Meten & Weten

Gerard Tuk

Meten & Weten

Gerard Tuk

Meten & Weten

Uitgave 2025/1
Eerste uitgave: 27 december 2023
Tweede uitgave: 26 augustus 2024
Derde uitgave: 7 augustus 2025
Auteur: Gerard Tuk

Druk: Pumbo.nl
Omslagontwerp: Gerard Tuk
Vormgeving binnenwerk: Gerard Tuk
Lettertype: Linux Libertine (tekst)
Amaranth (koppen)
Consolas (code)

Inhoud

Proeven

0. Lees wijzer

Toen ik in augustus 1986 werd gekeurd voor militaire dienst, vonden ze het bij Defensie nodig of nuttig om mijn lengte vast te stellen. Mijn naam werd afgeroepen en er bulderde een bevel: ik diende op blote voeten met mijn rug tegen de muur te gaan staan. Als dat maar goed afliep.

Zo begint straks hoofdstuk 1, en hoe het afloopt, zie je daarna wel.

Vóór hoofdstuk 1 komt hoofdstuk 0, met een leeswijzer. Dit boek is namelijk anders dan de meeste andere boeken. Het heeft een beetje de stijl van *Lessons Learned in Software Testing* van Cem Kaner. Die schrijft in zijn voorwoord:

Imagine that you are holding a bottle of 50-year-old port. There is a way to drink port. It is not the only way, but most folks who have enjoyed port for many years have found some guidelines that help them maximize their port-drinking experience. Here are just a few:

Lesson 1: Don't drink straight from the bottle. (...) Don't gulp down the port.

Lesson 2: Don't drink the entire bottle. If you are drinking because you are thirsty, put down the port and drink a big glass of water.

Lees wijzer! Dit boek is wel bedoeld om van voor tot achter door te lezen, maar niet in één adem. De opzet is dat je losse 'lesjes' leest en tot je door laat dringen. De oefeningen zijn ervoor om het je beter te laten snappen en onthouden. Zwemmen en pianospelen leer je ook niet uit een boek.

> Meten & Weten bestaat eigenlijk uit verhaaltjes en heeft hier en daar onbelangrijke uitstapjes. Die zijn grijs gemarkeerd, zoals deze alinea. Die mag je rustig overslaan, maar misschien zijn ze juist wel leuk. Of toch leerzaam, omdat ze je aanzetten tot verder nadenken over de stof.

Studieboeken gebruik je niet altijd voor je lol, maar het is mooi meegenomen als ze prettig zijn om te lezen. De schrijver van Meten & Weten heeft daarnaar gestreefd: serieuze inhoud op een luchtige manier gebracht.

Bij de opleiding Mechatronica aan de Haagse Hogeschool hebben er voor het vak *Error Budgeting* twee verschillende boeken op de lijst gestaan, namelijk dat van Taylor (1997) en van Hughes & Hase (2010). De ouverture van (het voorwoord van) het boek van Taylor luidt als volgt.

All measurements, however careful and scientific, are subject to some uncertainties.

Berendsen (2016) zet in hoofdstuk 1 in met een onmogelijkheid:

It is impossible to measure physical quantities without errors.

Ratcliffe & Ratcliffe (2014) steken op een soortgelijke manier van wal.

Every time anyone takes a measurement, the only certainty is that the measurement is not exact.

Dat klinkt nog wat cynischer, wanhopig bijna. Maar ook filosofisch. Het heeft iets van een milde spot in zich. Je hebt maar één zekerheid: het klopt niet helemaal. En dat is nog universeel ook.

Kirkup & Frenkel (2006) kiezen hun uitgangspunt in het dagelijkse leven en hebben het niet meteen over metingen.

We live with uncertainty every day. Will the weather be fine for a barbecue at the weekend? What is the risk to our health posed by a particular item of diet or environmental pollutant? Have we invested our money wisely? It is understandable that we would like to be able to eliminate, or at least reduce, uncertainty.

Is dit niet een heel ander soort metingen? De onzekerheid die ontstaat doordat je geen glazen bol hebt, is dat niet iets anders dan een meting doen in het hier en nu en nog niet precies weten wat je hebt? Tot op zekere hoogte wel. Maar het weer van morgen is niet zuiver toeval. De onzekerheden in de metingen van vandaag maken dat we niet alles weten over het volgende etmaal.

Onzekerheid is iets waar je last van hebt en wat ingeperkt moet worden. Je zou kunnen kiezen voor een hapje en een drankje binnenshuis in plaats van een barbecue. Je verstand gebruiken bij wat je eet en drinkt en accepteren dat ziekte bij het leven hoort. En je geen zorgen maken over beleggingen, maar blij zijn dat je geen schulden hebt.

Over schulden gesproken: Lyons (1991) begint zijn boek met

שְׁגִיאוֹת מִי־יָבִין

Mochten er lezers zijn die geen Hebreeuws beheersen — die heb je er altijd bij: dit is een Bijbeltekst. Psalm 19:13. *Maar wie kan al zijn fouten kennen? Spreek mij vrij van verborgen zonden.* ‘Zonden’ zijn ook fouten, maar wel van een heel andere categorie.

Dezelfde Louis Lyons schrijft op de achterkant van zijn boek het volgende.

It is usually straightforward to calculate the result of a practical experiment in the laboratory. Estimating the accuracy of that result is often regarded by students as an obscure and tedious routine, involving much arithmetic. An estimate of the error is, however, an integral part of the presentation of the results of experiments.

Schatten van nauwkeurigheden en/of onzekerheden is kennelijk nog niet eens leuk ook! De woorden *obscure* en *tedious* voorspellen niet veel goeds.

Taylor vervolgt zijn voorwoord wat dat betreft iets optimistischer.

The analysis of uncertainties, or "errors", is a vital part of any scientific experiment, and error analysis is therefore an important part of any college course in experimental science. The challenges of estimating uncertainties and of reducing them to a level that allows a proper conclusion to be drawn can turn a dull and routine set of measurements into a truly interesting exercise.

'De uitdagingen' maken wat saai is een 'waarlijk interessante oefening.'

Hughes & Hase (2010), de auteurs van het boek dat op de lijst stond vóór Meten & Weten, hebben het niet over 'saai' of 'leuk', maar geven al vroeg in hun eerste hoofdstuk aan dat het *belangrijk* is wat we doen. Je bent niet klaar met je meting als je niet gezegd hebt hoe (on)zeker je van je zaak bent.

We will emphasise in this book that an experiment is not complete until an analysis of the uncertainty in the final result to be reported has been conducted.

Important questions such as: "Do the results agree with theory?", "Are the results reproducible?" and "Has a new phenomenon or effect been observed?" can only be answered after appropriate error analysis.

De vraag is niet eens of het leuk is. Het *moet*, want anders weet je zo weinig.

Mijn gewaardeerde collega Mark Schrauwen schreef me tijdens het lezen van Editie 2023 van Meten & Weten:

Ik vind het lastig om te bepalen met wat voor boek ik te maken heb. Het ene moment denk ik het is een verhalend boek zoals Zoekt & Gijzelt, het andere moment is het iets meer een studieboek. Ik zou zelf gebaat zijn met een soort van mindmap die een grafische relatie tussen de hoofdstukken aanbrengt. Die ga ik sowieso voor mezelf maken.

Wat een goed idee! Het boek nemen zoals het is, en er dan *zelf* mee aan de slag gaan. Zonder leeswijzer is het devies: Lees wijzer!

Oefeningen

1. Vat in hooguit vijftig woorden samen wat de openingszinnen van al die boeken eigenlijk zeggen.
2. Bedenk wat je gaat doen om ervoor te zorgen dat je tijdens het lezen van dit boek kunt denken: "Ik lees wijzer".

1. “Tweeënzeventig zes!”

Toen ik in augustus 1986 werd gekeurd voor militaire dienst, vonden ze het bij Defensie nodig of nuttig om mijn lengte vast te stellen. Mijn naam werd afgeroepen en er bulderde een bevel: ik diende op blote voeten met mijn rug tegen de muur te gaan staan. Als dat maar goed afliep.

Iemand liet een stompe staaf op mijn — toen nog — donkere haren dalen, las iets af bij een of andere streep, en schreeuwde: “tweeënzeventig zes!”
De toon en het bijpassende volume maakten duidelijk dat eventuele tegenspraak geen gevolgen zou hebben. Of juist wel.

Dat “tweeënzeventig zes!” duurde, als je de galm niet meerekende, ongeveer 1,414 213 562 seconde. Beide woorden riepen bij mij een paar gedachten op, die ik strak zwijgend voor me hield. De eerste was dat er een *eenheid* ontbrak, maar die sprak kennelijk vanzelf. Een belangrijkere bespiegeling was dat die *zes* waarschijnlijk het aantal millimeters betrof. Het leek me onwaarschijnlijk dat je zó goed kon vaststellen hoe lang iemand was. Wat me nog het meest verbaasde: dat ze er *precies* een meter naast zaten…

Om goedgekeurd te worden, moest je in ‘mijn’ jaar minimaal 1 meter 60 zijn en maximaal 2 meter. Was dat toegestane bereik van 40 cm soms twee keer tweemaal de standaardafwijking? In dat geval was $\sigma = 10$ cm en viel ik in de middengroep. Dat had ik niet verwacht, omdat ik vroeger bij gym altijd tot de minst lange leerlingen behoorde. Als die dertig jongens op volgorde van lengte op de witte lijn moesten gaan staan, zat ik altijd bij de kleinste vijf.

Die millimeters leken mij wat overdreven en die 1 en de eenheid lieten ze waarschijnlijk weg voor het gemak. Verdedigbare keuzes, van Defensie. Een afstand van top tot teen druk je nu eenmaal uit in meters en centimeters. Waarom zou je het erbij zeggen, als het vanzelf spreekt? Bij veruit de meeste zeventienjarigen is het cijfer voor de komma een 1. Zo’n twee procent van de mensen met mijn geboortejaar haalt de twee meter.

Als vwo’ertje aan het begin van zijn laatste schooljaar heb ik die keuring zó serieus genomen dat ik de afgeronde 1,73 m jarenlang als mijn officiële lengte heb beschouwd. Paspoorten, rijbewijzen, ID-kaarten: meer dan dertig jaar lang hebben die het resultaat van deze ene meting vermeld. Nog steeds trouwens. Op het gemeentehuis werd ik bij het verlengen van een officieel document nooit eens een keertje nagemeten. Ze vroegen gewoon of het oude nog klopte en ik zei “ja”. Zonder dat zelf gecontroleerd te hebben. Ik hield me met haast militaire discipline aan wat die officier als officieel had vastgelegd.

In de zomer van 2022 besloot ik een poging te doen zelf vast te stellen hoe ver ik boven het maaiveld uitkwam. In tweevoud, omdat het herhalen van een meting vaak verstandig is. Beide keren kwam ik op 1,71 m...
Hadden ze dan niet goed gekeken toen, in die militaire meetkamer in Breda? Of gaan mensen al veel eerder krimpen dan ik dacht?

Studenten van nu zijn gemiddeld ruim een centimeter langer dan toen ik studeerde. (Bron: een oppervlakkig zoektochtje via Google. Doe dit nooit in je afstudeerverslag!) Zelf ben ik dus inmiddels ruim een centimeter kleiner. Zou het — zo vroeg ik me tijdens het schrijven van dit hoofdstuk af — niet zo kunnen zijn dat de *meters* van tegenwoordig gewoon kleiner zijn? De euro is ook minder waard dan toen hij werd ingevoerd.

Onzekerheden noteren we met een plusminusteken (±) en kleine delta (δ). Mijn lengte is gelijk aan $1{,}71 \pm 0{,}02\ m$, oftewel $l = 1{,}71\ m$ en $\delta l = 0{,}02\ m$.

Het schatten (meten) van die lengte l "voelt" als makkelijker dan het schatten van δl. Niet veel mensen zullen beweren dat ik 1,29 meter ben, behalve dan dommeriken met een meetlat van drie meter die ze per ongeluk ondersteboven houden. Toch scheelt dat nog geen 25% met de lengte die ik zelf opgeef. Die militair van toen overdreef met z'n millimeters, maar er zijn heus wel optimisten die denken dat je iemand met een marge van minder dan een hele centimeter kunt meten. Pessimisten zullen drie centimeter misschien wel te weinig vinden. Dat scheelt dus al 50%, beide kanten op.

Oefeningen

1. *"Zou het niet zo kunnen zijn dat de meters van tegenwoordig gewoon kleiner zijn?"* Is deze (niet serieus bedoelde) uitspraak een theoretisch mogelijke verklaring voor het verhaal hierboven over het verschil in lengtes toen en nu?
2. Wanneer is de meter voor het laatst veranderd?
3. Uit welke bron heb je het antwoord op de vorige vraag?
4. In dit hoofdstuk wordt gesproken over een afwijking van *twee keer tweemaal de standaardafwijking*. Is dat niet dubbelop?
5. *Dat "tweeënzeventig zes!" duurde, als je de galm niet meerekende, ongeveer 1,414213562 seconde.* Welke twee aspecten uit dit hoofdstuk zijn verwerkt in deze (ook al niet serieus bedoelde) uitspraak?
6. *Waarom zou je het erbij zeggen, als het vanzelf spreekt?*

2. Snelheidscontroles

Op de website van het Openbaar Ministerie staat de volgende tekst in het menu Home > Onderwerpen > Verkeer > Handhaving in het verkeer > Snelheid en te hard rijden > Marges en meetcorrecties.

Snelheidscontroles zijn een veelbesproken onderwerp. Als de politie een voertuig geflitst heeft, wordt er met twee zaken rekening gehouden:

1. Meetcorrectie

Om met 100% zekerheid te kunnen garanderen dat iemand te hard heeft gereden, trekt de politie van de gemeten snelheid een paar kilometer af. De meetcorrecties zijn:

- *bij minder dan 100 km/u: 3 kilometer meetcorrectie*
- *bij meer dan 100 km/u: 3 procent meetcorrectie*
- *Het Nederlands Meetinstituut ijkt alle flitsapparatuur.*

2. Ondergrens voor bekeuren

Er geldt een ondergrens van 4 km/u voordat wordt bekeurd.

De tekst roept wat vragen op. Dat het 'veelbesproken' is, zal waar zijn. Maar wat was de inhoud van die gesprekken? Is er soms discussie over de vraag of de controles eerlijk of nodig zijn? Vindt men dat er te veel of te weinig wordt gecontroleerd? Hoeveel mensen denken dat eigenlijk en *waarom* dan?

Over de eerste zin over de meetcorrectie valt ook nog wel wat te zeggen.

Om met 100% zekerheid te kunnen garanderen dat iemand te hard heeft gereden, trekt de politie van de gemeten snelheid een paar kilometer af.

Verderop in dit boek zullen we zien dat we het in dit vakgebied zelden over 100% zekerheid hebben, vanwege het simpele feit dat er voor 100% zekerheid extreem ruime marges of zeer goede middelen nodig hebt.

> Met 100% garanderen is dubbelop, tenzij je luchtig over garantie denkt.

De tekst noemt een ondergrens van 4 km/u… Word je dan nooit bekeurd als je langzamer dan een slakkengangetje van 4 km/u rijdt? Er zal wel een ondergrens *aan de overschrijding van de maximumsnelheid* bedoeld worden.

De meetcorrecties zijn:

- *bij minder dan 100 km/u: 3 kilometer meetcorrectie*
- *bij meer dan 100 km/u: 3 procent meetcorrectie*
- *Het Nederlands Meetinstituut ijkt alle flitsapparatuur.*

Voor een goede technicus doet dit pijn aan de ogen… Een correctie van 3 km op een snelheid van 100 km/u. Dat betekent dat je een *afstand* probeert af te trekken van een *snelheid*. Dat is zoiets als twee appels uit een doos met twaalf eieren halen. Het kan niet en het slaat nergens op.

Wat een goede technicus zich ook meteen afvraagt: en als je nu eens exact 100 km/u rijdt? Strikt genomen zegt de tekst daar niets over. Toch kun je het wel met gezond verstand invullen. Elk van beide regels geeft voor 100 km/u namelijk dezelfde correctie.

Eigenlijk zou de vraag moeten zijn: *Weten we voldoende zeker dat de bestuurder te hard reed?* Kennelijk hebben de makers van de flitspalen voldoende vertrouwen om 'ja' te zeggen als hun meetapparatuur aangeeft dat iemand meer dan 103 km/u reed op een weg waar je 100 km/u mag rijden.

Is het mogelijk dat je daardoor bestuurders *niet* bekeurt terwijl ze *wel* te hard reden? Jazeker! En naarmate je grotere marges neemt, wordt die kans groter.

We hebben hier te maken met vier mogelijkheden.

	bekeuring	geen bekeuring
Reed te hard	terecht	?
Reed niet te hard	onrechtvaardig	gewenst

Van dit kwartet aan opties hebben we er aan drie een waardeoordeel gehangen door dat in te vullen in het schema.

- Het is wenselijk dat niemand te hard rijdt en dat niemand een bekeuring krijgt. Eigenlijk zou je die hele flitspaal niet willen hebben.
- Het idee van bekeuringen is dat je iedereen een prent wilt geven die te hard rijdt. Vandaar het woord 'terecht'.
- Maar wat nu als je keurig 97,6 km/u reed en je kreeg toch die boete? Dat is oneerlijk en moet voorkomen worden.
- Iemand reed te hard en wordt niet bekeurd. 'Die heeft geluk gehad', zal iemand zeggen, maar is dat ook zo? Stel dat je telkens wegkomt zonder bekeuring en doorgaat met gevaarlijk rijgedrag…

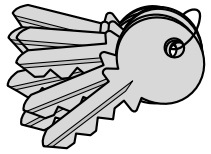
Oefeningen

1. Verbeter alle fouten in de cursieve tekst (punten 1 en 2) van de website van het Openbaar Ministerie.
2. Wordt iemand die 107,2 km per uur rijdt bekeurd als zij of hij geflitst wordt?

3. De sleutel

Toen ik in de vijfde klas van het vwo zat (1985/'86, Het Christelijk Lyceum Dordrecht) kregen we bij Wiskunde A een sommetje over statistiek.

Je hebt een sleutelbos met acht sleutels en eentje daarvan past op de deur waar je voor staat. Omdat je niet weet welke het is, ga je het domweg proberen. Hoe groot is de kans dat de vijfde sleutel die je probeert de juiste is?



Ik wist hoe je dat aanpakte. Je moest weten hoe groot de kans is dat een willekeurige andere sleutel paste. Dat was de kans dat de eerste sleutel niet goed was, en die moest je nou net hebben, want de som ging over het geval dat de *vijfde* goed was. Niet de eerste dus. Vervolgens berekende je van het setje dat overbleef opnieuw wat de kans op een verkeerde sleutel was, en dat vermenigvuldigde je daar dan mee. Per stap werd het steeds waarschijnlijker dat je de juiste sleutel had, want alles wat niet paste, viel af. (Tenzij je dom en/of dronken was, maar we gingen uit van iemand die niet dezelfde sleutel nog een keer probeerde. "Kijken of-ie het inmiddels wel doet.") Dat kon je netjes uitschrijven en dat deed ik ook. Het ging om de *vijfde* sleutel… We moeten dus eerst *vier keer* de kans nemen dat je een *verkeerde* sleutel te pakken hebt en dan *één keer* de kans op de *goede*. Nadat we vier keer misgegrepen hebben, blijven er nog vier sleutels over. De vijfde poging heeft dan een kans van 1 op 4 om de juiste te zijn.

$$P = \frac{7}{8} \times \frac{6}{7} \times \frac{5}{6} \times \frac{4}{5} \times \frac{1}{4}$$

Bij het vereenvoudigen van die breukberekening blijkt zo'n beetje alles tegen elkaar weg te delen! Alleen de noemer van die eerste factor (8) en de teller van die laatste factor (1) blijven over. Sterker nog: je zou dit voor willekeurige aantallen sleutels (N dus) kunnen doen. De kans dat de vijfde — of willekeurige andere – sleutel de goede is, blijkt gelijk te zijn aan $1/N$.

Ik herinner me nog de glimlach op mijn eigen gezicht destijds. Want wat had ik moeilijk gedaan over zo'n simpele vraag! De kans dat de vijfde sleutel aan een bos van acht de juiste was, is natuurlijk gewoon *een achtste*. Dat zie je in één keer als je beseft dat alle sleutels even waarschijnlijk zijn.

De omslachtige oplossing staat tegenover de aanpak van het gezonde verstand. In dit boek geven we aan beide zaken aandacht. De kans dat die gevraagde sleutel de juiste was, kun je bepalen

- via een formele som,
- met gezond verstand,
- in een Pythonprogramma.

Oefeningen

1. Hoe groot is de kans dat bij de sleutelbos uit dit hoofdstuk één van de eerste vijf sleutels die je probeert de juiste is?

2. Bepaal de kans dat de eerste sleutel de juiste is voor een bos van acht, twee, oneindig veel, één en nul sleutels. Gebruik in je antwoord niet alleen getallen, maar ook termen als 'onmogelijk' of 'zeker'.

3. Bij een hbo-opleiding met een instroom van zo'n 120 studenten per jaar en een uitvalspercentage van bijna 60% staan er 389 studenten ingeschreven. Hoe groot is de kans dat twee of meer studenten van deze opleiding op dezelfde dag jarig zijn?

Een man loopt 's nachts op straat te zoeken naar zijn huissleutel, die hij heeft laten vallen. Hij tuurt en tiert en loert en vloekt en baalt en haalt mismoedig zijn schouders op.

"Die ga ik nooit meer vinden in deze vieze, vuile, verrekte..."

Nog net voordat hij een héél lelijk woord wil gebruiken, hoort hij achter zich de stem van een goede vriend.

"Wat is er met jou aan de hand, joh?"

"Ik heb mijn sleutel laten vallen
en kan hem niet vinden."

"Zal ik je helpen zoeken?"

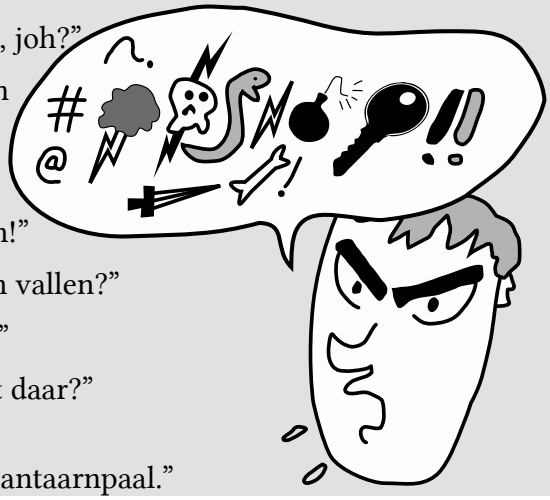
"Nou, als je dat zou willen doen!"

"Waar heb je 'm ongeveer laten vallen?"

"Een meter of zes die kant op..."

"Maar waarom zoek je dan niet daar?"

"Het is verderop nogal donker.
Hier heb ik meer licht van die lantaarnpaal."



4. Doellijntechnologie

Feyenoord heeft in Rotterdam een enerverende eredivisietopper tegen PSV met 2-1 gewonnen. Er zal nog lang nagepraat worden over het winnende doelpunt, dat acht minuten voor tijd op basis van doellijntechnologie werd toegekend. Nadat Jan-Arie van der Heijden na veel duw- en trekwerk los was gekomen van Héctor Moreno, zag hij zijn kopbal op de lijn gestopt worden door PSV-doelman Jeroen Zoet. Feyenoord claimde tevergeefs een goal: het horloge van scheidsrechter Bas Nijhuis (waarop doellijntechnologie zichtbaar is) gaf geen doelpunt aan. Bij het opstaan rolde Zoet de bal echter naar achter en daarbij passeerde die alsnog de doellijn. Toen pas ook ging het horloge van Nijhuis af. Het ging echter wel om millimeters.

Bron: NOS, 26 februari 2017

Ging het inderdaad om millimeters? Het artikel zegt van wel, maar het horloge van Bas Nijhuis deed daar geen uitspraak over. Daar stond het Engelse woord GOAL in hoofdletters, omdat het systeem een doelpunt vaststelde.

Wat zou er gebeurd zijn als dat doelpunt in Rotterdam-Zuid niet opgemerkt en toegekend was? Dan was het (op dat moment) geen 2-1 geworden en misschien zelfs 1-1 gebleven. Geen winst voor de thuisclub (en op dat moment koploper), maar een gelijkspel.

De eindstand in de competitie van het seizoen 2016/2017, waarin deze wedstrijd plaatsvond, zag er als volgt uit.

			W	V	G	+	−	+/-	pt.
1	Feyenoord	34	26	4	4	86	25	61	82
2	Ajax	34	25	3	6	79	23	56	81
3	PSV	34	22	2	10	68	23	45	77

Alle clubs hadden 34 wedstrijden gespeeld. Iedere overwinning (kolom W) leverde drie punten op en elk gelijkspel (kolom G) was één punt waard. Het aantal doelpunten voor en tegen en het doelsaldo (verschil tussen voor en tegen) was ook berekend. Hoe zou dit lijstje eruit hebben gezien als het 1-1 was gebleven? De nummer 1 had dan één voordoelpunt en één overwinning minder gehad en de nummer 3 kreeg juist één tegendoelpunt minder en leed één nederlaag minder. Beide clubs hadden dan één keer meer gelijkgespeeld.

De eindstand had er dan nogal anders uitgezien.

Ondertussen is het maar de vraag of het 1-1 gebleven zou zijn als dat doelpunt niet gevallen was. In de laatste minuten kan het soms raar lopen als de belangen groot zijn. En ook die tien wedstrijden die er nog te spelen waren, zouden wellicht anders verlopen zijn als deze uitslag anders was geweest.

'Wat er gebeurd zou zijn als' is een uitspraak die zelden te bewijzen of te controleren is. Had je geen lekke band gehad, dan was je toch nog op tijd gekomen voor het tentamen. Maar misschien was je dan wel geschept door een motorfiets die uit de bocht vloog. Meten is weten en zodra je iets meet dat van belang is, doet dat iets met je. Je gaat anders voetballen, wat harder fietsen, meer of minder rekening met dingen en mensen houden.

Wat was eigenlijk de onzekerheid in het meetsysteem? De KNVB claimde dat de foutmarge van doellijntechnologie een millimeter is. Peter de With, hoogleraar videotechnologie van de TU Eindhoven, trok dat in twijfel. Hij dacht eerder aan een centimeter. 'In een laboratoriumomgeving kun je het testen, maar dan zijn de doelpalen recht en is de doellijn nauwkeurig', zei hij tegen Omroep Brabant. 'In de werkelijkheid is dat anders. Een doellijn wordt getrokken met een kalkmachine, op gras dat niet glad is en waarvan de sprietjes alle kanten op staan. Doelpalen zijn ook nooit honderd procent verticaal en recht. Daardoor is de berekening van de doelpositie veranderlijk, waardoor de doellijn in het model verschuift.'

Een definitieprobleem dus. Waar staat het doel precies en hoe loopt die lijn? Niet zozeer het *meetapparaat*, maar *datgene wat je meet* is onzeker.

Een interessante vraag is wat je als technicus doet met die onzekerheid. In het geval van doellijntechnologie: als de bal een halve centimeter over de doellijn 'gemeten' wordt, en je onzekerheid is een hele centimeter, telt het doelpunt dan wel of niet? Het kampioenschap kan er zomaar van afhangen…

Oefening

Je werkt als mechatronicus voor de KNVB en krijgt de opdracht om een nieuw systeem voor doellijntechnologie te realiseren. Het systeem dat je ontwerpt, kent een onzekerheid van 2 cm. Je zou daarvoor kunnen corrigeren door de software zo te schrijven dat het horloge pas 'GOAL' roept als de bal minstens 2 cm over de lijn is. Je kunt ook *niet* corrigeren en de onzekerheid voor lief nemen. Of nog iets anders doen.

Maak een keuze tussen de mogelijke oplossingen en motiveer die.

5. Model & Werkelijkheid

"All models are wrong, but some are useful."

George Box

Bovenstaande tegeltjeswijsheid is afkomstig van George E. P. Box. Zijn uitspraak kwam voor het eerst op papier in 1976, in een wetenschappelijke publicatie in *the Journal of the American Statistical Association*.

Dat woord *wrong* doet denken aan fouten die je hebt bij metingen. En daar zit wel een overeenkomst in. Bij metingen zeggen we dat we te maken hebben met een *fout*, hoewel we het in dit boek meestal (en liever) over een *onzekerheid* hebben. Dat is niet hetzelfde.

De meting die je doet, komt niet exact overeen met de werkelijkheid. Ook als je niets verkeerd doet. Een telling kan exact kloppen, maar de meting van een lengte, massa, tijd, spanning, druk of wat dan ook, heeft altijd een beperking in zich. Je zou de uitspraak van Box een beetje kunnen aanpassen.

"All measurements are wrong, but some are useful."

Het *some* is in beide uitspraken een beetje cynisch of komisch bedoeld. Het is niet zo dat er maar een paar nuttige modellen of metingen bestaan. Maar het is belangrijk om te zien dat ze niet een exacte weergave van de werkelijkheid zijn.

Een leraar van me op de middelbare school zei het anders.

> *"Alle modellen zijn minder dan de werkelijkheid, behalve fotomodellen. Die zijn méér dan de werkelijkheid."*

Soms tref je in Meten & Weten hoofdstuktitels aan met twee woorden en een & ertussen. Beide helften beginnen dan met een hoofdletter.
In dat geval worden er twee begrippen naast elkaar gezet die écht iets anders betekenen en goed uit elkaar moeten worden gehouden.

Het Ene & Het Andere is niet hetzelfde, maar enigszins anders.
Of heel anders. Dat zit 'm in de verschillen…

6. Meten is weten dat…

Tegeltjeswijsheden zijn vaak onzinnig, maar soms raak. ‘Als het regent in mei, is april voorbij’, zei André van Duin ooit. Of, minstens zo waar: ‘Als de haan niet kraait bij morgenrood, krijgen we regen… of de haan is dood.’

‘Meten is weten’ is ook zo’n tegeltjeswijsheid. Maar het is NIET de titel van dit boek. En het is ook niet wat Kamerlingh Onnes ooit sprak in Leiden: ‘Door meten tot weten.’

Wat zou je op een tegeltje moeten zetten om het waarheidsgehalte wat op te krikken, tot het niveau van regen in mei en al dan niet dode hanen? Laten we een poging wagen. (Schrik niet als het een wat lange tekst wordt…)

Meten is weten dat
de werkelijke waarde
met een bepaalde
waarschijnlijkheid
hoogstens een zekere
onzekerheid
verschilt van je
meetresultaat

meten is weten

Meten *is* natuurlijk niet hetzelfde als weten, maar *door* te meten, weet je iets. Daar had Kamerlingh Onnes gelijk in: het is *door meten tot weten.*
Als je meet, moet je wel weten dat wat je *meent* te weten, betrekkelijk is.
Of in ieder geval een onzekerheid heeft.

werkelijke waarde

De *werkelijke waarde* is wat je weten wil en het is óók belangrijk om te beseffen dat je die niet kent én dat er soms niet eens sprake is van een werkelijke waarde, omdat datgene wat je meet *varieert* of *niet scherp genoeg gedefinieerd* is. Dat is niet erg, maar je moet je er wel van bewust zijn.

een bepaalde waarschijnlijkheid

Onzekerheden beschrijven geen *harde grenzen* waar een werkelijke waarde tussen ligt, maar geven slechts een *bereik* aan waar die waarde *waarschijnlijk* te vinden is. Dat lijkt gek: waarom geef je nou niet gewoon de uitersten op? Het antwoord: als je 100% garantie wilt, krijg je wel héél grote onzekerheden.

Het is een *bepaalde* waarschijnlijkheid omdat iets of iemand *bepaald* heeft dat pakweg 2/3 of 68% wel aardig is.

hoogstens

Het woord *hoogstens* kan verwarring scheppen. De onzekerheid geeft namelijk geen harde onder- of bovengrens. Je geeft alleen maar aan dat de waarde waarschijnlijk niet hoger is. *Waarschijnlijk hoogstens* dus.

een zekere onzekerheid

Dit wordt spelen met woorden, maar *een zekere* onzekerheid zegt eigenlijk dat je er helemaal niet zo zeker van bent. 'Een of andere onzekerheid…'

Onthoud: er is altijd onzekerheid over de werkelijke waarde en meestal een nog grotere onzekerheid over de onzekerheid.

verschilt

Op een eerder tegeltje stond 'afwijkt', maar dat is een beetje de wereld op z'n kop. De werkelijke waarde wijkt toch niet af? Het is *de meting* die afwijkt…

je meetresultaat

Dat woordje 'je' geeft aan dat anderen iets anders meten. En 'meetresultaat' is dubieus, want het is eigenlijk *de beste schatting*. Strikt genomen is een meetresultaat een *beste schatting* mét een *onzekerheid*. Het nadeel van de term 'beste schatting' is dat die niet duidelijk maakt dat meten iets anders is dan schatten.

Meten is geen schatten. Het is ook geen weten. Nou ja, het is *weten dat de werkelijke waarde met een bepaalde waarschijnlijkheid hoogstens een zekere onzekerheid verschilt van je beste meetresultaat, of beste schatting.*

7. SI: Standaardiseer Iets!

In het voorjaar van 1999 verongelukte de *Mars Climate Orbiter* kort na aankomst bij de rode planeet. Oorzaak: software in pounds en inches, hardware in newton en meter. NASA en Lockheed-Martin hadden niet afgesproken in welke eenheden ze rekenden. De sonde scheerde langs Mars en verdampte in de dampkring. Tachtig miljoen dollar foetsie.

Meten is vergelijken. Maar vergelijken lukt pas als je het eens bent over de maat. Als kind leer je al wat een meter is, of een seconde. 'Een kilo' zelfs. Maar waarom je niet gewoon meet in 'blokken', 'knopen', of 'eenheden per scherm', daar hoorde je niemand over. Terwijl *afspreken* de kern is van elk meetsysteem. Dat een kilogram hetzelfde betekent in Korea als in Kroatië. Dat een seconde in de ruimte net zo lang duurt als in Rotterdam. En dat het niet uitmaakt of je de hoogte van een satelliet of van een stapel boeken meet. Je vergelijkt die hoogte ergens mee, namelijk met *één meter*. "Die satelliet bevindt zich 35786 keer zo hoog boven de aarde als de lengte van een meter." En: "Die stapel boeken is de helft zo hoog als één meter lang is." Maar ja, wat is dan een meter? Daar moet je het *exact* over eens zijn.

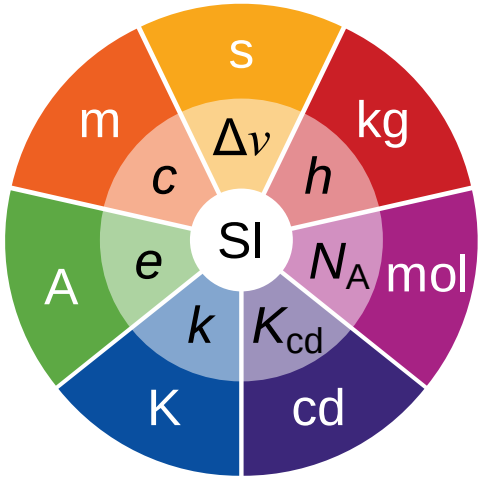
Sinds het einde van de achttiende eeuw proberen deskundigen het wereldwijd daarover eens te worden. Dat leidde tot afspraken die we nu kennen als het SI — het *Système International d'Unités*. Daarin zijn zeven grootheden als fundamenteel gekozen: lengte, massa, tijd, elektrische stroom, temperatuur, hoeveelheid stof en lichtsterkte. Elk van die zeven heeft een bijbehorende eenheid met een internationaal afgesproken naam en symbool: meter (m), kilogram (kg), seconde (s), ampère (A), kelvin (K), mol (mol) en candela (cd).

Je zou denken dat zeven weinig is, maar we redden het ermee. Alle andere grootheden zijn van deze zeven afgeleid. Denk aan kracht (N), energie (J), druk (Pa), spanning (V), weerstand (Ω), vermogen (W), versnelling (m/s^2)… Ze zijn allemaal te herleiden tot combinaties van die zeven.

> De kilogram is trouwens een raar geval. Alle andere basiseenheden heten gewoon "meter" of "seconde", zonder franje. Maar bij massa is het niet de gram die de basis vormt, maar de kilogram. Dat zit historisch zo. In 1795 werd in Frankrijk de gram gedefinieerd als het gewicht van één kubieke centimeter water. Later bleek dat veel te klein voor praktische toepassingen: je moest steeds met duizendtallen rekenen. Daarom werd de kilogram — duizend gram — de standaard. En toen in de twintigste eeuw het SI werd opgezet, kozen ze ervoor om dat zo te laten. Gevolg: het enige SI-basiselement dat begint met een voorvoegsel.

Toch blijft het vreemd, die zeven. Ze lijken willekeurig. Waarom *deze* zeven? Waarom geen coulomb in plaats van ampère? Waarom de candela, en geen lux of lumen? Het antwoord: omdat die zeven te baseren zijn op natuurconstanten. Die hoef je niet af te spreken. Die liggen al vast. Moeder Natuur of de Schepper heeft die ooit gekozen. Toen zij of Hij het Universum maakte.

Een ampère is gedefinieerd via de krachtwerking tussen twee evenwijdige draden met stroom. Een coulomb lijkt fundamenteler, maar is in feite abstracter: een hoeveelheid lading die je pas kunt meten als je weet hoeveel stroom er loopt en hoe lang. Zo is ook de volt een gevolg van de ampère, de meter, de kilogram en de seconde. Het SI is opgebouwd volgens wat je natuurkundig kunt vastleggen. Zo is de seconde gekoppeld aan de trilling van een cesiumatoom. Eenheden als *bar*, *liter* of *graad Celsius* zijn herkenbaar voor de meeste mensen, maar behoren niet tot het SI.



De basiseenheden zijn de afgelopen decennia vernieuwd en aangescherpt. Tot ver in de twintigste eeuw was de kilogram een echt bestaand voorwerp: een metalen cilinder in een kluis bij Parijs. En de meter was één bepaalde staaf van platina-iridium. Maar sinds 1983 is de meter gedefinieerd aan de hand van de lichtsnelheid, namelijk 1/299792458^{e} deel van de afstand die licht in vacuüm aflegt in één seconde. En sinds 2019 is de kilogram niet langer afhankelijk van dat ene cilinderblokje, maar van de Planckconstante. Daarmee zijn tegenwoordig alle zeven basiseenheden terug te voeren op natuurwetten. Je hebt er dus geen voorwerpen meer voor nodig. Alleen nog een verzameling begrippen en constanten. En goede apparatuur.

Het SI is een afspraak die werkt voor wie zich eraan houdt. Met zeven bouwstenen kun je een huis maken, een brug ontwerpen en een raket besturen. Maar als jij en je collega's allemaal een ander woordenboek gebruiken, komt die raket nooit van de grond. Of hij verdwijnt in de dampkring van Mars.

Hoe kan dat nou, dat iets wat wordt uitgedrukt in een joule per coulomb — een spanning dus — volledig terug te voeren is op niet meer dan zeven basiseenheden? Laten we beginnen met de definitie: een volt is dus een joule per coulomb. Energie per lading. Een joule is nog wel te volgen, want arbeid is kracht maal weg, en energie is de arbeid die je kunt leveren. Kracht is massa keer versnelling — dat dus vermenigvuldigd met de afstand. Uiteindelijk krijg je dan: *kilogram maal meter kwadraat per seconde kwadraat*. Maar die coulomb? Die staat helemaal niet in het rijtje van fundamentele grootheden! En toch kun je ook die herleiden. De truc zit in de ampère: die is namelijk wél een SI-basisgrootheid. Een ampère is een coulomb per seconde, dus een coulomb is niets anders dan een ampère maal een seconde. Lading als stroom × tijd. Vul je dat allemaal in, dan krijg je: volt = $(kg \cdot m^2 \cdot s^{-2}) / (A \cdot s) = kg \cdot m^2 \cdot s^{-3} \cdot A^{-1}$. Zo blijkt een volt, die zo vertrouwd klinkt, gewoon een samenstelling van vier van de zeven basisgrootheden: massa, lengte, tijd en stroomsterkte.

Nu we toch bezig zijn: hoe zit dat met de ohm? Weerstand is spanning gedeeld door stroom. Daar hoort als eenheid een volt per ampère bij, dus je deelt wat we voor de volt vonden door A: $kg \cdot m^2 \cdot s^{-3} \cdot A^{-2}$. Weerstand, spanning, vermogen, lading: het lijken zelfstandige begrippen, maar ze vallen allemaal terug op die ene compacte set van zeven afgesproken grootheden. Een ohm is eigenlijk een $kg \cdot m^2 \cdot s^{-3} \cdot A^{-2}$. Een ohm is een oom die je jarenlang niet hebt gezien. Onherkenbaar...

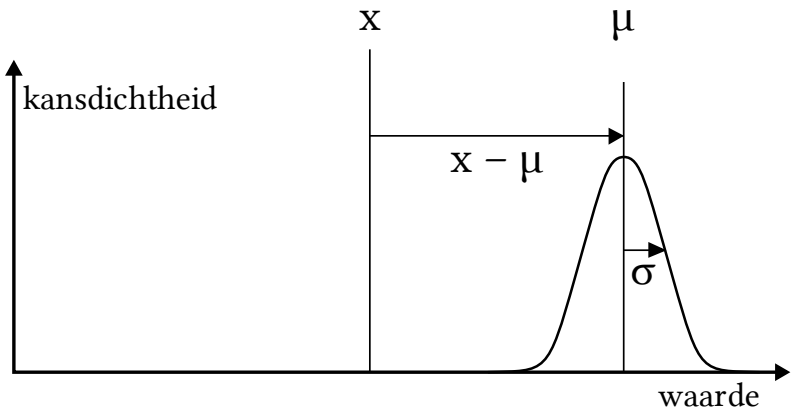
Zo'n ohm gaat nog wel, maar wat is eigenlijk een weber? Het klinkt als een merk barbecue, maar het is de SI-eenheid van magnetische flux. Wat dat precies is, weet je pas als je het hebt opgezocht. En dan nog blijft het vaag: magnetisme dat ergens doorheen stroomt, of zoiets. Maar ook hier zit een verhaal onder de motorkap. Een weber is gedefinieerd als een volt maal een seconde. Tijd en spanning in één pakketje.
Jetzt geht's los! We gaan voor de tesla! Niet die van Musk, maar die andere. Eerst de weber. (Niet Marianne...)

Je zou zeggen: weber = $(kg \cdot m^2 \cdot s^{-3} \cdot A^{-1}) \times s = kg \cdot m^2 \cdot s^{-2} \cdot A^{-1}$. En kijk eens aan: daar heb je ze weer: massa, lengte, tijd en stroomsterkte. Vier van de zeven. Helemaal volgens het recept. En de tesla? Die is één weber per vierkante meter. Dus: $kg \cdot s^{-2} \cdot A^{-1}$. Magnetische flux (weber) is veldsterkte (tesla) maal oppervlak, net zoals arbeid kracht maal afstand is. Al die begrippen blijken gewoon variaties op hetzelfde riedeltje. De weber hoort bij de volt zoals de joule bij de newton hoort. Een weber is geen mysterie meer als je hem uit elkaar haalt.

Feilen

8. Model van een meting

Figuur 8.1 toont een klassiek schema dat duidelijk maakt hoe systematische en toevallige fouten zich tot elkaar verhouden.

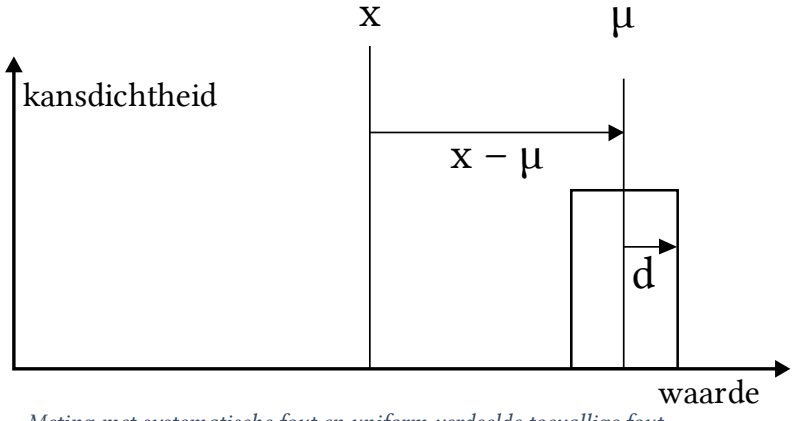


Figuur 8.1 Meting met systematische fout en normaal verdeelde toevallige fout

Deze figuur is overgenomen van *Wikipedia*, met aanpassingen: de woorden *Juistheid* en *Precisie* zijn vervangen door symbolen.

De waarde die je wilt meten (de werkelijke waarde of de *true value*) is x.

De meetwaarde zelf is een *kansvariabele*, normaal verdeeld met een gemiddelde μ en een standaardafwijking σ. In deze figuur wordt de meting beschouwd als een normaal verdeeld proces, maar die verdeling kan ook uniform zijn, bijvoorbeeld bij een digitaal meetapparaat dat een toevallige fout heeft die afhangt van de decimalen die je niet kunt aflezen.



Figuur 8.2 Meting met systematische fout en uniform verdeelde toevallige fout

Laten we voor het gemak eens uitgaan van een meting aan een dimensieloze (vage, abstracte) grootheid x met eenheid *nono* (symbool: n), waarvoor geldt

$$x = 1{,}98\, n$$

Het meetresultaat dat we vinden is

$$x_{m1} = 2{,}01 \pm 0{,}04\, n$$

De absolute onzekerheid in dit meetresultaat is 4 cn (centinono).
De relatieve onzekerheid in dit meetresultaat is 2%.

Weten we nu hoe groot het verschil is tussen x en μ? Nee. In de tekst hierboven is *gegeven* dat $x = 1{,}98\, n$. In een 'echte' situatie weet je dat vrijwel nooit. Je kent soms een *geaccepteerde waarde*, maar slechts zelden de *werkelijke waarde*. Waarom zou je nog meten als je de waarde al kent?

> De makers van de GUM schrijven op pagina 16:
>
> *NOTE The term "true value" is not used in this Guide for the reasons given in D.3.5; the terms "value of a measurand" (or of a quantity) and "true value of a measurand" (or of a quantity) are viewed as equivalent.*
>
> Dat is een vreemde bewering. Ze noemen datgene wat ze *niet* gebruiken niet alleen twee keer in deze ene alinea (om te zeggen dat ze het niet gebruiken), maar als je telt hoe vaak de term *true value* voorkomt in hun hele *manual*, kom je op 55 keer. Dat is toch niet heel weinig in een document van 134 pagina's. In ieder geval vaker dan nooit.

Kennen we μ eigenlijk wel? Nee. We hebben een waarde x_{m1} gevonden. Die x_{m1} is het steekproefgemiddelde $\bar{x}$ van deze ene steekproef van tien metingen. Doe je nog een keer tien metingen, dan vind je een x_{m2} die — als de toevallige fout niet nul is — anders zal zijn. Naarmate de standaardafwijking σ groter is, zal de spreiding in x_{m1}, x_{m2}, x_{m3}, etc. groter zijn.

Kennen we σ dan wél? Nee. Die is ook al niet bekend in praktische situaties. We kunnen wel een schatting ervan maken door vaak te meten en dan de standaardafwijking van de steekproef te bepalen. Als je oneindig vaak meet, wordt die gelijk aan de 'echte' σ. Is er eigenlijk wel een 'echte' σ? Nee. En er is ook geen 'echte' μ. We veronderstellen een normale verdeling, maar dat is slechts een model. Maar laten we toch maar doen alsof er wél een echte μ en σ bestaan. Dan kennen we die nog lang niet. En de echte x kennen we alleen maar doordat we die zelf verzonnen hebben in dit hoofdstuk. Op grond van de metingen weten we alleen: $x_{m1} = 2{,}01 \pm 0{,}04\, n$.

Hoe is die onzekerheid van $0{,}04\ n$ eigenlijk bepaald? Daar hebben we het nog niet over gehad. Er is een constant deel en een toevallig deel in die onzekerheid. Het constante deel kan niet uit de metingen worden afgeleid, maar zou gegeven kunnen zijn in specificaties van de meetapparatuur. Het toevallige deel kan volgen uit die specificaties. Het toevallige deel kan óók worden gevonden door de metingen te herhalen en te kijken wat de spreiding is. Door oneindig vaak te herhalen, kun je de kansverdeling bepalen.

Er is een belangrijk aspect in het soort plaatjes aan het begin van dit hoofdstuk dat meestal onbenoemd blijft. Die plaatjes geven namelijk de verdelingsfunctie van één bepaalde meting voor één bepaalde grootheid.

We hadden gevonden als meetresultaat: $x_{m1} = 2{,}01 \pm 0{,}04\ n$.

De absolute onzekerheid was 4 cn en de relatieve onzekerheid was 2%. Hoe groot zullen nu de absolute en relatieve onzekerheid zijn in een grootheid die twee keer zo groot is? En wéér dat irritante antwoord: dat weten we niet.

Terug naar Figuur 8.1 waar we mee begonnen. Zou onze meting echt normaal verdeeld zijn? Nee. Als je een tijdsduur meet en je komt uit op gemiddeld 1,5 s en een standaardafwijking van 0,3 s, dan zou er een (kleine) kans zijn dat die tijd negatief is. En negatieve tijdsduren bestaan niet.

> Dat zou mooi zijn, zeg! Een negatieve tijdsduur… Je zit op de fiets naar school en komt waarschijnlijk vijf minuten te laat voor het tentamen. Je snijdt een flink stuk af via een route waar je *een negatief kwartier* over doet… Dan heb je toch nog tien minuten speling!

Die verdeling is dus niet ‘echt’ normaal. En als die verdeling dat wel was, kenden we nog steeds die μ en σ niet. Er is nog meer aan de hand. Die onzekerheid kan *per gemeten waarde* verschillen. Op die 1,5 s zit een onzekerheid van 0,3 s. Dat is 20%. Zit er dan op een tijdsduur van 2,5 s óók een onzekerheid van 0,3 s? Of óók een onzekerheid van 20% en dus 0,5 s? Misschien is het ene waar, misschien het andere… en misschien ook allebei niet.

> Het zou kunnen dat de onzekerheid op die meting gelijk is aan 0,15 s + 10%. Het zou zelfs zo kunnen zijn dat er een kwadratisch of logaritmisch verband is tussen het gemiddelde en de standaardafwijking. Theoretisch kan er heel veel. In de praktijk zijn we al blij als het ongeveer normaal (of uniform) verdeeld is en we daar wat getallen bij kunnen schatten.

Wat is hier mis met het model dat we gebruiken? Op zich niet zo veel, maar je moet je altijd realiseren dat die normale verdelingen met z’n μ en σ *slechts een model zijn*. En dat George Box gelijk had: elk model is fout.

9. Systematisch & Toevallig

Als we uitgaan van een goed gedefinieerde grootheid, een correcte meting en geen vreemde (of grote) spreidingen in de gemeten grootheid, dan blijven er eigenlijk maar twee soorten onzekerheid over.

- de onzekerheid als gevolg van een toevallige fout
- de onzekerheid als gevolg van een systematische fout

Het toevallige deel verschilt per meting. Het systematische deel is altijd hetzelfde: je meet altijd te veel of te weinig en de hoeveelheid varieert niet.

Een misverstand dat je soms tegenkomt, is dat alleen het toevallige deel als 'onzekerheid' wordt gezien.

Laten we om dit te illustreren nog eens naar zo'n beginzin kijken, uit het boek van Barford (1985).

Words like error, accurate, truth, probable, likelihood, are part of everyday language. For most of us their meanings may not be mathematically precise, but the ideas behind them reflect experience we have gained from an early age.

Barford kiest zijn uitgangspunt bij de alledaagse ervaring van begrippen. Op zich niets mis mee. Maar dan gaat hij verder.

One way of approaching the subject matter of this book is to step aside from this experience and to talk of tossing coins and rolling dice, of the laws of probability that these activities demonstrate and the statistical predictions that may be derived from them.

En dat is een wat merkwaardige beperking. Munten, dobbelstenen, kansberekening en statistiek hebben allemaal met 'toeval' te maken. Maar er zit (soms) ook iets niet-toevalligs in de fout. Nou ja, misschien is het wel *toevallig*, maar *varieert het niet.* Als hij bij elke meting even groot is, kun je dat niet met statistische theorieën over munten en dobbelstenen komen. (Tenzij je dobbelstenen hebt waarmee je altijd 'zes' gooit.)

Dat alleen het toevallige deel soms ten onrechte als onzekerheid beschouwd wordt, zal twee oorzaken hebben.

- Iets wat de hele tijd gelijk blijft, komt niet over als onzeker.
- Welk deel fout is én de hele tijd gelijk blijft, wordt niet zichtbaar door de meting te herhalen.

Berendsen (2011) geeft in zijn boek een driedeling die we niet overnemen.

3.1 Classification of errors

There are several types of error in experimental outcomes:

(i) (accidental, stupid or intended) mistakes

(ii) systematic deviations

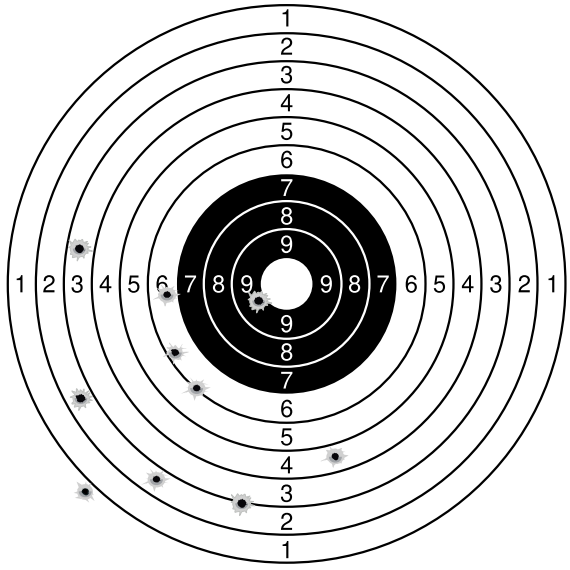
(iii) random errors or uncertainties

The first type we shall ignore. Accidental mistakes can be avoided by careful checking and double checking. Stupid mistakes are accidental errors that have been overlooked. Intended mistakes (e.g. selecting data that suit your purpose) purposely mislead the reader and belong to the category of scientific crimes.

> S*tupid*? Slim is een fout zelden, maar meestal ook niet stom. Niet toegeven is dat wel, maar een fout maken is vergefelijk en leerzaam en niet erg. De fout zelf kan wel stom zijn. Degene die hem maakt is dat niet.

Als mijn personenweegschaal steeds iets anders aangeeft als ik erop en eraf stap, weet ik heus wel dat mijn gewicht (massa) niet met een paar ons verandert in die korte tijd. Dat toevallige deel zie ik wel. Maar dat ik er de hele tijd een procent naast zit, blijkt nergens uit. (Behalve dan dat ik heus wel lichter ben dan wat dat ding zegt… Een paar jaar geleden gaf hij minder aan.)

Toevallige en systematische fouten worden wel eens in kaart gebracht en uitgelegd aan de hand van schietschijven en pijlen die in de roos belanden of juist niet. Die komen we in de loop van dit boek vaker tegen.



Figuur 9.1 Metafoor van toevallige en systematische fouten: de schietschijf

> In de vele plaatjes op het internet wordt er gebruik gemaakt van ronde, tweedimensionale schietschijven. Waarom dan? Meten doe je meestal aan een *scalaire* grootheid, niet aan een vector. Alleen de afstand tot de

roos heeft dus betekenis. Maar ja: een ééndimensionale schietschijf… Dan is de kans dat raak schiet wel erg klein.

Er zit een spreiding in de kogelgaten, veroorzaakt door de schutter en/of het geweer. Die spreiding duidt op een *toevallige fout*. Dat de meeste kogels in het kwadrant linksonder zitten, gemiddeld genomen een stuk van de roos af, duidt op een *systematische fout*. Bij de schietschijf kun je beide zien:

- de toevallige fout zit in de grootte van de spreiding
- de systematische zit in de gemiddelde afstand tot de roos

Bij een meting hebben we een probleem dat we met de schietschijf niet hebben: we zien bij metingen wel een spreiding en weten hoe groot die is. Maar we weten niet hoe ver die afstand tot de roos is. Hier zit wat verwarring. Als we een tijd meten en beweren dat we een systematische fout hebben van 0,2 seconde, dan ben je geneigd om te zeggen: 'Man (m/v), corrigeer daar dan voor!' Maar je weet *wel* dat de grootte ongeveer 0,2 seconde is en *niet* of het nou 0,1 s of -0,2 s of 0,15 s of 0,25 s of -0,005 s is. Het zijn in ieder geval geen tien seconden, maar heus wel meer dan een nanoseconde.

We hebben dus een ruwe 'maat' voor die systematische fout. Je weet bijvoorbeeld dat je er grofweg 2% naast zal zitten. Maar je weet niet hoeveel precies. Het is een soort stelpost: je weet dat er wat aan mankeert, maar niet hoeveel.

We weten van de systematische fout niet eens of hij positief of negatief is. Als we wisten dat de afwijking in de tijd ergens tussen de 0,1 s en 0,3 s zou zitten, zouden we bij alle waarden 0,2 s optellen. Omdat we dan dichter in de buurt komen.

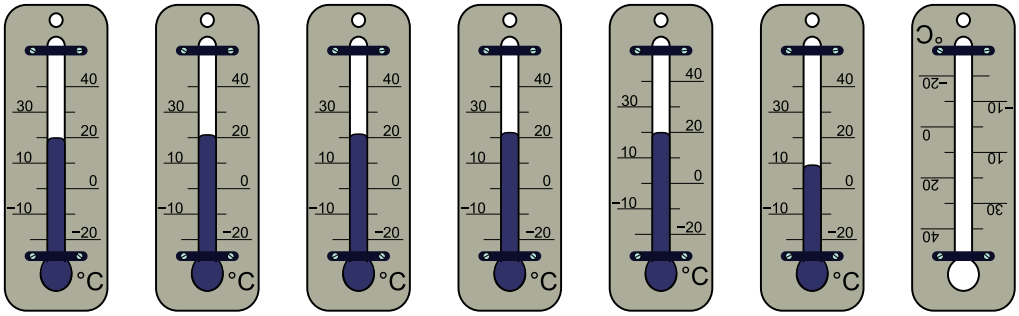
Er is nog iets te zeggen over systematische en toevallige fouten, namelijk dat je beter kunt spreken van een *deterministisch* en een *stochastisch* deel. Een stuk dat de hele tijd hetzelfde is een stuk dat bij elke meting anders kan zijn. Die systematische fout kan door toeval veroorzaakt zijn. Als dat toevallige deel de hele tijd hetzelfde is, noemen we het nog steeds systematisch. De waarde op zich is dan toevallig, maar wel constant. In het volgende hoofdstuk zien we daar een voorbeeld van.

Statistisch geformuleerd: als fouten normaal verdeeld zijn met een gemiddelde μ en een standaardafwijking σ, dan is

- de systematische fout gelijk aan μ
- de toevallige fout een normaal verdeeld proces met een gemiddelde van 0 en een standaardafwijking σ

10. Toevallige systematische fouten

Een systematische fout kan een toevallige grootte hebben. Die fout is dan bij elke meting hetzelfde (systematisch), terwijl de grootte een toevallige oorzaak heeft. Dat wordt duidelijker als we het volgende voorbeeld bekijken.



Figuur 10.1 Zeven in serie geproduceerde thermometers

We gaan een temperatuur meten met een thermometer. We hebben zeven exemplaren ter beschikking. Hoe zit het met de toevallige en systematische fouten bij het werken ermee?

Er zal een toevallige fout zijn als gevolg van het niet exact kunnen aflezen. Het stukje dat je 'ernaast zit', is willekeurig en daar zit dus spreiding in.

Er zal ook een systematische fout zijn als gevolg van het niet perfect kunnen fabriceren. De hoeveelheid vloeistof in het glazen buisje en de exacte positie van de streepjes van de temperatuurschaal zijn onvolmaakt. Het lukt de fabrikant niet om tot op het molecuul nauwkeurig af te vullen en lijntjes te trekken. Daar zullen procenten aan schommeling in zitten.

Die hoeveelheid vloeistof is dus iets te groot of iets te klein en we weten niet hoeveel. Maar hij is wel de hele tijd gelijk voor één bepaalde thermometer. De gevolgen van de fout zijn dus telkens even groot.

Laten we de zeven thermometers van Figuur 10.1 eens beter bekijken. De eerste vijf (van links) geven allemaal ongeveer 20 °C aan, maar er zitten wel kleine verschillen in, als gevolg van de hoeveelheid vloeistof die niet exact klopt en een bepaalde spreiding heeft. De zesde en zevende thermometer wijken veel af. De zesde staat op ongeveer 10 °C; dat komt doordat de vloeistof bijna op was toen deze gevuld werd. Bij de zevende was het helemaal op… Er zit dan ook geen druppel in. En tot overmaat van ramp zijn de streepjes ook nog eens op zijn kop afgedrukt!

Het voordeel van strepen die helemaal op hun kop staan, is dat je duidelijk ziet dat er iets misgegaan is bij de productie. Maar wat nu als de streepjes een klein stukje te laag staan, niet als gevolg van onvermijdbare onzekerheid, maar als gevolg van een fout bij het produceren? Eigenlijk heb je dan een meetinstrument dat gewoon niet goed is. Terugsturen en een nieuwe vragen. Maar hoe kom je erachter dat hij niet goed is, als het verschil zo klein is?

De twee meest rechtse thermometers kunnen we buiten beschouwing laten. Maar hoe zit het met die linkse vijf? Die zijn gewoon 'goed' en hebben een acceptabele spreiding. De specificaties geven ook aan dat de onzekerheid twee graden Celsius is. (Het waren goedkope thermometers, van de Action of AliExpress…)

Nu de essentie van dit verhaal. De hoeveelheid vloeistof verschilt per thermometer en is een gevolg van *toeval*. Maar ook al is de *oorzaak* toevallig, die hoeveelheid vloeistof *is voor elke individuele thermometer wel constant*. Je zou het afvullen als een stochastisch proces kunnen beschouwen, met de hoeveelheid vloeistof in elke thermometer als één concrete uitkomst daarvan.

De afwijking van de 'juiste' hoeveelheid vloeistof ligt na de fabricage vast, dus de fout is *systematisch*. Elke keer als je meet met dezelfde thermometer, is de onzekerheid als gevolg van deze afwijking hetzelfde. De situatie wordt anders als je vijf metingen doet en je gebruikt de linkse vijf thermometers alle vijf één keer. Elke meting heeft dan een andere afwijking ten gevolge van die variatie bij de productie. Dan is de fout die onveranderlijk is per thermometer opeens een toevallige fout.

> Als de hoeveelheid vloeistof te groot is, zal een te hoge temperatuur worden weergeven. Het aantal milliliters afwijking is vast, maar de afwijking in de meting als toenemen voor hogere temperaturen. Er zal een lineair verband zijn tussen de gemeten temperatuur en de werkelijke temperatuur (in kelvin!)

Oefeningen

1. Als je deze zeven thermometers ter beschikking hebt, hoe zou je daar dan het beste mee kunnen meten?
2. Als er in een thermometer 5% minder vloeistof zit dan erin zou moeten zitten, is er dan sprake van een onzekerheid of een productiefout?

11. Nauwkeurig & Precies

Als je wilt weten wat een woord betekent, helpt Google of een woordenboek. Dat laatste is betrouwbaarder, want via een zoekmachine weet je maar nooit waar je terechtkomt. Met wat mazzel kom je in de gratis *online Van Dale*.

We willen weten wat het verschil is tussen *nauwkeurig* en *precies*.

nauw·keu·rig (*bijvoeglijk naamwoord, bijwoord*) juist, zorgvuldig, stipt; = accuraat

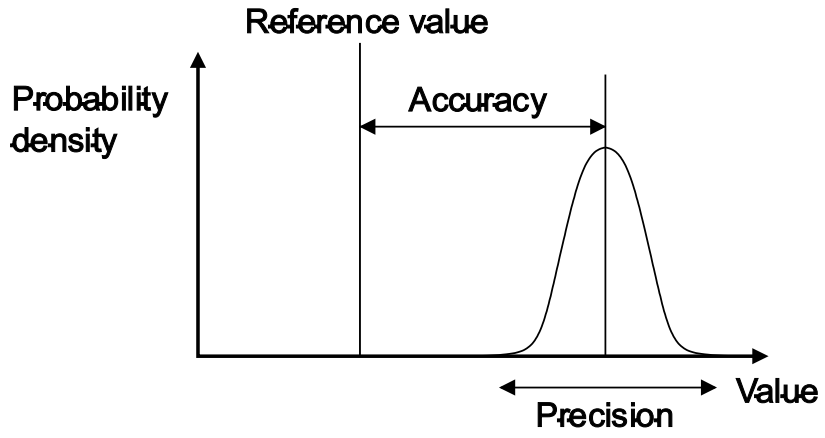
pre·cies (*bijvoeglijk naamwoord, bijwoord;* vergrotende trap: *preciezer*, overtreffende trap: *preciest*) zonder afwijking of verschil; juist, nauwkeurig, stipt: *precies honderd euro; hij is al te precies* nauwgezet

Waarom er bij 'precies' een vergrotende en overtreffende trap staat en bij 'nauwkeurig' niet, is niet duidelijk. Verwarrender is dat bij 'precies' het woord 'nauwkeurig' als synoniem staat, maar omgekeerd lijkt 'nauwkeurig' geen synoniem te zijn van 'precies'. Hoe zit dat precies?

In het vakgebied van Meten & Weten zijn nauwkeurigheid en precisie twee wezenlijk verschillende dingen.

- nauwkeurigheid is de mate waarin het gemiddelde van (oneindig) veel metingen overeenkomt met de werkelijke waarde
  anders gezegd: hoe dicht de metingen bij de werkelijke waarde liggen
- precisie is de mate waarin herhaalde metingen met elkaar overeenkomen
  anders gezegd: hoe dicht de metingen bij elkaar liggen

Een klassieker is het plaatje hieronder, van de Engelstalige *Wikipedia*.



Figuur 11.1 Het verschil tussen Accuracy en Precision volgens Wikipedia, grafisch weergegeven

De Engelse vertaling van 'nauwkeurigheid' is *accuracy* en van 'precisie' is het *precision*. Op grond van het plaatje zou je denken dat de nauwkeurigheid toeneemt als de verticale lijnen verder uit elkaar liggen en dat de precisie toeneemt als de klokvormige curve breder is. Dat is juist niet het geval. Het zou beter zijn om te spreken van onnauwkeurigheid en imprecisie. Als je zegt dat je 1,71 meter lang bent en dat vastgesteld hebt met een nauwkeurigheid van 1 cm, bedoel je dan dat je er maar liefst 1,70 meter naast kunt zitten? ("Misschien ben ik wel drie meter veertig, de nauwkeurigheid is nogal klein…")

Er is nog iets vervelends aan de hand. De termen in de Engelstalige afbeelding (en hun vertalingen in het Nederlands) zijn zeer gangbaar (geweest) en komen ook in veel boeken voor. De officiële termen zijn vastgelegd in ISO 5725. Wat in bovenstaand plaatje *accuracy* is (nauwkeurigheid), heet volgens die norm *trueness* ('juistheid'). Dat laatste overigens nog met een subtiele toevoeging: de pijl in het plaatje moet korter zijn, omdat we als getalswaarde aan de *precision* de standaardafwijking van de verdeling koppelen. De pijl bij het woord 'Precision' in Figuur 11.1 is meer dan vier keer zo groot.

Het wordt nog vervelender: de term *accuracy* (nauwkeurigheid) wordt in ISO 5725-1 gebruikt om te beschrijven hoe dicht een meting bij de werkelijke waarde ligt. Als er dan sprake is van een reeks metingen van dezelfde meetgrootheid, gaat het om

- een deel dat wordt veroorzaakt door willekeurige fouten en
- een deel dat wordt veroorzaakt door systematische fouten.

Dan is *juistheid* de mate waarin het gemiddelde van een reeks meetresultaten de feitelijke (ware) waarde benadert en *precisie* de mate van overeenstemming tussen een reeks resultaten.

We begonnen dit hoofdstuk met het opzoeken van *nauwkeurig* en *precies* in *Van Dale* (*online*). En we eindigen met het opzoeken van *exact*.

exact *(bijvoeglijk naamwoord, bijwoord)* nauwkeurig, stipt: exacte wetenschappen wis-, schei- en natuurkunde; *de exacte cijfers heb ik niet; zij heeft exact dezelfde jas als ik* precies dezelfde

> Gekker moet het toch niet worden! Als eerste betekenis van 'exact' wordt door *Van Dale* 'nauwkeurig' gegeven. En 'exact dezelfde jas' is volgens datzelfde gezaghebbende boek 'precies dezelfde'… Pak je een naslagwerk erbij, dan wordt het één pot nat. Ze nemen het niet zo nauw, keurig uitgedrukt. Terwijl dat nou juist precies de bedoeling is!

Naast *nauwkeurig* (gemiddeld genomen dicht bij de werkelijke waarde) en *precies* (weinig spreiding in de meetwaarden) hebben we ook nog een derde begrip: *exact*. Een meting is *exact* als de uitkomst in alle mogelijke decimalen

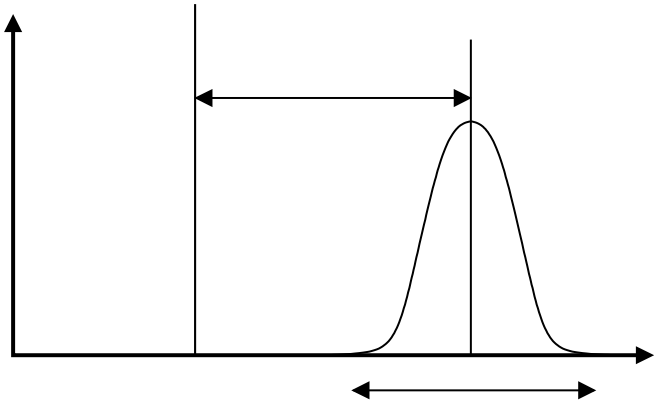
overeenkomt met de werkelijke waarde en er geen spreiding is. De onnauwkeurigheid en de imprecisie zijn nul. Er is dan geen verschil tussen meting en werkelijkheid. Bij een telling kan dat het geval zijn.

Voorbeelden van exacte dingen:

- een inch is exact 2,54 cm (want hij is zo gedefinieerd)
- de lichtsnelheid is exact 299 792 458 m/s (want ze is zo gedefinieerd)
- er gaan exact twaalf eieren in een dozijn (in een dozijn *eieren* dan…)

Oefeningen

1. Vul in Figuur 11.2 de juiste Nederlandse woorden in.



Figuur 11.2 Soorten onzekerheden, zonder tekst

2. Vul in Figuur 11.2 de juiste Engelse woorden in.

3. Iemand vraagt naar de lichtsnelheid afgerond op een geheel aantal inches per seconde. Is het mogelijk om een exact antwoord te geven?
4. Maak het woord ‘precision’ met de letters van ‘spreiding’.
5. Twee multimeters worden gebruikt voor het bepalen van een spanning. De eerste heeft als meetresultaat 2,36 ± 0,02 V.
   De tweede heeft als meetresultaat 2350 ± 5 mV.
   Welke van beide multimeters is nauwkeuriger?
6. Twee digitale afstandsmeters worden gebruikt voor het bepalen van een lengte. De eerste heeft een spreiding van 2% op een lengte van ongeveer acht meter. De tweede heeft een spreiding van 1 cm op een lengte van ongeveer zestig centimeter.
   Welke van beide afstandsmeters heeft de grootste nauwkeurigheid?

Een van mijn vroegste jeugdherinneringen — denk ik — gaat terug naar het antwoord dat ik kreeg van mijn vader op een in mijn ogen eenvoudige vraag. *"Papa, vijfentwintig gulden, is dat veel?"* Dat vroeg ik pakweg in 1973 en ik schreef dit hoofdstuk in 2023: precies een halve eeuw later. Gecorrigeerd voor inflatie van ƒ 25 in 1973 net zoveel als € 50 en een paar cent in 2023.

Mijn vraag werd beantwoord met: "Dat ligt eraan waarvoor." En dat is ook zo. Maar ik vond destijds dat je niets aan dat antwoord had. Vertel nou gewoon maar of het veel is, zonder daar omheen te draaien.

Van 1973 terug naar Meten & Weten en iets wat hierop lijkt. Of iets 'nauwkeurig' is of 'precies', is namelijk ook niet zomaar te beantwoorden. Je kunt alleen zeggen dat de ene manier van meten *nauwkeuriger* (of *preciezer*) is dan de andere. En je kunt ook nog kiezen of je dat absoluut of relatief neemt.

Nauwkeurig en precies zijn dus relatieve begrippen. En wat nu zo bijzonder is: ook de *relatieve* onzekerheden zijn getallen die je moet relateren aan iets anders. "Papa, is een relatieve onzekerheid van 5% klein?" "Dat ligt eraan…" Ook een relatieve fout is betrekkelijk.

12. Type A & Type B

De GUM *(Guide to the Expression of Uncertainty in Measurement*) zou je kunnen beschouwen als de standaard rondom het noteren van onzekerheden. In dat document legt de ISO een aantal zaken vast, waaronder ook terminologie.

