

4 De normale verdeling



Inhoudsopgave

4.0	Extreem weer	3
4.1	Vele kleintjes middelen uit	5
4.2	Wat is normaal?	7
4.3	Standaardiseren	15
4.4	Over continue verdelingen	23

Colofon

© 2011 cTWO

Experimentele uitgave Kansrekening en Statistiek, vwo, wiskunde A en C

versie 2 (aug 2012)

auteurs Leon van den Broek, Maris van Haandel

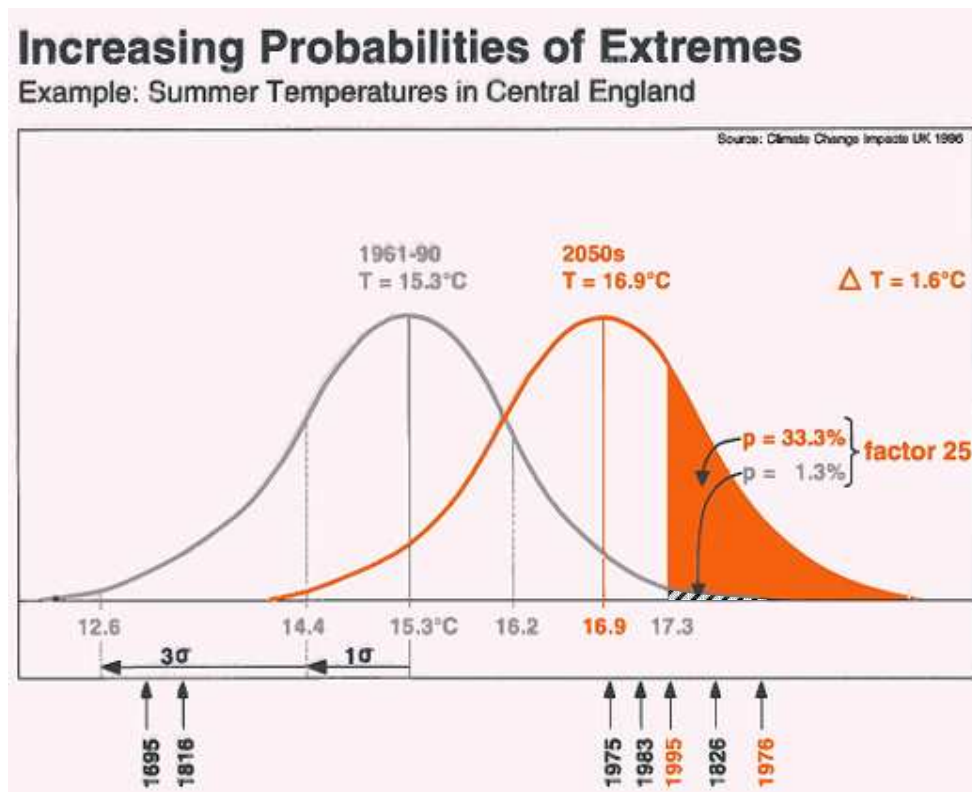
digiboek Carel van de Giessen

met medewerking van Simon Biesheuvel, Piet Versnel, Peter van Wijk

4.0 Extreem weer

- 1 Om te voorspellen hoe het klimaat verandert, is de gemiddelde temperatuurstijging van belang. Die veroorzaakt het smelten van ijsmassa's, zeespiegelstijging, verschuiving van vegetatiegordels (verwoestijning). In het bijzonder krijgen weersextremen aandacht. Denk aan langdurige perioden van extreme droogte (voedselschaarste) en extreme neerslag (overstromingen).

De figuur hieronder toont de statistische kansverdeling van de zomertemperaturen in Centraal Engeland (bron: Climatic Research Unit, University of Norwich). De meteorologische zomer bestaat uit de volle maanden juni, juli en augustus.



In de figuur zie je hoe een temperatuurstijging van 1,6 °C (resultierend in een verschuiving van de mediaan van de zomertemperatuur van 15,3 °C naar 16,9 °C) leidt tot een enorme toename van de kans op extreem hoge temperaturen.

In 1995 was de zomertemperatuur in Engeland gemiddeld 17,3 °C. Dat is extreem hoog.

- a. Wat is de kans op zo'n hoge of nog hogere zomertemperatuur
 - als de mediaan van de zomertemperatuur 15,3 °C is,
 - als de mediaan van de zomertemperatuur 16,9 °C is?

Misschien haal je je schouders op over een opwarming van slechts 1,6 °C; dat lijkt niet zo veel. Bovenstaand rapport brengt je misschien op andere gedachten. In de figuur is sprake van "factor 25".

b. Wat zegt die?

Dit model is wereldwijd toepasbaar en verklaart niet alleen dat extreem warme zomers veel vaker gaan voorkomen bij de “geringe” stijging van 1,6 °C, maar ook de veel kleinere kans op extreme kou, zodat bijvoorbeeld gunstige omstandigheden voor een Elfstedentocht aanzienlijk minder waarschijnlijk worden (*).

Het veel vaker voorkomen van extreem warme zomers heeft ernstige gevolgen. Droogten kunnen zorgen voor hongernood, tekort aan drinkwater en bosbranden. Ook neemt de schade die orkanen aanrichten kwadratisch toe met de kracht ervan, en die is afhankelijk van de temperatuur van zeewater.

De verdeling van de zomertemperaturen in Centraal Engeland is gemaakt op grond van historische data (van 1961 t/m 1990). De meeste waarden zitten in de buurt van de mediaan (het gemiddelde); grote afwijkingen komen weinig voor. De verdelingskromme is een zogenaamde klokkromme. Omdat hij zo vaak voorkomt, spreekt men van een “normale” verdeling. Daarover gaat dit hoofdstuk.

**c. Wat was de gemiddelde zomertemperatuur in Centraal Engeland tussen 1961 en 1990?
Wat is de standaardafwijking?**

Een vuistregel zegt hoeveel procent tussen het gemiddelde-min-de standaardafwijking en het gemiddelde-plus-de standaardafwijking ligt.

d. Hoeveel procent is dat? Als je de vuistregel niet meer weet, kun je het uit het percentage plaatje schatten.

Wat is extreem?

In het voorgaande werd de zomer van 1995 als extreem warm genoemd. Die van 1975 was ook heel warm, maar waarschijnlijk noem je die niet extreem warm. Daar zit iets subjectiefs in: niet iedereen zal dezelfde grenzen voor extreem hanteren. Een goede maatstaf hiervoor is de standaardafwijking.

2 Leg uit hoe je de standaardafwijking kunt gebruiken om te zeggen of een resultaat extreem is.


(*) Het zou ook kunnen zijn dat het Nederlandse klimaat verschuift naar een landklimaat. Dat heeft warmere zomers en strengere winters. Dan zou de kans op een Elfstedentocht juist toenemen!

4.1 Vele kleintjes middelen uit

Voorbeeld 1: Lichaamslengte

Mensen worden niet even lang. Voor een deel zit hem dat in de genen (aanleg), voor een deel ook in de voedingsgewoonten en voor een deel in de kwaliteit van de gezondheidszorg. Maar zelfs mensen die in eenzelfde gezin opgroeien worden niet allemaal even lang. Een toevallige blessure kan invloed hebben of een of andere ziekte of toevallige vrienden of een succesvolle sportcarrière. Dit zijn allemaal factoren die in het voordeel of in het nadeel van de uiteindelijke lengte werken en er zijn er ongetwijfeld nog veel meer. Sommige factoren zullen meer invloed hebben en andere minder. Eén belangrijke (genetische) factor is natuurlijk het geslacht, want mannen zijn nu eenmaal gemiddeld langer dan vrouwen. We richten ons in het volgende voorbeeld op de mannen; voor vrouwen geldt een soortgelijk verhaal.

We maken een (sterk vereenvoudigd) model van de te bereiken lengte van een volwassen man. We gaan uit van een basislengte van 180 cm en veronderstellen dat er 100 factoren meespelen, die allemaal positief (1), neutraal (0) of negatief (-1) kunnen zijn. Deze drie uitkomsten komen gemiddeld even vaak voor. Een positieve waarde verhoogt de lengte met 1 cm, een neutrale verandert de lengte niet en een negatieve waarde verlaagt de lengte met 1 cm. Een man met bijvoorbeeld 27 positieve, 35 neutrale en 38 negatieve factoren zal volgens dit model al met al 169 cm lang worden.

-  **3** Met de Random Generator (onder Simulaties) in VUStat kunnen we op deze manier snel 1000 lengtes maken (simuleren). Zet daarvoor *Model* op *Gehele getallen met teruglegging*, stel *Van* -1 tot 1, zet *Aantal getallen per keer* op 100 en *Aantal keer* op 1000.


Kies bij *Beeld* voor *Som* en vink de optie *Grafiek* aan. Start vervolgens de simulatie en zet deze op snel (anders duurt het erg lang). Bekijk het resultaat. Bedenk wel dat de waarnemingen met 180 moeten worden verhoogd om er lengtes van te maken.

Rechts vind je de knop \bar{x}, \dots , waaronder het gemiddelde, mediaan, minimum, maximum en standaardafwijking staan. Links ervan is een knop, waarmee je een frequentietabel kunt maken en erboven zijn knoppen om een histogram te maken.

- Noteer gemiddelde, mediaan, minimum, maximum en standaardafwijking.
- Maak in VUStat een frequentietabel met zes gelijke klassen, waarbij het minimum de uiterste linkergrens is en maximum de uiterste rechtergrens. (NB de klassenbreedte hoeft geen geheel getal te zijn.)
- Breid de frequentietabel in VUStat uit met procenten en neem deze over.


Voorbeeld 2: Een vulmachine


Een vulmachine maakt rollen beschuit. Elke rol bevat 13 beschuiten. De geproduceerde beschuiten zijn niet allemaal even zwaar: de gewichten variëren *uniform* van 9 tot 11 gram. Uniform wil zeggen dat alle gewichten tussen 9 en 11 gram even vaak voorkomen. Het percentage beschuiten met een gewicht tussen bijvoorbeeld 10,6 en 11,3 gram is 35%, want het interval van 10,6 tot 11,3 is 35% van het hele waardengebied. Met andere woorden: het percentage is evenredig met de breedte van het interval.


-  **4** Maak met de Random Generator een serie gewichten van 1000 van rollen beschuit. Verzamel daarbij dezelfde gegevens als bij de lengtes in opgave 3.


Voorbeeld 3: Een halve marathon


Een geofefende loper legt de halve marathon (21 km) af in een min of meer gelijkmatig tempo. Hij doet er 87,5 minuut over. Dit betekent dat hij over honderd meter gemiddeld 25 seconden doet. Maar hij loopt gedurende de halve marathon niet altijd precies even hard: zijn honderd-metertijden variëren uniform van 22 tot 28 seconden. Deze verschillen hangen af van toevallige factoren, dus het is niet zo dat zijn snelheid in de loop van de wedstrijd omlaag gaat vanwege vermoeidheid of zo.

-  **5** Maak met de Random Generator een serie van 1000 halve-marathontijden (in seconden). Verzamel daarbij weer dezelfde gegevens als bij de lengtes en de rollen beschuit.

-  **6** Vergelijk de gegevens van de simulaties in de drie voorbeelden. Iemand vindt dat ze op hetzelfde neerkomen. Verdedig dat standpunt.

-  **7** Bedenk een manier om met de Random Generator het totale aantal ogen bij 50 worpen met een zuivere dobbelsteen te simuleren en voer die simulatie uit.
 - a.** Hoeveel ogen gooi je gemiddeld met 50 dobbelstenen. Had je dat ook zonder simulatie kunnen voorspellen?
 - b.** Schat met behulp van het simulatieresultaat de kans op 190 of meer ogen in 50 worpen.
 - c.** Schat met behulp van het simulatieresultaat de kans dat het totaal aantal ogen in 50 worpen tussen 170 en 180 ligt (inclusief 170, zonder 180).

