat is. Dat kan ook niet, aangezien er twee centra zijn.

3 Geef een beknopte beschrijving van het eruptiegedrag van Old Faithful.

Blijkbaar zijn er twee 'soorten' intervallen tussen twee erupties. De 'korte' intervallen duren gemiddeld rond de 54 à 55 minuten (centrum van de linker piek), terwijl de 'lange' intervallen een gemiddelde hebben van ongeveer 80 minuten. Er zijn meer lange dan korte intervallen; bij de lange is de variabiliteit ook groter.

Opdracht 19 bladzijde 86

Als het gemiddelde van twee reële getallen gelijk is aan hun verschil, dan verhouden deze getallen zich als

A 3 : 2 **B** 2 : 1 **C** 5 : 2 **D** 3 : 1 **E** 7 : 2

(Bron © VWO eerste ronde, 2014)

Stel x_1 en x_2 de twee getallen, met $x_1 > x_2$. Uit $\frac{x_1 + x_2}{2} = x_1 - x_2$ volgt gemakkelijk dat $3x_2 = x_1$ zodat x_1 drie keer groter is dan x_2. Ze verhouden zich als 3 tot 1.
Antwoord D is dus het juiste.

Opdracht 20 bladzijde 86

De eigenaar van een bioscoopcomplex verhoogt de huurprijs voor het snoepkraam dat er wordt uitgebaat met 5 %. De uitbater van het snoepkraam wil zijn woekerwinsten niet zien slinken en besluit de prijs van alle snoepsoorten met € 0,50 per 100 g te verhogen.

1 Zal het gemiddelde van de snoepprijzen afnemen, toenemen of gelijk blijven?

Het gemiddelde zal ook met 0,50 euro per 100 g toenemen.

2 Zal de populatiestandaardafwijking van de snoepprijzen afnemen, toenemen of gelijk blijven?

De spreiding zal niet toenemen. Op een grafische voorstelling zouden alle prijzen 0,5 euro naar rechts opschuiven, zodat de onderlinge afstand tussen de prijzen niet zou veranderen.

Zie opdracht 23 voor een algebraïsche verantwoording.

Opdracht 21 bladzijde 86

Je rijdt 25 km met 50 km/h en dan 25 km met 90 km/h. Reed je dan gemiddeld met 70 km/h? Verklaar je antwoord met een berekening.

Het verband tussen de snelheid v, de positieverandering Δx en de bijbehorende verandering in de tijd Δt is $v = \frac{\Delta x}{\Delta t}$. Daarbij kunnen de eenheden respectievelijk m/s, m en s zijn, of km/h, km en h. Hieruit volgt: $\Delta t = \frac{\Delta x}{v}$. Het duurt dus $\frac{25}{50} = \frac{1}{2}$ h om de eerste 25 kilometer te rijden en $\frac{25}{90} = \frac{5}{18}$ h voor de laatste 25. In totaal is dus $\frac{1}{2} + \frac{5}{18} = \frac{7}{9}$ h gereden over 50 km, zodat de gemiddelde snelheid $\frac{50}{\frac{7}{9}} = \frac{450}{7} \approx 64$ km/h is. Dit is niet gelijk aan 70, het gemiddelde van 50 en 90. (Die $\frac{450}{7}$ is het *harmonisch gemiddelde* van 50 en 90.)

Opdracht 22 bladzijde 86

Een leraar wil voor elke leerling uit zijn klas de gemiddelde score *m* van drie toetsen bepalen. Hij berekent voor elk van de leerlingen eerst het gemiddelde van de twee minst goede scores en daarna het gemiddelde van dat resultaat en de beste score. Hij vindt voor elke leerling steeds een getal dat

A strikt kleiner is dan *m*,

B strikt groter is dan *m*,

C gelijk is aan *m*,

D niet kleiner is dan *m*,

E niet groter is dan *m*.

(Bron © VWO eerste ronde, 2000)

Stel $x_1 \leqslant x_2 \leqslant x_3$ de drie resultaten van een leerling. Het gemiddelde is $\frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3$.

De leraar berekent $\dfrac{\frac{x_1 + x_2}{2} + x_3}{2} = \frac{1}{4}x_1 + \frac{1}{4}x_2 + \frac{1}{2}x_3$.

In deze berekening heeft de beste toets een groter gewicht in het eindresultaat dan bij het gemiddelde. Zijn de drie resultaten niet gelijk, dan zal het getal dat hij berekent groter zijn dan het gemiddelde; zijn ze wel gelijk, dan zullen beide getallen even groot zijn.

Dit kan ook via een berekening aangetoond worden, door beide getallen van elkaar af te trekken:

$$\begin{aligned}\frac{1}{4}x_1 + \frac{1}{4}x_2 + \frac{1}{2}x_3 - \frac{1}{3}x_1 - \frac{1}{3}x_2 - \frac{1}{3}x_3 &= -\frac{1}{12}x_1 - \frac{1}{12}x_2 + \frac{2}{12}x_3 \\ &= \frac{1}{12}(x_3 - x_1) + \frac{1}{12}(x_3 - x_2)\end{aligned}$$

Van zodra x_3 strikt groter is dan een van de andere resultaten, is het verschil tussen wat de leerkracht berekent en het gemiddelde positief, anders is het nul.

Antwoord D is dus correct.

Opdracht 23 bladzijde 87

Bij Newcombs onderzoek naar de lichtsnelheid op p. 14 e.v., rekenden we niet met de werkelijk opgemeten tijden. Van de exacte tijden werd immers telkens 0,000 024 8 s afgetrokken. We onderzoeken het effect van die bewerking op het gemiddelde en de standaardafwijking.

Algemeen: beschouw een steekproef van meetwaarden x_i $(i = 1, 2, \ldots, n)$, met gemiddelde $\overline{x}$ en standaardafwijking s_x. Vervang alle meetwaarden x_i door $x_i + a = y_i$.

1 Bereken het gemiddelde $\overline{y}$ en de standaardafwijking s_y van deze y-waarden.

Het gemiddelde neemt toe met a, maar de standaardafwijking verandert niet.

$$\begin{aligned}\overline{y} &= \frac{y_1 + y_2 + \ldots + y_n}{n}\\ &= \frac{x_1 + a + x_2 + a + \ldots + x_n + a}{n}\\ &= \frac{x_1 + x_2 + \ldots + x_n + n \cdot a}{n}\\ &= \frac{x_1 + x_2 + \ldots + x_n}{n} + \frac{\cancel{n} \cdot a}{\cancel{n}}\\ &= \overline{x} + a\end{aligned}$$

$$\begin{aligned}s_y &= \sqrt{\frac{(y_1 - \overline{y})^2 + (y_2 - \overline{y})^2 + \ldots + (y_n - \overline{y})^2}{n-1}}\\ &= \sqrt{\frac{\left(x_1 + a - (\overline{x} + a)\right)^2 + \left(x_2 + a - (\overline{x} + a)\right)^2 + \ldots + \left(x_n + a - (\overline{x} + a)\right)^2}{n-1}}\\ &= \sqrt{\frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n-1}}\\ &= s_x\end{aligned}$$

2 Wat gebeurt er met het histogram van de waarden x_i wanneer je bij alle waarden de waarde a optelt? Maak een schets van x- en overeenkomstige y-waarden ter illustratie. Is het histogram van y-waarden in overeenstemming met je bewering over $\overline{y}$ en s_y?

Het histogram zal horizontaal verschuiven met a eenheden naar rechts ($a > 0$) of links ($a < 0$). De breedte blijft dus dezelfde, maar het centrum verschuift.

3 Zelfde vragen als bij **2**, maar nu is $y_i = a \cdot x_i$ (met $a \neq 0$).

Het gemiddelde wordt met a vermenigvuldigd; de standaardafwijking met de absolute waarde van a. In de meeste toepassingen zal a strikt positief zijn, zodat $s_y = a \cdot s_x$.

$$\overline{y} = \frac{ax_1 + ax_2 + \ldots + ax_n}{n} = a \cdot \frac{x_1 + x_2 + \ldots + x_n}{n} = a\overline{x}$$

$$\begin{aligned}s_y &= \sqrt{\frac{(ax_1 - a\overline{x})^2 + (ax_2 - a\overline{x})^2 + \ldots + (ax_n - a\overline{x})^2}{n-1}}\\ &= \sqrt{a^2 \cdot \frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n-1}}\\ &= \sqrt{a^2} \cdot \sqrt{\frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n-1}}\\ &= |a| \cdot s_x\end{aligned}$$

Opdracht 24 bladzijde 87

Een tabel met strikt positieve getallen bestaat uit 40 rijen en 75 kolommen. Het gemiddelde van de 40 rijsommen is gelijk aan A en het gemiddelde van de

75 kolomsommen is gelijk aan B. Wat is $\frac{A}{B}$?

A $\frac{1}{2}$ **B** $\frac{8}{15}$ **C** 1 **D** $\frac{3}{2}$ **E** $\frac{15}{8}$

(Bron © VWO tweede ronde, 2014)

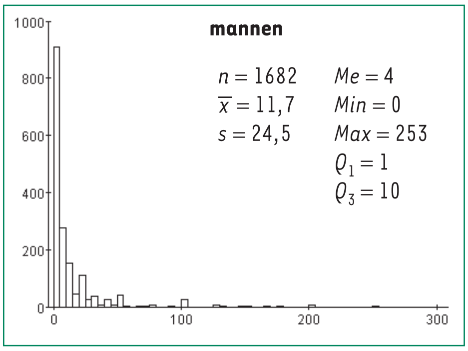
De som van alle getallen in de tabel is zowel gelijk aan 40A als aan 75B.