Ratcliffe & Ratcliffe (2014) beginnen hun boek met een hoofdstuk over terminologie en besteden daarin aandacht aan het feit dat de GUM over Type A en Type B spreekt. Ze merken daarbij op dat de GUM zelf óók gebruik maakt van de als verouderd beschouwde termen.

De tekst uit de GUM (ISO, 2008) waar het om gaat, is als volgt.

The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated:

A. those which are evaluated by statistical methods,

B. those which are evaluated by other means.

There is not always a simple correspondence between the classification into categories A or B and the previously used classification into "random" and "systematic" uncertainties. The term "systematic uncertainty" can be misleading and should be avoided.

Any detailed report of the uncertainty should consist of a complete list of the components, specifying for each the method used to obtain its numerical value.

Kirkup & Frenkel (2006) geven een aardig voorbeeld om dit subtiele verschil duidelijker te maken, aan de hand van een meetinstrument waarbij er sprake is van een *drift*. In het Nederlands wordt *drift* ook wel *verloop* genoemd. Het is een kleine continue verandering in de weergegeven meetwaarden in een bepaald tijdsverloop waarbij de te meten waarden constant blijven. Dat kan een gevolg zijn van een verloop in de temperatuur of het inzakken van de spanning van een batterij.

Als er sprake is van ruis in het meetinstrument en de drift is betrekkelijk klein, dan is die niet zomaar waar te nemen. Een weegschaal die per minuut een gram meer aangeeft en geen toevallige fout heeft, is een voorbeeld daarvan. Als er geen ruis was, zou je de drift dus vrij eenvoudig kunnen vaststellen én ervoor corrigeren. Maar als de ruis groot is ten opzichte van de drift, wordt dat lastiger.

Hoe bepaal je een drift? Met *statistische methoden*. Als je elke seconde één meting zou kunnen doen, kon je de ruis uitmiddelen. Dat kan niet altijd,

bijvoorbeeld omdat de omstandigheden van de meting maken dat je niet vaak genoeg kunt herhalen. Er is dan sprake van een systematische fout en een toevallige fout. Beide zou je onder Type A moeten scharen.

De GUM geeft elders ook nog een toelichting op het verschil.

In some publications, uncertainty components are categorized as "random" and "systematic" and are associated with errors arising from random effects and known systematic effects, respectively. Such categorization of components of uncertainty can be ambiguous when generally applied. For example, a "random" component of uncertainty in one measurement may become a "systematic" component of uncertainty in another measurement in which the result of the first measurement is used as an input datum. Categorizing the methods of evaluating uncertainty components rather than the components themselves avoids such ambiguity. At the same time, it does not preclude collecting individual components that have been evaluated by the two different methods into designated groups to be used for a particular purpose.

Een toevallige waarde ligt vast vanaf het moment dat je die gemeten hebt. Als die toevallige vaste waarde vervolgens wordt meegenomen in een reeks berekeningen, levert dat een systematische fout op. De toevallige fout is daarmee systematisch geworden.

Het kan raar lijken: er bestaat een officiële internationale norm voor het omgaan met meetonzekerheden — de Guide to the Expression of Uncertainty in Measurement (GUM) — en toch volgen we die in dit boek meestal niet. Een norm is er immers om gevolgd te worden.

De reden is historisch en praktisch. Veel van wat je in Meten & Weten tegenkomt, sluit aan bij de aanpak van de boeken die jarenlang de ruggengraat vormden van het vak *Error Budgeting* op De Haagse Hogeschool: Taylor (1997) en Hughes & Hase (2010). Veel bestaande studieboeken hanteren de GUM-terminologie niet, of slechts gedeeltelijk. De Type A/B-classificatie vind je dus vooral terug in normen, audits en rapportages die formeel herleidbaar moeten zijn.

In de Taylor-Hughes-traditie ligt de nadruk op het inzichtelijk en kritisch omgaan met meetresultaten, foutbronnen en onzekerheden. Dat past bij de manier waarop studenten leren denken over meten — en bij de praktijk op veel hogescholen en universiteiten.

Toch is het goed dat je de terminologie even gezien hebt. In een stageverslag of ISO-audit kan het je zomaar van pas komen. En als je later wél met GUM gaat werken, weet je in elk geval dat het geen synoniem is voor wat je hier over systematisch en toevallig hebt geleerd.

13. Digitaal & Analoog

Figuur 13.1 Analoge klok

Hoe laat is het op bovenstaande klok? Vijf voor twee.

Bij een twaalfuursschaal zijn er dan nog twee mogelijkheden: het kan middag of nacht zijn. Op een digitale schaal: 13:55 of 01:55.

Deze klok heeft geen secondewijzer. Kunnen we het aantal seconden dan schatten? Dat ligt eraan. Bij sommige klokken verspringt de 'minst significante' wijzer (minuten of seconden) met stappen en op andere klokken gebeurt dat in een vloeiende beweging. In het laatste geval kun je inderdaad schatten hoeveel seconden er voorbij zijn.

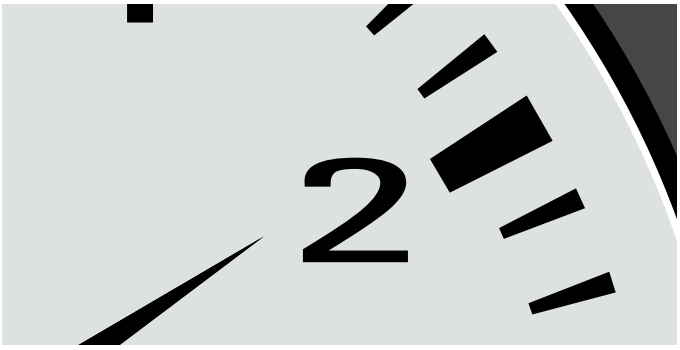
Deze klok heeft geen secondewijzer. Als een van beide wijzers afbreekt, dan mag je hopen dat het de minutenwijzer is. Die is het minst significant. Gek eigenlijk: de 'grote wijzer' is minder belangrijk dan de 'kleine wijzer'…

Hè, nou gebeurt het nog ook! Terwijl we nadenken over waarom ze de belangrijkste wijzer het kleinst gemaakt hebben, breekt die grote af.

Figuur 13.2 Klok met alleen een kleine wijzer

Kunnen we nu ook nog zien hoe laat het is? Ja. Tegen tweeën.

De kleine wijzer heeft in ieder geval wel een vloeiend verloop. Hij staat duidelijk dichter bij de 2 dan bij de 1. Sterker nog: er staan vier streepjes tussen de 1 en de 2, dus dat gebied is in vijven gedeeld. De afstand tussen twee strepen is 1/5 uur oftewel 12 minuten.



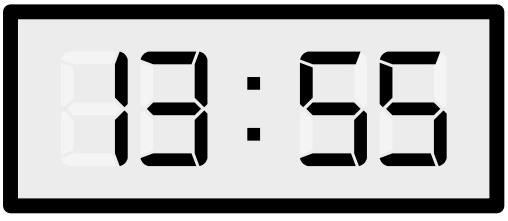
Figuur 13.3 Kleine wijzer, ingezoomd

Hè, was de kleine wijzer nou maar groter! Waren de klokkenbedenkers maar wijzer geweest. Dit is best nog lastig schatten. De wijzer staat dichter bij de dikke streep van de 2 dan bij de dunne streep van 12 voor 2…

Als ik zou moeten schatten, zeg ik: vijf voor twee.

En dat is nou juist zo lastig aan schatten. Het is bijna niet te doen om iets nog 'eerlijk' te schatten als je al weet hoeveel het is. De uitslag van een korfbalwedstrijd kun je na afloop exact goed voorspellen. "Ik had wel verwacht dat PKC zou winnen. Achteraf." Het psychologische aspect van *cognitieve bias* komt bij aflezen en meten ook voor.

Nu een andere klok.



Figuur 13.4 Digitale klok

Dat is andere koek: hier valt niks meer te schatten tussen streepjes.

Stel nu eens dat die analoge klok een minutenwijzer had die alleen maar precies op de streepjes kon staan. (Een wijzer die op z'n strepen staat dus.) Was het dan eigenlijk nog wel een analoge klok?

Kijk je in de *Van Dale Online*, dan vind je als betekenis van 'analoog':

niet werkend op basis van het binaire stelsel; gebruikmakend van wijzers bij het weergeven (tegenstelling: digitaal): een analoog horloge

Een klok met wijzers maakt gebruik van wijzers. *Wikipedia* zegt iets anders.

Analoog is de aanduiding voor een bepaald type technologie en de in sommige gevallen daarmee gepaard gaande signalen. Analoge meetinstrumenten geven een aanwijzing die in een continue relatie staat tot de te meten grootheid.

Als de secondewijzer sprongetjes maakt, is dat dus niet het geval. De te meten grootheid is de tijd die verstreken is sinds middernacht. Die kun je in seconden opgeven, maar ook in tienden ervan. Of in milliseconden. Of in nanoseconden of nog kleinere stapjes.

In dit hoofdstuk gaat het over klokken, maar de grote lijn is van toepassing op analoge en digitale aflezingen in het algemeen. Daarbij moet gezegd worden dat je er bij een klok vanuit gaat dat de minutenwijzer niet 'afrondt'. Hij springt pas op vijf voor twee als het vijf voor twee is. Bij een digitale meter die iets anders meet dan de tijd, zou het kunnen zijn dat er een afronding plaatsvindt naar de dichtstbijzijnde waarde.

We hebben overigens één belangrijke stap overgeslagen. Wie zegt dat de klok gelijkloopt?

Oefeningen

1. Geef een schatting van de onzekerheid die optreedt bij het aflezen van de tijd op bovenstaande klok met één wijzer. Maak gebruik van zowel Figuur 13.1 als Figuur 13.2.
2. Hoe ziet de kansverdelingsfunctie eruit van tijd die je afleest op de digitale klok?
3. Welke eenheden staan er langs de assen van de grafiek van bovengenoemde kansdichtheidsfunctie?

14. Aflezen van een analoge schaal

Was het niet veel handiger geweest als die klok digitaal was?

Misschien. Als de seconden erbij gestaan hadden, zouden we in één keer hebben kunnen aflezen. Als de seconden er niet bij gestaan hadden trouwens ook, maar dan hadden we geen flauw idee gehad van dat gedeelte van de tijd. Als je 15:02 afleest, is het misschien wel één milliseconde voor 15:03.

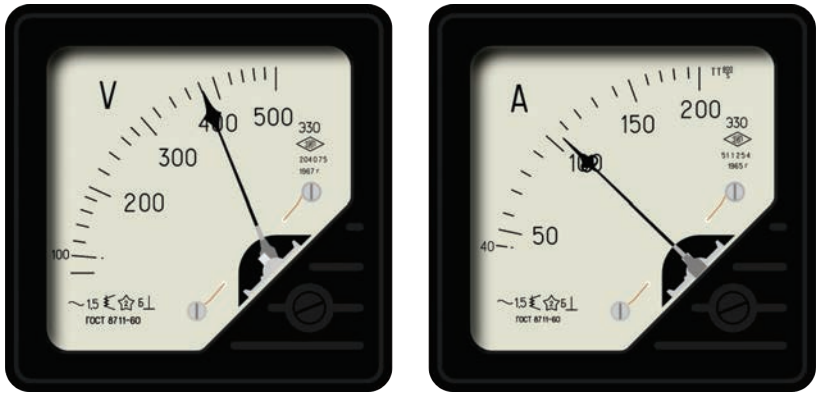
Sommige analoge klokken hebben dat trouwens ook. De secondewijzer beweegt met sprongetjes van een minuut.

Dat laatste brengt ons op een interessante kwestie. Is een analoge klok in dat geval eigenlijk niet ook digitaal? Hij is analoog in de zin dat hij wijzers heeft. Hij is digitaal in de zin dat er alleen maar discrete waarden uit kunnen komen. Er valt niets af te lezen of te schatten tussen de minuten in.

In feite heeft de analoge klok ondanks zijn vloeiend bewegende wijzer ook iets digitaals in zich. Stel dat je een foto van zo'n klok maakt en je zou zo nauwkeurig mogelijk aflezen. Dan kom je misschien op 15:02:48,2. Dat betekent dat je twee tienden van een seconde afleest door heel goed te schatten. Misschien lukt het je zelfs met een extreem scherpe wijzer en een foto met hoge resolutie om te zeggen dat het 15:02:48,21 is of 15:02:48,214. Maar dan houdt het echt wel op. Ook een continue schaal wordt een keer digitaal. Het aantal cijfers is nu eenmaal altijd eindig.

Hoeveel cijfers moet je zien af te lezen op een analoge of digitale schaal? Bij een digitale is het nogal makkelijk: alle cijfers die er maar staan. Er is geen goede reden om gegevens weg te laten die je hebt. Met de onzekerheid gaan we later pas rekening houden.

Bij een analoge meter zul je altijd moeten kiezen. Het aantal streepjes is eindig. Maar kun je ook nog iets zinnigs zeggen over wat er tussen de streepjes zit?



Figuur 14.1 Analoge meters voor spanning en stroom

Er is een vuistregel die zegt dat je tussen twee streepjes niet echt kunt aflezen, maar wel kunt zien welk streepje het dichtst in de buurt komt. Je kunt dus geen extra cijfer aflezen, maar je bent er wel zeker van dat je dichter bij het ene streepje zit dan bij het andere. De onzekerheid als gevolg van het aflezen zou dan een halve afstand tussen twee streepjes zijn.

Er valt wel wat af te dingen op die vuistregel. Wat als de wijzer van de schaal precies tussen twee streepjes lijkt te staan? Als je het dan niet helemaal goed ziet, zou je kunnen kiezen voor het streepje waar de wijzer juist iets verder vanaf staat. Zoals in het voorbeeld van de stroommeter hierboven, het rechter plaatje.

Aan de andere kant is het zo dat je bij een wat grove schaalindeling heus wel meer kunt zeggen dan bij welk streepje de wijzer het dichtst staat.

Figuur 14.2 Analoge spanningsmeters met verschillende schaalverdelingen

Neem de twee spanningsmeters hierboven.

De plaatjes zijn aan elkaar gelijk, het enige verschil is dat de kleine streepjes rechts weg zijn. Ja, je kunt zien dat de wijzer dichter bij de 400 V staat dan bij

de 300 V. Maar je kunt óók zien dat hij ergens halverwege ‘de tweede helft’ van dat gebied staat. Hij staat ergens tussen de 360 V en 390 V, laten we zeggen 375 ± 15 V. Een aanzienlijk kleinere onzekerheid dan 50 V dus.

Eigenlijk heb je gewoon niet zo veel aan die vuistregel. Kijk maar liever naar een onder- en bovengrens. Als je kunt schatten waar de werkelijke waarde tussen ligt, dan weet je het gebied en dus ook de onzekerheid.

Oefeningen

1. Leg uit waarom elke analoge meter in zekere zin ook digitaal is.
2. Waarvan hangt het af of je bij een analoge klok een schatting kunt maken van de tienden van de seconden?

15. Aflezen van een digitale schaal

Wat is de onzekerheid bij een digitale schaal?

Wat weet ik als het onderstaande getal op een digitale spanningsmeter staat, als die spanningen meet in volt?

Eigenlijk weet je niet wat je weet en niet weet, zolang dat niet in specificaties staat. Laat het apparaat de cijfers achter de komma gewoon weg? Of wordt er een afronding gedaan? Ligt de getalswaarde tussen 483,000… en 483,999…? Simpeler gezegd: tussen 483 en 484? Of is er een 'nette' afronding gedaan en ligt de getalswaarde tussen 482,5 en 483,5?

In beide gevallen geldt: het komt sowieso wat vreemd uit met de significante cijfers. Want 483,5 ± 0,5 V en 483,0 ± 0,5 V suggereren beide een cijfer meer dan je eigenlijk hebt.

Om die reden geven we de onzekerheid bij digitale meters als gevolg van het aflezen meestal gewoon op als 1 in het laatste weergegeven cijfer. De 483 op het display hierboven noteren we dus als 483 ± 1 V. Het is een vuistregel; niet meer dan dat, maar ook niet minder.

Oefeningen

1. Hoe groot is de onzekerheid in de meting als je niet weet of de meter afrondt op gehele getallen of afkapt na het laatste cijfer?
2. Geef in de correcte notatie en met het juiste aantal significante cijfers de beste schatting en onzekerheid als het display 483 aangeeft en de eenheden *volt* zijn.

16. Definitieprobleem

Taylor (1997) gebruikt in zijn boek het begrip 'definitieprobleem' (*problem of definition).* Hij geeft een voorbeeld van een timmerman (v/m) die wil meten hoe hoog een deur is. Dat doet hij of zij bijvoorbeeld met een rolmaat of duimstok, en die kun je vrij aardig aflezen. Op millimeters misschien wel. Maar het probleem met die deur is niet zozeer het aflezen en de eventuele onbetrouwbaarheid van datgene waarmee er wordt gemeten, maar de deur zelf. Die is nooit volmaakt recht en in oude huizen vaak zichtbaar scheef. Wat bedoel je dan precies met *de hoogte*? Het gemiddelde zou een aardig uitgangspunt kunnen zijn, maar de maximale hoogte is misschien wel nuttiger.

Zo'n definitieprobleem gaat nog een stap verder dan alleen scheef zijn. Er is niet alleen variatie in de *ruimte*, maar ook in de *tijd*. Zelfs als de deur niet meetbaar scheef is of als je besluit het midden van de deur aan te houden, kan bijvoorbeeld vocht ervoor zorgen dat de deur uitzet of krimpt.

Bij de meting van de lengte van een mens speelt dit soort verschijnselen ook een rol. Ons gewicht (natuurkundigen zouden zeggen: onze massa) varieert in de loop van een dag, van het jaar en van de jaren. Maar ook onze lengte is niet het hele etmaal door constant. Dat heeft te maken met de tussenwervelschijven die een mens in zijn of haar rug heeft. Terwijl de aarde om haar as draait, drukken die tussenwervelschijven wat in elkaar. Doordat een mensenlichaam er 23 heeft, kan dat over je hele rug genomen wel om een paar centimeter gaan. 's Morgens ben je langer dan 's avonds. Tenzij je dag- en nachtritme niet zo is dat je 's nachts op bed ligt.

Wat voor die timmerman met z'n deur en die zeventienjarige jongen en zijn keuring voor militaire dienst geldt, is van toepassing op veel metingen. Ook als er geen definitieprobleem is, hebben we vrijwel altijd te maken met een *onzekerheid*.

Als je het volledig en tot in detail met elkaar eens bent wat je wilt meten, kun je tot verschillende meetresultaten komen doordat je op een andere manier meet. Dat heeft te maken met *onzekerheden*. Die zijn onvermijdelijk, maar je kunt ze wel klein proberen te krijgen. *Definitieproblemen* zijn beter te vermijden, maar ook niet altijd helemaal. Wat bedoel je met de oppervlakte van een kamer? Meet je de afstand van wand tot wand of van plint tot plint? Neem je het midden van de wanden, of kijk je over de hele lengte en breedte?

Het ligt bij de oppervlakte van een kamer voor de hand om van de vloer uit te gaan. Toch is het niet echt fout om langs het plafond te meten…

17. Meetschalen

In het wetenschappelijke tijdschrift *Science* van vrijdag 7 juni 1946 werd er een artikel gepubliceerd met de titel *On the Theory of Scales of Measurement*. Het was afkomstig van de psycholoog Stanley Smith Stevens van de universiteit van Harvard en beschreef de verschillende niveaus waarin je metingen kunt onderverdelen.

Hij benoemde vier verschillende meetschalen.

- *nominaal*: je kunt zeggen dat iets tot een bepaalde categorie behoort
- *ordinaal*: je kunt zeggen dat het ene meer of minder is dan het andere
- *interval*: je kunt zeggen hoeveel het ene meer of minder is
- *ratio*: je kunt dat ook nog eens uitdrukken in een percentage

Noemenswaardig is ook nog de circulaire schaal. Als de ene klok op kwart voor drie staat en de andere op kwart over negen, dan kun je niet vaststellen of de ene voorloopt op de andere of omgekeerd.

Een bijzonder geval van de nominale schaal is een binaire schaal. Je kunt bijvoorbeeld zeggen dat een getal even of oneven is, terwijl even niet meer of minder is dan oneven.

Voorbeelden van een ordinale schaal kom je in de dagelijkse praktijk vaak tegen. Als je kiest tussen aardbeienijs en chocolade-ijs, dan kun je een sterke voorkeur voor het ene of het andere hebben. Misschien ook wel of het veel of weinig uitmaakt. Maar daar hoort niet zomaar een getal bij. Of het zou een jury-cijfer moeten zijn.

Bij een intervalschaal hebben de getallen wel een zinvolle betekenis, maar drukken ze niet meer uit dan een verschil. Een korfbalwedstrijd die met 21-12 gewonnen werd, kende een ruimere overwinning dan 16-14. Maar kun je zeggen dat 21-12 beter is dan 16-6? Of bij een sport waar minder gescoord wordt: 3-1 is beter dan 2-1. Maar is 3-1 ook beter dan 1-0?

Bij een ratioschaal kun je alle bewerkingen uitvoeren die er met getallen mogelijk zijn. Voorbeeld: Noa is 1,60 meter lang en Bo is 2 meter.

- Bo is dus langer.
- Het scheelt 40 centimeter.
- Bo is 25% langer dan Noa.
- Noa is 20% kleiner dan Bo.

Dat percentage hangt af van wie je als uitgangspunt neemt, maar de getallen zijn hard.

In tabelvorm ziet dat er als volgt uit.

Niveau		Kenmerkend	Volgorde	Verschillen	Nulpunt	Operatoren
Categorisch, Kwaliteit	nominaal	x				=, ≠
	ordinaal	x	x			>, <
Kardinaal, Numeriek, Kwantitatief	interval	x	x	x		+, −
	ratio	x	x	x	x	× , /

Het verhaal 'meetschaal' is lastig te combineren met onzekerheden. Als ik twee metalen staven heb met lengtes 50 ± 6 cm en 48 ± 4 cm, dan kan niet eens met zekerheid vaststellen welke van de twee groter is. Over de *beste schatting* kun je zeggen dat het verschil 2 cm of 4 % is: dat is een ratioschaal.

Een interessant voorbeeld van meetschalen is temperatuur. Uitgedrukt in graden Celsius of Fahrenheit is dat een intervalschaal. In kelvin is het een ratioschaal, omdat het absolute nulpunt een echt nulpunt is.

Oefeningen

1. Vandaag staat de thermometer op 20,2 °C en gisteren gaf hij 20 °C aan. Kun je zeggen dat het 1% warmer is dan gisteren?

2. Anne, Beau en Cor schaken tegen elkaar. Anne wint de partij tegen Beau. Beau wint daarna de partij tegen Cor. Cor wint tenslotte de laatste wedstrijd tegen Anne. Analyseer en beschrijf wat hier gebeurt.

3. Jan en Piet (uit Tilburg) stellen zich net als Kees (uit Eindhoven) verkiesbaar als voorzitter van Carnaval Brabant. Kees krijgt 37% van de stemmen, Piet 35% en Jan 27%. Omdat het verschil tussen Kees en Piet zo klein is, wordt er besloten om opnieuw te stemmen met alleen die twee kandidaten. Bij de tweede ronde krijgt Kees 41% van de stemmen en Piet 59%. Bedenk een verklaring hiervoor.

Ontwarren

18. Spreiding in ruimte & tijd

In hoofdstuk 1 verscheen het *definitieprobleem*. Wat bedoel je precies als je het hebt over de hoogte van een deur? De maximale hoogte? De gemiddelde hoogte? Meestal zit er in de praktijk niet veel verschil in. Maar als je in een huis van tweehonderd jaar oud woont, weet je dat het ook anders kan zijn.

Wat bedoel je precies als je vraagt naar de temperatuur in een kamer? Veel mensen kijken op een thermometer die aan de muur hangt en noemen het aantal graden Celsius. (Meestal laten ze het woord Celsius weg… Dat mag, zolang je het maar niet doet op je werk of als student op een tentamen, of tijdens de les.)

Dat het geen graden Fahrenheit zijn, nemen anderen meestal wel aan. Een belangrijkere vraag is: wat meet je precies? Als je een thermometer eerst op de grond zet en daarna boven op een kast, zul je zien dat je niet precies dezelfde temperatuur vindt, doordat warme lucht stijgt en koude lucht daalt. Misschien schijnt de zon aan de ene kant van de kamer iets meer, en verklaart dat het verschil. Het kan ook zo zijn dat het overal in de kamer wat warmer is geworden na die vorige meting.

De temperatuur in een kamer varieert dus in de ruimte en in de tijd.

Wiskundig gezien zou je het antwoord op de vraag ‘Wat is de temperatuur in deze kamer?’ alleen maar kunnen beantwoorden in de vorm van een functie die afhangt van de plaats (drie coördinaten) en de tijd (één variabele).

$$T(x, y, z, t)$$

Onder ideale omstandigheden zou dat kunnen. Maar dan ontstaat het volgende probleem. In hoeveel decimalen geef je de coördinaten weer? En in hoeveel cijfers achter de komma noteer je de tijd? Niet te vergeten: hoeveel cijfers gebruik je voor de temperatuur zelf?

Zelfs als je in micrometers en nanoseconden de temperatuur in een kamer zou kunnen bepalen en je voor de temperatuur vier of tien of honderdzestig cijfers achter de komma zet: er komt ergens een keer een moment dat het niet exact meer is. We hebben bij volmaakte meetinstrumenten dus al te maken met twee soorten onzekerheid als gevolg van wat we meten:

A. Alles gaat op één hoop, terwijl er een spreiding in ruimte en tijd is.
B. Er is een eindig aantal cijfers bij het noteren van de uitkomst.

Algemener gezegd:

1. Er zit een onzekerheid in datgene *wat* we meten.
2. Er zit een onzekerheid in datgene *waarmee* we meten.

Vaak gaat alle aandacht uit naar die tweede vorm van onzekerheid: de beperkingen van je meetinstrument en het aflezen daarvan. Daar heb je ook nog eens twee soorten in:

2a. Systematische fouten, die altijd hetzelfde zijn.
2b. Toevallige fouten, die per meting verschillen, maar gemiddeld wel nul zijn.

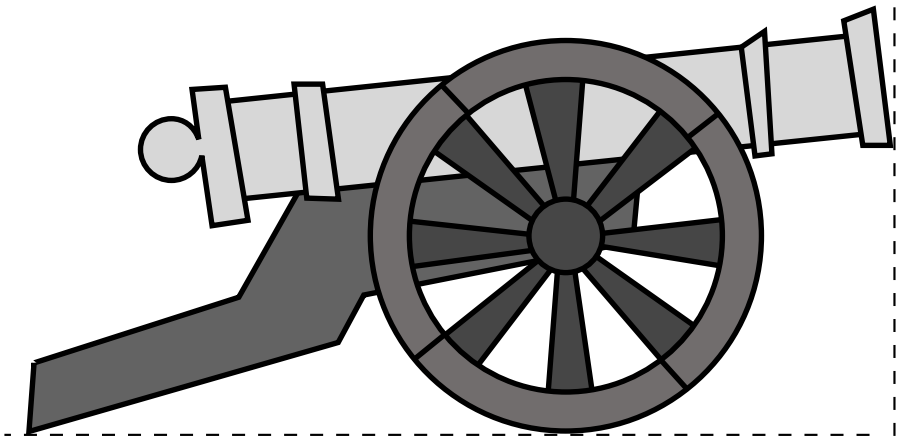
Veel lesboeken op dit vakgebied gaan uitgebreid in op punt 2 en het verschil tussen 2a en 2b, maar veel minder op punt 1. Helaas wordt er als het om onzekerheden gaat onevenredig veel aandacht besteed aan 2b. Want hoewel toevallige fouten 'lastiger' lijken, zijn ze dat juist niet. Ze vallen veel meer op dan systematische. Bovendien kun je er ook iets mee: statistiek bedrijven en ze daarmee verkleinen.

19. Spreiding in wat we meten

“Hoe ver kun je schieten met een kanon?”

Hebben we bij deze vraag een definitieprobleem? Ja. Op z’n minst eentje: welk kanon bedoel je? Het ene kanon is het andere niet. Laten we de vraag aanpassen.

“Hoe ver kun je schieten met dit ene kanon uit Figuur 19.1?”



Figuur 19.1 Kanon

Natuurlijk kun je met een getekend kanon niet schieten, maar dit is ook maar een boek. Laten we uitgaan van één specifiek werkelijk bestaand kanon.

Hebben we nu nog een definitieprobleem? Je zou in ieder geval goed moeten vastleggen welke afstand je precies bedoelt. Laten we meten vanaf het kruispunt van de stippellijnen: het uiteinde van de loop van het kanon en de grond zijn dan je uitgangspunten. En we trekken loodrecht een lijn omlaag.

Zijn die lijnen in de werkelijkheid exact gedefinieerd? De horizontale lijn is wat twijfelachtig: de ondergrond is namelijk nooit volmaakt vlak. Maar je zou de verticale lijn wel kunnen definiëren als de exacte richting van de zwaartekracht. Niet dat je die exact kunt *meten*, maar je hebt dan wel een exacte *definitie*.

Missen we iets? We hebben een werkelijkheid met drie ruimtelijke dimensies. (Misschien zijn er ook hogere dimensies, maar die zien we niet en kunnen we dus bij het meten buiten beschouwing laten.)

Toch hebben we nu aan één punt genoeg om het andere te vinden.

- vanaf het exacte midden van de loop van het kanon
- in de richting van de zwaartekracht
- naar de grond toe

We gaan voor het gemak — en voor onze veiligheid — uit van een speelgoedkanon dat een paar meter kan schieten. Dan kunnen we onze meting doen met een meetlint van 5 meter. Hoe groot is de onzekerheid in de meting? Een paar millimeter? Een hele centimeter? Daar kun je verschillend over denken. En dus is het erg belangrijk dát je er over nadenkt.

Dit zijn de afstanden die we meten:

2,60 m, 2,73 m, 2,57 m, 2,69 m, 2,56 m, 2,68 m, 2,72 m, 2,69 m, 2,60 m, 2,67 m.

Het verschil tussen de grootste gemeten afstand (2,73 m) en de kleinste gemeten afstand (2,56 m) is 17 cm. Dat is heus wel meer dan die onzekerheid in het aflezen van het meetlint.

We zien variaties in verschillende soorten:

1. De exacte definitie.
2. De kwaliteit van je meetlint.
3. Het aflezen van het meetlint.
4. De variatie in het schieten.

Het eerste punt valt hier op te lossen door eenduidig te definiëren. Het tweede en derde punt vormen de onzekerheid in je meting (*measurement*). Het vierde punt is de onzekerheid in de grootheid die je meet (*measurand*).

De variaties hebben te maken met luchtweerstand en toevallige windkracht, hoeveelheid kruit, positionering van de kogel en nog wel wat dingen. Die spreiding, is zo vreemd niet.

We hebben hier dus te maken met een spreiding in datgene wat je meet die groter is dan je onzekerheid in de meting.

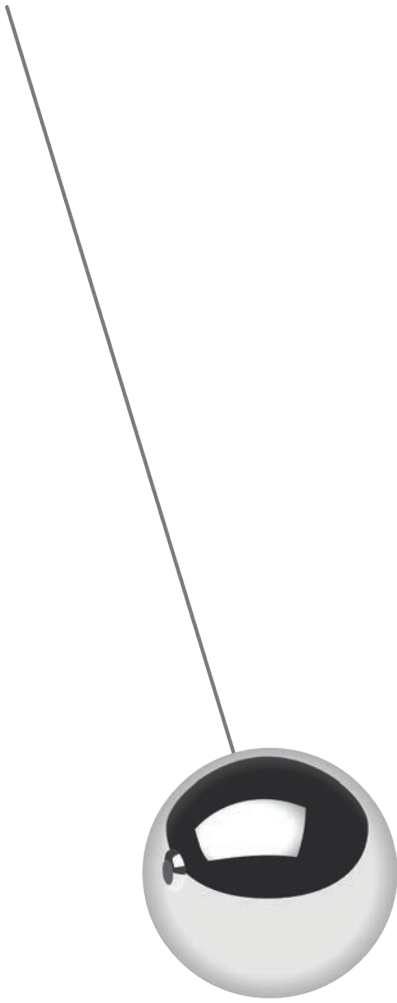
Oefeningen

1. Geef aan hoe je de onzekerheid in de meting zou kunnen schatten.
2. Geef een aanduiding voor de spreiding in de meting zelf.

20. Spreiding door het meten

De spreiding in de afstanden die we gemeten hebben bij dat speelgoedkanon uit het vorige hoofdstuk werd vooral veroorzaakt door dat kanon en de kogels zelf. Een beetje door de wind misschien. Het was wel telkens hetzelfde kanon, maar niet altijd dezelfde kogel.

Het meetlint had een onzekerheid bij het aflezen en ook het lint zelf is niet perfect. Zoals de Engelsen zeggen: *Nobody's perfect, even not a meetlint.*



Figuur 20.1 Perfecte kogel aan een perfect koordje

Laten we nu eens gaan kijken naar één zo'n kanonskogel. Een heel erg mooie kogel, mooi zwaar en symmetrisch. Opgehangen aan een flinterdun koordje.

We gaan nu de slingertijd meten van deze bijna volmaakte bol, in een ruimte waar alle ramen dicht zijn en de temperatuur en luchtdruk en ventilatie perfect constant is.

De tijd meten we op met een stopwatch. Een ouderwetse stopwatch met knoppen en wijzers, of een moderne app op een telefoon. In beide gevallen zul je zien dat je echt niet altijd precies dezelfde periodetijd meet.

Waar komen die verschillen vandaan?

Het is telkens dezelfde kanonskogel aan datzelfde koordje. Dat zal niet veel slijten of veranderen na één keer slingeren. Ook de hoek is nauwelijks verschillend. We houden de kogel strak tegen de tekst van dit hoofdstuk aan. Kijk maar in Figuur 20.1. Hij raakt nét niet de woorden op de volgende twee regels. Die hoek is telkens perfect gelijk aan de vorige keer. Nou ja, bijna perfect dan. *Nobody's perfect, even not Figuur 20.1.*

Hoe groot is de invloed van de tijdsmeting op de gemeten tijd?

Bij het kanon was het vooral het kanon en de kogel en de wind en het kruit of het onkruid

van het gras waarin de kogel landde. Dat meetlint trof geen blaam, dat kon je goed aflezen.

Figuur 20.2 Een stopwatch, met wijzers

Kun je een stopwatch goed aflezen? De wijzers wel en een digitale stopwatch kun je zelfs *perfect* aflezen. (Als je nog nooit een stopwatch met wijzers gezien hebt, er staat er eentje in Figuur 20.2.) Toch meet je niet elke keer hetzelfde. Dat heeft te maken met hoe goed en snel wij reageren en waarnemen. Onze ogen, hersenen, zenuwen, spieren en vingers zorgen voor de spreiding.

We hebben dus twee soorten spreidingen.

1. Spreidingen in datgene wat we meten.
2. Spreidingen door het meten,
    a. veroorzaakt door het meetapparaat, en
    b. veroorzaakt door de menselijke waarneming.

Nobody's perfect. Wat je meet is niet perfect, je meter is het niet én jijzelf bent het niet. Dat geeft niet, maar houd er rekening mee en schat de gevolgen ervan goed in.

Oefeningen

1. Als de kogel in Figuur 20.1 zo goed mogelijk (perfect lukt toch niet) tegen de tekst aan wordt gehouden, hoe groot zal dan de spreiding / onzekerheid zijn in de hoek?
2. Als je een reactietijd hebt van 0,3 seconde, hoe groot is dan de onzekerheid als gevolg hiervan in een tijdmeting met een stopwatch?

21. Beïnvloeden van wat we meten

Vaak meten we 'iets' en is dat iets niet scherp gedefinieerd. De hoogte van een deur varieert, dus je zult moeten aangeven of je maximale, gemiddelde of minimale hoogte bedoelt. Dat de hoogte varieert met de temperatuur en luchtvochtigheid, is nog zo'n probleem. We noemen dit een *definitieprobleem*.

Je kunt ervoor kiezen om het scherper te definiëren en al de genoemde factoren vast te leggen. Een alternatief is dat je het 'gewoon' over de hoogte hebt en die andere zaken meeneemt in je onzekerheid. Als er een spreiding van 3 mm is over het verloop van de deur en die ook nog eens 2 mm kan uitzetten of krimpen door temperatuurschommelingen en vocht, kun je dit natuurlijk ook meenemen in de onzekerheid zelf. De bijdrage van dit deel van de onzekerheid is dan groter dan die van het meetinstrument.

Onderschatting van fouten wordt vaak hierdoor veroorzaakt. We zien dat meetlint en denken: dit is wel vrij goed af te lezen. Maar er is veel meer onzeker dan dat. Dat is wel zeker.

In het genoemde voorbeeld lag het aan de deur zelf. Vaak 'doen' metingen ook iets met datgene wat we meten.

- De dikte van een zacht voorwerp wordt opgemeten met een schuifmaat die dat zachte voorwerp een klein beetje samendrukt.
- De spanning van een batterij zakt in als er stroom loopt en er loopt inderdaad stroom als je de spanning meet.
- De temperatuur van een vloeistof daalt of stijgt een klein beetje doordat de temperatuur van de thermometer die je erin stopt niet gelijk is aan de temperatuur van de vloeistof.

Ook hier geldt dat je de beïnvloeding zou moeten meenemen in het resultaat. In sommige gevallen kun je ervoor compenseren of rekening houden met het feit dat je weet in welke 'richting' de invloed is. Als de schuifmaat inderdaad het voorwerp een beetje indrukt, weet je dat de afgelezen waarde te klein is.

Als je weet dat je moet compenseren maar niet exact met hoeveel, zul je het moeten schatten. Die onzekerheid introduceert een nieuwe toevallige fout.

22. Herhalen & Reproduceren

Er is een Engels spreekwoord dat zegt: *Measure twice, cut once*. Je zou het kunnen vertalen met: 'twee keer meten, één keer zagen'. Of 'snijden'.

> Wie de spreuk bij Google intikt, vindt allerlei enorm grappige plaatjes en foto's van bordjes en fotolijstjes met deze tekst, die dan net niet past. Als je bijgekomen bent van het lachen: er zit een serieuze gedachte achter.

Wie iets moet opmeten en dan afzagen, is in de aap gelogeerd als hij of zij daarna pas erachter komt dat de maat niet klopte. Dan moet je nóg een keer zagen en dat kost je extra tijd en hout. En dus geld. Een tweede keer meten kost ook tijd. Maar meestal minder.

> Deze benadering heeft een keerzijde. Als je héél vaak twee keer meet zonder dat je daardoor een fout vindt, kost dat ook veel tijd.

Een belangrijke meting herhalen is verstandig. Maar als je datgene wat je deed op precies dezelfde keer nogmaals doet, is er een reële kans dat je dezelfde fout maakt. Hoe kun je iets over doen op een andere manier?

- Je zou een andere meetmethode of meetinstrument kunnen kiezen.
- Je zou het iemand anders kunnen laten doen.
- Je zou het op een andere dag of plaats kunnen herhalen.

De kans dat je iets minstens één keer fout doet, wordt overigens *groter* als je het twee keer doet. Statistisch gezien is de kans dat een handeling die in 90% van de gevallen goed gaat twee keer achter elkaar goed gaat 81%. De kans dat het tien keer achter elkaar goed gaat, is $0{,}9^{10} \approx 35\%$ Maar dat is niet het sommetje dat we maken. Als je voor 90% zeker weet dat het goed gaat, hoe groot is dan de kans dat het fout gaat? Inderdaad: 10%. En hoe groot is de kans dat het twee keer achter elkaar fout gaat? Inderdaad: 1%.

> De enige manier om zeker te weten dat je iets niet fout doet, is: er niet eens aan beginnen. Nul pogingen kunnen niet één keer falen.

Wat moet je doen als je de som van tien getallen wilt bepalen? Optellen, inderdaad. En wat als je zeker wilt weten dat er geen fout in de optelling zit?

- Nog een keer optellen.
- Nog een keer optellen, maar dan van onder naar boven.
- Nog een keer optellen, maar dan met een rekenmachine.
- De getallen overtikken in Excel en dan de som berekenen.
- Iemand anders vragen om het op te tellen.
- Nog iemand anders vragen om het op te tellen.

Helemaal zeker weten dat je geen fout in een meting of berekening maakt doe je eigenlijk nooit. Maar goede controles helpen wel.

Is 'reproduceren' niet gewoon een uit het Latijn afkomstig (en dus deftig of zelfs *chique*) woord voor 'herhalen'? Nee. De zuiver Hollandse uitdrukking voor 'reproduceren' is: *nabootsen*.

> Waarom heet dit hoofdstuk niet: Herhalen & Nabootsen? Omdat 'reproduceren' een gangbaar woord is als tegenhanger van 'herhalen'. De norm ISO 5725-1 gebruikt in de Engelse versie *repeatability & reproducibility*. In de Franse editie: *répétabilité* en *reproductibilité*.
>
> Als we het Engels (en het Frans) willen nabootsen, zou je in plaats van 'herhalen' beter 'repeteren' kunnen zeggen. Waarom is de titel van dit hoofdstuk dan niet: Repeteren & Reproduceren? Omdat *herhaalbaarheid* en *reproduceerbaarheid* nu eenmaal gebruikelijk zijn.

In de techniek zijn herhaalbaarheid (*repeatability*) en reproduceerbaarheid (*reproducibility*) twee verschillende dingen. Beide hebben betrekking op het vermogen om consistente resultaten te verkrijgen; het verschil zit in de omstandigheden waaronder de metingen worden uitgevoerd.

Herhaalbaarheid is de mate van overeenstemming tussen meetresultaten wanneer de test wordt uitgevoerd onder dezelfde omstandigheden. Dus:

- door dezelfde persoon
- met hetzelfde meetinstrument
- op dezelfde locatie
- niet al te veel later (tegelijk zou mooi zijn, maar kan niet)

Reproduceerbaarheid is de mate van overeenstemming tussen meetresultaten wanneer de test wordt uitgevoerd onder andere omstandigheden. Dus:

- door andere personen
- met andere meetinstrumenten
- op andere locaties
- op een later (of eerder) tijdstip (andere dag, een jaar later…)

> Van *herhalen* en *reproduceren* is alleen sprake als je dezelfde methode blijft gebruiken. Maar wat als je expres anders meet? Dan ben je bezig met *method comparison* of *methodologische validatie*.
>
> Het kan dus zijn dat twee metingen van hetzelfde wél van elkaar mogen verschillen. Omdat er geen fout is, maar omdat de methode zelf een andere gevoeligheid of systematiek kent.

23. Werkelijk & Geaccepteerd

Dit hoofdstuk is gedeeltelijk een herhaling van wat we al gezien hebben. Het gaat over iets wat vaak verwarring geeft. En die verwarring wordt een beetje in stand gehouden door de term *reference value* in het plaatje van *Wikipedia*.

Wat is nou weer een *reference value* (referentiewaarde)? Is dat dan nóg iets anders dan *true value* (werkelijke waarde) en *accepted value* (geaccepteerde waarde)?

Nee. Als het goed is, bedoelen de makers van het plaatje de *true value*. Maar die kennen we niet. Eigenlijk is dat een tekst om op een spandoek te drukken en dan heel groot op te hangen.

De werkelijke waarde ken je niet.

You do not know the true value.

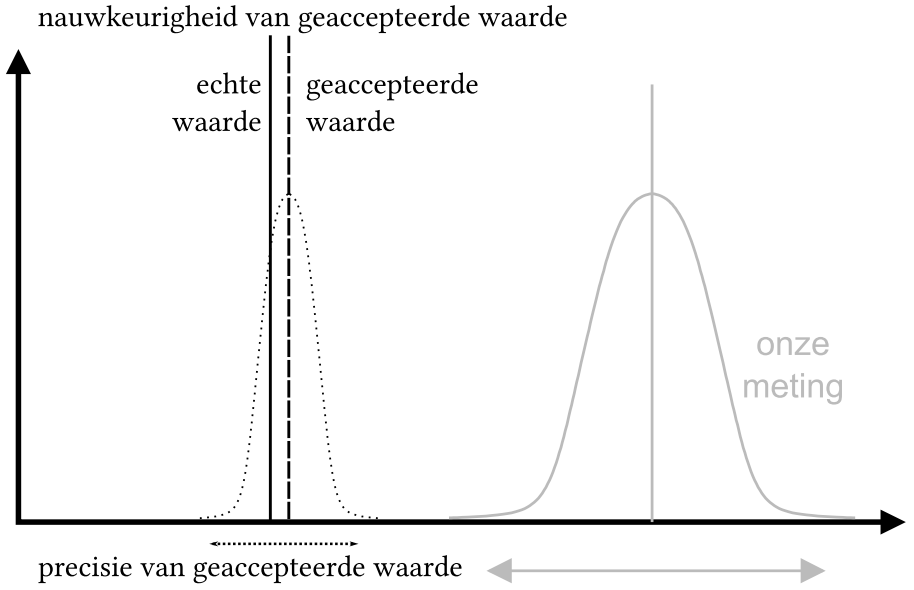
> Vreemd eigenlijk, dat je in het Engels het lijdend voorwerp niet zo mooi vooraan in de zin kunt zetten.

De werkelijke waarde ken je niet — anders hoefde je niet te meten. Wie meet er nou iets wat hij of zij al weet?

Wie de versnelling van de zwaartekracht op zijn plekje op aarde wil weten, kan die meten of opzoeken. De werkelijke waarde ken je niet. Wat je opzoekt, is de *accepted value*. Een of ander genormaliseerd instituut heeft metingen en berekeningen gedaan. Die mensen hebben dat — nemen we aan — héél goed gedaan. Onze eigen uitkomsten leggen we daarnaast en we zien in het verschil onze gebrekkigheid. Als we een strijdigheid vinden, concluderen we dat we zelf iets verkeerd deden. Of dat we te optimistisch waren over de onzekerheid.

De werkelijke waarde ken je niet, en omdat je nu eenmaal behoefte hebt aan iets om mee te vergelijken of rekenen nemen we dan maar iets uit een boek. Vaak is dat een *gemeten* waarde met een (kleine) onzekerheid. Die onzekerheid zal veel kleiner zijn dan onze huis-tuin-en-keukenmeting met een stukje touw dat we in de keukenla vonden en ‘iets zwaars’ om mee te slingeren.

Wat nu als wel de *accepted value* en de *true value* in dat *Wikipedia*-plaatje willen tonen? Dan krijg je Figuur 23.1.



Figuur 23.1 Echte en geaccepteerde waarde met onzekerheden

De werkelijke waarde ken je niet. Maar hij bestaat wel! Althans, onder bepaalde omstandigheden. Die valversnelling in *precies* het zwaartepunt van je slinger op die ene plek in het universum, die heeft een waarde die hooguit nog afhangt van hoe de massa van de aarde varieert: geen grote veranderingen dus. Voor een natuurconstante geldt nog sterker: het *is* een werkelijke waarde. Maar die ken je niet.

De geaccepteerde waarde komt dicht in de buurt, maar heeft wel een onzekerheid. Een onzekerheid die is opgebouwd uit een toevallig deel en een systematisch deel.

De verticale dichte lijn in Figuur 23.1 is de werkelijke waarde.

De gestreepte dichte lijn in Figuur 23.1 is de geaccepteerde waarde.

De gestippelde kromme geeft de kansverdeling van meetwaarden weer als je diezelfde meting van de geaccepteerde waarde herhaaldelijk uitvoerde onder precies gelijke omstandigheden. Die verdeling is smaller dan ‘de onze’ en het gemiddelde ligt dichter bij de werkelijke waarde. De metingen waren namelijk nauwkeuriger én preciezer. In dat onzekerheidsgebied ligt ergens de werkelijke waarde.

24. Beste schattingen

Een meting bestaat uit een beste schatting en een onzekerheid, zeggen we vaak. En dat is ook zo. Maar voor je het weet, verlies je uit het oog dat je dan met *twee* schattingen te maken hebt. Eigenlijk zelfs met *drie*.

- een beste schatting van de waarde (schatting 1)
- een beste schatting van de onzekerheid, bestaande uit
    - een deel dat altijd gelijk is (schatting 2)
    - een deel dat varieert (schatting 3)

> Eigenlijk is het vreemd dat we die twee delen waaruit de onzekerheid bestaat op één hoop gooien. Het is bepaald niet onbelangrijk of de onzekerheid door een *systematische* of een *toevallige* fout komt.

> En dan nog iets: bij 'onzekerheid' denken we, met het vak statistiek in ons achterhoofd, aan waarden die telkens anders zijn. Bij systematische fouten is dat niet het geval. Die zijn *constant*, maar we weten *niet* hoe groot. Dat we het niet weten, is dus een onzekerheid. Je weet alleen maar zeker dat de grootte steeds hetzelfde is.

Van de eerste weten we heus wel dat er een onzekerheid op zit. Daar gaat dit vakgebied zo'n beetje helemaal over. Over die tweede weten we *helemaal niets* op basis van de metingen.

Zoals we er bijna blind van uitgaan dat *het gemiddelde* de beste schatting is van de (werkelijke) waarde, nemen we aan dat de standaardafwijking de beste maat is van de spreiding. Een aanname die voor normale verdelingen zo gek nog niet is.

Barford (1985) laat op een aardige manier zien hoe we tot de gebruikelijke formule voor het schatten van de standaardafwijking komen.

Stel, je hebt één meting. Het is alleszins redelijk om het gemiddelde van die ene meting (de meting zelf dus!) als schatting te nemen voor de waarde. Maar het is *volstrekte onzin* om de spreiding in de meting(en) te nemen als schatting van de onzekerheid. Die ene meting is exact gelijk aan het gemiddelde… de spreiding is dus nul.

No one number can ever give us two pieces of independent information. One measurement gives no information whatsoever about the precision of the apparatus.

Stel, je hebt twee metingen. Het gemiddelde is de som van beide gedeeld door twee. Nu kunnen we wél een zinnige uitspraak doen over de spreiding.

$$\delta_1 = |x_1 - \overline{x}|$$
$$\delta_2 = |x_2 - \overline{x}|$$

Het lijkt nu alsof we te maken hebben met twee gegevens van de spreiding, maar dat is niet zo. Kijk maar naar de formule voor het gemiddelde.

$$\overline{x} = \frac{x_1 + x_2}{2}$$

Als je die formule invult in de twee voor de spreiding vind je:

$$\delta_1 = |x_1 - \overline{x}| = \left|x_1 - \frac{x_1 + x_2}{2}\right| = \left|\frac{2x_1 - x_1 - x_2}{2}\right| = \left|\frac{x_1 - x_2}{2}\right|$$

$$\delta_2 = |x_2 - \overline{x}| = \left|x_2 - \frac{x_1 + x_2}{2}\right| = \left|\frac{2x_2 - x_1 - x_2}{2}\right| = \left|\frac{x_2 - x_1}{2}\right|$$

Met andere woorden: $\delta_1 = |-\delta_2| = \delta_2$

> Voor wie niet zo veel met wiskunde heeft, kan het een stuk makkelijker. Als ik 5 V en 7 V heb gemeten, dan is het gemiddelde 6 V. Die 5 V en 7 V zitten beide 1 V van het gemiddelde af. Zo werkt dat nu eenmaal met het gemiddelde van twee getallen: dat zit precies in het midden, en dus even ver van allebei.

We zien nu al één belangrijk ding: we hebben geen twee waarden die aangeven hoe groot de spreiding is, maar slechts eentje.

Wat nu als je de gemiddelde spreiding van N metingen wilt hebben? Elk van die N metingen heeft een spreiding met $N - 1$ andere metingen. (Eigenlijk tel je nu dubbel, want het verschil (de spreiding) tussen x_i en x_j is natuurlijk hetzelfde als het verschil tussen x_j en x_i. Dat maakt voor de berekening niet uit, want als we straks gaan middelen, delen we ook door twee keer zo veel.)

Zoals de spreiding van twee (2) getallen maar één (2 – 1) waarde heeft, heeft de spreiding van N getallen maar $N - 1$ onafhankelijke afwijkingen ten opzichte van het gemiddelde.

> Anne zegt dat ze 1,77 m is. Maar hoeveel marge zit daar eigenlijk op? Wil denkt: ±3 cm. Anne zelf: hooguit 1 cm. Een vriend houdt het op 2.
>
> Haar lengte kennen we dus tot binnen 2%, maar hoe goed kennen we die onzekerheid nou?
>
> Als je zegt dat de onzekerheid 2 cm is, met een spreiding van ±1 cm, dan heeft die onzekerheid zélf een onzekerheid van 50%.
>
> Iemands lengte schatten is één ding. De onzekerheid daarvan inschatten is vaak veel moeilijker.

25. Systematische fout door toeval (1)

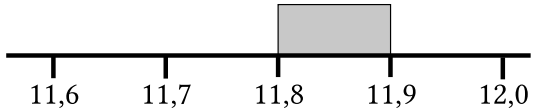
Als een *toevallige* fout in een meting veroorzaakt wordt door een *systematische* afronding omlaag (lees: door het weglaten van decimalen), dan bevat deze fout een systematische component. Dat komt doordat het gemiddelde van het stuk dat je weglaat, niet nul is.

> Dit is er eentje om in de gaten te houden. We kennen de kansverdeling van de fouten in werkelijkheid eigenlijk nooit; we maken er slechts schattingen van. Dat resulteert in normale of uniforme of andere verdelingsdichtheidsfuncties die altijd een gemiddelde van nul hebben. Het zou ook een beetje raar zijn als dat niet zo was. “Ik denk dat de fout die we maken gemiddeld −3 mV is…” Als je dat denkt, waarom trek je die waarde dan niet overal van af? De spreiding verandert er niet door en de gemiddelde fout wordt (in absolute zin) kleiner. Kleiner kan hij niet zijn: exact 0 mV.

Stel dat je een multimeter hebt die de cijfers die hij niet kan weergeven domweg weglaat. Wat betekent het dan als je 11,8 V meet? Ervan uitgaande dat de fout alleen wordt veroorzaakt door het weglaten van cijfers (wat in werkelijkheid natuurlijk niet zo is) weet je dat de spanning minstens 11 800 mV is en maximaal 11 899 mV. (Met achter de komma nog een heleboel negens.)

De waarde die weggelaten wordt, kan alles zijn. En alle mogelijkheden zijn ook even waarschijnlijk.

> Dat begrip *oneindig* blijft toch ook maar een raar ding. Er zijn oneindig veel mogelijkheden. Toch is er maar een relatief klein deel van echt mogelijk: alleen dat smalle stukje tussen 11,8 V en 11,9 V. Waarbij geldt dat 11,8 V *wel* een mogelijke uitkomst is en 11,9 *net niet*.
>
> Hoe groot is eigenlijk de kans op *exact* 11,9 V? Die is nul. Dat is ook weer zo vreemd met continue kansverdelingen. Doordat er oneindig veel waarden liggen in dat continue interval, is de kans op elke individuele uitkomst nul. Wiskundigen hebben het dan ook over een *kansdichtheid*. Er is namelijk wel een kans op een waarde tussen 11,80 V en 11,81 V. En er is ook een kans dat een waarde ligt tussen 11,80000 V en 11,80001 V. Terwijl daar ook weer oneindig veel getallen tussen liggen.
>
> Je wordt een beetje duizelig van lang over getallen nadenken.

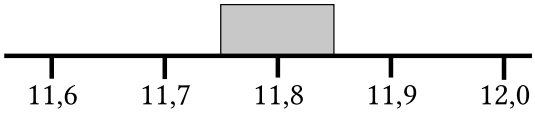


Die uniforme verdeling van de spanningen heeft een gemiddelde dat niet gelijk is aan 11,8 V, maar aan 11,85 V. Als je 11,8 V afleest, betekent dit dat de spanning gelijk is aan 11,85 ± 0,05 V.

> Ook vreemd: om de gemeten waarde met zijn onzekerheid op juiste wijze te noteren, gebruik je één significant cijfer meer dan het display aangeeft. Er zijn geen drie, maar vier significante cijfers in de beste schatting. (En eentje in de onzekerheid.)

We zien dus dat het weglaten van alle volgende decimalen twee soorten fouten tot gevolg heeft. Een toevallige fout van 0,05 V én een systematische fout van 0,05 V. Tenzij je ervoor corrigeert. Als je weet dat de afronding altijd naar beneden is en je noteert de zojuist genoemde 11,85 ± 0,05 V, dan ben je dat probleem kwijt.

De kansdichtheidsfunctie ziet er na corrigeren (of voor een meetinstrument dat het goed afrondt) als volgt uit.



Hier zit geen systematische fout meer in. Er is een stuk dat we nooit zullen weten. Het enige wat we wel weten, is dat alle waarden even waarschijnlijk waren.

> Een grijs kader! Sommige lezers slaan die altijd over. Andere lezers spitsen dan de ogen. (Kun je ogen wel spitsen eigenlijk? Nee. Maar zolang Meten & Weten geen luisterboek is…)
>
> Ergens in dit hoofdstuk stond:
>
> *Met achter de komma nog een heleboel negens.*
>
> Daar had moeten staan ‘oneindig veel’, en niet ‘een heleboel’. In een stageverslag krijg je opmerkingen als je zoiets doet. Dan zegt de docent dat je taalgebruik te informeel is. Waarom doe ik het hier dan wel? (En waarom schrijf ik ‘ik’? Dat is ook een vorm die afgeraden wordt.)
>
> Ik heb daar verschillende redenen voor. Eén daarvan is dat ik nu een aanleiding heb om een Uitstapje te maken over informeel taalgebruik in verslagen en boeken. Denk daar maar eens over na, nu de pagina vol is.

26. Systematische fout door toeval (2)

Een systematische fout kan ook uit een toevallige fout ontstaan als gevolg van de doorwerkingsformule. Bijvoorbeeld als je te maken hebt met een derde macht.

Voordat we dat gaan uitschrijven, even oefenen met haakjes uitwerken.

$$(a+b)^2 = (a+b) \times (a+b) = a^2 + 2ab + b^2$$

$$(a-b)^2 = (a-b) \times (a-b) = a^2 - 2ab + b^2$$

Deze kende je vast nog wel. Nu een macht hoger: geen kwadraat, maar tot de derde. We maken dan gebruik van de stap die we al hadden.

$$(a+b)^3 = (a+b)^2 \times (a+b) =$$

$$\left(a^2 + 2ab + b^2\right) \times (a+b) =$$

$$a^3 + 2a^2b + ab^2 + a^2b + 2ab^2 + b^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

Nu gaan we deze uitgeschreven vorm gebruiken voor de oppervlakte van een cirkel en het volume van een bol, beide met een straal r.

Formules:

$$A = 4\pi r^2$$

$$V = \frac{4}{3}\pi r^3$$

Stel dat er in de straal een fout (*error*) zit ter grootte van Δr.

> We nemen nu even een *fout* en niet een *onzekerheid*, en daarom ook een andere delta (de hoofdletter in plaats van de kleine letter). De reden daarvoor is dat het wat vreemd aanvoelt om te rekenen (vermenigvuldigen en delen) met een onzekerheid. Dat is een *bereik*, geen getal. Dat noteer je ook niet met een *plusteken*, maar met een *plusminusteken*. Een fout is gewoon een waarde (die we meestal niet kennen). Je kunt de gevolgen van die (hypothetische, verzonnen) waarden doorrekenen.

Wat komt er uit ons sommetje als we niet met r rekenen, maar met $r + \Delta r$?

Dat kunnen we eenvoudig bekijken dankzij die formules die we net uitschreven, met voor a de waarde r en voor b de waarde Δr.

$$a^2 + 2ab + b^2$$

$$a^3 + 3a^2b + 3ab^2 + b^3$$

Ingevuld en vermenigvuldigd met de factor die ervoor staat:

$$A = 4\pi(r^2 + 2r\Delta r + (\Delta r)^2)$$

$$V = \frac{4}{3}\pi(r^3 + 3r^2\Delta r + 3r(\Delta r)^2 + (\Delta r)^3)$$

De fout Δr kan positief of negatief zijn. We gaan uit van een symmetrisch verdeelde fout (normaal of uniform, meestal). Die symmetrie raken we kwijt door het doorrekenen. Vul in plaats van Δr maar eens $-\Delta r$ in.

$$A = 4\pi(r^2 - 2r\Delta r + (\Delta r)^2)$$

$$V = \frac{4}{3}\pi(r^3 - 3r^2\Delta r + 3r(\Delta r)^2 - (\Delta r)^3)$$

Als de fout in de oppervlakte van een cirkel en het volume van een bol symmetrisch zou doorwerken, dan zou de oppervlakte van een cirkel met straal $r + \Delta r$ (in positieve zin) net zoveel moeten afwijken van de oppervlakte van een cirkel met straal $r - \Delta r$ (in negatieve zin). Een fout 'de ene kant op' moet net zoveel invloed hebben als een fout 'de andere kant op'. Alleen dan heffen positieve en negatieve fouten elkaar op.

We zouden het verder kunnen uitschrijven door de oppervlakte en het volume zonder de fout Δr ervan af te trekken. Dan krijg je dit.

$$\Delta A = 4\pi(-2r\Delta r + (\Delta r)^2)$$

$$\Delta V = \frac{4}{3}\pi(-3r^2\Delta r + 3r(\Delta r)^2 - (\Delta r)^3)$$

Om het nog leuker te maken, zou je die verschillen kunnen berekenen voor een positieve én een negatieve fout van dezelfde grootte. Daar zou in het ideale geval nul uitkomen. Maar dat is niet zo.

> Dat zul je altijd zien: gevallen zijn maar zelden ideaal. Ja, in films. En in boeken. En in de verhalen van docenten als ze voor de klas staan. En tijdens diploma-uitreikingen, want dan is opeens alles goed. Maar nu dus niet. Meestal niet trouwens, let er maar eens op.

We nemen nu voor altijd een positieve waarde en vullen die in de formule in met een plus of een min ervoor.

$$\Delta A_{positief} = 4\pi(-2r\Delta r + (\Delta r)^2)$$

$$\Delta A_{negatief} = 4\pi(2r\Delta r + (\Delta r)^2)$$

$$\Delta V_{positief} = \frac{4}{3}\pi(-3r^2\Delta r + 3r(\Delta r)^2 - (\Delta r)^3)$$

$$\Delta V_{negatief} = \frac{4}{3}\pi(3r^2\Delta r + 3r(\Delta r)^2 + (\Delta r)^3)$$

Nu kunnen we ze van elkaar aftrekken.

$$\Delta A_{negatief} - \Delta A_{positief} =$$

$$4\pi\big((2r\Delta r + (\Delta r)^2) - (-2r\Delta r + (\Delta r)^2)\big) =$$

$$16\pi r\Delta r$$

$$\Delta V_{negatief} - \Delta V_{positief} =$$

$$\frac{4}{3}\pi\big((3r^2\Delta r + 3r(\Delta r)^2 + (\Delta r)^3) - (-3r^2\Delta r + 3r(\Delta r)^2 - (\Delta r)^3)\big) =$$

$$\frac{4}{3}\pi(6r^2\Delta r + 2(\Delta r)^3) =$$

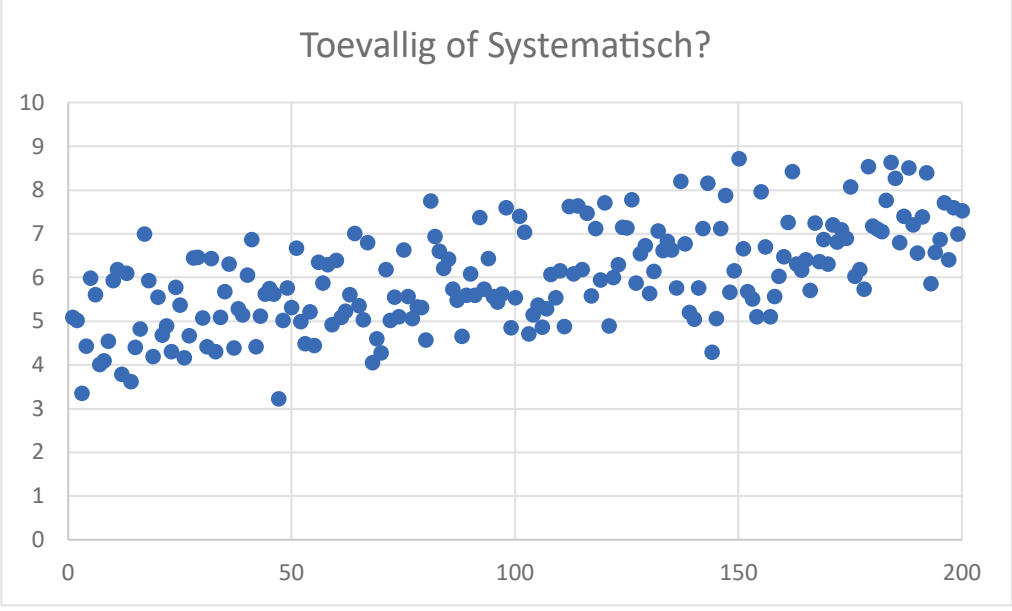
$$\frac{8}{3}\pi(3r^2\Delta r + (\Delta r)^3) =$$

We zien dus dat er zowel bij de oppervlakte van de cirkel als bij het volume van de bol een altijd positieve term overblijft: r en Δr zijn immers positief. De positieve term is groter naarmate de fout groter is. Voor beide geldt trouwens ook dat hij groter wordt naarmate de straal zelf groter is. Best opmerkelijk… De doorwerking van de *absolute fout* wordt dus groter als de *grootheid* zelf ook groter is. Het klopt, maar het is niet vanzelfsprekend.

Berendsen (2011) gebruikt in bijlage A2 van zijn boek ook het voorbeeld van het volume van een bol. Niks mis mee, maar hij had ook de oppervlakte van een cirkel kunnen nemen. Het zit namelijk niet in die derde macht.

27. Systematisch toenemende fout

Onderstaande grafiek toont de uitkomst van tweehonderd metingen aan een denkbeeldige grootheid.



Er staan geen eenheden bij. Het zou een spanning kunnen zijn in millivolt, een temperatuur in graden Celsius, of wat dan ook. Voor het gemak kijken we nu alleen even naar de getallen.

De metingen zijn gedaan in een periode van één uur. Precies om de achttien seconden werd er een meting verricht en omdat 200 × 18 s = 3600 s, is er dus exact een uur gemeten.

Kunnen we op basis van deze metingen iets zeggen over de toevallige of systematische fout?

Het moge duidelijk zijn dat er een flinke spreiding in de metingen zat. De laagste waarde was 3,23, de hoogste 8,71. Het gemiddelde was 6,02 en de standaardafwijking was 1,14. Je zou dus kunnen zeggen dat de uitkomst van de meting gelijk was aan 6,0 ± 1,1.

Bij het kijken naar de grafiek valt er iets op. De spreiding in de waarden is vrij groot, maar er lijkt toch een bepaalde trend zichtbaar te zijn. Gemiddeld genomen nemen de waarden toe in de loop van dat ene uur waarin er gemeten is. De waarden aan de rechterkant van de grafiek liggen gemiddeld hoger dan aan de linkerkant.

Excel kan ons een handje helpen met een trendlijn, maar die hebben we hier weggelaten. Als je die zou trekken, zag je inderdaad een stijgende lijn.

Omdat we de fictieve metingen zelf hebben gegenereerd, weten we hoe ze tot stand zijn gekomen. De tweehonderd getallen bestaan uit normaal verdeelde ruis met een gemiddelde van 0 en een standaardafwijking van 1, opgeteld bij een lineaire functie.

$$y(t) = 5 + 0{,}01t$$

Als het een temperatuurmeting was, zou je dus kunnen zeggen dat de temperatuur lineair toenam van 5,0 °C naar 7,0 °C. Een toename die gelijk is aan tweemaal de standaardafwijking van de ruis, de toevallige fout die naast de lineaire trend aanwezig was.

Kun je nu concluderen dat er sprake is van een systematische fout die van 0,0 °C naar 2,0 °C oploopt? Nee. Het is in ieder geval niet de enige mogelijkheid. Misschien is er wel een systematische fout die van 1,3 °C naar 3,3 °C oploopt. Of van −0,5 °C naar 1,5 °C.

De *verandering* in de *systematische* fout is gedeeltelijk te vinden door *statistische* analyses. En daarmee geldt dus hetzelfde als voor toevallige fouten. Veranderingen kun je waarnemen en analyseren, vaste afwijkingen zijn ‘ingekapseld’ in de metingen zelf.

28. Zwevendekommagetallen

Dit is de langste hoofdstuktitel die uit maar één woord bestaat. En het is nog niet eens zeker of het de juiste vertaling van *floating point numbers* is. Het Engelse *floating* kan inderdaad 'zweven' betekenen, maar ook 'drijven'.

Wikipedia gebruikt beide termen, en zelfs nog een derde.

Een zwevendekommagetal of drijvendekommagetal, verouderd ook vlottendekommagetal (Engels: floating-point number) is een gegevenstype dat een vaste geheugenruimte beslaat en een grote variëteit aan getallen kan bevatten, van zeer kleine tot zeer grote. Hoewel zwevendekommagetallen strikt genomen rationale getallen zijn, worden ze meestal gebruikt als benadering voor reële getallen. De relatieve nauwkeurigheid waarin getallen door zwevendekommagetallen worden gerepresenteerd, is min of meer gelijk over het gehele bereik. Het is de digitale versie van de wetenschappelijke notatie.

> Als je de woorden onder elkaar zet, krijg je een grappig effect.
>
> zwevendekommagetallen
> drijvendekommagetallen
>
> De eerste lijkt langer te zijn dan de tweede… terwijl het aantal letters juist eentje minder is: 21 in plaats van 22. Dat komt doordat 'zwe' (drie letters) breder is dan 'drij' (vier letters). De achttien letters daarna zijn hetzelfde.

In het geheugen van een computer wordt een zwevendekommagetal weergegeven in drie 'delen'.

- *s*: een bit dat het teken aangeeft
- *e*: een bitreeks die een macht van 2 aangeeft
- *m*: een bitreeks die een geheel getal aangeeft

Een teken (*sign*, s), exponent (e) en mantisse (m).

> Het woord 'mantisse' is ook een raar ding. Je komt het alleen hier tegen en in een paar verschillende betekenissen in de wiskunde. Het deel na de komma van een logaritmisch getal heet óók mantisse.

De representatie van een zwevendekommagetal ligt sinds 1985 vast in IEEE 754. De verdeling van de beschikbare bits verschilt bij enkele precisie (*single*) van 32 bits en dubbele precisie (*double*) van 64 bits.

- 32 bits: 1 voor *s*, 8 voor *e* en 23 voor *m*
- 64 bits: 1 voor *s*, 11 voor *e* en 52 voor *m*

Dat ene bit voor het teken is eenvoudig: een 1 is negatief, een 0 is positief.

De binaire exponent wordt zó gecodeerd dat een deel van het bereik negatieve machten van twee weergeeft. Je moet daarom — we gaan nu voor het gemak even uit van 32-bits getallen — 127 van de decimale waarde aftrekken.

Bij de mantisse is het even opletten. Het binaire getal is namelijk *het deel na de komma*. Vóór de komma staat altijd een 1. En die 1 wordt niet gecodeerd in die 23 of 52 bits. Dat zou zonde van de ruimte zijn: hij is toch altijd 1.

Een voorbeeld: het getal `01000100110011010100000000000000`

deel je op in die s, e en m: `0 10001001 00110101000000000000000`

Die eerste 0 is een plus. De exponent rekenen we om naar decimaal: 137. Van die decimale 137 moeten we 127 aftrekken en dan houden we dus 10 over.

Die 00110101000000000000000 is het deel achter de komma, en staat dus voor 1,100110101. Alle laatste nullen zijn weggelaten, want die voegen niets toe. Dit binaire getal moet nog worden omgerekend naar decimaal: 1,603515625.

Dat omrekenen kun je via een *online tool* doen, maar ook 'met de hand'. Het getal is $2^0 + 2^{-1} + 2^{-4} + 2^{-5} + 2^{-7} + 2^{-9} =$

$1 + 0{,}5 + 0{,}0625 + 0{,}03125 + 0{,}0078125 + 0{,}001953125 = 1{,}603515625$

In *Excel*:

1	0	1	1
1	-1	0,5	0,5
0	-2	0,25	0
0	-3	0,125	0
1	-4	0,0625	0,0625
1	-5	0,03125	0,03125
0	-6	0,015625	0
1	-7	0,0078125	0,0078125
0	-8	0,00390625	0
1	-9	0,001953125	0,001953125
			1,603515625

De waarde van het (positieve) getal kun je nu uitrekenen door de mantisse te vermenigvuldigen met 2 tot de macht die is uitgedrukt in de exponent.

$1{,}603515625 \times 2^{10} = 1{,}603515625 \times 1024 = 1642$

Nu andersom. We willen het getal pi opslaan als zwevendekommagetal. Pi is 3,14159265358979323846264338327950288419716939937510582097494459230 7.

Ongeveer dan. Het echte aantal decimalen is oneindig. Hoeveel kunnen we er kwijt? We hebben 23 bits voor de mantisse en $2^{-23} \approx 0{,}000000119$. De vijftien decimalen van *Excel* gaan we heus niet halen met 32 bits. En als we opslaan in 64 bits, houdt het ook een keer op. Dan zijn er 52 bits voor de mantisse en $2^{-52} \approx 2{,}22 \times 10^{-16}$. Dat komt in de buurt van *Excel*, maar is maar een kwart van wat hierboven staat op die ene regel. Dat lange getal, namelijk 3,14159265358979323846264338327950288419716939937510582097494459230 7.