-  **8** Bedenk een manier om met de Random Generator het totale aantal keer kop bij 100 worpen met een zuivere munt te simuleren en voer die simulatie uit.
 - a.** Hoeveel keer kop verwacht je bij 100 worpen?
 - b.** Wat is de standaardafwijking van het aantal keer kop?
 - c.** Wat is het minimum en wat is het maximum aantal keer kop in het simulatieresultaat?
 - d.** Schat de kans dat het aantal keer kop meer dan 10% afwijkt van het verwachte aantal.

-  **9** Simuleer het totale aantal keer kop bij 400 worpen met een zuivere munt.
 - a.** Hoeveel keer kop verwacht je bij 400 worpen op grond van de simulatie?
 - b.** Wat is de standaardafwijking van het aantal keer kop?
 - c.** Wat is het minimum en wat is het maximum aantal keer kop in het simulatieresultaat?
 - d.** Schat de kans dat het aantal keer kop meer dan 10% afwijkt van het verwachte aantal.

- 10** Vergelijk de antwoorden van opgave 8 en 9 per onderdeel. Wat valt je op?

4.2 Wat is normaal?

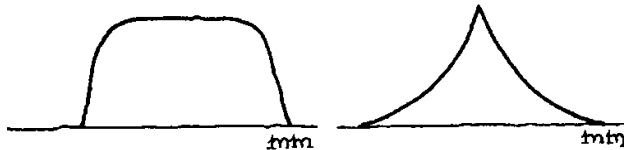
Klokvormig

- 11 Vanaf 1848 worden in Nederland systematisch allerlei gegevens over het weer bijgehouden. Gemiddeld valt er jaarlijks 780 mm neerslag in de Bilt. Grotere afwijkingen dan 380 mm van dit gemiddelde zijn nooit voorgekomen. Op grond van die jarenlange ervaring maakt men een plaatje van de kansverdeling van de neerslag voor het komend jaar.
- a. Hoe ziet dat plaatje eruit, denk je? Vermeld de relevante gegevens.

Stel dat er in een jaar 500 mm neerslag viel.

- b. Vind je dat extreem weinig of valt het wel mee? Waarom?

Hieronder staan twee mogelijke antwoorden op vraag a.



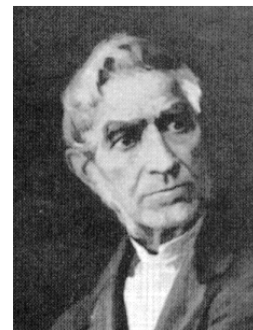
- c. Wat is je bezwaar tegen elk van deze antwoorden?

Over een klokvormige kromme

In de simulaties in de vorige paragraaf kwam je hetzelfde soort plaatje tegen: de zogenaamde **klokvorm**. Kenmerkend voor de klokvorm zijn:

- Symmetrie om het gemiddelde: afwijkingen naar boven zijn even waarschijnlijk als even grote afwijkingen naar beneden.
- Hoe groter de afwijking, des te kleiner is de kans dat die optreedt.
- Erg grote afwijkingen komen praktisch niet voor.

In veel voorbeelden in de natuur en bij het menselijk handelen komt de klokvorm voor (vaak bij benadering). Dat deze verdeling vaak voorkomt, is ontdekt in de negentiende eeuw. De belangrijkste onderzoeker was de Belg Quételet. In 1835 publiceerde hij een boek met statistisch materiaal over allerlei grootheden betreffende een mens (bijvoorbeeld de lengte van 18-jarige jongens). Hij merkte op dat de grootheden klokvormig verdeeld waren rond een gemiddelde. Een individuele afwijking van dat gemiddelde kwam door toevallige oorzaken. Hij voerde het idee van de “volmaakte” mens in: dat is de mens die alle grootheden gemiddeld heeft. Heel iets anders dan wat als ideaal gezien wordt!



Adolphe Quételet
1796 - 1874

- 12 Teken voor elk van de volgende voorbeelden een plaatje zoals hiernaast. Op de horizontale as wordt de genoemde grootheid uitgezet.



Schrijf bij de horizontale as de eenheid waarin je meet. (Bij het eerste voorbeeld is de grootheid *lengte* en is de eenheid *cm.*)

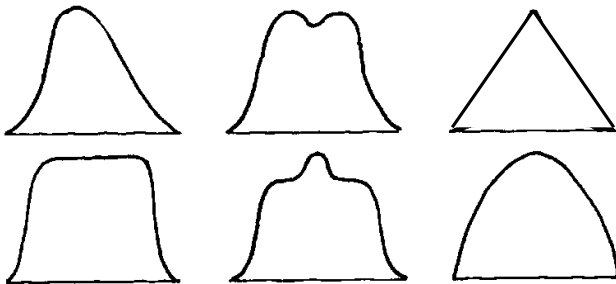
Schrijf bij de drie streepjes op de horizontale as getallen die redelijk kloppen met de werkelijkheid.

- Lengte van een Nederlandse jongen van 18 jaar.
- Leeftijd van een vrouw als ze moeder wordt (haar eerste kind krijgt).
- Tijdsduur van een autorit Arnhem-Nijmegen (18 km) in de ochtendspits.
- Het precieze gewicht in een zogenaamd kilopak suiker.
- Het aantal keer kop bij 100 worpen met een muntstuk.

Zoals gezegd, zijn veel verdelingen klokvormig, of ze lijken daar sterk op. We spreken wel van een **normale verdeling**. De term normale verdeling is ingevoerd door de Engelse statisticus Karl Pearson (1857-1936). Het plaatje bij opgave 12 staat model voor de normale verdeling.

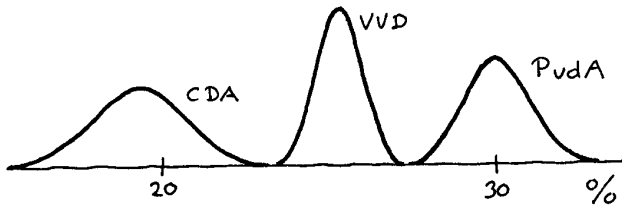
Maar niet alle verdelingen zijn normaal.

- 13 a. Geef zelf nog een paar praktijkvoorbeelden van (ongeveer) normale verdelingen.
 b. Geef zelf ook een paar voorbeelden waarbij de verdeling duidelijk niet normaal is.
- 14 Geen van de volgende verdelingen is normaal. Zeg van elke verdeling, waarom hij niet normaal is.



Zoals je gezien hebt, kun je uit een verdelingskromme kansen of percentages aflezen. Daarvoor is de *oppervlakte* onder de kromme bepalend.

- 15 Op 6 mei 1998 vonden er verkiezingen voor de Tweede Kamer plaats. De laatste dagen voor de verkiezingen werd de uitslag voorspeld. Er zijn twee grote bureaus die zich daarmee bezighouden: Inter/View en Nipo. Zij houden peilingen onder de Nederlandse bevolking. Op grond van die peilingen voorspellen de bureaus voor elke politieke partij een percentage van de stemmen. Maar dat voorspelde percentage klopt natuurlijk zelden precies: er is een onzekerheid. Welke percentages voorspeld werden voor de drie grote partijen, drie dagen voor de verkiezingen, kun je aflezen uit onderstaand plaatje. Bovendien kun je de erin zien in hoeverre de voorspellingen betrouwbaar zijn.

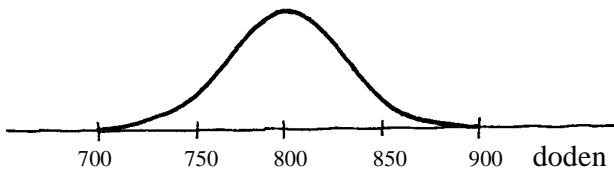


We bekijken het percentage voor de PvdA.

- Welk percentage is voorspeld voor de PvdA?
- Tussen welke twee grenzen ligt het percentage (met grote waarschijnlijkheid)?
- Bij welke partij is de onzekerheid van de voorspelling het grootst? Bij welke partij is de onzekerheid het kleinst? Waar zie je dat aan?
- De oppervlakte onder elk van de drie krommen is hetzelfde. Waarom moet dat zo zijn?
- Hoe groot ongeveer is volgens het plaatje de kans dat de PvdA meer dan 31% van de stemmen haalt?

16 Verkeersdoden

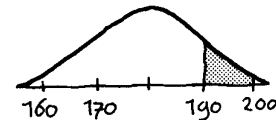
Het aantal verkeersdoden in een jaar in Nederland schommelt de laatste jaren rond de 800. Op grond van het verleden wordt het aantal verkeersdoden voor komend jaar voorspeld. De voorspelling en de onzekerheid van de voorspelling kun je aflezen uit het volgende plaatje.



- Schat hoe groot de kans is dat het aantal verkeersdoden onder de 750 ligt.
- Schat hoe groot de kans is dat het aantal verkeersdoden ligt tussen 800 en 830.

- 17 De lengte van 18-jarige jongens in Nederland is klokvormig verdeeld. Het percentage van de jongens die langer zijn dan 190 cm wordt gegeven door de grijze oppervlakte.

- Hoe groot schat jij dat dat percentage ongeveer is?
- Schat hoeveel procent van de jongens een lengte heeft tussen 170 en 180 cm.

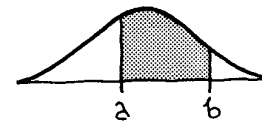


Een grootheid X is verdeeld volgens de kromme hiernaast. Op de horizontale as zijn twee mogelijke waarden van X aangegeven: a en b .

De kans dat de waarde van X ligt tussen a en b wordt gegeven door de oppervlakte onder de kromme tussen a en b .

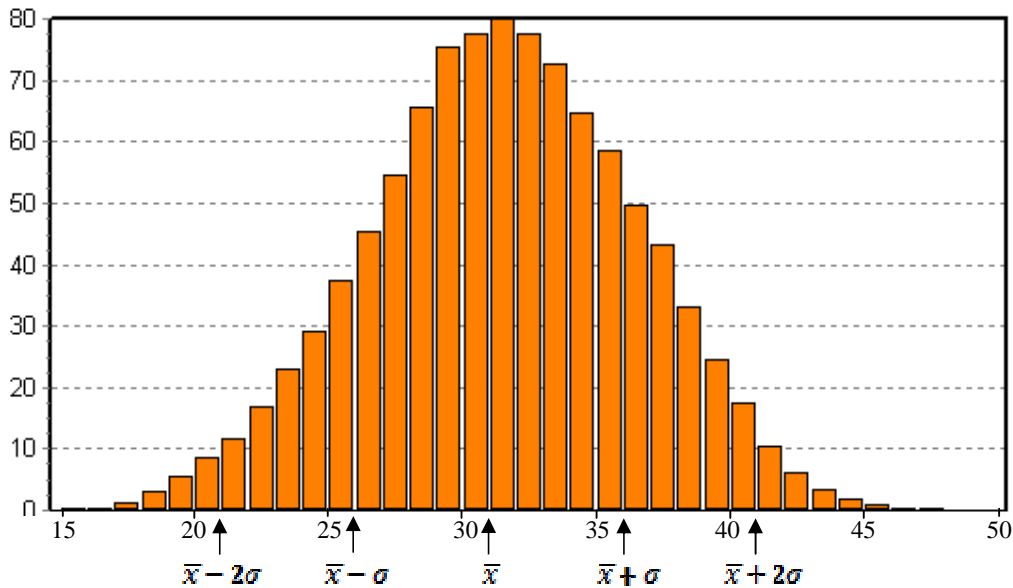
Preciezer: de volgende drie uitspraken komen op hetzelfde neer.