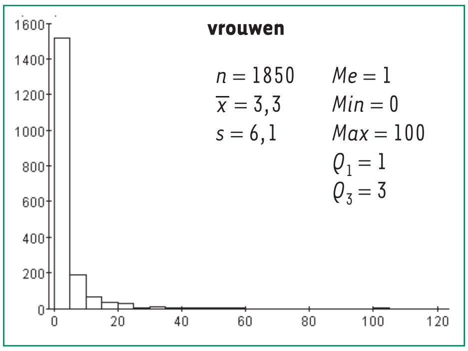
Uit 40A = 75B volgt dat $\frac{A}{B} = \frac{75}{40} = \frac{15}{8}$.

Opdracht 25 bladzijde 87

De aidsepidemie die in de jaren 80 van de vorige eeuw de kop opstak, zorgde ervoor dat het Amerikaanse Congres toestemming gaf voor uitgebreid onderzoek naar het gedrag van Amerikanen op het gebied van seks en drugs, aangezien het hiv-virus voornamelijk via seksueel contact en besmette naalden verspreid werd. In de jaarlijks georganiseerde General Social Survey, bedoeld voor Amerikanen van 18 of ouder, werden vanaf 1989 een aantal vragen i.v.m. seksueel gedrag toegevoegd. Aan de respondenten werd o.a. gevraagd hoeveel mannelijke en vrouwelijke seksuele partners ze hadden gehad sinds ze 18 waren. Dat gebeurde via een invulformulier dat de deelnemers zelfstandig konden invullen, om eventuele gêne tegenover de interviewer te verminderen.

Hieronder vind je een grafisch en numeriek overzicht van de aantallen voor mannen en vrouwen. De klassenbreedte is telkens 5; de eerste klasse is [0, 5[.





De verhouding mannen/vrouwen in de studie komt overeen met de Amerikaanse bevolking van 18 of ouder: 47,62 % bestaat uit mannen.

1 Wat kunnen redenen zijn voor een man om minder dan wel meer partners op te geven dan in werkelijkheid het geval is?

Redenen om minder partners op te geven: schaamte, mening 'dat dit niemand aangaat' … Redenen om meer partners op te geven: opschepperij, een zekere schaamte omdat het er zo weinig zijn …

2 En voor een vrouw?

Maatschappelijk is het wellicht ook in de V.S. zo dat vrouwen die veel seksuele partners hebben negatiever beoordeeld worden dan mannen in die situatie, wat vrouwen er misschien toe kan aanzetten om minder partners op te geven. Maar dezelfde redenen als bij mannen zijn ook mogelijk.

3 Eén man rapporteert precies 253 seksuele partners. Wat denk je van de betrouwbaarheid van dergelijk cijfer?

Dit cijfer is verdacht accuraat. Ofwel houdt die man een boekje bij waarin hij het aantal partners netjes registreert, ofwel heeft iemand een grapje willen uithalen. Geloofwaardiger ware 250 geweest, als een soort schatting. In het histogram is trouwens duidelijk dat een aantal mensen ronde schattingen opgaven: 50, 100, 200.

4 Je kunt de 'gemiddelde Amerikaan' voor wat betreft het aantal partners omschrijven d.m.v. het gemiddelde, de mediaan of de zgn. 'modus': dat is de waarde die het vaakst voorkomt. Zowel bij mannen als bij vrouwen blijkt uit de gegevens dat de modus 1 is. Welke van die drie centrummaten geeft volgens jou het beste de 'gemiddelde' Amerikaanse vrouw en man weer?

Het is duidelijk dat het gemiddelde geen goede centrummaat is, gezien de zeer rechtsscheve verdeling. Mediaan en modus zijn beide mogelijk. Beide geven een ander aspect van de 'gemiddelde Amerikaanse man of vrouw' weer.

Opdracht 26 bladzijde 88

Het gemiddelde van twee reële getallen a en b noteren we als $a \square b$. Bekijk de volgende uitspraken:

I. Voor alle reële getallen a, b en c geldt dat
$$a \square (b + c) = (a \square b) + (a \square c)$$
II. Voor alle reële getallen a, b en c geldt dat
$$a + (b \square c) = (a + b) \square (a + c)$$
III. Voor alle reële getallen a, b en c geldt dat
$$a \square (b \square c) = (a \square b) \square (a \square c)$$

Dan geldt:

A Precies één van deze uitspraken is correct.

B De enige verkeerde uitspraak is I.

C De enige verkeerde uitspraak is II.

D De enige verkeerde uitspraak is III.

E De drie uitspraken zijn correct.

(Bron © VWO tweede ronde, 2015)

Voor de drie uitspraken moeten zowel het linker- als het rechterlid berekend worden met de formule $a \square b = \dfrac{a+b}{2}$.

Voor de eerste vind je zo:

- $a \square (b+c) = \dfrac{a+(b+c)}{2} = \dfrac{a+b+c}{2}$
- $(a \square b) + (a \square c) = \dfrac{a+b}{2} + \dfrac{a+c}{2} = \dfrac{2a+b+c}{2}$

Beide leden zijn niet gelijk, zodat uitspraak I zeker al verkeerd is.

In uitspraak II zijn beide leden gelijk aan $\dfrac{2a+b+c}{2}$ en in uitspraak III zijn ze beide gelijk aan $\dfrac{2a+b+c}{4}$.

Het juiste antwoord is dus B.

Opdracht 27 bladzijde 89

De bissectrice (deellijn) van de scherpe hoek gevormd door de rechten $y = x$ en $y = 3x$ heeft een richtingscoëfficiënt die gelijk is aan

A het **rekenkundig gemiddelde** van 1 en 3, nl. $\frac{1+3}{2} = 2$.

B het **meetkundig gemiddelde** van 1 en 3, nl. $\sqrt{1 \cdot 3} = \sqrt{3}$.

C het **harmonische gemiddelde** van 1 en 3, nl. $\frac{2}{\frac{1}{1}+\frac{1}{3}} = \frac{3}{2}$.

D het **kwadratisch gemiddelde** van 1 en 3, nl. $\sqrt{\frac{1^2+3^2}{2}} = \sqrt{5}$.

E het getal van de gulden snede, nl. $\frac{1+\sqrt{5}}{2}$.

(Bron © VWO tweede ronde, 2012)

Stel α_1 en α_2 de hoeken die de rechten met vergelijking $y = x$ resp. $y = 3x$ maken met de positieve x-as. Dan geldt dat $\tan\alpha_1 = 1$ en $\tan\alpha_2 = 3$.

De bissectrice van de scherpe hoek gevormd door beide rechten maakt een hoek $\alpha = \frac{\alpha_1+\alpha_2}{2}$ met de positieve x-as. De gevraagde richtingscoëfficiënt is $\tan\alpha$.

Uit de somformule voor de tangens volgt dat $\tan(\alpha_1+\alpha_2) = \frac{\tan\alpha_1+\tan\alpha_2}{1-\tan\alpha_1\tan\alpha_2} = \frac{1+3}{1-3} = -2$.

Hieruit volgt dat $\tan(2\alpha) = -2$.

Met de halveringsformule voor de tangens kunnen we hieruit $\tan\alpha$ berekenen.

Uit $\tan(2\alpha) = \frac{2\tan\alpha}{1-\tan^2\alpha}$ volgt:

$$\frac{2\tan\alpha}{1-\tan^2\alpha} = -2 \Leftrightarrow 2\tan^2\alpha - 2\tan\alpha - 2 = 0$$

$$\Leftrightarrow \tan^2\alpha - \tan\alpha - 1 = 0$$

$$\Leftrightarrow \tan\alpha = \frac{1\pm\sqrt{5}}{2}$$

Aangezien α in het eerste kwadrant ligt, moet $\tan\alpha$ positief zijn, zodat de gevraagde richtingscoëfficiënt gelijk is aan $\frac{1+\sqrt{5}}{2}$. Het juiste antwoord is E.

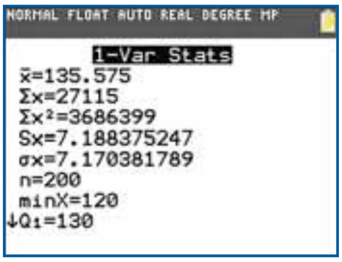
Opdracht 28 bladzijde 89

Hiernaast zie je de lengtes van 200 tienjarigen, uitgedrukt in cm.

120	121	121	122	122	122	123	123	124	124
124	124	125	125	125	126	126	126	127	127
127	127	128	128	128	128	128	128	128	128
128	128	128	129	129	129	129	129	129	129
129	129	129	130	130	130	130	130	130	130
130	131	131	131	131	131	131	131	131	132
132	132	132	132	132	132	132	132	132	132
133	133	133	133	133	133	133	133	133	133
133	133	134	134	134	134	134	134	134	134
135	135	135	135	135	135	135	135	135	135
136	136	136	136	136	136	136	136	136	136
136	137	137	137	137	137	137	137	137	137
137	137	137	137	137	138	138	138	138	138
138	138	138	139	139	139	139	139	139	139
139	140	140	140	140	140	140	140	140	140
140	141	141	141	141	141	141	141	141	141
141	141	141	142	142	142	142	142	142	142
142	142	143	143	143	143	143	144	144	144
145	145	145	145	145	145	146	146	146	146
148	149	150	151	152	153	153	157	157	157

1 Controleer dat $\overline{x} = 135{,}6$ en $s = 7{,}19$.

We lezen de gegevens in via het programma.

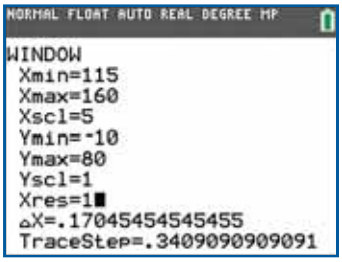


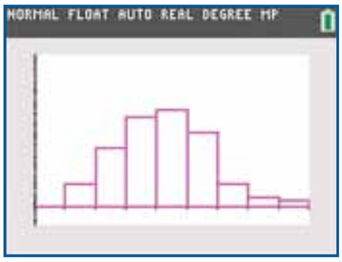
2 Onderzoek aan de hand van een grafische voorstelling of deze gegevens normaal verdeeld zijn.