De helft daarvan is (afgerond op diezelfde 66 decimalen):

1,57079632679489661923132169163975144209858469968755291048747229615 4.

We kunnen pi nu dus schrijven in de vorm 1,.... × 2^1: het product van een gebroken getal met een 1 voor de komma en een macht van twee. Dit getal gaan we nu binair schrijven. Niet exact, maar in 66 decimalen, één regel vol.

1,1001001000011111101101010100010001000010110100011000010001101001 10.

De mantisse bestaat uit de eerste 52 bits na de komma.

m = 0010010000111111011010101000100010000101101000110000

We hadden al s = 0 (positief) en e = 10000000000 ($2^1 \longrightarrow 1 + 1023$).
Achter elkaar gezet:

0 10000000000 0010010000111111011010101000100010000101101000110000

> Hoe gaan we nu controleren of dit klopt? Tsja… Dat valt nog niet mee. Misschien moeten we het maar terugschroeven naar 32 bits. Daarna moet de exponent in 8 bits worden gecodeerd en de mantisse in 23 bits. Dan kun je met een *online tool* controleren of er hetzelfde uitkomt.
> IEEE-754 Floating Point Converter (h-schmidt.net)
>
> 0 10000000 00100100000111111011011
>
> Joepie! Dat is hetzelfde.

29. Decimalen in Python & C

Een klassiek voorbeeld om te laten zien hoe het werken met *floats* mis kan gaan, is het optellen van 0,1 en 0,2.

Neem onderstaand Pythonprogramma.

```
a = 0.1
b = 0.2
c = a + b

print(a)
print(b)
print(c)

if 10 * a == 1:
   print("gelijk")
else:
   print("ongelijk")
```

Als je niet beter wist, zou je denken dat hier de getallen 0,1 en 0,2 en 0,3 worden afgedrukt, gevolgd door de conclusie dat tien keer een tiende gelijk is aan één.

Als je het ietsje beter weet, en het vorige hoofdstuk begrepen én onthouden hebt, denk je dat die a niet precies 0,1 kan zijn en b niet precies 0,2. Dan kan c dus ook niet precies 0,3 zijn. En waar je vergif op zou innemen, is dat tien keer net-niet-een-tiende net niet gelijk is aan één.

Volgens ons hulpje IEEE-754 Floating Point Converter (h-schmidt.net) zijn dit de exacte representaties die de computer intern gebruikt.

0,1	0,100000001490116119384765625
0,2	0,200000002980232238769531250
0,3	0,300000011920928955078125

Natuurlijk is het even schrikken als je ziet hoeveel cijfers er staan die daar helemaal niet horen. Maar kijk eens wat een lange lijst met nullen na die komma. Zo gek is dit nog niet. Het is alleen niet exact.

Nu de uitvoer van het programma.

```
0.1
0.2
0.30000000000000004
gelijk
```

Huh? Hij geeft die waarden van 0,1 en 0,2 helemaal goed weer.

> Niet helemaal natuurlijk: er staat een punt terwijl er een komma zou moeten staan. Dat kun je *Python* niet kwalijk nemen. De taal is bedacht door de in Den Haag geboren Hollander Guido van Rossum, maar die heeft dat op het Engels gebaseerd.

Dat tien keer een tiende gelijk is aan één, is op zich waar. Maar de uitkomst van 10 × 0,100000001490116119384765625 is 1,000000014901161193847656 25. En dat is net geen één. We hadden dus ongelijk met ons 'ongelijk'.

Hoe gaat dit in C?

```
int main()
{
   float a = 0.1;
   float b = 0.2;
   float c = a + b;

   printf("%f\n", a);
   printf("%f\n", b);
   printf("%f\n", c);

   if (10 * a == 1)
      printf("gelijk");
   else
      printf("ongelijk");

   return 0;
}
```

Dit zou fout moeten gaan, toch? Nee hoor. Kijk maar naar de uitvoer.

```
0.100000
0.200000
0.300000
gelijk
```

Omdat ik het niet vertrouw, voer ik het programma nog een keer uit, maar dan door meer decimalen te gebruiken en de waarde van tien keer een tiende af te drukken.

```
printf("%.20f\n", a);
printf("%.20f\n", b);
printf("%.20f\n", c);
printf("%.20f\n", 10 * a);
```

En maar weer eens naar de uitvoer kijken.

```
0.10000000149011612000
0.20000000298023224000
0.30000001192092896000
1.00000000000000000000
gelijk
```

Nou ja, zeg! Is hier nog een peil op te trekken? Een touw aan vast te knopen?

Terug naar het hulpje IEEE-754 Floating Point Converter (h-schmidt.net): wat verwachten we precies?

0,1	0,100000001490116119384765625
0,100000001490116119384765625	0,100000001490116119384765625
1,000000014901161193847656 25	1

Tsja. Een computer heeft dus moeite om 0,1 goed op te slaan. Of beter gezegd: het lukt niet. Het getal 0,100000001490116119384765625 gaat probleemloos, dankzij het feit dat het exact te schrijven is als het ietwat lange getal $1{,}600000023841858 \times 2^{-4}$ (= 1,600000023841858 / 16). Maar als je precies het tienvoudige van 'zijn' versie van 0,1 probeert op te slaan, wat dus gelijk is aan 1,000000014901161193847656 25, dan lukt 'm dat ook niet en zit hij er een klein stukje naast door exact 1 te nemen.

> Logisch: 1 is een macht van 2 en komt wel heel dicht in de buurt.

Het getal dat in C wordt afgedrukt (0,10000000149011612000) is overigens niet exact gelijk aan 0,100000001490116119384765625. De eerste zestien decimalen zijn gelijk, met de laatste elf gaat het mis. Waar dát verschil dan weer in zit, is een vraag voor wie zin heeft om nog dieper in de bits te spitten.

In Python is 0.1 + 0.2 gelijk aan 0.30000000000000004. Dat komt door die representatie in machten van 2. Gelukkig kent die taal ook een manier van opslaan waar 'gewoon' decimaal wordt geteld.

```
from decimal import Decimal
Decimal('0.1') + Decimal('0.2')
```

Dobbelen

30. Rekenkundig gemiddelde

In het vak Statistiek is er maar weinig zo eenvoudig als *het gemiddelde*. Je telt gewoon alles bij elkaar op en deelt de som door het aantal getallen dat je had. Het gemiddelde van 6 en 8 is 7.

In het Engels zeg je *average* of *mean*, maar let op: die woorden betekenen niet precies hetzelfde, al worden ze nogal eens door elkaar heen gebruikt. Een *average* is een maat voor de centrale tendens in een reeks getallen. Vaak is dat het gemiddelde (de *mean*), maar het kan ook de *median* (mediaan) of de *mode* (modus). Dat zijn namelijk óók getallen die maatgevend zijn.

Laten we deze drie begrippen eens bekijken aan de hand van een voorbeeld. Voor een tentamen zijn de volgende cijfers gehaald.

1, 4, 4, 5, 5, 6, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8, 8, 9, 9 en 10

Twintig resultaten; optellen en delen geeft een *gemiddelde (mean)* van 6,45.

De *modus (mode)* is het cijfer dat het vaakst voorkomt. Dat is de 6, het enige cijfer dat vijf keer gehaald is. In dit voorbeeld is de 6 het *modale* cijfer.

De *mediaan (median)* is de *middelste waarde* uit het rijtje als er sprake is van een oneven aantal, en het *gemiddelde van de middelste twee* als er sprake is van een even aantal. De cijfers staan al op volgorde en het zijn er twintig. Het tiende en elfde cijfer zijn een 6 en een 7. De mediaan is dus 6,5.

We hebben inmiddels drie *averages* gevonden:

- Het gemiddelde is een 6,45
- De modus is een 6
- De mediaan is 6,5

Die 6,45 is dus de *mean*. Maar zelfs de Engelstalige *Excel* doet dat niet correct.

Wat gemeen! In *Excel* zeggen ze *average* terwijl ze *mean* bedoelen. En om het nog bonter te maken: 'bedoelen' is in het Engels *to mean* en 'gemeen' is ook al *mean*... Dat kun je toch niet menen! Dat vreemde woord *average* veroorzaakt zomaar een rare maar ware vage ravage!

Het gemiddelde lijkt simpel — tot de Engelsen zich ermee gaan bemoeien. Ondertussen hebben wij wel een woord dat wij niet kunnen vertalen. Het is zoiets als 'doorsnee' of 'normgetal'. Engelsen zeggen *typical value*, wat in het Nederlands neerkomt op 'representatieve waarde'. An *average guy* is een gewone man, een doodnormale vent.

Zullen we *average* dan maar 'de doodnormale waarde' noemen? Nee.

Wat als het tentamen anders was uitgevallen en dit de cijfers waren geweest?

1, 3, 3, 3, 3, 3, 5, 5, 5, 7, 8, 8, 8, 9, 9, 9, 10, 10, 10, 10

Is dit tentamen nu beter of slechter gemaakt? Het hoogste en laagste cijfer is hetzelfde. Niemand haalde een 2. Verder zijn er opvallende verschillen. Maar liefst vier tienen en aardig wat negens, maar ook flink wat mensen met een 3. We hebben drie *averages* gevonden:

- Het gemiddelde is een 6,45
- De modus is een 3
- De mediaan is 7,5

Ook het gemiddelde is exact gelijk aan dat van de vorige toets. De mediaan is een heel punt hoger. Maar die modus is nogal schrikken: vorige keer nog een nette 6, nu maar de helft daarvan.

De modus is een wat wispelturige maat. Bij twintig tentamencijfers zou je theoretisch alle cijfers precies twee keer kunnen hebben. Dan is elk cijfer een modaal cijfer. Maar als de 10 maar één keer wordt gehaald en de 9 drie keer: dan is die 9 de modus. Het had ook de 4 kunnen zijn. Of de 1.

De mediaan is minder gevoelig voor uitschieters. Als de vier studenten die nu een 10 hadden een 8 haalden, zou de mediaan hetzelfde geweest zijn.

> Mochten er docenten zijn die Meten & Weten lezen: als u (of jij) en ik een stapel tentamens nakijken, kijken we dan naar het gemiddelde cijfer? Of de mediaan? De modus wellicht? Nee. Wij kijken naar het *slagingspercentage*. Dat is een maatgevend getal dat interessanter is. We hadden eerst 15 voldoendes en toen 11. De percentages geslaagden waren 75% tegenover 55%. Die waarden zeggen meer over het aantal herkansingen dat we over twee weken kunnen verwachten.

> Mochten er taalliefhebbers zijn die Meten & Weten lezen: het is vreemd dat we 'het rekenkundig gemiddelde' zeggen. We zeggen: 'de groene kaft van een groen boek'. Eerst 'groene' met een e, dan zonder. Als we de bepaalde en onbepaalde lidwoorden omwisselen, wordt het 'een groene kaft van het groene boek'. *Een* groen boek, *het* groen*e* boek. Na 'het' komt een 'e'. Waarom dan niet 'het rekenkundig*e* gemiddelde'? Je zegt toch ook niet: 'het laag gemiddelde'?
>
> In het Nederlands worden bijvoeglijke naamwoorden die een kenmerk beschrijven vaak niet verbogen wanneer ze een technische of abstracte betekenis hebben. Taal is soms raar. *Also in Dutch.*

31. Harmonisch & kwadratisch gemiddelde

Het gemiddelde laat zich het makkelijkst uitleggen in spreektaal: ‘alle getallen optellen en delen door het aantal’. Maar het kan ook in formulevorm.

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Hierin is: $\bar{x}$ het gemiddelde, N het aantal waarden en x_i één enkele waarde.

Eigenlijk moet je zeggen: *rekenkundig gemiddelde*. Er is nog minstens één andere manier om een gemiddelde uit te rekenen: het *harmonisch gemiddelde*.

Wat is het nut van het harmonisch gemiddelde? Stel, je fietst van het gebouw van de Haagse Hogeschool in Delft naar de hoofdvestiging in Den Haag. Je doet dat via de Lange Kleiweg. Dat is — zoals de naam al zegt — niet de kortste route, maar rekent wel handig: volgens Google Maps is het 10,0 km. Op de heenreis fiets je 20 km/u. Terug heb je tegenwind en haal je nog 15 km/u. Wat is nu je gemiddelde snelheid over de hele fietstocht?

Het lijkt een makkelijk sommetje: het gemiddelde van 20 en 15 is namelijk 17,5. Maar je gemiddelde snelheid blijkt geen 17,5 km/u te zijn.

Op de heenweg zat alles mee. De meewind, maar ook de ronde getallen. Een afstand van 10 km en een snelheid van 20 km/u: dat betekent precies een half uur fietsen. Terug deed je er langer over, namelijk 10/15 (oftewel 2/3) van een uur: 40 minuten.

Je gemiddelde snelheid is de afgelegde weg gedeeld door de tijd. Dat is dus 20 km gedeeld door 30 + 40 = 70 minuten. Omdat we in uren rekenen, moet je delen door 7/6 van een uur. De uitkomst is 120/7 = $17^1/_7$ uur. Voor wie liever niet exact rekent: 17,14 km/u. Minder dus dan wat we eerder uitrekenden.

Hoe kan dat nu? Waar is die 0,36 km/u gebleven? Het punt is: de *afstand* waarover je die 20 km/u haalde was gelijk aan de afstand waarover het langzamer ging. Dat was allebei precies 10 km. Maar *de tijd* die je bezig was met langzamer fietsen was wel langer. En dat hakt erin.

Vaak is het begrijpen van dit soort dingen makkelijker als je wat overdrijft. Bij getallen helpt het dan soms als je met 0 werkt. Wat zou er gebeurd zijn als je op de terugweg écht niet vooruitkwam en met die stevige zuidoostenwind op de heenreis 40 km/u had gehaald? Hoe laat zou je dan thuis geweest zijn die dag? De rit naar Den Haag toe was een kwartiertje geweest. Maar met 0 km/u zou je nooit in Delft gekomen zijn. Een oneindig lange terugreis.

Ook het harmonisch gemiddelde is in een formulevorm te geven.

$$h = \frac{N}{\sum_{i=1}^{N} \frac{1}{x_i}}$$

En hoewel het formulevormtechnisch minder netjes is, schrijf je ook wel:

$$\frac{1}{h} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{x_i}$$

Minder netjes, omdat je nu niet links van het isgelijkteken datgene hebt staan wat je wilt uitrekenen. Waarom dan toch die andere formule? Omdat die wat duidelijker laat zien wat het is. Je bepaalt namelijk het gemiddelde van alle *inverses* en uiteindelijk doe je 1 gedeeld door wat eruit komt.

De *inverse* van x is 1 gedeeld door x oftewel x^{-1}. Je noemt het ook wel *omgekeerde* of de *reciproque* (vaak op z'n Hollands geschreven als 'reciproke').

Het omgekeerde van een breuk ontstaat door teller en noemer onderling te verwisselen. Let op: het *omgekeerde* is iets anders dan het *tegengestelde*. Dat krijg je door iets met −1 te vermenigvuldigen.

> De vervangingsweerstand van twee weerstanden die parallel staan, is ook een mooi voorbeeld om het uit te leggen. In serie is het 'gewoon' optellen. Parallel is het 'gewoon' het harmonisch gemiddelde.

Het *kwadratisch gemiddelde* vind je door het optellen van getallen in het kwadraat en van de som de wortel nemen.

$$x_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2} \quad (31.1)$$

Het Engelse *RMS* staat voor *Root Mean Square*.

> Voor de taalliefhebbers: *root* is 'wortel', maar ook 'oorsprong', *mean* is 'gemiddeld', maar ook 'gemeen', *square* is 'kwadraat', maar ook 'vierkant', 'plein' of 'ouderwets'.
>
> Er staat niet wélke wortel je moet trekken: de tweedemachtswortel, vierkantswortel of kwadraatwortel. *Square Root Mean Square* dus eigenlijk. In het Nederlands ontbreekt het hele woord 'wortel' in 'kwadratisch optellen'. In het Duits is het *Quadratwurzel*, wat klinkt als iets wat je met zuurkool op een broodje kunt doen.

32. Standaardafwijking

De *standaardafwijking* is het kwadratisch gemiddelde van de afwijkingen van het gemiddelde.

Eigenlijk is dit hoofdstuk nu klaar, want meer is het niet. Je zou nog kunnen zeggen dat de standaardafwijking vaak *standaarddeviatie* wordt genoemd en dat het de wortel is van de variantie. Maar in wezen is dat alles.

Laten we nog eens kijken naar die tentamencijfers waarvan we eerder al het gemiddelde, de mediaan en de modus bepaalden.

1, 4, 4, 5, 5, 6, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8, 8, 9, 9 en 10

Het gemiddelde cijfer was een 6,45. Je zou van alle behaalde cijfers de afwijkingen van dat gemiddelde kunnen bepalen door die 6,45 ervan af te trekken. De student met een 10 week 3,55 punt af van het gemiddelde. De studenten met een 6 weken −0,45 punt daarvan af en die ene met een 1 zelfs −5,45.
Om het kwadratisch gemiddelde te bepalen, kwadrateren we die afwijkingen. Dat leidt tot grote getallen voor de uitersten — logisch, want die wijken het meest af — en geeft alleen maar positieve waarden, door dat kwadrateren.

Om het met de hand te kunnen uitrekenen pakken we vijf cijfers eruit: 4, 5, 7, 7, 9. De modus en de mediaan zie je in één oogopslag: allebei 7. De modus is 7 omdat dat cijfer het vaakst voorkomt. De mediaan is 7 omdat dat het middelste cijfer is als je ze van klein naar groot sorteert.

Het gemiddelde kan ook uit het hoofd: 4+5+7+7+9=32. Als je dat deelt door 5, kom je op 6,4.

De standaardafwijking wordt al gauw lastig om uit het hoofd te rekenen. Het gemiddelde aftrekken van het cijfer gaat nog wel. Kwadrateren is al pittig. Dan nog het gemiddelde van al die kwadraten… En een wortel trekken uit je blote bol, is de meeste mensen ook niet gegeven.

Cijfer	4	5	7	7	9
Afwijking	-2,4	-1,4	0,6	0,6	2,6
Kwadraat	5,76	1,96	0,36	0,36	6,76

Vooruit dan maar, toch de rekenmachine. Of nog makkelijker: *Excel*. De som van die kwadraten is 15,2. Het gemiddelde is die waarde gedeeld door 5 en is

dus 3,04. De wortel daarvan is in vier decimalen 1,7436. Omdat we met hele cijfers werken, lijkt het redelijk om één decimaal te nemen: 1,7 of 1,8.

> Welke malloot gaat er nou kwadrateren, optellen, delen en worteltjes trekken als hij met *Excel* een standaardafwijking berekent? Dat doe je toch met de formule die daarvoor gemaakt is?
>
> Met `=STDEV.P(A:A)` krijg je de standaardafwijking van kolom A.
>
> Python met NumPy kan ook: `waarde = np.std(getallen)`

We weten nu hoe we een standaardafwijking berekenen, maar de vraag wat dat *is* kun je ook beantwoorden door het te hebben over de *betekenis* ervan. Je rekent een getal uit, maar wat doe je ermee?

Omdat de standaardafwijking de wortel van de som van de kwadraten van de afwijkingen is, heeft hij altijd dezelfde dimensie als datgene waaraan je rekent. Stel dat je met snelheden werkt: het verschil met het gemiddelde is in meters per seconde, als je kwadrateert wordt dat m^2/s^2 en als je som neemt, blijft het m^2/s^2. Uiteindelijk trek je de wortel en ben je bij de snelheden terug.

De vraag was: wat is de betekenis van de standaardafwijking? Het is een maat voor de spreiding van de waarden. Als je voor die oorspronkelijke twintig tentamencijfers de standaardafwijking berekent, kom je (met één decimaal) op 2,0. Hoe groot zou die voor dat tweede tentamen zijn? Meer of minder dan 2,0? Laten we de resultaten er nog eens bij pakken.

1, 3, 3, 3, 3, 3, 5, 5, 5, 7, 8, 8, 8, 9, 9, 9, 10, 10, 10, 10

In de eerste toets was er ook een 1, maar er waren geen tweeën en drieën. En er was maar één 10. Bovendien waren er meer cijfers rond de 6 en 7. Veel cijfers die ver afwijken van het gemiddelde dus. En weinig die in de buurt van het gemiddelde zitten. Zonder het uit te rekenen kun je 'aanvoelen' (eigenlijk: beredeneren) dat de standaardafwijking groter zal zijn. Als je het uitrekent, blijkt dat te kloppen: 2,94. Bijna een punt meer.

Oefeningen

1. Als de hele klas een 1 had en één student haalde een 10, zou de standaardafwijking dan groter of kleiner zijn geweest?
   Geef eerste een beredenering, dan pas een berekening.
2. Wat is de hoogst mogelijke standaardafwijking bij twintig cijfers?
3. Wat is de standaardafwijking bij één cijfer?

33. Een klein aantal metingen

Als je één meting hebt, laten we zeggen: 22 °C, kun je daar dan statistiek op bedrijven? De modus, de mediaan en het gemiddelde bestaan wel voor één meting, maar ze zijn onzinnig. Beter gezegd: ze voegen niets toe aan wat je al hebt. Ze zijn voor de 'metingenreeks' in dit geval alle drie gelijk aan 22 °C.

Hoe zit dat met de standaardafwijking? Is die gedefinieerd voor één waarde? Het antwoord is ja, maar de uitkomst is nog gekker: de afwijking tussen één meting en het gemiddelde is gelijk aan 0. *Nee, dat is niet waar: er moet nog een eenheid bij.* In ons geval is de standaardafwijking exact gelijk aan 0 °C.

Is het raar om die waarde als beste schatting aan te nemen als je één meting hebt? Nee. Maar soms hebben metingen een spreiding, en die zie je niet. Wat als we twee metingen hebben? Het gemiddelde van 22 °C en 24 °C is 23 °C. Is het zinnig om dat als beste schatting te nemen? Ja. Is er een standaardafwijking te berekenen? Dat is lastiger te zeggen.

Beide wijken 1 °C af van het gemiddelde. De standaardafwijking vind je door beide afwijkingen te kwadrateren en op te tellen. Maar eigenlijk zijn het twee keer dezelfde waarden. Wiskundig kun je dat algemeen opschrijven.

$$\bar{x} = \frac{x_1 + x_2}{2}$$

$$\Delta x_1 = |x_1 - \bar{x}| = \left| x_1 - \frac{x_1 + x_2}{2} \right| = \left| \frac{x_1 - x_2}{2} \right|$$

$$\Delta x_2 = |x_2 - \bar{x}| = \left| x_2 - \frac{x_1 + x_2}{2} \right| = \left| \frac{x_2 - x_1}{2} \right|$$

De verschillen zijn aan elkaar gelijk. Twee metingen wijken altijd evenveel af van hun gemiddelde, dat precies tussen die twee waarden in zit.

De standaardafwijking van twee getallen is gelijk aan het kwadratische gemiddelde van hun afwijkingen, die allebei gelijk zijn aan dezelfde waarde Δx.

$$\sigma = \sqrt{\frac{2(\Delta x)^2}{2}} = \Delta x$$

Op die manier bereken je dus een spreiding door het (kwadratische) gemiddelde te nemen van de afwijkingen. Maar het is helemaal geen gemiddelde. Het is gewoon de afwijking zelf. Er wordt niet opgeteld en gedeeld, want er is er maar eentje.

Is dat een vreemde keuze als waarde? Nee. Het is niet vreemd om op grond van die twee metingen te zeggen dat de temperatuur gelijk is aan:

$$T = (23 \pm 1)\ ℃$$

> Moet daar niet een significant cijfer bij? We zouden achter een 1 toch altijd nog een cijfer hebben? Nee, niet altijd. Dat doen we alleen maar als er een cijfer voorhanden is dat significant is.

Met één meting kun je wel een beste schatting doen van het gemiddelde (namelijk die ene waarde), maar je kunt *helemaal niets* zeggen over de spreiding.

Oefeningen

1. Maak een werkblad in *Excel* en neem deze getallen over in kolom A.

   22 24 23 23 23 24 24 23 21 23 24 23 24 21

2. Neem in kolom B in je werkblad een formule op waarmee je het gemiddelde van de getallen berekent. Stel de formule zodanig op dat je het aantal elementen in kolom A kunt veranderen, door gebruik te maken van een formule waarin de hele kolom wordt meegenomen.

   ```
   =GEMIDDELDE(A:A)
   ```

3. Voeg in kolom B in je werkblad formules toe waarmee je onderstaande uitkomsten kunt berekenen.
   a. de som
   b. het aantal waarden
   c. het gemiddelde (som gedeeld door aantal waarden)
   d. de mediaan
   e. de modus
   f. de grootste waarde
   g. de kleinste waarde
   h. de grootste afwijking van het gemiddelde
   i. de gemiddelde absolute afwijking
   j. het aantal waarden dat groter is dan het gemiddelde
4. Bedenk wat er verandert aan de tien uitkomsten in kolom B als je het getal 23 zou toevoegen aan de veertien waarden die je al hebt.
5. Bedenk wat er verandert aan de tien uitkomsten in kolom B als je het getal 20 zou toevoegen aan de veertien waarden die je al hebt.
6. Als de kans dat een willekeurige wereldbewoner iets heeft of doet één op miljard is, hoe groot is dan de kans dat er ergens op de wereld iemand is dat heeft of doet?

34. Eindig aantal mogelijkheden

Aan het tentamen hadden twintig studenten deelgenomen en je kon alleen maar hele cijfers halen. Het aantal mogelijke uitkomsten was dus eindig. Er waren veel mogelijke cijferlijstjes denkbaar, maar dat hield wel een keer op.

Elke student had tien mogelijke resultaten om te halen, want er stonden geen cijfers achter de komma. Hoeveel mogelijkheden zijn dat? Voor die eerste student tien, voor de tweede ook. Voor hen samen dus honderd: naast elk cijfer dat student A haalde, waren er voor B tien mogelijkheden.

Als het tentamen nu eens met 'voldoende' of 'onvoldoende' was beoordeeld? Dan waren er twee mogelijke 'cijfers' geweest. Bij twee studenten zijn er dan vier mogelijkheden. Bij twintig studenten worden dat er 2^{20} = 1.048 576. Ruim een miljoen. Best veel, maar wel een *eindig* aantal. Bij 33 studenten zou je al meer mogelijkheden hebben dan er mensen op de wereld zijn.

Als de toets door 100 studenten gedaan was, zou het aantal mogelijke cijferlijsten 2^{100} zijn geweest. Dat is 1 267 650 600 228 229 401 496 703 205 376, wat je liever schrijft als een afgeronde macht van 10: $1{,}268 \times 10^{30}$. Een getal met dertig cijfers… Nee, met 31 cijfers, want 10 tot de macht N heeft N+1 cijfers. Ga maar na: 10^2 = 100 en dat zijn drie cijfers.

Bij 150 studenten krijg je een getal dat niet meer op één regel past en als je alle cijfers van alle studenten die ooit een bepaalde studie deden zou verzamelen, praat je al gauw over een paar duizend: getallen met pakweg 300 of 400 cijfers. Ja, dat zijn flinke jongens. Maar het aantal blijft eindig.

We hadden het over de situatie waar je voldoende of onvoldoende scoort. Met tien mogelijke cijfers gaat het veel sneller. Wat is eigenlijk het grootste getal dat nog op een regel past?

10 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000

Dat is 10 tot de macht 61. Hoeveel studenten heb je nodig om zoveel mogelijkheden te hebben? Inderdaad: 61. En hoeveel studenten zorgen samen voor dit aantal als je toch weer die o/v-score invoert? Dan moet je de logaritme nemen van het getal. De $^2\log$ om precies te zijn. Hoe moest dat ook alweer?

De $\log_g$ van x reken je uit door $\log_{10}(x)$ te delen door $\log_{10}(g)$. Bereken dus twee keer de $\log_{10}$ op je rekenmachine en deel de uitkomsten. $^{10}\log(10^{61}) = 61$ en $\log_{10}(2) \approx 0{,}3010$. Delen en afronden op een geheel getal naar boven(!) leert ons dat je met 203 studenten op dit aantal mogelijkheden zit.

Stel nu eens dat we cijfers zouden geven die niet in gehele getallen werden uitgedrukt, maar één decimaal hadden. Dat je ook een 5,4 kon halen in plaats van een 5. Of een 9,9 in plaats van een 10.

Als je hele cijfers geeft, heb je tien mogelijkheden. Heb je er met één decimaal dan honderd? Nee. Als je een 10 haalt, is het cijfer achter de komma altijd een 0, omdat er hoger niet gegeven wordt. De cijfers 1 t/m 9 kunnen tien verschillende 'varianten' hebben, maar een 10 is een altijd een 10. Er zijn dus in totaal $9 \times 10 + 1 = 91$ mogelijkheden.

> Wie heeft ooit verzonnen om van 1 t/m 10 te tellen? Met een tientallig stelsel is 0 t/m 9 is veel logischer. Bovendien is 10 geen *cijfer*, maar een *getal*. Nog onlogischer wordt het als je bedenkt dat 1 het minimum is. Wie niets invult op het antwoordvel of alle meerkeuzevragen fout doet, zou ook niets (niks, nul, noppes) als waardering moeten krijgen.

Als je van 0 t/m 9 zou tellen, zou een cijfer achter de komma wél tot honderd opties leiden. Hierbij is er één probleem: je moet dan wel accepteren dat je alle cijfers naar beneden afrondt. Iemand met een 9,9 (alles goed!) zou in hele cijfers dan een 9 hebben. Die voelt zich ook bekocht.

Hoeveel mogelijke uitslaglijstjes heb je met 120 studenten en 91 mogelijke cijfers? Inderdaad: 91^{120}. Dat is $1{,}2161 \times 10^{235}$. Ja, dat is een groot getal. Hoeveel ruimte heb je nodig om al die mogelijke lijstjes op te slaan. Laten we zeggen dat je één byte nodig hebt voor een cijfer, Vooruit: een halve byte per cijfer, want met vier bits kun je ook wel tien (hele) cijfers opslaan. (Zestien zelfs.) Uitgaande van een harde schijf van één miljard terabyte heb je dan nog steeds $6{,}1 \times 10^{126}$ van die schijven nodig. Sterker nog: als elke wereldbewoner zo'n schijf ter beschikking zou stellen, red je het nog niet. De exponent daalt met negen of tien, maar dat zet geen zoden aan de dijk. Al zou elke ster in het heelal (naar schatting zo'n 10^{22}) een miljard planeten bevatten waarop alle bewoners hun kolossale harde schijven ter beschikking stelden: het schiet nog steeds tekort.

> Sommige getallen zijn domweg te groot om te bevatten. Hoeveel energie hebben we eigenlijk nodig voor zoveel computers? De zon levert een vermogen van zo'n 4×10^{22} watt. Dat is aardig wat.

Is het aantal mogelijke uitkomsten van een toets met 120 studenten en 91 mogelijke cijfers oneindig? Nee. Erg veel, laten we het daar op houden.

Oefening

Hoe zou je op een beknopte manier tentamencijfers in een database opslaan?

35. Kansvariabele

In de kansrekening is een stochastische variabele of stochastische grootheid een grootheid waarvan de waarde een reëel getal is dat afhangt van de toevallige uitkomst in een kansexperiment.

De stochastische variabele, ook toevalsvariabele, kansvariabele of stochast, is een eigenschap van de uitkomst die in een getal is uit te drukken, zoals de leeftijd of het inkomen van een toevallige voorbijganger. Het toeval bepaalt de uitkomst van het experiment, en bijgevolg is de waargenomen waarde van de stochastische variabele ook afhankelijk van het toeval. Bij een onderzoek naar de verdeling van de leeftijd is niet de toevallige voorbijganger zelf, de uitkomst, van belang, maar z'n leeftijd, een eigenschap van de uitkomst. Die leeftijd is in dit geval de stochastische variabele. Zo zijn ook het inkomen van een willekeurig gekozen Nederlander en het aantal keren dat 'kruis' gegooid wordt in een serie van 100 worpen met een munt, stochastische variabelen. Hoewel voor elke mogelijke uitkomst de waarde van de stochastische variabele vastligt, hangt de waargenomen waarde af van het toeval, als gevolg van de toevallige uitkomst. Formeel is een stochastische variabele daarmee een functie die aan elke uitkomst een getal, de waarde van de bedoelde eigenschap, toevoegt.

Bron: Wikipedia

De uitleg op *Wikipedia* van het begrip kansvariabele beslaat een halve pagina in dit boek. Is het echt zo'n lastig begrip? Of is het alleen maar moeilijk voor mensen die het niet begrijpen?

Het lijkt een beetje op de begrippen *object* en *klasse* in objectgeoriënteerd programmeren. Is 'fiets' een object of een klasse? Dat ligt er maar aan. Als je 'de fiets in z'n algemeenheid' bedoelt, dan is het een *klasse*. Als je die ene fiets bij jou thuis in de schuur bedoelt, is het een *object*.

"De fiets staat in de belangstelling van veel mensen."

Is 'de fiets' in deze zin een object of een klasse? Waarschijnlijk een klasse, want we bedoelen het *verschijnsel*, het *begrip*. Het *fenomeen*, het 'algemene ding'. Niet één bepaalde fiets. Tenzij je natuurlijk wél één heel bepaalde fiets bedoelt.

Een gek voorbeeld. In een woonwijk wordt er om onverklaarbare redenen telkens een fiets bij iemand in de tuin aangetroffen. Telkens weer bij een ander huis, niemand weet van wie die fiets is of wie dat ding daar neerzet. Het gaat al weken zo en op een gegeven moment is De Fiets een onderwerp van gesprek. Die ene fiets dus, dat rare stuk schroot op twee wielen dat overal

opduikt. In dat geval bedoel je met 'De Fiets' die in de belangstelling staat toch een specifiek voorbeeld, een heel concrete instantie van de klasse 'fiets'.

Hou dit in je achterhoofd: er is een verschil tussen het algemene (abstracte) geval en dat ene specifieke (concrete) geval.

Laten we teruggaan naar de kansvariabele, met een voorbeeld van *Wikipedia*.

Bij het gooien met twee dobbelstenen bestaat de uitkomstenruimte uit de $6^2 = 36$ paren mogelijke ogenaantallen:

$$\begin{aligned}\Omega &= \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} \\ &= \{(1,1), (1,2), \ldots, (1,6), (2,1), \ldots, (6,6)\}\end{aligned}$$

De stochastische variabele X geeft het totale aantal geworpen ogen aan:

$$X(\omega_1, \omega_2) = \omega_1 + \omega_2$$

Het waardenbereik van X is {2, 3, …, 12}.

Het totale aantal geworpen ogen is dus de stochastische variabele. Dat zijn elf (geen twaalf!) verschillende uitkomsten. Het zijn elf verschillende reële getallen, elk met een eigen kans van voorkomen.

Toch is het geen stochastische variabele meer zodra je iets anders bedoelt. In een klas zitten twaalf studenten. We willen door loting een bepaalde student ('Student 1') koppelen aan een willekeurige andere student ('Student 2' t/m 'Student 12'). Dat doen we door met twee dobbelstenen te gooien. Nadat we gegooid hebben, liggen de twee stenen op tafel. We zijn benieuwd wie het geworden is!

Het totale aantal geworpen ogen blijkt 7 te zijn. Is dat nu nog wel een kansvariabele? Nee. Als er eenmaal gegooid is, zijn er geen kansen meer. Het is zoals het is. Je zou hooguit nog kunnen zeggen dat het kansvariabele is waarbij de uitkomst 7 een kans van 1 heeft en alle andere uitkomsten een kans van 0.

Iets ingewikkelder wordt het als die twee dobbelstenen in een niet-doorzichtige beker zitten. We schudden de beker en zetten die met een behendige zwaai — zonder dat de dobbelstenen eruit vallen — omgekeerd op tafel.

Is *het totale aantal geworpen ogen* nu een kansvariabele? Of toch niet?

Je ziet aan dit voorbeeld dat het begrip kans op zich wat lastig en abstract is. Dat *totale aantal geworpen ogen* ligt na het omkeren van die beker gewoon vast! Je kunt wel hopen of bidden dat het een 7 is, maar die stenen liggen daar. Het lot is geworpen en we richten alle ogen op de dobbelsteen. We *weten* alleen het aantal nog niet. De *kans* dat je goed zo *raden* hoeveel het er zijn, zou je weer wél als een stochastische variabele kunnen zien.

“Je kunt wel hopen of bidden dat het een 7 is…”

Dit is nog een aspect van begrippen als ‘toeval’ en ‘kans’. Bestaat er wel zoiets als toeval? Is het echt willekeur wat er uit die worp komt, of beschikken God, je Karma, je Sterrenbeeld of Het Universum daarover?

Veel mensen denken dat die dobbelstenen toeval zijn, net als kop of munt. Bij het krijgen van een ernstig ongeluk of het tegenkomen van een bijzonder persoon op een bijzonder moment ervaren sommige van die mensen dat anders. “Dat kan geen toeval zijn!”

Als 60% van de studenten elk jaar het tentamen in één keer haalt, heb jij dan een slagingskans van 0,6? Of ben jij die ene die altijd alles (of nooit iets) in één keer haalt?

Ook als je niet in de invloed van God en Het Universum gelooft is dat kansbegrip lastig. Die dobbelstenen vallen niet *toevallig* zo. Dat hangt af van hoe je gooit en hoe de lucht stroomt en tot op microscopisch niveau de massa binnen de dobbelstenen verdeeld is. Toch hoef je niet te discussiëren over levensbeschouwelijke onderwerpen en je visie of dingen toeval zijn of niet om met kansen te rekenen. Dat je die ene bijzondere persoon op een bijzonder moment tegenkwam, is een eenmalige gebeurtenis. Was dat toeval? Was het voorbestemming? Statistiek leent zich niet voor eenmalige gebeurtenissen. Maar als je duizend keer met een dobbelsteen gooit, zul je zien dat het gemiddelde echt niet ver van de 3,5 ligt. En als je het honderdduizend keer doet, wordt dat nog duidelijker.

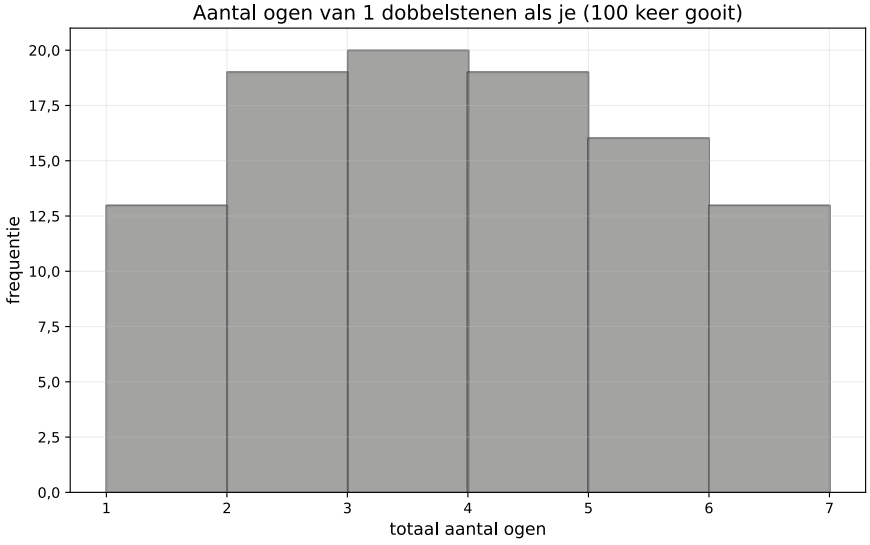
36. Normale verdeling

Een *normale verdeling*, die we (nergens anders in dit boek dan hier voor de volledigheid één keertje) ook *gaussverdeling* noemen, is een continue kansverdeling die wordt beschreven met twee parameters: het gemiddelde μ en de standaardafwijking σ . De kansdichtheid wordt gegeven door:

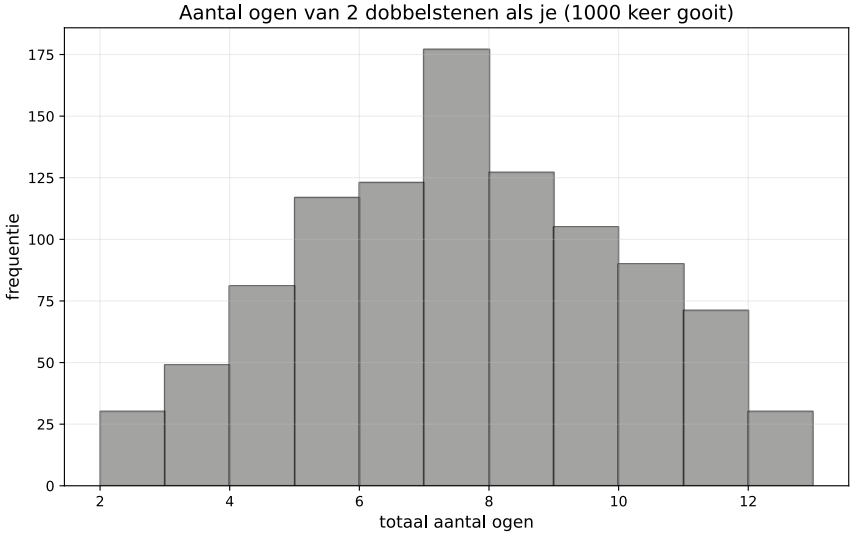
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

Een toverformule zou je haast zeggen, als je niet erg wiskundig onderlegd bent. Hij wordt wel wat makkelijker als $\mu = 0$ en $\sigma = 1$. Dan heet het een standaardnormale verdeling. Een normale verdeling ontstaat wanneer je de som neemt van een (oneindig) groot aantal willekeurige variabelen, ongeacht hun oorspronkelijke verdeling, mits je dit proces oneindig vaak herhaalt en het gemiddelde neemt. (Die bewering heet de *centrale limietstelling*.)

Dat kun je vrij eenvoudig laten zien. Als je honderd keer met een dobbelsteen gooit, krijg je histogram dat er bijvoorbeeld zo uitziet.

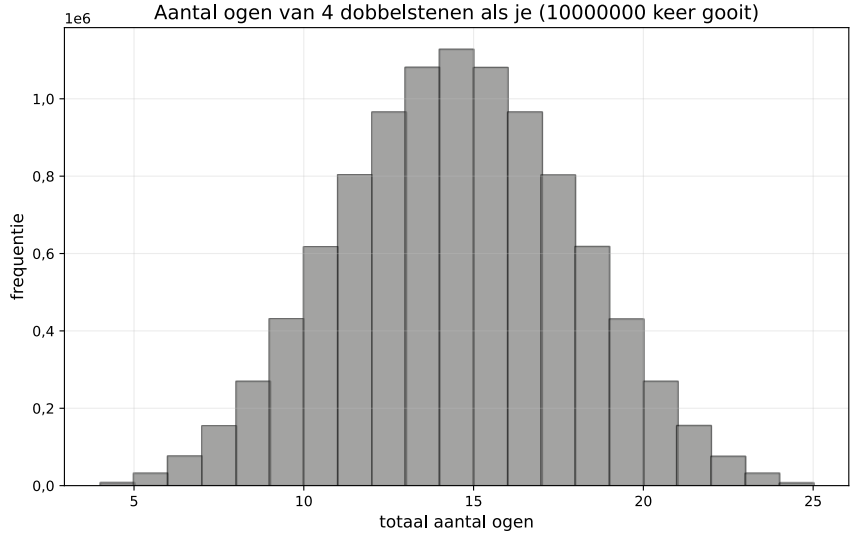


Alle mogelijkheden 1 t/m 7 komen grofweg even vaak voor. Doe je dit duizend keer met twee dobbelstenen en neem je dan de som, dan krijg je een verdeling die wat lijkt op een tweezijdige trap.

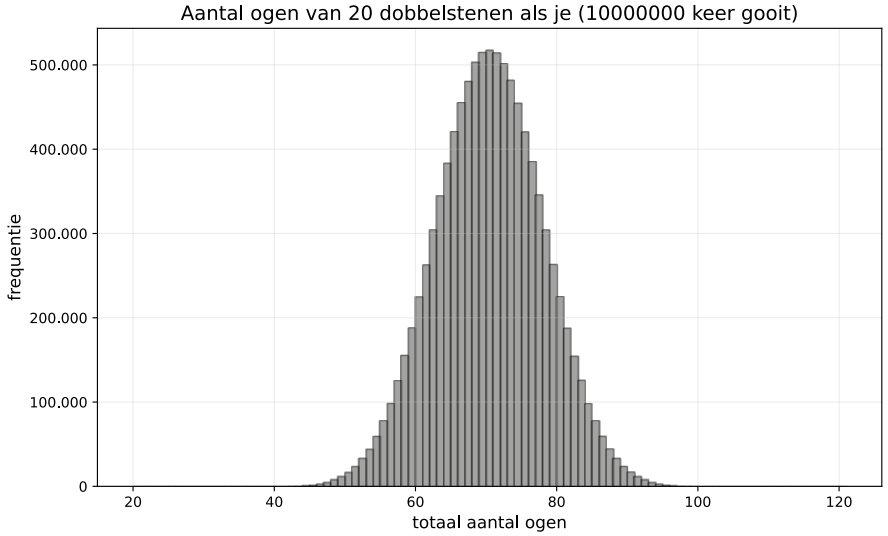


Nu zijn computers geduldig, dus waarom zou je niet tien miljoen keer gooien? Dan krijgt de trap al bijna perfect de vorm die je theoretisch kon beredeneren. De waarden 2 en 12 kun je maar op één manier gooien, 3 en 11 op twee manieren, 4 en 10 op drie manieren. Het aantal 7 kan op zes manieren. Anders gezegd: het midden komt vaak voor, de uitersten veel minder.

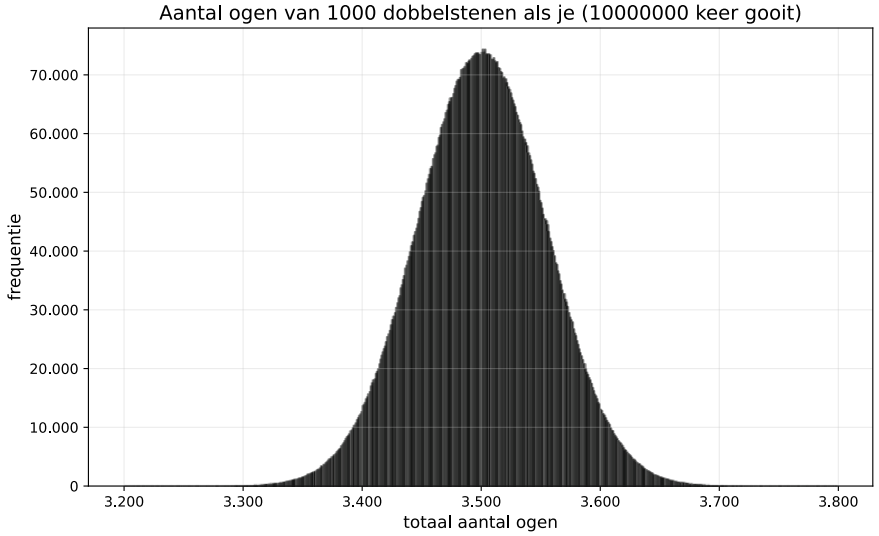
Bij vier dobbelstenen wordt het histogram wat ronder. De uitersten gaan omlaag. Twee keer 6 gooien doe je niet zomaar, en de kans dat je vier keer 6 gooit is kwadratisch kleiner.



Naarmate je meer dobbelstenen neemt (laten we zeggen twintig), krijg je steeds meer die klokvormige kromme van de normale verdeling.



Dat kun je nog veel verder opvoeren. Bijvoorbeeld met duizend dobbelstenen tien miljoen keer gooien.



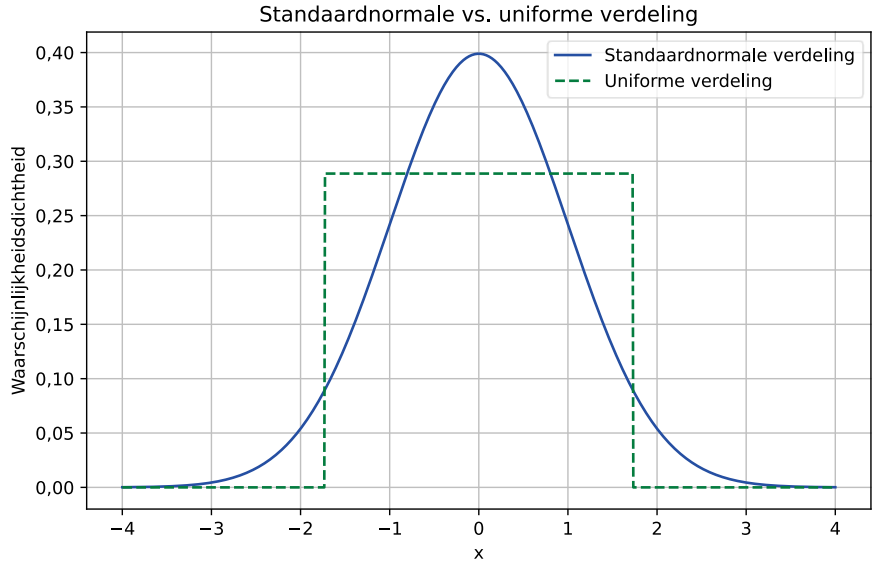
Wat je met dit eenvoudige voorbeeld ziet, is dat de uiterste waarden maar heel weinig voorkomen. Met tweehonderd dobbelstenen zou je 200 × 6 = 1200 ogen kunnen gooien, maar die kans is extreem klein.

Hoe klein dan? 1/6 tot de macht 200 en dat is kleiner dan 10^{-155}.

37. Normaal & Uniform

Barford (1985) laat in zijn boek, meteen nadat hij ‘gemiddelde’ en ‘standaardafwijking’ behandeld heeft, zien dat je met deze twee belangrijke kentallen nog niet alles gezegd hebt. Een gemiddelde alleen zegt niet heel veel, omdat je de spreiding niet kent. Maar de spreiding als enige extra gegeven, zegt ook nog niet alles.

In onderstaande figuur zijn een normale verdeling en een uniforme verdeling weergegeven die beide een gemiddelde van 0 en een standaardafwijking van 1 hebben.



Zoek de verschillen! De ‘piek’ van de normale verdeling is hoger. Hoe hoog is die eigenlijk? Pakweg 0,4… Dat kun je aflezen in de grafiek. Ja. Maar kunnen we het niet gewoon uitrekenen? Ja, en niet eens zo moeilijk ook. Kijk nog maar eens naar die kansdichtheidsfunctie van de normale verdeling.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

We hadden al gezien dat voor een standaardnormale verdeling geldt: $\mu = 0$ en $\sigma = 1$. Nu gaan we ook nog eens als $x = 0$ invullen. Dan wordt die exponent dus 0, $e^0 = 1$. Dan wordt het simpel: $f(0) = \frac{1}{\sqrt{2\pi}} \approx 0{,}3989422804$.

Het goede nieuws: die 0,4 van ons klopte aardig. Het slechte nieuws: we hebben weinig nut van dit sommetje, want dat de kansdichtheid ongeveer 0,4 is in het maximum van deze functie, is niet iets met veel praktisch nut.

Nu die uniforme verdeling. Heeft die ook een standaardafwijking? Ja. De standaardafwijking is voor een continue kansverdeling kunnen we vinden door de variantie te berekenen:

$$\sigma^2 = \int_{x=-\infty}^{\infty} f(x) \cdot (x-\mu)^2 dx$$

Als we kijken naar een uniforme verdeling met een gemiddelde van 0 en een 'breedte' van 2b, dan hebben te maken met een kansdichtheidsfunctie die 0 is voor waarden kleiner dan -b en groter dan b, en 1/2b is voor waarden tussen -b en b in. Die integraal wordt dus:

$$\int_{x=-b}^{b} \frac{1}{2b} x^2 dx = \left[\frac{x^3}{6b}\right] = \frac{b^3}{6b} - \frac{-b^3}{6b} = \frac{1}{3}b^2$$

In ons plaatje hadden we een variantie van 1, en als $\frac{1}{3}b^2 = 1$ kun je ook zeggen dat $b^2 = 3$. De b die we zoeken is dus $\sqrt{3} \approx 1{,}7321$. De kansdichtheid is gelijk aan $\frac{1}{2b} = 0{,}2886$.

Nu terug naar het plaatje met die twee verdelingen. Van een normale verdeling is bekend dat 68% van alle waarden minder dan één keer de standaardafwijking van het gemiddelde afwijkt en 95% minder dan twee keer. Hoe zit dat bij een uniforme verdeling? Die som is vrij makkelijk:

$$2 \times \frac{1}{2b} = \frac{1}{b} = \frac{1}{\sqrt{3}} = \frac{1}{3}\sqrt{3} \approx 58\%$$

Dat is dus minder dan bij de normale verdeling. Maar voor een uniforme verdeling geldt dat alle waarden minder dan $\sqrt{3} \approx 1{,}7321$ van het gemiddelde afwijken. En dus ook minder dan twee keer. Geen 95%, maar 100%.

Waar hadden we het ook alweer over? O ja. Barford (1985) laat in zijn boek zien dat je met 'gemiddelde' en 'standaardafwijking' nog niet alles gezegd hebt. Een gemiddelde alleen zegt niet heel veel, omdat je de spreiding niet kent. Maar een uniforme of normale verdeling liggen wel helemaal vast met deze twee waarden.

> Bij die laatste ware bewering moet wel worden opgemerkt dat weinig mensen een uniforme verdeling zullen beschrijven met de standaardafwijking. De onder- en bovengrens zijn wat eenvoudiger.
>
> Het gemiddelde hoef je dan al niet eens meer te noemen.

38. Verwachtingswaarde

Iemand vraagt of je lootjes wilt kopen. De opbrengst is voor een goed doel. Er worden er honderd verkocht en ze kosten één euro per stuk. Valt er iets te winnen dan? Ja! Eén prijs van € 50, drie prijzen van € 20 en vijf van € 10.

Vraag 1: Is het een goed idee om lootjes te kopen?

Vraag 2: Hoeveel dan?

> Als het antwoord op vraag 1 'nee' was, is het antwoord op vraag 2: 'nul'. Eigenlijk is het in dat geval maar één vraag dus.

Om tot je keuze te komen, kun je gebruikmaken van de *verwachtingswaarde*.

De verwachtingswaarde van een stochastische variabele is wat je 'gemiddeld' kunt verwachten als je het experiment oneindig vaak uitvoert — of, formeler gezegd: het gewogen gemiddelde van alle mogelijke uitkomsten, met de kansen als gewichten.

> Pascal & Fermat bedachten in 1654 dit begrip bij hun oplossing van het zogenaamde *puntenprobleem*. Twee partijen spelen een spel waarbij je een gelijke kans hebt om te winnen. De partij die als eerste zes keer heeft gewonnen, wint de pot met zestig muntstukken. Stel dat het spel afgebroken moet worden bij een stand van 5-3, hoe moet de pot dan verdeeld worden? Oplossing: zie *Wikipedia* ('puntenprobleem').

Bij die loterij is er een kans van 1% dat je die € 50 wint, 3% dat je € 20 wint en 5% dat je € 10 wint. De verwachtingswaarde van de 'winst' bij één lot dus:

0,01 × € 50 + 0,03 × € 20 + 0,05 × € 10 = € 0,50 + € 0,60 + € 0,50 = € 1,60

Dat is niet de *winst*, maar de *opbrengst*. Om de winst te berekenen, moet je de *kosten* (€ 1) ervan aftrekken. De verwachtingswaarde van de winst is € 0,60.

Is de kans dat je geld wint groter dan de kans dat je geld verliest? Nee! De kans dat je iets wint is namelijk 1% + 3% + 5% = 9% en de kans dat je verliest is dus 91%. Maar als je verliest, verlies je niet veel. Als je wint, win je veel meer dan je zou verliezen als je verloor.

Er is overigens wel nog een probleem. De opbrengst is voor een goed doel… Als jouw kans op winst groter is, is de kans op verlies voor dat goede doel kleiner. Je kunt dus beter geen loten kopen en een tientje overmaken aan dat goede doel. Een definitieprobleem? Nee. De begrippen zelf zijn wel duidelijk. We hebben alleen het *doel* van het kopen van de lootjes niet bepaald.

Als je het doel niet goed kent, weet je niet wat je het beste kunt doen.

Iemand laat twee enveloppen zien waar geld in zit. Je mag er eentje hebben. Het enige wat je weet, is dat in de ene envelop *twee keer zoveel* zit als in de andere. Je maakt de envelop open: € 10. 'Wil je nog ruilen?'

Je denkt even na. In die andere envelop zit óf de helft (€ 5) óf het dubbele (€ 20). Beide mogelijkheden zijn even waarschijnlijk (50%).
De verwachtingswaarde van de inhoud van de envelop die je krijgt door te ruilen is dus € 12,50… Toch?

Maar wat nou als je die andere als eerste had genomen?

De enveloppenparadox.

Oefeningen

1. Wat is de verwachtingswaarde van het kwadraat van het totaal aantal ogen dat je gooit met twee dobbelstenen?

2. Waar zit de redeneerfout in de enveloppenparadox?

De oefeningen hieronder gaan uit van de in dit hoofdstuk genoemde loterij.

3. Bekijk twee verschillende situaties.
   A. Je doet twee keer mee met de loterij.
   B. Je doet één keer mee en je koopt twee lootjes.

   Wanneer is de verwachtingswaarde van de winst groter?

4. Wat is de verwachtingswaarde van de winst bij vijftig loten?

5. Bij welk aantal loten is de winst geen stochastische variabele?

6. Bekijk twee verschillende situaties.
   A. Je doet twee keer mee met de loterij en koopt beide keren tachtig lootjes.
   B. Je doet één keer mee en je koopt 92 lootjes.

   Wanneer is de verwachtingswaarde van de winst groter?
   Wanneer is de kans dat je een prijs wint groter?

7. Bedenk argumenten om wel of niet mee te doen met de *Nationale Postcode Loterij*.

39. Permutaties & Combinaties

De filosoof en wetenschapper Leibniz schreef in 1714 in een brief dat het even makkelijk is om 11 als 12 te gooien met twee dobbelstenen. Er was namelijk volgens hem maar één manier om elk van beide getallen te gooien.

Het klopt dat je 11 met twee dobbelstenen alleen maar kunt gooien met een 5 en 6. En het klopt ook dat 12 alleen maar kan met twee zessen. Maar wat Leibniz — toch echt geen domme jongen — over het hoofd zag, is dat er *twee* manieren zijn om 5 en 6 te gooien: eerst een 5 en dan een 6, of eerst een 6 en dan een 5.

Wanneer zou je die denkfout niet maken? Misschien als je heel goed nadacht over alle mogelijkheden. Met één dobbelsteen heb je zes mogelijke uitkomsten. Met twee dobbelstenen heb je voor elk van die zes uitkomsten opnieuw zes mogelijkheden. Je hebt er dus $6 \times 6 = 36$. En als je die alle 36 in een tabelletje zet, zie je ook dat niet alleen 5 en 6, maar ook 6 en 5 de gevraagde 11 oplevert.

Je ziet dan ook meteen dat 7 de meest waarschijnlijke uitkomst is. Die kun je namelijk vormen uit 1+6, 2+5, 3+4, 4+3, 5+2 en 6+1. De kans om 7 te gooien is maar liefst zes keer zo groot als de kans op 12.

> Dat Leibniz zich vergiste, is menselijk. Misschien heeft hij heel hard gelachen toen iemand hem op die domme fout wees. Het zou ook een grap geweest kunnen zijn trouwens. Er zijn mensen die beweren dat de kans op het winnen van de Staatsloterij 50% is, om het simpele feit dat er maar twee mogelijke uitkomsten zijn. Je wint 'm wel of je wint 'm niet.

Dat iets moeilijk is, maakt het makkelijker om jezelf niet te verwijten dat je het fout doet. Daar schiet je sowieso niet veel mee op. Maar met *voorkomen* dat je fouten maakt en/of ze zo snel mogelijk vinden, schiet je wél iets op.

Bij statistische dingen kan het helpen dat je een simulatie schrijft met een computerprogramma. Bijvoorbeeld in *Python*. Even zoeken met Google en je weet hoe je een *random number* genereert. In dit geval: een geheel willekeurig getal van 1 t/m 6, want het is een dobbelsteen.

Zou dit 'm zijn?

```
import random

dobbelsteen1 = random.randint(1,6)

print(dobbelsteen1)
```

Nee! Als je dit programma test, lijkt het werken. Er komt 3 uit of zo. Of 1. Of 5, of 4, of nog een keer 3. Maar nooit 6, want het tweede argument van `randint()` is de bovengrens en die wordt nooit bereikt. En om dat aan te tonen, doe ik het honderd keer.

2 5 5 1 1 3 1 1 3 4 5 5 2 6 3 6 2 1 2 1 4 2 2 3 5 4 5 4 6 3 2 6 2 2 1 1 6 4 5 3 5 6 5 4 1 4 3 2 4 4 4 6 1 6 1 6 3 4 2 2 4 5 6 1 6 1 1 5 4 1 3 5 6 4 3 6 3 3 2 1 4 2 3 2 3 1 4 3 4 3 2 1 1 5 1 4 3 4 1 4

Hé, daar staat dus toch 6 tussen… Waarom dacht ik dan van niet? Omdat in sommige programmeertalen en functies die bovengrens niet meedoet. Neem het Java-programma hieronder. Als je denkt dat die 6 ervoor zorgt dat de opties 1 t/m 6 voorkomen, dan heb je het mis. En dan nog iets: deze code kan ook de waarde 0 opleveren. Er zijn wel zes mogelijke uitkomsten, maar dat zijn dus de getallen 0 t/m 5. Er is niet veel voor nodig om het correct te krijgen. Maar dat geldt voor zo veel bugs.

```
public static void main( String args[] ) {
    Random rand = new Random();
    int upperbound = 6;
    using nextInt()
    int int_random = rand.nextInt(upperbound);
    System.out.println("Random integer value from 0 to" +
          (upperbound-1) + " : " + int_random);
    }
```

Hoe voorkom je nu deze fout? Geloof het of niet: deze vijf dingen helpen.

1. De specificaties van de functie lezen.
2. Goed nadenken van tevoren.
3. Je concentreren tijdens het schrijven van de code.
4. Heel veel testen op een slimme manier.
5. Iemand anders ernaar laten kijken.

Dat ik het programma honderd keer uitvoerde, was geen bewijs dat het goed werkte. Maar ik had aan de uitvoer wel een aantal dingen gezien die me een beetje geruststelden.

1. Het getal 0 kwam nooit voor
2. Het getal 1 wel
3. Het getal 6 ook
4. Alle getallen daartussenin ook
5. Geen enkel getal opvallend vaak

Dat vierde geloofde ik sowieso wel, maar toch goed om te zien.

Nu ik weet hoe je willekeurige getallen maakt in *Python*, kan ik mijn programma schrijven. Ik kom op deze code.

```
import random
import numpy as np

uitkomst = np.zeros(13)

for i in range(100):
    dobbelsteen1 = random.randint(1,6)
    dobbelsteen2 = random.randint(1,6)
    worp = dobbelsteen1 + dobbelsteen2
    uitkomst[worp] = uitkomst[worp] + 1

print(uitkomst)
```

Ik test het programma en kom op deze uitvoer.

```
[ 0.  0.  0.  2.  9. 14.  9. 17. 18. 13. 14.  4.  0.]
```

Tsja. Wat moet je ervan zeggen? Het 'valt op' dat de waarden aan de rand niet of weinig voorkomen en in het midden veel meer. Dat had ik al verwacht, vanwege die vele manieren om 7 te gooien. Maar bewijst dit nu iets?

Nee. Iets bewijzen doet een simulatie zelden of nooit. Maar waarom herhaal ik het eigenlijk 100 keer en geen 1000 keer, of 10000 keer of 100.000 keer? Zelfs een miljoen keer kan mijn computer nog wel in een paar seconden aan. Tien miljoen lukt ook nog binnen een minuut.

```
[      0.       0.  277714.  556593.  831557. 1111509. 1387420
 1666589. 1391111. 1109824.  834771.  556060.  276852.]
```

Bewijst dit nu wél wat? Nee. Bewijzen doet zoiets zelden, zeiden we al. Maar als alle uitkomsten meer dan honderdduizend keer voorkomen en 0 en 1 niet, dan doet het wel iets vermoeden. Het bevestigt wat ik al dacht: dat je nooit minder dan 2 kunt gooien met twee dobbelstenen. Dat 11 twee keer zo vaak voorkomt als 12, begint er ook aardig op te lijken. (Deel de twee getallen op elkaar en je komt op 2,0087. Dat wijkt minder dan 1% af van 2.)

Maar het bewees toch allemaal niets? Nee, niet echt. Maar ik heb nu twee dingen gedaan.

1. goed nadenken over statistiek
2. mijn best doen op een goed programma dat het simuleert

Als wat daar uitkomt met elkaar overeenstemt, dan klopt het misschien wel. Meer weet je niet, maar het is al veel meer dan niets.

40.Simulatie van een kansvariabele

Hoe groot is het aantal ogen van een dobbelsteen die je op tafel gooit? Voor een gewone dobbelsteen is het antwoord: 21. Die heeft namelijk zes kanten, met in totaal 1 + 2 + 3 + 4 + 5 + 6 = 21 ogen.

Meestal bedoelen we met het aantal ogen niet het totale aantal ogen, maar het aantal ogen van de kant die boven ligt. Dan zijn er zes mogelijke antwoorden: 1, 2, 3, 4, 5 of 6. Als je het aantal ogen van één bepaalde worp bedoelt tenminste! Het aantal ogen *in het algemeen*, het aantal ogen van *een toevallige worp*, een worp die je *niet doet maar zou kunnen doen*; dat is een ander verhaal. Die vraag heeft maar één antwoord: 1, 2, 3, 4, 5 of 6, met elk van die uitkomsten een kans van 1/6.

We hebben nu dus verschillende antwoorden op dezelfde vraag.

- 21: het totale aantal
- 5: het aantal ogen van de worp die ik zojuist deed
- 2: het aantal ogen van de worp die ik daarvoor deed
- 6: het aantal ogen dat ik hoopte te gooien
- P(X): een uniforme discrete kansverdeling van 1 t/m 6

Veel mensen die statistiek lastig vinden, lijken moeite te hebben met het begrip *kansvariabele*. Zo'n soort variabele heeft namelijk niet een bepaalde waarde, maar beschrijft alle mogelijke uitkomsten en de kans op elk van die mogelijke uitkomsten.

> *In de kansrekening is een stochastische variabele of stochastische grootheid een grootheid waarvan de waarde een reëel getal is dat afhangt van de toevallige uitkomst in een kansexperiment. De stochastische variabele, ook toevalsvariabele, kansvariabele of stochast, is een eigenschap van de uitkomst die in een getal is uit te drukken, zoals de leeftijd of het inkomen van een toevallige voorbijganger.* *Wikipedia*
>
> Er in dit soort definities altijd wel een mogelijkheid om ergens moeilijk over te doen. Een toevallige voorbijganger… bedoel je daarmee *die ene* toevallige voorbijganger, die vijf minuten geleden voorbijkwam? Die heeft heus wel één bepaalde leeftijd. Bijvoorbeeld 35 jaar. Of 21 jaar. Misschien wel 64 jaar of 11 jaar. Of 111 jaar, maar dat is minder waarschijnlijk. Hoewel… Als die ene voorbijganger toevallig 111 jaar was, dan is daar niets onwaarschijnlijk meer aan. Het is gewoon waar.

Het lastige van kansvariabelen is dat je hun uitkomst niet als een hard getal kunt berekenen. Het hoogst mogelijke aantal ogen met twee dobbelstenen is

makkelijk: dat is 12. Het gemiddelde aantal ogen met twee dobbelstenen is ook niet moeilijk: dat is 7. Maar dat zijn geen kansvariabelen, ook al is het gemiddelde van een verdeling wel iets wat in de statistiek wordt gebruikt.

De schatting van het gemiddelde op basis van een aantal worpen is wél een kansvariabele. Als je één keer gooit met één steen, komt er volstrekt willekeurig een 1, 2, 3, 4, 5 of 6 uit. Dat zegt maar heel weinig over het gemiddelde. Je zult nooit 0 of 7 gooien en ook geen -3 of 42. En geen 9,8 ook. Maar naarmate je vaker gooit, zal de uitkomst gaandeweg dichter bij de 3,5 liggen.

> Wat gebeurt er met het gemiddelde aantal ogen als je na N keer gooien nog een keer gooit? Komt dat dichter bij het echte gemiddelde te liggen? Antwoord: dat weet je niet. Stel dat het gemiddelde tot nu toe 3,75 was. Dan komt het verder van het echte gemiddelde te liggen als 4, 5 of 6 gooit. En als je 1 gooit? Dan komt het er dichter bij. Hoewel… Als die 3,75 berekend is uit de vier worpen 6 + 2 + 3 + 3 = 14, dan is dat juist niet zo. Want 14 / 4 = 3,75. En 15 / 5 = 3, dat verder van 3,5 ligt dan 3,75.

We zouden het programma kunnen uitbreiden door vaker dan één keer te werpen en de resultaten te bewaren in een array. Dan kunnen we niet alleen zien of alle waarden voorkomen, maar ook of dat even vaak gebeurt.

```
import random
import numpy as np

aantal_worpen = 10
uitkomst = np.zeros(7, dtype=int)

totaal = 0
for i in range(aantal_worpen):
    ogen = random.randint(1,6)
    uitkomst[ogen] = uitkomst[ogen] + 1

print(uitkomst)
```

Hé, een array met een lengte van zeven … Er zijn toch maar zes mogelijke uitkomsten? Ja. Maar het is handig om het aantal gegooide ogen als index te nemen. Dan moet de index lopen van 0 t/m 6, omdat een array bij 0 begint.

Tien keer gooien is wel erg weinig. Doe het liever honderd, duizend of nog liever honderdduizend keer. Het bewijst weinig, maar helpt je wel fouten te vinden. De 0 mag niet voorkomen en 1 t/m 6 ongeveer even vaak. Er zullen verschillen zijn. Mocht de 6 ontbreken, dan heb je een programmeerfout.

Of hij kwam echt niet voor. Dat zou wel héél toevallig zijn.

We gaan nu naar een ander programma kijken, dat hier een beetje op lijkt. We dobbelen nog steeds, maar delen het aantal ogen door 10 en tellen er dan telkens een bepaald getal bij op. Laten we zeggen: 42.

```
import random

ogen = random.randint(1,6)
meetwaarde = 42 + ogen / 10
print(meetwaarde)
```

Wat is de uitkomst van dit programma? Toen ik het deed, kwam er 42,3 uit.

> Op mijn scherm stond een punt, maar in het Nederlands gebruik je een komma als scheidingsteken en Meten & Weten is een Nederlands boek. Beter gezegd: in het Nederlands *hoor je een komma als scheidingsteken te gebruiken*. Lang niet iedereen doet dat.

Deze vraag lijkt op waar we mee begonnen:

Wat is de uitkomst van dit programma?

lijkt op

Hoe groot is het aantal ogen van een dobbelsteen die je op tafel gooit?

De antwoorden

42,3

en

een willekeurig getal van 42,1 t/m 42,6 met één decimaal

zijn beide juist. In het eerste geval bedoel je die ene specifieke uitkomst van die ene simulatie van mij, op een willekeurige dinsdagmorgen. In het tweede geval heb je het dus over de *kansvariabele*.

De uitkomsten van dat dobbelen en van die 'meetwaarde' zijn allebei toevallig. Bij de dobbelsteen lijkt het zuiver toeval, bij die meetwaarde is het toevallige deel veel kleiner dan het 'vaste' deel. In beide gevallen geldt dat er zes mogelijke uitkomsten zijn die onderling even waarschijnlijk zijn.

Metingen bestaan in ons model uit een vast en een variabel deel. Dat variabele deel heeft als gemiddelde de waarde 0. Je zou kunnen zeggen dat er bij de werkelijke waarde een kansvariabele wordt opgeteld. Misschien was de werkelijke waarde van mijn programma wel 41 en werd daar een kansvariabele bij opgeteld (de fout!) die als uitkomsten 1,1 t/m 1,6 had. Dan hadden we te maken met een systematische fout van 1 en een uniform verdeelde toevallige fout met de discrete uitkomsten 0,1 t/m 0,6.

41. Betere schatting van spreiding

Als je een steekproef neemt uit een normaal verdeeld proces, dan kun je uit die steekproef een gemiddelde en variantie (en standaardafwijking) bepalen. We berekenen dan op grond van tien willekeurige getallen van het (in theorie) normaal verdeelde proces met gemiddelde μ en standaardafwijking σ. Dat gemiddelde van die ene steekproef duiden we aan met $\overline{x}$ en s_x.

Vaak doe je maar één steekproef en zorg je ervoor (of ga je er zonder nadenken van uit) dat die groot genoeg is om de waarden $\overline{x}$ en s_x dicht in de buurt te laten komen van wat je eigenlijk zou willen weten, namelijk μ en σ.

Wanneer zijn het er genoeg? Twee of drie is wel erg weinig en duizend klinkt als erg overdreven. Maar welke waarde daartussenin is goed genoeg?

Laten we voor de aardigheid eens tien verschillende steekproeven van tien gesimuleerde grootheden doen. De uitkomsten daarvan zetten we in een tabel. De kolommen zijn de steekproeven, de bovenste tien waarden zijn de gevonden metingen en onderaan staan de $\overline{x}$ en s_x per steekproef.