- De grijze oppervlakte is $p\%$ van de totale oppervlakte onder de kromme.
- Bij een groot aantal herhalingen zal X in ongeveer $p\%$ van de gevallen een waarde tussen a en b hebben,
- De kans dat $a \leq X \leq b$ is $\frac{p}{100}$.



- 18 In 2009 werden in Nederland in totaal 184.915 kinderen levend geboren. Hieronder staan een frequentietabel en een frequentiehistogram van de leeftijd van de moeders. De gegevens zijn ontleend aan het CBS StatLine. De verdeling is klokvormig.

leeftijd moeder	15	16	17	18	19	20	21	22	23	24	25	26
frequentie (‰)	0,2	0,4	1,2	3,1	5,6	8,5	11,8	16,8	23,0	29,1	37,3	45,4
leeftijd moeder	27	28	29	30	31	32	33	34	35	36	37	38
frequentie (‰)	54,7	65,6	75,3	77,6	80,0	77,5	72,5	64,7	58,5	49,8	43,2	33,0
leeftijd moeder	39	40	41	42	43	44	45	46	47	48	49	
frequentie (‰)	24,5	17,4	10,4	6,1	3,4	1,9	0,8	0,3	0,2	0,1	0,1	



Voer op de GR de leeftijden in.

- Controleer het histogram door het op de GR te tekenen.
- Bereken op de GR het gemiddelde en de standaardafwijking van de leeftijd van de moeders.

Als je als leeftijden 15, 16, ... invoert, krijg je als gemiddelde ongeveer 31,0 jaar. In werkelijkheid is het gemiddelde echter 31,5 jaar.

- Kun je dat halve jaar verschil verklaren?
- Ga na dat in het frequentiehistogram \bar{x} , $\bar{x} + \sigma$ en $\bar{x} - \sigma$ goed zijn aangegeven.
- Hoeveel procent van de moeders is ouder dan \bar{x} , hoeveel procent is ouder dan $\bar{x} + \sigma$ en hoeveel procent is ouder dan $\bar{x} + 2\sigma$?

Klokvormige verdelingen

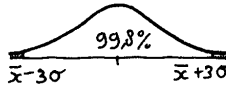
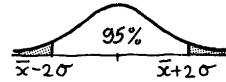
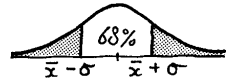
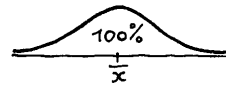
Veel data zijn zo ongeveer klokvormig verdeeld zoals hiernaast schetsmatig is aangegeven. Er is één top en de verdeling is symmetrisch en loopt geleidelijk af tot 0. Zie opgave 18, maar je kunt ook denken aan juli-temperaturen, lengte van rekruten, de eindexamencijfers voor het vak wiskunde A of het aantal jongens bij 1000 geboortes in een zekere gemeente. Bij zo'n verdeling zijn grote afwijkingen van het gemiddelde zeldzaam.



Vuistregels (herhaling)

Bij klokvormige verdelingen gelden de volgende vuistregels voor de afwijkingen van het gemiddelde:

- afwijkingen van meer dan σ zijn heel gewoon: dit gebeurt in ongeveer 32% van de gevallen,
- afwijkingen van meer dan 2σ zijn tamelijk zeldzaam: dit gebeurt in ongeveer 5% van de gevallen,
- afwijkingen van meer dan 3σ zijn uiterst zeldzaam: dit gebeurt in ongeveer 0,2% van de gevallen.



- 19 Kijk nog eens naar de plaatjes die je hebt getekend bij opgave 12. Schat bij elk van de voorbeelden de standaardafwijking met behulp van het plaatje en de derde vuistregel.
- Lengte van een Nederlandse jongen van 18 jaar.
 - Leeftijd van een vrouw als ze moeder wordt (haar eerste kind krijgt).
 - Tijdsduur van een autorit Arnhem-Nijmegen (18 km) in de ochtendspits.
 - Het precieze gewicht in een zogenaamd kilopak suiker.
 - Het aantal keer kop bij 100 worpen met een muntstuk.

Kansen bij de normale verdeling

Bij de volgende opgaven mag je er telkens van uitgaan dat er sprake is van een **normale** verdeling.

- 20 Op een weg binnen de bebouwde kom (waar 50 km/u de maximumsnelheid is) wordt vaak te hard gereden. Controles wijzen uit dat de gemiddelde snelheid 60 km/uur is en de bijbehorende standaardafwijking 5 km/uur.
- Hoeveel procent van de passerende voertuigen rijdt te hard?
 - Hoeveel procent van de passerende voertuigen rijdt tussen 55 en 70 km/uur?
- 21 Een melkkoe geeft in een lactatieperiode (dat is de periode waarin ze melk geeft) gemiddeld 25 liter melk per dag. Een koe wordt gemiddeld 315 dagen per jaar gemolken. Dit komt neer op een gemiddelde productie van 7875 liter melk per koe per jaar. Sommige koeien produceren tijdens hun leven wel 100.000 liter melk. Nederland telt ongeveer 4 miljoen melkkoeien.

De standaardafwijking van de dagelijkse hoeveelheid melk van melkkoe Bertha is 3 liter.

- Hoeveel dagen per jaar zal Bertha naar verwachting minder dan 19 liter melk geven?

De standaardafwijking van de jaarlijkse hoeveelheid melk van Nederlandse melkkoeien is 125 liter.

- Hoeveel melkkoeien in Nederland zullen naar verwachting meer dan 8000 liter melk per jaar geven?

- 22 Klas 4V2 (met 28 leerlingen) heeft bij Engels een schriftelijke overhoring gehad over een groot aantal woordjes. Gemiddeld had een leerling 17,5 fout. De standaardafwijking van het aantal fouten was 2,5. De toegepaste normering is één punt eraf per vijf fouten.
- Schat het laagste en het hoogste cijfer in deze klas.
 - Schat het aantal onvoldoendes ($\leq 5,5$) in deze klas.
- 23 Neem aan dat de leeftijd van leerlingen in vwo 5 aan het begin van een schooljaar (1 september) normaal verdeeld is met een gemiddelde van 16,3 jaar en een standaardafwijking van 0,8 jaar. (Hierbij rekenen we de leeftijd als voortdurend oplopende tijd sinds de geboorte, zodat iemand bijvoorbeeld 15,7 jaar kan zijn.)
- Is de leeftijd van deze leerlingen aan het eind van het schooljaar (op 1 juli) ook normaal verdeeld? Wat is dan het gemiddelde en wat is de standaardafwijking?
 - Is de leeftijd van deze leerlingen gerekend in maanden (op 1 september) ook normaal verdeeld? Wat is dan het gemiddelde en wat is de standaardafwijking (beide in maanden)?
- 24 De lengte van 21-jarige Engelse jongemannen is normaal verdeeld met een gemiddelde van 178 cm en een standaardafwijking van 7 cm. Hun gewicht is ook normaal verdeeld en wel met een gemiddelde van 78 kg en een standaardafwijking van 11 kg.
- Wat wordt het gemiddelde en wat de standaardafwijking als de lengtes in “feet” en “inches” worden weergegeven? (1 foot = 30,48 cm en 1 inch = 2,54 cm)
 - Wat wordt het gemiddelde en wat de standaardafwijking als de gewichten in “stones” worden weergegeven? (1 stone = 6,35 kg)
- 25 Temperaturen kunnen zowel in graden Celsius als in graden Fahrenheit worden uitgedrukt. Omrekenen gaat met de formule: graden Fahrenheit = $1,8 \times (\text{graden Celsius}) + 32$.
Wat gebeurt er met een gemiddelde julitemperatuur van $17,4^\circ\text{C}$ en een bijbehorende standaardafwijking van $1,5^\circ\text{C}$ als de temperaturen naar Fahrenheit worden overgezet?

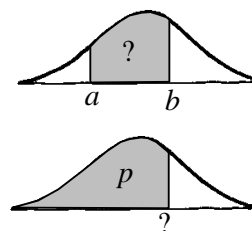
Een normale verdeling wordt volledig bepaald door twee parameters:

- het gemiddelde,
- de standaardafwijking.

Het gemiddelde en de standaardafwijking worden in de statistiek met Griekse letters aangegeven:

- μ (spreek uit *mu*) voor het gemiddelde,
- σ (spreek uit *sigma*) voor de standaardafwijking.

Stel dat een grootheid normaal verdeeld is en je weet het gemiddelde en de standaard-afwijking. De kans dat de grootheid een waarde tussen a en b aanneemt ligt dan in principe vast. De GR heeft een optie om die kans te berekenen. De GR heeft ook een optie om bij gegeven kans p de waarde van a te vinden, zo dat p de kans is op een waarde kleiner dan a .



- 26** In 1986 werd van 103.370 dienstplichtige 18-jarigen de lengte opgemeten. Hun gemiddelde lengte bleek 181,8 cm te zijn en de standaardafwijking 7 cm.
We willen weten hoeveel jongens 190 cm of langer zijn.
- Schets een normale kromme en geef daarbij de gegevens en het gevraagde aan.
 - Bereken met de GR hoeveel van de jongens naar verwachting 190 cm of langer waren.

Jongens die langer waren dan 200 cm of korter dan 160 cm werden op grond van hun lengte afgekeurd.

- Teken weer een bijpassend plaatje.
 - Bereken hoeveel jongens er in 1986 op grond van hun lengte werden afgekeurd.
 - Bereken vanaf welke lengte een jongen tot de 1% langste behoorde.
 - Bereken tot welke lengte een jongen tot de 5% kleinste behoorde.
- 27** Een ei weegt gemiddeld 63 gram met een standaardafwijking van 4 gram. Men verdeelt eieren in verschillende gewichtsklassen: S staat voor Small, dat is met een gewicht onder de 53 gram, M betekent Medium, dat is met een gewicht van 53 tot 61 gram, de L staat voor Large, dat is met een gewicht van 61 tot 73 gram en XL is Extra Large; dat zijn eieren die zwaarder zijn dan 73 gram.
- Bereken hoeveel procent van de eieren in de verschillende gewichtsklassen terecht komt.

Stel dat de eieren worden ingedeeld in vier opeenvolgende gewichtsklassen die elk 25% van het totaal bevatten.

- Bereken dan de bijbehorende gewichtsgrenzen.

- 28** Twee fabrikanten brengen voor dezelfde prijs eenzelfde type lamp op de markt. Merk A heeft een gemiddelde van 1250 branduren en een standaardafwijking van 300 uur. Merk B heeft een gemiddelde van 1200 uur en een standaardafwijking van 250 uur.
Je wilt een lamp kopen die minstens 1000 uur moet branden.
Welk merk heeft jouw voorkeur ?

- 29** Een tomatenkweker heeft geoogst. De vruchten variëren in grootte en gewicht. Het gewicht is normaal verdeeld met $\mu = 90$ gram en $\sigma = 15$ gram. In totaal zijn 60.000 tomaten geoogst. De oogst wordt op gewicht gesorteerd.

De drie gewichtsklassen zijn:

- klasse A: tot 70 gram,
- klasse B: van 70 tot 100 gram,
- klasse C: meer dan 100 gram.

- Hoeveel procent van de oogst komt in elk van de klassen terecht?

De opbrengst van een tomaat hangt af van zijn gewichtsklasse:

- klasse A: 20 cent,
- klasse B: 25 cent,
- klasse C: 30 cent.

- Welke opbrengst mag de kweker voor zijn hele oogst verwachten?

- 30** Nederlandse euromunten worden in Utrecht geslagen bij de Koninklijke Nederlandse Munt. De afmetingen en gewichten zijn aan zeer strikte regels gebonden.