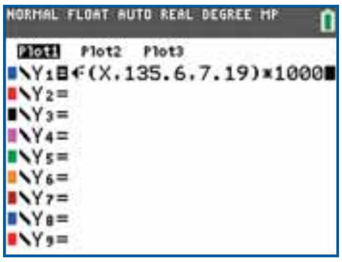
We bekijken de vorm van de verdeling.

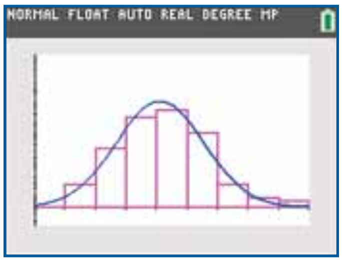
Het histogram ziet er tamelijk klokvormig uit.

Zoals vermeld in opdracht 9, kunnen we een dichtheidshistogram tekenen door met een lijst van frequenties te werken. Aangezien we hier enkel moeten nagaan of de normale verdeling een goed model is, kiezen we voor de snellere manier.









We plotten de normale kromme met parameters 135,6 en 7,19. Omdat het histogram hier geen oppervlakte 1 heeft, maar 200 · 5, zullen we de waarden bij de normale kromme ook met 1000 vermenigvuldigen. Het normale model sluit vrij goed aan bij de verdeling van de gegevens.

Opdracht 29 bladzijde 89

Je wil de verdelingen Norm(2, 3) en Norm(5, 2) beide weergeven in een assenstelsel. Welke grenzen moet je kiezen op de *x*-as indien je van beide verdelingen minstens 3 standaardafwijkingen links en rechts van het gemiddelde wil zien?

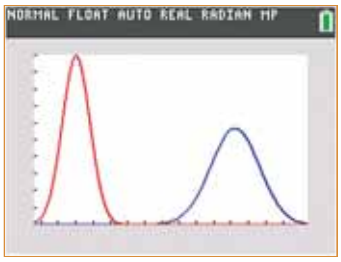
Voor Norm(2, 3): $x_{min} = -7$ en $x_{max} = 11$

Voor Norm(5, 2): $x_{min} = -1$ en $x_{max} = 11$

De grenzen op de x-as zijn: –7 en 11.

Opdracht 30 bladzijde 90

In de figuur hiernaast zie je de grafieken van de verdelingen van de lengtes van 7-jarige meisjes en van 25-jarige vrouwen.



1 Waarom is de normale verdeling die hoort bij de 7-jarigen smaller dan de grafiek bij de 25-jarige vrouwen?

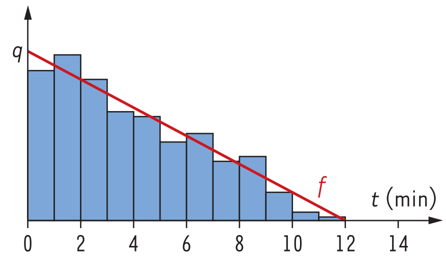
σ en de breedte zijn evenredig, dus een kleine standaardafwijking wil zeggen minder spreiding. Bij 7-jarigen zijn de onderlinge verschillen in lengtes minder groot dan bij 25-jarige vrouwen, want de 7-jarigen wijken minder af van het gemiddelde.

2 Waarom ligt bij de smallere grafiek de top hoger?

De totale oppervlakte onder de kromme is 1, dus als de kromme smaller wordt, is ze ook hoger. Uit het voorschrift leerden we: σ en de hoogte zijn omgekeerd evenredig.

Opdracht 31 bladzijde 90

Een statisticus onderzocht de wachttijden, uitgedrukt in minuten, bij een helpdesk.
De tijden variëren van 0 tot 12 minuten. Het histogram suggereert op het interval [0, 12] een eenvoudige dichtheidskromme: een rechte door de punten $(0, q)$ en $(12, 0)$. Noem de bijbehorende dichtheidsfunctie f.



1 Welke waarde moet q krijgen opdat f een dichtheidsfunctie zou zijn?

De totale oppervlakte onder de kromme moet 1 zijn, dus moet

oppervlakte driehoek $= \dfrac{12 \cdot q}{2} = 1 \Leftrightarrow q = \dfrac{1}{6}$.

2 Geef het voorschrift van f op het interval [0, 12].

De grafiek is een rechte door (12, 0) en $\left(0, \dfrac{1}{6}\right)$, de vergelijking van de rechte is:

$y = \dfrac{\frac{1}{6}}{-12}(t-12)$, dus $y = -\dfrac{1}{72}t + \dfrac{1}{6}$. Het voorschrift over [0, 12] is: $f(t) = -\dfrac{1}{72}t + \dfrac{1}{6}$.

3 Gebruik de dichtheidsfunctie om de volgende vragen te beantwoorden.

a In hoeveel procent van de gevallen was de wachttijd meer dan 5 minuten?

De oppervlakte onder de rechte over [5, 12] kan op verschillende manieren berekend worden, bijvoorbeeld als oppervlakte van een driehoek:

$$\frac{7 \cdot f(5)}{2} = \frac{7 \cdot \left(-\frac{5}{72} + \frac{1}{6}\right)}{2} = \frac{49}{144} \approx 0{,}34.$$

In ongeveer 34 % van de gevallen was de wachttijd meer dan 5 minuten.

b In hoeveel procent van de gevallen was de wachttijd minder dan 3 minuten?

De oppervlakte onder de rechte over [0, 3] kan bijvoorbeeld berekend worden als oppervlakte van een trapezium: $\dfrac{(f(0)+f(3)) \cdot 3}{2} = \dfrac{\left(\frac{1}{6} - \frac{3}{72} + \frac{1}{6}\right) \cdot 3}{2} = \dfrac{7}{16} = 0{,}4375.$

In ongeveer 44 % van de gevallen was de wachttijd minder dan 3 minuten.

c In hoeveel procent van de gevallen lag de wachttijd tussen 3,5 en 5,5 minuten?

Ook hier kan de oppervlakte onder de rechte over [3,5; 5,5] bijvoorbeeld berekend worden als oppervlakte van een trapezium:

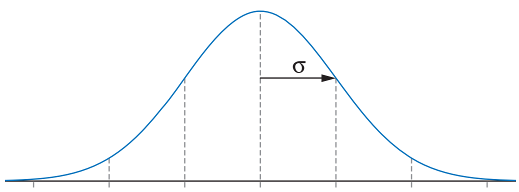
$$\frac{(f(3,5)+f(5,5))\cdot 2}{2} = -\frac{3,5}{72}+\frac{1}{6}-\frac{5,5}{72}+\frac{1}{6}=\frac{5}{24}\approx 0,208.$$

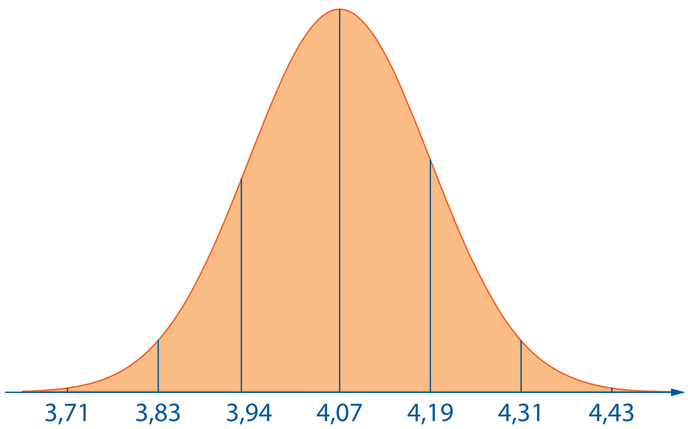
In ongeveer 21 % van de gevallen lag de wachttijd tussen 3,5 en 5,5 minuten.

Opdracht 32 bladzijde 91

Een supermarkt verkoopt waspoeder in pakken van 4 kg. Een consumentenorganisatie controleert de massa van 500 pakken. Het gemiddelde is $\overline{x} = 4,07$ kg, met een standaardafwijking $s = 0,12$ kg. De massa's zijn normaal verdeeld.

1 Maak een grafiek zoals die hiernaast en plaats op de horizontale as de massa die met elk ijkstreepje overeenkomt. De afstand tussen twee streepjes is telkens σ.



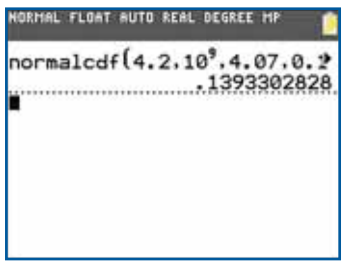


2 Hoe groot is het percentage pakken met een massa van minstens 4,20 kg?

Op de grafiek kun je met behulp van de vuistregel schatten dat het percentage pakken met een massa van minstens 4,20 kg een beetje minder dan de helft van

$\frac{100-68}{2}\,\% = 16\,\%$ zal zijn.

Met het rekentoestel vind je ongeveer 13,9 %.



3 Leg uit aan de hand van de grafiek dat 16 % van de pakken een massa zal hebben van minder dan 3,95 kg.

3,95 ligt op een standaardafwijking onder het gemiddelde, dus ongeveer

$\frac{100-68}{2}\,\% = 16\,\%$ zal nog een kleinere massa hebben.