We zien in de tabel dat $\overline{x}$ varieert tussen 4,27 en 5,46 (een onderling verschil van 28 %) en dat s_x varieert tussen 0,73 en 1,21 (onderling verschil: 66 %).

	4,59	5,55	3,76	4,56	6,54	5,60	5,91	5,65	4,76	4,37
	4,43	6,74	6,94	5,21	3,75	5,87	4,96	5,77	5,77	4,34
	3,44	6,30	4,92	5,85	5,23	5,86	3,57	4,64	5,82	4,20
	3,59	5,48	6,01	3,67	5,23	5,47	4,26	3,86	5,27	4,76
	3,15	6,80	3,43	6,79	6,02	6,27	3,93	3,41	5,14	3,53
	5,74	4,98	5,18	3,25	4,52	7,04	5,16	4,87	4,09	4,79
	5,01	3,89	5,41	3,95	5,81	3,56	7,31	6,29	4,91	5,15
	4,14	4,58	5,64	5,31	5,44	5,90	6,41	5,97	4,88	5,38
	5,32	5,09	5,41	3,50	4,57	4,11	6,15	4,56	8,37	4,00
	3,27	4,53	4,77	6,02	5,33	4,88	5,96	5,32	5,37	2,98
$\overline{x}$	4,27	5,39	5,15	4,81	5,24	5,46	5,36	5,03	5,44	4,35
s_x	0,90	0,98	1,02	1,21	0,81	1,03	1,19	0,94	1,15	0,73

Het gemiddelde van die tien gemiddelden is overigens 5,05 en het gemiddelde van die tien standaardafwijkingen is 0,99. Dat wijkt allebei 1 % af van 5 en 1… Laten 5 en 1 nu precies de μ en σ zijn waarmee deze getallen gegenereerd zijn.

Als je diezelfde tien steekproeven nog een keer doet, zul je tien andere gemiddelden en tien andere standaardafwijkingen vinden. Het gemiddelde daarvan zal waarschijnlijk dicht in de buurt liggen van 5 en 1.

Nu de wiskunde erachter. Wat hieronder volgt, is eigenlijk niet echt moeilijk, maar het is wat lastig om goed onder woorden te brengen.

Als x een stochastische variabele is die normaal verdeeld is met een gemiddelde van μ en standaardafwijking σ, dan zijn de $\overline{x}$ en s_x die uit een steekproef berekend worden zelf óók normaal verdeeld. De ene steekproef levert een ander gemiddelde op dan de andere en ook standaardafwijkingen verschillen. Het zijn zelf ook *kansvariabelen*, die afhangen van de toevallige uitkomst van die tien (of hoeveel dan ook) metingen.

Wat is het gemiddelde van al die gemiddelden? Dat weet je niet precies, maar de verwachtingswaarde is wel gelijk aan μ. De normale verdeling van die steekproefgemiddelden heeft namelijk dezelfde μ. Eigenlijk ook wel logisch, want de kans op een waarde die afwijkt naar boven is even groot als de kans op dezelfde afwijking naar beneden. Als je oneindig vaak zou middelen, is de gemiddelde waarde van $\overline{x}$ — net als die van x — gelijk aan μ.

Hoe groot is de *standaardafwijking* van dat gemiddelde? Die vraag is wat minder simpel, maar gevoelsmatig zeg je dat die kleiner zal zijn dan de standaardafwijking van de verdeling zelf. In een gemiddelde van tien metingen zal minder spreiding zitten dan in de meting zelf.

Als je de lengte van studenten in één klas op een rij zet, zul je een flinke spreiding zien. Afhankelijk van wat je 'flink' noemt. Maar vaak zit er wel iemand van 2 meter of langer tussen en je hebt ook studenten van 1,60 meter of kleiner. Als je de gemiddelde lengte van alle studenten in één klas vergelijkt met de gemiddelde lengte van alle studenten in andere klassen, zal die spreiding veel minder groot zijn. Je vindt nooit een klas waar studenten gemiddeld 2 meter lang zijn, tenzij je ze heel bewust geselecteerd hebt. De spreiding van de gemiddelden is nu eenmaal kleiner dan de spreiding van de individuele waarden.

Wiskundig gezien kun je bewijzen dat de standaardafwijking van het gemiddelde gelijk is aan

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}$$

De formule is wat groot afgedrukt, omdat hij nogal belangrijk is. Daarom is het ook zo belangrijk dat je snapt wat er staat.

Om met de makkelijkste te beginnen: N is de grootte van de steekproef. Dan die ene die we al wat langer kennen: σ_x, de standaardafwijking van alle mogelijke metingen die je kunt doen. (Van de onderliggende verdeling dus.)

De lastigste is $\sigma_{\bar{x}}$, die met dat streepje boven de x. Dat is namelijk de *standaardafwijking van het steekproefgemiddelde*, in het Engels vaak aangeduid als *Standard Deviation of the Mean* of kortweg *Standard Error*.

Zoals σ_x een handige maat is om aan te geven hoe groot de spreiding is van de *individuele metingen*, zo is $\sigma_{\bar{x}}$ een handige maat voor de spreiding in de *gemiddelden van de steekproeven*. Het is meestal een getal dat je lekker klein wilt krijgen. Want als de spreiding in de gemiddelden van de steekproeven klein is, dan maakt het niet zo veel meer uit welke steekproef je neemt. Er komt toch telkens ongeveer hetzelfde uit.

Die σ_x die in de teller van de breuk staat, daar doe je meestal niet zo veel aan. Dat is gewoon de spreiding die in je metingen zit. Die N is een ander verhaal. Vaak kun je zelf kiezen hoe groot je steekproef is.

Het is belangrijk om in te zien dat die N een getal is waarvan eerst de wortel wordt getrokken voordat er gedeeld wordt. Als je zou willen dat de spreiding twee keer zo klein wordt, zul je dus vier keer zoveel metingen moeten doen.

> Stel, je staat voor een zaal met honderd volwassen mensen, evenveel mannen als vrouwen. Je vraagt één man en één vrouw om op het podium te komen. Zal die man dan de langste van de twee zijn? Waarschijnlijk wel, maar het hoeft natuurlijk niet. Als je toevallig een klein mannetje of een Grote Vrouw uitkoos, dan is zij langer dan hij.
>
> Wat nu als je de gemiddelde lengte bepaalt van die honderd mannen én van die honderd vrouwen. Zal dat gemiddelde van die mannen dan groter zijn? Waarschijnlijk wel. ZEER waarschijnlijk wel. Of je moet héél toevallig veel klein mannetjes en Grote Vrouwen in de zaal hebben. Met honderd mannen en vrouwen is die kans erg klein.

Oefeningen

Stel dat volwassen mannen gemiddeld 1,80 meter en volwassen vrouwen gemiddeld 1,70 meter. De standaardafwijkingen: 8 cm en 6 cm.

1. Schat de kans dat een willekeurige vrouw groter is dan een willekeurige man.
2. Schat de kans dat de gemiddelde lengte van 100 willekeurige vrouwen groter is dan de gemiddelde lengte van 100 willekeurige mannen.

42. De spreiding in de spreiding

In het vorige hoofdstuk zagen we dat het gemiddelde van een steekproef van normaal verdeelde metingen zelf ook normaal verdeeld is. Het gemiddelde van die verdeling is μ. Als de gemiddelde lengte in een groep vrouwen 1,70 m is, dan zal het gemiddelde van een willekeurige steekproef ook 1,70 m zijn.

De spreiding van de steekproefgemiddelden is kleiner en kan worden berekend door te delen door de wortel van de steekproefgrootte.

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}$$

Als de standaardafwijking in de lengte van die groep vrouwen 5 cm is, dan zal de standaardafwijking in het gemiddelde van een steekproef van 25 gemiddeld genomen maar 1 cm zijn.

Uiteraard geldt ook hier: de ene groep van 25 vrouwen is de andere niet.

De uitspraak is natuurlijk niet te 'bewijzen' met een simulatie, maar het is niet heel moeilijk om (bijvoorbeeld in *Excel*) een tabel te genereren met tien 'steekproeven' onder 25 fictieve vrouwen die een normaal verdeelde lengte hebben met een gemiddelde van 1,70 m en een standaardafwijking van 5 cm. De resultaten van de gemiddelden en standaardafwijkingen per steekproef zijn als volgt.

171	171	170	169	169	171	169	172	170	169
5,1	6,0	5,1	5,6	4,7	3,8	7,0	6,3	5,5	3,9

Natuurlijk kun je uit deze twee regels weer een gemiddelde bepalen: het gemiddelde van de steekproefgemiddelden is 1,702 m en de standaardafwijking van de steekproefgemiddelden is 0,90 m. (We verwachtten 170,0 en 1,0 m. Het komt in de buurt. En het bewijst niets, maar geeft een goed gevoel en misschien wat meer begrip.)

Nu de laatste stap, dus hou nog even vol. Niet alleen het steekproefgemiddelde verschilt per steekproef, maar ook de spreiding (standaardafwijking) is voor elke steekproef anders. De standaardafwijkingen van de steekproeven verschillen van elkaar, en vertonen dus zelf ook een spreiding. Eigenlijk zou je die spreiding ook willen kennen. Alleen als de spreiding in die standaardafwijkingen klein is, volstaat het om er slechts één te bepalen.

Ben je er nog? Mooi. Hier komt de formule.

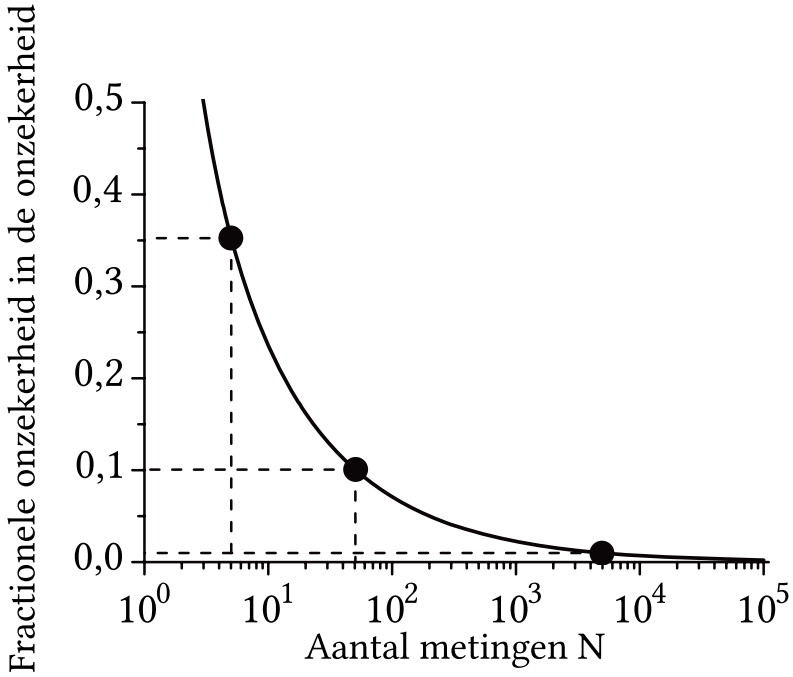
$$\sigma_s = \frac{\sigma}{\sqrt{2N-2}}$$

In ons geval was σ gelijk aan 5 cm en N gelijk aan 25. Invullen:

$$\sigma_s = \frac{\sigma}{\sqrt{2N-2}} = \frac{5}{\sqrt{2 \cdot 25 - 2}} \approx 0{,}72\ cm$$

Wat het een beetje ingewikkeld maakt: we kennen wel de N in de formule, maar in de praktijk meestal niet de σ. Als we niet wisten dat hij 5 cm was (wat we wel weten, we hebben immers zelf die simulatie geschreven), dan zouden we schatten dat hij 5,3 cm was, omdat dat het gemiddelde is van de standaardafwijkingen die we uit de steekproeven haalden. Die 0,72 cm helpt ons om een beeld te vormen van hoe betrouwbaar die 5,3 cm was.

Hughes & Hase nemen in hun boek een grafiekje op dat een indruk geeft van het verband tussen het aantal metingen en de (fractionele) onzekerheid in de onzekerheid. Zij hebben het dan over de *fractional error in the error*, maar omdat ze het over de spreiding (standaardafwijking) hebben, bedoelen ze niet de *fout*, maar de *onzekerheid*.



Figuur 42.1 De afname van de onzekerheid bij een toename van het aantal metingen

Getallen in een tabel kunnen ook helpen om een gevoel te krijgen.

N	$\frac{1}{\sqrt{N}}$	$\frac{1}{\sqrt{2N-2}}$
1	100%	
2	71%	71%
3	58%	50%
4	50%	41%
5	45%	35%
10	32%	24%
20	22%	16%
50	14%	10%
100	10%	7%
500	4%	3%
1000	3%	2%
10000	1%	1%

Wie één meting doet, heeft een spreiding die net zo groot is als de spreiding in de verdeling zelf. Vier keer meten en dan middelen, halveert de spreiding. Je hebt er honderd nodig om tot 10% te komen en tienduizend voor 1%.

We hebben nu verschillende woorden voorbij zien komen die iets zeggen over spreiding en allemaal beginnen met ‘standaard’. Hoog tijd voor een overzicht:

standaardafwijking (verdeling)	spreiding van de theoretische verdeling (alleen bekend in wiskundesommen)	σ
standaardafwijking (steekproef)	spreiding in de gemeten waarden (verschilt per reeks metingen)	s
standaardfout	spreiding van gemiddelden bij herhalen (term uit de statistiek)	$\frac{s}{\sqrt{N}}$
standaardonzekerheid	beste schatting van onzekerheid (GUM) (term uit de metrologie, ‘ons’ vak dus)	u
standaarddeviatie	synoniem van standaardafwijking (wordt in Meten & Weten niet gebruikt)	σ

Standaardfout en *standaardonzekerheid* hebben vaak dezelfde formule, maar niet hetzelfde doel. De eerste zegt iets over spreiding van gemiddelden, de tweede over onzekerheid van één gerapporteerde waarde. Dat laatste is waar we bij metingen mee bezig zijn.

43. Normaal & Uniform verdeelde fouten

In dit hoofdstuk kijken we naar twee veelvoorkomende bronnen van meetonzekerheid: de 'algemene' toevallige fout en de digitale afleesfout. Ze lijken op het eerste gezicht best op elkaar: in beide gevallen heb je een onnauwkeurigheid, maar de wiskunde erachter is verschillend.

Eerst de 'algemene' toevallige fout. Die ontstaat als je een paar keer hetzelfde meet, en telkens een net iets andere uitkomst krijgt. Dat kan komen door trillingen, temperatuurverschillen, luchtstromen, het al dan niet goed aandrukken van een schuifmaat, enzovoort. Zulke invloeden zijn onvoorspelbaar en veranderen van moment tot moment. Je meet vijf keer, en krijgt vijf waarden. Daaruit bereken je het gemiddelde en de standaardafwijking s van de metingen. De onzekerheid u van het gemiddelde is dan:

$$u = s/\sqrt{N}$$

waarbij N het aantal metingen is. Deze u noemen we de standaardonzekerheid ten gevolge van toevallige fouten.

> Volgens de GUM is dit een type A-onzekerheid. Duidelijk, want de grootte is op een *statistische analyse* gebaseerd.

Als de meetfouten normaal verdeeld zijn, dan ligt de werkelijke waarde met ongeveer 68 procent kans binnen het interval $\bar{x} \pm u$

> Het blijft een vreemde keuze, die standaardafwijking. Of eigenlijk: dat we die één keer nemen. Het is lang niet zeker dat de werkelijke waarde erbinnen ligt. Die kans is maar 68,3%. Bij $2\sigma$is die kans al 95,4%. Bij 3σ wordt dat 99,7%, bij 4σ zelfs 99,9937%. bij 9σ staan er zoveel negens achter de komma, dat *Excel* het met z'n vijftien decimalen niet meer aankan. Toch is ook dat niet zeker. Dat wordt het nooit.
>
> Dan is één keer de standaardafwijking toch wel logisch. Twee keer zou misschien wel logischer zijn. Als je iets niet zeker weet, is niet niet raar om dan maar van 95% zekerheid uit te gaan. Eén op twintig keer zet je fout. Dat voelt redelijker dan één op drie.
>
> Voor de liefhebber: als je $1{,}96\sigma$ zou nemen, kwam je op 95,00045% uit.

Heel anders is het bij een digitale schaal. Stel, je meet iets met een digitale schuifmaat met resolutie 0,01 mm. De maat springt van 10,01 naar 10,02, en dan naar 10,03. Als jij 10,02 afleest, dan weet je: de werkelijke waarde ligt ergens tussen 10,015 en 10,025. Je weet niet waar precies, maar we gaan ervan uit dat alle waarden binnen dat interval even waarschijnlijk zijn.

Vaak is dat ook zo, maar niet altijd. Neem een digitale klok die uren en minuten aangeeft. Als je die op 16:42 ziet springen, weet je dat het aantal seconden vlak daarna klein is. Je kunt niets aflezen, maar 'weet' wel iets.

Volgens de GUM is dit een type B-onzekerheid. Er is wel sprake van een kansverdeling, maar die is *niet op een statistische analyse* gebaseerd. Je weet wat de grenzen van de waarden zijn, op basis van de schaal.

De halve resolutie van die uniforme verdeling noem je a (dus bij 0,01 mm resolutie is a = 0,005 mm), en de standaardonzekerheid die daarbij hoort is:

$$u = a/\sqrt{3}$$

Leuke oefening: bewijs dat. Niet eens zo heel moeilijk. De variantie is.

$$\sigma^2 = \int_{-a}^{a} (x-h)^2 f(x)dx$$

De functie $f(x)$ is in dat hele bereik $1/2a$. Die integraal is te doen.

Dat getal heeft niet dezelfde 'betekenis' als bij de toevallige fout. Bij een normale verdeling ligt ongeveer 68,3%van de waarden binnen $\pm u$, maar bij een uniforme verdeling ligt slechts 57,7% van de waarden binnen $\pm u$. De kans dat de werkelijke waarde binnen $\bar{x} \pm u$ ligt is dus kleiner dan je zou denken als je de term 'standaardonzekerheid' al te letterlijk neemt.

Toch gebruiken we ook hier een standaardonzekerheid. Waarom? Omdat we dan de verschillende fouten wiskundig met elkaar kunnen combineren. De totale onzekerheid wordt dan:

$$u_{tot} = \sqrt{{u_1}^2 + {u_2}^2 + \ldots + {u_N}^2}$$

Dat kan alleen als al die termen hetzelfde type onzekerheid uitdrukken: een standaardafwijking die hoort bij een kansverdeling. Vandaar dat we zelfs voor een keihard rechthoekige verdeling zoals bij een digitale schaal een standaardafwijking definiëren, ook al is die wat kunstmatig.

Het is dus belangrijk om te beseffen: bij een uniforme verdeling heeft u een andere betekenis dan bij een normale verdeling. Bij een digitale meting ligt de werkelijke waarde gegarandeerd binnen het interval $\bar{x} \pm a$, maar niet noodzakelijk binnen $\bar{x} \pm u$. Bij een reeks herhaalde metingen ligt de werkelijke waarde met ongeveer 68 procent kans binnen $\bar{x} \pm u$, maar je weet niet of dat ook echt zo is in jouw geval.

Als je de 68-95-99,7 regel al kent (de makkelijke versie van 68,27%, 95,45% en 99,93%), kun je als kers op de taart 58% toevoegen. Of 57,74%.

Vertrouwen

44. Voorbeelden van (on)afhankelijkheid

Voordat het we het over 'gewoon' en kwadratisch optellen gaan hebben in deze sectie, bekijken we eerst twee voorbeelden die duidelijk maken wat we eigenlijk met *afhankelijk* bedoelen.

We willen de oppervlakte van een tafelblad berekenen en gebruiken een rolmaat om de lengte en de breedte te bepalen. De oppervlakte volgt uit de formule

$$A = l \times b$$

Bij een vermenigvuldiging tel je de relatieve fouten bij elkaar op. Dus:

$$\frac{\delta A}{A} = \frac{\delta l}{l} + \frac{\delta b}{b}$$

De modulusstrepen zijn hier weggelaten, omdat we met alleen maar positieve waarden te maken hebben.

Hoe zit het nu met de *afhankelijkheid* tussen l en b? Wat als onze rolmaat systematisch 1% te veel aangeeft? Dat een tafelblad met een lengte van 120,0 cm door ons opgemeten wordt als 121,2 cm? Dat zal dan doorwerken in de oppervlakte. Volgens de formule hierboven is de bijdrage 1%.

Diezelfde rolmaat zal bij het opmeten van de breedte hetzelfde euvel vertonen. Ook daar is het 1% te veel. Het tafelblad van 80,0 cm wordt door ons opgemeten als 80,8 cm.

> Voor de zekerheid meten we de lengte en de breedte nog twee keer. Geloof het of niet: we vinden nóg twee keer 121,2 cm en 80,8 cm. Is dat even toevallig? Nee. De rolmaat heeft een *systematische* fout. Die vind je niet door te herhalen. Hij is niet zichtbaar, en zou gevonden moeten worden door een kalibratie.

De werkelijke oppervlakte van de tafel zou je kunnen uitrekenen als je de werkelijke waarden van de lengte en de breedte wist.

$$A = l \times b = 1{,}200 \times 0{,}800 = 0{,}960\ m^2$$

Reken je met de door ons gemeten waarden, dat vind je net iets anders.

$$A = l \times b = 1{,}212 \times 0{,}808 = 0{,}979296\ m^2$$

Dat is natuurlijk een te groot aantal significante cijfers, maar we moeten de onzekerheid nog uitrekenen. Die was 1% en dat weten we dan toch? Nee! De

systematische fout is in ons zelfverzonnen voorbeeld gelijk aan 1%. In werkelijkheid weet je niet dat een rolmaat exact 1% te veel aangeeft. (Dan zou je ervoor corrigeren. Bovendien maakt geen enkele fabrikant een rolmaat die exact 1% te veel aangeeft. Als hij dat wist, zou hij de streepjes beter zetten.)

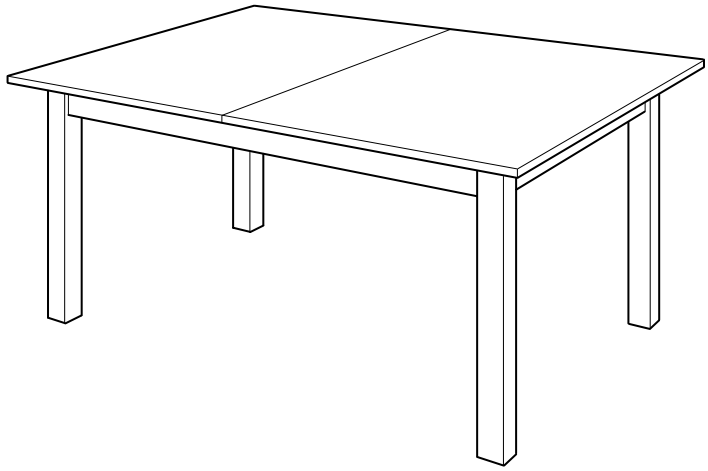
Laten we aannemen dat de fabrikant opgegeven heeft dat de rolmaat een onzekerheid heeft van 2%. Dan is de onzekerheid in de oppervlakte het dubbele daarvan. En 4% van 0,979296 is 0,03917184. De onzekerheid wordt opgegeven in één significant cijfer. Dat resulteert in de volgende uitkomst.

$$A = 0{,}98 \pm 0{,}04\ m^2$$

De uitkomst is niet strijdig met de werkelijke waarde.

Hoe zit het met de afhankelijkheid? Wat nu als we een rolmaat hadden gehad die precies 1% te weinig had aangegeven? Dan waren zowel de lengte als de breedte te klein geweest.

Het moge duidelijk zijn dat hier sprake is van een zeer sterke afhankelijkheid. Als de gemeten lengte te groot is, is de gemeten breedte ook te groot. En als de lengte te klein is, is de gemeten breedte ook te klein. Het meten met hetzelfde instrument geeft een sterke afhankelijkheid.



Figuur 44.1 Zomaar een tafel, om de lege ruimte op de pagina te vullen (afbeelding van IKEA)

Zomaar een tafel? Welnee. Dit is een BJÖRNA. Kenners zien dat zo.

Nu een ander voorbeeld. Stel dat je het rendement van een motor wilt berekenen. De motor tilt een blok met een gegeven massa op tot een bepaalde hoogte, in een bepaalde tijd. Op de motor staat een spanning en erdoorheen

loopt een stroom. Het rendement is de benodigde energie om het hoogteverschil te overbruggen gedeeld door het vermogen van de motor maal de tijd.

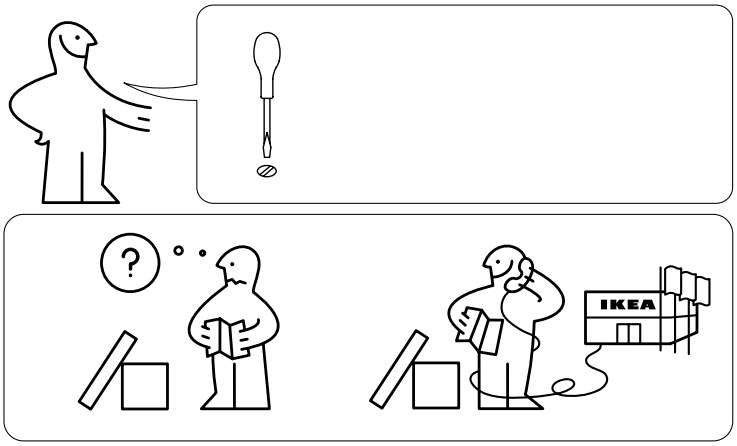
$$e = \frac{geleverde\ arbeid}{verbruikte\ energie} = \frac{mgh}{VIt}$$

Rechts van het isgelijkteken staan maar liefst zes grootheden waarvan er eentje wordt opgezocht in de literatuur (de gravitatieversnelling ter plaatse) en de andere vijf door vijf verschillende instrumenten wordt opgemeten. De relatieve fout in de uitkomst zou je kunnen berekenen door

$$\frac{\delta e}{e} = \frac{\delta m}{m} + \frac{\delta g}{g} + \frac{\delta h}{h} + \frac{\delta V}{V} + \frac{\delta I}{I} + \frac{\delta t}{t}.$$

Wat denk je, zal hier een te grote m ook zorgen voor een te grote g of h? Of zal de fout in je stopwatch doorwerken op de multimeter? Nee. En dus is het domweg optellen van de relatieve onzekerheden veel te pessimistisch.

Van belang is dat je inziet dat die metingen bij dat tafelblad *afhankelijk* waren. De grootheden in het voorbeeld van de motor zijn *onafhankelijk*.



Dat mannetje van de IKEA-handleidingen maakt altijd een wat sukkelige indruk op me. Kijk eens naar die domme grijns als hij uitroept dat hij alleen maar een schroevendraaier nodig heeft. En dan komt hij er nóg niet uit. Nou ja, dan maar bellen naar de zaak waar het vandaan komt. Met een telefoon die kennelijk *in de winkel* staat. Volg het snoer maar en je komt bij de klantenservicebalie uit. De Klåntförvirringsbøreubilly.

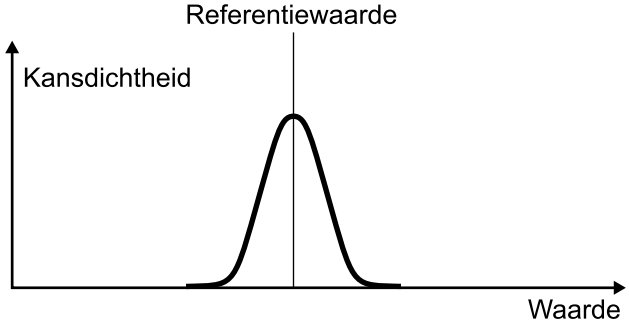
45. Extreme voorbeelden

Stel, we hebben twee (fictieve) multimeters, die op het eerste gezicht sterk op elkaar lijken. Vooral aan de buitenkant. Zoek de verschillen!

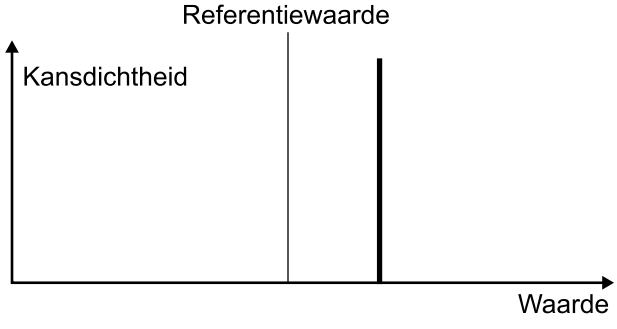
Er is maar één verschil:

- meter A (de bovenste) geeft meetwaarden die extreem nauwkeurig zijn (maar niet erg precies)
- meter P (de onderste) geeft meetwaarden die extreem precies zijn (maar niet erg nauwkeurig)

In het *Wikipedia*-plaatje waarin *accuracy and precision* worden geschetst, zien hun kansverdelingen er als volgt uit.



Figuur 45.1 Metingen van Accurate Digital Multimeter A



Figuur 45.2 Metingen van Precise Digital Multimeter P

Gemiddeld genomen geeft meter A (de nauwkeurige, *accurate*) de juiste waarde, maar er zit spreiding in de uitkomsten. Meter

P, de *precieze* variant geeft bij elke herhaalde meting hetzelfde resultaat. Maar die meting wijkt wel een (onbekend) stuk af van de werkelijke waarde.

Beide apparaten hebben bij het meten van een weerstand een onzekerheid van 10 Ω. Die komt op een verschillende manier tot uiting.

- meter A geeft normaal verdeelde metingen met een standaardafwijking van 10 Ω en een gemiddelde dat gelijk is aan de werkelijke waarde
- meter P geeft telkens dezelfde afwijking, die met een kans van ongeveer twee derde binnen 10 Ω van de werkelijke waarde ligt.

Voor meter P zou de afwijking van de werkelijke waarde 10 Ω kunnen zijn, maar ook 8 Ω of 11 Ω. Of 7 Ω, of 9 Ω, of 12 Ω… Het moet in die orde van grootte liggen, maar je weet het niet exact. Anders zou je ervoor kunnen compenseren. (En een *bekende* correctie is geen *onzekerheid* meer.)

Voor meter A geldt dat je de onzekerheid net zo klein kun krijgen als je maar wilt, door veel metingen uit te voeren. Je bepaalt dan namelijk gewoon het gemiddelde van een groot aantal metingen. We weten dat de standaardafwijking van de weerstanden $\sigma_R = 10\ \Omega$. Het gemiddelde van een (groot) aantal metingen is op zichzelf ook normaal verdeeld, maar wel met een kleinere standaardafwijking, namelijk

$$\sigma_{\bar{R}} = \frac{\sigma_R}{\sqrt{N}}$$

Bij honderd metingen heb je dus in het gemiddelde maar een onzekerheid van 1,0 Ω. Als de onzekerheid 0,10 Ω zou moeten zijn, moet je nóg honderd keer zo vaak meten: tienduizend keer.

Het is niet erg realistisch dat je twee multimeters zou hebben waarvan de ene geen systematische fout heeft en wel een toevallige en de andere juist wel een systematische en geen toevallige. Maar het is wel nuttig om er dit gedachtenexperiment mee te doen.

Stel dat je een (ook weer fictieve) weerstand zou hebben van exact 200 Ω. Je meet die tien keer met Meter P (die een volmaakte precisie heeft) en vindt dan: 207 Ω, 207 Ω, 207 Ω, 207 Ω, 207 Ω, 207 Ω, 207 Ω, 207 Ω, 207 Ω en 207 Ω. Ja, het is best een saaie Piet, die Meter P.

> Tien keer hetzelfde: je zou bijna gaan denken dat er geen onzekerheid op de meting zit. Maar dat is schijn: de fout zit verborgen in de constante afwijking. De onzekerheid van 10 Ω betekent dat de werkelijke waarde waarschijnlijk (met een kans van ongeveer 2 op 3) ligt tussen 197 Ω en 217 Ω.

Nu de andere multimeter, Meter A. Die vindt wél tien verschillende waarden.

188 Ω, 195 Ω, 202 Ω, 226 Ω, 202 Ω, 200 Ω, 186 Ω, 196 Ω, 185 Ω en 195 Ω.

> Meter A is fictief en we willen wel 'echte' normaal verdeelde metingen vinden. Het is lastig om die te bedenken. 'Zomaar een getal onder de 10… Ergens op internet stond dat mensen meestal 'zeven' zeggen.
>
> Hoe komen we nu aan tien willekeurige getallen die normaal verdeeld zijn met een gemiddelde van 200 en een standaardafwijking van 10? Daar hebben we *Excel* voor! Die heeft het bovenstaande rijtje verzonnen.

Nu we de beide reeksen hebben, komen we tot de kern van dit hoofdstuk.

- De metingen van Meter A zijn onderling *volledig onafhankelijk*.
- De metingen van Meter P zijn onderling *volledig afhankelijk*.

Dat heeft gevolgen voor het doorwerken van onzekerheden bij optellen.

Als we twee weerstanden van exact 200 Ω in serie zetten, dan is de vervangingsweerstand gelijk aan 400 Ω.

Meter P zal voor de eerste weerstand zeggen dat die 207 Ω is... *en voor de tweede weerstand ook*! De meter heeft een afwijking die de gebruiker niet exact kent, maar we weten wel dat die constant is en (absoluut) in de orde van grootte van 10 Ω ligt. Maar de afwijking is altijd hetzelfde en versterkt dus zichzelf bij optellen.

Meter A zal voor de eerste weerstand zeggen dat die 188 Ω is en voor de tweede weerstand dat die 195 Ω is. Een eventuele derde is 202 Ω en ga zo het rijtje van willekeurige afwijkingen maar af. Telkens is de volgende weerstand gebaseerd op *toeval*. De waarde is gemiddeld gelijk aan 200 Ω en is de ene keer lager en de andere keer hoger. Het is dus wat pessimistisch om te zeggen dat die spreiding bij twee weerstanden twee keer zo groot is.

Wiskundig kun je bewijzen dat de som van twee normaal verdeelde variabelen ook normaal verdeeld is. Het gemiddelde van die 'opgetelde' verdeling is gelijk aan het gemiddelde van beide gemiddelden. De variantie van de 'opgetelde' verdeling is gelijk aan de som van beide varianties.

In ons voorbeeld is de verwachtingswaarde van de meting voor de eerste en de tweede weerstand gelijk aan 200 Ω. De variantie is voor beide gelijk aan $(10\ \Omega)^2 = 100\ \Omega^2$. De som is dus $200\ \Omega^2$ en de standaardafwijking is daar de wortel van: $10\sqrt{2}\ \Omega \approx 14{,}142135623731\ \Omega$. In twee significante cijfers: 14 Ω.

46. Onafhankelijke fouten optellen

Als je een weegschaal hebt die ergens tussen en 0% en 2% te veel aangeeft, maar je weet niet hoeveel, wat moet je dan?

Over dat 'ergens' weet je niet zo veel. Domweg overnemen wat er op het display staat, zou niet slim zijn. Daar kun je voor corrigeren. Standaard 1% eraf trekken en een onzekerheid aanhouden van 1%, dat is geen slecht idee.

We wegen 406 gram suiker af. Dan zeggen we dat het 1% minder is en geven de bijbehorende onzekerheid op.

$$m_{suiker} = 0{,}402 \pm 0{,}004\ kg$$

Het zou kunnen, dat het inderdaad 406 gram was. Maar 398 gram is ook mogelijk.

Intussen hebben we een nieuwe weegschaal gekocht en de fabrikant is zo slim geweest om die ene procent al te corrigeren. Dat rekensommetje hoeft dus niet meer, maar die onzekerheid zit er nog steeds in. De schaal kan nu te veel of te weinig aangeven: we weten alleen niet hoeveel precies. (De fabrikant ook niet, anders hadden ze daar ook wel voor gecorrigeerd.)

We wegen opnieuw vier ons suiker af, maar ook nog eens twee ons bloem en een pond boter. (Toe maar!) Dat gooien we allemaal bij elkaar… hoeveel hebben we dan?

$$m_{suiker} = 0{,}404 \pm 0{,}004\ kg$$
$$m_{bloem} = 0{,}199 \pm 0{,}002\ kg$$
$$m_{boter} = 0{,}503 \pm 0{,}005\ kg$$

Het vervelende van die weegschaal is: we weten wel dat-ie afwijkt, maar niet precies hoeveel en ook niet of het te veel of te weinig is. (Hadden we nou die oude weegschaal nog maar! Die week meer af, maar je wist tenminste zeker dat je nooit te veel nam.)

Die schaal geeft dus structureel te veel of te weinig aan en we ontkomen er niet aan om de onzekerheden bij elkaar op te tellen. Het goede nieuws is: relatief blijft de onzekerheid gewoon die ene procent.

$$m_{alles} = 1{,}106 \pm 0{,}011\ kg$$

Als we twee keer zo veel hadden genomen, hadden we een significant cijfer moeten inleveren. Eigenlijk is dat niet logisch, maar dat zijn de regels.

$$m_{dubbel} = 2{,}21 \pm 0{,}02\ kg$$

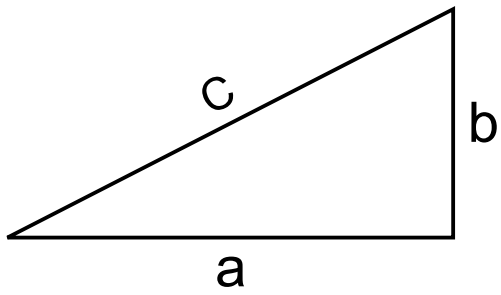
Nu gaan we iets anders doen. We lenen de weegschaal van onze buren. Twee verschillende weegschalen, van hetzelfde merk en met dezelfde specificaties. Die geven beide een procent te veel of te weinig aan, maar het probleem is: het kan best zijn dat de ene altijd te veel aangeeft en de andere altijd te weinig. Of andersom. Of beide te veel, of beide te weinig…

We gaan nu de suiker en de bloem in de ene weegschaal wegen en de boter in de andere. Hoe zit dat nu met het totaal? We moesten eerst de onzekerheden domweg optellen omdat het als maar erger werd. Maar het zou nu kunnen zijn dat de ene te veel aangeeft en de andere te weinig. Het is wel erg pessimistisch om ervan uit te gaan dat je het moet optellen.

Wiskundig gezien kun je bewijzen dat het *kwadratisch optellen* van de onzekerheden een betere maat geeft dan 'gewoon' optellen. Dat heeft iets te maken met Taylorreeksen en fouten die 'haaks' op elkaar staan, maar dat bewijs besparen we je. Hughes & Hase (2010) bewijzen het wel. Dus voor wie wil…

> Wiskundigen noemen 'haaks' anders: 'orthogonaal'. Een rechte hoek dus, van 90° of een half pi. Engelsen noemen een halve taart een *half pie*.
>
> Dat heeft niets met elkaar te maken en dat brengt ons terug bij 'orthogonaal': in dit verband betekent het: niets met elkaar te maken hebben. We zeggen wel dat dingen haaks op elkaar staan als ze tegenstrijdig zijn. Vreemd eigenlijk, want dan gaan ze toch juist recht tegen elkaar in?

Het bewijs hoef je niet te kennen of te begrijpen. Maar het is wel handig om in te zien dat een kwadratische optelling altijd minder oplevert dan een gewone optelling (tenzij een van beide nul is, maar dan tel je niets op).



Alleen als b of c gelijk is aan 0 is a gelijk aan de som van die twee. In alle andere gevallen is de waarde kleiner. Een rechte lijn tussen twee punten is nu eenmaal altijd een kortere weg dan eentje met een haakse bocht erin.

47. Met & zonder afhankelijkheid

De bewegingsvergelijking van een vallende steen luidt als volgt.

$$s = \frac{1}{2} g t^2$$

We kiezen omlaag als positief en het punt van loslaten als oorsprong. De massa van de steen is zo groot dat de wrijvingskracht te verwaarlozen is ten opzichte van de zwaartekracht op de steen.

We willen de versnelling van de zwaartekracht berekenen en herschrijven daarom de formule in onderstaande vorm.

$$g = 2\frac{s}{t^2}$$

Hoe moeten we nu de onzekerheid bepalen? Is er hier sprake van afhankelijke of onafhankelijke onzekerheden? De tijd wordt gemeten met een stopwatch of een camera en de afstand met een rolmaat of laserafstandmeter. Die metingen hebben niet veel met elkaar te maken. Je zou dus kunnen zeggen dat de fouten onafhankelijk zijn. De formule zou dan als volgt luiden.

$$\frac{\delta g}{|g|} = \sqrt{\left(\frac{\delta s}{s}\right)^2 + 2\left(\frac{\delta t}{t}\right)^2}$$

De factor 2 kent geen onzekerheid, dus zie je in de formule niet terug.

Als je pessimistischer bent en met afhankelijke onzekerheden wilt rekenen, kom je niet met een kwadratische optelling, maar een ‘gewone’.

$$\frac{\delta g}{|g|} = \frac{\delta s}{|s|} + 2\frac{\delta t}{|t|}$$

Als de onzekerheid in de afstand 1% is en de onzekerheid in de tijd is 2%, dan zou uit deze formule volgen dat de relatieve onzekerheid in g gelijk is aan 1% + 2 × 2% = 5%. Tel je die termen kwadratisch op, dan kom je op andere waarden.

$$\frac{\delta g}{|g|} = \sqrt{(0{,}01)^2 + 2 \cdot (0{,}02)^2} = \sqrt{0{,}0001 + 2 \cdot 0{,}0004}$$

$$= \sqrt{0{,}0009} = 3\%$$

Welke berekening is correct? Het antwoord is: geen van beide.

De fout in s is onafhankelijk van de fout in t, maar de fout in t is *volledig* afhankelijk van zichzelf.

> Een beetje filosofische kwestie misschien… Kan iets afhankelijk zijn van zichzelf? Als we zeggen dat x afhankelijk is van y, dan bedoelen we dat een verandering in x 'automatisch' een verandering in y betekent.
> Als x en y beide gelijk zijn aan t, dan is dat meeveranderen wel erg sterk aanwezig.

Hoe lossen we dit op? Het beste is om even uit te gaan van een andere grootheid, die gelijk is aan het kwadraat van de tijd.

$$T = t^2$$

Daarvoor geldt:

$$\frac{\delta T}{|T|} = 2\frac{\delta t}{|t|}$$

Als we dit invullen in de oorspronkelijke formule, dan staat er:

$$g = 2\frac{s}{T}$$

Voor de onzekerheid geldt dan

$$\frac{\delta g}{|g|} = \sqrt{\left(\frac{\delta s}{s}\right)^2 + \left(\frac{\delta T}{T}\right)^2} = \sqrt{\left(\frac{\delta s}{s}\right)^2 + \left(2\frac{\delta t}{t}\right)^2} = \sqrt{\left(\frac{\delta s}{s}\right)^2 + 4\left(\frac{\delta t}{t}\right)^2}$$

$$= \sqrt{(0{,}01)^2 + 4(0{,}02)^2} = \sqrt{17}\% \approx 4{,}1\%$$

De waarheid ligt dus in het midden. Nou ja, niet precies in het midden, maar wel tussen die 3% en 5%. Rekenen alsof het helemaal afhankelijk is, zou wat te pessimistisch zijn. Veronderstellen dat het volledig onafhankelijk is, wordt weer te optimistisch. De waarheid ligt ertussenin — een mooie les voor het leven, maar ook voor de meetkunde.

48. (On)afhankelijke normale verdelingen

Dat van die afhankelijkheid van verdelingen zou je zomaar voor waar kunnen aannemen, maar er is wiskundig iets eenvoudigs over te zeggen dat het inzichtelijker maakt.

Hughes & Hase (2010) vermelden op hun formuleblad (binnenzijde van de achterkant van het boek) dat de onzekerheid in de som van twee grootheden gelijk is aan de som van die twee onzekerheden, *kwadratisch opgeteld.*

Taylor (1997) vermeldt op zijn formuleblad (binnenzijde van de voorkant van het boek) hetzelfde, maar zegt daarbij dat dit alleen geldt voor *independent random errors*: onafhankelijke toevallige fouten. Wat altijd geldt, is dat die onzekerheid maximaal gelijk is aan de 'gewone' som van de onzekerheden.

Berendsen (2011) geeft in bijlage A1 met als titel *Combining uncertainties* een wiskundige analyse. De vetgedrukte vraag luidt: *Why do squared uncertainties add up in sums*?

> "Waarom tellen fouten kwadratisch op?" luidde het Nederlandse origineel in zijn boek uit 1997. Toen was het *cursief* en niet **vetgedrukt**, maar dat is slechts vormgeving.
>
> Taalkundig rammelt het beide: de fouten tellen niet *zelf* op: ze *worden* opgeteld. In het Engels doen ze daar niet zo moeilijk over, maar in het Nederlands is dat toch echt vreemd: 'de fouten tellen op".
>
> Er zit nog wat rarigheid in de Engelse kop. Letterlijk staat er namelijk: "Waarom tellen gekwadrateerde onzekerheden op in sommen?" Dat is een rare vraag… In sommen tel je toch altijd op? Ja. Aantallen appels en peren tel je op. Dat geldt ook voor gekwadrateerde onzekerheden.
> De vraag is dus niet "Waarom tel je *gekwadrateerde onzekerheden* op?", maar: "Waarom tel je onzekerheden *kwadratisch* op?"
>
> Opvallend is dat Berendsen het in het oorspronkelijke boek over *fouten* had en zijn Engelse uitgave van veertien jaren later over *onzekerheden.*

Als je twee normaal verdeelde grootheden hebt, geldt voor de verwachtingswaarden en de spreidingen:

$$E[x] = \mu_x \qquad E[(x-\mu_x)^2] = {\sigma_x}^2$$

$$E[y] = \mu_y \qquad E\left[\left(y-\mu_y\right)^2\right] = {\sigma_y}^2$$

Je kunt die twee normaal verdeelde grootheden optellen.

$$z = x + y$$

Voor het gemiddelde van de verdeling van die z geldt dan:

$$\mu_z = \mu_x + \mu_y.$$

Nu de variantie. Die is gedefinieerd als het gemiddelde van het kwadraat van de afwijking van het gemiddelde. Dat is te schrijven als een verwachtingswaarde.

$${\sigma_z}^2 = E[(z - \mu_z)^2]$$

Die z en μ_z kunnen we vervolgens vervangen door de formules hierboven.

$${\sigma_z}^2 = E\left[\left(x + y - (\mu_x + \mu_y)\right)^2\right]$$

Verander nu de volgorde van de termen tussen de ronde haken (en let op de plus- en mintekens).

$${\sigma_z}^2 = E\left[\left((x - \mu_x) + (y - \mu_{y)})\right)^2\right]$$

Nu is er een handige rekenregel voor kwadrateren van een optelling.

$$(A + B)^2 = A^2 + B^2 + 2AB$$

Pas die toe op bovenstaande, en je krijgt de uitkomst.

$${\sigma_z}^2 = E\big[(x - \mu_x)^2 + (y - \mu_y)^2 + 2(x - \mu_x)(y - \mu_y)\big]$$

Deze uitkomst bevat twee ‘delen’: de eerste twee termen zijn de som van de beide varianties, de laatste term is het dubbele van de covariantie van beide grootheden. (En dan vermenigvuldigd met een factor 2.)

$${\sigma_z}^2 = {\sigma_x}^2 + {\sigma_y}^2 + 2cov(x, y)$$

De variantie van de som van x en y is dus afhankelijk van twee termen die we kennen als de normale verdelingen beschreven zijn. Een normale verdeling ligt immers vast als je de μ en σ kent.

Hoe groot kan die derde term zijn, de covariantie?

$$cov(x, y) = E\big[(x - \mu_x)(y - \mu_y)\big]$$

Wat is de verwachtingswaarde van het product van de afwijkingen van het gemiddelde van beide grootheden?

Als de waarden van x en y op zuiver toeval berusten, dan zijn hun afwijkingen van het gemiddelde volstrekt ongecorreleerd. Gemiddeld zijn de afwijkingen gelijk aan nul, soms zijn ze positief en soms zijn ze negatief. Soms zijn ze groot en soms zijn ze klein. Er zit geen relatie tussen en gemiddeld is hun product gelijk aan nul.

Maar wat nu als de waarden van x en y van elkaar afhankelijk zijn? De grootst denkbare afhankelijkheid is dat ze altijd gelijk zijn aan elkaar. In dat geval zijn ook hun gemiddelden en standaardafwijkingen aan elkaar gelijk en mag je voor y ook x invullen en μ_x voor μ_y.

$$cov(x,y) = E[(x-\mu_x)(x-\mu_x)] = E[(x-\mu_x)^2] = {\sigma_x}^2$$

Dan geldt dus dat

$${\sigma_z}^2 = {\sigma_x}^2 + {\sigma_y}^2 + 2cov(x,y) = {\sigma_x}^2 + {\sigma_x}^2 + 2{\sigma_x}^2 = 4{\sigma_x}^2$$

Wat overblijft, is heel eenvoudig: $\sigma_z = \sqrt{4{\sigma_x}^2} = 2\sigma_x = \sigma_x + \sigma_y$.

We hebben nu bewezen dat de variantie van z

- minimaal gelijk is aan de kwadratische optelling en
- maximaal gelijk aan de gewone optelling

van de varianties van x en y.

Bij volstrekte onafhankelijkheid geldt het eerste.

Bij volstrekte afhankelijkheid geldt het tweede.

Bij gedeeltelijke afhankelijkheid zit het ertussenin.

49. Het criterium van Chauvenet

Soms heb je een meting die nogal ver afwijkt van de rest. Dat voedt het vermoeden dat die meting niet goed is, maar hoe weet je dat?

William Chauvenet (24 mei 1820 – 13 december 1870) was een professor in de wiskunde, astronomie, navigatie en landmeetkunde. Hij bedacht een criterium op grond waarvan je — op basis van statistiek — kon besluiten of je een waarde zou mogen verwerpen.

Simpel gezegd komt zijn methode hierop neer.

1. Schat het gemiddelde en de standaardafwijking van de normale verdeling die ten grondslag ligt aan de metingen
2. Bepaal de kans dat de 'verdachte waarde' optreedt.
3. Bereken de verwachtingswaarde van hoe vaak zoiets gebeurt bij het aantal metingen dat je hebt.
4. Verwerp de meting als die verwachtingswaarde kleiner is dan een half. Een half zit immers dichter bij nul dan bij één.

Hoe werkt het criterium van Chauvenet?

Je veronderstelt dat de metingen (inclusief de verdachte waarde!) afkomstig zijn van een normale verdeling. Wat je vervolgens wilt weten, is hoe die ene uitschieter (of een grotere) voorkomt in dit aantal metingen. Beter gezegd: wat de verwachtingswaarde is van het aantal keren dat die waarde voorkomt. Als dat minder is dan een half (0,5), dan verwerp je de meting. De verwachtingswaarde zit dan namelijk dichter bij 0 dan bij 1.

> Het is belangrijk dat je bij de berekening uitgaat van de juiste formule voor de standaardafwijking, namelijk die voor een *steekproef*, met de N – 1 in de noemer.

Taylor (1997) gebruikt in zijn boek het voorbeeld van iemand die slingertijden meet. De gevonden periodetijden (in seconden) zijn als volgt:

3,8, 3,5, 3,9, 3,9, 3,4, 1,8.

Afgerond op één decimaal vind je: $\overline{x} = 3{,}4\ s$ en $s_x = 0{,}8\ s$.

De verdachte waarde van 1,8 s wijkt precies twee keer de geschatte standaardafwijking af van het geschatte gemiddelde. Vermoedelijk heeft Taylor die waarde opzettelijk zo gekozen, want het verschil tussen 3,4 en 1,8 is precies twee keer die 0,8. En de kans dat een waarde minstens twee keer de

standaardafwijking afwijkt van het gemiddelde is bij een normale verdeling ongeveer 5%. Je kunt dus uit je hoofd uitrekenen dat de verwachtingswaarde van het aantal metingen met een dergelijke afwijking gelijk is aan 0,05 × 6 = 0,3. En dat is minder dan 0,5.

Het is een beetje vreemd om statistiek te bedrijven met zulke kleine aantallen. Zes metingen is niet echt veel. En bij de bepaling van het gemiddelde en de standaardafwijking neem je ook nog eens de verdachte waarde mee! In het toetsen van hypotheses is dat gebruikelijk: je bewijst dat iets onwaarschijnlijk is door eerst aan te nemen dat het waar is en daarna te bepalen dat de kans dit die aanname klopt klein is. Als je de verdachte waarde niet mee zou nemen, dan werk je eigenlijk al naar het antwoord toe: je stopt namelijk je eigen vooroordeel erin dat er verworpen gaat worden.

Wat nu als er meer metingen geweest waren? Je verwerpt een verdachte meting als de verwachtingswaarde kleiner is dan 0,5, maar als je meer metingen hebt, wordt de verwachtingswaarde toch ook groter?

In plaats van lang na te denken, maken we nu gewoon een sommetje.
We laten de vijf 'gewone' metingen allemaal twee keer voorkomen en die ene verdachte meting één keer, door die eerste vijf waarden achter de reeks te plakken. De metingenserie is dan (met weglating van de eenheden):

3,8, 3,5, 3,9, 3,9, 3,4, 1,8, 3,8, 3,5, 3,9, 3,9, 3,4.

Dat resulteert in een ander gemiddelde en een andere standaardafwijking: 3,5 en 0,6. De kans op een afwijking van 1,9 of hoger is lastiger uit te rekenen, omdat het niet meer exact een factor 2 maal de standaardafwijking is: 1,6/0,8 was precies 2, maar 1,7/0,6 ≈ 2,83.Deze overschrijdingskans is al een stuk kleiner, namelijk 0,73 %. De verwachtingswaarde is 0,08: veel minder dan 0,5.

> Als je niet weet hoe je een overschrijdingskans bij een normaal verdeel proces berekent: in een volgend hoofdstuk wordt dat uitgelegd.

We gaan nu de tussenstappen in een tabel zetten, met veel cijfers achter de komma. Dat doen we voor de zes metingen (vijf plus één verdachte) én voor een set van elf metingen (twee keer die vijf plus één verdachte) en 21 metingen (vier keer die vijf plus één verdachte). We voegen bovendien een extra kolom toe, namelijk eentje waar 101 metingen gedaan zijn: twintig keer het setje van die vijf 'goede' metingen en die ene verdachte ertussendoor.

Het klinkt aannemelijk dat je bij een groot aantal waarnemingen die ene verdachte meting eerder durft te verwerpen dan bij een klein aantal, omdat die

uitschieter zelf meedoet in de bepaling van het gemiddelde en de standaardafwijking. Die worden bij kleine aantallen sterker 'gekleurd' door de uitschieter. Bij honderd metingen vormt die ene vreemde eend in de bijt maar één procent van het geheel. Dat zie je nauwelijks meer terug in $\bar{x}$ en s_x.

In de tabel staat bij alle variabelen ook hun betekenis weergegeven. Het voorbeeld helpt dus ook als je beter wilt snappen hoe het werkt.

$\bar{x}$	gemiddelde van de steekproef, inclusief verdachte waarde
s_x	standaardafwijking van de steekproef (inclusief…)
x_v	verdachte waarde
$\lvert\bar{x} - x_v\rvert$	[1]afwijking (verschil gemiddelde en verdachte waarde)
t_v	[1]afwijking gedeeld door de standaardafwijking
P_v	kans op een overschrijding in de normale verdeling
N	aantal metingen
n	[2]verwachtingswaarde van aantal keer zo'n uitschieter

[1] *Normering maakt er een standaardnormale verdeling van.*
[2] *Zo'n uitschieter: net zo groot of groter, beide kanten op.*

	6 metingen	11 metingen	21 metingen	101 metingen
$\bar{x}$	3,3833	3,5273	3,6095	3,6812
s_x	0,8035	0,6101	0,4647	0,2824
x_v	1,8	1,8	1,8	1,8
$\lvert\bar{x} - x_v\rvert$	1,5833	1,7273	1,8095	1,8812
t_v	1,9705	2,8313	3,8943	6,6617
P_v	0,0488	0,0046	$9{,}84 \times 10^{-05}$	$2{,}71 \times 10^{-11}$
N	6	11	21	101
n	0,2927	0,051	0,0021	$2{,}73 \times 10^{-9}$

De eenheden hebben we weggelaten. De bovenste vier zijn in seconden, de onderste vier zijn dimensieloos.

In de laatste tabel wordt zichtbaar wat het belang van veel metingen is als je statistiek bedrijft. De verwachtingswaarde op grond waarvan er besloten wordt, neemt af van 0,3 naar 0,05 als je van zes naar elf metingen gaat, en naar 0,002 bij 21 metingen. De kans op zo'n afwijkende waarde bij 101 metingen is extreem klein. Die 1,8 kan in principe nog steeds géén uitschieter zijn, maar die kans is drie op miljard. Dat mag je verwaarlozen.

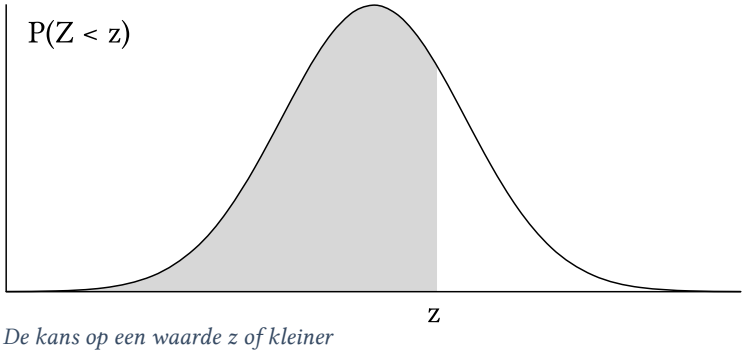
50. Overschrijdingskansen

Je weet vast wel wat een gemiddelde en een standaardafwijking is, en snapt ongetwijfeld waarom je niet de μ en de σ, maar de $\bar{x}$ en de s gebruikt. Ook een verdachte meting is niet zo moeilijk: dat er eentje wat afwijkt, dat zie je zo. Maar dat die ene waarde zoveel afwijkt dat hij volgens Chauvenet verworpen moet worden, daar zul je eerst nog aan moeten rekenen.

Het gaat om de kans dat minimaal die waarde voorkomt, als je veronderstelt dat er sprake is van een normaal verdeeld proces met een gemiddelde van μ en een standaardafwijking σ.

> Ai... Die zouden we niet gebruiken toch? Nee. We zouden het best willen, maar we *kennen* de μ en de σ niet. Daarom gebruiken we *bij gebrek aan beter* de $\bar{x}$ en de s: het gemiddelde en de standaardafwijking uit de steekproef. Je moet toch wát...
>
> Hoe groot is de kans dat een getal z de uitkomst is van een normaal verdeeld proces met een bepaalde μ en de σ? Die kans is nul. Een normaal verdeeld proces heeft namelijk oneindig veel mogelijk uitkomsten. De kans op elk van die uitkomsten is 0. Maar we kunnen het wel hebben over de kans dat een waarde wordt gevonden in een bepaald bereik, of groter of kleiner is dan z.

Eigenlijk klopt de formulering niet. Het gaat niet om de kans dat die ene uitschieter voorkomt, maar dat een uitschieter minstens zo veel afwijkt van het gemiddelde. Die kun je bepalen door de overschrijdingskans te berekenen.



Figuur 50.1 De kans op een waarde z of kleiner

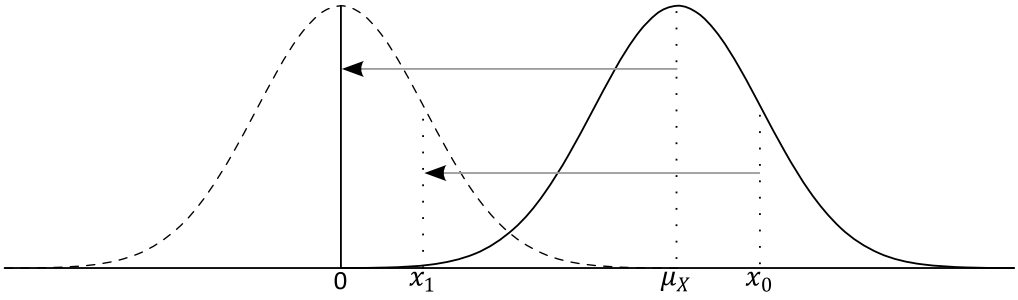
Ter illustratie kijken we naar een normaal verdeelde grafiek, met daarin een waarde z. De totale oppervlakte onder de kromme is gelijk aan 1: de som van alle kansen op alle mogelijke waarden. De kans op 'iets' is nu eenmaal 100%. De oppervlakte van het grijze gebied is de kans op een waarde kleiner dan z.

De oppervlakte van het witte gebied is de kans dat er een waarde optreedt die groter is dan z.

Hoe vinden we de oppervlakte van het grijze gebied in Figuur 50.1? De klassieke manier is: aflezen uit een tabel die al die kansen weergeeft voor een standaardnormale verdeling.

Maar wacht even... De waarde z komt toch niet uit een *standaardnormale* verdeling? Die heeft een μ en een σ die makkelijk zijn: $\mu = 0$ en $\sigma = 1$. Met twee slimme trucjes kunnen we elke normale verdeling omtoveren in een standaardnormale verdeling. Je trekt eerst het gemiddelde μ_x van jouw 'eigen' verdeling af van alle waarden, en je deelt die vervolgens door 'jouw' σ_x.

Die twee stappen kun je in een plaatje weergeven. Door van alle waarden μ_X af te trekken, verschuift het gemiddelde naar 0 en krijgen we een normale verdeling die symmetrisch is rondom 0.

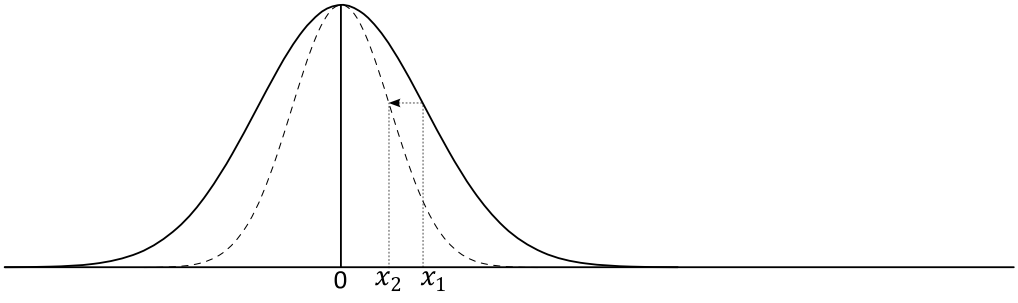


Figuur 50.2 'Opschuiven' van de verdeling zodat het gemiddelde 0 wordt

Wat gebeurt er dan met een willekeurige waarde x_0? Die wordt met μ_x verlaagd. In formulevorm:

$$x_1 = x_0 - \mu_X$$

Door alle waarden door σ_X te delen, wordt de grafiek (in dit geval) smaller en wordt de standaardafwijking gelijk aan 1.



Figuur 50.3 'Samendrukken' van de verdeling om een standaardafwijking van 1 te krijgen

Wat gebeurt er dan met onze waarde x_1? Ook die wordt gedeeld door σ_X. In formulevorm:

$$x_2 = \frac{x_1}{\sigma_X} = \frac{x_0 - \mu_X}{\sigma_X}$$

> Wiskunde is leuk, maar af en toe word je dol van al die indexjes, tekentjes & streepjes en kronkels & bocheltjes. Waarom staat hier nou soms een hoofdletter X en soms een kleine letter x? Antwoord: omdat die twee niet hetzelfde zijn. De hoofdletter X duidt een *kansvariabele* aan. De kleine letter x is één bepaalde waarde, en dus geen kansvariabele meer. Hij is gewoon 1,2 of 1,8 of −0,4 of welk ander getal dan ook.

Dankzij dat opschuiven en delen is de kansdichtheidsfunctie genormeerd tot een *standaardnormale verdeling*. En voor die bijzondere normale verdeling met $\mu = 0$ en $\sigma = 1$ vind je tabellen in wiskundeboeken.

Nu terug naar het voorbeeld. Zoals gezegd: we kennen het echte gemiddelde en de echte standaardafwijking niet, en gebruiken daarom de *schattingen*.

$$x_{norm} = \left|\frac{x_{ver} - \overline{x}}{s_x}\right| \approx \left|\frac{1{,}8 - 3{,}3833}{0{,}8035}\right| \approx 1{,}9705$$

Deze genormeerde x_{norm} is precies wat we bij Chauvenet aanduiden als *het aantal keren de standaardafwijking dat de verdachte waarde afwijkt van het gemiddelde*. En we willen weten hoe groot de kans is dat zo'n afwijking zoveel keer (of meer) voorkomt.

Hoe lees je dat af uit die tabel? We hoeven niet alle rijen te bekijken, alleen de waarden die rond de 2 zitten.

In de tabel vind je de kansen in vier decimalen (vier significante cijfers) voor getallen met twee decimalen (maximaal drie significante cijfers). Daarom moeten we 1,9705 afronden op 1,97. Die 1,9 vind je in de linker kolom. De tweede decimaal, die 7 dus, vind je in de bovenste rij. We zijn op zoek naar het getal dat in de rij staat die met 1,9 begint en de kolom waar 7 boven staat. Dat getal is 0,9756.

Wat is de betekenis daarvan ook alweer? Het is de kans dat één waarde van een standaardnormaal verdeeld proces kleiner is dan 1,97. Die kans is dus 97,56%. De kans dat die waarde groter is, is dus 2,44%. Maar let op: we zijn geïnteresseerd in uitschieters *naar beide kanten*. We moeten dus ook de gevallen meenemen waar die ene waarde kleiner is dan −1,97. De gevonden kans moet daarom nog met 2 worden vermenigvuldigd. Maar dan zijn we er!

$$P(X > 1{,}97) = 2 \times (1 - 0{,}9756) = 2 \times 0{,}0244 = 0{,}0488$$

De *Excel*-functie `NORM.VERD(x;gem;σ;WAAR)` berekent de kans dat een normaal verdeelde waarde kleiner is dan `x`, bij een verdeling met:

- `x`: de waarde waarom het gaat,
- `gem`: het gemiddelde,
- `σ`: de standaardafwijking,
- `WAAR`: omdat je de *opgetelde kans* wilt, niet de kansdichtheid

Stel: je wilt weten hoeveel procent van de Nederlandse mannen langer is dan 1,71 meter. Dan moet je in Excel juist berekenen hoeveel procent **korter** is dan 1,71 meter. Dat is wat `NORM.VERD` doet. De kans op 'langer dan' volgt dan als 1 − die uitkomst.

Voor deze berekening neem je aan dat:

- het gemiddelde gem gelijk is aan 183 cm,
- de standaardafwijking σ ongeveer 10,5 cm is,
- je werkt in centimeters (want dat sluit aan bij de gegevens).

De formule wordt dan:

```
=NORM.VERD(171;183;10,5;WAAR)
```

Dat levert ongeveer 0,127 op: 12,7% is korter dan 1,71 meter. De kans dat iemand langer is dan jij, is dan 1 − 0,127 = 0,873, oftewel 87,3%. Dat zou best kunnen: 1 op de 8 mannen kleiner dan jij (als jij die man van die lengte bent) betekent ook dat 7 op de 8 langer zijn.

Let op dit soort vergissingen:

- Als je de eerste twee getallen verwisselt, krijg je een ander verhaal met wél een uitkomst.
- Als je `ONWAAR` invult in plaats van `WAAR`, krijg je de kansdichtheid (de hoogte van de kromme bij dat punt), geen kans. Zo'n getal lijkt op een percentage, maar betekent in dit verband niets.

Denk ook goed na over eenheden: *Excel* ziet geen verschil tussen meters en centimeters. Je moet dus zorgen dat je gegevens allemaal in dezelfde eenheid staan, en dat die eenheid logisch is. Hier: centimeters.

En tenslotte: als de uitkomst niet klopt met je gezonde verstand, zit er waarschijnlijk een fout in. Maar omgekeerd: als het wel *zou kunnen kloppen*, is dat nog geen bewijs. Controleer je formule. Denk na. Altijd.

51. Chauvenet in Excel

Moet het nou echt zo moeilijk? Nee. Vroeger moest het zo, toen we nog geen computers hadden. Tegenwoordig doe je het hooguit om meer inzicht in de lesstof te krijgen. Hoe pak je dit aan in *Excel*?

```
=2*NORM.VERD(1,8;3,8333;0,7661;WAAR)
```

En klaar ben je! Volgens *Excel* is de uitkomst 0,04878. In vijf decimalen (vier significante cijfers) net niet hetzelfde als wij 'met de hand' gevonden hadden. In drie significante cijfers is het exact hetzelfde.

In plaats van het invullen van die harde getallen, kun je in *Excel* ook met cellen werken. Dan komt de formule er zo uit te zien.

```
=2*NORM.VERD(C8;C1;C2;WAAR)
```

Die C8, C1 en C2 zijn de cellen in mijn werkblad waar de formules staan voor de verdachte waarde, het gemiddelde en de standaardafwijking. WAAR betekent dat we cumulatief werken. Nu vinden we 0,048785389886591.

Is het nou niet een beetje overdreven om *cellen* op te nemen in de formule in plaats van *getallen*? En een antwoord te geven met alle decimalen waarmee *Excel* rekent? Om met dat laatste te beginnen: ja, dat is overdreven. Maar het rekenen met cellen in plaats van getallen verdient wel echt de voorkeur. Niet alleen omdat je dan minder snel fouten maakt, maar vooral omdat je wijzigingen in de dataset meteen terugziet.

In onderstaande tabel staat de inhoud van een *Excel*-werkblad met dertien rijen (1 t/m 13) en vier kolommen (A t/m D). De kolommen bevatten:

A. De gegevens waarop de analyse wordt uitgevoerd.
B. Een omschrijving van de berekening in die kolom.
C. De uitkomst van die berekening. Er worden nul, een of twee decimalen getoond, afhankelijk van de waarde die er staat.
D. De uitkomst van die berekening, maar dan in alle decimalen die *Excel* standaard toont.

In kolom D staan nog steeds niet alle decimalen die *Excel* heeft. De uitkomst van de verwachtingswaarde is 0,292712339319546. *Excel* rekent namelijk met vijftien decimalen.

In het kader rechts staan de gebruikte formules. Het gemiddelde (1), de standaardafwijking van de steekproef (2), de maximale waarde (3) en de minimale waarde (4) worden allemaal gehaald uit kolom A. Dat er tussen de haken A:A staat in plaats van A1:A6 zorgt ervoor dat je gegevens kunt toevoegen of

verwijderen zonder dat je deze formules hoeft aan te passen. Let op dat je STDEV.S neemt en niet STDEV.P. (Geen *populatie* maar *steekproef*.)

Cellen C5, C6 en C7 zijn eenvoudige berekeningen op C3 en C4. C8 is een bijzondere functie die lijkt op de vraagtekenoperator in de taal C. Je test of C5 groter is dan C6. Als dat waar is, is de uitkomst gelijk aan C5. Als dat onwaar is, is het gelijk aan C6. Zo vind je de grootste waarde van de twee.

Met AANTALARG wordt geteld hoeveel metingen er in kolom A staan.

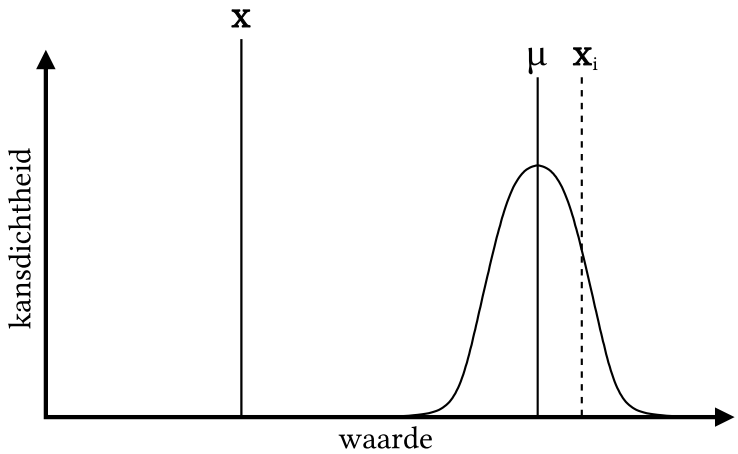
	A	B	C	D	
1	3,8	gemiddelde	3,38	3,383333333	=GEMIDDELDE(A:A)
2	3,5	standaardafwijking	0,80	0,803533862	=STDEV.S(A:A)
3	3,9	hoogste	3,9	3,9	=MAX(A:A)
4	3,9	laagste	1,8	1,8	=MIN(A:A)
5	3,4	afwijking 1	0,52	0,516666667	=C3-C1
6	1,8	afwijking 2	1,58	1,583333333	=C1-C4
7		grootste afwijking	1,58	1,583333333	=MAX(C5:C6)
8		verdachte waarde	1,8	1,8	=ALS(C5>C6;C3;C4)
9		aantal metingen	6	6	=AANTALARG(A:A)
10		aantal keer standaardafwijking	1,97	1,97046249	=C7/C2
11		kans kleiner / groter	0,0241	0,024392695	=NORM.VERD(C8;C1;C2;WAAR)
12		kans grotere afwijking	4,8%	0,04878539	=2*C11
13		verwachtingswaarde	0,29	0,292712339	=C9*C12

52. Kansverdeling resultaat & meting

Kansvariabelen zijn lastige dingen. Gemiddelden ook trouwens. Bedoel je het ‘echte’ gemiddelde van de verdeling, of dat ene ‘toevallige’ gemiddelde van die ene steekproef?