Muntsoort	metaal	middellijn in mm	gewicht in mg
twee euro	koper/nikkel/messing	25,75	8500
een euro	koper/nikkel/messing	23,25	7500
vijftig cent	Nordic gold	24,25	7800
20 cent	Nordic gold	22,25	5740
10 cent	Nordic gold	19,75	4100
5 cent	staal/koper	21,25	3920
2 cent	staal/koper	18,75	3060
1 cent	staal/koper	16,25	2300

Het gewicht van een nieuw geslagen euromunt is normaal verdeeld met $\mu = 7500$ mg en $\sigma = 6$ mg. Munten die meer dan 15 mg afwijken van het vereiste gewicht mogen niet in omloop worden gebracht.

- Waarom gelden er zulke strikte eisen voor het toegestane gewicht?
- Bereken welk percentage van de nieuw geslagen één-euromunten niet in omloop zal worden gebracht.
- Per jaar zijn er 25 miljoen nieuwe één-euromunten nodig.
Hoeveel moeten er geslagen worden om aan die vraag te kunnen voldoen?

- 31** In een fabriek worden blikken gevuld met (gemiddeld) 1 liter verf. De vulmachine levert niet blikken van precies 1 liter. De inhoud van de blikken is normaal verdeeld met een standaardafwijking van 15 milliliter.

We willen weten hoeveel procent van de blikken meer dan 30 ml verf te weinig bevat.

- Schets een normale kromme en kleur de oppervlakte die hoort bij deze vraag.
- Bereken hoeveel procent van de blikken meer dan 30 ml verf te weinig bevat.

Een liter verf weegt 2 kg.

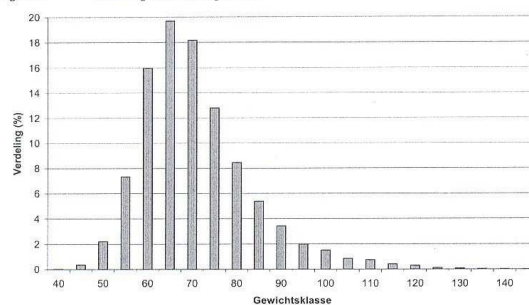
- Bereken hoeveel procent van de blikken minder dan 1980 gram verf bevat.

Begrepen?

Van 22000 zwangere vrouwen werd het gewicht bepaald. Hiernaast staat het histogram.

- Is het gewicht bij benadering normaal verdeeld?

Figuur 4 Verdeling maternaal gewicht



De diastolische bloeddruk (ofwel onderdruk) van mensen tussen 30 en 70 jaar is ongeveer normaal verdeeld met een gemiddelde van 85 (mm Hg) en standaardafwijking 13 mm Hg.

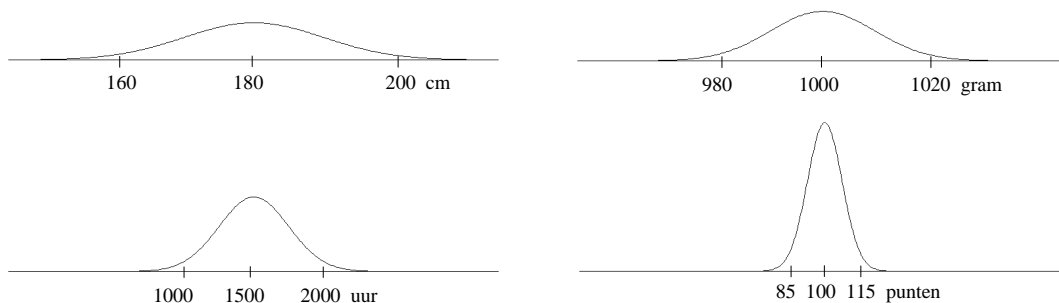
12,4% van de mensen tussen 30 en 70 jaar hebben een diastolische bloeddruk boven 100 mm Hg en 12,4% van de mensen hebben een bloeddruk onder ... mm Hg.

- Welke grens is dat?
- Tussen welke waarden ligt de bloeddruk van de middelste 50% van de mensen?

4.3 Standaardiseren

De verwantschap van normale verdelingen

We vergelijken vier normale verdelingen. Omdat de oppervlakte onder een normale kromme alle gevallen vertegenwoordigt (100%), is er zo geschaald dat de oppervlaktes onder de krommen gelijk zijn.

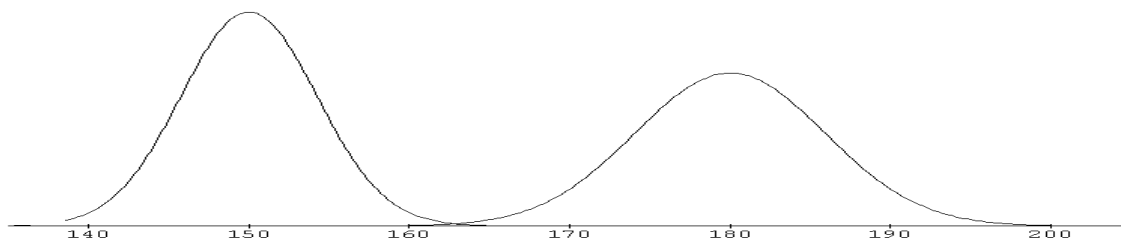


De verdelingen gaan over:

- de lengte van volwassen Nederlandse mannen (in cm),
- het gewicht van een kilopak suiker (in gram),
- de levensduur van TL-buizen (in uren),
- het IQ van Nederlandse scholieren (in punten).

- 32 a. Controleer dat bij de vier krommen geldt: als een verdeling n keer zo breed is, is hij ook n keer zo laag.
b. Leg uit waarom dat zo is.

Nederlanders zijn de langste mensen ter wereld, pygmeeën in Kameroen de kortste. De volwassen Nederlandse man is gemiddeld 1,80 met sd 7 cm. Een volwassen pygmeeman is gemiddeld 1,50 met sd 5 cm. Van beide lengtes staat hieronder de verdeling, met gelijke oppervlakte onder de krommen.



- 33 De verdelingen lijken veel op elkaar. Je kunt die van de pygmeeën in die van de Nederlanders overvoeren door naar rechts te verschuiven en vervolgens te verbreden (op te rekken in horizontale richting ten opzichte van het midden). Dan verandert de hoogte automatisch mee. Over welke afstand moet je naar rechts verschuiven en met welke factor moet je verbreden?

Normale verdelingen verschillen alleen wat gemiddelde en wat standaardafwijking betreft. Dat betekent dat je de ene in de andere kunt overvoeren door een horizontale verschuiving (van het ene gemiddelde naar het andere), gevolgd door een horizontale verbreding of versmalling (van de ene sd naar de andere).

- 34 Stel dat het aantal grammen suiker S in een pak normaal verdeeld is met gemiddelde 1000 en sd 10 en dat de levensduur L (uren) van TL-buizen normaal verdeeld is met 1500 en sd 50.
- a. Hoeveel moet je de verdelingskromme van S verschuiven en verbreden om die van L te krijgen?

Je kunt ook zeggen dat $\dots S - \dots$ dezelfde verdeling heeft als L .

- b. Welke getallen horen op de stippelijntjes?
c. Evenzo heeft S dezelfde verdeling als \dots (uitdrukken in L).

Opmerking

Iets dergelijks is ook aan de hand bij parabolen. Als twee parabolen een verticale symmetrieas hebben, kun je de ene zo verschuiven en oprekken dat hij precies op de andere valt; eventueel moet je nog spiegelen in de x -as.

Uitzonderlijk

- 35 Gemiddeld bedraagt de temperatuur in De Bilt in de maand juli 17,4 °C. In 2010 was de gemiddelde juli-temperatuur in De Bilt 19,2 °C.
- a. Is dat uitzonderlijk hoog? Wat denk jij?

Anneke simuleert op de computer het gooien met een dobbelsteen. De computer “gooit” 1000 keer. Anneke verwacht ongeveer 167 keer zes ogen te krijgen, met een standaardafwijking van 12. Bij de simulatie krijgt ze 150 keer zes ogen.

- b. Is dit uitzonderlijk weinig? Wat vind jij?

De consumentenbond controleert 10 kilopakken suiker. Gemiddeld behoren de pakken 1000 gram te bevatten. In de steekproef bleken acht pakken minder dan 1000 gram te bevatten.

- c. Vind jij dit uitzonderlijk?

Vaak is het lastig om, zo op het oog, te beoordelen of een waarneming uitzonderlijk is. Daarom gebruiken we een methode:

Bepaal het gemiddelde en de standaardafwijking. Kijk hoeveel sd's de waarneming boven/onder het gemiddelde ligt. Hoe hoger dit aantal sd's, des te uitzonderlijker is de waarneming.

Het aantal keer de sd dat een waarneming afwijkt van het gemiddelde, noemen we de **z-waarde** van die waarneming.

In formule:
$$\mathbf{z-waarde} = \frac{\text{waarneming} - \text{gemiddelde}}{\text{sd}}$$

- 36 De lengte van een Nederlandse volwassen man van 2,00 meter heeft z-waarde $(200 - 180) / 7 \approx 2,86$.
- Hoe lang is een volwassen Nederlander met z-waarde van -2,86?
 - Hoe lang is een volwassen Nederlander met z-waarde 1?
 - Hoe lang is een volwassen Nederlander met z-waarde 0?

- 37 We bekijken de lengte in een groep 16-jarige jongens en in een groep 16-jarige meisjes. Bij de jongens is de gemiddelde lengte 178 cm en de sd 7 cm. Bij de meisjes is de gemiddelde lengte 168 cm en de sd 6 cm.

Een jongen en een meisje uit deze groepen krijgen verkering. Ze zijn beiden erg lang: de jongen 196 cm en het meisje 186 cm.

Bereken de z-waarde van de lengte van de jongen en van de lengte van het meisje om te bepalen wie van de twee de grootste uitschieter is qua lengte binnen zijn/haar groep.

Als de waarneming zelf normaal verdeeld is, is de z-waarde dat ook, en wel met gemiddelde 0 en sd 1. We zeggen dat de z-waarde **standaardnormaal** verdeeld is.

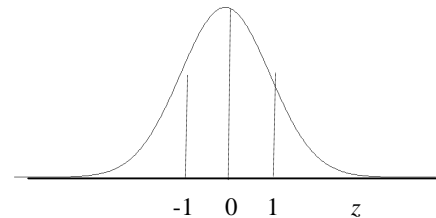
De normale verdeling met gemiddelde 0 en sd 1 is de “standaard” onder de normale verdelingen. Elke normale verdeling is daarop terug te voeren door van de waarneming de gemiddelde waarneming af te trekken en vervolgens door de standaardafwijking te delen.

[Vergelijk dit met de “standaard” $y = x^2$ onder de parabolen.]

In formule

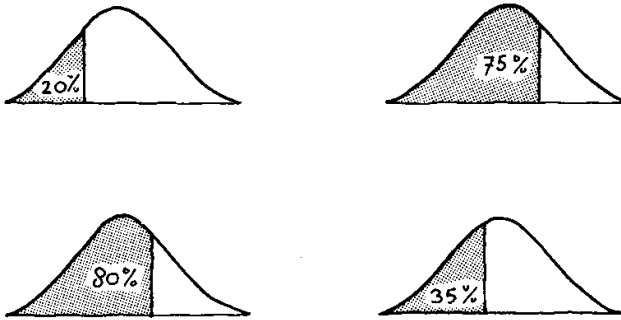
Als X normaal verdeeld is met gemiddelde μ en standaardafwijking σ , is $Z = \frac{X - \mu}{\sigma}$ **standaardnormaal** verdeeld.

Het overstappen op de standaardnormale verdeling noemen we **standaardiseren**.