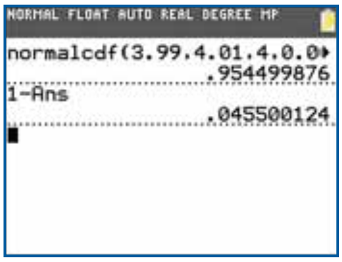
Opdracht 33 bladzijde 91

In een bedrijf worden elke dag 10 000 kogellagers gemaakt, die een diameter van 4 mm moeten hebben. Ten gevolge van onvermijdelijke afwijkingen zijn de diameters normaal verdeeld met een gemiddelde van 4 mm en een standaardafwijking van 0,005 mm.

1 Kogellagers met een diameter kleiner dan 3,99 mm of groter dan 4,01 mm worden afgekeurd. Hoeveel kogellagers worden er gemiddeld per dag afgekeurd?

Er wordt ongeveer 95,45 % goedgekeurd, dus ongeveer 4,55 % afgekeurd.

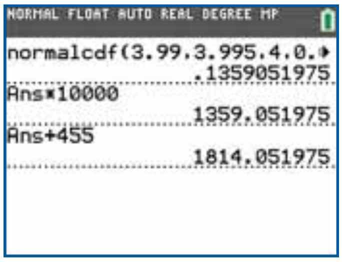
Ongeveer 455 kogellagers worden gemiddeld per dag afgekeurd.



2 Voor een bepaalde bestelling moeten ook kogellagers met een diameter tussen 3,99 mm en 3,995 mm afgekeurd worden. Hoeveel kogellagers zullen nu in totaal gemiddeld afgekeurd worden per dag?

Ongeveer 1359 kogelladers worden ook nog afgekeurd,

dus in totaal dagelijks ongeveer 1814.



Opdracht 34 bladzijde 91

Een IQ-test wordt altijd zo opgesteld dat een gemiddelde persoon uit de doelgroep 100 scoort en zo dat de standaardafwijking 15 is.

Lien is 33 en haar moeder is 61. Beiden leggen ze dezelfde IQ-test af, bedoeld voor de leeftijdsgroep van 16 tot 90 jaar. Het gemiddelde resultaat van iemand uit die groep is dus 100. Maar aangezien bepaalde mentale capaciteiten afnemen met het verouderen, varieert het gemiddelde resultaat met de leeftijd.

De leeftijdsgroep 60-64 scoort gemiddeld 90 met een standaardafwijking van 25; de leeftijdsgroep 20-34 heeft als gemiddelde 110 met standaardafwijking 15. Lien heeft een score van 125; haar moeder heeft volgens de test een IQ van 110.

Wie scoort, relatief gezien, het best ten opzichte van de eigen leeftijdsgroep?

$$z_{\text{Lien}} = \frac{125-110}{15} = 1$$

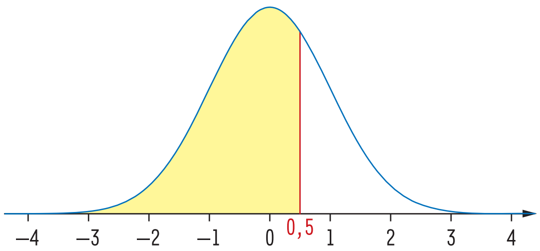
$$z_{\text{moeder}} = \frac{110-90}{25} = 0{,}8$$

Lien scoort het beste ten opzichte van de eigen leeftijdsgroep.

Opdracht 35 bladzijde 92

Hiernaast zie je een grafiek van de normale verdeling met gemiddelde 0 en standaardafwijking 1.

Van de totale oppervlakte tussen de kromme en de horizontale as is 69 % ingekleurd.



Het gewicht (de massa) van kisten mandarijnen in een supermarkt is normaal verdeeld met een gemiddelde van 3 kg. De uitbater van de supermarkt stelt vast dat 69 % van de kisten minder weegt dan 3050 g.

Wat is de standaardafwijking van het gewicht van de kisten mandarijnen? Rond af op 1 g.

(Bron: Peiling wiskunde in de derde graad aso, kso en tso)

69 % komt overeen met een x-waarde van 3050 en een z-waarde van 0,5.

$$0{,}5 = \frac{3050 - 3000}{\sigma} \Rightarrow \sigma = \frac{50}{0{,}5} = 100$$

De standaardafwijking is 100 g.

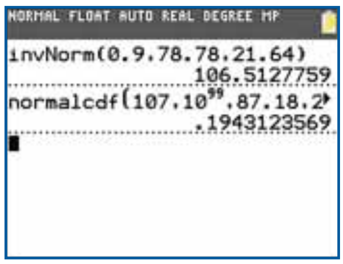
Opdracht 36 bladzijde 92

De scores op de eerste ronde van de Vlaamse Wiskunde Olympiade waren in 2017 nagenoeg normaal verdeeld, met een gemiddelde van 78,78 en een spreiding van 21,64 in het vijfde jaar en 87,18 respectievelijk 22,99 in het zesde jaar.

1 Welke score hadden de 10 % beste leerlingen van het vijfde jaar minstens? Rond af op een geheel getal. Hoeveel procent van de zesdejaars haalde eveneens minstens dat resultaat?

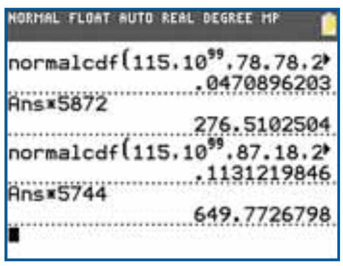
De 10 % beste leerlingen van het vijfde jaar hadden minstens 107 punten.

Ongeveer 19,4 % van de zesdejaars behaalden minstens 107 punten.



2 Leerlingen die 115 of meer haalden, mochten naar de tweede ronde. Aan de eerste ronde deden 5872 vijfdejaars- en 5744 zesdejaarsleerlingen mee. Maak een schatting van het totale aantal leerlingen dat naar de tweede ronde ging.

Ongeveer 277 vijfdejaars en ongeveer 650 zesdejaars, dus in totaal ongeveer 927 leerlingen, mochten door naar de tweede ronde.



Opdracht 37 bladzijde 92

In een compostbedrijf worden zakken potgrond machinaal gevuld. De machine kan men op een gemiddelde vulmassa instellen. Bij iedere instelling zijn de vulmassa's normaal verdeeld met een standaardafwijking van 0,3 kg. De machine wordt zo ingesteld dat 4 % van de zakken minder dan 11,0 kg weegt.

Bereken op welke gemiddelde vulmassa de machine ingesteld is.

De z-score die overeenkomt met 4 % is –1,75.

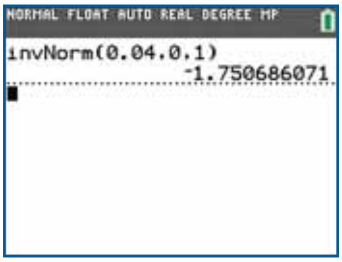
Ze komt overeen met de x-waarde 11.

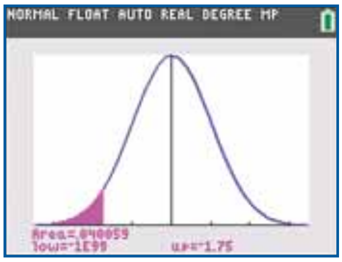
$$-1{,}75 = \frac{11-\mu}{0{,}3}$$

$$\Downarrow$$

$$\mu = 11 + 1{,}75 \cdot 0{,}3 = 11{,}525$$

De machine is ingesteld op een gemiddelde vulmassa van 11,5 kg.





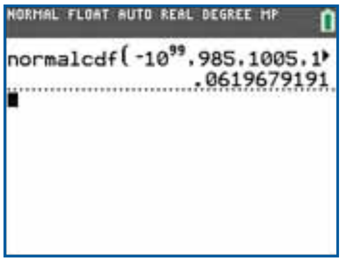
Opdracht 38 bladzijde 92

Op pakken suiker staat ' 1 kg ℮'. Die ℮ komt van 'estimate'; een EU-norm vereist dan dat hoogstens 2 % van de pakken meer dan 15 gram minder weegt dan 1 kg.

1 Stel dat een fabrikant een vulmachine heeft, die pakken suiker vult volgens een Norm(1005, 13)-verdeling. Hoeveel procent van de pakken zal minder dan 985 g wegen? Is hij in orde met de EU-norm?

Ongeveer 6,2 % van de pakken suiker weegt minder dan 985 g.

Hij is dus niet in orde met de EU-norm.



2 De standaardafwijking van een vulmachine kan niet gewijzigd worden, maar het gemiddelde kan dat wel. Op welk gemiddelde moet de fabrikant zijn vulmachine met $\sigma = 13$ g instellen, opdat hij aan de EU-norm zou voldoen? Onderbouw je redenering met een schets.

De z-score die overeenkomt met 0,02 is –2,054.

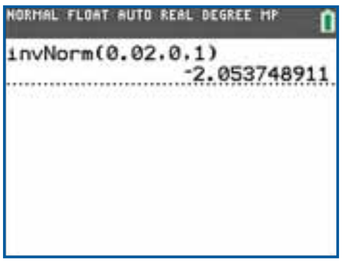
Deze moet overeen komen met de x-waarde 985.

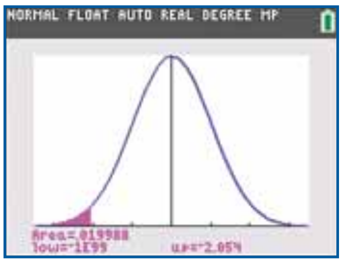
$$-2{,}054 = \frac{985-\mu}{13}$$

$$\Downarrow$$

$$\mu = 985 + 2{,}054 \cdot 13 = 1011{,}702$$

Hij moet het gemiddelde instellen op 1012 g.