Dan heb je ook nog *standaardafwijkingen*. De kansverdeling heeft een standaardafwijking, die je de ‘echte’ zou kunnen noemen. De steekproef heeft er ook eentje. En net zoals je per steekproef een ander gemiddelde vindt, vind je per steekproef ook een andere standaardafwijking.

Het woord ‘echt’ kan verwarring geven. Die kansverdeling is niet iets wat werkelijk ‘bestaat’. In die zin is het gemiddelde van de kansverdeling helemaal niet *echt*, maar *theoretisch* en verzonnen. Het is een *model* van de werkelijkheid. In de werkelijkheid hebben we wel concrete metingen, maar geen abstracte kansverdelingen. Die heb je alleen in de wiskunde en de statistiek.



Figuur 52.1 Kansverdeling van één meting

De waarde x in Figuur 52.1 is de werkelijke waarde (*true value*). Bestaat de werkelijke waarde ook echt? Ja, maar we kennen hem meestal niet. (Anders hoefden we niet te meten.) Wat we soms wel kennen, is de geaccepteerde waarde. Die ligt meestal dicht bij x, maar is zelden precies gelijk. De werkelijke waarde x (die we niet kennen) heeft geen onzekerheid.

De geaccepteerde waarde (die niet in de grafiek is getekend) heeft wel een onzekerheid. Die zal meestal veel kleiner zijn dan die van de meting. Maar we kennen x dus niet. Wat we wel kennen, is de uitkomst van één bepaalde meting x_i. Die ene x_i beschouwen we als de toevallige uitkomst van een kansvariabele. Hij had ook een andere waarde kunnen hebben.

De kromme in Figuur 52.1 is de kansdichtheidsfunctie van de meting. Dat is een niet 'echt bestaande', maar theoretische wiskundige beschrijving (modellering) van de werkelijkheid van het meten.

> In Figuur 52.1 is die verdeling normaal. Dat hoeft niet zo te zijn. Uniform kan ook. Of één bepaalde waarde, als er geen toevallige fout is. Dan is de uitkomst van de meting altijd hetzelfde, al herhaal je nog zo vaak. Maar nog steeds is de meting x_i dan niet de echte x. Het is een veelgemaakte denkfout: "Als er geen spreiding is, dan is er geen onzekerheid…" Die is er wél dus! Je kunt die alleen niet schatten door te herhalen.

Als je één enkele meting doet en je hebt geen aanvullende gegevens (uit specificaties of kalibraties), dan heb je niets anders dan één getal en een eenheid.

- We weten niet hoe ver x_i van μ vandaan ligt.
- We weten niet hoe ver μ van x vandaan ligt.
- We weten niet hoe groot σ is.

We hebben alleen een meting x_i. Die heeft een waarde (die we bepaald hebben) en een onzekerheid (waar we geen flauw idee van hebben.)

Nu gaan we de stap maken van een *meting* naar een *meetresultaat*. We doen nog negen metingen en hebben dan tien uitkomsten. Die middelen we en we bepalen de standaardafwijking van de steekproef. Het gemiddelde vind je door alles op te tellen en te delen door (in dit geval) tien. De standaardafwijking van de steekproef vind je met de formule die we daarvoor kennen.

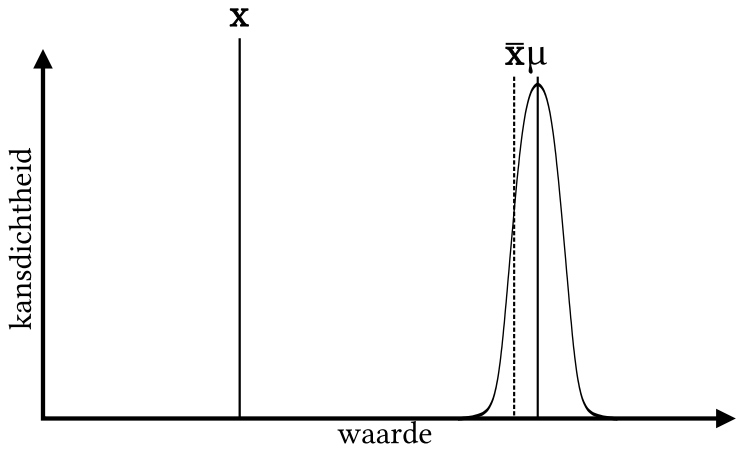
$$s = \sqrt{\frac{1}{N-1}\sum (x_i - \bar{x})^2}$$

In deze formule staan geen μ en σ. Die kennen we namelijk niet. Er staan wel een s (die we berekenen) en een $\bar{x}$ (die we berekend hebben).

> Alle symbolen nog even op een rij.
>
> μ: het gemiddelde van de (theoretische) verdeling
> σ: de standaardafwijking van de (theoretische) verdeling
> $\bar{x}$: het gemiddelde van de (praktische) steekproef
> s: de standaardafwijking van de (praktische) steekproef

Weten we nu meer? Zijn we iets opgeschoten door de meting tien keer uit te voeren? Jazeker! Een belangrijke winst is dat we met s een mooie schatting in handen hebben van σ, en dus van de grootte van de toevallige fout. Een andere belangrijke winst is dat $\bar{x}$ naar verwachting dichter bij de echte μ zal liggen dan die ene x_i die we hadden.

Dat brengt ons op Figuur 52.2. De normale verdeling hier is niet de kansverdeling van een meting, maar van een meetresultaat van tien metingen.



Figuur 52.2 Kansdichtheidsfunctie van een meetresultaat (gemiddelde van een aantal metingen)

Het gemiddelde van (in dit geval) tien normaal verdeelde metingen is zelf ook normaal verdeeld. De standaardafwijking ervan is een stuk kleiner. Die is te berekenen met de formule die we eerder hebben gezien.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

Kennen we nu de standaardafwijking van de kansverdeling? Nee. Kennen we het gemiddelde? Ook niet. Hoe goed kennen we dan het gemiddelde van de verdeling? Het is de uitkomst van een normaal verdeeld proces met μ en σ. Die kennen we niet. We kennen wel s en $\bar{x}$. Dat zijn de beste schattingen die we hebben, en als het aantal metingen groot genoeg is, zijn die niet eens zo slecht. De onzekerheid op het gemiddelde schatten we dus gelijk aan s.

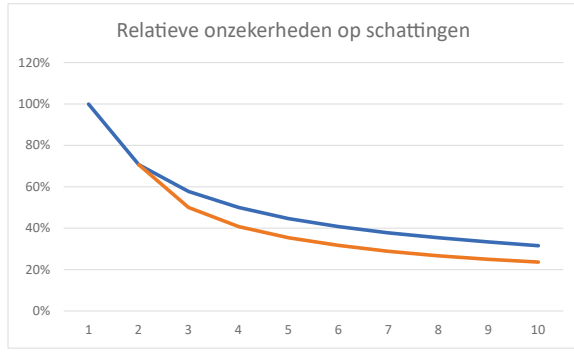
Hoe goed kennen we nu de standaardafwijking van de verdeling? Volgens de wiskunde volgt de relatieve onzekerheid hierin uit onderstaande formule.

$$\frac{1}{\sqrt{2N-2}}$$

Als N dus ‘aardig groot’ is, wordt deze term klein. Dan kunnen we de onzekerheid met een kleine onzekerheid vaststellen.

Om een indruk te krijgen van wat die formules in de praktijk voorstellen, zetten we wat waarden in een tabel en in een grafiek. Merk op dat in de tweede formule de waarde voor $N = 1$ niet gedefinieerd is.

	$\frac{1}{\sqrt{N}}$	$\frac{1}{\sqrt{2N-2}}$
1	100%	
2	71%	71%
3	58%	50%
4	50%	41%
5	45%	35%
6	41%	32%
7	38%	29%
8	35%	27%
9	33%	25%
10	32%	24%



Om het gemiddelde en standaardafwijking goed te schatten, is het doen van tien metingen nog wat weinig. Je zit er dan zomaar een derde of een kwart naast (wat soms overigens goed genoeg is). Voor een serieuzere statistische analyse moet je eerder denken aan honderd metingen.

Oefeningen

1. Hoeveel metingen moet je doen om het gemiddelde met een onzekerheid van maximaal 5% te kunnen schatten?
2. Hoeveel metingen moet je doen om de standaardafwijking met een onzekerheid van maximaal 10% te kunnen schatten?
3. Welke waarde ligt (in het algemeen) dichter bij het gemiddelde van de verdeling: x_i of $\bar{x}$?
4. Vanaf hoeveel metingen is de onzekerheid in de standaardafwijking kleiner dan de onzekerheid in het gemiddelde?

   Bij de laatste opgave is een lange uitwerking beschikbaar in de bijlage met de Antwoorden. Misschien helpt die uitleg je verder om de stof in dit hoofdstuk beter te begrijpen.

53. Overschrijdingskans & significantie

Als een verdeling volmaakt normaal is, zou je ook 3, 4, 5 of 6 keer de standaardafwijking kunnen afwijken van het gemiddelde. Laten we zeggen dat basketballers gemiddeld 2,00 meter lang zijn, met een standaardafwijking van 5 cm. Is er dan ook een kans dat iemand 4,05 meter lang is? Of een negatieve lengte heeft: −5 cm… Het antwoord is: ja. Er bestaat een kans dat je 41 keer 5 cm afwijkt van het gemiddelde. Alleen is die extreem klein.

Je kunt dat afdrukken in een tabel. In de eerste kolom zetten we een lengte in meters. In de tweede kolom het aantal keren de standaardafwijking boven het gemiddelde. In de derde, vierde en vijfde kolom noteren we de kans dat iemand minstens die lengte heeft. In elke van die drie kolommen staat hetzelfde getal (de fractie), maar telkens in een andere notatie.

l [cm]	N	P	P	P
200	0	0,50000000000000000000	50,0%	$5{,}0 \times 10^{-1}$
202,5	0,5	0,30853753872598800000	30,9%	$3{,}1 \times 10^{-1}$
205	1	0,15865525393145800000	15,9%	$1{,}6 \times 10^{-1}$
207,5	1,5	0,06680720126885750000	6,68%	$6{,}7 \times 10^{-2}$
210	2	0,02275013194817910000	2,23%	$2{,}3 \times 10^{-2}$
212,5	2,5	0,00620966532577616000	0,621%	$6{,}2 \times 10^{-3}$
215	3	0,00134989803163010000	0,135%	$1{,}3 \times 10^{-3}$
217,5	3,5	0,00023262907903554000	0,023%	$2{,}3 \times 10^{-4}$
220	4	0,00003167124183312000	0,003%	$3{,}2 \times 10^{-5}$
222,5	4,5	0,00000339767312473871	0,0%	$3{,}4 \times 10^{-6}$
225	5	0,00000028665157192354	0,0%	$2{,}9 \times 10^{-7}$
227,5	5,5	0,00000001898956247803	0,0%	$1{,}9 \times 10^{-8}$
230	6	0,00000000098658770042	0,0%	$9{,}9 \times 10^{-10}$
232,5	6,5	0,00000000004015998645	0,0%	$4{,}0 \times 10^{-11}$
235	7	0,00000000000127986510	0,0%	$1{,}3 \times 10^{-12}$
237,5	7,5	0,00000000000003186340	0,0%	$3{,}2 \times 10^{-14}$
240	8	0,00000000000000000000	0,0%	$0{,}0 \times 10^{0}$

> Huh? Die laatste uitkomst van $0{,}0 \times 10^{0}$ is toch zeker een tikfout in het boek? Nee. Dan hadden we die fout er namelijk wel uit gehaald in plaats van een grijze alinea eraan te besteden. Het is de uitkomst die *Excel* berekent. Kennelijk lukt het niet meer met zulke kleine getallen.

Deze tabel is een goede oefening in het kiezen van notaties. Wat is een handige weergave? Wat geeft snel overzicht en hoeveel cijfers gebruik je?

Percentages zijn handig om een snelle indruk te geven. Dat gaat mis als de getallen heel klein worden is. Voor wetenschappelijke notatie geldt het omgekeerde. Een fractie $5{,}0 \times 10^{-1}$ is wel erg cryptisch voor 50% of een kans van ½. En $6{,}2 \times 10^{-3}$ leest ook lastiger dan 0,62% of 0,6% of 6‰. Vooral in mondelinge communicatie: 'Die kans is maar zes promille.' Misschien zou je zelfs liever 'een half procent' zeggen, of simpelweg 'minder dan één procent'.

De wetenschappelijke notatie is handig voor alle getallen in de onderste helft van de tabel, waar de numerieke waarden niet-alledaags klein worden. Een kans van 0,00000001898956247803 zegt niet zoveel, ondanks het grote aantal cijfers. En 0,0% laat informatie weg. Het voordeel van $1{,}9 \times 10^{-8}$ is dat alleen al die macht voldoende is. Twee significante cijfers zijn vrij standaard. Eentje geeft al goed aan hoe groot het getal ongeveer is, en met twee heb je dat al aardig scherp vastgelegd. Voor minder wetenschappelijk lezerspubliek is het beter om in plaats van 2×10^{-8} te schrijven dat de kans 1 op 50 000 000 is.

Wanneer is een kans verwaarloosbaar? Eigenlijk nooit. Althans: er is geen algemeen getal voor te geven, want het hangt af van *het belang* of iets gebeurt. Te laat komen in een les is vervelend, maar minder erg dan dat een patient overlijdt als gevolg van een onnodig risico. Hoeveel speling neem je om de bus te halen? Nooit zoveel dat de kans 100,00% is dat je hem haalt. Als 5% van de mensen die een nieuwe pijnstiller slikt eraan overlijdt, kun je beter paracetamol gebruiken. (Je hoofdpijn is trouwens wel weg als je overlijdt.) Zelfs als het 1% of 1‰ is. Maar als 5% van de patiënten milde bijwerkingen heeft terwijl de kans op genezing van een ernstige ziekte halveert, is er reden om het wél te gebruiken.

Oefeningen

Je bent voor de Vierdaagse van Nijmegen verantwoordelijk voor de veiligheid van de wandelaars. In het parcours ontdek je een ontbrekende tegel die slecht zichtbaar is. Je schat de kans dat een deelnemer erdoor ten val komt en iets breekt erg klein. De kans dat iemand dit op tijd ziet, schat je op 99,99%.

1. Bereken de kans dat alle 47 000 wandelaars niet ten val komen door die ontbrekende tegel.
2. Als de kans exact 99,99% zou zijn, hoeveel wandelaars mogen er dan meedoen om de kans kleiner dan 99% (of 1%) te laten zijn dat er iemand valt over die ontbrekende tegel?

Gek eigenlijk… Je kunt toch niet vallen over een ontbrekende tegel?
Nee. Je valt waarschijnlijk over de volgende tegel (die *niet* ontbreekt).

Rekenen

54. Absolute & relatieve onzekerheden

De woorden ‘absoluut’ en ‘relatief’ spelen in het dagelijkse taalgebruik een nogal verschillende rol. Iemand kan ‘absoluut!’ als antwoord geven op een vraag. Dan betekent het ‘ja’ in een sterke vorm. Maar als je zegt: ‘alles is relatief’, dan is er juist weinig sprake van stelligheid.

Als je houdt van zo oorspronkelijk mogelijk Nederlands, zul je in plaats van ‘absoluut’ liever ‘volkomen’, ‘volstrekt’ of ‘zeker’ zeggen. In plaats van ‘relatief’ kies je dan voor ‘betrekkelijk’ of ‘verhoudingsgewijs’. Toch hebben die woorden niet precies dezelfde betekenis. ‘Absoluut’ komt van het Latijnse *absolvere*, dat ‘losmaken’ betekent. Daar zit het wezenlijke verschil met ‘relatief’, waar je iets nou juist niet losmaakt van andere dingen, maar er iets bij betrekt of een verhouding aangeeft.

De meetcorrectie die we eerder zagen, bedroeg 3% bij snelheden vanaf 100 km/u. De correctie staat dus niet los van de snelheid, maar hangt ervan af. Je kunt je afvragen of 3% veel of weinig is. Stel dat men in buurland Duitsland een andere norm heeft (wat overigens niet het geval is), dan zou je kunnen zeggen dat het relatief veel of weinig is. Technisch gezien is het interessanter om het te relateren aan de onzekerheid van het meetsysteem. Als die slechts 2% zou zijn, dan is 3% relatief veel. Hoewel… het scheelt maar een factor anderhalf: 3% is 50% meer dan 2%, om het maar eens ingewikkeld te maken. Ook als je iets in verhouding ziet tot iets anders, blijft het betrekkelijk of je het veel of weinig vindt.

Wat nu als je die 3% zou toepassen op een trajectcontrole? De apparatuur stelt vast dat een auto een afstand van 850 meter aflegt in exact één minuut. Wordt de bestuurder van die auto dan bekeurd? Deze kun je uit je hoofd uitrekenen, want er gaan zestig minuten in een uur. De gemiddelde snelheid is dus 51 km/u. Minder dan 3% te hard, maar is 3% niet relatief veel voor een meting over zo’n grote afstand en tijd? Op 850 meter is 3% gelijk aan 25,5 meter en op één minuut is 3% gelijk aan 1,8 seconde. Gevoelsmatig zeg je dat dat wat veel is.

Nog even over beide vormen van meten: die flitspalen stellen de snelheid op één bepaald moment vast, de trajectcontrole geeft een gemiddelde snelheid aan over een wat langere tijd. In geen van beide gevallen is het onmogelijk dat een automobilist op een bepaald moment te hard reed en toch niet bekeurd werd. Het omgekeerde is wel het geval: wie bekeurd wordt, reed sowieso te hard. Absoluut!

Wiskundig zijn absolute en relatieve fouten (relatief) makkelijk weer te geven. Als we een werkelijke waarde x_w willen meten, vinden we iets anders, namelijk de gemeten waarde x_g met een onzekerheid δx.

$$x_w = x_g \pm \delta x$$

De *absolute* onzekerheid is δx, een positief getal (of nul, maar dan zou je een perfecte meting hebben) met (meestal) een eenheid. De *relatieve* onzekerheid δ_{rel} is een dimensieloos getal, meestal uitgedrukt in een percentage.

$$\delta_{rel} = \frac{\delta x}{|x_g|}$$

Stel: er komt een nieuwe Tripel op de markt. "Nu met 10% meer alcohol!" staat er op het etiket. De vorige versie had 8%. Dus nu... 18%? Nee hoor. 10% méér betekent: 10% van 8% = 0,8 procentpunt erbij. Het nieuwe alcoholpercentage is dus 8,8%. Dat verschil kun je op twee manieren beschrijven:
– 0,8 procentpunt (het verschil tussen twee percentages)
– 10 procent meer (de relatieve toename)

Veel mensen halen die begrippen door elkaar. Toch is het verschil groot. Van 2% naar 1% betekent een daling van 1 procentpunt, maar wel 50% minder. Een *procentpunt* zegt iets over het *verschil* ('min' dus), *procent* over de *verhouding* ('gedeeld door').

Oefeningen

Iemand meet de lengte van een tafelblad en vindt een lengte en een onzekerheid in centimeters. Daarna rekent hij of zij dit om naar *inches*.

1. Wordt de *relatieve fout* door die omrekening groter of kleiner?
   Of blijft hij gelijk?

2. Wordt de *absolute fout* door die omrekening groter of kleiner?
   Of blijft hij gelijk?

3. *Die flitspalen stellen de snelheid op één bepaald moment vast, de trajectcontrole geeft een gemiddelde snelheid aan over een wat langere tijd.*
   In geen van beide gevallen is het onmogelijk dat een automobilist op een bepaald moment te hard reed en toch niet bekeurd werd.
   Leg uit waarom dit zo is.

55. Kleine waarde, grote onzekerheid

In het vorige hoofdstuk zagen we modulusstrepen verschijnen voor de relatieve fout. De grootheid die je meet, zou namelijk negatief kunnen zijn. Als je een weegschaal zodanig kalibreert dat hij exact 0 g aangeeft als er een leeg bakje op staat, dan zal een meting van datzelfde lege bakje heus wel zoiets kunnen opleveren als −0,2 g.

Een spanningsmeting is ook een aardig voorbeeld. Als je een spanning (potentiaalverschil) van ongeveer 0 V zou verwachten, kan het heel goed zijn dat je een negatief aantal millivolt op je scherm krijgt. Niks mis mee: dat is gewoon nul met een onzekerheid. $V_g = -0{,}01 \pm 2\ mV$ Mijn absolute fout bedraagt dan 2 mV, de relatieve onzekerheid is… 200%. TWEEHONDERD PROCENT! Is dat geen waardeloze meting dan? Nee. Het heeft alleen geen zin om over relatieve fouten te spreken als je 'nul' verwacht. Alles is relatief veel vergeleken met niks. Zelfs weinig is al veel als je het vergelijkt met niets.

Hoe groot is eigenlijk de relatieve fout als de onzekerheid 2 mV is, de werkelijke waarde 4 mV en de gemeten waarde 5 mV? Dat is een wat vreemde vraag, want we kennen de werkelijke waarde niet. We meten het volgende.

$$V_g = 5 \pm 2\ mV$$

De relatieve onzekerheid is dus 40%. Maar onze meting wijkt slechts 1 mV af van de werkelijke waarde van 4 mV. Dat is maar 25%. En eigenlijk zou je willen weten hoeveel procent de onzekerheid bedraagt ten opzichte van de *werkelijke* waarde. Hoe onzeker ben ik ten opzichte van wat ik wil weten?

Het probleem is: de werkelijke waarde kennen we niet. Het beste wat we hebben, is de meting, die normaal gesproken aardig in die richting zal liggen. Dit 'probleem' doet zich trouwens alleen maar voor bij onzekerheden die groot zijn ten opzichte van de gemeten waarde. Als je er 100 mV bij optelt, valt het allemaal wel mee. 2 mV op 105 mV is 1,90 % en 2 mV op 104 mV is nog geen promille meer dan dat: 1,92 %.

Terug naar die snelheidsmeting. Er zijn plekken waar je 15 km/u mag rijden. Het voelt onrechtvaardig om een boete te krijgen als je daar 16 km/u rijdt. Toch is dat 6,7 % te snel, die ene kilometer per uur.

Je ziet trouwens maar zelden flitspalen op plekken waar je maar 15 km/u mag rijden.

56. De spreiding schatten

De *standaardafwijking* is het *kwadratisch gemiddelde* van de *afwijkingen van het gemiddelde.*

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2}$$

Om die te bepalen, heb je dus nodig:

- alle individuele waarden
- hun gemiddelde μ
- hun aantal n

Het goede nieuws: die laatste twee volgen uit de eerste.
Het slechte nieuws: meestal heb je niet alle individuele waarden.

Oorzaak van het slechte nieuws: het zijn er meestal gewoon te veel, vaak zelfs *oneindig* veel. Als het om meetwaarden gaat, zijn de uitkomsten die je hebt dan *de populatie*? Eigenlijk niet. Ze vormen *een steekproef* van een andere populatie, namelijk alle mogelijke meetwaarden. Je kunt namelijk oneindig vaak meten, en je doet dat slechts een aantal keer.

Omdat de meetwaarden slechts een steekproef vormen, is het berekende niet de *populatiestandaardafwijking* σ maar de *steekproefstandaardafwijking* s.

Je hebt niet alle n (oneindig veel?) individuele waarden. Wat nu? Het ligt voor de hand om dan maar uit te gaan van wat je wel weet. Je hebt slechts N (tien?) waarden. Je kunt het gemiddelde van de populatie (de theoretische verdeling) schatten door het gemiddelde $\bar{x}$ van de steekproef te nemen. Als je daarbij de formule gebruikt voor de standaardafwijking, dan krijg je dit:

$$s_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

Hier gaat iets mis… Het probleem is namelijk dat je eigenlijk met het ‘echte’ gemiddelde zou moeten werken: het gemiddelde van die normale verdeling. Die hebben we niet, en dus werken we met een schatting ervan: het steekproefgemiddelde. Dat klopt niet, maar we hebben niets beters. Die schatting van de spreiding kan bij voldoende metingen best aardig zijn. Al kan het wel beter.

Stel je een plein voor met een aantal mensen erop. Ze staan kriskras verspreid. In het midden staat een paal. Die staat precies in het midden en is dus 'het werkelijke gemiddelde'. Je meet de afstand van iedereen tot die paal en berekent de spreiding.

Stel nu dat die paal er niet stond? En dat je niet wist wat het midden van het plein was. Iedereen is op een volstrekt willekeurige plek gaan staan, zonder voorkeur voor een bepaald deel van het plein. Elk plekje had evenveel kans. En *gemiddeld* stonden al die mensen dus midden op het plein. Als er oneindig veel mensen stonden — die passen er nooit op — zou hun gemiddelde het midden van het plein zijn.

Wat doen we als die paal er niet staat? Dan *schatten* we het midden. We nemen de gemiddelde positie van al die mensen zelf en zetten daar vervolgens die paal neer. Niet precies op het midden van het plein, maar wel ongeveer. Als er heel veel mensen stonden, zou het best nog aardig kloppen. Vervolgens meten we alle individuele afstanden tot die nieuwe paal.

Wat blijkt? Die nieuwe paal ligt *gemiddeld dichter* bij de mensen dan de oorspronkelijke paal. Geen wonder — hij is immers op hún positie gebaseerd. Als er twee mensen waren, stond die paal precies tussen hen in. Als er maar één persoon aanwezig was, zouden we op zijn of haar plek een paal planten. Die ene persoon heeft geen afstand tot het gemiddelde. Hij of zij ziet zichzelf als het middelpunt van het hele plein. Bij meer aanwezigen wordt dat effect kleiner, maar de afstanden lijken kleiner dan ze werkelijk zijn. *Je onderschat dus de spreiding.*

Dat is precies wat er gebeurt als je de standaardafwijking berekent uit een steekproef: je gebruikt het gemiddelde van diezelfde steekproef, en dat ligt automatisch een beetje gunstig. Om die onderschatting te corrigeren, deel je niet door N, maar door N − 1. Dat heet de Bessel-correctie. Het bewijs daarvan staat in een bijlage van dit boek. Dat het gevoelsmatig klopt en een beetje logisch is, volgt uit het paalverhaal.

Laten we eens kijken wat er gebeurt met een gemiddelde uit een steekproef. Bijvoorbeeld door te gooien met een dobbelsteen. Het liefst met een vijfzijdige dobbelsteen, want het gemiddelde aantal ogen is dan een geheel getal: 3. Nu valt het niet mee om vijfzijdige dobbelstenen te vinden die 'eerlijk' zijn, maar als je Dungeons & Dragons speelt, ken je wel de Twintigzijdige & Tienzijdige. Voor de kenners: een d20 & d10.

Vervang op je tienzijdige dobbelsteen de 6, 7, 9, en 10 door 1, 2, 3, 4 en 5 en je het een tienzijdige steen met alle cijfers 1 t/m 5 twee keer. De kans op elk van die vijf uitkomsten is dan gelijk. Plakkers klaar? Dobbelen maar!

Ho, wacht… Voordat we met stenen gaan gooien: ik reken liever met de *varianties* in plaats van met de *standaardafwijkingen*. 'Variantie' is een korter woord, maar belangrijker nog: daar zit geen wortel in. De variantie σ^2 is namelijk het kwadraat van de standaardafwijking σ.

De vijfzijdige dobbelsteen heeft als waarden 1, 2, 3, 4 en 5. Het gemiddelde is dus 3 (en niet die rare 3½ van gewone dobbelstenen) en de afwijkingen zijn:

aantal ogen	1	2	3	4	5
afwijking van 3	2	1	0	1	2
kwadraten	4	1	0	1	4

> Ik begin nu te snappen waarom die vijfzijdige steen zo lekker werkt. Kijk eens wat een mooie gehele getallen! Zelfs het optellen van de kwadraten en het delen door 5 lukt me nog uit het hoofd: 4 + 1 + 0 + 1 + 4 = 10. Dat kan ik heel makkelijk delen door 5.

De variantie van het aantal ogen van een vijfzijdige dobbelsteen is 2.

De variantie van drie willekeurige waarden, zou die gemiddeld niet gelijk moeten zijn aan de theoretische variantie die we berekenden? Die 2, weet je nog... Misschien denk je van wel. Maar wedden van niet?

Mijn simulatie komt (bij één miljoen herhalingen, ik kon het niet laten) op gemiddeld 1,333. En als ik het overdoe op 1,332. En nog een keertje, om het af te leren: 1,332. Ongeveer 4/3 dus… maar geen 2.

> Even kijken of ik nog met breuken kan rekenen.
> Hoeveel is 2 gedeeld door 4/3? Delen door een breuk: vermenigvuldigen met het omgekeerde. Dan is het 2 × 3/4 = 6/4 = 3/2. Anderhalf dus.

Nu naar de formule die je in wiskundeboeken altijd ziet voor het schatten van de variantie van de verdeling, op grond van een steekproef. Die luidt:

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

En dus niet wat wij doen:

$${s_N}^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

Die twee formules kun je eenvoudig door elkaar delen. Het som-deel is het zelfde. En delen door $1/N$ is hetzelfde als vermenigvuldigen met N. De deling levert dus $N/(N-1)$ op. Als N gelijk is aan 3, is dat dus 3/2.

Hé daar is die 3/2 die we net hadden, toch? Ja. Toen we heel veel simulaties deden met vijfzijdige dobbelstenen, kwamen we op 1,332 en 1,333. Vermenigvuldig je dat met 3/2, dan vind je 1,998 en 1,9995. Een promille (of minder) verschil met 2.

Hebben we nu een wiskundig bewijs? Nee. Maar we hebben wel gezien dat het waar is: een willekeurige steekproef uit de getallen 1 t/m 5 geeft ons gemiddeld wel die gemiddelde 3, maar niet die gemiddelde variantie 2. Dat komt doordat we ons gemiddelde schatten uit dezelfde steekproef als waar de variantie uit komt.

Je zou natuurlijk voor de aardigheid dat programma kunnen aanpassen. Stel je dat je een eenendertigzijdige (huh?) dobbelsteen hebt en je gooit daar dertien keer mee, klopt het verhaal dan ook? Ja.

Mijn simulatie komt op 73,840 bij een miljoen simulaties.

De factor waarmee we nu moeten corrigeren:

$$\frac{N}{N-1} = \frac{13}{13-1} = \frac{13}{12} \approx 1{,}0833$$

De variantie zou dus 73,840 × 1,0833 = 79,991 zijn.

De werkelijke waarde vind je door in Excel de waarden 1 t/m 31 in kolom A te zetten en de formule `=VAR.P(A:A)` toe te passen, of in Python met NumPy de functie `np.var(range(1,32))` uit te voeren.

> Let op dat je 32 (en dus niet 31) neemt als bovengrens die niet meedoet.

Beide berekeningen geven 80 als variantie; dat scheelt niet veel met 79,991.

Hebben we nu een wiskundig bewijs? Nee. Maar het zou me niet verbazen als het waar is, van die $N - 1$.

Oefeningen

Domme Dodo heeft de pech dat zijn **Dungeons & Dragons** dobbelstenen niet compleet meer zijn: de d10 en de d20 zijn foetsie.

1. DD kiest ervoor om dan maar met twee stenen te gooien, namelijk de d4 en de d6. De som van het aantal ogen is dan ook maximaal 10. Leg uit wat er fout gaat door het histogram van de kansen te geven.
2. Hoe zou je dit probleem aanpakken als je DD was, maar dan minder dom. (En zeg nou niet: nieuwe dobbelstenen kopen.)

57. De wortel van het gemiddelde

De *wortel van het gemiddelde*, is die gelijk aan het *gemiddelde van de wortel*?

We zouden met 9 en 25 een simpel testje kunnen de doen. De wortels van die twee getallen zijn 3 en 5. Het gemiddelde van 3 en 5 is 4. Als de wortel van het gemiddelde van 9 en 25 nou ook eens 4 was…

Helaas: het gemiddelde van 9 en 25 is 17 en de wortel van 17 is meer dan 4. De wortel van 16 is al 4. Die van 17 is 4,123 als je afrondt op drie decimalen.

> Het verschil tussen 9 en 25 is trouwens 16 en dat is het kwadraat van 4, maar dat is toeval. En dat is meteen een leuk voorbeeld dat één sommetje niet bewijst dat iets altijd waar is. Eén sommetje kan wel bewijzen dat iets *niet waar* is. Tenzij je sommetje fout is, want dan bewijst het weer niets… Hoe bewijs je dat je sommetje klopt? En waarom zijn dingen soms zo moeilijk in het leven? Alleen wiskunde is al lastig.

Als je bij getallen iets optelt, kun je bij hun gemiddelde hetzelfde optellen en dan krijg je dat nieuwe gemiddelde.

Als je getallen ergens mee vermenigvuldigt, kun je hun gemiddelde met datzelfde getal vermenigvuldigen en dan krijg je dat nieuwe gemiddelde.

Als het gemiddele van a en b gelijk is aan c, dan is het gemiddelde van a + d en b + d gelijk aan c + d. (Hun som is 2d groter, en dat deel je door 2.)

Als het gemiddele van a en b gelijk is aan c, dan is het gemiddelde van e × a en e × b gelijk aan e × x. (Hun som is namelijk e keer zo groot.)

Bij wortels en kwadraten gaat dat niet op. Bij delen en aftrekken wel.

Waarom is dat belangrijk? Omdat we net hebben aangetoond dat we een zuivere schatter hebben voor de *variantie* van een steekproef. Maar daarmee hebben we helaas geen zuivere schatter voor de *standaardafwijking* van een steekproef. De verwachtingswaarde van de wortel is namelijk iets anders dan de wortel van de verwachtingswaarde. Of anders gezegd: de wortel van het gemiddelde is niet het gemiddelde van de wortel.

> Het aantal letters in ‘de wortel van het gemiddelde’ is wel gelijk aan het aantal letters in ‘het gemiddelde van de wortel’. Maar daar koop je niks voor. Het staat hier alleen om de bladzijde vol te maken. Daarom moet je die teksten in de grijze kaders ook niet lezen. Daar staan zelden belangrijke dingen in. *The root of the mean* in het Engels is niet hetzelfde als *the root of all evil*. Want *mean* is wel gemeen, maar niet slecht.

58. Waarom RMS?

Waarom RVS? Omdat het niet gaat roesten. Maar waarom dan nog RMS? Je moet dan nog steeds middelen. En kwadrateren. En worteltrekken, wat klinkt als een combinatie van een kies trekken en een wortelkanaalbehandeling.

Het ligt voor de hand: als je wilt weten hoe ver getallen gemiddeld van het gemiddelde af liggen, dan neem je gewoon de afwijkingen. Alleen: daar komt altijd nul uit. Tenzij je ze eerst allemaal positief maakt (de grootte neemt, ofwel de absolute waarde) en dan pas middelt. Dat is de gemiddelde absolute afwijking, ook wel bekend als MAD: *mean absolute deviation*. Makkelijker uit te leggen is er haast niet. En toch gebruiken we hem zelden.

Vrijwel alle natuur- en gedragswetenschappen, alle technische handboeken en alle softwarepakketten kiezen voor een andere maat: de wortel van het gemiddelde van de kwadraten van de afwijkingen. De *standaardafwijking*. Of, in zijn notatiegedaante: de *wortel van de variantie*. Dat voelt misschien omslachtig. Waarom zou je eerst kwadrateren, dan middelen, en daarna ook nog de wortel? Dat is toch ingewikkelder dan gewoon de absolute waarde?

Ja. Maar juist het kwadrateren levert iets bijzonders op. Als je getallen kwadrateert, verdwijnen negatieve verschillen op een manier die wiskundig milder is. Kwadraten zijn namelijk continu, glad, en differentieerbaar — eigenschappen waar de calculus dol op is. Als je twee onafhankelijke meetfouten bij elkaar optelt, kunnen hun varianties worden opgeteld. Niet ongeveer of bij benadering, maar *exact*. Lekker, toch?

Bij absolute afwijkingen is dat niet zo. Daarom is de standaardafwijking heel prettig: die maakt het mogelijk om onzekerheden uit verschillende bronnen netjes te combineren. Het is dus niet zomaar dat het foutdoorwerkingsrecept uitgaat van varianties. Zoals Kenney & Keeping (1951) het al opschreven:

"The use of the mean square deviation is preferred not because it is more intuitive, but because it is more tractable. The squared deviations lead to simpler algebraic manipulation and have desirable statistical properties."

We kiezen niet voor kwadraten omdat ze zo mooi zijn met die tweetjes in hun macht, maar omdat ze beter werken.

Een ander voordeel is dat de standaardafwijking naadloos aansluit bij de normale verdeling. De klokvormige vaste gast in Statistiek & Natuurkunde heeft in zijn formule een term staan met een kwadraat. En die formule is niet willekeurig. Ze komt voort uit het optellen van veel kleine onafhankelijke storingen — precies het terrein waarop de standaardafwijking schittert.

Veel statistiek rondom toetsen, intervallen en betrouwbaarheidsgrenzen is gebaseerd op dit idee. Vandaar ook dat David Moore (2009) schrijft:

"Although the mean absolute deviation is easier to understand, the standard deviation has superior mathematical properties and is the basis for most inferential statistics."

Toch is het goed om te weten dat er alternatieven bestaan. In robuuste statistiek, waarbij je uitschieters of verstoringen door kleine aantallen juist negeert, komen absolute afwijkingen weer om de hoek kijken. John Tukey en Frederick Mosteller (1977) schreven daar al over:

"The choice of squared deviations is historically driven and statistically justified. But in applications where robustness is essential, alternatives like absolute deviations deserve attention."

Zo'n zin doet je even knikken. Want dat is de kern: de standaardafwijking is gekozen omdat zij goed rekent, goed combineert en goed aansluit bij wat we verder willen doen. Ze is een keuze die voortkomt uit de eisen van de praktijk én de kracht van de theorie.

Dus waarom nemen we de wortel van het gemiddelde van de kwadraten? Omdat het werkt. Niet omdat het intuïtief is, niet omdat het mooier klinkt — maar omdat het ons verder helpt. In berekeningen, in foutdoorwerking, in modellen en in verwachtingen. En omdat het ons, ironisch genoeg, een stevigere standaard geeft dan de gemiddelde absolute afwijking kan bieden.

Er zijn meer manieren om spreiding weer te geven. Bijvoorbeeld met *kwartielen*: een soort dubbele mediaan. Bij kwartielen verdeel je alle data in vier even grote groepen (kwarten dus).

Hoe doe je dat? Bepaal eerst de mediaan en noem die Q_2. Neem dan alle waarden onder die mediaan, en bepaal daar opnieuw de mediaan van: dat is Q_1. Doe hetzelfde alles boven de mediaan: die geven Q_3. Zo krijg je drie getallen die de groep in vier hokjes opdelen — van laag naar hoog.

Zomaar een ander voorbeeld: je kunt tentamencijfers splitsen in voldoendes en onvoldoendes, en van beide groepen het gemiddelde nemen. Dat geeft ook een rijker beeld dan 'gewoon' een *average*.

> Bonus: noteer het minimum en het maximum als Q_0 en Q_4.
> Met vijf kwartielen krijg je een *vijf-getallensamenvatting*.

De moraal van dit uitstapje: er zijn vele vormen waarin je de grootte en de spreiding van een verzameling gegevens samen kunt vatten — met twee of drie kentalen die elk op hun manier iets laten zien.

59. Foutdoorwerking bij optellen

Stel dat je de hoogte van een blikje wilt meten. Je doet dat met een duimstok en komt op 113,4 mm. Daar hoort een onzekerheid bij als gevolg van het aflezen. Laten we aannemen dat die onzekerheid een halve millimeter is. Wat is dan de hoogte van een stapel van zes blikjes?

Je zult eerst iets meer over de blikjes moeten weten. Bij frisdrank bijvoorbeeld zijn de blikjes zo gemaakt dat ze een beetje 'in elkaar schuiven' als je ze opstapelt. Dat is handig, want dan vallen ze niet zo snel om. Maar voor het optellen van lengtes komt dat minder goed van pas.

Laten we ervan uitgaan dat we blikjes hebben waar dat niet zo is. De hoogte van de stapel is dan 6 × 113,4 = 680,4 mm. Hoe groot is de onzekerheid dan?

We hebben maar van één blikje de hoogte gemeten. De onzekerheid bedroeg 0,5 mm. Het zou dus kunnen dat de hoogte niet 113,4 mm was, maar 113,9 mm. Of 113,2 mm. Of 113,0 mm. Het zou ook 114,1 mm kunnen zijn, ook al is dat meer dan de beste schatting (113,4 mm) plus de onzekerheid (0,5 mm). Een onzekerheid geeft immers geen *harde grenzen* aan, maar een gebied waarin de werkelijke waarde *waarschijnlijk* ligt. Voor dit rekenvoorbeeld doen we het even anders. We veronderstellen een uniforme verdeling tussen de genoemde grenzen en gaan uit van tienden van millimeters. Uitgedrukt in tienden van millimeters zijn er elf mogelijke 'echte' waarden: 112,9 mm, 113,0 mm, 113,1 mm, 113,2 mm … 113,8 mm, 113,9 mm.

> Mmmmm… Hierboven staat wel erg vaak een letter 'm'. Hoe vaak eigenlijk? Dat ligt eraan wat je onder 'hierboven' verstaat.
>
> Maar er is nog iets: we hebben 'zomaar' gezegd dat de onzekerheid 0,5 mm bedroeg. Hoe weten we dat? Wij als lezer weten het omdat het al eerder in de tekst staat. Maar in de praktijk zijn er metingen waar je over de onzekerheid zelf in het ongewisse verkeert. Verkeerde aannames liggen dan op de loer en dan klopt de rest van je berekening ook niet.

Stel dat we er 0,2 mm naast zaten. Dan zitten we er bij twee blikjes op elkaar 2 × 0,2 mm = 0,4 mm naast en bij zes blikjes wordt dat 6 × 0,2 mm = 1,2 mm. Niet alleen de blikjes, maar ook de *fouten* stapelen zich op.

We hebben het nu over *fouten*, alsof we *weten* hoe groot het verschil is tussen de werkelijke waarde (*true value*) en onze beste schatting (*best estimate*), wat in werkelijkheid natuurlijk niet zo is. Maar met een simpel rekensommetje willen we kijken of we inzien hoe zo'n fout doorwerkt bij het optellen.

De totale hoogte is volgens deze redenering maximaal (680,4 ± 3,0) mm. Dat is een significant cijfer te veel! Ook hier geldt: het gaat nu even over het snappen van hoe die fout doorwerkt in de optelling.

> Wat is erger: een significant cijfer te veel, of een significant cijfer te weinig? In het eerste geval suggereer je meer dan je weet. In het tweede geval gooi je kennis weg.
>
> Iets heel anders: veel mensen laten altijd de spatie weg in 'te veel'. Waarom toch? Omdat er een zelfstandig naamwoord 'teveel' bestaat?
>
> Tegeltjesspreuk: *'Liever een cijfer te veel, dan een spatie te weinig.'*

Wat nu als we twee blikjes met verschillende afmetingen hebben? Dan zijn er twee beste schattingen en twee onzekerheden. De redenering wordt daar niet anders door: de onzekerheden tel je bij elkaar op. Als het tweede blikje een hoogte heeft van (162,7 ± 0,5) mm, is de totale hoogte van beide blikjes minimaal 113,4 − 0,5 + 162,7 − 0,5 = 275,1 mm en maximaal 113,4 + 0,5 + 162,7 + 0,5 = 277,1 mm, wat je dus kunt schrijven als (276,1 ± 1,0) mm.

Zijn we nu niet wat te pessimistisch over de onzekerheid in die optelling? Voor het eerste blikje gold dat we een hoogte hadden geschat. Die hoogte kon ook iets groter of iets kleiner zijn, net als voor het tweede blikje. Het is dus ook mogelijk dat de werkelijke waarde (*true value*) van het eerste blikje 0,2 mm *groter* was dan wat we gemeten hebben en dat de waarde van het tweede blikje 0,2 mm *kleiner* was dan wat we gemeten hebben. Dan zouden de twee afwijkingen elkaar precies opheffen.

Voor dit moment volstaan we met het domweg optellen van de onzekerheden. In hoofdstuk 46 komen we terug op het feit dat fouten elkaar kunnen versterken, maar ook (gedeeltelijk) compenseren.

Oefeningen

1. Twee weerstanden van 1,0 kΩ hebben een onzekerheid van 5%. Hoe groot (in procenten) is de relatieve onzekerheid van beide weerstanden samen als je ze in serie zet?
2. Aan 2,4 liter vloeistof (onzekerheid: 8%) wordt 600 cm^3 vloeistof toegevoegd. Hoe groot mag de absolute onzekerheid in de toegevoegde vloeistof zijn om de relatieve onzekerheid in de totale vloeistof kleiner dan 10% te houden?
3. In dit hoofdstuk staat de vraag "Hoe vaak eigenlijk?" Bedenk een manier om die vraag te beantwoorden.
4. Is de vorige vraag niet raar voor een boek als Meten & Weten?

60. Doorwerking bij aftrekken

Bij het optellen van metingen met een onzekerheid moet je, uitgaande van het meest ongunstige geval, de onzekerheden ook optellen. Twee ijzeren staven hebben een lengte van een meter en een onzekerheid van 2 cm. Hoe lang zijn ze samen? Twee meter. Maar wat is hun onzekerheid? In het meest ongunstige geval waren ze 'dezelfde kant op' en waren ze allebei 102 cm lang (of 98 cm) en zijn ze dus samen 2,04 m of 1,96 m. Een verschil van acht centimeter. Je zou de totale lengte moeten opgeven als 2,00 ± 0,04 m.

Hoe zit dat met het *verschil* in twee metingen? Stel, je gebruikt beide staven om een opstelling te maken en zet ze daarom naast elkaar. In het ideale geval is het verschil in hoogte exact nul. Hoe zit dat in het meest ongunstige geval? Die ene staaf kon weleens 102 cm zijn terwijl de andere misschien 98 cm is. Dan verschillen ze dus 4 cm in lengte. Als je het lengteverschil berekent en correct noteert, kom je op 0 ± 4 cm of 0,00 ± 0,04 m.

> Je ziet dat beide antwoorden één significant cijfer hebben in de onzekerheid. Hoe groot is eigenlijk het aantal significante cijfers in 0,00 m? Daar hebben we het in hoofdstuk 67 niet echt over gehad, maar het valt wel te beredeneren. 0,00 is namelijk net zoiets als 0,01 of 0,07: het is een cijfer met twee nullen ervoor. Die laatste nul is dus een bijzonder geval: dat cijfer is wél significant.

Het zou ook een beetje vreemd zijn: een getal zonder significante cijfers heeft namelijk geen enkele betekenis.

Doordat we een verschil berekenden tussen waarden waarvan de beste schatting gelijk is aan nul, krijgen we de bijzondere situatie dat we een relatieve fout hebben die oneindig groot is. Beter gezegd: ongedefinieerd. De relatieve fout is namelijk 4 / 0 en je kunt niet door 0 delen. Je zou ook kunnen zeggen dat je 0,00 beschouwt als maximaal 0,005. In dat geval kun je spreken van een relatieve fout van 4 / 0,005. Dan heb je een factor 800, oftewel 80000 %. Niet oneindig, maar wel bizar groot.

Je ziet met dit eenvoudige voorbeeld dat wisselen tussen absoluut en relatief soms rare dingen oplevert. Het is voor een technicus wel belangrijk om telkens door beide glazen van dezelfde bril naar de dingen te kijken. Trouwens, niet alleen als technicus. Wie een winkeltje begint en op de eerste dag 14 artikelen verkoopt, moet niet zeggen dat hij of zij een oneindig grotere verkoop heeft dan de dag ervoor. Het is zelfs maar de vraag of het zo zinvol is om te spreken van ruim 50% stijging als je er de dag erna 21 verkoopt. Alleen grotere aantallen op langere termijnen zijn geschikt voor dergelijke analyses.

Terug naar de verschilmeting. Een verschil tussen twee positieve getallen is altijd kleiner dan de grootste van de twee. Maar de onzekerheid van het verschil is groter dan de grootste onzekerheid. Absolute fouten nemen toe bij het berekenen van een verschil. Relatieve fouten kunnen enorm groot worden, als de verschillen klein zijn.

Oefeningen

Twee weerstanden van 1,0 kΩ hebben beide een onzekerheid van 5%.
Het verschil tussen die weerstanden zou nul moeten zijn (0 Ω), maar daar zit een onzekerheid op.

1. Hoe groot is die absolute onzekerheid?
2. Hoe komt het dat de relatieve onzekerheid in dit verschil veel groter is dan 10 %?

'Doorwerking' kwamen we in het vorige hoofdstuk al tegen. Het is een vertaling van het Engelse *propagation*, en dat kom je in Meten & Weten alleen tegen in dit grijze kader en de literatuurlijst. Peralta (2012) heeft het in de titel van zijn boek staan. *Propagation Of Errors: How To Mathematically Predict Measurement Errors.*

Er staan bar weinig Nederlandse boeken in die lijst. Berendsen (1997), de uitzondering, heeft het over 'doorwerking van fouten'. We zitten op één lijn. Maar ja, met twee standpunten lukt dat al snel. In de wiskunde liggen twee punten *altijd* op één lijn.

Het woord *propagation* komt van het Latijnse *propagare*, dat zoiets betekent als 'voortplanten', 'uitbreiden' of letterlijk: 'aan een wijnstok hechten om een nieuwe loot te laten groeien'. In de plantkunde heeft *propageren* nog steeds die inhoud. In bredere zin kreeg het de betekenis van voortzetting, uitbreiding en verspreiding. Zo kon ook een gedachte zich voortplanten: via de *Congregatio de Propaganda Fide* — letterlijk: de congregatie voor het verspreiden van het geloof.

Vanuit dat kerkelijke gebruik verschoof *propaganda* langzaam naar de betekenis die we nu kennen: de gerichte verspreiding van een boodschap, meestal met een bijsmaak. Maar in de natuurkunde bleef het oorspronkelijke idee behouden: een verstoring (een foutje) plant zich voort. Niet via flyers, niet via sociale media, maar via een formule.

Toch voelt 'voortplanting van fouten' in het Nederlands wat vreemd aan. Het klinkt als een stel losgeslagen metingen die zich lustig parend explosief vermenigvuldigen tot een bonte familie van vergissingen.
En hoewel dat soms ook best klopt, is het niet wat bedoeld wordt.

61. Doorwerken bij vermenigvuldigen

Zou je bij een vermenigvuldiging de fouten ook gewoon bij elkaar mogen optellen? Bij optellen was dat de oplossing, en bij aftrekken kwam dezelfde uitkomst naar voren… Het leven zou wel makkelijk zijn als je alleen maar hoefde op te tellen en dan de gevraagde uitkomst vond.

Het antwoord op de vraag is 'nee', en dat kun je al zien als je alleen maar naar de dimensies kijkt. Een vermenigvuldiging zou namelijk uitgevoerd kunnen worden met twee grootheden met verschillende dimensies. Een moment bijvoorbeeld is het product van een kracht (in newton) en een arm (in meters). En iets met de eenheid newton kun je niet optellen bij een grootheid die in meters wordt uitgedrukt.

Stel dat je een moment berekent. De kracht is 8,0 ± 0,4 N en de bijbehorende arm is 20 ± 2 cm. We hebben dan één relatieve onzekerheid van 5% en eentje van 10%. Wat is dan de onzekerheid in het product van die twee? Dat je de onzekerheden onderling moet vermenigvuldigen, zal ook wel niet… Dan kom je namelijk op 0,5% en dat is veel kleiner dan de kleinste van de twee.

Laten we eens uitrekenen wat er maximaal en minimaal uit de berekening kan komen. Of eigenlijk: de maximaal en minimaal *waarschijnlijke* waarde. De onzekerheid van 0,4 N garandeert niet dat je nooit boven de 8,4 N zult meten. Maar het klinkt wel erg redelijk om aan te nemen dat we de onzekerheden kunnen optellen bij en aftrekken van de beste schattingen om een marge te vinden waarbinnen het product ligt.

$$M_{min} = F_{min} \times r_{min} = 7{,}6 \times 0{,}18 = 1{,}368\ Nm$$

$$M_{max} = F_{max} \times r_{max} = 8{,}4 \times 0{,}22 = 1{,}848\ Nm$$

Als we het verschil tussen deze twee waarden door 2 delen, komen we uit op 0,240 Nm. Zonder rekening te houden met significante cijfers krijgen we dus 1,608 ± 0,240 Nm als uitkomst. De relatieve onzekerheid zou dan 14,93% zijn.

Is het toevallig dat we op de som van 5% en 10% uitkomen? Zou kunnen. Laten we het eens uitschrijven in symbolen. Dat is veel algemener en geeft vast een beter inzicht. En laten we ook met de groottes van kracht en arm rekenen en niet met de vectoren. (Dat deden we al niet.)

$$M_{max} = M + \Delta M = F_{max} \cdot r_{max} = (F + \Delta F) \cdot (r + \Delta r) =$$
$$F \cdot r + F \cdot \Delta r + \Delta F \cdot r + \Delta F \cdot \Delta r$$

Als we nu één slimme zet doen, zijn we bijna bij de oplossing. Deel links en rechts van het isgelijkteken beide door het moment, dan krijg je rechts het volgende:

$$\frac{M + \Delta M}{F \cdot r} = \frac{F \cdot r + F \cdot \Delta r + \Delta F \cdot r + \Delta F \cdot \Delta r}{F \cdot r}$$

Dat kun je ook schrijven als

$$1 + \frac{\Delta M}{M} = 1 + \frac{\Delta r}{r} + \frac{\Delta F}{F} + \frac{\Delta F \cdot \Delta r}{F \cdot r}$$

Aan beide zijden van het isgelijkteken kun je de 1 weglaten. Dat we na die ene slimme zet van de deling bijna bij de oplossing zijn, zit in die laatste term. Als we die zouden weglaten, dan stond hier dat de relatieve fout in M gelijk is aan de relatieve fout in r plus de relatieve fout in F.

Mogen we die laatste term dan zomaar weglaten? Nee, niet zomaar. Maar meestal wel. Onzekerheden zijn vaak (lang niet altijd!) veel kleiner dan de waarden waarop ze betrekking hebben. Daardoor is het product van twee relatieve fouten meestal heel klein. In ons voorbeeld: 5% van 10% is 0,5%. Door die term weg te laten, kom je op 15% in plaats van 15,5%. Daar ligt niemand wakker van.

Die 14,93% was overigens ook al niet gelijk aan 15,5% of 15%. Dat heeft te maken met het feit dat we niet met $F \cdot r$ zijn gaan rekenen als beste schatting: we gebruiken 1,608 N in plaats van 1,6 N.

We hadden bovenstaande formule trouwens ook kunnen uitschrijven voor de minimale uitkomst van het moment.

$$M_{min} = M - \Delta M = F_{min} \cdot r_{min} = (F - \Delta F) \cdot (r - \Delta r) =$$
$$F \cdot r - F \cdot \Delta r - \Delta F \cdot r + \Delta F \cdot \Delta r$$

Ook dan kunnen we één slimme zet en links en rechts van het isgelijkteken door het moment delen.

$$\frac{M - \Delta M}{F \cdot r} = \frac{F \cdot r - F \cdot \Delta r - \Delta F \cdot r + \Delta F \cdot \Delta r}{F \cdot r}$$

Dat kun je ook schrijven als

$$1 - \frac{\Delta M}{M} = 1 - \frac{\Delta r}{r} - \frac{\Delta F}{F} + \frac{\Delta F \cdot \Delta r}{F \cdot r}$$

Ook hier kun je de laatste term verwaarlozen en dezelfde enen kunnen schrappen. Je komt dan op dezelfde uitkomst, maar dan met overal minnen.

Wiskundig is het geen echt bewijs en het klopt ook niet exact. Maar de algemene formule is wel wat we gebruiken bij het doorrekenen van onzekerheden bij vermenigvuldigen.

Als $z = x \times y$

dan is

$$\frac{\delta z}{|z|} = \frac{\delta x}{|x|} + \frac{\delta y}{|y|}$$

We vinden dus óók bij vermenigvuldigen de onzekerheid in het product via optellen. We tellen nu alleen niet de *absolute* onzekerheden bij elkaar op, maar de *relatieve* onzekerheden.

Oefeningen

Een rechthoekig stalen vat heeft een hoogte van 1 m, een breedte van 4 m en een lengte van 8 m. Het is voor ongeveer de helft (je kunt daar nog 5% naast zitten) gevuld met water. De onzekerheden in de afmetingen van het vat bedragen 6 cm.

1. Welke onzekerheid levert de grootste bijdrage in de onzekerheid van het volume van het water?
2. Hoe groot zou het vat moeten zijn om ervoor te zorgen dat de onzekerheid in het hoogtepeil van het water evenveel invloed heeft op de onzekerheid in het volume als de afmetingen van het vat?
3. Bereken het volume van het water in het vat.
4. Het volume van een bol wordt gegeven door onderstaande formule.

   $$V = \frac{4}{3}\pi r^3$$

   Hoe groot mag de absolute onzekerheid in de straal van een bol met een diameter van 1,00 dm zijn om de onzekerheid in het volume minder dan 6% te laten zijn?

62. En als je het niet mag verwaarlozen?

Stel, we willen een vermogen berekenen door een spanning en een stroom te meten. Het vervelende is dat die spanning en die stroom beide nogal klein zijn en daardoor een grote relatieve fout hebben.

$$U = 0{,}2 \pm 0{,}1\ V$$

$$I = 5 \pm 3\ A$$

Uit je hoofd reken je uit dat het vermogen 1,0 W is. En dat de relatieve fouten 50% en 60% zijn. Bij vermenigvuldigen tel je de relatieve fouten bij elkaar op en kom je op 110%. Nog steeds uit het hoofd: de absolute fout is 1,1 W.

$$P = 1{,}0 \pm 1{,}1\ W$$

Huh? Mijn sommetje zegt dat ik misschien wel een negatief vermogen heb! En dat terwijl de waarden van de spanning en de stroom beide positief zijn.

> Ze hadden trouwens ook beide negatief mogen zijn. Eén van beide negatief en de andere positief, kan natuurkundig gezien niet. Dan zou de stroom tegen de stroom in zwemmen…

Een negatief vermogen, dat lijkt me sterk. En dat is het ook. Want die spanning is waarschijnlijk niet kleiner dan 0,1 V en de stroom is waarschijnlijk niet kleiner dan 2 A. Het vermogen zal dus niet onder 0,2 W liggen. De hoogste waarschijnlijke waarde vind je door 0,3 V te vermenigvuldigen met 8 A. Dat levert 2,4 W op. Een aardigere uitkomst zou dus zijn:

$$P = 1{,}3 \pm 1{,}1\ W$$

Het vervelende is alleen dat je beste schatting nu niet het product is van je beste schattingen. Dat is namelijk 1,0 W.

Waar komt nu toch dat negatieve vermogen vandaan? Hoe kun je ooit op −0,1 W uitkomen? Het antwoord is: die formules helpen je niet aan de exacte antwoorden, maar geven — zelfs als je corrigeert voor de term die we verwaarlozen — een grove schatting die alleen maar klopt voor kleine onzekerheden. (Dan klopt die schatting ook niet, maar maakt het niet zo veel uit.)

Als we in een tabel uitschrijven wat alle mogelijke uitkomsten zijn voor de combinaties van zeven verschillende stroomsterktes en vijf verschillende spanningen, dan vinden we deze waarden.

Tabel 62.1 Getalswaarden (zonder eenheden dus!) van vermogen, spanning en stroomsterkte

	2	3	4	5	6	7	8
0,10	0,20	0,30	0,40	0,50	0,60	0,70	0,80
0,15	0,30	0,45	0,60	0,75	0,90	1,05	1,20
0,20	0,40	0,60	0,80	1,00	1,20	1,40	1,60
0,25	0,50	0,75	1,00	1,25	1,50	1,75	2,00
0,30	0,60	0,90	1,20	1,50	1,80	2,10	2,40

In de linker kolom staan spanningen in volt (V). In de bovenste rij staan stroomsterktes in ampère (A). Alle andere cellen zijn vermogens in watt (W). In het midden vind je keurig die 1,0 W die je zou verwachten. Maar die extreme waarden die uit de formule volgen, komen nergens voor.

Oefening

1. Bereken de gemiddelde waarde van de 35 mogelijke uitkomsten uit de tabel, en ga na of die gelijk is aan de waarde van 1,00 die precies in het midden van de tabel staat. (Derde rij, vierde kolom.)
   Geef een verklaring voor de overeenkomst of het verschil.

2. Lees de beide vragen hieronder door en beantwoord ze zonder eerst een berekening uit te voeren.
   (Je mag niet rekenen, wel nadenken en schatten. Dat is iets anders.)

3. Bereken de gemiddelde waarde van het kwadraat van het aantal ogen dat je gooit met één dobbelstenen.

4. Bereken de gemiddelde waarde van het product van het aantal ogen dat je gooit met twee dobbelstenen.

63. Doorwerken in het algemeen

Niet alle grootheden zijn te berekenen zijn via optellen, aftrekken, vermenigvuldigen en delen. In dit hoofdstuk een 'lastigere' functie. Het leuke daarvan is dat de aanpak zo algemeen is, dat hij op alle functies toepasbaar wordt.

We nemen als voorbeeld de wet van Bragg.

> Die wet is opgesteld door Bragg en zijn zoon. Die heette niet toevallig ook Bragg, dus eigenlijk is het de wet van Bragg & Bragg. We laten een van beiden weg, kijk zelf maar of je de vader of de zoon overhoudt.
> Je mag ook zeggen: door de heer Bragg en zijn vader.

De wet van Bragg beschrijft hoe met diffractie van röntgenstralen de afstanden tussen de atomen in een kristalrooster worden berekend. Het deel van de stralen dat wordt gereflecteerd, vertoont interferentiepatronen waaruit de afstanden kunnen worden afgeleid. De formule daarvoor luidt als volgt.

$$n\lambda = 2d \cdot sin\theta$$

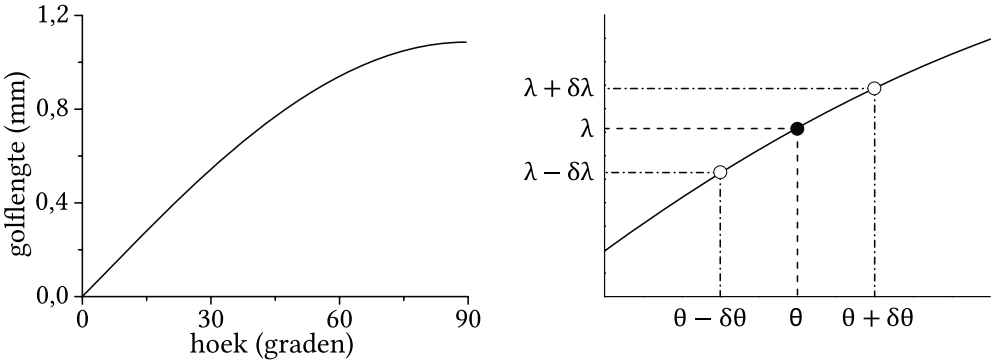
Hierin is

- n een geheel getal dat de diffractie-orde aangeeft;
- λ de golflengte van de röntgenstraling;
- d de afstand tussen de vlakken in het kristalrooster;
- θ de diffractiehoek tussen de stralen en de vlakken.

In dit voorbeeld gaan we uit van een eerste orde, dus $n = 1$. Als er in bovenstaande formule een grootheid s had gestaan waar nu $sin\theta$ staat, dan hadden we te maken met een eenvoudige vermenigvuldiging van 2, d en s. Dan kon je gewoon de relatieve fouten optellen in die drie factoren. (Eigenlijk alleen in de laatste twee, want een factor 2 heeft geen onzekerheid.)

Als we aannemen dat de onzekerheid in d verwaarloosbaar klein is, dan houden we een functie over waaruit blijkt dat de onzekerheid in λ alleen maar afhangt van de onzekerheid in θ. Het lastige is: daar zit nog een sinus in. Hoe pakken we dat aan?

Kijk eens naar onderstaande twee figuren. Links staat een grafiek waarin de golflengte is uitgezet tegen de hoek. Rechts staat een klein stukje van diezelfde grafiek, sterk ingezoomd op het punt waarin we geïnteresseerd zijn.

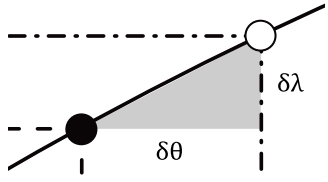


Figuur 63.1 Golflengte als functie van de hoek

Omdat we ver ingezoomd zijn, is het getoonde gedeelte van de grafiek bijna een rechte lijn geworden. Een rechte lijn, die dus te beschrijven is met een eenvoudige vergelijking:

$$\lambda = a\theta + b$$

Hoe groot is dan a en hoe groot is dan b? Om met de laatste te beginnen: die interesseert ons hier niet. Laten we nog verder inzoomen op het plaatje.



Figuur 63.2 Golflengte als functie van de hoek (ingezoomd)

De lijn tussen de zwarte en de witte bol is net niet helemaal recht, maar het scheelt niet veel. Als we de lijn als recht beschouwen, dan geldt dat

$$\delta\lambda = |a|\delta\theta$$

Van a nemen we de absolute waarde, zodat het verhaal ook opgaat voor een dalende lijn, die een negatieve a heeft.

We zouden hetzelfde kunnen doen voor de andere kant van de zwarte bol. Als de lijn echt recht is, maakt dat niet uit. Als er een kromming in zit die niet symmetrisch is, kom je op andere waarden uit.

Hoe groot is de waarde van a? Die is gelijk aan de richtingscoëfficiënt van de raaklijn waarin we geïnteresseerd zijn. Die waarde kun je vinden door de

afgeleide te bepalen in dat punt. We moeten dan de afgeleide bepalen van de functie, in ons geval van

$$f(\theta) = 2d \cdot sin\theta$$

Die afgeleide is gelijk aan

$$\frac{df}{d\theta}(\theta) = 2d \cdot cos\theta$$

In het punt $\theta = 60°$ is de cosinus gelijk aan 0,5. Daar is de onzekerheid in de golflengte dus

$$\delta\lambda = 2d \cdot cos\theta \cdot \delta\theta = d \cdot \delta\theta$$

> Zoals bij elke formule vragen we ons af of de dimensies kloppen. De golflengte is in meters (of millimeters of nanometers, maar in ieder geval is het een lengte). De afstand d tussen de vlakken in het kristalrooster is ook een lengte. Maar $\delta\theta$ is toch in graden? Nee! Dit soort berekeningen kun je alleen maar uitvoeren als je in radialen rekent. Dan ben je namelijk af van de tamelijk willekeurige keuze dat een rechte hoek 90° is. Een rechte hoek is een kwart van een hele cirkel, die een omtrek heeft van 2π maal de straal. Rekenen in radialen geeft het enige juiste verband tussen de lengte langs de rand van de cirkel, de hoek en de straal.

Wat gebeurt er in de uiterste gevallen? Die sinus kan (in absolute waarde) variëren van 0 tot 1. Dat betekent dat de waarde van $\delta\lambda$, afhankelijk van de hoek maximaal gelijk is aan $2d \cdot \delta\theta$, en minimaal gelijk aan… nul! Dat kan toch niet? De onzekerheid in iets wat afhankelijk is van een grootheid waar onzekerheid in zit, kan toch niet nul zijn? Inderdaad. De afgeleide van de sinus (de cosinus dus) is *alleen maar* exact gelijk aan 0 bij een hoek van 90° (en 270°). De raaklijn loopt in dat punt precies horizontaal. Dat kun je ook zien in de grafiek die we eerder dit hoofdstuk bekeken. Vlak voor (en na) dat punt is de afgeleide ook erg klein, maar uiteraard niet nul.

In dit voorbeeld met de wet van Bragg geldt dus dat de onzekerheid bij kleine hoeken sterker afhangt van de onzekerheid in die hoek dan voor hoeken van bijna 90°. De onzekerheid is geen constante waarde (absoluut) en ook niet een bepaald percentage (relatief), maar een functie van de hoek zelf.

64. Doorwerken bij delen

Voordat we kijken naar doorwerken bij delen, doen we het sommetje met vermenigvuldigen nogmaals, maar dan via een andere aanpak, namelijk de algemene benadering van het partiële differentiëren.

Als

$$z = x \cdot y$$

dan is

$$\delta z = \frac{\partial z}{\partial x} \cdot \delta x + \frac{\partial z}{\partial y} \cdot \delta y$$

In woorden: de onzekerheid in z is gelijk aan de som van de onzekerheden in x en y, vermenigvuldigd met hun partiële afgeleiden. Die zijn vrij eenvoudig bij een vermenigvuldiging, en dus kom je tot

$$\delta z = y \cdot \delta x + x \cdot \delta y$$

Delen we links en rechts door z, dan hebben we een uitdrukking voor de relatieve onzekerheid.

$$\frac{\delta z}{z} = \frac{y}{z} \cdot \delta x + \frac{x}{z} \cdot \delta y$$

Dan zijn we er al bijna, want

$$\frac{y}{z} = \frac{1}{x} \text{ en } \frac{x}{z} = \frac{1}{y}$$

Er staat dus precies wat we al wisten:

$$\frac{\delta z}{|z|} = \frac{\delta x}{|x|} + \frac{\delta y}{|y|}$$

De verticale strepen om de absolute waarde te nemen mag je hier toevoegen, omdat onzekerheden altijd positief zijn.

Diezelfde benadering met partiële afgeleiden kun je toepassen op het quotient (de deling) van twee grootheden. Dan hebben we te maken met het volgende uitgangspunt.

$$z = \frac{x}{y}$$

De doorwerking van de onzekerheid berekenen we via

$$\delta z = \frac{\partial z}{\partial x} \cdot \delta x + \frac{\partial z}{\partial y} \cdot \delta y$$

wat leidt tot

$$\delta z = \frac{1}{y} \cdot \delta x + \frac{x}{y^2} \cdot \delta y$$

Omdat

$$y = \frac{x}{z}$$

geldt dus

$$\delta z = \frac{z}{x} \cdot \delta x + \frac{1}{y} z \cdot \delta y$$

Links en rechts delen door z leidt tot

$$\frac{\delta z}{|z|} = \frac{\delta x}{|x|} + \frac{\delta y}{|y|}$$

Zoals de formules voor optellen en aftrekken aan elkaar gelijk bleken te zijn, zijn de formules voor vermenigvuldigen en delen dat ook.

Te onthouden:

- Bij optellen (of aftrekken) van twee grootheden wordt de *absolute* onzekerheid in de som (of het verschil) gevonden door de *absolute* onzekerheden in beide grootheden bij elkaar op te tellen.
- Bij vermenigvuldigen (of delen) van twee grootheden wordt de *relatieve* onzekerheid in het product (of het quotiënt) door de *relatieve* onzekerheden bij elkaar op te tellen.

65. Doorwerken bij een exacte factor

Hoe zou de formule eruitzien voor de doorwerking van fouten als je een grootheid vermenigvuldigt met een exact getal?

Formule:

$$z = Bx$$

Eigenlijk is deze functie een bijzonder geval van een vermenigvuldiging.

Stel nu eens dat B niet een exact getal was. Dan had B een onzekerheid. Dan zou de relatieve onzekerheid in z gelijk zijn aan de som van de relatieve onzekerheden in B en x. Als B echter een exact getal is zonder onzekerheid, dan is die bijdrage nul. Conclusie: de relatieve onzekerheid in z is dan gelijk aan die in x. En aangezien z precies B keer zo groot is als x, zal δz precies B keer zo groot zijn als δx. Dat is ook heel logisch: als zowel z als δz beide B keer zo groot worden, verandert hun verhouding niet.

Eén ding moeten we nog in de gaten houden: we moeten in de formule wel de absolute waarde van B nemen. Een relatieve en een absolute fout zijn namelijk allebei altijd positief, ook als B negatief zou zijn.

Oefeningen

1. Toon voor het geval $B = 2$ aan dat de formule voor vermenigvuldigen met een exact getal consistent is met de formule voor doorwerken bij optellen.

66. Doorwerken bij machtsverheffen

Dit wordt weer een lekker kort hoofdstuk. Want wat is machtsverheffen méér dan gewoon vermenigvuldigen? Het kwadraat van een getal is dat getal keer zichzelf. Vermenigvuldig je de uitkomst nog een keer met dat getal, dan heb je de derde macht. Als je N dezelfde getallen met elkaar vermenigvuldigt, heb je de N^e macht.

$$z = x^2 = x \times x$$

$$\frac{\delta z}{|z|} = \frac{\delta x}{|x|} + \frac{\delta x}{|x|} = 2\frac{\delta x}{|x|}$$

$$z = x^n = x \times x \times \ldots \times x$$

$$\frac{\delta z}{|z|} = \frac{\delta x}{|x|} + \frac{\delta x}{|x|} + \cdots + \frac{\delta x}{|x|} = |n|\frac{\delta x}{|x|}$$

Je ziet dat van n de absolute waarde wordt genomen. Een negatieve macht betekent een deling, en voor delingen geldt hetzelfde verhaal.

Oefening

1. Beredeneer of deze formule ook geldt voor $n = 0$.
2. Welke absolute onzekerheid is er groter: de absolute onzekerheid in het volume van een bol met een straal van x, of de absolute onzekerheid in het volume van een kubus met een zijde van x?

Beantwoord de vorige vraag opnieuw, maar nu voor een bol met een *diameter* van x en een kubus met een zijde van x.