- 38 X is normaal verdeeld met gemiddelde 1000 en sd 25.
 Y is standaardnormaal verdeeld.
- Leg uit dat de kans dat $X > 1035$ gelijk is aan de kans dat $Y > 1,4$.
 - De kans dat $X > a$ is gelijk aan de kans dat $Y > \dots$ (vul in).
- 39 De vuistregels zeggen dat bij een normale verdeling de kans op een afwijking kleiner dan $1 \times \text{sd}$ gelijk is aan 68% en een afwijking kleiner dan $2 \times \text{sd}$ gelijk is aan 95%.
- Wat betekent dit voor de z-waarde?
 - Controleer deze percentages voor de standaardnormale verdeling met de GR .
 - Wat is de kans dat een afwijking kleiner is dan $3 \times \text{sd}$?
- 40 X is bij benadering normaal verdeeld met gemiddelde 222 euro en sd 14 euro.
Tussen welke waarden ligt X met 95% kans (symmetrisch om 222)?

41 Welke z-waarden passen bij de volgende oppervlakten?



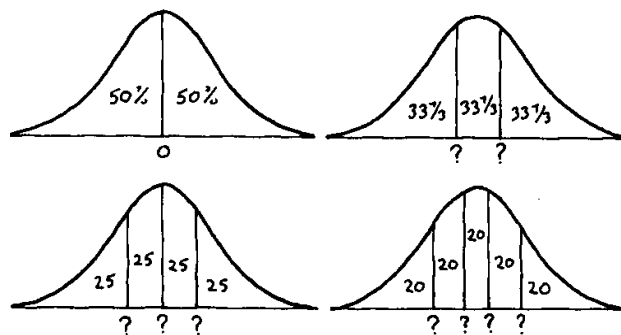
42 Bij gegeven percentages tussen twee waarden kunnen de bijbehorende z-waarden meestal niet gevonden worden. Twee situaties:



In het linkerplaatje liggen de linker- en rechtergrens even ver van het midden. Bij het rechter plaatje is dat niet zo.

- Bepaal de z-waarden in het linker plaatje.
- Kun je de z-waarden ook in het rechter plaatje bepalen?

43 De lijn bij $z = 0$ deelt het gebied onder de standaardnormale kromme in twee symmetrische helften, elk met oppervlakte 50%.



- Bepaal de z-waarden die de oppervlakte verdelen in drie gelijke stukken.
- Bepaal ook de z-waarden die de oppervlakte verdelen in vier gelijke stukken en in vijf gelijke stukken.

44 Een machine vult pakken met 1000 gram suiker. Althans dat is de bedoeling. Als de vulmachine ingesteld staat op 1000 gram, zal het werkelijke gewicht van een pak normaal verdeeld zijn met gemiddelde 1000 gram en standaardafwijking 10 gram.

- Ga na dat bijna 7% van de pakken een gewicht heeft van 985 gram of minder.

- b. Op welk gemiddelde moet de machine worden ingesteld, opdat slechts 2% van de pakken een gewicht heeft van 985 gram of minder? Neem aan dat de standaardafwijking 10 gram is, onafhankelijk van het gemiddelde waarop de machine is ingesteld.
- 45 Uit een onderzoek bleek dat de score van leerlingen bij het CSE wiskunde A1 vwo in 2000 bij benadering normaal verdeeld was. Het gemiddelde was 62 punten en 28% van de leerlingen had een onvoldoende (54 punten of minder).
- Welke z-waarde hoort bij 28%?
 - Bereken de standaardafwijking.
 - Bereken hoeveel punten je moet hebben om bij de 20% besten te horen.
- 46 Een slijter gaat op examen om het biercertificaat te halen. Het examen bestaat uit 40 drie-keuzevragen. Om te slagen moet hij ten minste 24 vragen goed beantwoorden, dat is 60%. De slijter heeft zich in het geheel niet voorbereid en zal de vragen dan ook op de gok beantwoorden. X is het aantal vragen dat de slijter goed beantwoordt. Dit is binomiaal verdeeld.
- Wat is de kans dat hij slaagt?
 - Bereken $E(X)$ en $sd(X)$.

Vako Drankenopleidingen neemt het examen af. Vako wil dat de kans dat iemand die de vragen puur op de gok beantwoordt hoogstens 1% kans heeft om te slagen.

c. Wat is de z-waarde die hoort bij het slagingspercentage van 1%?

Vako kan met minder dan 40 vragen volstaan om aan de eis te voldoen dat een gokker hoogstens 1% kans heeft om te slagen. De slagingseis blijft dat 60% van de vragen goed moet zijn beantwoord. We willen weten hoeveel vragen Vako minstens moet stellen; noem dat aantal vragen n . We noemen het aantal vragen dat een pure gokker goed heeft X .

d. Druk $E(X)$ en $sd(X)$ uit in n .

X is bij benadering normaal verdeeld.

- e. Leg uit dat de z-waarde van het minimum aantal goede antwoorden $\frac{0,6n - \frac{1}{3}}{\sqrt{\frac{2}{9}}}$ is.

Laat algebraïsch zien dat dit gelijk is aan $0,5657 \sqrt{n}$.

- f. Bereken n waarvoor $0,5657 \sqrt{n}$ gelijk is aan de in c gevonden waarde. Hoeveel vragen moet Vako minimaal stellen?
- g. Controleer het antwoord van vraag f door bij het gevonden aantal vragen de kans uit te rekenen dat een gokker slaagt.

Bij vraagstukken rond de normale verdeling draait alles om drie grootheden: het gemiddelde μ , de standaardafwijking σ en een percentage (dat is de oppervlakte boven een interval onder de normale kromme). De drie grootheden zijn gekoppeld: als er twee bekend zijn, kun je de derde uitrekenen. In principe zijn er dus drie verschillende soorten vragen mogelijk. Van elk soort volgt nu een voorbeeld.

47 μ en σ zijn bekend

Auto's worden op de lopende band in elkaar gezet. Een robot heeft voor het monteren van een wiel gemiddeld 96 seconden nodig met een standaardafwijking van 5 seconden. Er treedt vertraging op in de hele Montagelijijn als de robot meer dan 110 seconden nodig heeft.

a. Bereken bij hoeveel procent van de auto's er vertraging zal optreden.

μ en percentage zijn bekend

Een robot heeft gemiddeld 80 seconden nodig voor het bevestigen van een bumper. Bij zo'n 20% van de auto's is hij al na 77 seconden klaar.

b. Bereken hoe groot de standaardafwijking is.

σ en percentage zijn bekend

De robot die de deuren inzet, heeft daarvoor bij 8 op de 1000 auto's meer dan 105 seconden nodig. De standaardafwijking van de benodigde tijd bedraagt 4 seconden.

c. Bereken hoeveel seconden de robot gemiddeld doet over zijn karwei.

Andere mogelijkheden om μ of σ te vinden

Stel dat je het percentage p weet dat de waarde tussen a en b ligt. Als je bovendien μ weet, kun je σ vinden en omgekeerd. Daarvoor kun je de z -waarde goed gebruiken. Er zijn ook allerlei andere methodes. We noemen er een paar, waarbij we μ bekend veronderstellen en σ zoeken. Als omgekeerd σ bekend is en μ moet worden gezocht, gaat het net zo.

1. Proberen

Doe een gok voor σ en bereken bij die gok hoe groot het percentage is bij de bekende μ en gegokte σ tussen a en b . Pas de gok σ aan zodat dat percentage dichter bij p komt te liggen, net zo lang totdat je tevreden bent.

2. Met een grafiek

Teken de grafiek van het percentage tussen a en b bij de bekende μ , als functie van σ . Kijk voor welke invoer σ er p uitkomt.

3. Met de optie solver

Gebruik de optie op de GR om vergelijkingen op te lossen.

48 In een land is de gemiddelde lengte van de volwassen vrouwen onbekend; die noemen we μ . In dat land is de gemiddelde lengte van de volwassen mannen 190 cm. De standaardafwijkingen van de volwassen mannen en van de volwassen vrouwen zijn beide 7 cm.

a. Maak hierbij plaatjes van de verdelingen van de lengtes:

- een voor de volwassen vrouwen; geef daarin het percentage groter dan 190 aan,
- een voor de volwassen mannen; geef daarin het percentage kleiner dan μ cm aan.

Het percentage van de volwassen vrouwen die langer zijn dan 190 cm is gelijk aan het percentage van de volwassen mannen dat korter is dan μ cm.

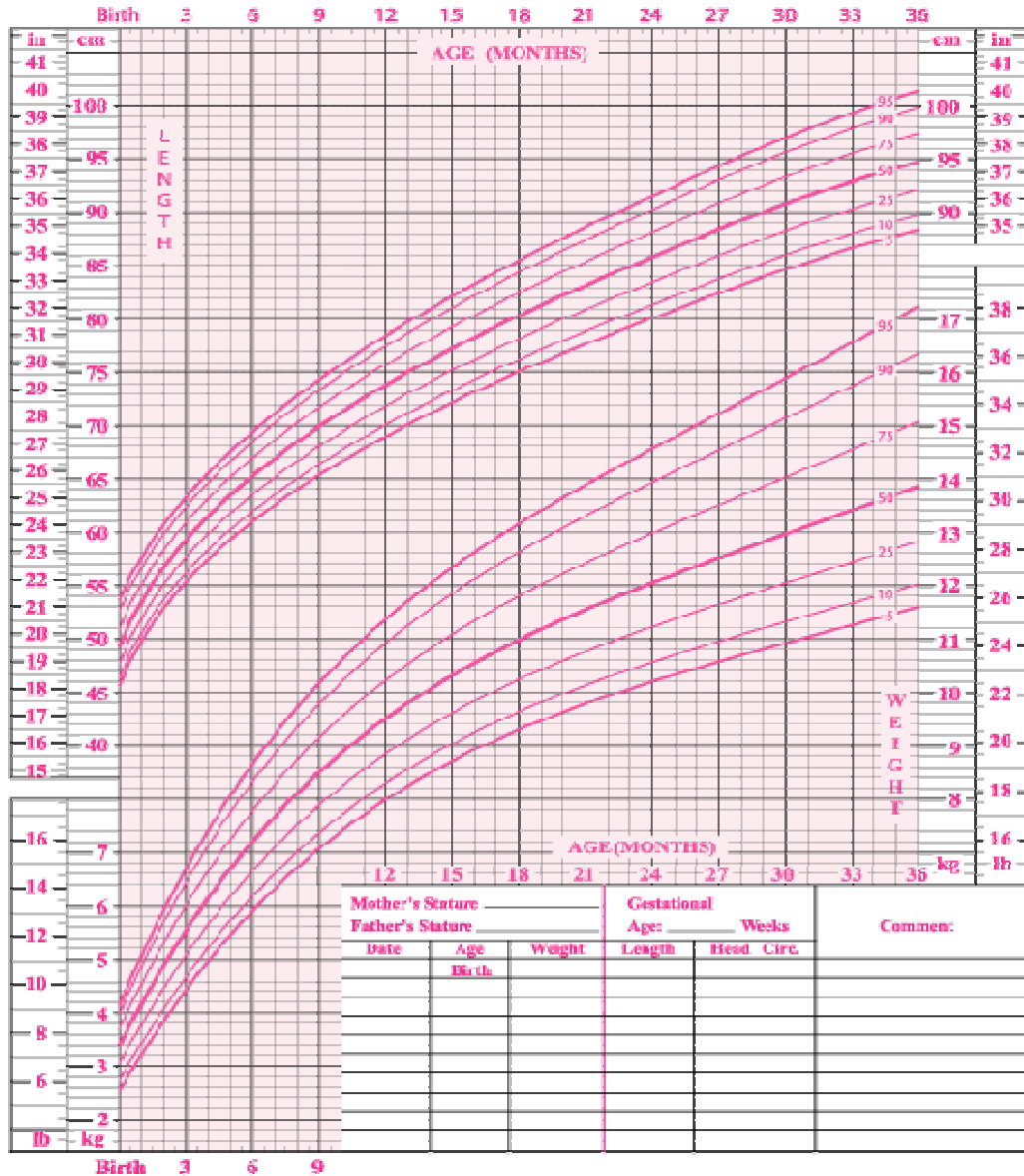
b. Leg dat uit met behulp van z -waarden.