3 Een kleinere standaardafwijking betekent nauwkeuriger vullen en dus minder verlies voor de fabrikant. Stel dat hij zijn toestel wil afstellen op $\mu = 1000$, wat moet dan de standaardafwijking zijn, opdat hij aan de EU-norm zou voldoen?

Zelfde redenering:

$$-2{,}054 = \frac{985 - 1000}{\sigma} \Rightarrow \sigma = \frac{-15}{-2{,}054} \approx 7{,}3$$

Afgerond tot op een gram: de standaardafwijking is 7 g.

Opdracht 39 bladzijde 93

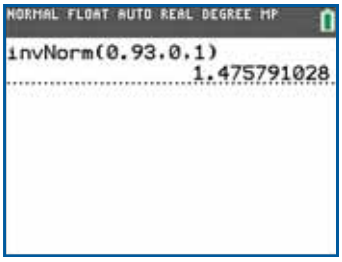
Op een autosnelweg staat een flitspaal. De maximumsnelheid is er 120 km/h. Uit de snelheidscontroles blijkt dat 7 % van de automobilisten er meer dan 130 km/h rijdt. Verder blijkt dat de snelheid normaal verdeeld is met gemiddelde $\mu = 113$ km/h.

1 Wat is de standaardafwijking van deze verdeling?

De z-score die overeenkomt met 93 % is 1,476.

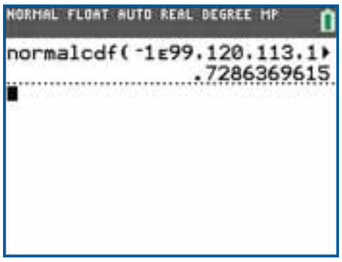
$$1{,}476 = \frac{130 - 113}{\sigma} \Rightarrow \sigma = \frac{17}{1{,}476} \approx 11{,}5$$

De standaardafwijking is 11,5 km/h.



2 Hoeveel procent van de automobilisten houdt zich aan de maximumsnelheid van 120 km/h?

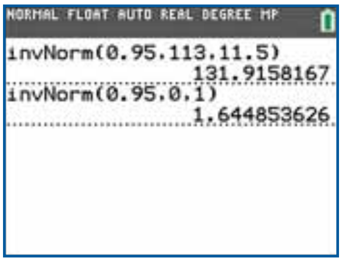
Ongeveer 73 % van de automobilisten houdt zich aan de maximumsnelheid van 120 km/h.



3 Hoe snel rijden de 5 % snelste wagens? Wat is hun *z*-score dan minstens?

De 5 % snelste wagens rijden minstens 132 km/h.

De overeenkomstige z-score is 1,64.



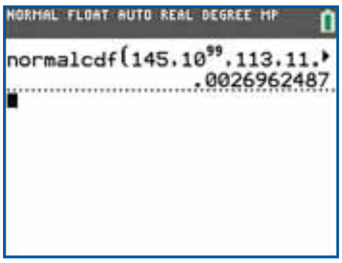
4 Op een dag zijn er op één uur tijd 16 voertuigen geflitst met een snelheid van meer dan 145 km/h. Hoeveel auto's zijn er in die tijdsspanne ongeveer voorbij de flitspaal gepasseerd?

Ongeveer 0,27 % van de wagens reed meer dan 145 km/h.

$$0{,}0027 \cdot x = 16 \Rightarrow x = \frac{16}{0{,}0027} \approx 5926$$

$$0{,}002696 \cdot x = 16 \Rightarrow x = \frac{16}{0{,}002696} \approx 5935$$

Er zijn ongeveer 5930 auto's gepasseerd.

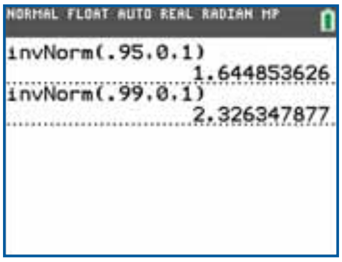


Opdracht 40 bladzijde 93

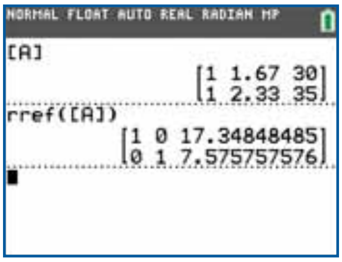
Amin moet elke weekdag de trein van 7.45 uur halen. De tijd die hij nodig heeft om van zijn huis naar het station te fietsen kan goed gemodelleerd worden door een normale verdeling. Indien hij thuis om 7.15 uur vertrekt, is hij één keer op twintig dagen te laat aan het station. Vertrekt hij om 7.10 uur, dan is hij één keer op honderd dagen te laat.

Stel dat hij per jaar 240 dagen naar zijn werk moet. Hoe laat moet hij thuis vertrekken om hoogstens één keer per jaar de trein te missen?

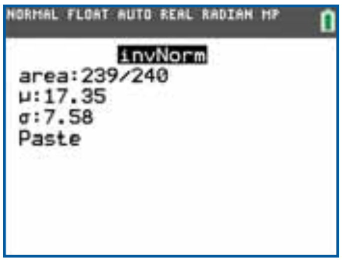
De z-score die overeenkomt met een fietstijd van 30 minuten is 1,64; deze bij een tijd van 35 minuten is 2,33.



De parameters μ en σ van de normale verdeling voldoen dus aan de vergelijkingen $\mu + 1{,}64\sigma = 30$ en $\mu + 2{,}33\sigma = 35$. Dit stelsel oplossen geeft: $\mu = 17{,}35$ en $\sigma = 7{,}58$.



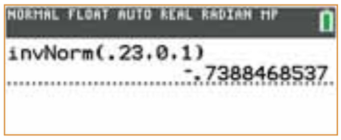
Hiermee kunnen we de fietstijd bepalen die 239 van de 240 dagen volstaat. Dit blijkt ongeveer 37 minuten te zijn (iets meer). Hij moet dus ten laatste om 7.08u vertrekken.



Opdracht 41 bladzijde 93

Percentielen van een normale verdeling

Het 23e percentiel van de standaardnormale verdeling is die z-waarde waarvoor geldt dat de oppervlakte links van z gelijk is aan 23 % of dus 0,23. M.b.v. een rekentoestel vinden we $z = -0{,}74$.

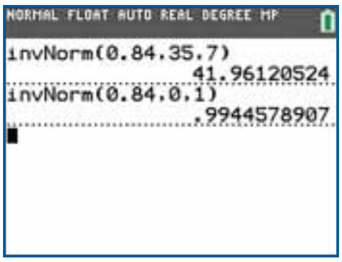


Algemeen

Het p-de percentiel van een normale verdeling Norm(μ, σ) is die x-waarde x_p waarvoor geldt dat de oppervlakte onder de verdeling links van x_p gelijk is aan p %. Percentiel komt van een verdeling in honderdsten (per 100, of 'per cent').

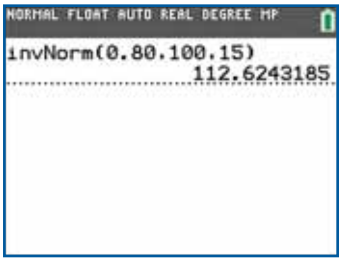
1 Bepaal het 84e percentiel van de verdeling Norm(35, 7). Welke z-waarde komt hiermee overeen?

Het 84e percentiel is ongeveer 42 en de bijbehorende z-score is ongeveer 0,99.



2 Linh behoort tot het 80e percentiel wat haar IQ betreft. De IQ-score is normaal verdeeld met $\mu = 100$ en $\sigma = 15$. Wat is haar IQ?

Haar IQ is ongeveer 113.



3 Bepaal zonder grafisch rekentoestel de volgende percentielen van Norm(μ, σ): 16e, 50e, 84e.

Op basis van de 68-95-99,7-regel:

16e percentiel: $\mu - \sigma$

50e percentiel: μ

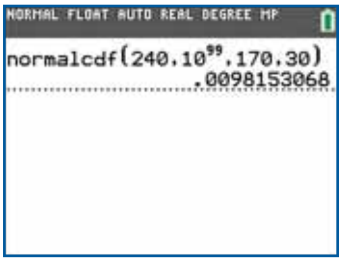
84e percentiel: $\mu + \sigma$

Opdracht 42 bladzijde 94

Het cholesterolgehalte bij mensen van dezelfde leeftijd is nagenoeg normaal verdeeld. Voor 14-jarige jongens is $\mu = 170$ milligram per deciliter bloed (mg/dl) met een standaardafwijking $\sigma = 30$ mg/dl. Cholesterolgehaltes boven de 240 mg/dl kunnen het best medisch worden opgevolgd.

Je neemt lukraak een jongen van 14 uit de populatie. Wat is de kans dat hij een cholesterolgehalte van 240 mg/dl of meer zal hebben?

De kans is ongeveer 0,0098.

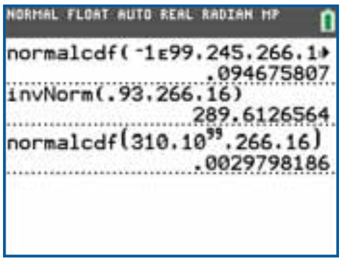


Opdracht 43 bladzijde 94

De duur van een voldragen zwangerschap bij de vrouw is normaal verdeeld met gemiddelde 266 dagen en met standaardafwijking 16 dagen.

1 Gebeurt de geboorte minstens 3 weken vroeger dan het gemiddelde, dan spreekt men van een 'premature' geboorte. Wat is de kans hierop?

De kans dat een zwangerschap hoogstens 245 dagen duurt, is 0,0947.



2 Laat de geboorte te lang op zich wachten, dan wordt ze ingeleid. Dit gebeurt in 7 % van de gevallen. Vanaf hoeveel dagen wordt de geboorte dus ingeleid?