Weergeven

67. Beste schatting & onzekerheid

Als iemand zegt: “Ik ben Twee Zeven”, wat weet je dan?

Niet veel, maar je kent wel de voor- en achternaam van de heer of mevrouw T. Zeven. Er waren volgens de Nederlandse Familienamenbank in het jaar 2007 in Nederland 72 mensen met de achternaam ‘Zeven’. In de hoofdstad waren er precies zeven Zevens.

Misschien bedoelde degene die zei “Ik ben Twee Zeven” wel iets anders, namelijk “Ik ben twee meter en zeven centimeter lang.”

In het dagelijkse leven doen we dat vaak zo. “Het wordt morgen dertig graden” kan betekenen dat de *verwachting* volgens *het KNMI* is dat het morgen dertig graden *Celsius* wordt. In wetenschap en techniek (en op de hogeschool en op het tentamen en in je verslag en je latere baan) leggen we de lat hoger. Eenheden mag je niet weglaten. Onzekerheden ook niet.

Natuurlijk weet je in de praktijk wel iets van iemand die zegt ‘twee zeven’ te zijn. Dat zijn heus wel meters. En ook als je de onzekerheid niet kent: die persoon kan best wel ‘twee zes’ of ‘twee vijf’ of ‘twee vier’ zijn, maar heus geen ‘één vijfentachtig’. Ook al is de onzekerheid niet vermeld: daar heb je bij een lengtemeting wel een gevoel bij. De onzekerheid is zeker groter dan een millimeter en zeker kleiner dan een decimeter.

Toch is de norm in dit boek strenger.

Als je de eenheid weglaat, is je meetresultaat FOUT.

Soms kan het echte verwarring geven, soms is het alleen maar onhandig en snapt de ander je ook wel. Maar het moet gewoon echt kloppen wat er staat en dat doet het niet zonder die eenheden.

Als je de onzekerheid weglaat, weet je NIETS.

Stel dat de spanning op de pin van een microcontroller doorslaggevend is voor de vraag of het brandalarm moet afgaan. De spanning kan 0 V zijn of 5 V. Maar in de praktijk natuurlijk nooit helemaal exact, en dus wordt alles onder de 2,5 V beschouwd als 0 V en wordt alles boven de 2,5 V beschouwd als 5 V.

Je doet een meting en je meet: ‘Twee zeven’. Dat is fout, want je moet een eenheid opgeven: 2,7 V. Maar hoe zit dat met de onzekerheid? Als die 0,3 V is, zou de werkelijke spanning ook 2,4 V kunnen zijn. En als die 1,4 V is, misschien zelfs maar 1,3 V… Dat is heus wel minder dan 2,5 V. En om het nog

erger te maken: misschien is de onzekerheid wel tien keer zo groot als die 0,3 V. Dan zou de spanning op de pin zelfs negatief kunnen zijn.

We meten dus 'twee zeven' en nemen aan dat het om een hoge spanning gaat. Maar misschien was de werkelijke waarde wel negatief.

Terug naar die militaire lengtemeting uit hoofdstuk 1. Hoe noteer je de meetuitkomst correct? Laten we aannemen dat de onzekerheid anderhalve centimeter is. Dan geef je de *beste schatting* op met daarachter een *plusminusteken* en de onzekerheid, gevolgd door de eenheid.

$$l = (1{,}726 \pm 0{,}015)\ m$$

De haakjes mag je weglaten, omdat het vanzelf spreekt dat de eenheid betrekking heeft op beide getallen.

Volgens de GUM, de *Guide to the expression of uncertainty in measurement*, is de officiële notatie anders.

$$l = 1{,}726\ (15)\ m$$

Het getal tussen ronde haken geeft de onzekerheid aan en moet worden gelezen als de onzekerheid in de laatste cijfers. In dit boek volgen we die meestal niet, omdat die niet gebruikelijk is, al is het dan de officiële standaard.

> Barford (1985) stelt in hoofdstuk 2 van zijn boek dat de modus niet uniek gedefinieerd is. Althans niet in alle gevallen. Als je twee keer 5 V meet en twee keer 7 V, dan zijn er twee waarden die het vaakst voorkomen.
>
> *The mean, however, is always uniquely defined and this is another reason, in the absence of arguments to the contrary, for using this and defining the best estimate, derived from n measurements of equal reliability* x_1, x_1, …, x_n *as the mean,* $X_n = (x_1 + x_1 + \ldots + x_n)/n$
>
> Dat is een zinsnede om over na te denken. We nemen het gemiddelde "bij gebrek aan beter". De mediaan zou ook geen vreemde keuze zijn.

> We hebben het vaak over de *beste schatting* en *de onzekerheid*. Correcter zou zijn: de beste schatting *van de waarde én van de onzekerheid*.

Oefeningen

Als de hierboven genoemde gemeten spanning op het display werd afgelezen als 2,697 V en de onzekerheid was 0,3 V…

1. Hoe noteer je dan de spanning volgens de conventie in dit boek?
2. Hoe noteer je dan de spanning volgens de richtlijnen van de GUM?
3. Noem twee voordelen en twee nadelen van de GUM-notatie.

68. Rekenkundig & statistisch afronden

Er zijn verschillende manieren om een getal af te ronden. Welke methode je kiest, hangt af van het doel. We noemen het laatste cijfer van het afgeronde getal het *relevante cijfer*.

Een gebruikelijke manier van afronden kijkt naar het eerste cijfer na het relevante cijfer. Als dat een 5 of hoger is (een van de cijfers 5 t/m 9), dan wordt het relevante cijfer met 1 verhoogd. Als het lager is dan 5 (een van de cijfers 0 t/m 4), dan blijft het relevante cijfer onveranderd. Zo worden cijfers op schoolrapporten vastgesteld. Een 7,4 wordt een 7. Een 7,5 wordt een 8. Bij lage cijfers is die afronding relatief groot. Wie een 1,4 haalde krijgt een 1. Dat wijkt 29% af van het oorspronkelijke cijfer. Wie een 9,4 ziet afgerond worden op een 9 krijgt maar 4,3% minder. Toch 'voelt' juist dat verschil heel groot. Als je één tiende meer had gehad, was het een 10 op je eindlijst geworden.

Eigenlijk is het vreemd om dit soort voorbeelden met percentages te berekenen. Als je minder dan een 2 hebt, wat doet het cijfer er dan nog toe? Maar aan de 'bovenkant' van de scores is de sores om de afrondingsmores des te groter. Wie een 9,4 had, zat maar 6% van een 10 af. Die afronding maakt dat het 10% wordt: meer dan anderhalf keer zo veel.

Soms rond je alles af *naar boven*. Voor een excursie met 204 studenten wil je bussen huren en er kunnen 60 personen in één voertuig. Dan zou je er 3,4 nodig hebben. Als je dat zou afronden naar beneden, kunnen er 24 studenten niet mee. Daarom moet je afronden naar boven.

Als je projectgroepen wilt maken die uit minstens vijf studenten bestaan — maar liever niet veel meer dan dat — deel je het aantal studenten in de klas door 5 en rond je af *naar beneden* als je wilt uitrekenen hoeveel groepen er komen. Een klas van 32 studenten wordt dan in 6 groepen verdeeld. Je hebt dan vier groepen van vijf studenten en twee groepen van zes.

De 'gewone' manier van afronden heeft een nadeel. De 5 zit namelijk precies tussenin. Als je het getal 7,5 naar 7 afrondt, is het verschil met de oorspronkelijke waarde net zo groot als wanneer je naar 8 afrondt. Voor een 7,6 geldt dat niet: die zit 0,4 punt van de 8 vandaan en meer (0,6) van de 7. Ook voor 7,5000001 geldt dat het verschil met 8 kleiner is dan met 7. Maar met exact 7,5 is het even groot.

Als je uitgaat van een getal met één decimaal dat je wilt afronden op een geheel getal, dan zijn er tien mogelijkheden. Die ene decimaal waar je van af wilt, kan namelijk 0 t/m 9 zijn.

Voor een 0 geldt dat er niets af te ronden valt. Bij een 1 als decimaal raak je 0,1 punt 'kwijt', maar dat wordt gecompenseerd door de 9 als decimaal, want daarbij 'win' je juist 0,1 punt. Zo compenseert de 2 ook de 8 en de 3 de 7, net als de 4 tegenover de 6. Alleen voor de 5 geldt dat — als je die altijd omhoog zou afronden — nooit gecompenseerd wordt.

Bij *statistisch afronden* (of: *afronding naar even*) wordt het cijfer 5 op een aparte manier gebruikt. Je rondt namelijk niet af naar boven, maar naar het *dichtstbijzijnde even cijfer*. Een 5,5 wordt dus een 6, maar een 4,5 wordt een 4. Het voordeel van deze aanpak is dat de verwachtingswaarde van het gemiddelde van afgeronde getallen (uniform verdeeld) gelijk zal zijn aan de verwachtingswaarde van het gemiddelde van niet-afgeronde getallen.

> *Statistisch afronden* is in het Engels *round-to-even* of als je zorgvuldiger bent: *round half to even*. Het wordt in sommige vakgebieden *banker's rounding* of ook wel *Dutch rounding* genoemd.

Hoe groot is het verschil ten opzichte van 'gewoon' afronden? Om dat te berekenen, moet je inzien dat er twintig verschillende mogelijkheden zijn: alle tien cijfers 0 t/m 9 kunnen als eerste cijfer achter de komma staan. Het getal vóór de komma heeft ook tien mogelijkheden, maar het enige wat van belang is, is of er een even of een oneven getal staat. Er zijn tien mogelijke cijfers achter de komma en twee mogelijke pariteiten (even of oneven) van het getal vóór de komma, dus twintig combinaties in totaal. Bij negentien van die twintig is er geen verschil tussen rekenkundig en statistisch afronden; alleen een 5 met een even cijfer ervoor maakt verschil. Die wordt namelijk *omlaag* afgerond in plaats van omhoog. Het verschil in de afgeronde getallen is dan 1: de 4,5 wordt 4 in plaats van 5. Omdat dit in één op de twintig gevallen voorkomt, zal het gemiddelde bij statistisch afronden $1/20 = 0{,}05$ punt lager liggen dan bij rekenkundig afronden. Het gemiddelde van alle getallen is bij statistisch afronden precies gelijk voor onafgeronde en afgeronde getallen.

> Er wordt in boeken verschillend omgegaan met afrondingen. Taylor (1997), Berendsen (2011) en Rabinovich (2005) ronden het cijfer 5 en hoger altijd af naar boven. Hughes & Hase (2010), Squires (2001), Drosg (2007), Kirkup & Frenkel (2006) en Bevington & Robinson (2002) gebruiken *statistisch afronden*.
>
> Grabe kiest ervoor om onzekerheden *altijd* naar boven af te ronden. Niet alleen de 5 dus, maar *alle* cijfers. Een voorbeeld in zijn boek is het getal (met onzekerheid) 119748,8 ± 123,7. Hij noteert dat als 119750 ± 130. In Meten & Weten zouden we kiezen voor $(1{,}1975 \pm 0{,}0012) \times 10^5$. Dat ziet er iets ingewikkelder uit, maar volgt wel de regels voor significante cijfers. Bovendien is er een verschil in afronding: wij maken 12 van 12,37. Taylor (1997) doet het ook zo, net als Hughes & Hase (2010).

Strikt genomen is *statistisch afronden* het meest zuiver. Aan de andere kant: waarom moeilijk doen over (1,45 ± 0,32) A dat eigenlijk zou moeten worden afgerond op (1,4 ± 0,3) A en niet op (1,5 ± 0,3) A, terwijl de onzekerheid in het cijfer achter de komma al zo groot is?

Sommige boeken kiezen zelfs voor (1,5 ± 0,4) A, omdat je de onzekerheid 'voor de zekerheid' naar boven afrondt. Die keuze gaat ervan uit dat de onzekerheid een soort absolute grens aangeeft. Maar als je met onzekerheid bedoelt dat er een redelijk grote *kans* is dat de werkelijke waarde in dat bereik ligt, dan is het zinvol om zo weinig mogelijk af te ronden.

De eerste uitgave van Meten & Weten sprak altijd over 'gewoon afronden' en *banker's rounding*. Oude tentamens van het vak Error Budgeting doen dat ook. Maar er is niets gewoons aan rekenkundig afronden (en ook niets raars aan statistisch afronden) en de term banker's rounding kom je in de metrologie maar weinig tegen. De (meestal Engelse) boeken hebben het vaak over *round-to-even*.

Je komt banker's rounding, bankers rounding en bankers' rounding alle drie tegen trouwens. Zuiver afronden is lastig, correct spellen ook.

Oefeningen

1. Bereken of beredeneer wat het effect zou zijn als *statistisch afronden* zou worden toegepast op tentamencijfers. Wat voor verschil zou dat maken op de gemiddelde cijfers van studenten?

Wat u door 'n domme ezelsbrug te kennen, immer met gemak onthoudt.

Als je het aantal letters van elk van deze twaalf woorden opschrijft en een komma noteert na het eerste cijfer, krijg je pi in elf decimalen: 3,14159265358. Als je het zou afronden, moet die laatste een 9 zijn.

Engelstaligen schijnen onderstaande zin te gebruiken als ezelsbrug.

How I need a drink, alcoholic of course, after the heavy chapters involving quantum mechanics.

2. Je kunt het getal 3,14159265358 afronden op 1 of 2 decimalen, maar ook op 3, 4, 5, 6, 7, 8, 9 of 10. Voor welke van die tien opties geldt dat het uitmaakt of je rekenkundig of statistisch afronden gebruikt?
3. Rond het hexadecimale kommagetal 1A,F84 af op één cijfer achter de komma. Doe dat via rekenkundig én via statistisch afronden.
4. Rond het binaire kommagetal 101,010 af op één cijfer achter de komma. Doe dat via rekenkundig én via statistisch afronden.

69. Significante cijfers

Om metingen met hun onzekerheden goed te kunnen noteren, moeten we weten wat *significante cijfers* zijn. In Vlaanderen noemen ze dat *beduidende* cijfers. Het zijn cijfers die 'van betekenis zijn', en dus iets zeggen over de onzekerheid. Alle cijfers behalve *voorafgaande nullen* (nullen die vanaf links gezien vóór een ander cijfer staan) zijn significant.

Het getal 0 heeft één significant cijfer. Het staat niet vóór een ander cijfer.

Voor nullen 'aan het eind' geldt dat ze altijd significant zijn als ze na de komma staan. Als je geen komma hebt, is er eigenlijk geen duidelijkheid over nullen aan het eind. Je kunt van 1000 zeggen dat het vier significante cijfers heeft, maar één is ook verdedigbaar. Van 1,000 geldt dat het sowieso vier significante cijfers heeft. Anders hadden we wel wat nullen weggelaten.

Tabel 69.1 Voorbeelden van getallen met significante cijfers

getal	niet-nul	significante nullen	significante cijfers	getal grijs en onderstreept
2025	3	1	4	2025
3,14	3	0	3	3,14
0,98	2	0	2	0,98
1,0001	2	3	5	1,0001
0,0060	1	1	2	0,0060
7,50	2	1	3	7,50
0,000001	1	0	1	0,000001
$7{,}20 \times 10^6$	2	1	3	$7{,}20 \times 10^6$
0	0	1	1	0
0,00	0	1	1	0,00
1000	1	0, 1, 2 of 3	1, 2, 3 of 4	1000
$1{,}00 \times 10^3$	1	2	3	$1{,}00 \times 10^3$

De laatste twee regels in deze tabel geven samen een voorbeeld van het geval dat niet duidelijk is. Als iemand het getal '1000' opschrijft, weet je zonder toevoeging niet welke cijfers significant zijn. Wellicht heeft iemand 1,0 kΩ en zijn het tweede en derde cijfer achter de komma helemaal niet bekend. Als je dan liever in ohm rekent, moet je van die 1,0 kΩ toch 1000 Ω maken. Het onderste voorbeeld is de oplossing voor dit probleem als je geen gebruik wilt maken van voorvoegsels als kilo, mega etc.

Er staat in deze tabel ook twee keer 'nul'. Niks dus, noppes. Eén keer in de vorm '0' en één keer in de vorm '0,00'. Hoeveel significante cijfers zijn dat?

Als de Belastingdienst zegt dat je € 0 terugkrijgt, kun je vermoeden dat er centen in je nadeel zijn afgerond. (Wat niet waar is, want er wordt—volgens de regels althans — in je voordeel afgerond.) Als je leest dat je € 0,00 terugkrijgt, gaat het hooguit om tienden van centen.

Als je een spanning meet van 0 V, dan kunnen er nog tienden van een volt (positief of negatief) aanwezig zijn. Die zie je niet op de meter. Maar wat als je 0,000 V meet? Of 0 mV? Dan weet je dat er hooguit een spanning van een paar tienduizendste volt aanwezig is.

Toch is het aantal significante cijfers bij al deze spanningen gelijk aan 1. De metingen verschillen wel in onzekerheid, maar niet in significante cijfers.

Hoeveel significante cijfers heeft 1/7? De vraag klopt niet helemaal, want we zien dat niet als een getal, maar als een sommetje. En hoeveel cijfers je ook gebruikt: het decimale getal gaat nooit exact kloppen.

Tot slot: onze (over)grootouders rekenden nog zonder computer en rekenmachine. De breuk was toen gebruikelijker en gaf een extra mogelijkheid om iets exact weer te geven. Significante cijfers deden dan niet meer ter zake.

Rekenkundig kun je een repeterende breuk noteren door het cijfer of het groepje cijfers dat zich oneindig lang herhaalt te voorzien van een streep erboven. Wat je ook wel ziet, is haken eromheen. Decimaal én exact dus.

$$\frac{1}{7} = 0,\overline{142857} = 0,[142857] \approx 0,142857142857142857142857$$

Oefeningen

Noem het aantal significante cijfers in onderstaande waarden of grootheden.

1. 0 A
2. 1
3. 0,1 A
4. 0,10 mA
5. 0,0000000000000000000010
6. € 1,11
7. € 0,00
8. 0,01100
9. 1000000
10. $1{,}0 \times 10^6$
11. 1 000 000,00

70. Significante cijfers in berekeningen

Wat moet je doen met significante cijfers in berekeningen? In tussentijdse resultaten laat je in principe alle cijfers staan. Je neemt er in elk geval zoveel mogelijk van mee in je berekening. Maar wat doe je met het eindresultaat?

Stel je wilt een snelheid meten. Je meet hoe ver een voorwerp zich verplaatst en komt op 2,235 m. Een extra cijfer aflezen lukt niet, maar die derde decimaal klopt wel. Het aantal significante cijfers is dus vier. Je meet de tijd die het voorwerp nodig heeft voor de verplaatsing en komt op 2,5 s. Je gebruikte tijdmeting staat niet toe om nog een cijfer achter de komma te zetten. Je hebt dus twee significante cijfers. Nu kun je de snelheid uitrekenen.

$$v = \frac{d}{t} = \frac{2{,}235\ m}{2{,}5\ s} = 0{,}894\ m/s$$

In dit geval levert een deling van een grootheid met vier significante cijfers door een grootheid met twee significante cijfers een resultaat op met drie decimalen. Maar als je 2,4 s had gemeten, zou er 0,93125 m/s uitgekomen zijn. En als je in milliseconden had kunnen meten en je was op 2,235 s uitgekomen, dan was je antwoord 1 m/s geweest: één cijfer dus.

Los van het feit dat het aantal cijfers in het antwoord niet telkens overeenkomt: het is een wat vreemde gedachte dat je door te delen door iets met weinig significante cijfers er net zo veel of meer overhoudt. Als je niet 'beter' dan in hele seconden had kunnen meten, zou je door 2 s hebben gedeeld. Dan was de uitkomst 1,1175 m/s geweest. Maar liefst vijf significante cijfers.

Hoe goed ken je de waarden waarmee je rekent eigenlijk? Ervan uitgaande dat alle significante cijfers een correcte waarde hebben, is de echte afstand minstens 2,2345 m en hoogstens 2,23549999 m. Anders zou je niet op deze cijfers hebben afgerond. Voor de gemeten 2,5 s geldt dat het minimaal 2,45 s was en maximaal 2,549999 s. (Na die vier negens nog oneindig veel.)

Je zou kunnen zeggen dat de relatieve onzekerheid in de afstand gelijk is aan:

$$\frac{\delta d}{d} = \frac{0{,}0005\ m}{2{,}235\ m} = 0{,}02237\%$$

De relatieve onzekerheid in de tijd is gelijk aan:

$$\frac{\delta t}{t} = \frac{0{,}05\ s}{2{,}5\ s} = 2\%$$

Kirkup & Frenkel (2006) geven als vuistregel voor het aantal significante cijfers dat je moet kiezen voor een aantal dat zorgt dat de relatieve onzekerheid zo dicht mogelijk in de buurt van de grootste van die twee relatieve onzekerheden komt te liggen. Als je opties in een tabel weergeeft, kom je dan op deze waarden.

Waarde	Onzekerheid	relatief
0,894 m/s	0,0005 m /s	0,056%
0,89 m/s	0,005 m/s	0.56%
0,9 m/s	0,05 m/s	5,6%

De auteurs geven dezelfde aanpak voor optellen en aftrekken.

Rule 3: When numbers are added or subtracted, quote the answer to the number of significant digits that implies a proportional uncertainty closest to the greater of the component proportional uncertainties.

Ze gebruiken als rekenvoorbeeld het optellen van de massa van een koperen vat met water (1,5778 kg) en de massa van alleen het vat (0,582 kg). Ze komen dan op percentages van 0,003% en 0,05%.

Wat nu als je een olifant hebt gewogen met een muis op zijn rug? Je meet dat ze samen 4162 kg wegen. De muis kun je eraf halen en daarna wegen op een andere weegschaal. Je stelt vast dat hij een massa heeft van 19 g. De relatieve onzekerheden in de metingen zijn 0,5 / 4162 ≈ 0,012 % en 0,5 / 19 = 2,6%. Dat zou je dus volgens Kirkup & Frenkel van die 2,6% moeten uitgaan…

De olifant zonder de muis weegt 4162 – 0,019 = 4161,981 kg. Met een onzekerheid van 2,6% zou de onzekerheid in de massa van de olifant alleen dus 110 kg zijn. Hier lijkt iets niet te kloppen: je gezonde verstand zegt dat de massa van de muis te verwaarlozen is.

Een veel eenvoudigere — en vaak gebruikte — vuistregel is:

- Gebruik bij optellen en aftrekken het kleinste aantal *cijfers achter de komma.*
- Gebruik bij vermenigvuldigen en delen het kleinste aantal *significante cijfers.*

Bij het voorbeeld van de olifant en de muis had de eerste nul cijfers achter de komma en de tweede drie. Die 4161,981 kg rond je dus af op 4162 kg. Wat we al dachten: de massa van de muis is verwaarloosbaar.

Bij het voorbeeld van de snelheid hadden we twee significante cijfers in de tijd en vier in de afstand. Volgens de simpele vuistregel kom je dan op twee significante cijfers. Die 0,894 m/s moet dus worden genoteerd als 0,89 m/s.

Wat nu als we in milliseconden hadden gemeten en die 2,235 s hadden gevonden? Zou die 1 m/s dan nog steeds 1 m/s zijn? Ja, maar wel met meer decimalen. Als tijd en afstand beide in vier significante cijfers waren opgegeven, zou het dus 1,000 m/s moeten zijn. Let op: dat zijn *drie* decimalen en *vier* significante cijfers.

Oefeningen

Bereken onderstaande sommetjes en geef de antwoorden in het juiste aantal significante cijfers.

1. 456 + 3,5
2. 100 – 99,9
3. 3,1415 × 100
4. 12,0 / 2,54
5. 7,21 × 8,66 × 2
6. $1{,}2345^2$
7. de lichtsnelheid in kilometer per uur
8. een versnelling van 9,87654321 m/s^2, maar dan in *miles per square quarter*.

Huh? Wat een rare vraag, die achtste… Mag je die ook overslaan? Ja hoor. Maar het is wel een grappige oefening. Rare vragen zijn soms een handige manier om te kijken of je het echt snapt.

71. Significante cijfers in onzekerheden

Hoeveel significante cijfers wil je gebruiken voor een onzekerheid? Stel dat je de versnelling van de zwaartekracht hebt gemeten en je komt uit op

$$g = (9{,}81643552 \pm 0{,}03244913)\ m/s^2$$

Hoeveel zin heeft het dan om acht cijfers achter de komma te geven?
De waarde ligt in Nederland overigens tussen de 9,78 en 9,84 m/s^2.
Daarom zeggen we meestal 9,8 m/s^2.

> "Ik haw de fersnelling fan de swiertekrêft mjitten en kaam út op njoggen komma acht", zeggen ze in Friesland. "Ich höb de versnelking van de zwaartekrach gemete en kwam oet op nege komma acht", zeggen ze in het zuiden van Limburg. Of je nou njoggen of nege zegt, maakt niet zoveel uit. Maar wat je in heel Nederland zou moeten doen, is de eenheden vermelden. *Altied en euveral.*

Je zou ook kunnen uitrekenen wat de relatieve fout in g is. Als je bovenstaande twee getallen op elkaar deelt, kom je op 0,33%. Neem je de onzekerheid met één significant cijfer, dan kom je op 0,3%. Als je naar die twee getallen kijkt (0,33% en 0,30%), lijkt dat een relevant verschil: die 0,03% (procentpunt) is namelijk 10% (procent) van 0,30%. Maar heeft het nu zin om te zeggen dat de waarde ligt tussen 9,784 en 9,848 m/s^2, in plaats van tussen 9,78 en 9,84 m/s^2 ? Of de absolute onzekerheid nu 0,064 of 0,06 m/s^2, maakt niet veel verschil. Het gaat om een indicatie, een indruk van hoe groot het gebied is waarin de waarde waarschijnlijk ligt.

Wat zegt de GUM, de Enige Echte Norm, over het aantal decimalen?

7.2.6 The numerical values of the estimate y and its standard uncertainty $u_c(y)$ or expanded uncertainty U should not be given with an excessive number of digits. It usually suffices to quote $u_c(y)$ and U [as well as the standard uncertainties $u(x_i)$ of the input estimates x_i] to at most two significant digits, although in some cases it may be necessary to retain additional digits to avoid round-off errors in subsequent calculations. In reporting final results, it may sometimes be appropriate to round uncertainties up rather than to the nearest digit.

Geen 'excessief aantal', niet te veel dus. 'Het is meestal voldoende' om hooguit twee significante cijfers te geven. En 'soms' is het beter om omhoog af te ronden, in plaats van op het dichtstbijzijnde cijfer. Best vaag, voor een norm.

Eén ding is wel duidelijk: het aantal cijfers achter de komma in de beste schatting in de onzekerheid moet met elkaar overeenkomen.

In Meten & Weten hanteren we een strakke regel. We geven onzekerheden op met één significant cijfer, tenzij dat ene cijfer een 1 is. Dan voegen we een tweede significant cijfer toe.

> Deze notatie wordt gebruikt door Hughes & Hase (2010) en door Taylor (1997), maar die zijn niet helemaal 'dwingend'. Fornasini (2008) zegt dat het er nooit meer dan twee moeten zijn en dat één soms al voldoende is. Volgens Kirkup & Frenkel (2006) is het *ongebruikelijk* om meer dan twee significante cijfers op te geven bij een onzekerheid. Grabe (2014) kiest ervoor om zowel bij een 1 als bij een 2 een extra cijfer op te geven, terwijl Squires (2001) en Morris & Langari (2020) datzelfde in overweging geven. Rabinovich (2005) wil ook twee cijfers als er 3 staat. Merrin (2017) gebruikt er altijd twee. Berendsen (2011) gebruikt meestal één significant cijfer en raadt aan om naar boven af te ronden. Hij vindt meer cijfers acceptabel als je voldoende metingen hebt en een toevallige fout die je nauwkeurig genoeg kent. Is hier dan geen norm voor? Nee. Van de zeven boeken die we hierboven aanhalen, doen Hughes & Hase (2010) en Taylor (1997) precies hetzelfde. De andere boeken zijn daar niet strijdig mee, in die zin dat ze meestal pleiten voor weinig cijfers.
>
> Drosg (2007) noemt geen harde één of twee, maar zegt wel iets aardigs.
>
> *A good rule is to determine an uncertainty in such a way that one additional decimal digit (when compared to the data value) is necessary for the (implicit) presentation of it. If done so the data value will have one additional decimal digit, which is insignificant.*
>
> Je geeft dus eigenlijk een cijfer te veel op, maar dat maak je dan ook duidelijk in je onzekerheid.

Dat je bij een 1 een extra cijfer wilt opgeven, heeft te maken met de relatieve invloed van de afrondingen. Stel je voor dat je een lengtemeting hebt met een onzekerheid van 0,0149952 m… Als je dat noteert als 0,01 m, dan heb je in feite 15 mm afgerond naar 10 mm. De berekende onzekerheid is dan maar liefst 50% meer dan wat je opgeeft. Dat probleem zou te verhelpen zijn door naar boven af te ronden op 2, maar dan is het nog steeds 25%. Bij hogere cijfers is dat relatieve verschil kleiner. Overigens kan een 2 ook een afronding zijn van 2,499 en dan zit je ook met die 25%. Toch volgen we in dit boek de lijn van Taylor en van Hughes & Hase.

Natuurconstanten worden in de literatuur vrijwel altijd met twee significante cijfers in hun onzekerheden opgegeven, ook als de eerste geen 1 is. Nu zijn de metingen daarvan vaak zo extreem goed, dat twee cijfers in de onzekerheid ook niet heel vreemd zijn.

Stappenplan

Hoe krijg je de correcte notatie voor de beste schatting en de onzekerheid?

1. Zorg dat de eenheden en voorvoegsels met elkaar kloppen.
2. Maak eventuele machten van tien aan elkaar gelijk.
    - Haal de macht van tien eventueel buiten haakjes
3. Kies het juiste aantal significante cijfers in de onzekerheid:
    a. twee cijfers als het eerste cijfer een 1 is,
    b. één cijfer in alle andere gevallen.
4. Rond de onzekerheid correct af op het juiste aantal cijfers.
5. Pas het aantal significante cijfers in de beste schatting zodanig aan dat het laatste cijfer op dezelfde positie staat.

Er hoeven geen haakjes te staan om beste schatting en onzekerheid samen te nemen. Dat moet wel voor eventuele van machten van tien.

Oefeningen

Noteer de volgende onzekerheden met het juiste aantal significante cijfers.

1. 5,94 s
2. 14,56 m
3. $1{,}55 \times 10^6$ kg
4. 2,51 A
5. 1,91 K
6. 0,1999 V

72. Noteren van onzekerheden

Je hebt een meting pas goed genoteerd als de volgende punten correct zijn.

1. de eenheid van de beste schatting en de onzekerheid
2. de waarde van de onzekerheid in het juiste aantal cijfers
3. de waarde van de beste schatting in het juiste aantal cijfers
4. eventuele machten van 10 en voorvoegsels zoals 'milli'
5. een punt of een komma op de juiste plaats
6. eventuele spaties om de waarde duidelijker te maken
7. eventuele haakjes om het geheel duidelijker te maken

Bewust is in dit lijstje de eenheid op nummer 1 gezet. Zonder een eenheid heb je namelijk niets, ook al kun je het in de praktijk vaak zelf wel invullen. Als iemand zegt dat hij of zij een houten ondergrond van 'twee bij twee' gaat aanleggen onder een meetopstelling, dan zullen dat heus wel meters zijn. Als dezelfde persoon zegt: "Ik ben 34", dan is het vast wel een leeftijd in jaren. "Ik ben één zesentachtig" is meestal een lengte in meters (aantal: 1) en centimeters (aantal: 86).

Met stip op 2: de waarde van de onzekerheid. Ook daarvoor geldt: als je die niet kent, weet je ook weinig. "Ik heb een voorlopig cijfer voor je tentamen: je hebt een 6,2." "Joepie!" "En de onzekerheid in het cijfer is 3,5 punt."

Er is nog een reden om met de onzekerheid te beginnen. Het juiste aantal cijfers daarin is namelijk bepalend voor de notatie van de beste schatting. Dat is ook wel logisch: de onzekerheid zegt namelijk iets over de waarde ervoor, namelijk hoe goed je die kent. Het omgekeerde is niet het geval.

Stel dat je een multimeter hebt met te veel cijfers achter de komma. Je meet daarmee een weerstand en komt uit op de volgende waarde.

$$R = (122{,}1437668 \pm 0{,}0034220575623)\ \Omega$$

Hoe ga je dit in de juiste vorm schrijven? Begin met de eenheden (die kloppen al) en de onzekerheid. Die noteer je met één significant cijfer.

$$\delta R = 0{,}003\ \Omega$$

Let op: de onzekerheid in de gemeten weerstand is hier dus in *één significant cijfer* en in *drie decimalen*. Om de beste schatting daarmee te laten kloppen, moet die waarde ook in drie decimalen worden genoteerd. We krijgen dus:

$$R = (122{,}144 \pm 0{,}003)\ \Omega$$

Het derde cijfer achter de komma wordt een 4, omdat er na de 3 een 7 stond.

Is de weerstand zelf nu ook in één significant cijfer gegeven? Nee, de weerstand zelf heeft zes significante cijfers. Als de onzekerheid tien keer zo klein was geweest (0,0003 Ω dus), waren dat zelfs zeven cijfers geweest. De beste schatting had dan namelijk als 122,1438 Ω moeten worden genoteerd. En was de onzekerheid honderd keer zo groot (0,3 Ω dus), dan bleven er vier significante cijfers over. Drie voor de komma en één erachter.

Stel nu dat het eerste significante cijfer van de onzekerheid in de weerstand niet een 3 maar een 1 was geweest, en de andere cijfers hetzelfde; wat zou dan het resultaat zijn geworden? De onzekerheid zou in één extra decimaal worden gegeven (twee significante cijfers) en de beste schatting dus ook. De afronding pakt dan wel anders uit. In dat geval komen we op:

$$R = (122{,}1438 \pm 0{,}0014)\ \Omega$$

Dit zijn twee voorbeelden met nogal kleine onzekerheden. Stel nu eens dat de onzekerheid veel groter was.

$$\delta R = 10{,}00\ \Omega$$

We moeten dan de onzekerheid in twee significante cijfers schrijven, omdat het eerste cijfer in δR een 1 is. Dan blijft er een waarde over zonder cijfers achter de komma. Voor de beste schatting geldt dan hetzelfde.

$$R = (122 \pm 10)\ \Omega$$

Maar wat nu als de onzekerheid nog drie keer zo groot was? Het is verleidelijk om dan in plaats van 10 gewoon 30 te schrijven, maar dat mag niet: een 3 is geen 1, en moet dus met één significant cijfer worden genoteerd. Een oplossing die je in sommige boeken wel ziet, is dat de laatste 2 van 122 wordt weggelaten. Of beter gezegd: omgezet in een 0.

$$R = (120 \pm 30)\ \Omega$$

Het nadeel hiervan is dat je volgens de regels van significantie nog steeds te veel cijfers hebt. We kunnen dit oplossen door met machten van 10 te werken, of te kiezen voor een voorvoegsel in de eenheid.

$$R = (0{,}12 \pm 0{,}03)\ \mathrm{k\Omega}$$

Wie niet van onnodige komma's houdt, wordt hier niet blij van. Toch heeft deze notatie een voordeel: de kΩ maakt duidelijk dat je het zo heel zeker niet weet.

We gaan nu een spanning meten met de multimeter. Dit is het meetresultaat.

$$V = 3{,}7502\ V \pm 42\ mV$$

De eenheden kloppen nu dimensioneel wel met elkaar, maar het voorvoegsel m (milli, een duizendste) maakt dat de getallen niet met elkaar overeenkomen. Er moet dan gekozen worden tussen volt en millivolt. Vanwege de grootte van de beste schatting ligt de eenheid volt hier het meest voor de hand. Om fouten te voorkomen, voeren we die eerste stap apart uit.

$$V = 3{,}7502\ V \pm 0{,}042\ V$$

Het lijkt een simpele zet, maar je gaat met het verschuiven van de komma heel snel de mist in. Van millivolt naar volt is een factor 1000. Je moet de komma dus drie plaatsen naar links zetten. Daarna is het nog een kwestie van het juiste aantal significante cijfers kiezen en de haakjes om de getallen.

$$V = (3{,}75 \pm 0{,}04)\ V$$

Naast voorvoegsels kunnen er ook machten van 10 zijn die maken dat de beste schatting en de onzekerheid nog moeten worden 'uitgelijnd'. Dat zie je in onderstaand voorbeeld.

$$l = 0{,}01773452 \cdot 10^{-2}\ mm\ \pm 6{,}222 \cdot 10^{-3} \mu m$$

Een raar geval, maar als oefening des te aardiger. Waar begin je mee? De eerste stap zou kunnen zijn om die machten van 10 weg te werken. Dan krijg je veel cijfers achter de komma, maar dat komt daarna wel goed.

$$l = 0{,}0001773452\ mm\ \pm 0{,}006222\ \mu m$$

Om ze aan elkaar gelijk te maken, moet er worden gekozen tussen mm en µm. Natuurlijk zou je een ander voorvoegsel kunnen kiezen, maar als je dat doet, maak je twee stappen tegelijk en dat vergroot de kans op fouten. Gegeven de beide getallen liggen micrometers het meest voor de hand. De beste schatting moet met een factor 1000 worden vermenigvuldigd. De komma gaat dus drie plaatsen naar rechts.

$$l = 0{,}1773452\ \mu m\ \pm 0{,}006222\ \mu m$$

$$l = (177{,}3452\ \pm 6{,}222)\ nm$$

De laatste stap is het correct noteren van het aantal *significante cijfers* in de *onzekerheid* en het daarmee laten overeenkomen van het aantal *decimalen* in de *beste schatting*. Let in die stap goed op de afronding: in dit geval omlaag, omdat er na de komma een 3 staat.

$$l = (177 \pm 6)\ nm$$

Op toetsen van het vak *Error Budgeting* in de opleiding Mechatronica van de Haagse Hogeschool (locatie Delft) wordt al jarenlang dezelfde vraag 1 gesteld. Die vraag luidt als volgt.

Herschrijf de volgende (berekende) grootheden zodanig dat het aantal significante cijfers klopt met de regels zoals die gegeven zijn in het vak Error Budgeting.

Studenten krijgen vervolgens vijf metingen en onzekerheden die ze op de correcte manier moeten opschrijven. De vraag is één punt waard en per fout gaat er een halve punt af. Als je twee of meer (van de vijf) vragen fout doet, heb je dus nul punten… Streng? Misschien. Maar studenten weten van tevoren dat ze deze vraag krijgen! En de regels zijn gegeven en uitgelegd.

Oefeningen

Herschrijf de volgende (berekende) grootheden zodanig dat het aantal significante cijfers klopt met de regels zoals die gegeven zijn in dit hoofdstuk.

1. lengte = 1,71 m ± 1½ cm
2. breedte = 200 mm ± een half inch
3. oppervlakte = 60,55 m^2 ± 450 cm^2
4. volume = 40,3020 cl ± 10 cc
5. tijd = 59,9876 s ± 5 ms
6. massa = 72,55 kg ± 391 g
7. hoek = 3,1415279 ± 0,025 rad
8. volume = 98,357 cm^3 ± 246 mm^3
9. frequentie = 50 Hz ± 5%
10. temperatuur = 20 °C ± 2 °F
11. elektrische stroom = 1,101011001 A ± 1,49 mA
12. volume = 0,44 × 10^{-2} mm^3 ± 7,1 × 10^{12} nm^3
13. spanning = 230,00000500025 V ± 45 µV
14. elektrische stroomdichtheid = 0,04 A/m^2 ± 400 mA/mm^2
15. warmte = 6033 kJ ± 4499 × 10^2 J
16. oppervlakte = 7659 × 10^{-6} μm^2 ± 8,834 × 10^{-1} nm^2
17. kracht = 99,9 kN ± 0,99 × 10^4 N
18. vermogen = 0,22 kW ± 33 W
19. druk = 200 Pa ± 3 N/dm^2
20. weerstand = 2 × 10^4 Ω ± 0,3 kΩ

73. De Grote Negen

Iemand heeft een slingerproef gedaan en heeft daarmee geprobeerd om de versnelling van de zwaartekracht op een bepaalde plaats te bepalen.

“Zo dat is aardig gelukt! Ik kom uit op 9.876…”

Au. Hier gaan wat dingen mis.

De berekening resulteerde in 9,87654321 m/s^2. Wat toevallig! Nee hoor. Het voorbeeld is verzonnen. Het zou erg toevallig zijn als je bij een echte meting op precies alle tien cijfers van het decimale talstelsel uitkomt. Maar in boeken en films gebeuren wel vaker toevallige dingen.

Nee, het punt is *die punt*. In het Nederlands gebruik je een *komma*. (Fout 1)

Maar ook de getalswaarde klopt niet: als je 9,87654321 wilt afronden op drie decimalen, dan kom je uit op 9,877. (Fout 2)

Nu kan een getalswaarde zonder eenheid natuurlijk nooit de versnelling van de zwaartekracht zijn.

“Ja, die liet ik voor het gemak even weg.”
“Wat was je aan het meten dan?”
“De Gravitatieconstante. Dat is hoe sterk de zwaartekracht is. Dus als jij zo nodig eenheden erbij wilt: het was in Newtons.”

Au. Ook hier gaan wat dingen mis. En dan bedoel ik nog niet eens die hoofdletter bij de gravitatieconstante. Dat is een natuurconstante die wel te maken heeft met de zwaartekracht, maar is iets anders dan wat we willen meten: de versnelling van de zwaartekracht op een bepaalde plaats.

De eenheid is niet *newton* (met een kleine letter, dus Fout 3a en enkelvoud Fout 3b), maar m/s^2. Als je de eenheid weglaat (Fout 4), zeg je eigenlijk niets. Of iets verkeerds.

De persoon die het experiment uitvoerde is zo eigenwijs om geen komma te gebruiken in een Nederlandse tekst en probeert dat daarom nu op te lossen door een macht van tien te gebruiken.

$$g = 9877 \pm 25 \cdot 10^{-3}\ m/s^2$$

Leuk geprobeerd, maar als je het zo opschrijft, heeft die macht van tien alleen maar betrekking op de *onzekerheid*. Hier moeten toch echt haakjes omheen. (Fout 5) Tenzij je de macht van tien gewoon niet gebruikt. Je hoeft geen haakjes te zetten vanwege die eenheid: die hoort namelijk

vanzelfsprekend bij *beide*. Een absolute onzekerheid heeft altijd dezelfde eenheden als de meetwaarde zelf.

$$g = 9{,}88 \pm 0{,}25\ m/s^2$$

En hier gaat wéér iets mis. Je maakt heel snel fouten bij het omschrijven van machten en factoren 10 en milli-, kilo- en nano-dingen. De getalswaarde van de meting is door 1000 gedeeld en de onzekerheid door 100. (Fout 6)

$$g = 9{,}877 \pm 0{,}025 m/s^2$$

Zucht. Het gaat de goede kant op, maar tussen een waarde en een eenheid hoort toch echt een *spatie*. (Fout 7) En je geeft onzekerheden op met één significant cijfer, tenzij dat cijfer een 1 is: dan geef je er twee. (Fout 8)

$$G = 9{,}87 \pm 0{,}02\ m/s^2$$

Soms heb je zo'n dag dat álles misgaat: beide afrondingen zijn verkeerd gedaan. (Fout 9 en 10) Bovendien staat er opeens een hoofdletter G, maar die betekent iets anders (Fout 11).

$$g = 9{,}88 \pm 0{,}03\ m/s^2$$

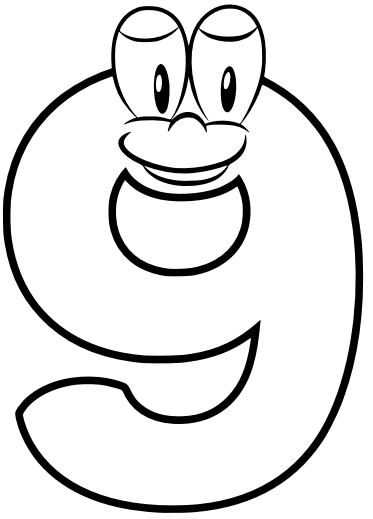
En nu zijn we er bijna! Deze notatie klopt wel. Er is alleen nog één probleem. Met een zo kleine onzekerheid is het verschil tussen de gevonden waarde en de literatuur meer dan twee keer de onzekerheid. Het "Zo dat is aardig gelukt" was een beetje te optimistisch. Er is namelijk een strijdigheid gevonden. En dus een verkeerde conclusie getrokken (Fout 12).

Het zou al extreem zijn om precies 9,87654321 te meten, maar nóg extremer is dat iemand daarbij *twaalf* fouten maakt. Maar misschien zijn er ook nog wel fouten die we niet gezien hebben. De slingerproef zelf kan verkeerd uitgevoerd zijn, waardoor de getalswaarde niet klopte. Er kan een rekenfout in hebben gezeten. De opgegeven waarde van de onzekerheid is wellicht te gunstig ingeschat.

Er moet wel nog een fout zijn, want de versnelling van de zwaartekracht ligt heus niet in het opgegeven bereik.

Twaalf fouten is wat overdreven, maar dit extreme voorbeeld laat wel zien hoe veel er mis kan gaan bij alleen al het noteren van de beste schatting en de onzekerheid. In dit boek noemen we al de mogelijke foutoorzaken samen *De Grote Negen*.

1. **eenheid**
   Vermeld de juiste SI-eenheden (en dus ook de correcte hoofdletters).
2. **beste schatting**
   Als je geen waarde opgeeft, heb je geen meetresultaat.
3. **onzekerheid**
   Laat de onzekerheid niet weg, anders kun je geen conclusies trekken.
4. **voorvoegsel**
   Gebruik voorvoegsels als *milli* en *kilo* (met correcte hoofdletters).
5. **factor**
   Ga correct om met machten van 10.
6. **significantie**
   Geef de onzekerheid in het juiste aantal significante cijfers en de beste schatting met hetzelfde aantal decimalen.
7. **afronding**
   Rond het cijfer 5 en hoger af naar boven en het cijfer 4 of lager naar beneden.
8. **haakjes**
   Zet haakjes om getallen als dat nodig is voor de rekenregels.
9. **komma**
   Gebruik komma's in Nederlandse teksten (*and points in English*).
0. **spaties**
   Zet een spatie tussen een getal en de eenheid. (Deze tiende van *De Grote Negen* heeft nummer 0, omdat nul niks is. Net als een spatie.)



Dit rijtje is zo belangrijk dat er nog wel een extra toelichting bij mag.

- De **eenheid** is het belangrijkste van alles. De lichtsterkte van een lamp druk je uit in lumen. Het vermogen ervan in watt. De massa in kilogram. Wie geen eenheid vermeldt, vertelt niet eens wat er gemeten wordt. "Een gloeilamp van 40 wát? Van 40 watt? Of van 40 lumen? Dat scheelt nogal wat…"
- De **beste schatting** en de **onzekerheid** erop zijn de getallen die aangeven wat je gemeten hebt. De kale cijfers dus, die allebei een keiharde onzekerheid hebben.
- Het **voorvoegsel** en de **factor** zijn twee verschillende dingen, maar ze hebben veel met elkaar te maken. Als je van 10^{-4} cm^3 naar 10^{-6} m^3 gaat, heb je met twee verschillende factoren (machten van 10) te maken, maar ook met een voorvoegsel 'centi' dat bij de ene term wel staat en bij de andere niet.

- De **significantie** in de beste schatting hangt af van de significantie in de onzekerheid. Niet *het aantal significante cijfers* is gelijk, maar *het aantal decimalen.*
- De **afronding** komt pas na het vaststellen van de significantie en kan zowel bij de beste schatting als bij de onzekerheid fout zijn: de beste schatting kan verkeerd afgerond zijn en de onzekerheid ook. (Eigenlijk twee mogelijke fouten… De Grote Negen is een tiental.)
- **Haakjes** hoef je niet te zetten om het getallenpaar dat de beste schatting en de onzekerheid aanduidt, maar zijn wel belangrijk als je met factoren 10 werkt.
- Een **komma** in een Engelse tekst zetten of een punt in een Nederlandse: het is allebei onjuist als je het decimale scheidingsteken bedoelde. (Deze staat wat laag in het lijstje, omdat de fout bewust of onbewust door de lezer genegeerd wordt.)
- De **spatie** is heus wel verplicht tussen getal en eenheid, maar net niet afkeurenswaardig genoeg om hem tot *De Grote Negen* te rekenen. Als ik het heb over een afstand van 1609km, dan weet iedereen wat ik bedoel. Bij 1.609 km ligt dat anders.
  What do you mean? One mile or a thousand miles?

Oefeningen

Een nieuwe meting komt uit op een onafgeronde waarde van 9,84452 m/s^2 en een onzekerheid van 0,17622 m/s^2. Noem de fouten in de notaties hieronder.

1. 9,8 ± 0,2 m/s^2
2. (9,84 ± 0,17) m/s^2
3. 9,84 ± 0,176 m/s^2
4. 9,85 ± 0,18m/s^2
5. 9.84 ± 0.18 m/s^2
6. (9,8 ± 0,18) m/s^2
7. 9.844

8. Leid uit onderstaande formule af (zonder ergens op te zoeken dus!) wat de eenheid is voor de Gravitatieconstante G.

$$F = G\frac{m_1 \times m_2}{r^2}$$

74. Maak er maar geen punt van

Moet je in de uitkomst van 2 gedeeld door 5 een punt of een komma zetten?

Is het 0.4 of 0,4? Mijn rekenmachine zei 0.4…

> Even voor alle duidelijkheid: 2 door 5 delen lukt me ook wel zonder rekenmachine. Ik gebruikte hem alleen maar om te kijken of er een punt of een komma zou verschijnen. Echt waar!

Mijn rekenmachine is beter in rekenen dan in taal. Helaas heb ik geen taalmachine. Dan maar met Google aan de slag. Scribbr, Taaladvies.net en Wikipedia zeggen allemaal stellig dat je in het Nederlands een komma gebruikt als decimaalteken, terwijl Engelstaligen een punt gebruiken. Onze Taal lijkt een klein beetje ruimte te laten om in Nederlands het net als in het Engels te doen. *In een getal met een of meer decimale cijfers wordt in het Nederlands traditioneel een komma gebruikt: 36,8 °C, 1,98 m, 3,5 miljoen. De punt komt onder invloed van het Engels wel steeds meer voor.*

Toch schrijven het *Bureau voor Normalisatie (NBN*) uit Vlaanderen en het *Nederlands Normalisatie-instituut (NEN)* een komma als decimaalteken voor, evenals een spatie om de cijfers in groepjes van drie van elkaar te scheiden.

De Nederlandstalige versie van Excel en de Rekenmachine van Windows vertonen ook een komma in getallen.

Berendsen besteedt in zijn boek aandacht aan die schrijfwijze, in de Nederlandse als in de Engelse uitgave. Terecht merkt hij op dat die punt in het Engels juist een uitzondering is. Naast China, Japan, Israel en Zwitserland zijn er niet veel niet-Engelstalige landen waar ze ook een punt gebruiken. In veruit de meeste Europese talen en landen is het de komma die als decimaalscheiding dient.

Het punt met die komma is dat hij in de praktijk problemen geeft. In het Nederlands is de komma het *decimaalteken*, maar óók het *scheidingsteken* in opsommingen. En daar wringt het. Neem deze zin:

De metingen waren 1,71, 2,5, 3, 6,7, 8, 9 en 4,1.

Hoeveel getallen staan hier? In het Engels zou je schrijven:

The measurements were 1.71, 2.5, 3, 6.7, 8, 9, and 4.1.

In die taal heeft de punt een dubbelrol, maar is niet verwarrend.

> Zag je dat er in het Engels een extra komma stond? Vóór het woord '*and*'. Je noemt dat een *Oxford comma* (NL: Oxfordkomma) of *serial comma*. In het Nederlands doen we dat niet. Nou ja, soms wel, als het verwarring voorkomt.

Voor wie correct Nederlands wil gebruiken, is er wel een oplossing: gebruik witruimte, óf zet je waarden in een nette tabel. Voor grote getallen gebruik je liever een spatie als duizendtalscheiding, zodat die punt geen kwaad kan.

Fijnproevers gebruiken dan een *smalle spatie*.
(In *Word*: tik de Unicode 2009 in en druk op Alt+X.)

Zodra je met software werkt, wordt het lastiger. Op je numerieke toetsenbord zit een punt, maar in Nederland verschijnt daar vaak een komma. Dat komt doordat Windows kijkt naar je taalinstellingen (*Tijd en taal > Regio > Extra instellingen*). Daar kun je zelf het decimaalteken kiezen.

In Excel komt daar nog een sausje eigenwijsheid bij: Excel gebruikt z'n eigen logica, afhankelijk van je Windowsinstellingen, je Office-versie én het bestandsformaat. Het gevolg: formules doen het ineens niet meer, of getallen worden als tekst gezien.

Programmeertalen als C, Python, Java en JavaScript zijn daar eenvoudiger in: die gebruiken altijd een punt in de code én de in- en uitvoer, omdat die talen op het Engels gebaseerd zijn.

Hoe doe je dat als Windows-programmeur? Dit zijn de drie smaken.

Teken	Symbool	Unicode	ASCII	Toetscode (Windows)
Punt	.	U+002E	46	VK_OEM_PERIOD
Komma	,	U+002C	44	VK_OEM_COMMA
Numpad (.)	. of ,*	wisselend	46 of 44	VK_DECIMAL

De toets op je NumPad levert altijd VK_DECIMAL, maar wat er op het scherm verschijnt hangt af van je Windows-instellingen. In het Nederlands meestal een komma, in het Engels een punt.

75. Cijfers in letters

Als je niet goed bent in Nederlands en er beter in wilt worden, zijn er een paar eenvoudige tips.

1. Raadpleeg bronnen als *Onze Taal* en stel vragen aan ChatGPT.
2. Geef veel aandacht aan werkwoordspelling. Oefen erin.
3. Schrijf samenstellingen aan elkaar (en niet los, zoals in het Engels).
4. Maak niet de bekende fouten in ‘te veel’ en ‘door middel van’.
5. Zorg dat je de regels kent rondom al dan niet het uitschrijven van getallen in letters.

Over dat laatste punt gaat dit hoofdstuk. Wanneer schrijf je getallen nou in letters en wanneer in cijfers?

Daarover zijn verrassend veel adviezen geschreven, maar gelukkig zijn ze het grotendeels met elkaar eens. Taaladvies.net, Onze Taal, de *Schrijfwijzer* en zelfs de website van de Rijksoverheid volgen dezelfde lijn. Dat is niet alleen een harde regel. Het advies is ook: wees redelijk, wees leesbaar, wees consequent. Gebruik je gezonde verstand.

De stelregel is simpel: schrijf hele getallen tot en met twintig in letters, en gebruik cijfers vanaf 21. Dus:

- Ik zag drie vogels.
- Hij gaf haar vijftien euro.
- In het verslag stonden 27 aanbevelingen.

Schrijf je *2 leerlingen* in plaats van *twee leerlingen*, dan lijkt het alsof je nadruk legt op het getal — en dat is meestal niet de bedoeling.

Nummeringen doe je ook met cijfers. Dus niet: *hoofdstuk negentien.*

Als er een maat, tijdsduur of andere eenheid achter staat, gebruik je cijfers:

- 7 liter water, 2,5 maand, 4 uur en 36 minuten
- niet: zeven liter water — dat oogt als een kinderboek

In technische context is dit de standaard. Het leest niet alleen vlotter, het voorkomt ook fouten.

Let ook op het enkelvoud bij kommagetallen. Je schrijft *2,5 maand*, *1,5 liter* en *0,8 seconde* — niet *2,5 maanden* of *0,8 seconden*. Dat voelt wat onnatuurlijk, maar is correct. Volgens Taaladvies.net: “Bij getallen groter dan twee is het enkelvoud gebruikelijk, zeker in combinatie met maten of eenheden.”

De 7 liter ketels worden al 2.5 maanden gebruikt.

De drie fouten in bovenstaande zin zijn hieronder verbeterd.

De 7-literketels worden al 2,5 maand gebruikt.

Daarmee zit alles goed: het getal is in cijfers, de komma is Nederlands, de eenheid is gekoppeld, en het zelfstandig naamwoord staat in het enkelvoud.

> Toen ik ChatGPT vroeg om de spelling in Meten & Weten grondig te controleren, wees hij me erop dat ik 'in volts' en 'in millivolts' schreef. Die eenheden moeten inderdaad in het enkelvoud. Maar 'een lengte in meters' is weer wel goed. Ik vroeg ChatGPT naar bronnen en kwam zo op betrouwbare voorbeelden. Maar of die zo echt zo logisch waren?

Breuken als *één derde* of *twee vijfde* schrijf je uit als het om een verhouding gaat: *Twee derde van de klas stemde tegen.* Maar zodra het om een meetgetal gaat, gebruik je cijfers met een komma: 1,5 liter, 0,4 seconde, 2,75 meter.

Voor duizendtallen gebruik je cijfers, met een (smalle) spatie als scheiding. Dus: 1 000 euro, 15 000 bezoekers, 3 200 000 kilometer. Dit is de officiële notatie volgens het SI-stelsel. Een duizendtalscheiding met een punt (*3.200.000*) is niet correct, hoe vaak je het ook tegenkomt.

> In Word typ je een smalle spatie met Unicode 2009 gevolgd door Alt+X.

'Wees consequent' is ook een soort regel. Als je *vijftien leerlingen* noemt en even later *23 leerkrachten*, voelt dat onevenwichtig. Het advies: schrijf álle getallen in cijfers zodra één ervan boven de twintig uitkomt:

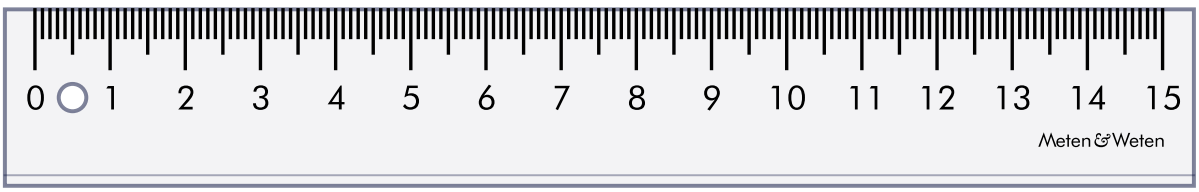
- niet: z*even leerlingen en 23 leerkrachten*
- wel: *7 leerlingen en 23 leerkrachten*
- ook: *zeven dwergen en één Doornroosje*

Dat oogt rustiger en voorkomt onbedoelde nadruk.

> Bij decimale getallen is het soms lastig kiezen tussen enkelvoud en meervoud. Toch is er een duidelijke regel:
> - Kies bij getallen onder 1 enkelvoud: 0,8 seconde, 0,3 meter, 0,5 liter.
> - Vanaf 1,1 gebruik je meervoud: 1,8 seconden, 3,7 liters, 6,2 maanden.
> - 1,0 blijft nog enkelvoud: 1,0 seconde.
>
> Zo is *0,99 seconde* nog net enkelvoud, maar 1,001 seconden (11 ms meer) meervoud — het draait om de harde grens van 1.
>
> Waarom zeggen we *drie uur* en *drie minuten*? En *tien jaar* ervaring, maar *twee maanden* vakantie? Dat zit niet in de logica, maar in het gebruik. Sommige eenheden (zoals jaar, uur, liter) blijven vaak enkelvoud, andere (zoals seconde, minuut, maand) worden meervoud. In het Nederlands althans: Engelsen schudden verbaasd hun hoofd. *That makes sense.*

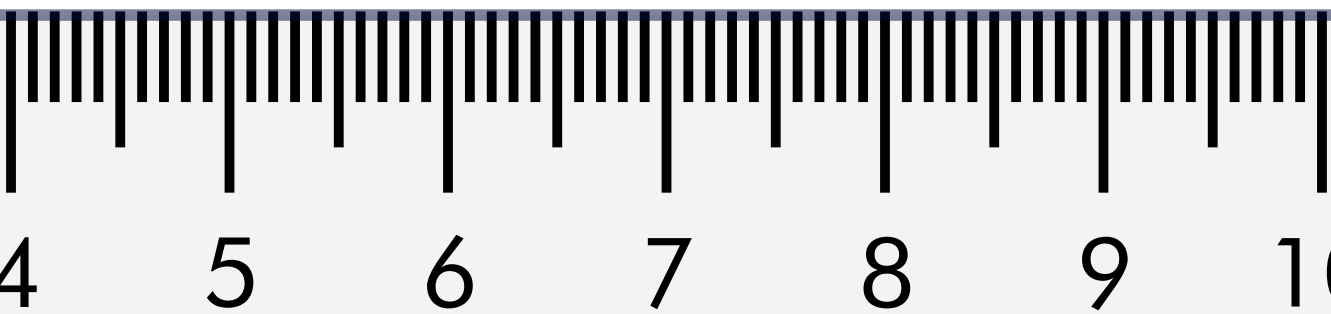
76. Verschillen meten

Als je de lengte van dit potlood zou willen meten, hoe doe je dat dan heel nauwkeurig?



Met een liniaal natuurlijk. Maar dan ben je er nog niet. Het potlood zal ook recht moeten liggen, en dit potlood ligt een graad of zes schuin. Je kunt zien dat het meer dan acht centimeter lang is... Maar hoe lang dan precies?

Laten we hem recht leggen en er dan eens flink op inzoomen.



Wow! Dat is nog eens inzoomen zeg! Het potlood blijkt toch een stukje langer te zijn dan ik dacht. Maakt het dan zoveel uit, dat hij recht ligt?

Maar wacht eens even...

Als we nauwkeurig willen meten, zullen we ook moeten controleren of de 0 exact bij het andere uiteinde van het potlood ligt. Dat kunnen we helemaal niet zien op deze twee afbeeldingen.

We kunnen niet zien of het potlood exact bij de nul ligt, maar we kunnen ook niet zien of hij er een dikke centimeter naast ligt. Alles kan, zolang we het niet kunnen controleren.

Er kunnen twee dingen aan de hand zijn.

- Het potlood ligt zo goed mogelijk bij de nul, maar natuurlijk nooit helemaal exact. We kunnen daar een half streepje naast zitten. In dat geval hebben we te maken met een *onzekerheid* in de meting als gevolg van een onvermijdelijke fout.
- Het potlood ligt helemaal niet bij de nul, als gevolg van onoplettendheid of opzet. We kunnen er dan een flink stuk naast zitten. In dat geval hebben we te maken met een *verkeerde meting* als gevolg van een al dan niet opzettelijke fout, die te vermijden was.

In dit vakgebied moeten we die twee dingen altijd goed uit elkaar houden.

- Je hebt fouten die een gevolg zijn van verkeerd meten.
- Je hebt fouten die zelfs optreden als je onberispelijk gemeten hebt.

Is die eerste categorie te vermijden? Nou ja, je ontkomt er als mens niet aan dat je fouten maakt. Dat is niet erg. Maar wie z'n best doet, maakt er minder. En als je het werk dat je zelf doet controleert én laat controleren, blijven er veel minder over.

Die tweede categorie, daar kom je niet onderuit. Wat je wél kunt doen is er rekening mee houden.

Als je de lengte van (bijvoorbeeld) een potlood opmeet, doe je eigenlijk *twee* metingen. Beter gezegd: je doet twee *aflezingen*. Aan beide uiteinden van het potlood lees je iets af. Aan de ene kant probeer je dat precies nul te laten zijn. Aan de andere kant schat je zo goed mogelijk tussen twee streepjes.

Een vuistregel is dat de onzekerheid gelijk is aan de helft van de afstand tussen twee streepjes. Omdat je dat twee keer hebt, stellen we de onzekerheid gelijk aan de hele afstand tussen twee streepjes. In dit geval: 1 mm.

77. Lijnen door punten

Als je één punt hebt, hoeveel lijnen kun je daar dan doorheen trekken?

Oneindig veel!

Als het er *oneindig veel* zijn, betekent dat dan ook dat *alle* lijnen door dat punt heen gaan? Nee. Oneindig veel betekent niet allemaal. Voor elke lijn die wel door het punt gaat, kun je oneindig veel lijnen bedenken die er niet doorheen gaan.

> Het begrip *oneindig* is lastig, maar ga je beter bevatten als je het verhaal kent van die beheerder van een oneindig groot hotel. Alle kamers van het hotel (oneindig veel) zijn bezet: er zijn dus oneindig veel gasten.
> Nu komt er iemand binnen die vraagt of er nog een kamer vrij is. Nee, natuurlijk. Maar de beheerder bedenkt een list. Hij boekt de gasten die verblijven in kamer 1 om naar kamer 2. Daardoor komt kamer 1 vrij! Maar ja… waar laat je dan de gasten uit kamer 2? Gewoon, in kamer 3. Want degenen die daar slapen, kunnen óók een kamer opschuiven, naar kamer 4. En op dezelfde manier kunnen de mensen uit kamer 4 naar 5 en die uit N naar N + 1. Alle gasten (oneindig veel dus) schuiven één kamer op. Dat kan zonder problemen, want het aantal kamers eindigt niet. Na elke kamer komt er weer een kamer. En daarna weer eentje.
>
> Duurt dat opschuiven dan niet oneindig lang? Nee hoor. Het gaat wel oneindig door, maar het duurt niet lang. Alle kamerwisselingen vinden tegelijkertijd plaats.

Er kunnen oneindig veel lijnen door één punt. En je kunt al die lijnen wiskundig nog beschrijven ook.
Voor elke rechte lijn geldt de volgende vergelijking.

$$y = ax + b$$

Voor elk punt (x_i, y_i) geldt dat er voor elke mogelijke waarde van a (en dat zijn er oneindig veel) een $b = y_i - ax_i$ bestaat waardoor die lijn door het punt (x_i, y_i) gaat.

Als je twee punten hebt, hoeveel lijnen kun je daar dan doorheen trekken? Eentje! Altijd precies één, tenzij de punten samenvallen. Dan kunnen er oneindig veel punten door, omdat je eigenlijk maar één punt hebt. Door twee

niet-samenvallende punten (x_i, y_i) en (x_j, y_j) is er altijd één lijn mogelijk. De richtingscoëfficiënt a ligt vast door die twee punten.

$$a = \frac{y_j - y_i}{x_j - x_i}$$

De waarde van b ligt ook vast.

$$b = y_i - ax_i$$

> In plaats van het punt (x_i, y_i) mag je ook (x_j, y_j) invullen om b te bepalen. Er komt namelijk toch hetzelfde uit.

Als je drie punten hebt, hoeveel lijnen kun je daar dan doorheen trekken?

Nul of één!

Als dat derde punt exact op de lijn ligt die door de eerste twee gaat, dan past er precies één lijn door alle drie. (Er zijn trouwens oneindig veel punten op die lijn, maar er zijn nog veel meer punten die er *niet* op liggen.) Als dat derde punt net niet of helemaal niet op de lijn door de eerste twee ligt, dan gaat er geen lijn meer door de drie punten.

We hebben nu dus dit.

- één punt → oneindig veel lijnen
- twee punten → één lijn
- drie punten → nul of één lijn

Voor vier en meer punten geldt hetzelfde als voor drie. Ze kunnen allemaal op één lijn liggen. (Naarmate je meer punten hebt is de 'kans' dat dat zo is kleiner.)

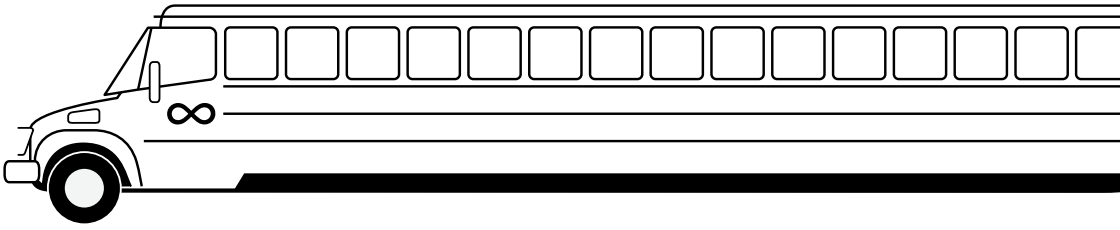
> Over de situatie met nul punten valt er ook iets te op te merken. Hoeveel lijnen gaan er door nul punten heen? Nul, zou je zeggen. Aan de andere kant: geef mij een willekeurig aantal punten en er is een lijn te bedenken die er niet doorheen gaat. Misschien zelfs oneindig veel. Misschien is het zelfs zinloos om daarover na te denken.

Door drie punten kun je wel altijd een parabool tekenen. En door vier punten past hoe dan ook een derdegraads polynoom. En door N punten een polynoom met de graad N — 1.

Bij metingen hebben we vaak veel meer punten dan nodig om er een lijn doorheen te trekken. Wat je dan zoekt, is de ‘best mogelijke’ lijn. Maar wat is de best mogelijke? Daarover gaat het in het volgende hoofdstuk.

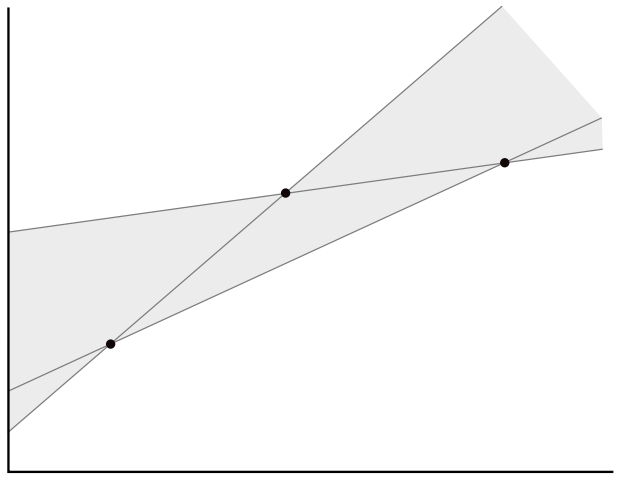
Nog even over die beheerder van dat hotel. Stopt er opeens een oneindig grote touringcar met oneindig veel mensen erin. Die willen allemaal een kamer hebben. Dat gaat natuurlijk nooit passen. Hoewel… Als de gasten uit kamer 1 nu eens naar kamer 2 gaan, en die uit kamer 2 niet naar 3, maar naar 4. Dan komen kamer 1 én 3 alvast vrij, want die uit kamer 3 gaan naar 6. De gasten uit 4 gaan naar 8, die uit 5 naar 10. En die uit N naar 2N. Zo komen alle oneven kamers beschikbaar. En dat zijn er oneindig veel. Want voor elke oneven kamer geldt: twee kamers verderop is er weer eentje.

Gelukkig is de parkeerplaats oneindig groot en kan die bus gewoon achter het oneindig grote hotel staan.



78. Kleinste kwadraten

Je kunt dus door drie punten één lijn trekken, maar dan moeten ze wel precies in elkaars verlengde liggen. En meestal doen ze dat niet. Bijvoorbeeld in Figuur 78.1.



Figuur 78.1 Het "grijze gebied" van lijnen door drie punten

Als je toch een lijn wilt trekken, dan moet je ergens concessies doen. Ergens ertussenin gaan zitten. Niet alles is even waarschijnlijk: normaal gesproken zul je toch iets kiezen in het grijze gebied.

Laten we een iets specifieker geval kiezen, met coördinaten erbij. Welke lijn zou je trekken door de punten (2,2), (6,3) en (8,4)? Dat eerste punt ligt ergens links, de andere twee juist rechts. Als we er van de rechter twee eentje achterwege laten, vind je een oplossing.

$$y = ax + b$$

$$a = \frac{y_2 - y_1}{x_2 - x_1}$$

$$b = y_2 - ax_2 \qquad \text{of} \qquad b = y_1 - ax_1$$

Vul je hier de punten (6,3) en (2,2) in, dan vind je dat a = $^1/_4$ en b = $^3/_2$.

Vul je hier de punten (8,4) en (2,2) in, dan vind je dat a = $^1/_3$ en b = $^4/_3$.

Ergens daartussen moet het toch liggen, zou je zeggen. Maar wat is nou het beste? De beide a's en b's middelen? Dan moet je wel met breuken kunnen rekenen… Laten we alles in twaalfden doen.

Het gemiddelde van $^1/_4$ en $^1/_3$ is het gemiddelde van $^3/_{12}$ en $^4/_{12}$. Dat is $^7/_{24}$.

Het gemiddelde van $^3/_2$ en $^4/_3$ is het gemiddelde van $^{18}/_{12}$ en $^{16}/_{12}$. Dat is $^{17}/_{12}$.