49 Op de volgende bladzijde staan de groeicurves voor meisjes van 0 t/m 36 maanden. Je kunt er bijvoorbeeld uit aflezen dat 25% van de meisjes van 26 maanden 85 cm of korter is. En dat 90% van de meisjes van 18 maanden 12,6 kg of minder weegt.

Birth to 36 months: Girls

Length-for-age and Weight-for-age percentiles

NAME _____



SOURCE: Developed by the National Center for Health Statistics in collaboration with the National Center for Chronic Disease Prevention and Health Promotion (2000). <https://www.cdc.gov/growthcharts>

We letten op het gewicht na 36 maanden. Je kunt zien dat dan het gewicht niet zuiver normaal verdeeld is.

a. Hoe zie je dat? Laat in een schets zien hoe de verdelingskromme van het gewicht afwijkt van de normale verdeling.

We letten op de lengte na 36 maanden. Veronderstel dat dan de lengte van meisjes normaal verdeeld is.

b. Bepaal de sd van de lengte.

- 50** Zijn jongens slimmer dan meisjes of omgekeerd? Er is veel onderzoek gedaan naar eventueel verschil in intelligentie van jongens en meisjes. De resultaten spreken elkaar soms tegen; bovendien ligt het onderwerp politiek en sociaal gevoelig. Op Wikipedia is onder andere te vinden: *Analysing data from the international PISA student evaluation study, Machin and Pekkarinen found higher variance in boys' than girls' results on mathematics and reading tests in most OECD countries....* en *A study by Rosalind Arden and Robert Plomin from 2006 found greater variance among boys than among girls.*

Laten we eens van het volgende uitgaan:

- * het IQ van jongens is normaal verdeeld met gemiddeld 100,
 - * het IQ van meisjes is normaal verdeeld met gemiddelde 100,
 - * de sd van het IQ van jongens is 16, die van het IQ van meisjes is 14.
- a.** Schets de normale krommen van de IQ's in één figuur.

Het valt op dat er bij w4kangoeroe veel meer jongens onder de prijswinnaars zijn dan meisjes. Neem maar aan dat er evenveel jongens als meisjes meedoen aan w4kangoeroe. Gezien bovenstaande is het logisch dat er meer jongens onder de prijswinnaars zijn.

- b.** Leg dat uit.
- c.** Stel dat je 500 jongens en 500 meisjes hebt. Hoeveel jongens en hoeveel meisjes hebben een IQ boven 128?
- d.** Hoeveel procent van de leerlingen met een IQ boven 128 is jongen?

Begrepen?

Anne heeft op haar computer maar één plaatje van een normale curve staan. Dat gebruikt ze bij elke opgave waarin sprake is van een normale verdeling.

- Geef hierop commentaar.

Anne heeft bij een normale verdeling de z-waarde bepaald van de waarde waarboven 12% ligt en ook de z-waarde van de waarde waaronder 12% ligt.

- Wat is het verband tussen deze twee z-waardes?

Stel je weet dat bij twee normale verdelingen met dezelfde standaardafwijking de de percentages onder 50 respectievelijk onder 60 gelijk zijn.

- Wat weet je dan van hun gemiddeldes?

Stel je weet dat bij twee normale verdelingen met gemiddeldes 50 en 60 de percentages onder 70 gelijk zijn.

- Wat weet je dan van de standaardafwijkingen?



4.4 Over continue verdelingen

De uniforme verdeling

Op een computer en op de GR zit een *randomgenerator*. Die produceert “toevalsgetallen”, getallen die zuiver door toeval tot stand komen. Dat wil zeggen dat elk getal (in een zeker aantal decimalen, bijvoorbeeld tien) evenveel kans heeft.

- 51 Een *randomgenerator* produceert getallen van tussen 0 en 1 in zes decimalen, inclusief 0 zelf. Zo’n getal noemen we X . Dus $0,000000 \leq X \leq 0,999999$.
- Wat is $P(X = 0,123456)$?
 - Wat is $P(0,123 \leq X < 0,777)$?
 - Wat is de kans dat X uit alleen even cijfers bestaat?
 - Wat is de kans dat X eindigt op een 7?

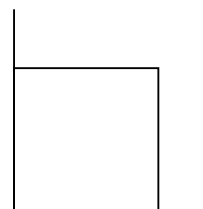
In opgave 51 was er sprake van 1 miljoen mogelijke uitkomsten voor X (tussen 0 en 1, inclusief 0), die allemaal even waarschijnlijk zijn. In opgave 52 gaan we *alle* uitkomsten X tussen 0 en 1 bekijken. X is dan dus een willekeurig getal uit $[0,1>$.

- 52
- Teken de verdelingskromme van X op het interval $[0,1>$.
 - Hoe groot is de kans dat $0,5 \leq X < 0,77$?
 - Hoe groot is de kans dat $a \leq X < b$, uitgedrukt in a en b ?

We zeggen dat X (uit opgave 52) **uniform verdeeld** is op het interval $[0,1>$.
uniform = gelijkmatig

- 53 Anne kijkt rond het middaguur op haar analoge horloge. S is het aantal seconden dat de secondenwijzer aangeeft.
- Hoe is S verdeeld?
 - Wat is de kans dat S tussen a en b ligt?
 - Weet je nog een ander voorbeeld van een uniform verdeelde grootheid?
- 54 Kijk nog even naar de randomgenerator X van opgave 51.
- Welke waarden neemt $6 \cdot X$ aan?
 - En $1 + \text{INT}(6 \cdot X)$, waarbij INT naar beneden afrondt op een geheel getal.
 - Leg uit dat je met de randomgenerator het werpen met een dobbelsteen kunt simuleren.
 - Hoe zou je het werpen met een munt kunnen nabootsen met behulp van X ?
- 55 We laten een randomgenerator twee getallen produceren, X_1 en X_2 .
- Wat is de kans dat $X_1 < 0,2$ en bovendien $X_2 < 0,5$?
 - Wat is de kans dat X_1 en X_2 beide met een 7 achter de komma beginnen?
 - Wat is de kans dat $0,31 \leq X_1 < 0,68$ en $0,36 \leq X_2 < 0,84$?

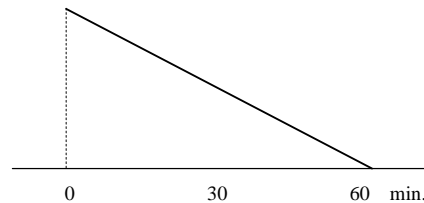
We kunnen het paar (X_1, X_2) in een plaatje weergeven door een punt in een vierkant met zijde 1.



- d. Waardoor wordt $P(0,31 \leq X_1 < 0,68 \text{ en } 0,36 \leq X_2 < 0,84)$ dan in dat plaatje voorgesteld?
- e. Teken een plaatje in een eenheidsvierkant bij $P(X_1 < X_2)$. Hoe groot is die kans?
- f. Teken een plaatje in een eenheidsvierkant bij $P(X_1 \text{ en } X_2 \text{ verschillen hoogstens } 0,2)$. Hoe groot is die kans?

- 56** Arie en Gré werken 's nachts. Na hun werk komen ze, onafhankelijk van elkaar, tussen middernacht en 1:00 uur aan bij een bushalte.
Er vertrekken in het eerste uur van de dag drie bussen: om 0:15 uur, om 0:30 uur en om 1:00 uur.
- a. Hoe groot is de kans dat Arie en Gré allebei de bus van 0:30 uur hebben?
 - b. Hoe groot is de kans dat Arie en Gré dezelfde bus hebben. Tip: teken zo nodig een passend plaatje in een eenheidsvierkant.
 - c. Hoe groot is de kans dat Arie en Gré niet meer dan 10 minuten na elkaar bij de bushalte arriveren?
 - d. Hoe groot is de kans dat Arie en Gré minder dan 30 minuten van elkaar bij de bushalte arriveren?

- 57** Het verschil in aankomsttijd bij de bushalte tussen Arie en Gré (uit opgave **56**) noemen we V (in minuten).
Hiernaast zie je hoe V verdeeld is.



Volgens deze kansverdeling is $P(V = 0) = 0$ en $P(V \leq 60) = 1$.

- a. Klopt dat met de werkelijke kansen?
- b. Wat is $P(V \leq 30)$ volgens deze kansverdeling?
Klopt dat met je antwoord op vraag **56d**?
- c. Controleer zo ook de kansverdeling voor het geval in **56c**.

Discreet en continu

Bij een kansexperiment hangen de uitkomsten van een grootheid af van toeval. We kunnen daarbij twee soorten grootheden onderscheiden. Die twee soorten leggen we uit aan de hand van de volgende voorbeelden.

- Het aantal ogen bij het werpen met een dobbelsteen.
- De lengte van een willekeurige Nederlandse volwassen man.
- Een willekeurig tijdstip tussen 0:00 en 1:00 uur.
- Het aantal mensen dat zichzelf trekt bij lootjestrekken voor Sinterklaas.
- Het aantal goede antwoorden bij het puur op de gok beantwoorden van een test met zes driekeuzevragen.

In het eerste voorbeeld hebben we te maken met een uniforme verdeling.

- 58**
- a. Met welk type verdeling hebben we te maken bij het tweede, derde en vijfde voorbeeld?
 - b. Welke waarden kunnen de grootheden in elk van de vijf voorbeelden aannemen?

In het tweede en derde voorbeeld kan de grootte een *continue* range van waarden aannemen. We spreken dan van een **continue verdeling**.

In het eerste, vierde en vijfde voorbeeld kan de grootte alleen *losse* waarden aannemen. We spreken dan van een **discrete verdeling**.

Variabelen of uitkomsten heten continu als zij in een bepaald interval iedere waarde kunnen aannemen. Tussen twee willekeurige waarden ligt altijd nog een tussenliggende waarde. Voorbeelden: tijd, lengte en massa.

Variabelen of uitkomsten heten discreet, als hun mogelijke waarden slechts een beperkt aantal getallen of klassewaarden zijn. De tussenliggende waarden hebben geen betekenis. Voorbeelden: aantallen bij tellingen, bloedgroepen.

Het woord *continu* in het dagelijks taalgebruik betekent *doorlopend, onafgebroken*.

Het woord *discreet* is bekender in de betekenis van *onopvallend, kies*. In de medische wereld noemt men ontstekingen discreet als ze *van elkaar gescheiden* zijn.

- 59 De plaatjes bij een continue en een discrete verdeling zijn verschillend. Bij een continue verdeling hoort een vloeiende kromme.

Welk type plaatje hoort bij een discrete verdeling?

- 60 De jaarlijkse hoeveelheid neerslag in Nederland is normaal verdeeld met gemiddelde 810 mm en standaardafwijking 152 mm.

a. Teken een plaatje bij de kans dat de hoeveelheid neerslag in een jaar tussen 700 en 800 mm ligt.

b. Hoe groot is die kans?

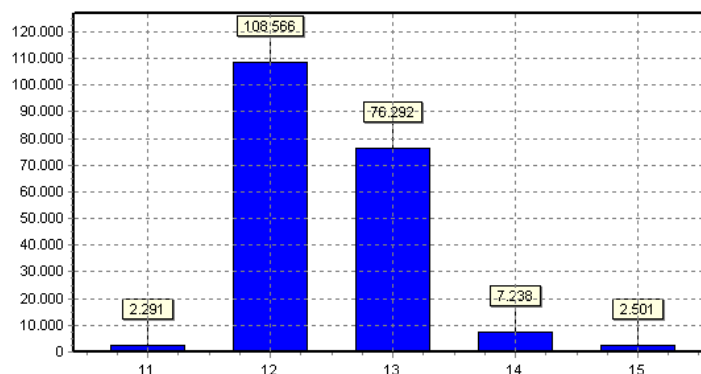
c. Wat is de kans dat in twee van de drie komende jaren de neerslag tussen 700 en 800 mm ligt?