Vanaf 290 dagen wordt ze ingeleid.

3 In 1973 beweerde een vrouw dat ze 310 dagen zwanger was geweest: op de dag van de bevalling was haar man immers al 310 dagen als marinier van huis. Wat is de kans dat een zwangerschap minstens 310 dagen duurt?

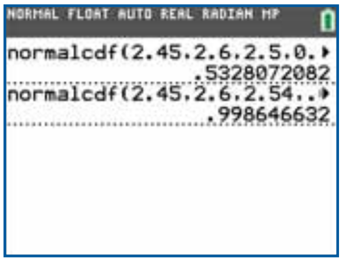
Die kans is 0,0030.

Opdracht 44 bladzijde 94

Twee toestellen snijden kurken voor wijnflessen. De eerste snijdt kurken met een diameter die normaal verdeeld is met een gemiddelde van 2,50 cm en een standaardafwijking van 0,10 cm. Ook bij de tweede is de diameter normaal verdeeld, met gemiddelde 2,54 cm en standaardafwijking 0,02 cm.

Enkel kurken met een diameter tussen 2,45 en 2,60 cm worden aanvaard. Welk toestel heeft de grootste kans om goede kurken te fabriceren?

Het tweede toestel heeft de grootste kans om een goede kurk te maken.

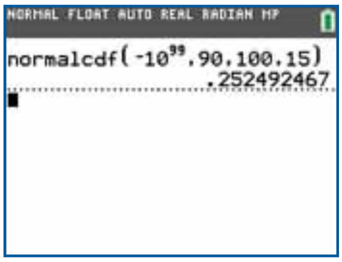


Opdracht 45 bladzijde 94

De *Wechsler Intelligence Scale for Children* is een intelligentietest voor kinderen tussen 6 en 16 jaar, die zo is gestandaardiseerd dat voor een bepaalde leeftijd het gemiddelde kind een IQ heeft van 100, met een standaardafwijking van 15. IQ-scores zijn normaal verdeeld.
Je kiest lukraak drie kinderen tussen 6 en 16.

1 Wat is de kans dat er precies één een IQ van 90 of minder heeft?

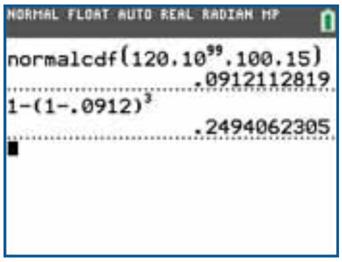
De kans dat een lukraak gekozen kind een IQ van 90 of minder heeft, is 0,2525. Met behulp van een kansboom kan snel nageteld worden dat er drie takken zijn waarin één kind een IQ van 90 of minder heeft (kans 0,2525) en de andere twee een IQ van meer dan 90 (kans 0,7475). De kans op precies één kind is $3 \cdot 0{,}2525 \cdot 0{,}7475^2 = 0{,}4233$.



(Deze vraag kan ook met behulp van de binomiale kansverdeling beantwoord worden, indien die al werd behandeld.)

2 Wat is de kans dat er minstens één een IQ van 120 of meer heeft?

De kans op een IQ van 120 of meer is 0,0912. Minstens één kind in die situatie is het complement van geen enkel kind in die situatie, met andere woorden drie kinderen met een IQ van minder dan 120 (voor een willekeurig kind is die kans $1 - 0{,}0921 = 0{,}9088$). Met behulp van de complementregel vinden we de gevraagde kans als $1 - 0{,}9088^3 = 0{,}2494$.

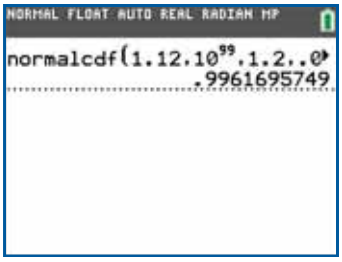


Opdracht 46 bladzijde 95

De diameter van cilindervormige staven is normaal verdeeld met gemiddelde 1,20 cm en standaardafwijking 0,03 cm. Is de diameter groter dan of gelijk aan 1,12 cm, dan is de kans dat de staven breken 0,001. Is de diameter kleiner dan 1,12 cm, dan is de kans op breken 0,1.

1 Bereken de kans dat een willekeurig gekozen staaf zal breken.

De kans dat zo'n staaf een diameter van minstens 1,12 cm heeft, is 0,9962; de kans dat de diameter kleiner is, is $1 - 0{,}9962 = 0{,}0038$



De gevraagde kans kan snel met behulp van een kansboom berekend worden. Door middel van kanswetten gebeurt de berekening als volgt:

P(staaf breekt)

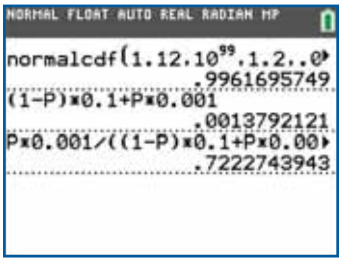
$$\begin{aligned}
&= P\big((\text{diam.} < 1{,}12 \text{ en staaf breekt}) \text{ of } (\text{diam.} > 1{,}12 \text{ en staaf breekt})\big)\\
&= P(\text{diam.} < 1{,}12 \text{ en staaf breekt}) + P(\text{diam.} > 1{,}12 \text{ en staaf breekt})\\
&= P(\text{diam.} < 1{,}12) \cdot P(\text{staaf breekt} \mid \text{diam.} < 1{,}12)\\
&\quad + P(\text{diam.} > 1{,}12) \cdot P(\text{staaf breekt} \mid \text{diam.} > 1{,}12)\\
&= 0{,}0038 \cdot 0{,}1 + 0{,}9962 \cdot 0{,}001\\
&= 0{,}0014
\end{aligned}$$

2 Stel dat een staaf gebroken is. Gebruik de regel van Bayes om de kans te berekenen dat de staaf een diameter groter dan 1,12 cm had.

De regel van Bayes geeft:

$$\begin{aligned}
P(\text{diam.} > 1{,}12 \mid \text{staaf is gebroken}) &= \frac{P(\text{diam.} > 1{,}12 \text{ en staaf is gebroken})}{P(\text{staaf is gebroken})}\\
&= \frac{0{,}9962 \cdot 0{,}001}{0{,}0014}\\
&= 0{,}7116
\end{aligned}$$

Ten gevolge van afrondingen in de tussenresultaten, en vooral door het kleine aantal beduidende cijfers in de noemer op de voorlaatste lijn, ontstaan afrondingen. Rekenen we voor alle kansen met het maximale aantal cijfers na de komma, dan vinden we als kans 0,7223.



Opdracht 47 bladzijde 96

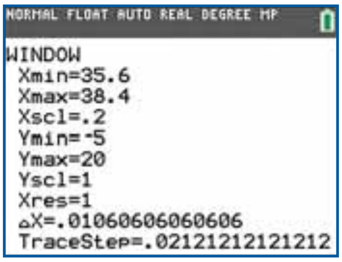
In 1992 verscheen in het Journal of the American Medical Association een onderzoek naar de lichaamstemperatuur bij 65 mannen en 65 vrouwen.

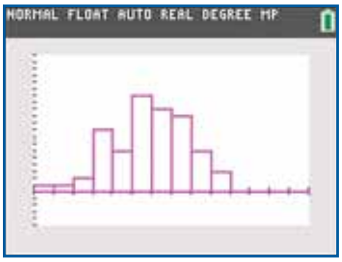
Temperatuur (°C) bij 65 mannen				
35,7	36,4	36,7	36,9	37,1
35,9	36,4	36,7	36,9	37,1
36,1	36,4	36,7	36,9	37,1
36,1	36,4	36,7	36,9	37,1
36,2	36,4	36,7	36,9	37,2
36,2	36,5	36,7	36,9	37,2
36,2	36,6	36,7	37,0	37,2
36,2	36,6	36,8	37,0	37,2
36,3	36,6	36,8	37,0	37,3
36,3	36,6	36,8	37,0	37,3
36,3	36,6	36,8	37,0	37,4
36,3	36,6	36,8	37,0	37,4
36,3	36,7	36,8	37,1	37,5

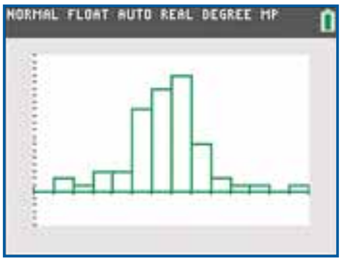
Temperatuur (°C) bij 65 vrouwen				
35,8	36,6	36,8	37,0	37,1
35,9	36,6	36,8	37,1	37,2
36,0	36,7	36,8	37,1	37,2
36,2	36,7	36,8	37,1	37,2
36,2	36,7	36,9	37,1	37,3
36,3	36,7	36,9	37,1	37,3
36,4	36,7	36,9	37,1	37,3
36,5	36,7	36,9	37,1	37,3
36,5	36,8	36,9	37,1	37,4
36,6	36,8	36,9	37,1	37,4
36,6	36,8	37,0	37,1	37,7
36,6	36,8	37,0	37,1	37,8
36,6	36,8	37,0	37,1	38,2

1 Geef een beschrijving van de verdeling van de temperaturen bij de mannen en de vrouwen. Zijn de gegevens normaal verdeeld?

We lezen de gegevens in via het programma.