Toch wordt zo'n 'middelste' lijn niet zomaar als de beste oplossing beschouwd. Meestal willen we de lijn zo trekken dat de verschillen tussen de lijn en de punten zo klein mogelijk zijn. En met 'zo klein mogelijk' bedoelen

we dan dat de kwadraten van de verschillen minimaal zijn. De waarden van a en b waarvoor dat geldt, volgen uit onderstaande formules.

$$a = \frac{N\sum xy - \sum x \sum y}{N\sum x^2 - \left(\sum x\right)^2}$$

$$b = \frac{\sum x^2 \sum y - \sum x \sum xy}{N\sum x^2 - \left(\sum x\right)^2}$$

Ja, daar word je een beetje duizelig van. Maar toch valt het wel mee. Tenminste: als je de moeite neemt om die termen en factoren even allemaal in een tabelletje te zetten.

	x	y	xy	x^2
	2	2	4	4
	6	3	18	36
	8	4	32	64
Σ	16	9	54	104

Vier kolommen met daarin de drie punten en hun x en y en daarnaast het product xy en x^2. De waarden in elk van die vier kolommen tel je op en je vindt vier getallen: 16, 9, 54 en 104. Met die vier getallen en de waarde voor het aantal punten N (in dit geval dus drie) kun je alles uitrekenen.

$$a = \frac{N\sum xy - \sum x \sum y}{N\sum x^2 - \left(\sum x\right)^2} = \frac{3 \times 54 - 16 \times 9}{3 \times 104 - 16^2} = \frac{162 - 144}{312 - 256} = \frac{18}{56}$$

$$b = \frac{\sum x^2 \sum y - \sum x \sum xy}{N\sum x^2 - \left(\sum x\right)^2} = \frac{104 \times 9 - 16 \times 54}{3 \times 104 - 16^2} = \frac{936 - 864}{312 - 256} = \frac{72}{56}$$

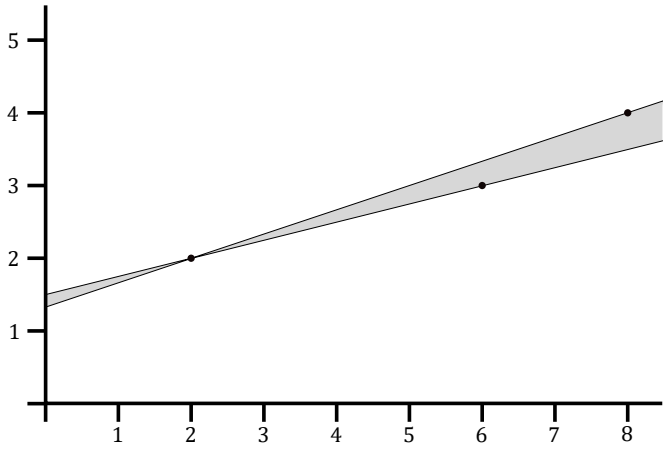
Het is eigenlijk niet moeilijk, maar het zijn stappen die je wel netjes één voor één moet kunnen uitvoeren. Zorgvuldig werken, daar gaat het om.

De waarden van a en b kun je ook als decimale getallen schrijven.

$$a = 0{,}3214\overline{285714}$$

$$b = 1{,}\overline{285714}$$

Het zijn exacte waarden, omdat de strepen boven de laatste cijfers aangeven dat het repeterende breuken zijn. Onafgerond dus. Exact.



Figuur 78.2 Twee mogelijke lijnen door drie punten

Maar moet het nou echt zo moeilijk? Nee.

Tik bij Google maar eens 'least squares online' in en je vindt heel wat websites die dit voor je kunnen uitrekenen. Als je daar de waarden invult, vind je hetzelfde. Meestal in een paar decimalen en/of significante cijfers:

$$a = 0{,}3214$$

$$b = 1{,}286$$

Van de repeterende breuk is niets meer te zien. Van de notatie in breuken al helemaal niet. Maar voor praktische situaties is het meestal goed genoeg.

In *Excel* is het ook al niet zo moeilijk. Zet de y-waarden in een kolom en x-waarden ernaast. (Let op: eerst de y en dan de x). De waarden van a en b vind je met deze formule. (De waarden van a en b worden in twee cellen getoond.)

```
=LIJNSCH(B2:B4;A2:A4)
```

En in *Python*? Is dat ingewikkeld? Nee. Je moet even de code weten.

```
from numpy.polynomial import polynomial as P

x = [2,6,8]
y = [2,3,4]

c, stats = P.polyfit(x,y,1,full=True)
print(c)
```

Deze code is trouwens voor het fitten van een willekeurige polynoom, dus niet per se en rechte lijn zoals hier. Door drie punten gaat altijd precies een parabool, dus je zou voor die 1 ook een 2 kunnen invullen.

79. Résidu

In het vorige hoofdstuk probeerden we zo goed mogelijk een rechte lijn te trekken door drie punten die niet op één lijn zaten met elkaar. We zagen dat er verschillende mogelijkheden waren, en dat er eentje is die we vaak als beste zien: de lijn waarvoor geldt dat de *som van de kwadraten van de residuen* zo klein mogelijk is.

Residuen? Ja. Zo hadden we ze nog niet genoemd, maar zo heten ze wel.
Het verschil tussen de y-waarde van een punt en de y-waarde van de lijn die we gefit hebben bij de x-waarde van datzelfde punt, is een *residu*.

> Ik vroeg me af wat de Nederlandse vertaling van residu is, omdat ik dacht dat het Frans was. Dus ik tikte 'residu vertaling' in bij Google en belandde (zoals ik gehoopt én verwacht had) bij Google Translate.
> En wat zei die? Die zei dat de *Franse* vertaling van het *Nederlandse* woord 'residu' 'résidu' is. Het wordt dus Frans door dat accentje op de é.
> Het 'echte' Nederlandse woord is 'overblijfsel'.

De residuen van drie punten reken je uit door de waarde op de gefitte rechte lijn $y_{fit} = 0{,}353214\,x + 1{,}286$ af te trekken van de 'werkelijke' y.

x	y	y_{fit}	residu	residu²
2	2	1,9288	-0,0712	0,005069
6	3	3,2144	0,2144	0,045967
8	4	3,8572	-0,1428	0,020392

Laten we voor de gein onze 'eigen' rechte lijn hier eens mee vergelijken. We hadden zelf al een keer een rechte lijn bedacht die door het linker punt ging en zo'n beetje midden tussen die andere twee door. Dat paste ook wel aardig, toch? We kregen er exacte breuken uit.
En een exacte breuk is wel zo leuk.

Om het residu in elk punt te vinden, gaan we dus

$$y_{onze\ eigen(wijze)\ fit} = \frac{7}{12}\,x + \frac{24}{17}$$

aftrekken van de 'werkelijke' y.

x	y	$y_{o.e.w.f.}$	residu	residu²
2	2	2	0	0
6	3	3,166667	0,166667	0,027778
8	4	3,75	-0,25	0,0625

Past die lijn van ons *beter* dan die van de kleinste kwadraten methode? De residuen zijn kleiner... Zowel in absolute waarden als in de kwadraten. Kijk maar in de tabel. Voor het punt (2,2) is 'ons' residu exact *nul*. Voor het punt (6,3) is het kwadraat ongeveer de helft. Alleen voor dat derde punt is hij iets groter. Het tweede past beter, het derde slechter, het eerste is *exact* goed. Een twee derde meerderheid van de stemmen zegt dat onze lijn beter is.

Maar ja, wanneer is iets *beter*? Als je naar de harde feiten kijkt en de som van de kwadraten optelt, is die van de door *Excel* en *Python* of wat dan ook gefitte lijn inderdaad beter. Tel ze maar bij elkaar op. De som van 'onze' lijn is afgerond 0,09. Die van 'hun' lijn is afgerond 0,07. Het scheelt ruim een kwart.

0,005069	0
0,045967	0,027778
0,020392	0,0625
0,071429	0,090278

Maar waarom kijken we eigenlijk steeds naar de *kwadraten* van de verschillen? Als ik wil weten of A dichter bij B ligt dan bij C, dan ligt het toch voor de hand om naar de *absolute* verschillen te kijken?

> Arend haalt een 6,1 voor het tentamen en Bert een 9,6. Cor had een 5,4. Bij wie komt Arend het dichtst in de buurt? Bij Cor, want dat scheelt maar een halve punt en het verschil met Bert is drieënhalve punt. Dat is een factor zeven groter! Toch zitten Arend en Bert in dezelfde categorie, namelijk de club die de taart kan aansnijden. Die arme Cor is gezakt. Zo zie je hoe betrekkelijk getallen en cijfers zijn. De meest relevante vraag is in dit geval of het een voldoende was.

Als we de absolute waarden van de residuen nemen, krijgen we een andere vergelijking. Het scheelt heel weinig (zo'n 2,8%), maar 'onze' simpele fit is volgens de som van de absolute waarden iets kleiner dan de lijn die via de methode van de kleinste kwadraten gevonden is.

0,0712	0
0,2144	0,1667
0,1428	0,25
0,4284	0,4167

Zouden er nog andere manieren en criteria te bedenken zijn om te kijken naar een 'beste' lijn? Jazeker. Of het de beste is, is altijd nog een vraag, maar op zijn minst kun je naar een zinvolle lijn zoeken.

Wat te denken van de lijn die het grootste residu zo klein mogelijk maakt? Uitgaande van de keuze om de lijn links door dat ene punt te laten gaan en

rechts ergens tussen die andere twee door, kunnen we beginnen met vast te stellen dat het residu van het eerste punt per definitie gelijk is aan nul.

Voor een punt (x_i, y_i) geldt:

$$r_i = y_i - ax_i - b$$

Het kleinst mogelijke maximale verschil krijg je als de verschillen qua grootte aan elkaar gelijk zijn. Het ene positief en het andere negatief. In alle andere gevallen zijn de verschillen namelijk groter: als het ene residu afneemt, neemt het andere toe.

$$y_2 - ax_2 - b = -(y_3 - ax_3 - b)$$

$$y_2 + y_3 - a(x_2 + x_3) = 2b$$

$$b = \frac{y_2 + y_3 - a(x_2 + x_3)}{2} = \frac{3 + 4 - a(6 + 8)}{2} = 3{,}5 - 7a$$

We hebben nu één vergelijking met twee onbekenden, en hebben er dus nog eentje nodig. Die volgt uit dat andere punt (helemaal links) waarvoor geldt dat het residu nul is.

$$0 = y_1 - ax_1 - b$$

$$b = y_1 - ax_1 = 2 - 2a$$

Deze vergelijking kunnen we combineren met de vorige.

$$2 - 2a = 3{,}5 - 7a$$

$$5a = 1{,}5$$

$$a = 0{,}3$$

$$b = 2 - 2a = 1{,}4$$

Een andere lijn dus, die volgens een ander criterium ‘beter’ is.

80. Past het model?

Hughes & Hase hebben een alleraardigst raadseltje, dat aangeeft dat het fitten van een hoge orde functie op een eindig aantal getallen altijd wel een perfecte pasvorm oplevert, maar niet altijd zinnig is.

Voordat ze hun raadseltje opgeven, staan ze eerst stil bij het begrip *vrijheidsgraad*. Als je N onafhankelijke datapunten hebt en je fit een functie met P parameters, dan is het aantal vrijheidsgraden v (*degrees of freedom*)

$$v = N - P.$$

> Ten onrechte denken sommige mensen dat het aantal vrijheidsgraden bij voorkeur zo klein mogelijk is. Natuurkundig gezien wil je juist veel vrijheidsgraden hebben: een fit met zo weinig mogelijk parameters.

Nu het raadseltje: wat is het volgende getal in de reeks 1, 2, 4, 6, 10?

Je bedenktijd is voorbij! Het juiste antwoord is: 12. Waarom? Omdat de reeks gevormd wordt uit 'de priemgetallen min 1'. De priemgetallen zijn 2, 3, 5, 7, 11, 13, 17, 19 etc. Als je daar 1 van aftrekt, vind je 1, 2, 4, 6, 10, 12, 16, 18 etc.

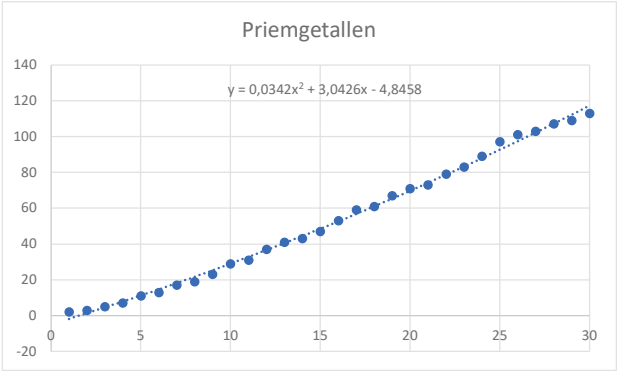
Het lukt niet om een lineaire functie door deze punten trekken. Kwadratisch gaat al iets beter en een derde macht begint ergens op te lijken. Omdat het om vijf punten gaat, kun je er altijd een vierdegraads polynoom exact doorheen trekken — net zoals een rechte lijn precies door twee punten past, en een parabool door drie.

Deze functie past volgens Hughes & Hase.

$$f(x) = 5 - 8{,}5833x + 5{,}875x^2 - 1{,}4167x^3 + 0{,}125x^4$$

Hier kun je $x = 6$ invullen en dan komt er 21 uit. Dat antwoord volgt perfect uit de gefitte functie, maar het is niet wat de bedenker van het raadseltje in gedachten had.

In een grafiek zien de eerste dertig priemgetallen er als volgt uit.



Nee, ze liggen niet op een rechte lijn en er is geen simpele formule om de volgende waarde exact te voorspellen. Maar zelfs een simpele parabool kan zodanig gefit worden dat alle dertig verschillen onder de 4,5 blijven. Een vierdegraadsvergelijking om het een beetje beter te doen, is overdreven.

Wat nu, als je een polynoom probeert te fitten op data waar wél een wiskundig verband achter zit? Hughes & Hase gebruiken — na dat raadseltje van het vorige hoofdstuk — de wiskundige beschrijving die het verband aangeeft tussen slingertijd, slingerlengte én beginhoek.

$$T = 2\pi\sqrt{\frac{l}{g}}\left(1 + \frac{1}{16}\theta^2\right)$$

Dit verband is een vereenvoudiging van de eerste-orde correctie op de formule die je meestal ziet en die 'geldig' is voor kleine hoeken.

$$T = 2\pi\sqrt{\frac{l}{g}}$$

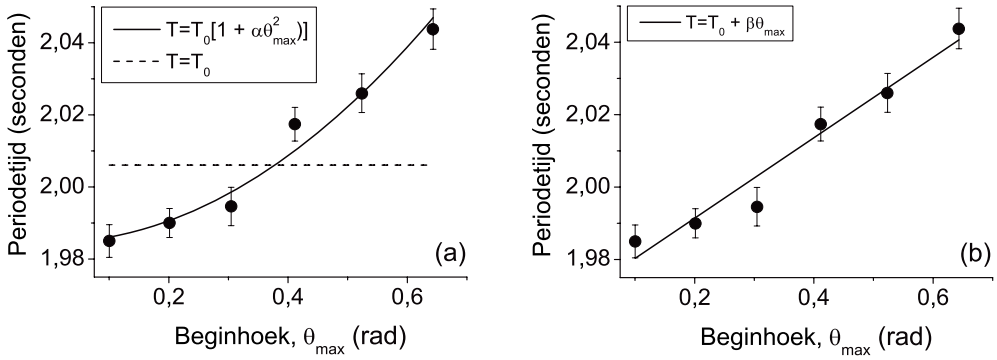
De 'echte' analytische oplossing bevat oneindig veel termen.

$$T = 2\pi\sqrt{\frac{l}{g}}\left(1 + \frac{1}{16}\theta^2 + \frac{11}{3072}\theta^4 + \frac{173}{737280}\theta^6 + \frac{22931}{132120570}\theta^8 + \cdots\right)$$

Je bent al aardig op weg als je alleen de eerste twee neemt.

Die formule met die ene term met θ^2erin is dus kwadratisch afhankelijk van de beginhoek als je de hogere-orde termen negeert. Dat kun je theoretisch afleiden, maar blijkt het ook uit metingen?

Hughes & Hase geven een voorbeeld waarbij ze niet de meetdata zelf geven, maar wel een grafiek met meetresultaten. Grafisch ziet het er als volgt uit.



In beide plaatjes staan dezelfde meetpunten. Links is een poging gedaan om een parabool door de foutgebieden te trekken. Dat lukt net niet. Rechts is geprobeerd om dat met een rechte lijn te doen. Ook dat slaagt net niet helemaal. Vijf van de zes foutgebieden omvatten de lijn, eentje niet.

En dan is er ook nog die stippellijn (gestreepte lijn) die uitgaat van een constante waarde. Het lukt je niet om een horizontale lijn te trekken die door meer dan twee foutgebieden gaat. Belangrijker dan dat: je 'ziet meteen' dat de zes meetpunten een oplopende volgorde hebben. Voor alle zes punten geldt: als de hoek groter is, dan is de periodetijd ook groter.

Hughes & Hase hebben de Chi-kwadraattoets gebruikt om te toetsen hoe het zit. De wiskunde die daarachter zit, komt in Meten & Weten niet aan de orde. In het kort: een Chi-kwadraattoets levert een getal op dat in de buurt van 1 moet liggen. Als het 2 of hoger is (bij dit aantal vrijheidsgraden), moet je de zogenaamde nulhypothese verwerpen.

Drie nulhypotheses worden er door Hughes & Hase getoetst, namelijk de hypotheses dat de data beschreven worden door de volgende drie modellen.

$$T = T_0$$

$$T = T_0(1 + a\theta)$$

$$T = T_0(1 + b\theta^2)$$

De *waarden* (die dus rond het getal 1 moeten liggen en in ieder geval kleiner moeten zijn dan 2) voor deze drie modellen zijn achtereenvolgens 21,4, 0,9 en 1,1. Op basis daarvan wordt het eerste model duidelijk verworpen en zijn de andere twee heel goed mogelijk.

Een model dat Hughes & Hase níét hebben meegenomen in hun toetsing, maar dat wel verdedigbaar is, is het volgende.

$$T = T_0(1 + c\theta + d\theta^2)$$

Terwijl je met zekerheid kunt zeggen dat dit model óók past: als c of d gelijk is aan 0, dan hebben we namelijk die andere twee, waarvan we al weten dat ze pasten.

Oefening

1. Bereken of beredeneer dat het verdedigbaar is om de termen die van een hogere orde zijn dan de kwadratische term te verwaarlozen zijn.

Bespiegelen

81. Onzekerheid

Misschien wel het belangrijkste begrip in dit hele boek is ‘onzekerheid’.

> Helemaal zeker is dat niet, maar *waarschijnlijk* is onzekerheid het belangrijkste begrip. Hoe weet je eigenlijk of een begrip het belangrijkste is? Je zou kunnen zeggen dat een begrip in een boek het belangrijkste is als je geen ander begrip kunt bedenken dat belangrijker is. Maar wat nu als van zes wijze mensen er vijf zonder enige twijfel stellen dat *onzekerheid* het belangrijkste is, en eentje beweert glashard dat *weten* belangrijker is. Ze weten het alle zes zeker, zeggen ze…
>
> Wat is eigenlijk zekerheid? Heb je dat nog nooit meegemaakt dat je iets zeker meende te weten wat later toch niet waar bleek te zijn?
> “Ik zou toch zweren dat…”

De GUM besteedt heel paragraaf 2.2 hieraan: *The term “uncertainty”*.

The word “uncertainty” means doubt, and thus in its broadest sense “uncertainty of measurement” means doubt about the validity of the result of a measurement. Because of the lack of different words for this general concept of uncertainty and the specific quantities that provide quantitative measures of the concept, for example, the standard deviation, it is necessary to use the word “uncertainty” in these two different senses.

We gaan hier vaak aan voorbij, maar ze hebben wel gelijk. Je pakt een meetlat van 100 cm en stelt vast dat de lengte van een ijzeren staaf 45,2 cm is. De onzekerheid schat je op 1 mm. Maar als je naar huis fietst en je meetresultaat wilt meenemen in een berekening, ga je opeens twijfelen of je de meetlat niet achterstevoren had. Dan was die staaf dus 100 − 45,2 = 54,8 cm.

Er is nu *onzekerheid over de meting*. Misschien is die niet goed uitgevoerd en klopt het voor geen meter. Een andere mogelijkheid: je hebt je blind zitten staren naar die twee millimeterstreepjes en niet in de gaten had dat het geen 45,2 maar 46,2 was. Het kan ook zo zijn dat je afgeleid was toen je het opschreef en 45,2 noteerde terwijl je 54,2 had afgelezen. Maar ook als je het wel zeker weet (wanneer is dat zo) blijft er een stuk onzekerheid over. Dat is het soort onzekerheid dat we meestal bedoelden.

Twee soorten onzekerheid dus.

1. Of het wel klopt.
2. Hoeveel je er naast zou kunnen zitten.

The formal definition of the term "uncertainty of measurement" developed for use in this Guide and in the VIM [6] (VIM:1993, definition 3.9) is as follows:

> ***uncertainty (of measurement)***
>
> *parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand*

Wat is de vertaling daarvan volgens Google Translate?

> *ölçülen büyüklüğe makul bir şekilde atfedilebilecek değerlerin dağılımını karakterize eden, bir ölçümün sonucuyla ilişkili parametre.*
>
> Die vertaling çlüpt als een büş, maar omdat dit boek in het Nederlands geschreven is, kunnen we beter een andere doeltaal kiezen.
>
> *parameter, geassocieerd met het resultaat van een meting, die de spreiding karakteriseert van de waarden die redelijkerwijs aan de meetgrootheid kunnen worden toegeschreven*

De onzekerheid is dus een getal (meestal met een eenheid), maar kan ook als percentage of fractie worden uitgedrukt.

Nu moeten we niet in de valkuil van dat woord 'spreiding' (*dispersion*) trappen. Hier staat NIET dat er DUS een spreiding in de METINGEN is. Het kan best zijn dat je de meting tien keer over doet en tien keer exact hetzelfde vindt. In het geval van onze ijzeren staaf: 45,2 cm. Maar er is wel een spreiding in de waarden die *redelijkerwijs de werkelijke waarde kunnen zijn.* Misschien is die staaf wel 45,194720 cm ± 5 µm. Dat blijkt misschien wel als we een meetapparaat hebben in micrometers. Maar er zijn oneindig veel meer waarden die redelijk zijn: 45,226 cm kan ook. Of 45,24 cm. Er zit een spreiding in de mogelijkheden. Niet per se in de meting als we die herhalen.

The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.

Dit is even wennen voor wie zich voor het eerst in dit vakgebied verdiept. De parameter kan een standaardafwijking zijn, maar ook een op het oog harde grens. Als je zeker denkt te weten dat een waarde tussen twee streepjes zit (tussen 45,1 en 45,3 cm), dan kan het niet zo zijn dat de werkelijke waarde 45,4 cm is. Maar stel nu eens dat die lengte uit een grote reeks normaal verdeelde metingen kwam. Dan zou die 0,1 cm de standaardafwijking zijn. Het is niet onmogelijk dat de werkelijke waarde dan 45,4 cm was. De kans dat je twee keer de standaardafwijking van het gemiddelde zit, is zo'n 5%. Niet heel groot, het komt gemiddeld toch één op de twintig keer voor.

Laten we nog even verder lezen in de GUM.

The definition of uncertainty of measurement given in 2.2.3 is an operational one that focuses on the measurement result and its evaluated uncertainty. However, it is not inconsistent with other concepts of uncertainty of measurement, such as

- *a measure of the possible error in the estimated value of the measurand as provided by the result of a measurement;*
- *an estimate characterizing the range of values within which the true value of a measurand lies (VIM:1984, definition 3.09).*

Although these two traditional concepts are valid as ideals, they focus on unknowable quantities: the "error" of the result of a measurement and the "true value" of the measurand (in contrast to its estimated value), respectively. Nevertheless, whichever concept of uncertainty is adopted, an uncertainty component is always evaluated using the same data and related information.

Volgens Google Translate:

De definitie van meetonzekerheid gegeven in 2.2.3 is een operationele definitie die zich richt op het meetresultaat en de geëvalueerde onzekerheid ervan. Het is echter niet inconsistent met andere concepten van meetonzekerheid, zoals

- *een maatstaf voor de mogelijke fout in de geschatte waarde van de maatstaf, zoals blijkt uit het resultaat van een meting;*
- *een schatting die het bereik van waarden karakteriseert waarbinnen de werkelijke waarde van een meetgrootheid ligt (VIM:1984, definitie 3.09).*

Hoewel deze twee traditionele concepten geldig zijn als idealen, richten ze zich op onkenbare grootheden: respectievelijk de 'fout' van het resultaat van een meting en de 'echte waarde' van de meetgrootheid (in tegenstelling tot de geschatte waarde ervan). Niettemin wordt, welk concept van onzekerheid ook wordt gehanteerd, een onzekerheidscomponent altijd geëvalueerd met behulp van dezelfde gegevens en gerelateerde informatie.

Op die *onkenbare grootheden* 'werkelijke waarde' en 'fout' komen we terug.

De GUM heeft het over *the measurement result and its evaluated uncertainty.* Dat zou betekenen dat een meetresultaat een onzekerheid *heeft*… Maar is het niet correcter om te zeggen dat de onzekerheid *deel uitmaakt* van het meetresultaat? Stephanie Bell (1999) citeerden we in een eerder hoofdstuk al. *Uncertainty is a quantification of the doubt about the measurement result.* Het is net als met een fiets: je kunt zeggen dat een fiets wielen heeft, maar het is netter om te zeggen dat wielen onderdeel van die fiets zijn.

meetresultaat = beste schatting (+ onzekerheid)

82. Fout gebruik van de term ‘error’

Stephanie Bell (1999) beschrijft in haar boek *A Beginner’s Guide to Uncertainty of Measurement* glashelder op wat het verschil is tussen een fout (*error*) en onzekerheid (*uncertainty*). Dat het verschil belangrijk (*important*) genoemd wordt, is terecht.

2.3 Error versus uncertainty

It is important not to confuse the terms ‘error’ and ‘uncertainty’. Error is the difference between the measured value and the ‘true value’ of the thing being measured. Uncertainty is a quantification of the doubt about the measurement result. Whenever possible we try to correct for any known errors: for example, by applying corrections from calibration certificates. But any error whose value we do not know is a source of uncertainty.

Een fout is wel een *bron van onzekerheid*, maar het is niet *de onzekerheid zelf*. De fout kan heel klein zijn terwijl de onzekerheid groot is, als een toevallige fout toevallig klein is. Een fout kan groter zijn dan de onzekerheid, bijvoorbeeld bij een toevallige fout die normaal verdeeld is. In dat geval ligt 68% van de meetresultaten binnen één standaardafwijking rond het gemiddelde — dus ongeveer één op de drie valt daarbuiten.

We hebben nu twee mogelijke misverstanden.

- Een fout is niet per se ‘iets verkeerd gedaan hebben’. Soms (meestal) ontkom je er niet aan dat er een verschil zit tussen de werkelijke waarde en je meting.
- Een fout geeft aan hoeveel je ernaast zit, een onzekerheid is daar slechts een indicatie van. De eerste ken je niet, de tweede is een schatting. De eerste is een getal dat per meting verschilt, de tweede is een kansvariabele.

> Je kunt je afvragen of Stephanie Bell zelf helemaal correct is in haar woordkeuze. In Meten & Weten verstaan we onder een *meetresultaat* een *beste schatting met een onzekerheid*. Strikt genomen is het beter om te spreken over de twijfel over de *meetwaarde* in plaats van het *meetresultaat*.

Wat Stephanie Bell er niet bij zegt, is dat het vaak door elkaar wordt gehaald. Hughes & Hase benoemen dat wel.

You should note that despite many attempts to standardise the notation, the words 'error' and 'uncertainty' are often used interchangeably in this context— this is not ideal — but you have to get used to it!

Je zult er inderdaad aan moeten wennen dat veel mensen (en boeken) fouten maken. Maar het is niet handig om daar zelf onnodig aan mee te doen of er-overheen te lezen als iets niet klopt.

Herman Berendsen (1997) gaf zijn oorspronkelijk in het Nederlands uitgegeven boek de titel: *goed meten met fouten*. Je ontkomt niet aan de fout, zelfs niet als je alles goed doet. Maar je kunt het natuurlijk óók verkeerd doen.

In dit vakgebied is een *error* (fout) **NIET** een vergissing of een misser.

Kirkup & Frenkel (2006) zetten in hoofdstuk 3 van hun boek wat begrippen op een rij. In de tabel hieronder staan de Nederlandse vertalingen van die begrippen. De onderste is toegevoegd en staat niet in de genoemde tabel.

measurement	meting
measurand	meetgrootheid
value	waarde
true value	werkelijke waarde
best estimate	beste schatting
error	fout
random error	toevallige fout
systematic error	systematische fout
accuracy	nauwkeurigheid
precision	precisie
uncertainty	onzekerheid
accepted value	geaccepteerde waarde

Een paar van die begrippen halen we nu naar voren.

De waarde (*value*) bestaat uit een getalswaarde én (meestal) een eenheid. Het getal zelf is nog niet de waarde. (Sterker nog: zonder eenheid is de uitkomst in feite waardeloos én zoals eerder gezegd: FOUT.)

> Als iemand op een feestje aan je vraagt hoe oud je bent, mag je heus wel 'twintig' antwoorden, als dat waar is. Zolang je in de context van dit vakgebied maar netjes '20 jaar' zegt.

De werkelijke waarde (*true value*) is datgene dat we willen weten en meten, en dus niet kennen, zelfs niet na de meting. De fout (*error*) is het verschil tussen de beste schatting en de werkelijke waarde. De GUM zegt daarover:

NOTE 2 When it is necessary to distinguish "error" from "relative error", the former is sometimes called absolute error of measurement. This should not be confused with absolute value of error, which is the modulus of the error.

Als de fout (*error*) het verschil is tussen de gemeten waarde en de werkelijke waarde, ligt het voor de hand dat ook voor de fout geldt dat je die niet kent. De fout is: hoeveel je afwijkt van de werkelijke waarde. Je meetresultaat ken je exact, want je schrijft het in een eindig aantal decimalen op. Maar die *true value*… kenden we die maar.

In hoofdstuk 3.2, Errors, effects, and corrections, is de GUM zorgvuldig in het gebruik van het woord 'error'.

3.2.1 In general, a measurement has imperfections that give rise to an error (B.2.19) in the measurement result. Traditionally, an error is viewed as having two components, namely, a random (B.2.21) component and a systematic (B.2.22) component.

NOTE Error is an idealized concept and errors cannot be known exactly.

> Zucht. Ik vind het zo vermoeiend lezen als er dingen als *(B.2.19) en (B.2.22)* in een zin staan. Hadden ze daar nou geen voetnoten voor kunnen gebruiken? Ja, dat had gekund. Maar ze hebben het niet gedaan. In Meten & Weten staan trouwens nooit voetnoten.[1]

Het zijn dus toevallige en systematische fouten die zorgen voor een onzekerheid. Het verschil tussen de werkelijke waarde en je meetwaarde kan per geval verschillen, maar ook de hele tijd hetzelfde zijn.
Je kent het verschil niet. Vandaar die onzekerheid.

3.2.2 Random error presumably arises from unpredictable or stochastic temporal and spatial variations of influence quantities. The effects of such variations, hereafter termed random effects, give rise to variations in repeated observations of the measurand. Although it is not possible to compensate for the random error of a measurement result, it can usually be reduced by increasing the number of observations; its expectation or expected value (C.2.9, C.3.1) is zero.

---

[1] Behalve dan deze ene.

Je kunt niet systematisch compenseren voor een toevallige fout, want die is telkens anders. Maar je kunt óók niet compenseren voor een systematische fout als je die niet kent.

NOTE 1 The experimental standard deviation of the arithmetic mean or average of a series of observations (see 4.2.3) is not the random error of the mean, although it is so designated in some publications. It is instead a measure of the uncertainty of the mean due to random effects. The exact value of the error in the mean arising from these effects cannot be known.

Ook dit is waar. Je kunt de standaardafwijking van de theoretische verdeling natuurlijk schatten uit herhaalde metingen, maar je weet niet exact hoe groot die in werkelijkheid is. Als je heel veel metingen doet, kom je zeer waarschijnlijk wel heel dicht in de buurt. Er zit niet alleen een fout in je beste schatting, maar ook in je onzekerheid.

NOTE 2 In this Guide, great care is taken to distinguish between the terms "error" and "uncertainty". They are not synonyms, but represent completely different concepts; they should not be confused with one another or misused.

Taylor (1997) schrijft in de eerste paragraaf van hoofdstuk 1 het volgende.

1.1 Errors as Uncertainties

In science, the word error does not carry the usual connotations of the terms mistake or blunder. Error in a scientific measurement means the inevitable uncertainty that attends all measurements. As such, errors are not mistakes; you cannot eliminate them by being very careful. The best you can hope to do is to ensure that errors are as small as reasonably possible and to have a reliable estimate of how large they are. Most textbooks introduce additional definitions of error, and these are discussed later. For now, error is used exclusively in the sense of uncertainty, and the two words are used interchangeably.

Tsja. *The two words are used interchangeably.* Dat is ook nogal andere taal dan de GUM gebruikt.

In this Guide, great care is taken to distinguish between the terms "error" and "uncertainty". They are not synonyms, but represent completely different concepts; they should not be confused with one another or misused.

Oefening

1. Geef van de begrippen in het lijstje van Kirkup & Frenkel aan of ze een onzekerheid hebben.

83. Verificatie & Validatie

Het verschil tussen *verificatie* en *validatie* zou je kunnen opzoeken in een woordenboek, maar dan kom je misschien niet veel verder.

- *Verifiëren* is 'de juistheid (of waarheid) aantonen'.
- *Valideren* is 'geldig verklaren'.

Beide woorden komen uit het Frans, vijftiende en zestiende eeuw. Maar ja, wat heb je eraan om dat te weten?

Wikipedia geeft fraai weer wat het verschil is tussen *verificatie* en *validatie*.

- *Verificatie: men controleert of het systeem juist is gemaakt (zijn alle eisen verwerkt?)*
- *Validatie: men controleert of het juiste systeem is gemaakt (doet het wat de klant wil?)*

Dat kun je nog korter zeggen:

- Verificatie: maken we het *juist*? (klopt het *met* de eisen?)
- Validatie: maken we *het juiste*? (kloppen *de eisen*?)

Ik kan de breedte van een rechthoekige tafel opmeten met een rolmaat. Mijn meetresultaat kan ik laten verifiëren door iemand anders te vragen om het met een duimstok na te meten. Vindt hij of zij hetzelfde? Dan is er sprake van *verificatie* van de meting. Als die ander nagaat of ik wel de kleinste van beide zijden heb genomen, is er sprake van *validatie*. De langste zijde moet je niet hebben: dat is niet de breedte, maar de lengte.

Er is een verschil tussen *nagaan of* en *nagaan dat*. In het eerste geval kijk je *of iets zo is* (of niet!), in het tweede geval zie je *dat het zo is*.

"Heb je gecontroleerd of de deur op slot zit?"
"Ja."
"En? Zat-ie op slot?"
"Nee."
"Moet je hem niet op slot doen dan?"
"Nee."
"Waarom niet?"
"Omdat ik dat gedaan heb, toen ik zag dat-ie niet op slot zat…"
"Toen'? Bedoel je niet: 'nadat'?"

84. Relatief & Absoluut

De *absolute onzekerheid* is een grootheid die wordt uitgedrukt in dezelfde eenheden als de gemeten grootheid. Die geeft dus het bereik aan waarin de werkelijke waarde waarschijnlijk zal liggen. De werkelijke waarde kan zich buiten dat bereik bevinden.

> Let op dat het bereik waarover we het nu hebben dus *twee keer zo groot* is als de grootheid die we onzekerheid noemen. En we weten nog niet eens voor 100% zeker dat de grootheid daarbinnen ligt.

De *relatieve onzekerheid* is de absolute onzekerheid gedeeld door de gemeten grootheid zelf. Dat is dus een dimensieloos getal, vaak aangeduid als een percentage.

> Let op dat dit percentage groter kan zijn dan 100%.
> Een standaardvoorbeeld is een spanning die heel klein is. Die spanning zou bijvoorbeeld −2 mV kunnen zijn terwijl de onzekerheid 5 mV is. Dan is de relatieve onzekerheid 40%.
> (Ook een relatieve onzekerheid is een absolute waarde, nooit negatief.)

Kan een relatieve fout ook *relatief* groot of klein zijn? Ja. In hoofdstuk 47 (titel: *Met & zonder afhankelijkheid*) wordt gerekend aan de onzekerheid in de gravitatieversnelling die je vindt via een valproef. Gekeken wordt of ze afhankelijk of onafhankelijk van elkaar zijn en hoe dat dan doorwerkt. Je ziet daar ook dat de onzekerheid volgt uit een formule waarin de relatieve fout in de tijd twee keer meetelt en de relatieve fout in de afstand één keer.

Laten we voor het gemak even doen alsof die onzekerheid in de lengte en in de tijd afhankelijk van elkaar zijn.

> Waarschijnlijk zijn ze dat *niet*. Die lengte meet je met een rolmaat en de tijd met een stopwatch. Het is niet logisch dat de onzekerheid van de ene iets met de andere te maken heeft. Het klopt dat een langere slinger een langere slingertijd geeft. Maar we hebben het niet over de afhankelijkheid in de werkelijke waarden, maar over de afhankelijkheid in de onzekerheden.

Dan is dit de formule.

$$\frac{\delta g}{|g|} = \frac{\delta s}{|s|} + 2\frac{\delta t}{|t|}$$

Kun je zeggen welke onzekerheid meer bijdraagt in de totale onzekerheid? Nee. Je weet wel dat het *gewicht* bij de tijd twee keer zo groot is. Maar als de

relatieve onzekerheid in de afstand 4% is en de relatieve onzekerheid in de tijd 0,1%, dan draagt die eerste een factor twintig keer zo veel bij.

Kun je zeggen dat de ene relatieve fout groter is dan de andere? Ja. Die 4% is veertig keer zo veel als 0,1%.

> Dat is het aardige aan relatieve onzekerheden. Ze geven niet alleen een indruk van hoe de onzekerheid zich verhoudt tot de waarde, maar je kunt ze ook vergelijken met andere grootheden met een andere dimensie. Door het relatieve ben je van die dimensies af.

Relatieve onzekerheden hebben ook iets vervelends. Voor kleine waarden van de gemeten grootheden worden ze al gauw erg groot. Als ik wil vaststellen dat de spanning 'bijna nul' is, zeg ik liever dat de onzekerheid 2 mV is dan dat de onzekerheid 40% is. Dat zegt namelijk weinig.

Kun je zeggen dat de ene relatieve fout relatief groot is? Ja. Hij is bijvoorbeeld veertig keer zo groot als die andere. Nu kun je óók nog van mening verschillen over de vraag hoeveel keer zo groot relatief groot is. Alles is betrekkelijk.

85. Graden & Ontraden

Wat is het verschil tussen een rechte hoek en het kookpunt van water?

Als je zegt dat het tien graden is, dan zie je het onzinnige van de vraag niet in. Of juist wel. Een graad op je geodriehoek is iets heel anders dan een graad (Celsius of Fahrenheit) op de thermometer. Er is wel een overeenkomst: ze worden allebei als maateenheid voor wetenschap en techniek ontraden.

Dat een graad in de temperatuur iets anders is dan een graad in een hoek, zie je aan de C of F die erachter staat. En nu we het er toch over hebben: het woord *graden* (of het tekentje °) wordt niet gebruikt als we het over de eenheid kelvin hebben. Dat is pas in 1967 besloten, dus in boeken van voor die tijd zie je het soms nog anders. Eén kelvin (dat is dus 1 K) is gedefinieerd als het 1/273,16 deel van het verschil tussen het absolute nulpunt en het *tripelpunt* van water: de temperatuur waarop een stof tegelijkertijd kan voorkomen als vaste stof, vloeistof en gas.

Dat tripelpunt van water ligt op 0,01 °C: nét iets boven wat we in de volksmond het vriespunt noemen. Het goede nieuws is dat het verschil zo klein is dat het in de praktijk geen betekenis heeft. Geen enkele weerman verwacht dat het morgen 18,01 °C wordt. Ze zeggen allemaal 'achttien graden', zonder cijfers achter de komma en zonder 'Celsius' erbij.

Als het 20 °C is, dan kun je dat noteren als

$$T = (20{,}0 \pm 0{,}5)\ °C$$

> Let op dat je een spatie zet tussen getal en eenheid. Volgens de internationale ISO-normen (en de Nederlandse norm NEN EN ISO 80000-1:2013) hoort dat zo, maar taalkundigen raden soms aan om 20°C te schrijven. Bij de 90° van een rechte hoek noteer je het gradenteken trouwens wel achter het getal.

De eenheid is met een hoofdletter K, maar als je het als woord uitschrijft in de tekst is het juist met een kleine letter: *kelvin*. Net als meter en volt en ampère: dat is ook met kleine letters en zonder graden.

Een graad Celsius is zo leuk omdat je de temperatuur uitdrukt in een soort percentage van een heel alledaagse schaal. Als 0 °C het vriespunt en 100 °C het kookpunt van water is, dan zijn de getallen daartussenin waarden die betekenis voor je hebben. Als je dan ook nog weet dat kamertemperatuur 20 °C is, je koelkast op 5 °C staat, dat je in water van 50 °C nog net je handen kunt houden zonder dat het pijn doet en lichaamstemperatuur iets minder dan 37 °C, dan krijg je een 'gevoel' bij die waarden.

Waarom rekenen natuurkundigen dan liever in kelvin? Omdat je dan een absolute schaal hebt die bij nul begint, en dus kunt vermenigvuldigen en delen. Als het de ene dag 10 °C is en de andere dag 20 °C, dan is het niet 'twee keer zo warm'. Maar als je in kelvin rekent, kun je wel over een verdubbeling van de temperatuur spreken bij het toepassen van de algemene gaswet.

$$pV = nRT$$

Als de temperatuur (in kelvin) 10% hoger wordt en het volume blijft gelijk, dan wordt de druk ook 10% hoger.

We rekenen makkelijker in kelvin dan in graden Celsius (of Fahrenheit of graden). En voor hoeken liever met radialen dan graden: 90° is een handig getal, want je kunt het delen door 2, 3, 5, 6, 9, 10, 15, 18, 30, 45 en 90. Dat zijn maar liefst elf getallen, terwijl je 100 maar door acht getallen kunt delen.

Maar waarom dan toch rekenen met radialen? Graden zijn toch veel makkelijker om je iets bij voor te stellen? Als 90° een rechte hoek is, dan is 81° bijna een rechte hoek: je zit maar 10% van de 90° af, daar kunnen mensen zich iets bij voorstellen. Maar wat zegt een hoek van 1,4 radialen? Dat moet je dan door π delen en met 180° vermenigvuldigen om te zien dat het tussen de 80° en 81° zit. Je kunt je daar minder goed iets bij voorstellen.

Toch doen wiskundigen het niet zomaar. Het leuke van radialen is dat je de oppervlakte van een taartpunt kunt berekenen door de hoek (in radialen) te vermenigvuldigen met de straal in het kwadraat en dat te delen door 2.

Hoe zit dat? Een hele taart is rond en heeft een oppervlakte van πr^2. Die hele taart is een 'hoek' van 360° oftewel 2π radialen. Het deel dat je daarvan hebt (de hoek α) is $\alpha/2\pi$. De oppervlakte van een taartpunt is dus:

$$A = \frac{\alpha}{2\pi} \cdot \pi r^2 = \alpha \cdot \frac{r^2}{2}$$

Oefeningen

1. Als het de ene dag 10 °C is en de andere dag 20 °C, dan is het niet 'twee keer zo warm'.
   Hoeveel procent warmer is het dan wel?
2. Jeroen Jansen wil graag een eigen temperatuurschaal bedenken. Hij twijfelt of hij de eenheden *graden Jansen* of *graden Jeroen* wil gaan noemen. In ieder geval gaat het genoteerd worden in °J. Het absolute nulpunt noemt hij 0 °J en het tripelpunt van water noemt hij 100 °J. Bereken hoeveel graden op 'zijn' schaal (°J) overeenkomt met 20 °C.

86. Discrepantie

Taylor (1997) besteedt in zijn boek een aparte paragraaf aan het begrip *discrepancy*, in het Nederlands: discrepantie. En over Nederlands gesproken: laten we eens opzoeken wat het woordenboek van *Van Dale* erover zegt.

> ***dis·cre·pan·tie***
> *zelfstandig naamwoord • de v • discrepanties*
> *1 onderlinge afwijking, het uiteenlopen*
> *2 tegenstrijdigheid, tegenspraak*

Twee verschillende betekenissen: een onderlinge afwijking is nog geen tegenstrijdigheid. Als iemand zegt dat iets nog ruim een kwartier duurt en een ander heeft het over bijna een half uur, dan zit er een afwijking tussen die onderlinge uitspraken. Maar er is nog geen sprake van een tegenstrijdigheid.

2.3 Discrepancy

Before I address the question of how to use uncertainties in experimental reports, a few important terms should be introduced and defined. First, if two measurements of the same quantity disagree, we say there is a discrepancy. Numerically, we define the discrepancy between two measurements as their difference.

discrepancy = difference between two measured values of the same quantity

More specifically, each of the two measurements consists of a best estimate and an uncertainty, and we define the discrepancy as the difference between the two best estimates.

Het is een beetje opletten met deze definitie. Want wat is het verschil tussen twee gemeten waarden? Strikt genomen is dat niet alleen het verschil tussen de waarden zelf, maar ook een berekening van de onzekerheid in dat verschil. Dat hebben we besproken in hoofdstuk 60. Als definitie houden we aan wat Taylor schrijft in de tekst achter '*More specifically, …*'

Discrepantie is het verschil tussen twee beste schattingen van twee verschillende metingen aan dezelfde grootheid.

Het voorbeeld dat Taylor noemt, geeft meer duidelijkheid.

For example, if two students measure the same resistance as follows

> *Student A: 15 ± 1 ohms*

and

Student B: 25 ± 2 ohms,

their discrepancy is

discrepancy = 25 – 15 = 10 ohms.

Recognize that a discrepancy may or may not be significant.

> O ja, dat is waar ook: Engelsen zeggen 10 ohms, Nederlanders 10 ohm.
> *Ten is also a singular, according to those strange Dutch people.*

Is er sprake van een discrepantie tussen de metingen? Ja, en die discrepantie bedraagt 10 ohm. Maar stel nu eens dat er een nog een student een meting doet. Die noemen we Student C (niet genoemd door Taylor) en die vindt:

Student C: (15,5 ± 1,5) Ω

(Student C zet altijd haakjes om de getallen, plaatst de eenheid tussen haakjes, en gebruikt het gestandaardiseerde symbool Ω in plaats van het woord *ohm*. Niet dat het fout is wat Taylor doet, maar Student C doet het netter.)

Nu de vraag: is er een discrepantie tussen de metingen van Student A en Student C. Het antwoord: ja, en die discrepantie bedraagt 0,5 Ω. Er is dus wél een *discrepantie* maar geen *onderlinge tegenstrijdigheid*. We gebruiken betekenis 1 uit *Van Dale*.

Hughes & Hase (2010) besteden ook aandacht aan *discrepantie*, maar hebben geen apart hoofdstuk met dat woord als titel en de term ontbreekt ook in de inhoudsopgave. Ook dat is niet fout, maar het vestigt minder de aandacht op de betekenis van het begrip.

3.3.4 Comparing experimental results with an accepted value

If we are comparing an experimentally determined value with an accepted value, we can use Table 3.1 to define the likelihood of our measurement being accurate. The procedure adopted involves measuring the discrepancy between the experimentally determined result and the accepted value, divided by the experimentally determined standard error. If your experimental result and the accepted value differ by:

- *up to one standard error, they are in excellent agreement;*
- *between one and two standard errors, they are in reasonable agreement;*
- *more than three standard errors, they are in disagreement.*

Hughes & Hase definiëren het begrip *discrepancy* dus niet, maar gebruiken het gewoonweg in de betekenis die Taylor hanteert en die wij in dit boek ook gebruiken.

Opmerkelijk is dat de auteurs een soort tussenvorm tussen strijdig en in overeenstemming hanteren. Beter gezegd: ze hebben gradaties in overeenstemming, namelijk *uitstekend* en *redelijk*.

Zijn ze dan niet veel te soepel? Zijn Hughes & Hase van die docenten die bij een 5,3 altijd wel een paar tienden zien te vinden? Nee, maar ze geven wel rekenschap van het feit dat metingen en onzekerheden meer met schattingen en kansen te maken hebben dan je wellicht denkt. Om een vergelijking te maken met die 5,3: een student die nét een onvoldoende haalt, beheerst de stof misschien toch wel goed genoeg. Daar staat tegenover dat je bij een student met 5,7 zo je vraagtekens kunt plaatsen. Eén vraag anders of één keer minder goed gegokt en het was onvoldoende geweest. Bij iemand die een 7,5 haalt, ben je veel zekerder van je zaak.

87. Relevant & Significant

Anne wil graag weten of het sierradendoosje dat ze kwijt is op de kast ligt. Dat kan ze niet zien, omdat ze niet lang genoeg is om zonder ergens op te gaan staan op de kast te kijken. Daarom vraagt ze het aan Wil, want die is langer. Helaas: Wil wil wel, maar hij kan het ook niet zien.

Is Wil dan niet langer dan Anne? Jawel. Maar het verschil is te klein om van belang te zijn in dit geval. Ze schelen ongeveer zes centimeter en ook als Wil nóg zes centimeter langer was, kon hij nog steeds niet op de kast kijken.

Het verschil is niet *relevant*: het maakt niet uit.

Of iets relevant is of niet, hangt meestal af van de context. Er zijn heus wel situaties waar het wél uitmaakt dat Wil langer is. Bijvoorbeeld als ze iets uit de kelder nodig hebben. Die is 1 meter 80 hoog en Anne kan daar staan zonder te hoeven bukken. Lekker makkelijk! En dus ook van belang.

Van Dale online vermeldt zowel *significant* als *relevant.*

re·le·vant (bijvoeglijk naamwoord)

1 van belang, van betekenis
maatschappelijk relevant: nuttig voor de samenleving

sig·ni·fi·cant (bijvoeglijk naamwoord, bijwoord)

1 statistisch niet aan toeval toe te schrijven en dus betekenisvol

Blijkbaar is ‘van betekenis’ niet hetzelfde als ‘betekenisvol’...
We *weten vrijwel zeker* dat Wil langer is, maar *het maakt niet uit* in de casus van de kast.

Hoe belangrijk zijn kleine verschillen in meetresultaten? Lord Rayleigh bewees eind negentiende eeuw dat een zorgvuldige analyse ervan iets moois kan opleveren. In 1892 gebruikte hij twee methoden om de dichtheid van stikstof te meten. In de ene methode werd stikstof volledig uit de atmosfeer verkregen, door lucht over gloeiend koper te laten stromen. Dat verwijderde de zuurstof. In de andere methode werd lucht eerst door ammoniak geleid en het mengsel vervolgens ook over gloeiend koper gestuurd. Zo verdween ook daar de zuurstof, maar de stikstof uit de lucht werd 'verontreinigd' met stikstof uit de ammoniak.

De stikstof uit de tweede methode was ongeveer 0,1% minder dicht dan die uit de eerste. Eén promille… dat is bijna hetzelfde als niks. Het scheelt maar 1‰.

Maar je kunt ook zeggen dat het 100% scheelt: het is *alles* meer dan 0. Of een verschil groot of klein is, hangt af van de onzekerheden in de metingen. Een kleine discrepantie kan een strijdigheid blootleggen. Dan is het verschil — hoe klein ook — best groot.

Die kleine verschillen zaten Rayleigh niet lekker. Hij deed wat een goede onderzoeker dan doet: het experiment aanpassen om het effect te vergroten. In 1894 verving hij de lucht in de chemische methode door pure zuurstof, zodat alle stikstof uit de ammoniak kwam. Deze aangepaste methode leverde stikstof op die 0,5% minder dicht was dan die uit atmosferische lucht.

De conclusie: atmosferische stikstof moest vermengd zijn met een ander gas. En dat was zo. De lucht bestaat voor 78% uit stikstof en bevat ongeveer 1,2% argon, dat 1,4 keer zo dicht is als stikstof. Rayleigh had, samen met chemicus William Ramsay, een nieuw edelgas ontdekt: argon.

Voor deze ontdekking kreeg Rayleigh in 1904 de Nobelprijs voor Natuurkunde; Ramsay kreeg die voor Scheikunde. Zijn meetonzekerheid was slechts 0,03%. Met slordiger werk was het verschil onopgemerkt gebleven. En de Nobelprijs aan zijn neus voorbij gegaan.

Oefening

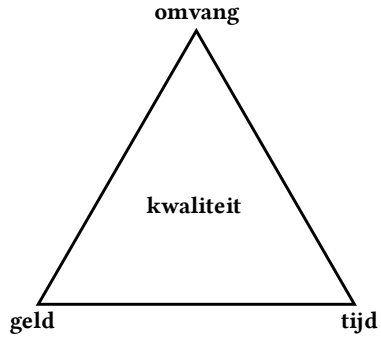
Een student haalt op een tentamen een 5,4 en twee weken later op het hertentamen een 5,5.

1. Leg uit waarom het verschil wel of niet *significant* is.
2. Leg uit waarom het verschil wel of niet *relevant* is.

88.Kwaliteit

Als je met *Google Afbeeldingen* zoekt naar 'duivelsdriehoek', kom je niet in satanische duistere hoeken van het internet terecht, maar wel in plaatjes van driehoeken waar 'tijd', 'geld' en 'kwaliteit' bij staan, zij het niet altijd bij dezelfde hoeken. Sterker nog: ze staan niet altijd bij de *hoeken*, maar soms bij de *zijden*. Op *Wikipedia* vind je een plaatje.

EEN PLAATJE DAT NIET EENS EEN VECTORFORMAAT IS. FOEI!
Zou het daarom een duivelsdriehoek heten? Er wordt in de tekst verteld dat het ook wel een *project management triangle* genoemd wordt. Zonder duivel erin dus, maar wel met een extra woord: *scope*. Laten we dat plaatje maar overnemen. (En eerst in het Nederlands vertalen.)



Figuur 88.1 Projectmanagementdriehoek

Het verhaal achter het plaatje is: als je kwaliteit wilt verhogen, zul je meer geld moeten uitgeven. Of extra tijd moeten besteden. Of minder doen (een kleinere *scope*), zodat je meer tijd en geld hebt voor wat je overhoudt.

In plaats van 'geld' (*money*) zou je ook 'kosten' (*cost*) kunnen zeggen. En tijd is tijd (*time*). Het Engelse woord *scope* is iets lastiger. *Van Dale* zegt: gebied, omvang, terrein, sfeer, gezichtsveld, reikwijdte, draagwijdte. Maar we snappen heus wel wat er bedoeld wordt. 'Kwaliteit' (*quality*) is het lastigste in dit plaatje, omdat het een veel te vaag of te veelomvattend begrip is.

Mocht de lezer een student zijn die nog wil afstuderen (die kans is vrij groot, gezien de doelgroep), let dan tijdens de rest van dit hoofdstuk goed op. Wie weet, krijg je de auteur van dit boek wel als examinator. Die gaat vervelende vragen stellen als je zegt dat iets 'beter' is dan iets anders, of dat je 'de kwaliteit wilt verhogen'. Het woord 'kwaliteit' zegt namelijk weinig. Iets is alleen maar beter dan iets anders als je weet wat je doel is. *Quality is fitness for use*, zeggen de Engelsen. (Sommige dan.)

Een handig hulpmiddel om zinige uitspraken te doen over kwaliteit, is de standaard ISO 25010. Die beschrijft namelijk 31 *kwaliteitskenmerken van software en systemen*. Wanneer is software 'goed'? Wanneer is het 'beter'? Het is zinvoller om te spreken van 'sneller', 'beter onderhoudbaar', 'beter leesbaar', 'veiliger', 'foutbestendiger'.

Meten & Weten gaat over *meten*. (En over *weten*.) Kun je zeggen dat de ene meting *beter* is dan de andere? Ja, maar dan moet je wel weten wat je doel is. De ene meting is sneller, makkelijker, prettiger of goedkoper uit te voeren dan de andere. Nauwkeuriger of preciezer. Minder foutgevoelig en...

Het gaat, zoals elders in dit boek benoemd wordt, om de vragen 'Wat, Hoe en Waarom?' Je zult vaak moeten weten *waarom* je iets meet om te weten hoe je het maar beter kunt meten.

> Een meting kan effectief zijn, of efficiënt. Soms allebei, soms geen van beide. Het is een verschil dat je bij verslagen, projecten en metingen vaak tegenkomt.
>
> **Effectief** betekent dat het doel bereikt is.
>
> **Efficiënt** betekent dat het doel bereikt is op een zuinige manier: snel, goedkoop of met weinig moeite.
>
> Je kunt dus prima iets heel efficiënt doen wat totaal niet werkt. Je kunt ook effectief zijn op een omslachtige manier. Stel: je wilt de temperatuur van een ruimte weten. Vijf studenten meten elk op een ander tijdstip de temperatuur met een thermometer, en je neemt het gemiddelde. Je komt redelijk uit bij de juiste temperatuur. Effectief dus. Maar efficiënt?
>
> Andersom: je hebt een digitale sensor die je elke seconde een waarde geeft, maar die blijkt structureel 1,5 graad te hoog te meten. Efficiënt, maar niet effectief.
>
> In deze zin is het streven naar 'kwaliteit' dus dubbelzinnig. Bedoel je dan een meting die klopt (effectief)? Of een meting die prettig werkt (efficiënt)? In het ideale geval beide, maar pas op dat je ze niet verwart.

Oefeningen

1. Noem tien verschillende kwaliteitsattributen uit de ISO-norm.
2. Hoeveel kwaliteitsattributen bevat die ISO-norm?
3. Welk nummer heeft die ISO-norm?

89. Falsifieerbaarheid

Falsifieerbaarheid of falsificeerbaarheid is een eigenschap van een wetenschappelijke of andere theorie, en wel dat er criteria kunnen worden aangegeven op grond waarvan de theorie zou moeten worden verworpen (niet te verwarren met een theorie die daadwerkelijk is verworpen).
Het falsificationisme is een wetenschapstheorie bedacht door Karl Popper, waarin falsificatie centraal staat.

Bron: Wikipedia

Wat heeft falsifieerbaarheid met Meten & Weten te maken? In ieder geval de gedachte dat je niet zo makkelijk kunt vaststellen dat je iets *weet*, of beter gezegd: dat je *weet dat iets waar is*. In de wetenschap, softwareontwikkeling en het schrijven van teksten heb je zelden de mogelijkheid om te bewijzen dat iets klopt, laat staan voor 100% foutloos is.

Popper stelde dat je in de wetenschap de dingen zo moet formuleren dat het in principe mogelijk is om de uitspraak te weerleggen als die niet waar is. Een klassiek voorbeeld is: "Alle zwanen zijn wit." Dat is te weerleggen door één zwaan te vinden die niet wit is. Ga naar het park, zoek naar zwanen en kijk welke kleur ze hebben. Als je een witte vindt, bewijst dat niets. Als je twintig of honderd witte vindt, ook niet. Naarmate je er meer vindt, wordt het *aannemelijker* dat de stelling klopt, maar *bewezen* is het niet. Zodra je één zwaan vindt met een andere kleur, is het klaar. Dan klopt de stelling niet.

> Een alternatief is: ga alle voorwerpen en levende wezens af die niet wit zijn. Zodra je er een zwaan tussen treft, is ook bewezen dat de stelling niet waar is.

De stelling klopt ook niet trouwens. Er bestaan ook zwarte zwanen.

Nu de stelling: "Er bestaan ook paarse zwanen."

Is die te bewijzen? Ja, als je één paarse zwaan vindt, is het nog een twijfelgeval, want de stelling formuleert een meervoud. Maar bij twee is het al bingo: dat is voldoende voor een bewijs.

Is die laatste stelling falsificeerbaar? Nee. Wie zich suf zoekt naar zwanen en nergens een paarse ziet, gaat er misschien wel aan twijfelen, maar bewezen is er niets. Het vervelende is: al zou je een leven lang dagelijks overal naar zwanen speuren en je treft nooit een paarse aan, dan heb je nog steeds niet bewezen dat er geen zijn. Misschien bevinden ze zich altijd in een ander land dan jij. Dan kom je ze nooit tegen, al trok je de hele wereld rond.

Het verschil tussen de eerste en tweede stelling zit niet zozeer in de kleur, maar in ‘alle’ en ‘eentje’. De stelling dat alle zwanen wit zijn, kun je weerleggen door één tegenvoorbeeld. De stelling dat er minstens één bestaat die paars is, kun je alleen maar weerleggen als je *alle* zwanen op de hele wereld gezien hebt. Dat lukt je niet.

Wetenschappers formuleren stellingen die *falsificeerbaar* zijn. Vervolgens doen zij hun best om te *weerleggen* dat ze zelf gelijk hebben. Dat gaat een beetje tegen je gevoel in, want meestal willen mensen gelijk hebben. Maar de kans dat je gelijk hebt wordt juist groter als je een uiterste inspanning levert (samen met anderen) om je *ongelijk* te halen en daar niet in te slagen.

Softwareontwikkelaars doen dat (als het goed is) ook. Ze schrijven code en in plaats van heel hard “Het werkt!” te roepen (dat doen de verkopers wel), doen ze hun uiterste best om een fout te vinden. Gewoon door creatief te zijn in het bedenken van gevallen die mis zouden kunnen gaan. Een tester doet niet zijn of haar best om aan te tonen dat het programma geen fouten bevat. Hij of zij zet zich (binnen zekere grenzen van tijd en middelen) zo goed mogelijk in om juist wél fouten te vinden. En als dat dan echt niet lukt, is het programma misschien wel goed.

Tekstschrijvers doen het (als het goed is) ook. Ze schrijven een tekst en in plaats van heel hard “Hij is foutloos!” te roepen, speuren ze de regels en alinea’s af om te kijken of er niet een woordje of lettertje ontbreekt. Pas als dat na lang zwijgen, zweten en zwoegen niet lukt, is er een reële kans dat de tekst vrij goed is.

Het lastige van metingen en onzekerheden is dat we nogal eens uitgaan van een normale verdeling. Met een lengte van 0,80 meter en een onzekerheid van 2 cm bedoelen we vaak dat de kans groot is dat de echte lengte tussen 78 en 82 cm ligt, en dat de kans heel groot is dat die tussen 76 en 84 cm ligt. Een grens met een absolute zekerheid is er wiskundig gezien niet.

90. Inschatten wat verwaarloosbaar is

Bij Proeflokaal vandeStreek in Utrecht hebben ze de bierprijs ook naar boven moeten bijstellen. De kosten voor mout zijn flink omhoog gegaan en ***ook het papier voor etiketten*** *is duurder geworden, zegt commercieel directeur Ronald van de Streek. "In de afgelopen twaalf maanden zijn de kosten zo'n 50 procent gestegen. Toch is het ons tot nu toe gelukt om de prijs van bier slechts met 5 tot 6 procent te verhogen, terwijl de grote merken wel 15 procent duurder werden. Daarmee zijn we in prijs dichter bij de grote brouwerijen in de buurt gekomen."*

Bron: Trouw, 27 januari 2023

De prijs van een fles bier hangt van allerlei factoren af. De kosten van mout en gerst bijvoorbeeld, maar ook van water. En van glas en papier, en metaal voor de kroonkurk. Arbeidsloon en energiekosten zijn ook al van invloed, net als accijnzen, reclamespots en transportkosten.

Bij veel berekeningen verwaarloos je bepaalde invloeden. Dat is handig, omdat het alles wat *eenvoudiger* maakt. Verwaarlozing zorgt ervoor dat je snel kunt *schatten*. En hoewel schatten veel minder zekerheid verleent dan berekenen, geeft het door de eenvoud en snelheid wel veel inzicht.

Om te weten hoe de kosten van een flesje bier afhangen van de stijging van de papierprijs, moet je een aantal dingen weten.

1. De prijs van papier.
    a. Hoeveel was het eerst?
    b. Hoeveel is het nu?
2. Hoeveel papier er op één fles zit.
    a. De oppervlakte van de etiketten bij elkaar.
    b. Het gewicht (de massa) per oppervlakte.

'Ergens op het internet' staat dat papier in prijs stijgt van € 50 naar € 75 per ton en dat de afmetingen van een etiket 180 mm × 90 mm zijn. Het etiket op de achterkant is meestal kleiner en vaak zit er nog wel wat papier om de hals. Een stijging van € 50 per ton en een oppervlakte van 500 cm^2 is dus een ongunstige schatting. Het papier zal minder wegen dan 100 g/m^2.

Laten we werken met wetenschappelijke notatie: eerst alles omzetten euro's en SI-eenheden, machten van 10 en één significant cijfer per grootheid.

Prijsstijging:	€ 50 per ton	=	5×10^{-2} euro/kg.
Oppervlakte van het papier:	0,05 m^2	=	5×10^{-2} m^2
Massa papier per oppervlakte:	0,1 kg/m^2	=	1×10^{-1} kg/m^2

Het lijkt overdreven om deze stappen zo lang gescheiden te houden en zo strikt in eenheden en machten van 10 te noteren, maar naarmate je daar voorzichtiger in bent, is de kans op fouten kleiner. Een nulletje meer of minder, de komma verkeerd zetten: op een tentamen zijn het tienden in je cijfer. In de werkelijkheid van de ingenieur kan het tonnen aan schade betekenen. Getallen met vijf of meer cijfers (en een euroteken) voor de komma.

Nu kunnen we deze drie getallen met elkaar vermenigvuldigen. We komen dan uit op $25 \times 10^{-5} = 2{,}5 \times 10^{-4}$ euro.

In een artikel, verslag, e-mail of app-bericht is het handig om nog één slag te maken, namelijk je afvragen hoe het voor de doelgroep het meest overtuigend overkomt. Die $2{,}5 \times 10^{-4}$ euro is voor veel lezers niet handig. Je kunt die noteren als ‘minder dan 0,1 eurocent’ of als € 0,001.

Hoe kwamen we hier ook alweer op? Als je een antwoord hebt, moet je altijd even bedenken wat de eigenlijke vraag ook alweer was. Dat was *niet* wat een los etiket kost; het ging om de invloed van de prijs*stijging* van het papier op een fles bier. En de concrete vraag was of het *verwaarloosbaar* is of niet.

Het antwoord: ja, dat is verwaarloosbaar. Tenzij je héél goedkoop bier hebt.

Oefening

1. Stelling: “Dit boek bevat 292 pagina’s.”
    a. Is deze stelling juist?
    b. Wat denk je dat de auteur gedaan heeft om het (al dan niet) juiste aantal in de stelling op te nemen?
    c. Als er op elke drie pagina’s van dit boek één spelfout staat, hoeveel procent van de woorden bevat er dan een spelfout?

2. Dit boek bevat 292 pagina’s en is tegen kostprijs te koop. Dat bedrag is te berekenen via de website van Pumbo.nl. Maak een schatting van de bijdrage van de papierprijs in de kostprijs van dit boek en geef aan of dit verwaarloosbaar is.

91. Vroeg vinden van verbeter(d)ingen

Software-ontwikkelaars zijn bekend met de wetmatigheid dat een fout duurder wordt naarmate je die later vindt. Je tikt een regel code in, kijkt ernaar, en ziet dat er iets niet correct is. De schade valt dan mee. Een keer met je muis klikken en de *Backspace* indrukken en je bent er al bijna. Dat kost een minuut van je tijd en dus weinig geld.

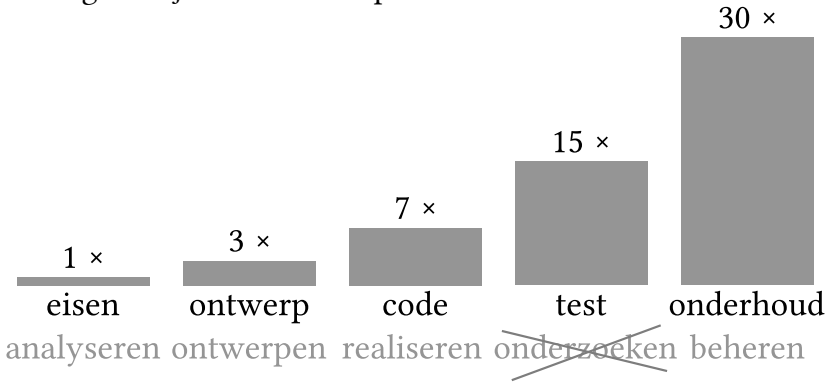
Als je de fout vindt tijdens de controles die je doet voordat je de software oplevert aan een tester, is het al wat duurder. Je moet dan langer zoeken naar waar het zit. Je bent dan gelukkig nog steeds de enige die er tijd aan kwijt is.

Als de tester 'm vindt, gaat er meer tijd in zitten. Die heeft er zelf werk aan en moet ook nog eens een *bug report* schrijven. Daar moet de programmeur dan ook weer wat aan toevoegen en bij de volgende oplevering aan de tester moet bevestigd worden dat de bug eruit is.

Als een tester een bug vindt, moet je nooit boos op hem of haar worden. Een dikke zoen is — als de persoon in kwestie dat op prijs stelt — een betere reactie. Een bosje bloemen, een kop koffie, een doos *Merci*: wees blij dat de tester die fout gevonden heeft.

> Natuurlijk kost het je tijd om dan die fout te gaan oplossen. Maar het is niet zo dat de fout er niet in zou zitten als hij niet gevonden was. De honderden of duizenden gebruikers van je software zijn veel meer uren met dat programma bezig. Als die de fout vinden, zijn ze soms niet blij. Omdat ze er last van hebben. Ze kunnen een schadeclaim indienen. Of naar de concurrent overstappen. In het gunstigste geval melden ze het beleefd. En dan ben je veel tijd kwijt om het alsnog op te lossen.

Er zijn wel onderzoeken naar die toename van de kosten gedaan. Die resulteren in grafiekjes waar een exponentiële toename in zit.



Dit plaatje is gebaseerd op gegevens van het *National Institute of Standards and Technology*. De rechter kolom zou eigenlijk hoger moeten zijn. Het NIST stelt namelijk dat die laatste een factor 30 tot 100 groter is dan de eerste.

> Onder de bekende fases van software-ontwikkeling heb ik de competenties van een mechatronicus gezet. Het klopt niet één-op-één en voor testen is er niet een aparte competentie. Eigenlijk hoort testen gewoon bij het realiseren (schrijven) van code.

Geldt wat hier over software staat ook voor metingen? Zijn de kosten van een meetfout ook hoger als je die niet meteen zelf ontdekt? Ja. Het is alleen wat lastiger in kaart te brengen.

- Wat als je een fout maakt in een meting of een berekening of een bepaling van een onzekerheid en het wordt niet ontdekt voordat je afstudeerverslag ingeleverd wordt, maar tijdens de examenzitting?
- Wat gebeurt er als een metalen constructie veel te sterk wordt uitgevoerd als gevolg van een verkeerde berekening?
- Wat zijn de gevolgen als een patiënt een factor duizend te veel morfine krijgt toegediend? (“Echt, alleen de komma stond verkeerd! Op school kreeg je dan een paar tienden puntenaftrek…”)

Er zit ook een andere kant aan het verhaal. Ja, een fout is inderdaad goedkoper op te lossen naarmate je die eerder vindt. Maar er is ook een soort omgekeerde wet.

Door heel even een stuk programmacode, een tekst of een berekening na te kijken, vind je al heel snel een fout.

- Door langer te kijken vind je meer fouten, maar niet heel veel meer.
- Door heel lang te kijken vind je er nog meer.
- Om de laatste fout te vinden, moet je extreem lang zoeken.

Hoe eerder je de fout vindt, hoe goedkoper het is. Maar als je foutloos wilt werken, gaat je werk wel heel erg duur worden.

> Een student zei eens: liever een zes zonder stress, dan een zeven zonder leven. Dat rijmt. Of de afweging terecht is, is een andere vraag. Studenten die altijd maar op een 6 mikken, komen in hun afstudeerwerk minder beslagen ten ijs dan anderen die probeerden om minstens een 8 te halen. Die gaan met minder stress hun examenzitting in. Hebben gewoon meer bagage dan anderen.
>
> Van wie gaat voor een acht, kan veel goeds worden verwacht.

92. Pas op

Het schijnt zo te zijn dat veel lezers over het hoofd zien dat ‘pas op voor de de hond’ een fout bevat.

Antwoorden

Hoofdstuk 0

1. Elke meting heeft een onzekerheid. Daar moet je een schatting van geven, omdat de uitkomst van de meting anders weinig betekenis heeft. Nadenken over hoe je die onzekerheid kleiner kunt krijgen, kan nuttig of zelfs nodig zijn.

2. Niet alles in één keer lezen. Oefeningen doen. Erover praten met anderen, de discussie in de lessen volgen.

Hoofdstuk 1

1. Nee. Als de meter in 1986 een centimeter *langer* was dan tegenwoordig, zou dat wél een verklaring kunnen zijn waarom jongeren nu zo'n anderhalve centimeter langer zijn volgens 'de nieuwe meetlat'. Als een meter vroeger 'meer' was, was je zelf volgens die meetlat korter.

2. In 1983.

3. Met het schaamrood op de kaken: van *Wikipedia*.
   Dat zal heus wel kloppen, maar kun je in je afstudeerverslag niet maken en dus als docent in een les of boek ook niet doen.
   Op de *Wikipedia*-pagina zelf kun je wel betrouwbare bronnen vinden.

4. Nee. De schrijver bedoelt een afwijking van 2σ in beide richtingen.
   Je mag maximaal 2σ 'te klein' zijn en maximaal 2σ 'te groot'.

5. Twee aspecten:
   a. De waarde 1,414 213 562 — het zijn trouwens de eerste decimalen van $\sqrt{2}$ — heeft veel meer cijfers dan redelijk is. 'Anderhalve seconde' zou zinniger zijn.

   b. Dat je de galm wel of niet meerekent, moet je vermelden om een definitieprobleem te voorkomen.