De kansen in b. en c. betreffen verschillende typen verdelingen.

d. Welke verdeling is continu en welke discreet?

In de praktijk worden de gemeten waarden van een continue grootte vrijwel altijd afgerond. Een jaarlijkse neerslag van 812 mm betekent dat die hoeveelheid tussen 811,5 en 812,5 mm ligt. En daarmee wordt de continue grootte eigenlijk discreet!

- 61 Hieronder staat een histogram van de frequentieverdeling van de leeftijden van de leerlingen in de brugklas in Nederland (peildatum 21 juni 2010).



- a. Is de variabele *leeftijd* hierbij discreet of continu?

De leeftijd is afgerond op hele jaren!

- b. Wordt er naar boven of naar beneden afgerond?

We kunnen de leeftijd ook continu meten. Dan kan iemand bijvoorbeeld 12 jaar + 216 dagen oud zijn. Ofschoon ook nu weer wordt afgerond (nu op dagen). Dus helemaal continu is de leeftijd ook nu niet!

Als je de leeftijd continu meet, kun je de volgende vraag stellen: hoeveel procent was op 21 juni 2010 tussen 11,5 en 12,5 jaar oud?

- c. Schat dit aantal zo goed mogelijk.

- 62 Nogmaals de lengte van een volwassen Nederlandse man. Die lengte is normaal verdeeld met gemiddelde 180 cm en standaardafwijking 7 cm. We vragen ons af wat de kans is dat een Nederlandse volwassen man 180 cm is?

- a. Bereken die kans als die lengte betekent dat zijn werkelijke lengte tussen 179,5 en 180,5 ligt.
b. Wat is de kans als die lengte betekent dat hij 180,000... is.

De lengte van een Nederlandse volwassen man noemen we L (in cm). L is continu verdeeld.

- c. Wat is het grootst, $P(L = 180)$ of $P(L = 179)$?
d. Wat is het grootst, $P(L \leq 180)$ of $P(L < 180)$?

Als je L afrondt op hele cm, krijg je X .

- e. Wat is het grootst, $P(X = 180)$ of $P(X = 179)$?
f. Wat is het grootst, $P(X \leq 180)$ of $P(X < 180)$?

63 Gemengd normaal

De lengte van 18-jarige jongens is normaal verdeeld met gemiddelde 180 en standaardafwijking 7 cm; de lengte van 18-jarige meisjes is normaal verdeeld met gemiddelde 170 en standaardafwijking 6 cm.

- a. Teken de twee verdelingskrommen in één figuur.

We bekijken nu een grote groep van 18-jarigen, evenveel jongens als meisjes. We kiezen een willekeurig persoon uit de groep. Zijn/haar lengte in cm noemen we L .

- b. Teken met een andere kleur de verdelingskromme van L .
c. Is L normaal verdeeld, denk je? Waarom?
d. Bereken $P(L < 175)$ als de gekozen persoon een jongen is en ook als de gekozen persoon een meisje is.
Wat is dus $P(L < 175)$ in de gemengde groep?
e. Is L normaal verdeeld?

Het is mogelijk de verdelingskromme van L op de GR te tekenen. Dan kun je zien dat L niet normaal verdeeld is. In het algemeen is “gemengd normaal” dus niet normaal.

64 Het gemiddelde van normale verdelingen

In een land leven twee stammen, de Langen en de Korten. Van beide stammen is de lichaamslengte van een volwassen man normaal verdeeld: van de Langen met gemiddelde 185 cm, van de Korten met gemiddelde 160 cm. Beide verdelingen hebben standaardafwijking 6 cm.

In het land behoort 20% van de volwassen mannen tot de Langen en 80% tot de Korten.

Als er bij de Korten evenveel volwassen mannen zouden zijn als bij de Langen, dan zou de gemiddelde lichaamslengte van alle volwassen mannen 172,5 cm zijn. In dit geval is dat niet zo: de gemiddelde lengte van alle volwassen mannen is 165 cm.

a. Toon dit aan.

Er geldt dat de lichaamslengte van meer dan 60% van de volwassen mannen in het land kleiner is dan de gemiddelde lengte van alle volwassen mannen.

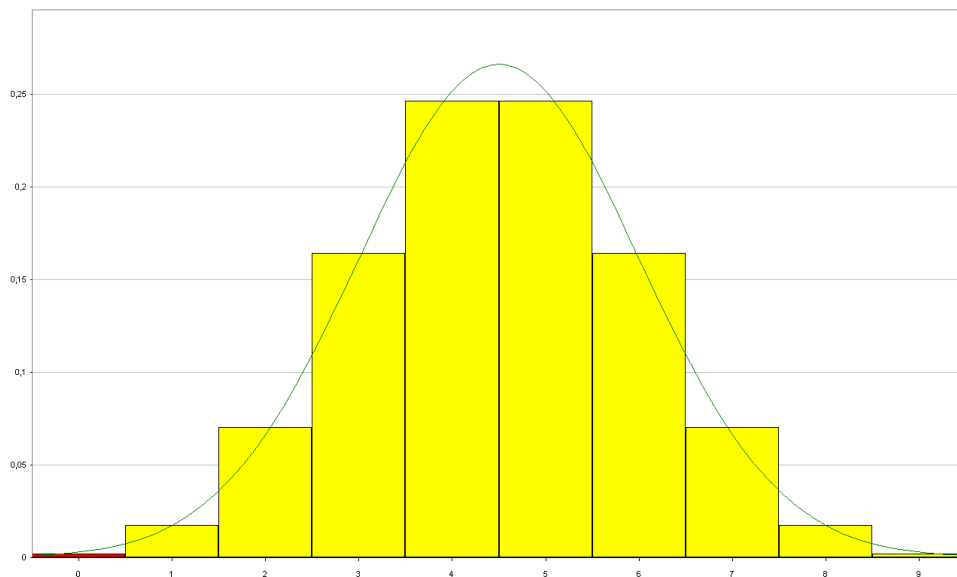
b. Toon dit aan.

c. Is de lichaamslengte van de totale groep van de volwassen mannen in het bewuste land normaal verdeeld? Licht je antwoord toe.

Examen vwo wiskunde b 2009, 2^{de} tijdvak

Binomiaal \approx normaal

65 X is het aantal keer kop bij negen worpen met een muntstuk. Hieronder staat een kans-histogram van X , met daar overheen de normale kromme die daar het best bij past. Dat is de normale verdeling met hetzelfde gemiddelde en dezelfde standaardafwijking als X . De bijbehorende grootte noemen we U .



a. Wat zijn μ en σ van X en U ?

b. Bereken $P(X = 3)$.

c. Bereken $P(2,5 \leq U \leq 3,5)$.

d. Vul in: $P(X = 5) \approx P(\dots \leq U \leq \dots)$
 $P(4 \leq X \leq 7) \approx P(\dots \leq U \leq \dots)$



- e. Ga naar VuStat/Kansverdelingen/Binomiale Verdeling. Teken het kanshistogram voor het aantal kop bij opeenvolgende aantallen worpen n met een muntstuk. Je ziet dat het kanshistogram steeds meer op een normale kromme gaat lijken. Teken de normale kromme erbij (aanvinken linksonder).

66 Y is het aantal keer zes bij 18 worpen met een dobbelsteen. V is de normale grootte die daar het best bij past, dat wil zeggen die hetzelfde gemiddelde en dezelfde standaardafwijking heeft.

- a. Wat zijn dat gemiddelde en die standaardafwijking?
b. Vergelijk $P(Y = 2)$ en $P(1,5 \leq V \leq 2,5)$.

c. Dezelfde vragen voor 180 in plaats van 18 worpen en 20 in plaats van 2 zessen.



- d. Ga naar VuStat/Kansverdelingen/Binomiale Verdeling. Teken het kanshistogram voor het aantal keer zes bij 18 en bij 180 worpen met een dobbelsteen. Teken er de normale kromme bij.

Algemeen

Een binomiale verdeling kan goed benaderd worden met een normale verdeling. Vooral als de succeskans p niet te ver van 0,5 afwijkt. Als p bijvoorbeeld 0,1 of 0,9 is, moet het aantal herhalingen n van de binomiale verdeling groter gekozen worden (ten minste tien).

Dat de binomiale verdeling voor grote waarden van n steeds meer op een normale verdeling gaat lijken werd omstreeks 1720 ontdekt door Abraham de Moivre. Later is de volgende algemene stelling bewezen.

Als onafhankelijke grootheden met dezelfde kansverdeling bij elkaar opgeteld worden, gaat de som steeds meer lijken op een normale verdeling.

Dit staat bekend als de **centrale limietstelling**.



Abraham de Moivre 1667 - 1754

De som van normale verdelingen

In opgave 63 bekeken we de lengte van 18-jarige jongens en 18-jarige meisjes. De lengte van 18-jarige jongens is normaal verdeeld met gemiddelde 180 cm en standaardafwijking 7 cm; de lengte van 18-jarige meisjes is normaal verdeeld met gemiddelde 170 cm en standaardafwijking 6 cm.

Kies nu een willekeurig 18-jarige jongen en een willekeurig 18-jarig meisje. Wat is dan de kans dat het meisje langer is dan de jongen? Of wat is de kans dat de jongen ten minste 17 cm langer is dan het meisje? Over dit soort vragen gaat het volgende gedeelte.

Om deze vragen te beantwoorden, willen we weten of het lengteverschil L (= lengte jongen – lengte meisje) normaal verdeeld is, en zo ja, wat het gemiddelde en de standaardafwijking van L is. We gaan eerst een en ander over de standaardafwijking herhalen.

Herhaling

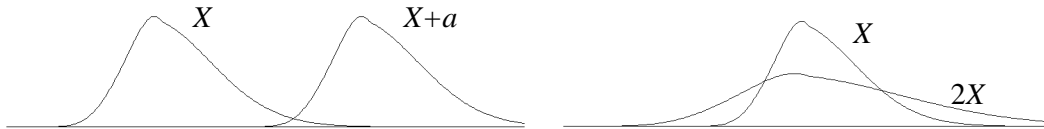
- * $E(X+Y) = E(X) + E(Y)$
- * $E(a+X) = a + E(X)$
- * $E(aX) = a \cdot E(X)$

en

- * $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$, mits X en Y onafhankelijk zijn.
- * $\text{Var}(a+X) = \text{Var}(X)$
- * $\text{Var}(aX) = a^2 \cdot \text{Var}(X)$

De formule voor $\text{Var}(X+Y)$ is lastig te bewijzen. In opgave **88** van Discrete Verdelingen is die aan de hand van een voorbeeld aangetoond.

Bij de formules voor $a+X$ en aX tekenen we plaatjes:



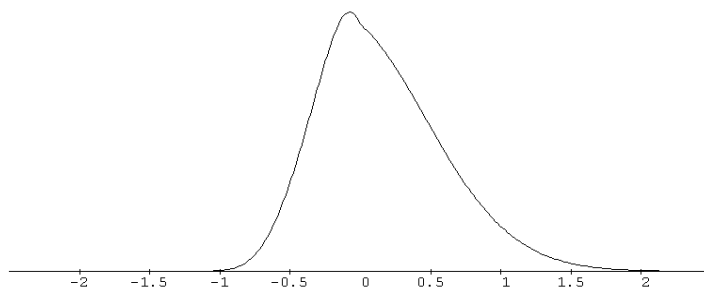
- 67 a.** Ga na dat de formules $E(a+X) = a + E(X)$ en $\text{Var}(a+X) = \text{Var}(X)$ in overeenstemming zijn met het linker plaatje.
- b.** $\text{Var}(2X) = 4 \text{Var}(X)$. Wat is dus het verband tussen $\text{sd}(2X)$ en $\text{sd}(X)$?
- 68** Xander en Yono spelen allebei een avond in een casino. Xander begint met 100 euro, Yono met 200 euro. Allebei zetten ze (onafhankelijk van de ander) twintig keer in, Xander zet steeds 5 euro in, Yono 10 euro. Veronderstel dat de kans $\frac{1}{2}$ is op verlies (dan ben je je inzet kwijt) en de kans ook $\frac{1}{2}$ is op winst (dan wordt de dubbele inzet uitbetaald). X is Xanders bedrag na de twintig keer spelen, Y is Yono's bedrag na afloop. X en Y zijn binomiaal verdeeld. $X+Y$ is het bedrag dat Xander en Yono samen na afloop hebben.
- a. Wat zijn $E(X)$, $E(Y)$ en $E(X+Y)$?
- b. Wat zijn $\text{Var}(X)$, $\text{Var}(Y)$ en $\text{Var}(X+Y)$?