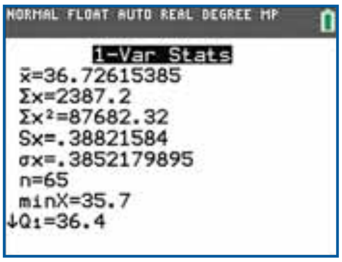


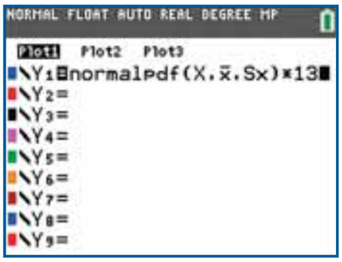


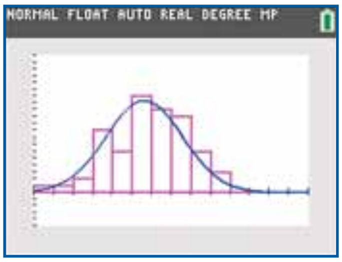
Geen van beide verdelingen zijn echt duidelijk symmetrisch, bij de verdeling van de temperatuur van de vrouwen zijn er ook uitschieters.

We plotten de normale kromme bij elk histogram. Omdat de oppervlakte van het histogram $65 \cdot 0{,}2 = 13$ is, moeten we de waarden van de normale kromme vermenigvuldigen met 13. Zie ook opdracht 9 voor het werken met een lijst van frequenties.

Mannen:

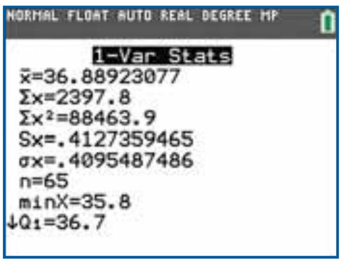


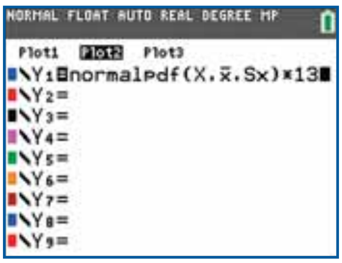


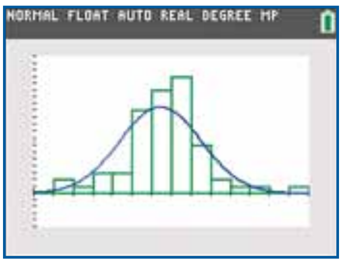


Hier blijkt dat de gegevens min of meer normaal verdeeld zijn.

Vrouwen:





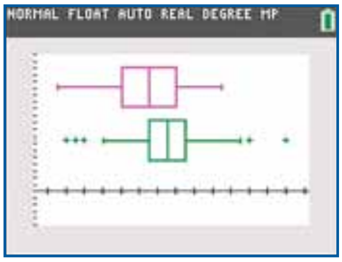


De normale verdeling is voor deze gegevens een minder goed model.

2 Welke verschillen stel je vast tussen de temperaturen bij de mannen en de vrouwen?

Uit de histogrammen en uit de centrum- en spreidingsmaten volgt dat de lichaamstemperaturen van de vrouwen gemiddeld iets hoger zijn. Bovendien lijken er bij de vrouwen enkele uitschieters te zijn.

Eventueel kan hier van de gelegenheid gebruik gemaakt worden om beide datasets samen voor te stellen als boxplots met uitschieters, zoals vaak werd gedaan in de 2e graad.



3 Bevestigen de gecombineerde gegevens dat de gemiddelde lichaamstemperatuur bij de mens 37 °C is?

Omdat beide verdelingen evenveel gegevens bevatten, zal het rekenkundig gemiddelde van beide gemiddelden de gemiddelde lichaamstemperatuur zijn van de 130 mannen en vrouwen samen.

$$\frac{36{,}7 + 36{,}9}{2} = 36{,}8$$

De gemiddelde lichaamstemperatuur bij de mens is inderdaad ongeveer 37 °C.

4 Schat met behulp van de 68-95-99,7-regel vanaf welke temperatuur een vrouw een 'uitzonderlijke' temperatuur heeft.

Een uitzonderlijke lichaamstemperatuur ligt twee standaardafwijkingen boven of onder het gemiddelde. Bij de vrouwen is dit dus voor waarden kleiner dan $(36{,}9 - 2 \cdot 0{,}4)$ °C = 36,1 °C of groter dan $(36{,}9 + 2 \cdot 0{,}4)$ °C = 37,7 °C.

Opdracht 48 bladzijde 97

Het gewicht van de Aarde

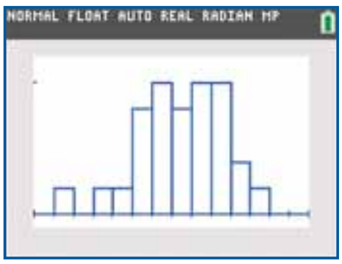
In 1797 begon Henry Cavendish (1731-1810) aan een reeks ambitieuze experimenten om "de wereld te wegen". In 1798 had hij 29 zorgvuldige herhalingen van zijn experiment uitgevoerd. De cijfers hieronder drukken de dichtheid van de Aarde uit als veelvouden van de dichtheid van water.

5,50	5,57	5,42	5,61	5,53	5,47	4,88	5,62	5,63	5,07
5,26	5,44	5,46	5,55	5,34	5,30	5,36	5,79	5,75	5,29
5,58	5,27	5,85	5,65	5,39	5,68	5,34	5,10	5,29	

1 Maak een histogram om op een eenvoudige manier na te gaan of de gegevens normaal verdeeld zijn. Zijn er uitschieters?

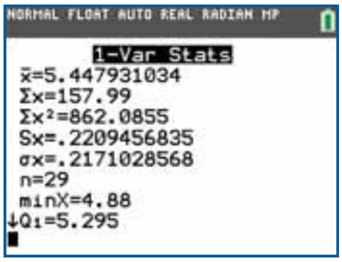
Uit een histogram blijkt dat de gegevens vrij symmetrisch verdeeld zijn. Hoewel links één waarde wat verder van de andere ligt, is de afstand niet groot genoeg om die als een uitschieter te beschouwen.

Hoewel dat op basis van een kleine dataset altijd moeilijker is om uit te maken, lijkt een normale verdeling hier een goed model te zijn.



2 Het gemiddelde van de bovenstaande waarden is een preciezere schatting van de relatieve dichtheid van de Aarde. Waaraan is die gelijk? Wat is het effect van eventuele uitschieters?

Het gemiddelde is, afgerond op drie cijfers na de komma, gelijk aan 5,448.



3 Als je weet dat de Aarde een diameter van ongeveer 12 753 km heeft, bereken dan de massa van de Aarde op basis van Cavendish' resultaat. Doe dit met en zonder eventuele uitschieters. Vergelijk met de recentere schatting van $5{,}972 \cdot 10^{24}$ kg

Het volume van de aarde is $V = \frac{4}{3}\pi r^3$ met r = 6 376 500 m. Uit het gemiddelde van de berekeningen van Cavendish volgt dat de dichtheid ρ van de aarde ongeveer 5448 $\frac{\text{kg}}{\text{m}^3}$ is. De massa m is dan gegeven door $m = V \cdot \rho$.

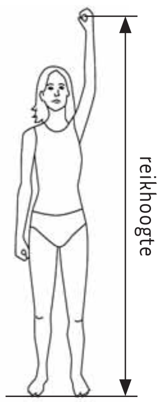
Invullen van alle waarden geeft: $m = 5{,}917 \cdot 10^{24}$ kg.

Deze waarde wijkt minder dan één procent af van de recentere schatting!

Opdracht 49 bladzijde 97

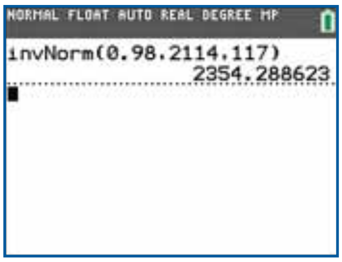
Een ontwerp moet niet alleen mooi, maar ook functioneel zijn. Bij veel ontwerpen wordt daarom rekening gehouden met de maten van het menselijk lichaam. Ontwerpers maken daarom vaak gebruik van antropometrietabellen. Dit zijn tabellen waarin het gemiddelde en de standaardafwijking van allerlei afmetingen van het menselijk lichaam staan. Al deze lichaamsmaten zijn (bij benadering) normaal verdeeld.

1 Om te zorgen dat een kamer als comfortabel ervaren wordt, moet de hoogte ervan minimaal gelijk zijn aan de reikhoogte (zie figuur).
Bij de bouw van een nieuwe studentenflat wil men dat de kamers door minstens 98 % van de studenten als comfortabel ervaren worden. De reikhoogte van Nederlandse studenten is gemiddeld 2114 mm met een standaardafwijking van 117 mm.



Bereken hoe hoog men de kamers minimaal moet maken.

De kamers moeten minimaal 2354 mm hoog zijn.

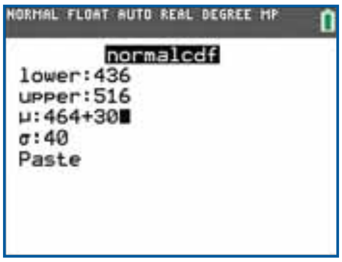


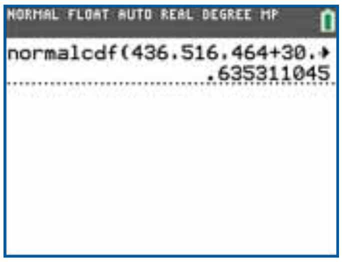
2 Ook bij het inrichten van een optimale werkplek houdt men rekening met lichaamsmaten. Een bureaustoel heeft precies de goede zithoogte als de zithoogte gelijk is aan de knieholtehoogte van een persoon plus 30 mm voor de schoenzool.
Van een bureaustoel is de zithoogte verstelbaar van 436 tot 516 mm. De knieholtehoogte is gemiddeld 464 mm met een standaardafwijking van 40 mm.

Bereken voor hoeveel procent van de mensen deze stoel op precies de goede zithoogte ingesteld kan worden.

(Bron: Examen VWO, Wiskunde C, Nederland, 2010)

Voor ongeveer 63,5 % van de mensen kan de stoel goed ingesteld worden.





Opdracht 50 bladzijde 98

Onderwijsonderzoekers nemen bij een grote groep kinderen in het derde en het zesde leerjaar een test af om hun leesvaardigheid te meten. De gemiddelde score voor de derdejaars is 75 en de standaardafwijking is 10. Bij de zesdejaars is het gemiddelde 82 en de standaardafwijking 6.

Samira zit in de derde klas en haalde 88 voor de test. Jente zit in de zesde klas en haalde ook 88. Wie van hen heeft de beste score binnen haar klas?

$$z_{\text{Samira}} = \frac{88-75}{10} = 1{,}3$$

$$z_{\text{Jente}} = \frac{88-82}{6} = 1$$

Samira heeft de beste score binnen haar klas.

Opdracht 51 bladzijde 98

Glaucoom is een oogziekte die in de meeste gevallen gepaard gaat met een te hoge druk binnen het oog. Door die verhoogde oogdruk kan beschadiging van de oogzenuw optreden, met als gevolg uitschakeling van een deel van het gezichtsveld.

Rechts hoe een persoon met glaucoom het landschap links zou kunnen waarnemen.

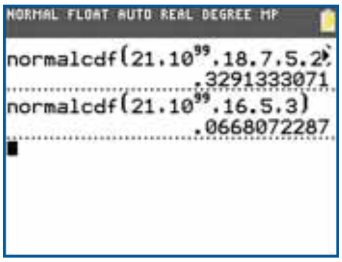
Uit verschillende studies blijkt dat de oogdruk, uitgedrukt in mmHg, ongeveer normaal verdeeld is. Er is sprake van een verhoogde oogdruk vanaf 21 mmHg.

1 Op het Caraïbische eiland Barbados bleek eind jaren 90 de prevalentie (dat is het aantal gevallen per duizend of per honderdduizend) van glaucoom bij de zwarte bevolking vrij hoog te zijn. Uit een lokale studie bleek bij zwarten de gemiddelde oogdruk 18,7 mmHg te zijn, met een standaardafwijking van 5,2 mmHg; bij blanken was dat 16,5 respectievelijk 3,0 mmHg.

Bereken voor beide groepen benaderend hoeveel procent een verhoogde oogdruk had.

Zwarten: ongeveer 33 %

Blanken: ongeveer 7 %



2 Uit een studie in Noord-Italië bleek dat de oogdruk binnen de bestudeerde groep gemiddeld ongeveer 15 mmHg bedroeg. De diagnose van glaucoom vereiste onder andere dat de personen een oogdruk van minstens 22 mmHg hadden. In de studie bleek 2,1 % van de personen daaraan te voldoen.

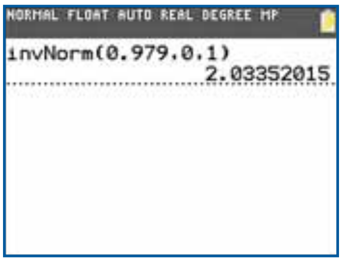
Wat was de standaardafwijking van de oogdruk?

De z-score die overeenkomt met 97,9 % (100 % – 2,1 %) is 2,0335.

Ze komt overeen met de x-waarde 22.

$$2{,}0335 = \frac{22-15}{\sigma} \Rightarrow \sigma = \frac{7}{2{,}0335} \approx 3{,}442$$

De standaardafwijking van de oogdruk is ongeveer 3,4 mmHg.



Opdracht 52 bladzijde 99

Een transportonderneming brengt elke dag over een vast traject verse vlaaien van Limburg naar Twente. De tijd die daarvoor nodig is, is normaal verdeeld met een gemiddelde van 2,5 u en een standaardafwijking van een kwartier. De vlaaien moeten om half negen geleverd zijn.

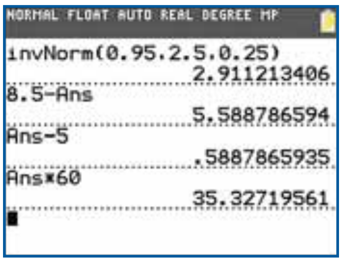
Enerzijds wil de directeur de loonkosten van de chauffeur beperken door hem niet te vroeg te laten vertrekken. Anderzijds kan de directeur zich niet permitteren om op meer dan 5 % van de dagen de vlaaien te laat af te leveren.

Bereken, in minuten nauwkeurig, hoe laat de chauffeur moet vertrekken.

(Bron: Examen VWO, Wiskunde B1, Nederland, 2003)

De maximum tijd voor het leveren van de vlaaien is 2,911 uren.

Vermits de vlaaien ten laatste om half negen moeten geleverd zijn, moet de chauffeur om 5.35 u vetrekken.



Opdracht 53 bladzijde 99

De gemiddelde leeftijd van de leerkrachten wiskunde van onze school is 44 jaar. Er zijn 12 leerkrachten wiskunde inclusief de coördinator. Laten we die laatste buiten beschouwing, dan stijgt die gemiddelde leeftijd met 1 jaar. Hoe oud is de coördinator?

A 32 **B** 33 **C** 44 **D** 55 **E** 56

(Bron © VWO tweede ronde, 2012)

Stel x de leeftijd van de coördinator.

De totale leeftijd van de 12 leerkrachten samen is $44 \cdot 12$. Verminderen we die totale leeftijd met x, dan vinden we de totale leeftijd van de 11 leerkrachten, zonder de coördinator, en die moet gelijk zijn aan $43 \cdot 11$. In symbolen: $44 \cdot 12 - x = 43 \cdot 11$.

Daaruit kan snel berekend worden dat $x = 528 - 473 = 55$.

Opdracht 54 bladzijde 99

Het gemiddelde $\overline{x}$ van 11 gegevens is 8 en de standaardafwijking s is 2. Minstens één van de gegevens heeft als waarde 8. Je verwijdert een gegeven met waarde 8 (en vervangt het niet door een ander gegeven).

De standaardafwijking zal dan

A afnemen;

B gelijk blijven;

C toenemen;

D er is te weinig informatie om hierover iets te besluiten.

Beargumenteer je antwoord aan de hand van de betekenis of de formule van de standaardafwijking.

De standaardafwijking zal toenemen.

De standaardafwijking is immers een maat voor de spreiding van de gegevens rond het gemiddelde. Doordat een centraal gelegen waarde wordt verwijderd, met andere woorden een waarde met kleine afwijking, is de gemiddelde afwijking van de overblijvende waarden iets groter geworden.

Door te redeneren op de formule van de standaardafwijking, kun je de toename als volgt inzien. In de teller verandert er niets door het verwijderen van die waarde 8 (de afwijking van de waarde 8 was immers nul) en de noemer neemt met 1 eenheid af. Dus zal de breuk groter worden.

Opdracht 55 bladzijde 100

Je hebt alle begrippen van dit en vorig hoofdstuk goed verwerkt indien je de onderstaande vragen kunt beantwoorden:

1 Wat is het voordeel of zijn de voordelen van het invoeren van dichtheidskrommen?

Een eerste voordeel is dat dichtheidskrommen het globale patroon van een verdeling weergeven, zonder rekening te houden met kleine onregelmatigheden of eventuele uitschieters (zie p. 53).

Verder kunnen bepaalde praktische problemen op basis van het functievoorschrift van de dichtheidsfuncties gemakkelijk (via software) opgelost worden: het berekenen van relatieve frequenties of kansen (als oppervlakte onder de dichtheidskromme) en van grenswaarden voor frequenties of kansen (zie paragraaf 2.3).

2 Hoe kun je louter grafisch μ en σ aflezen op een normale kromme?

μ is de x-coördinaat van de top en σ is de horizontale afstand van de symmetrieas tot het buigpunt (zie p. 58).

3 We ontmoetten op verschillende plaatsen het woord 'gemiddelde'. Wat is het verschil tussen het 'gemiddelde' $\overline{x}$ en het 'gemiddelde' μ? En wat is het verband?

Het gemiddelde μ is in eerste instantie louter een parameter in het voorschrift van de normale dichtheidsfunctie (zie p. 55), net zoals a een parameter is in het voorschrift van de functie $f : x \rightarrow ax^2 + bx + c$. A priori heeft μ dus niets te maken met gegevens, in tegenstelling tot het gemiddelde $\overline{x}$ dat het rekenkundig gemiddelde $\frac{1}{n}\sum_{i=1}^{n} x_i$ is van n gegevens $x_1, x_2, \dots, x_n$ (p. 44).

Het verband tussen beide zagen we als gevolg van de meetkundige betekenis van de getallen, in het geval de gegevens normaal verdeeld zijn. In dat geval is $x = \overline{x}$ immers de vergelijking van de rechte die min of meer als symmetrieas van de gegevens kan doorgaan (zie p. 47). De rechte met vergelijking $x = \mu$ is de symmetrieas van de normale verdeling. Voor normaal verdeelde gegevens ligt het dan voor de hand μ gelijk te stellen aan $\overline{x}$ (zie p. 56-58).

(In een meer gevorderde cursus kansrekening wordt het begrip rekenkundig gemiddelde uitgebreid voor continue verdelingen en kan aangetoond worden dat μ het gemiddelde is van de normale verdeling, volgens deze nieuwe definitie.)

4 Wat is een *z*-score? Geef minstens twee toepassingen van het gebruik van *z*-scores.

Zie p. 69 voor wat een z-score is. Opdracht 13 en het voorbeeld op p. 68 geven een eerste toepassing van z-scores: gegevens uit normale verdelingen vergelijken. De toepassing op p. 70 geeft een tweede toepassing: de parameters van een normale verdeling bepalen.