Op grond van de centrale limietstelling zijn X en Y bij benadering normaal verdeeld, en om dezelfde reden is $X+Y$ bij benadering normaal verdeeld. En de verdeling van $X+Y$ zou nog beter op een normale verdeling hebben geleken als Xander en Yono (veel) vaker dan twintig keer hadden ingezet.

Algemeen

Als X en Y onafhankelijk zijn en beide normaal verdeeld zijn, dan is ook $X+Y$ normaal verdeeld.

- WB 69** Stel dat X de onderstaande verdelingskromme heeft. Zoals je ziet neemt X de waarden van -1 tot 2 aan. Stel dat $E(X) = 0,2$ en $\text{sd}(X) = 0,6$.



Bij X maken we $T = -X$ (het tegengestelde van X ; bijvoorbeeld als X de winst is bij een zeker spel, is T het verlies bij dat spel.)

- a. Teken op het werkblad de verdelingskromme van T .
- b. Hoe groot zijn $E(T)$ en $sd(T)$?

70 Hoe zit het met het lengteverschil L van een willekeurige 18-jarige jongen en een willekeurig 18-jarig meisje? De lengte van de jongen noemen we X , die van het meisje Y . Dus $L = X - Y$.

- a. Omdat Y normaal verdeeld is, is $-Y$ dat ook. Leg uit dat L normaal verdeeld is.

18-jarige jongens zijn gemiddeld 180 cm lang met een standaardafwijking van 7 cm; 18-jarige meisjes zijn gemiddeld 170 cm lang met een standaardafwijking van 6 cm.

- b. Bereken $E(L)$, $Var(L)$ en $sd(L)$.

Algemeen

Als X en Y onafhankelijk zijn en beide normaal verdeeld zijn, dan is ook $X - Y$ normaal verdeeld.

71 We kiezen twee willekeurige Nederlandse mannen en nemen hen de maat. De lengte van de eerste die wordt gekozen noemen we X , die van de tweede Y . Het gemiddelde van zowel X als Y is 180 cm en de standaardafwijking is 7 cm.

We moeten zorgvuldig kiezen, anders zijn X en Y niet onafhankelijk.

- a. Noem omstandigheden waarbij X en Y zeker niet onafhankelijk zijn.

We zijn geïnteresseerd in de som S van X en Y .

Anneke stelt voor in plaats van een tweede man te kiezen gewoon de lengte van de eerste man twee keer te nemen. Dus met $D = X + X$ te werken in plaats van $S = X + Y$.

- b. Zijn D en S gelijk?
- c. Wat zijn de gemiddelden en standaardafwijkingen van D en S ?

Ten slotte bekijken we ook nog de gemiddelde lengte van de twee gekozen mannen:

$$G = (X + Y)/2.$$

- d. Wat zijn het gemiddelde en de standaardafwijking van G ?
- e. Wat zijn het gemiddelde en de standaardafwijking van de gemiddelde lengte van negen onafhankelijk van elkaar gekozen mannen?

De wortel- n -wet

We bekijken n onafhankelijke herhalingen van een toevalsexperiment. Zeg dat de resultaten zijn: $X_1, X_2, X_3, \dots, X_n$.

De som van de resultaten is $S = X_1 + X_2 + X_3 + \dots + X_n$, het gemiddelde is $G = \frac{S}{n}$.

Er geldt: $E(S) = n E(X)$, $sd(G) = \sqrt{n} sd(X)$ en $E(G) = E(X)$, $sd(G) = \frac{sd(X)}{\sqrt{n}}$.

Bovendien geldt:

Als $X_1, X_2, X_3, \dots, X_n$ normaal verdeeld zijn, zijn S en G zijn dat ook.

- 72** In een diepvriespak lekkerbekjes zitten volgens de verpakking 4 tot 6 wijtingfilets in beslag die samen een gewicht van 500 gram hebben. Neem aan dat het gewicht van zo'n pak normaal verdeeld is met gemiddelde 500 gram en standaardafwijking 15 gram.



- a. Waarom zal de standaardafwijking van het gewicht van een pak lekkerbekjes waarschijnlijk groter zijn dan de standaardafwijking van het gewicht van bijvoorbeeld een pondspak suiker?

Iemand koopt zo'n pak lekkerbekjes.

- b. Hoe groot is de kans dat het gewicht daarvan meer dan 10% afwijkt van wat de verpakking belooft?

Iemand koopt drie van deze pakken.

- c. Hoe groot is de kans dat het totale gewicht van de drie pakken meer dan 10% afwijkt van het te verwachten totale gewicht?

De kans bij c is beduidend kleiner dan bij b.

- d. Kun je dat verklaren?

- 73** In veel winkels wordt bij het afrekenen afgerond op veelvoud van 5 eurocent. Door het afronden betaalt de klant meestal iets te veel of te weinig. Het aantal eurocent dat hij te veel betaalt noemen we X ; in het geval dat de klant te weinig betaalt is X negatief.

- a. Bereken $E(X)$ en $\text{Var}(X)$.
b. Welke veronderstelling heb je gedaan bij vraag a?

Op een dag heeft een supermarkt – waar op bovenstaande manier wordt afgerond – 200 klanten gehad. T is het aantal eurocent dat de supermarkt die dag door het afronden teveel ontvangt.

- c. Op grond van welke stelling is T bij benadering normaal verdeeld?
d. Bereken de kans dat T groter dan 0,50 euro is.
e. Tussen welke grenzen ligt T met 95% zekerheid?

74 Twee koplampen

De levensduur van een halogeenkoplamp van een auto is normaal verdeeld met een gemiddelde van 2500 branduren en een standaardafwijking van 450 uur.

Neem aan dat de levensduur van de linker koplamp van een auto en de levensduur van de rechter koplamp onafhankelijk van elkaar zijn.

- a. Bereken de kans dat zowel de linker als de rechter koplamp binnen 2100 branduren kapot gaat.

De levensduur van de rechter koplamp noemen we R en die van de linker koplamp L .

Om R en L met elkaar te vergelijken, bekijken we het verschil V , gedefinieerd door $V = R - L$. Als bijvoorbeeld $V = -100$, dan brandt de linker koplamp 100 uur langer dan de rechter koplamp.

- b. Wat zijn het gemiddelde en de standaardafwijking van V ?
c. Bereken de kans dat het verschil in levensduur van de beide koplampen kleiner is dan 20 uur.

Naar: examen vwo wiskunde B, 2007, eerste tijdvak

75 Heupoperaties

Patiënten lopen na een operatie in het ene ziekenhuis veel meer gevaar een infectie te krijgen dan in het andere. In het jaar 2003 werden in een bepaald ziekenhuis 120 heupoperaties uitgevoerd, waarna 6 patiënten een infectie kregen. De directie vond het percentage van 5% infectiegevallen te hoog en nam extra preventieve maatregelen. In 2004 werden 154 heupoperaties uitgevoerd, met nu 2 infectiegevallen. Men vroeg zich af of dit betere resultaat toeval was of dat het door de extra preventieve maatregelen kwam.

- a. Bereken de kans op hoogstens 2 infectiegevallen bij 154 operaties voor het geval dat de kans op infectie per operatie 0,05 is.

Omdat de zojuist berekende kans klein is, neemt men aan dat na de extra preventieve maatregelen de kans op infectie na een operatie is afgenomen.

De kans op infectie na een operatie na de extra preventieve maatregelen noemen we p .

- b. Zoek de waarde van p waarvoor geldt: de kans op hoogstens 2 infectiegevallen bij 154 patiënten is 0,1.

De afgelopen vijf jaar was de verpleegduur in Nederlandse ziekenhuizen voor heupoperaties ongeveer normaal verdeeld met een gemiddelde van 4,5 dagen en een standaardafwijking van 1,8 dagen.

Van 100 patiënten wordt de gemiddelde verpleegduur bepaald.

- c. Bereken de kans dat de gemiddelde verpleegduur groter is dan 5,0 dagen. (Dat zou voor de directie aanleiding zijn om maatregelen te nemen.)

Naar: examen vwo wiskunde B, 2008, eerste tijdvak

- 76 De diameter van de bout DIN931 is normaal verdeeld met gemiddelde 6,0 mm en standaardafwijking 0,02 mm. De diameter van de bijbehorende moer is normaal verdeeld met gemiddelde 6,5 mm en standaardafwijking 0,2 mm. Iemand pakt willekeurig een bout en een moer van dit type. We willen weten wat de kans is dat de moer te klein is voor de bout.



Noem de diameter van de bout B en die van de moer M (in mm).

- Hoe is $M - B$ verdeeld?
- Wat betekent het voor $M - B$ dat de moer te klein is voor de bout?
- Wat is de kans dat de moer te klein is voor de bout?

77 Anne moet elke ochtend op halte Terminus overstappen van lijn 5 op lijn 9. De aankomsttijd van lijn 5 is normaal verdeeld met gemiddelde 7:30 uur en standaardafwijking 3 minuten. De vertrektijd van lijn 9 is normaal verdeeld met gemiddelde 7:35 en standaardafwijking 4 minuten.

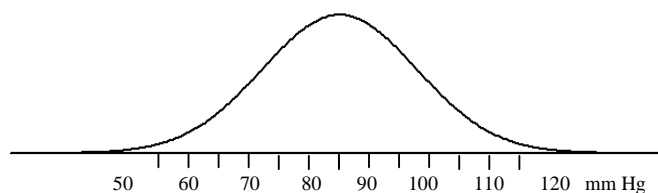
Wat is de kans dat Anne de aansluiting mist?

78 Easy is startloopster in een estafetteploeg 4 maal 100 meter. Haar tijd over de 100 meter is normaal verdeeld: gemiddeld 12,4 seconden, met een standaardafwijking van 0,6 seconden. De andere drie loopsters in Easy's ploeg doen korter over de 100 meter, want zij hebben een vliegende start. Hun tijd is normaal verdeeld met gemiddelde 10,8 seconde en standaardafwijking 0,4 seconden.

Wat is de kans dat hun totaal tijd voor de 4 keer 100 meter onder de 44 seconden ligt?

Begrepen?

De diastolische bloeddruk (ofwel onderdruk) van mensen tussen 30 en 70 jaar is ongeveer normaal verdeeld met een gemiddelde van 85 (mm Hg) en standaardafwijking 13 mm Hg.



- Is de bloeddruk een continue of een discrete variabele?
- Maak een histogram van de bloeddruk met klassengrenzen 50, 60, 70, 80, 90, 100, 110, 120.

Een arts noteert de bloeddruk als geheel getal.

- Bij hoeveel procent van de mensen zal hij 85 noteren?

Hiernaast staan de gemiddelde temperaturen per seizoen in De Bilt en de bijbehorende standaardafwijkingen.

- Bereken de gemiddelde jaartemperatuur in De Bilt en de bijbehorende standaardafwijking.

	<u>gem</u>	<u>sd</u>
winter	2,7	1,8
lente	8,7	1,0
zomer	16,4	1,0
herfst	10,0	1,0