6. Het spreekt niet vanzelf, een eenheid moet je altijd vermelden. Ook als die voor de hand ligt.

Hoofdstuk 2

1. Verbeteringen:

   a. "Om met 100% zekerheid te kunnen garanderen dat iemand te hard heeft gereden" moet zijn: "Om te kunnen garanderen dat iemand *alleen wordt bekeurd als* hij te hard heeft gereden"

   b. "trekt de politie van de gemeten snelheid een paar kilometer af" moet zijn "trekt de politie van de gemeten snelheid een paar km/u af"

   c. "3 kilometer meetcorrectie" moet zijn: "3 km/u meetcorrectie"

   d. "Er geldt een ondergrens van 4 km/u voordat wordt bekeurd" moet zijn: "Er wordt pas bekeurd als er minstens 4 km/u te hard wordt gereden."

   e. Het is niet duidelijk of een snelheid precies op de ondergrens als te snel wordt beschouwd.

f. Strikt genomen wordt er niets gezegd over een snelheid van exact 100 km/u, alleen over meer of minder.
Logisch gezien is dat geen probleem, want 3% is bij die snelheid precies gelijk aan 3 km/u.

2. Dit is een gemene vraag, maar er staan in Meten & Weten wel vaker gemene vragen, omdat het om de *oefening* gaat en niet om het goed hebben van het antwoord. Het is geen tentamen: daar horen geen gemene vragen te worden gesteld.

   Wat is er gemeen aan?

   a. Er staat niet bij wat de maximumsnelheid is, dus je kunt de vraag helemaal niet beantwoorden. Sommige lezers zullen automatisch uitgegaan zijn van 100 km/u als maximumsnelheid, maar het is niet gegeven.

   b. Er wordt gesproken over *iemand die 107,2 km per uur rijdt.*
   Als dat de *echte* snelheid is, hoeft er niet gecorrigeerd te worden. Als dat de *gemeten* snelheid is, moet er eerst 3% worden afgetrokken. Maar er staat niets over 'gemeten'.
   Mocht je de vraag toch zo opvatten: na correctie van 3% kom je op 103,984 km/u. Dat is (bij een maximumsnelheid van 100 km/u) minder dan 4 km/u te snel.

Hoofdstuk 3

1. Alle sleutels zijn even waarschijnlijk. De kans dat de eerste goed is, is 1/8. De kans dat een van de eerste twee goed is, is 2/8. Een van de eerste drie: 3/8. Een van de eerste vijf: 5/8. (Een van alle acht: 8/8).

2. 8: de kans is 1/8, dat kun je 'klein' noemen (en zeer klein als het 5% is: 1/20)
   2: de kans is 1/2, informeel: 'half kans' of 50%
   ∞: de kans is nul, maar het is niet onmogelijk
   1: de kans is 1: het is zeker
   0: er is geen sprake van een kans, want er is geen eerste sleutel

3. Het moet toch gek lopen willen alle 389 studenten op een *andere* dag jarig zijn… Er zijn maar 366 verschillende verjaardagen mogelijk. We kunnen 367 of meer studenten meer zo verdelen dat iedereen op een andere dag jarig is. Zoveel dagen zijn er namelijk niet. (De kans dat iemand op dezelfde dag *als jij* jarig is, is overigens wel kleiner dan 100%. Maar dat was de vraag niet.)

Hoofdstuk 4

Belangrijk is de vraag of je een van beide partijen bevoordeelt. Wat is erger? Dat het een doelpunt was en toch niet telt, of dat het geen doelpunt was terwijl hij wel telt?
Hoe dan ook gelden de regels / het systeem voor beide ploegen.

Hoofdstuk 10

1. Thermometers 6 en 7 laat je buiten beschouwing, want die zijn niet goed. Van de andere vijf zou je in het beste geval het gemiddelde kunnen nemen. Dat verkleint de

toevallige fout. Eenvoudiger (en praktisch even goed) is de middelste thermometer nemen. Die heeft ongeveer het gemiddelde vloeistofniveau.

2. Deze vraag kun je alleen maar beantwoorden als je meer weet over de tolerantie die gekozen is. Als de fabrikant weet dat er 5% spreiding in kan zitten en dat acceptabel vindt én in de specificaties vermeldt, dan is er sprake van een onzekerheid. Als de norm is dat er maar 2% afwijking mag zijn, dan valt deze buiten de norm. Dan hoort het product niet verkocht te worden.

Hoofdstuk 11

1. Zie figuur op Nederlandstalige pagina van *Wikipedia*.
2. Zie figuur op Engelstalige pagina van *Wikipedia*.
3. Gevraagd wordt om een afronding op gehele inches. Dat is 11 802 852 677, een getal op met elf cijfers. Weliswaar afgerond, *maar dat was de vraag ook*. Deze elf cijfers zijn exact.
4. Het kan natuurlijk niet helemaal. Op twee na kun je de letters een plekje geven. Je zou van wat je over hebt (de d en de g) een c en een o kunnen maken met wat krassen. (Misschien een flauwe oefening, maar nu vergeet je het tenminste nooit meer.)
5. Strikvraag! Er wordt gevraagd naar een *nauwkeurigheid*, maar wat geven is, is de *onzekerheid* van twee multimeters. Die onzekerheid bestaat uit een toevallig deel en een systematisch deel en je weet niet welke van beide het meeste bijdraagt. Dus kun je ook niet veel zeggen over de nauwkeurigheid. (Hooguit dat die niet groter kan zijn dan een bepaalde waarde. Voor de tweede is de onzekerheid kleiner, dus als je moet gokken, neem je die. Maar je weet het niet.)
6. Strikvraag! Er wordt alleen iets gezegd over de spreiding. Die uitspraken hebben dus betrekking op de precisie, niet op de nauwkeurigheid.

Hoofdstuk 14

1. Ook bij een analoge klok is het aantal cijfers of decimalen eindig. Het laatste cijfer kun je schatten, maar daarna komt er niet nóg een cijfer. De mogelijke waarden die je kunt aflezen zijn dus discreet.
2. Of je kunt schatten tussen de secondestreepjes hangt af van de vraag of de secondewijzer continu loopt of met stapjes verspringt.

Hoofdstuk 15

1. De onzekerheid bedraagt 0,75 V. De werkelijke waarde ligt namelijk tussen 482,5 V en 483,9999… V.
2. De spanning is 483,25 ± 0,75 V, wat in het juiste aantal significante cijfers gelijk is aan 483,3 ± 0,8 V.

Hoofdstuk 17

1. Je kunt het zeggen, maar er is nogal wat op aan te merken.

a. Om met de belangrijkste te beginnen: een Celsiusschaal is geen ratioschaal. We vergelijken de waarden 293,15 K en 293,35 K met elkaar. Dat verschilt niet 1%, maar slechts 0,07%.
b. De gegevens zijn onzorgvuldig en onvolledig geformuleerd. Er wordt niet eens vermeld of de meting op hetzelfde tijdstip is uitgevoerd. Het kan bovendien zo zijn dat de minimumtemperatuur *lager* is en de maximumtemperatuur *hoger*. Bij helder weer is dat mogelijk. Vaak kijken we naar het maximum, maar het gemiddelde ligt meer voor de hand. Dat het 'vandaag' warmer is, kun je ook zeggen na een heel koude nacht.
c. Of het *warmer* is, is niet alleen een kwestie van temperatuur, maar ook van hoeveel zonneschijn er is, de windkracht, luchtdruk en luchtvochtigheid.
d. Er is geen uitspraak gedaan over de onzekerheid in de genoemde temperaturen. Op een huis-, tuin- en keukenthermometer zal een verschil van 0,2 °C niet significant zijn.

2. In symbolen kun je zeggen: A > B, B > C, C > A.
Wiskundig gezien kan dat niet. Als A groter is dan B en B is groter dan C, dan moet A groter zijn dan C.

   Er is bij een schaakwedstrijd sprake van een *ordinale* schaal. Je kunt winnen of verliezen, de wedstrijd kan ook in remise eindigen: gelijk. Het probleem met deze ordinale schaal is dat die alleen maar geldt *voor die ene wedstrijd*. Je kunt niet zomaar zeggen dat de ene schaker 'beter' is dan de andere. Tenzij die ene altijd wint.

3. Een goede analyse begint met de harde cijfers.
Kees 37%, Piet 35%, Jan 27%.
In totaal is dat 99% en geen 100%.
Er zijn twee mogelijke oorzaken:

   a. De percentages zijn afrondingen van (bijvoorbeeld) 37,2%, 35,5% en 27,3%. De som daarvan is wel 100%.
   b. Er kunnen blanco stemmen zijn. Die tellen wel mee in het percentage (in tegenstelling tot ongeldige stemmen), maar worden niet uitgebracht op de drie kandidaten.

   Tweede ronde:
   Kees 41% (4% meer) en Jan 59% (32% meer)

   Hoe kan het dat Piet in de tweede ronde veel meer stemmen kreeg dan in de eerste, terwijl dat verschil bij Kees zo klein is? Een mogelijke verklaring is dat de mensen die hun stem uitbrachten het liefst stemden op iemand uit een bepaalde plaats. Wie een voorkeur had voor een inwoner van Tilburg moest kiezen tussen twee kandidaten. Er waren mensen die op Piet stemden terwijl ze liever Jan hadden dan Kees. In de eerste ronde was dat niet zichtbaar, in de tweede wel.

   De stemming was een verkapte verkiezing tussen Eindhoven en Tilburg. Een binaire en ordinale schaal liepen door elkaar heen.
   Theoretisch is het mogelijk dat Jan in een tweede ronde óók gewonnen zou hebben als hij de enige kandidaat tegenover Kees was geweest. Sterker nog: misschien zou Jan wel gewonnen hebben als hij de enige kandidaat tegenover Piet was geweest! Als alle Eindhovenaren gedacht zouden hebben: 'Dan altijd nog liever Jan…'

Hoofdstuk 18

1. We hebben eigenlijk veel te weinig informatie, maar laten we uitgaan van een standaard meetlint / rolmaat. Die heeft een millimeteraanduiding. Theoretisch zou je dan op halve millimeters kunnen aflezen. Dat doe je aan twee kanten (bij de nul en bij de waarde die je afleest) en de afleesfout zou dus 1 mm zijn. Lukt dat over een afstand van ruim 2,5 m? Je hebt misschien ook nog te maken met een ongelijke ondergrond. Daarom is 2 mm of zelfs 5 mm reëler.
   In de praktijk zou het aardig zijn om verschillende mensen dezelfde meting te laten doen. De spreiding geeft dan een aardig beeld van de werkelijke afleesfout.

2. De steekproef-standaardafwijking is afgerond op drie significante cijfers gelijk aan 0,0626 m. Het is gebruikelijk om die als aanduiding voor de spreiding in de meting zelf te geven.
   Je ziet in deze opgave (en in dit specifieke voorbeeld) dat de spreiding in datgene wat we meten een factor tien groter kan zijn dan de onzekerheid ten gevolge van de meting zelf.

Hoofdstuk 20

1. Het schatten van de hoek is niet eens zo moeilijk. Gewoon een geodriehoek erlangs leggen is een optie. Beter is om de hoogte en de breedte van de driehoek te meten. Dat is makkelijker aflezen, maar je moet dan nog wel een rekenmachine pakken voor die hoek.
   Als je gebruikmaakt van een PDF (geen gedrukt boek of print) zou je ook op *Print Screen* kunnen drukken en de schermafdruk zodanig uitknippen dat je de hoogte en breedte in pixels kunt opvragen. Mijn eigen aanpak leidde tot 615 × 188 pixels.
   Lastiger is de vraag wat de onzekerheid is. Dat is de essentie van deze oefening. Iets zo goed mogelijk doen is makkelijker dan weten hoe goed je het doet. Zelf denk ik dat drie pixels in hoogte en breedte redelijk is. Iemand anders zal één pixel zeggen, of vijf…

   De twee uiterste hoeken krijg je dus voor
   - 612 × 191 pixels ⟶ $\tan\theta = 191/612 \longrightarrow \theta \approx 17{,}3327°$
   - 618 × 185 pixels ⟶ $\tan\theta = 185/618 \longrightarrow \theta \approx 16{,}6652°$

   De helft van het verschil tussen deze waarden zou een redelijke schatting zijn van de onzekerheid. Afgerond op één significant cijfer kom je dan op 0,3°.

2. Deze vraag is eigenlijk niet te beantwoorden. Als onze reactietijd altijd exact 0,3 seconde zou zijn, dan ben je zowel bij het starten als bij het stoppen van de tijdsmeting exact 0,3 seconde te laat. Het tijdsinterval zou dan geen onzekerheid hebben: die twee keer te laat drukken heft elkaar precies op.
   Eigenlijk zou je moeten weten hoe groot de *spreiding* is in onze reactietijd. Die kun je bepalen door de meting vaak te herhalen.

Hoofdstuk 32

1. De beredenering is niet eens zo heel eenvoudig. Uitschieters hebben namelijk een heel groot gewicht. Die ene 10 (uitslover!) wijkt enorm af van al die enen. Het

gemiddelde is grofweg een 1,5. (Precies: 1,45.) Daar zitten bijna alle cijfers nog geen halve punt vandaan. Maar die ene 10 zit er heel veel bij vandaan, en die gaat kwadratisch. Uitrekenen leidt tot (in twee decimalen) 1,96.

2. De hoogst mogelijke standaardafwijking bij twintig cijfers krijg je als de ene helft van de klas een 1 had en de andere helft een 10. Dan wijken alle cijfers veel af van het gemiddelde. Kun je dat bewijzen? Je zou kunnen kijken wat er gebeurt als er 11 enen waren en 9 tienen, of omgekeerd.
3. De standaardafwijking bij één cijfer is gelijk aan 0. Dat heeft wel een *zinnige*, maar geen *zinvolle* betekenis. In je eentje ben je altijd gelijk aan het gemiddelde. Daar wijk je nooit van af.

Hoofdstuk 33

1. Geen uitwerking: controleer of je spreadsheet klopt.
2. =GEMIDDELDE(A:A)
   Deze formule bepaalt het gemiddelde van een hele kolom. Daardoor is voldaan aan de eis dat het aantal elementen in kolom A veranderd kan worden.
3. De *Excel*-formules en waarden staan in onderstaande tabel.

=GEMIDDELDE(A:A)	23
=SOM(A:A)	322
=AANTALARG(A:A)	14
=MEDIAAN(A:A)	23
=MODUS.ENKELV(A:A)	23
=MAX(A:A)	24
=MIN(A:A)	21
=ALS(B6-B1>B1-B7;B6-B1;B1-B7)	2
=GEM.DEVIATIE(A:A)	0,714285714
=AANTAL.ALS(A:A;">"&B1)	5

4. De veranderingen staan in onderstaande tabel.

gemiddelde	blijft 23
som	wordt 345 (23 meer)
aantal waarden	wordt 15 (1 meer)
mediaan	blijft 23
modus	blijft 23
grootste waarde	blijft 24
kleinste waarde	blijft 21
grootste afwijking van het gemiddelde	blijft 2
gemiddelde absolute afwijking	wordt 0,666667
aantal waarden groter dan gemiddelde	blijft 5

5. Het gemiddelde, de som, de kleinste waarde en de gemiddelde absolute afwijking veranderen.
   Eén waarde verandert spectaculair: het aantal waarden groter dan het gemiddelde wordt ‘opeens’ gelijk aan 11, terwijl het 5 was. Dat lijkt vreemd, maar het gemiddelde wordt een klein beetje kleiner. Daardoor zijn de zes waarden van 23 ‘opeens’ groter dan het gemiddelde van 22,8.

6. Als de kans dat iets gebeurt P is, is de kans dat het *niet* gebeurt 1 – P. De kans dat *iemand* iets heeft of doet, bereken je door eerst te bepalen hoe groot dat kans is dat *niemand* iets heeft of doet. Die kans is $(1 - P)^N$ als N het aantal personen is.
   Als we rekenen met $P = 10^{-9}$ en $N = 8 \times 10^9$, vinden we dat de kans dat niemand iets heeft of doet maar 0,034% is. (Als we hadden gerekend met één miljard mensen, zouden we op ongeveer 37% zijn uitgekomen.)
   De kans dat iemand het heeft of doet is dus ruim 99,9%.

Hoofdstuk 34

Je zou een byte kunnen splitsen en de eerste vier bits gebruiken voor het cijfer voor de komma en het laatste deel van vier bits voor het cijfer na de komma. Met vier bits kun je zestien cijfers aan. Je hebt er tien nodig voor de komma (1 t/m 10) en tien na de komma (0 t/m 9).

Hoofdstuk 38

1. Er zijn 36 mogelijkheden met een gelijke kans. Bereken het kwadraat van elk van die mogelijkheden en bepaal het gemiddelde. Dat geeft (in twee decimalen) 54,33.

   De verwachtingswaarde van het totaal per worp is overigens 7. Het kwadraat van de verwachtingswaarde is dus niet gelijk aan de verwachtingswaarde van het kwadraat. Om over na te denken: de verwachtingswaarde bij één dobbelsteen is 3½, maar je kunt nooit 3½ gooien.

2. Het klopt dat in de ene envelop het dubbele zit van het bedrag van de andere. Maar dat is dan wel het dubbele *van een ander bedrag*. In de ene envelop zit X euro en in de andere 2X euro. Je weet niet welke van de twee je hebt. Heb je die met 2X, dan verlies je X door te ruilen. Heb je die met X, dan win je X door te ruilen.
   In de 'hoogste' envelop zit het dubbele van X. In de 'laagste' zit de helft van 2X.
   Je wint dus X óf de helft van 2X.

3. De verwachtingswaarde is gelijk.
   In beide gevallen het dubbele van één keer meedoen met één loterij.

4. De vraag is niet te beantwoorden, omdat hij niet goed gedefinieerd is. 'Bij vijftig loten' kan op twee manieren worden geïnterpreteerd:
   a. als je vijftig loten *koopt* (en niet één of twee)
   b. als er vijftig loten *te koop zijn* (en niet honderd)
5. Bij 0 en 100: geen enkel lot, of alle loten. In beide gevallen ligt de uitkomst vast en is de winst geen kansvariabele meer.

6. De verwachtingswaarde van de winst per lot is vast. *De verwachtingswaarde van de totale winst* hangt dus af van het aantal lootjes dat je koopt. Twee keer tachtig is meer dan 92.

   De kans dat je een prijs wint, is een ander verhaal. Als je 92 lootjes koopt, ben je zeker van een prijs. Er zijn namelijk 1 + 3 + 5 = 9 lootjes waarop een prijs valt. Het kan niet zo zijn dat je die alle negen niet hebt, want je hebt er 92 van de 100. Als je extreem veel pech hebt, geef je € 92 uit aan lootjes en heb je één prijs van € 5.

   Bij twee keer tachtig lootjes is het mogelijk dat je geen prijs hebt.

7. Argumenten om wel of niet mee te doen zijn heel subjectief. Wie vindt dat de naam *Nationale Postcode Loterij* fout gespeld is (postcode-loterij is één woord), heeft gelijk. Is dat een argument om niet mee te doen? Voor sommige mensen wel. Wie principieel tegen gokken is, doet sowieso niet mee. Wie zich geen raad zou weten met een grote prijs, ook niet. Wie het 'gewoon leuk' vindt, heeft een heel eenvoudige reden. Het steunen van een goed doel kan ook een reden zijn.

   Rationeel gezien kan een argument zijn: de kosten van een lot zijn niet heel hoog. Dat geld kun je wel missen. Dat je een heel groot geldbedrag kunt winnen, is een aantrekkelijke gedachte. Dat de winstverwachting negatief is, wil op zichzelf niet zeggen dat je het niet moet doen.

Hoofdstuk 41

Schatten? Kun je het niet gewoon uitrekenen dan? Ja. Althans: als je verstand hebt van normale verdelingen en overschrijdingskansen. Dus: als je het vak Statistiek gehaald hebt en/of dit hele boek hebt gelezen. Het gaat namelijk om de *overschrijdingskans* van het *verschil*. Het verschil in lengte tussen mannen en vrouwen is het verschil tussen twee normale verdelingen. Die normale hebben een gemiddelde van 180 cm − 170 cm = 10 cm en een standaardafwijking die gelijk is aan de wortel van de som van de kwadraten van beide standaardafwijking. Komt dat even mooi uit! 6 in het kwadraat is 26 en 8 in het kwadraat is 64. De som daarvan is 100 en de wortel dáárvan is 10. We hebben dus een normale verdeling met een gemiddelde van 10 cm en een standaardafwijking van 10 cm. Dit gaat misschien wel uit ons hoofd lukken…

1. Hoe groot is de kans dat het verschil in lengte tussen een vrouw en een man positief is? Dat is de kans dat een waarde minder dan één keer de standaardafwijking onder het gemiddelde zit. We (nou ja, sommigen van ons) weten uit ons hoofd dat 68% van de waarden 'binnen plus of min sigma zit'. Dat betekent dat 32% erbuiten zit en de helft ervan dus 'aan de ene kant'. Dat is 16%.

   In *Excel*: NORM.VERD(0; 10; 10; WAAR)

   Wat nauwkeuriger: 15,8655253931457 %.

   Schaam je! Schattingen geef je niet op in vijftien significante cijfers. Twee is zat.

2. Zelfde sommetje, maar nu komt de essentie van de oefening.
   - Het gemiddelde van het verschil in lengte is 10 cm.
   - De standaardafwijking van het verschil in lengte is (toevallig) óók 10 cm.
   - De standaardafwijking van het *gemiddelde* verschil in lengte is gelijk aan die standaardafwijking gedeeld door de wortel van het aantal. De wortel van 100 is 10, dus de standaardafwijking van het *gemiddelde* verschil in lengte is 1 cm.

   Nu het sommetje, maar dat kan niet meer uit het hoofd.

   We hadden in Excel: NORM.VERD(0; 10; 10; WAAR) Dat is *de kans van meer dan één sigma boven het gemiddelde*. Nu zoeken we *de kans van meer dan tien sigma boven het gemiddelde*. Die kans is in twee significante cijfers: $7{,}6 \times 10^{-22}$ %.

   Als je niet van exponenten houdt: $7{,}6 \times 10^{-22}$ % kun je schrijven als

   0,000 000 000 000 000 000 000 76 %.

   En als je niet van machten, nullen en percentages houdt, kun je gebruik maken van de SI-voorvoegsels uit de bijlagen. De kans is namelijk 1 op 1 gedeeld door die

waarde die Excel uitrekende: 1 op 1,3 × 10^{-23} dus. Daar zijn geen woorden voor…
Toch wel: De kans is kleiner dan één op honderd triljard.

Hoofdstuk 49

1. De essentie van deze oefening is dat je laat zien dat je (in de context van dit boek) *altijd* een onzekerheid opgeeft. De slingertijd is niet domweg 3,7 seconde, maar correct weergegeven: 3,7 ± 0,2 s. Om de lengte van de slinger te berekenen herschrijf je de formule in de vorm waarbij datgene wat je wilt uitrekenen links staat.

$$T_0 = 2\pi\sqrt{\frac{l}{g}}$$

wordt dus omgeschreven in de vorm

$$l = \frac{{T_0}^2}{4\pi^2}g$$

De relatieve fout in π is verwaarloosbaar. Ervan uitgaande dat de slingerproef in Nederland gedaan is, ligt g tussen 9,8110 en 9,8136 m/s². Daar zit hooguit een promille onzekerheid op. De formule voor de onzekerheid is simpel:

$$\frac{\delta l}{|l|} = 2\frac{\delta T_0}{|T_0|}$$

De slinger is 3,4018 m met een onzekerheid van 0,3676 m.
In het juiste aantal cijfers: 3,4 ± 0,4 m.

In plaats van de onzekerheidsformule, kun je ook de hoogste en laagste waarde invullen die horen bij 3,7 ± 0,2 s, namelijk 3,5 s en 3,9 s. De lengtes die je dan vindt zijn 3,044 m en 3,780 m. Je komt met die aanpak op 3,4118 m met een onzekerheid van 0,3678 m. In het juiste aantal cijfers: 3,4 ± 0,4 m.

Hoofdstuk 52

1. Bereken wanneer je die waarde exact bereikt.

$$\frac{1}{\sqrt{N}} = 0{,}05 \implies \sqrt{N} = 20 \implies N = 400$$

Het juiste antwoord moet nog wel het woord *minimaal* bevatten.
Je moet namelijk *minimaal* 400 metingen doen. Meer is ook goed.

2. Bereken wanneer je die waarde exact bereikt.

$$\frac{1}{\sqrt{2N-2}} = 0{,}1 \implies \sqrt{2N-2} = 10 \implies 2N-2 = 100 \implies N = 51$$

Het juiste antwoord moet nog wel het woord *minimaal* bevatten.
Je moet namelijk *minimaal* 51 metingen doen. Meer is ook goed.

3. Dit zou geen goede tentamenvraag zijn, vanwege de wat misleidende formulering. Die is hier bewust gekozen, omdat het een *oefening* is en geen *toetsing* (wat een tentamen wél is).
"In het algemeen" ligt $\bar{x}$ dichter bij μ, in die zin dat de spreiding van een gemiddele

van tien kleiner is dan de spreiding van één waarde. De standaardafwijking is ruim drie keer zo klein. Maar het is niet *algemeen* waar dat één meting verder van het gemiddelde ligt dan het gemiddelde van een steekproef.

> Neem als voorbeeld onderstaande tien metingen (zonder eenheid). Ze zijn normaal verdeeld met $\mu = 8$ en $\sigma = 1$.
>
> 7,99, 8,50, 5,81, 8,13, 7,75, 8,59, 7,72, 8,21, 9,22, 8,35, 8,52
>
> Het gemiddelde van deze tien metingen is 8,03. (De steekproef-standaardafwijking is 0,894, maar doet niet ter zake.) Het gemiddelde van de metingen wijkt (in dit specifieke geval!) drie keer zoveel af van het gemiddelde van de verdeling als de eerste meting. Ja, dat is toeval. En meestal is het andersom. Je kunt dus niet zeggen dat $\bar{x}$ "in het algemeen" dichter bij μ ligt dan die ene meting.

4. Dit zou óók geen goede tentamenvraag zijn. Niet zozeer vanwege de inhoud, maar vanwege de formulering. Het gemiddelde en de standaardafwijking hebben namelijk geen onzekerheid — althans, niet als je μ en σ bedoelt: de parameters die de normale verdeling definiëren. Dat zijn vaste (maar onbekende) getallen. In de praktijk werk je echter met steekproefwaarden: het gemiddelde $\bar{x}$ en de standaardafwijking s. En die hebben wél onzekerheid. Het steekproefgemiddelde $\bar{x}$ is (bij voldoende metingen) normaal verdeeld, met een verwachtingswaarde μ en een standaardafwijking $\sigma/\sqrt{N}$. De steekproefstandaardafwijking s is niet normaal verdeeld, maar volgt een chi-verdeling. Ook s heeft dus een eigen verwachte waarde en spreiding.

   Los hiervan is het een strikvraag: je hebt niet voldoende gegevens. Als je kijkt naar de relatieve onzekerheid, zie je dat die voor beide grootheden bij N = 2 ongeveer gelijk is: zo'n 71%. Voor grotere steekproeven is de relatieve onzekerheid in s kleiner dan in $\bar{x}$. Dat gaat tegen je gevoel in: het lijkt moeilijker om een spreiding te schatten dan een gemiddelde. En meestal is de standaardafwijking ook kleiner dan het gemiddelde, dus relatief zou je het omgekeerde verwachten.

   > Een voorbeeld: volgens het CBS zijn volwassen vrouwen in Nederland gemiddeld 1,69 meter lang, met een standaardafwijking van 6,5 cm. Het gemiddelde is dus ongeveer 25 keer zo groot als de standaardafwijking. En toch is de relatieve onzekerheid in die standaardafwijking (bij voldoende metingen) kleiner dan die in het gemiddelde.

Hoofdstuk 54

1. De getallen (zonder de eenheden) worden beide door 2,54 gedeeld, dus de relatieve fout verandert niet.

2. Dit is een beetje een strikvraag, maar hij klopt wel.

   Ook de absolute fout verandert niet! De *getallen* veranderen wel (en de absolute fout dus ook), maar de *grootheid zelf* verandert niet.
   Voorbeeld: de tafel is 2,5400 m lang en de onzekerheid is 1,27 cm. (Het aantal significante cijfers in de onzekerheid is in dit voorbeeld voor het gemak even te groot gekozen.) Als je dat omrekent naar inches, kom je op een lengte van 100,0 inch en een onzekerheid van 0,5 inch. Het getal 0,5 is natuurlijk wel kleiner dan het getal 1,27, maar 0,5 inch is niet kleiner dan 1,27 cm. Het is precies hetzelfde.

3. Zowel voor de flitspalen als voor de trajectcontrole geldt dat het maar een zeer beperkte meting is. De flitspaal meet maar op één moment, dus de auto kan *daarvoor en daarna* te hard rijden.
   Voor de trajectcontrole geldt dat de onzekerheid veel kleiner is, maar ook daar weet je niets over ervoor en erna. Ook op het traject waar gemeten is, kan de snelheid op een bepaald moment te hoog geweest zijn. Je weet alleen *de gemiddelde snelheid* op dat traject.

Hoofdstuk 54

1. De relatieve onzekerheid van beide weerstanden samen (in serie) is ook 5%. De absolute onzekerheid is twee keer zo groot.
2. Reken alles om naar ml (cm^3) en absolute maten.
   8% van 2400 ml = 192 ml
   2400 ml + 600 ml = 3000 ml
   Als de relatieve onzekerheid in het totaal maximaal 10% mag zijn, mag de absolute onzekerheid in het totaal maximaal 300 ml zijn.
   De absolute onzekerheden van beide 'delen' worden bij elkaar opgeteld. In het eerste deel zit al een onzekerheid van 192 ml. In het tweede deel mag dus nog een onzekerheid van 108 ml zitten.
   Eigenlijk moet je dat schrijven in twee significante cijfers.
   Dan wordt het 0,11 liter.
3. Er zijn vast meer manieren, maar wat ik zelf zou doen, is: kopieer de tekst naar *Word* en voer de functie *Zoeken en vervangen* uit. Vervang dan de 'm' door een ander teken (of toch een 'm', dat mag ook) en *Word* vertelt je hoeveel keer hij die letter vervangen (en dus gevonden heeft.)
4. Een aantal aspecten speelt een rol.
   a. Het *definitieprobleem*. Wat bedoel je met "hierboven"?
      De alinea? De regel? De tekst van het hoofdstuk (of de pagina of het boek) tot aan dat punt?
      Bedoel je alleen de kleine letter 'm' of ook de hoofdletter 'M'?
   b. Meten & Weten heeft veel te maken met *nadenken* hoe je tot een uitkomst of waarde komt. Ook van zoiets doms als het tellen van de letter 'm'.

   Of het rare vraag is? Tsja…

Hoofdstuk 60

1. De absolute onzekerheid van een *verschil* tussen twee grootheden is gelijk aan *de som* van de absolute onzekerheden. Een relatieve onzekerheid van 5% op een weerstand van 1,0 kΩ = 1000 Ω betekent een absolute onzekerheid van 50 Ω. Twee keer die waarde is: 100 Ω.

2. De vraag is in feite niet te beantwoorden. Het verschil tussen 1,0 kΩ en 1,0 kΩ is 0,0 kΩ = 0 Ω. De absolute onzekerheid = 100 Ω. Het verschil zou je dus moeten noteren als 0 ± 100 Ω.

   Eigenlijk zeggen we dus: de weerstanden zijn aan elkaar gelijk, maar ze kunnen 100 Ω van elkaar verschillen.
   De relatieve onzekerheid is rekenkundig gezien ongedefinieerd. Je zou ook kunnen zeggen: die is oneindig groot.

Hoofdstuk 61

Er is in dit sommetje sprake van alleen maar vermenigvuldigingen. Het volume van het stalen vat bedraagt lengte × breedte × hoogte en het volume van het water erin vind je door het te vermenigvuldigen met een factor.

> Eigenlijk klopt de vraag niet. Wat bedoel je met "de afmetingen" van het vat? De buitenkant of de binnenkant? Als het de buitenkant is, moeten de dikte van de wand(en) nog weten.
>
> De relatieve fout wordt gevonden door de relatieve fout in die *vier* factoren op te tellen. Het zijn er vier omdat de onzekerheid in de hoogte is opgegeven als *fractie* van de hoogte. Was dat een *absolute* onzekerheid geweest, dan zouden er slechts *drie* factoren van belang zijn geweest.

De relatieve onzekerheden zijn 0,06 / 1 = 6% en 0,06 / 4 = 1,5 % en 0,06 / 8 = 0,75 % voor de drie afmetingen en 5 % voor de hoogte van het waterpeil. De grootste bijdrage wordt dus geleverd door de 6 % in de hoogte.

Als voor het hoogtepeil van het water in plaats van de *fractie* van de hoogte van het vat een *absolute onzekerheid* van 5 cm was opgegeven, dan zou de hoogte niet meer ter zake hebben gedaan.

1. De relatieve onzekerheid van 6% in de hoogte van het vat levert de grootste bijdrage. Het is de grootste term van de vier bijdragen.

2. Dit is een gemene vraag. Want wat is de onzekerheid in het hoogtepeil van het water? Is die 5%? Nee. Althans: je zou de vraag kunnen lezen als: "De onzekerheid in de hoogte is 5% van exact 1 m." Maar je kunt met evenveel recht beweren dat die onzekerheid gelijk is aan 5% van (1,00 ± 0,06) m. Dan is de onzekerheid dus een gevolg van die 5% én die 6%... Het gaat dan om het product van twee factoren met een onzekerheid. De totale relatieve onzekerheid is dan 5% + 6% = 11%.

   Laten we de makkelijkste variant nemen. We bedoelen met 5% gewoon 5% van 1 m. Dat is (absoluut gezien) 5 cm.

   > Hoe zit dat nu met de vorige vraag? Hebben we die dan niet goed beantwoord? Dat ligt er ook weer aan hoe je het bekijkt. Die 5% als onzekerheid op "de helft" is niet de grootste bijdrage. Maar als je zegt dat de hoogte van het waterpeil gelijk is aan 0,50 ± 0,06 m, dan hebben we daarin de grootste bijdrage.

3. We beginnen altijd eerst met de onzekerheid. In de meest eenvoudige interpretatie van de vraag is de relatieve onzekerheid gelijk aan 6% + 1,5% + 0,75% + 5% = 13,25%.

   Het volume is:
   - lengte × breedte × hoogte × fractie =
   - 1 m × 4 m × 8 m × 0,5 =
   - 16 m^3.

   Let (zoals altijd) op de eenheden. Lengte, breedte en hoogte zijn in meters. De fractie is dimensieloos.

   De absolute onzekerheid is 13,25% van 16 m^3 = 2,12 m^3. In het juiste aantal significante cijfers is het volume van het water dus 16 ± 2 m^3.

   > Het werd niet expliciet gevraagd, maar dat hoeft ook niet: de onzekerheid moet ALTIJD worden opgegeven als je die weet.

4. Het volume van een bol wordt gegeven door onderstaande formule.

$$V = \frac{4}{3}\pi r^3$$

De relatieve onzekerheid in het volume is drie keer de relatieve onzekerheid in de straal. Die relatieve onzekerheid in de straal mag dus 2% zijn: één derde van 6%.

De straal van de bol is 0,0500 m (de helft van de diameter van 1,00 dm). De absolute onzekerheid mag dus 2% × 0,0500 m = 0,000100 m zijn. In decimeters: 1,0 dm.

Hoofdstuk 62

```
som = 0
aantal = 0
for steen in range(1, 7):
    product = steen * steen
    som = som + product
    aantal = aantal + 1
gemiddelde = som / aantal

print(gemiddelde)

som = 0
aantal = 0
for steen1 in range(1, 7):
    for steen2 in range(1, 7):
        product = steen1 * steen2
        som = som + product
        aantal = aantal + 1
gemiddelde = som / aantal

print(gemiddelde)
```

Hoofdstuk 65

1. De formule voor optellen zegt dat je de absolute fouten bij elkaar optelt. En $\delta B + \delta B = 2\delta B$.

Hoofdstuk 66

1. Voor $n = 0$ heb je een grootheid tot de macht 0. De waarde daarvan is 1. Omdat het exact 1 is, is de absolute onzekerheid gelijk aan 0 en de relatieve onzekerheid dus ook. Die is namelijk 0 gedeeld door 1.
2. Noteer eerst de formules voor de volumes van een bol en een kubus.

$$V_{bol} = \frac{4}{3}\pi x^3$$

$$V_{kubus} = x^3$$

Uit de derde macht in de formules blijkt dat de relatieve onzekerheid van beide volumes hetzelfde is, namelijk drie keer de relatieve onzekerheid van x.

$$\frac{\delta V_{bol}}{|V_{bol}|} = \frac{\delta V_{kubus}}{|V_{kubus}|} = 3\frac{\delta x}{|x|}$$

De absolute onzekerheden zijn gelijk aan de relatieve onzekerheden maal de volumes zelf.

$$\delta V_{bol} = 3\frac{\delta x}{|x|} \times V_{bol} = 3 \times \frac{4}{3}\pi x^3 \times \frac{\delta x}{|x|} = 4\pi x^2 \delta x$$

$$\delta V_{kubus} = 3\frac{\delta x}{|x|} \times V_{kubus} = 3 \times x^3 \times \frac{\delta x}{|x|} = 3x^2 \delta x$$

Om die twee uitkomsten met elkaar te vergelijken, is het handig om ze op elkaar te delen.

$$\frac{\delta V_{bol}}{\delta V_{kubus}} = \frac{4\pi x^2 \delta x}{3x^2 \delta x} = \frac{4}{3}\pi \approx 4{,}19$$

De absolute onzekerheid in het volume de bol is dus ruim vier keer zo groot als de absolute onzekerheid in het volume de kubus.

3. Het maakt niet uit of je naar de straal of de diameter kijkt, omdat in beide formules een kwadraat staat. Beide onzekerheden worden een factor vier groter.

Hoofdstuk 67

1. 2,7 ± 0,3 V
2. 2,7(3) V
3. Voordelen:
    a. De notatie is korter als je veel decimalen hebt. Je noteert alleen een spatie, twee haakjes en één of twee cijfers.
    b. Je kunt geen inwendige tegenstrijdigheid krijgen in het aantal decimalen.

    Nadelen:
    a. De notatie maakt niet expliciet duidelijk dat er sprake is van een spreiding, wat het plusminusteken wel doet. Je zou het kunnen interpreteren als twee extra decimalen.
    b. Het getal dat de onzekerheid aangeeft kun je niet 'los' van de beste schatting lezen, omdat je de grootte niet kent.

Hoofdstuk 68

1. Om dat te kunnen uitrekenen, moet je eerst bedenken hoeveel mogelijke cijfers er zijn. Alle cijfers van 1 t/m 9 kunnen tien verschillende decimalen (0 t/m 9) hebben. Alleen een 10 wordt altijd gevolgd door een 0. Er zijn dus 9 × 10 + 1 = 91 mogelijke cijfers. De verschillen zitten in alle even cijfers die gevolgd worden door een 5. Dat zijn dus de 2,5, de 4,5, de 6,5 en de 8,5. Vier van de 91 cijfers pakken met rekenkundig afronden precies een punt hoger uit dan met statistisch afronden. Het gemiddelde voordeel dat studenten hebben met afronding is dus 4/91 = 0,044 punt. Is dat veel? Het valt wel mee, maar er zijn in Nederland meer dan achthonderdduizend studenten aan hogescholen en universiteiten. Als die allemaal twaalf cijfers per jaar

halen, dan worden er jaarlijks toch bijna een half miljoen punten cadeau gegeven. Daar hoor je nou nooit eens iemand over klagen…
Voor studenten die het nog niet helemaal begrijpen en bang zijn dat de afronding ooit zal veranderen: een 5,5 is sowieso een 6.

2. De enige situatie waar het uitmaakt, is het cijfer 5 dat voorafgegaan wordt door een *even* cijfer. In dat geval gaat statistisch afronden omlaag (naar dat even cijfer toe dus) en rekenkundig afronden omhoog.
In het getal 3,14159265358 staan drie vijven. Alleen de tweede wordt voorafgegaan door een even cijfer. Bij afronden op zeven decimalen krijg je dus een verschil: 3,1415926 voor statistisch afronden en 3,1415927 voor rekenkundig afronden.

3. Voorwaarde om deze opgave te kunnen maken, is dat je snapt hoe kommagetallen werken als je hexadecimaal rekent. Het gaat niet anders dan bij decimaal rekenen. Alleen staat werk je nu niet met 10^{-1} en 10^{-2} voor het eerste en tweede cijfer achter de komma, maar met 16^{-1} en 16^{-2}.
Ook het afronden gaat op dezelfde manier. Kijk naar het cijfer dat moet "verdwijnen". Dat is een 8. Een 8 is te vergelijken met de 5 bij decimaal tellen: het is namelijk de helft van het grondtal. Een 8 wordt naar boven afgerond bij rekenkundig afronden en naar het even cijfer bij statistisch afronden. Is het cijfer ervoor even of oneven? Een hexadecimale F is *oneven.* (Het is namelijk niet deelbaar door 2.) Het afronden van 1A,F84 gaat dus in beide gevallen naar boven.

4. Een binair kommagetal werkt óók gewoon als een decimaal getal. De cijfers achter de komma hebben als waarde niet 10^{-1} en 10^{-2} en 10^{-n}, maar 2^{-1} en 2^{-2} en 2^{-n}. Dat getal 101,01 is dus gewoon gelijk aan 5 (die 101 voor de komma) plus 0,25 (die 01 na de komma).
Nu het afronden. Je hebt maar twee cijfers, namelijk 0 en 1. Die 1 is te vergelijken met de 5 bij een decimaal talstelsel. Het is namelijk de helft van het grondtal.

   Wat misschien een beetje raar "aanvoelt", is dat een binair talstelsel maar twee cijfers heeft. Bij een decimaal talstelsel heb je 0 t/m 4 (die omlaag worden afgerond) en 5 t/m 9 (die omhoog worden afgerond). Bij een binair talstelsel bestaat de eerste helft van de cijfers uit alleen die 0 en de tweede helft uit alleen die 1. Bij rekenkundig afronden wordt elke 0 omlaag afgerond en elke 1 omhoog. Bij statistisch afronden wordt die 1 afgerond naar het even cijfer… naar de 0 dus.

   Het principe is niet anders dan bij decimale getallen. Sterker nog: het is misschien wel duidelijker. Een 0 hoef je niet af te ronden: die laat je gewoon weg. Een 1 moet je wel afronden en dat wil je dus in de helft van de gevallen omlaag doen en in de andere gevallen omhoog.

   Het getal 101,01 wordt in één decimaal dus 101,1 bij rekenkundig afronden en 101,0 bij statistisch afronden wordt het 101,0.

Hoofdstuk 69

1. 1
2. 1
3. 1
4. 2
5. 2

6. 3
7. 1 (en dat 'voelt' een beetje raar, want je geeft het geldbedrag net zo nauwkeurig weer als in de vorige oefening)
8. 4
9. 1, 2, 3, 4, 5, 6 of 7 (en die niet-eenduidigheid zou te voorkomen zijn door wetenschappelijke notatie)
10. 2 (en meteen een voorbeeld van de wetenschappelijke notatie)
11. 9 (of misschien toch minder?)
    De punten in het getal maken niet uit, die zijn alleen voor de leesbaarheid. Maar het feit dat er twee nullen achter de komma staan suggereert dat die andere nullen ook significant waren.

Bij de laatste oefeningen ging het telkens om 'een miljoen'. De wetenschappelijke notatie is eenduidig.

Hoofdstuk 70

1. 460
2. 0
3. 314
4. 4,72
5. 1×10^2
6. 1,5240
7. De lichtsnelheid is exact gedefinieerd als 299 792 458 m/s.
   Een uur is exact 60 × 60 = 3600 seconde.
   Een vermenigvuldiging levert op:
   299 792 458 m/s × 3600 s/u = 1 079 252 848 800 m/u
   Hoeveel significante cijfers zijn er? Negen? Daar lijkt het op, maar zowel voor de getallen 299 792 458 als 3600 geldt dat ze exact zijn.

   De exacte uitkomst is 1 079 252 848,8 km/u.

8. Deel de versnelling door het aantal meters per mijl (1609,344) in mijlen te rekenen. Vermenigvuldig dan met het aantal seconden in een kwartier (15 × 60 = 900) in het kwadraat. Het aantal significante cijfers is waar het in deze vraag om draait: dat moeten er namelijk negen zijn. De beide omrekenfactoren zijn exact.
   $$\frac{9{,}87654321 \times 900^2}{1609{,}344} = 4.970{,}969538\ \textit{miles per square quarter}$$

Hoofdstuk 71

1. 6 s
2. 15 m
3. $1{,}6 \times 10^6$ kg
4. 3 A
5. 1,9 K
6. 0,2 V
   Door de afronding wordt het eerste cijfer een 2 en dus geen 1.

Daarom noteren we maar één significant cijfer.
Het is verleidelijk om er 0,19 V van te maken, maar dan klopt je afronding niet.

Hoofdstuk 72

1. lengte = 1,710 ± 0,015 m
2. breedte = 200 ± 13 mm
3. oppervlakte = 60,55 ± 0,05 m^2
4. volume = 403 ± 10 ml
5. tijd = 59,988 ± 0,005 s
6. massa = 72,6 ± 0,4 kg
7. hoek = 3,14 ± 0,03 rad
8. volume = 98,4 ± 0,2 cm^3
9. frequentie = 50 ± 3 Hz
10. temperatuur = 20 ± 4 °C
11. elektrische stroom = 1101,0 ± 1,5 mA
12. volume = (4400 ± 7) × 10^3 μm^3
13. spanning = 230,00001 ± 0,00005 V
14. elektrische stroomdichtheid = 0,0 ± 0,4 A/m^2
15. warmte = (6,0 kJ ± 0,4) × 10^3 kJ
16. oppervlakte = 7659,0 ± 0,9 nm^2
17. kracht = 100 ± 10 kN
18. vermogen = 0,22 kW ± 0,03 kW
19. druk = 200,00 ± 0,03 N/m^2 (of Pa)
20. weerstand = 20,0 kΩ ± 0,3 kΩ

Hoofdstuk 73

De juiste notatie is: 9,84 ± 0,18 m/s^2.
Er mogen ook haken omheen: (9,84 ± 0,18) m/s^2.

De zeven opgaven samen bevatten *dertien* fouten. Had je ze *alle dertien*?

1. verkeerde significantie: fout moet in twee decimalen, want het eerste significante cijfer is een 1
2. onzekerheid verkeerd afgerond
3. onzekerheid heeft een significant cijfer te veel
4. beste schatting is verkeerd afrond én
   er ontbreekt een spatie voor de eenheid
5. punten gebruikt in plaats van komma's
6. beste schatting heeft een significant cijfer te weinig
7. punten in plaats van komma's
   verkeerd afgerond
   geen onzekerheid opgegeven
   geen eenheid opgegeven
   significant cijfer te veel
8. Schrijf eerst de formule zo dat G links van het isgelijkteken komt te staan.

$$G = F\frac{r^2}{m_1 \times m_2}$$

Daarna kun je de eenheden invullen in deze formule

$$N\frac{m^2}{kg \times kg} = kg \times m \times s^{-2}\frac{m^2}{kg \times kg} = m^3 kg^{-1}s^{-2}$$

Hierin is gebruik gemaakt van het feit dat een newton ter herleiden is tot basiseenheden kg, m en s.

Hoofdstuk 74

1. Schattingen als deze zijn subjectief, maar dat neemt niet weg dat er we wat over te zeggen is. De vraag is eigenlijk: wat is de "vroegste" tijd die je kunt aflezen en wat is de "laatste" tijd. De helft van het verschil kun je als de onzekerheid beschouwen. Anders gezegd: hoe goed kun je schatten tussen twee strepen op de klok.
   Zelf denk ik dat een kwart of een vijfde wel redelijk is. De afstand tussen twee strepen staat (wat de kleine wijzer betreft) voor één zestigste deel van twaalf uur en dus voor vijf minuten. Een onzekerheid van een minuut zou dan redelijk zijn.
   Ondertussen is het nog maar de vraag of die wijzer wel zo exact aangeeft hoe laat het is! Is de wijzerplaat exact rond? (Nee.) Zit het draaipunt van de wijzer exact in het midden? (Nee.) Zit er een beetje speling in het mechanisme dat de wijzer aandrijft? (Ja.)

   De vraag was hoe groot de onzekerheid is *ten gevolge van het aflezen*. Die is ongeveer een minuut.

2. Dat is een uniforme verdeling die loopt tussen twee grenzen: de ondergrens is het aangegeven tijdstip in hele minuten, de bovengrens is één minuut later dan dat.

3. Tsja… Hoeveel is één gedeeld door vijf voor twee? Je zult de tijd moeten uitdrukken in minuten (of seconden) als eenheid. Seconden zijn "netter" omdat de seconde een SI-eenheid is. Minuten liggen meer voor de hand, gezien de aanduiding op de klok.
   Langs de horizontale as staat de eenheid minuut.
   Langs de verticale as staat de eenheid 1/minuut (of minuut^{-1}).

Hoofdstuk 82

1. Dit is een wat vreemde opgave, maar het gaat om de *oefening*.
   De bedoeling is dat je heel goed nadenkt over de *exacte* betekenis van de begrippen.

   a. *measurement* (meting)
      Deze lijkt nogal makkelijk: elke meting heeft een onzekerheid.
      Maar een meting bestaat uit een beste schatting én de onzekerheid. De beste schatting zelf is niet de meting.

      Voorbeeld: we meten een tijd en komen op 2,6 ± 0,3 s.
      Hierin is 2,6 s de *beste schatting* en 0,3 s de *onzekerheid*.
      De meting is de combinatie van beide.

      Wat is nu het antwoord op de vraag? "Eet je appeltaart met slagroom *met slagroom*?" Ja. Tenzij je bedoelt dat je bij appeltaart met slagroom apart

slagroom bestelt. Dan kan wel, maar is niet gebruikelijk. Je hebt dan een dubbele portie.

b. *true value* (werkelijke waarde)

Deze vraag is wat eenduidiger. De *true value* (als die bestaat) heeft geen onzekerheid. Het is een exact getal, dat je meestal niet kent. Dat onbekende getal zelf heeft geen onzekerheid.

c. *best estimate* (beste schatting)
Zie onderdeel a hierboven. Je zou kunnen zeggen dat de beste schatting van 2,6 s een onzekerheid heeft van 0,3 s. Maar je kunt ook zeggen dat die beste schatting gewoonweg de uitkomst van een sommetje of meting is. Wat je afleest van het meetinstrument of de rekenmachine kan in principe volstrekt ondubbelzinnig zijn.

d. *accuracy* (nauwkeurigheid)
precision (precisie)
uncertainty (onzekerheid)
Inhoudelijk zijn deze drie het meest interessant.
Een onzekerheid bestaat uit twee delen: een deel dat een gevolg is van een systematische afwijking (nauwkeurigheid) en een deel dat een gevolg is van een toevallige afwijking (precisie). Voor alle drie geldt dat het berekende en/of geschatte waarden zijn. Die hebben zelf ook een onzekerheid.

e. *accepted value* (geaccepteerde waarde)
Belangrijk is om het verschil in te zien tussen *true value* en *accepted value*. De *accepted value* van bijvoorbeeld een natuurconstante wordt nogal eens gebruikt als de 'echte' waarde, omdat je de *true value* niet kent.

Hoofdstuk 85

1. Je kunt alleen maar vermenigvuldigen en delen als je met een absolute temperatuurschaal werkt. De temperaturen dienen daarom te worden omgerekend van graden Celsius naar kelvin. Daarna kun je ze op elkaar delen.

$$\frac{T_2}{T_1} = \frac{273{,}15 + 20}{273{,}15 + 10} = 1{,}03531697$$

Omdat beide temperaturen zonder onzekerheid zijn opgegeven in twee significante cijfers, kiezen we in het antwoord ook voor twee significante cijfers.

Antwoord: 3,5%

2. Eigenlijk is de opgave niet compleet. Er zou nog bij moeten staan dat Jeroen wil werken met een lineaire schaal. Er zijn namelijk oneindig veel functies die door de twee gegeven punten lopen. Een lineair verband ligt wel het meest voor de hand.

$$T_J = aY_C + b$$

$$a = \frac{T_{J2} - T_{J1}}{T_{C2} - T_{C1}} = \frac{100 - 0}{0{,}01 - (-273{,}15)} = \frac{10000}{27316}$$

$$b = T_{J1} - aT_{C1} = 0 - \frac{100}{273{,}16} \cdot (-273{,}15) = \frac{2731500}{27316}$$

In de berekening van a en b dienen alle onafgeronde getallen te blijven staan. Afronden op (bijvoorbeeld) drie significante cijfers zou maken dat er geen exact antwoord meer gegeven wordt. Afronden doe je pas als je getallen invult met een eindige nauwkeurigheid. Bij het programmeren van de oplossing in *Python* of het invoeren van een formule in *Excel* laten we bij voorkeur de breuken staan, tenzij er redenen zijn om dat niet te doen. In dat geval luidt de formule als volgt.

$$T_J = 0{,}366 \cdot Y_C + 100$$

Het getal 100 is een afronding van 99,996. Omdat we met twee significante cijfers rekenen, zou deze formule bruikbaar zijn. Invullen van 20 °C resulteert in 107,32 °J. Als je strikt bent in het werken met twee significante cijfers, zou je dat moeten noteren als 1,1 × 10^2 °J. Verdedigbaar is om niet naar het aantal significante cijfers te kijken, maar in hele graden te rekenen.

Antwoord: 20 °C komt overeen met 107 °J

Hoofdstuk 87

1. Het verschil is waarschijnlijk niet significant, omdat een tentamen een meting is die niet op één tiende punt nauwkeurig kan vaststellen hoe goed een student de stof beheerst. Wie een 5,4 haalde, had met een net iets andere vraag in de toets misschien wel een 5,7 gehaald. Of een 4,9.
   Je kunt een tentamencijfer beschouwen als een experiment met een normale verdeling. Het gemiddelde van die verdeling is hoger bij iemand die de stof beter beheerst. De standaardafwijking is aanzienlijk groter dan 0,1 punt. Hoe groot die is, hangt af van het vak en de docent en de manier van toetsen.
2. Het verschil is waarschijnlijk wel relevant. Het is namelijk het verschil tussen slagen en zakken.
   Eigenlijk kun je zonder de context de vraag niet beantwoorden. Voor iemand die besluit om te stoppen met de studie of gewoon nooit af te studeren geeft het niet dat het cijfer onvoldoende is.

Hoofdstuk 88

1. Zie ISO 25010 - Wikipedia
2. 31
3. 25010

Hoofdstuk 90

1. Hier is misschien wel sprake van een definitieprobleem. Tellen alle pagina's mee, of alleen de bedrukte? Waar begin je met nummeren en wat doe je met de omslag van het boek?
   Dat aantal van 292 is gebaseerd op het aantal dat *Microsoft Word* zelf aangeeft.

a. Met bovenstaande definitie is de stelling juist. (Tenzij er iets misgegaan is bij het aanmaken van de PDF.)
b. De auteur weet natuurlijk niet wat de lezer denkt, maar wat hij gedaan heeft, is het volgende. In *Word* zit de mogelijkheid om een *veld* op te nemen dat het aantal pagina's toont. Als je daarvan gebruik maakt, hoef je zelf niet te tellen. Belangrijker dan dat: als het aantal pagina's in een volgende uitgave verandert, blijft het gewoon kloppen.
In het menu *Invoegen* zit een knopje *Snelonderdelen* en daaronder *Veld*. In het lijstje waaruit je kunt kiezen, zit een veldnaam *NumPages*.
c. Het aantal even pagina's is het aantal pagina's (292) gedeeld door 2, afgerond naar beneden. Het aantal woorden is (volgens het veld *NumWords*) gelijk aan 81089. (Hoe weten we of dat getal klopt?)

Aantal pagina's	292	*NumPages*
Aantal fouten	97	=ROUND(B1/3;0)
Aantal woorden	81089	*NumWords*
Percentage	0,11962%	=B2/B3*100

De vraag is hoe je met dat aantal fouten moet omgaan. Delen door 3 inderdaad, maar moet het een geheel getal zijn voordat je verder rekent? En rond je dat getal dan "gewoon" af of naar beneden? Het maakt voor het eindantwoord niet veel uit, maar het is wel het verschil tussen exact en (net) niet exact. Het verschil is minder dan één procent van nog geen twee promille.

Waarom heb je deze oefening eigenlijk gedaan? Wat was het nut, het *leerdoel* zoals docenten zeggen?

- Je hebt als het goed is geleerd om in *Excel* met *cellen* te werken in plaats van met harde getallen.
- Je hebt gezien dat je in *Word* ook kunt rekenen met *velden*.
- Je bent je gaan realiseren dat een boek met weinig fouten (eentje per drie pagina's is toch best netjes) er toch al gauw meer dan honderd bevat.
- Je ziet in dat we met zo'n groot aantal nog steeds over *promillen* spreken.

2. Het gaat wat te ver om deze berekening uit te voeren met velden en die kostprijs zal jaarlijks veranderen. Laten we voor het gemak uitgaan van 320 pagina's, 25 euro en 100 grams papier. Het boek is grofweg A5-formaat. A5 is de helft van A4, een kwart van A3, een achtste van A2, een zestiende van A1, een tweeëndertigste van A0. Een A0 is één vierkante meter. Het boek is dus 10 m^2 papier van 100 g/m^2. Dan zou het binnenwerk dus 1 kg wegen. Als papier € 50 per ton kost, dan zit er dus voor 1/1000 van € 50 aan kosten in het papier. Dat is vijf eurocent…

Je kunt ook een ander sommetje maken. Een pak papier van 500 vel kost in de winkel een eurootje of zes. Je kunt een vel papier aan twee kanten bedrukken en A4'tjes doormidden snijden om er A5 van te maken. Met één pak papier druk je dus vier boeken van 500 pagina's. Dan zijn de kosten niet tien eurocent, maar € 1,50.

Bijlagen

A. Alt-codes & Unicodes

Tekens die niet op het toetsenbord zitten, kun je in *Word* invoegen door de Alt-toets in te drukken en tegelijkertijd de ASCII-code van dat teken in te tikken op het numerieke toetsenbord. Als het teken niet in de ASCII-tabel zit, bestaat er vaak een code voor die met een 0 begint. Een alternatief is de Unicode intikken en daarna op Alt+X drukken. Als het teken niet los staat van een ander teken, moet je de code (het getal) eerst apart selecteren en dan op Alt+X drukken. Merk op dat de unicodes hexadecimale getallen zijn.

Het vaste streepje en de vaste spatie (streepjes en spaties die ervoor zorgen dat de delen die er links en rechts van staan op verschillende regels komen) kun je ook maken door Ctrl+Shift in te drukken en de spatie of het streepje in te tikken.

Een *vaste spatie* wordt ook wel een *harde spatie* genoemd.

De 'belangrijkste' twee uit deze tabel zijn het vermenigvuldigingsteken en het minteken. Bij dat minteken helpt *Word* soms een handje. Die maakt (afhankelijk van instellingen) in de som 9 – 3 = 4 × 1½ van de - een — en van 1/2 maakt hij ½. Maar van een * (sterretje) of een x (kleine letter X) maakt hij niet een 'net' vermenigvuldigingsteken.

Maar… van de – maakt hij geen 'net' minteken. *Word* neemt een *half kastlijntje* (Unicode 2213). Je ziet het verschil bijna niet. In het lettertype van Meten & Weten is het verschil 3% en de breedte en 1% in de hoogte.

In de tabel hieronder staan de breedte en hoogte volgens *Inkscape* bij *Linux Libertine 12*.

		Unicode	breedte	hoogte	
9 – 3 = 6	echt minteken	2212	1,821	0,211	
9 – 3 = 6	half kastlijntje	2013	1,881	0,209	Ctrl+'-'
9 — 3 = 6	heel kastlijntje	2014	2,811	0,209	Ctrl+Alt+'-'
9 - 3 = 6	koppelteken	002D	1,091	0,254	'-'

In *Word* heb je ook nog het *zacht afbreekstreepje* (U+00AD, Ctrl + -), waarmee je expliciet aangeeft dat een woord op die plaats mag worden afgebroken en het *vast afbreekstreepje* (U+00AD, Ctrl + Shift + -), waarmee je aangeeft dat daar *niet* mag worden afgebroken. (Je zou in 's-Gravenhage liever een vast afbreekstreepje hebben gehad…)

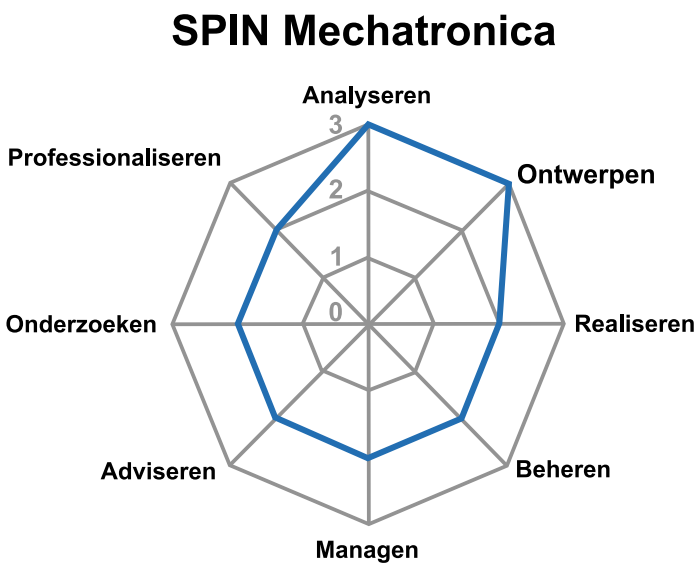
	Alt	Unicode	omschrijving
	8201	2009	smalle spatie (*thin space*)
	0160	00A0	vaste spatie (*non-breaking space*)
-	8209	2011	vast streepje (*non-breaking hyphen*)
–	0150	2013	half kastlijntje *(en dash)*
—	0151	2014	heel kastlijntje *(em dash)*
−	8722	2212	min *(minus)*
×	0215	00D7	vermenigvuldigen
÷	0247	00F7	delen
±	241	00B1	plusminus
‰	0137	2030	promille
·		2219	puntvermenigvuldiging (*bullet operator)*
·	0183	00B7	hoge punt (*interpunct*, voor woordscheidingen)
¼	0188	00BC	kwart
½	0189	00BD	half
¾	0190	00BE	driekwart
≡	240	2261	identiek gelijk aan
≈	247	2248	ongeveer gelijk aan
≤	242	2264	groter dan of gelijk aan
≥	243	2265	kleiner dan of gelijk aan
√	251	221A	wortel
^{n}	252	207F	tot de macht n
^{1}	0185	00B9	tot de macht 1
^{2}	0178	00B2	tot de macht 2
^{3}	0179	00B3	tot de macht 3
π	227	03C0	pi
°	248	00B0	graden
∞	236	221E	oneindig
µ	230	00B5	mu (micro-)
σ		03C3	sigma (kleine letter) / standaardafwijking
Σ	228	2211	sigma (hoofdletter) / som
𝔼		1D53C	verwachtingswaarde *(expected value)*
𝕍		1D54D	variantie *(variance)*

B. Bekende afgeleiden

$f(x)$	$\frac{df(x)}{dx}$
$\frac{1}{x}$	$-\frac{1}{x^2}$
e^x	e^x
$\ln(x)$	$\frac{1}{x}$
$\log(x)$	$\frac{1}{\ln(10)\,x}$
x^n	nx^{n-1}
10^x	$10^x \ln(10)$
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$
$\tan(x)$	$1 + \tan^2(x)$

C. Competenties bij het meten

Bij een hbo-studie Mechatronica horen competenties, die weer te geven zijn in een *spindiagram*. De cijfers daarin geven het niveau aan waarop een afgestudeerd mechatronicus minimaal eindigt.



Op het terrein van Meten & Weten komen de genoemde competenties alle acht voor.

Analyseren Misschien wel de belangrijkste, omdat voor een meting een analyse van de onzekerheid nodig is. Onzekerheden zijn zelden vanzelfsprekend.

Ontwerpen Voordat je een meting uitvoert, moet je nadenken over hoe je dat gaat doen. Het kan de simpele keuze tussen een rolmaat en een duimstok zijn. Maar het bepalen van de valversnelling op een bepaalde plaats brengt heel wat meer keuzes met zich mee. Je ontwerpt een meetopstelling én een meetmethode.

Realiseren Iedereen die iets maakt of doet, realiseert iets. De uitvoering van een meting is een typisch geval van een realisatie. Wie dat deel niet goed voor elkaar heeft, komt tot onjuiste resultaten.

Beheren Een competentie die belangrijker en moeilijker is dan hij lijkt. Natuurlijk leg je de metingen ergens vast. Met pen en papier op een blaadje, of door in te tikken in een spreadsheet of ander programma. Hoe nummer je

de gegevens en hoe voorzie je ze van metadata (datum, tijd, door wie uitgevoerd) zodat je alles weet wat je later weten wilt? Hoe zorg je voor *backups* en voldoende toegankelijkheid? Hoe zorg je voor eventuele geheimhouding en de zekerheid dat data niet per ongeluk of expres veranderd worden?

Managen De uitvoering van een meting vraagt soms om *planning*. Wanneer heb je de data nodig en hoe 'goed' moet de meting zijn? Wanneer zijn de personen en de ruimte(s) beschikbaar waarin de metingen moeten worden uitgevoerd? En hoe zit dat met eventuele meetapparatuur?

Adviseren Misschien wel de meest algemene competentie, want adviseren kun je over bijna alles. Je bent pas competent op dit terrein als je voldoende kennis van zaken hebt. Maar dat alleen is niet voldoende. Het goed overbrengen op de doelgroep is net zo belangrijk.

Onderzoeken Voordat je een meetmethode ontwerpt, is er soms kennis nodig die je nog niet hebt. Bij een nieuwe of meer complexe meting kan dat het geval zijn. Het verwerven van die nieuwe kennis via (literatuur)onderzoek kan dan nodig zijn.

Professionaliseren Wie metingen verricht heeft, zal daar doorgaans over moeten communiceren. Hoe presenteer je de uitkomst zo dat de essentie van wat je te zeggen hebt goed overkomt? Dat een meting al dan niet strijdig is met iets anders, moet je expliciet vermelden. Wat de gevolgen daarvan zijn, moet je ook duidelijk maken. Veel metingen zullen aanleiding zijn tot reflecteren en leren. Volgende keer nóg beter.

D. Doorwerking van fouten

De basis: optellen, aftrekken, vermenigvuldigen, delen

functie	onzekerheid (afhankelijk)	onzekerheid (onafhankelijk)
$Z = A + B$	$\delta Z = \delta A + \delta B$	$\delta Z = \sqrt{(\delta A)^2 + (\delta B)^2}$
$Z = A - B$	$\delta Z = \delta A + \delta B$	$\delta Z = \sqrt{(\delta A)^2 + (\delta B)^2}$
$Z = A \times B$	$\frac{\delta Z}{\lvert Z\rvert} = \frac{\delta A}{\lvert A\rvert} + \frac{\delta B}{\lvert B\rvert}$	$\frac{\delta Z}{Z} = \sqrt{\left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta B}{B}\right)^2}$
$Z = \frac{A}{B}$	$\frac{\delta Z}{\lvert Z\rvert} = \frac{\delta A}{\lvert A\rvert} + \frac{\delta B}{\lvert B\rvert}$	$\frac{\delta Z}{Z} = \sqrt{\left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta B}{B}\right)^2}$

De formules voor aftrekken zijn hetzelfde als die voor optellen.
De formules voor delen zijn hetzelfde als die voor vermenigvuldigen.

Afgeleid: machten en factoren

functie	onzekerheid
$Z = A^n$	$\frac{\delta Z}{\lvert Z\rvert} = \left\lvert n \frac{\delta A}{A}\right\rvert$
$Z = kA$	$\delta Z = \lvert k\rvert \delta A$

Machtsverheffen is vermenigvuldigen met het getal zelf.
Vermenigvuldigen met een exact getal is vermenigvuldigen met iets zonder onzekerheid.
Voor beide formules geldt dat er geen afhankelijke en onafhankelijke variant bestaat. Er is immers geen grootheid om van afhankelijk te zijn.

Combinaties

functie	onzekerheid (afhankelijk)	onzekerheid (onafhankelijk)
$Z = k\frac{A}{B}$	$\frac{\delta Z}{\lvert Z\rvert} = \frac{\delta A}{\lvert A\rvert} + \frac{\delta B}{\lvert B\rvert}$	$\frac{\delta Z}{Z} = \sqrt{\left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta B}{B}\right)^2}$
$Z = k\frac{A^n}{B^m}$	$\frac{\delta Z}{\lvert Z\rvert} = \left\lvert n\frac{\delta A}{A}\right\rvert + \left\lvert m\frac{\delta B}{B}\right\rvert$	$\frac{\delta Z}{Z} = \sqrt{\left(n\frac{\delta A}{A}\right)^2 + \left(m\frac{\delta B}{B}\right)^2}$

Combinaties (vervolg)

functie	onzekerheid (onafhankelijk)
$Z = A + B - (C - D)$	$\delta Z = \sqrt{(\delta A)^2 + (\delta B)^2 + (\delta C)^2 + (\delta D)^2}$

Combinaties (vervolg)

functie	onzekerheid (onafhankelijk)
$Z = \frac{A \times B}{C \times D}$	$\frac{\delta Z}{Z} = \sqrt{\left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta B}{B}\right)^2 + \left(\frac{\delta C}{C}\right)^2 + \left(\frac{\delta D}{D}\right)^2}$
$Z = \frac{A^n \times B^m}{C^p \times D^q}$	$\frac{\delta Z}{Z} = \sqrt{\left(n\frac{\delta A}{A}\right)^2 + \left(m\frac{\delta B}{B}\right)^2 + \left(p\frac{\delta C}{C}\right)^2 + \left(q\frac{\delta D}{D}\right)^2}$

Omdat de formules lang worden, is alleen de kolom met de formules voor onafhankelijkheid weergegeven.

Wiskundige functies

functie f(x)	afgeleide	onzekerheid
$\frac{1}{x}$	$-\frac{1}{x^2}$	$\frac{\delta x}{x^2}$
e^x	e^x	$e^x \delta x$
$\ln x$	$\frac{1}{x}$	$\frac{\delta x}{x}$
$\log x$	$\frac{1}{\ln(10)\, x}$	$\frac{\delta x}{x}$
x^n	nx^{n-1}	$\lvert nx^{n-1}\rvert \delta x$
10^x	$10^x \ln(10)$	$10^x \ln(10)\, \delta x$
$\sin x$	$\cos x$	$\lvert\cos x\rvert \delta x$
$\cos x$	$-\sin x$	$\lvert\sin x\rvert \delta x$
$\tan x$	$1 + \tan^2 x = \frac{1}{\cos^2 x}$	$\frac{\delta x}{\cos^2 x}$

Kettingregel

Als $f(x) = g(h(x))$ $f'(x) = g'(h(x)) \cdot h'(x)$

E. Enige SI-voorvoegsels

macht	voor	letter	naam	decimaal getal
10^{24}	yotta	Y	quadriljoen	1000000000000000000000000
10^{21}	zetta	Z	triljard	1000000000000000000000
10^{18}	exa	E	triljoen	1000000000000000000
10^{15}	peta	P	biljard	1000000000000000
10^{12}	tera	T	biljoen	1000000000000
10^{9}	giga	G	miljard	1000000000
10^{6}	mega	M	miljoen	1000000
10^{3}	kilo	k	duizend	1000
10^{2}	hecto	h	honderd	100
10^{1}	deca	da	tien	10

macht	voor	letter	naam	decimaal getal
10^{-24}	yocto	y	een quadriljoenste	0,000000000000000000000001
10^{-21}	zepto	z	een triljardste	0,000000000000000000001
10^{-18}	atto	a	een triljoenste	0,000000000000000001
10^{-15}	femto	f	een biljardste	0,000000000000001
10^{-12}	pico	p	een biljoenste	0,000000000001
10^{-9}	nano	n	een miljardste	0,000000001
10^{-6}	micro	µ	een miljoenste	0,000001
10^{-3}	milli	m	een duizendste	0,001
10^{-2}	centi	c	een honderdste	0,01
10^{-1}	deci	d	een tiende	0,1

Deze bijlage heeft als titel “Enige SI-voorvoegsels”. Dat suggereert dat er meer zijn, maar dat is niet zo. Dit zijn de enige. Maar ik vond het leuk als elke bijlage een titel had met de beginletter van A, B, C, D, E, F en G.

F. Foutenvinders

De volgende personen hebben me blij kunnen maken door bij het maken van dit boek een of meer fouten te melden.

Tijdens het schrijven van de eerste uitgave in 2023:

Yannick van den Arend, Tommy Baas, Kaj van Beest, Tom Bos, Lars Hartman, Ilse van den Heuvel, Harm Jan van Holst, Guus Jansma, Bram Nijhoff, Daan Sijnja, Jimmy Scheltema, Koos Schoo, Jasper Waaijer, Michael van Weelde.

In de eerste gedrukte uitgave van 2023:

Bas van den Heuvel, Matthias Paul, Mark Schrauwen, Gerard Tuk.

In de tweede gedrukte uitgave van 2024:

Arend Brandt, Pepijn Brozius, Moos Crebolder, Dirk Doedens, Gyan Fransen, Bram Heerschap, Steijn 't Hoen, Elise van der Knaap, Robin Suurs, Job Valkenburg, Matthijs Verburg.

G.Gebruikte boeken

Barford, N. C. (1985). *Experimental measurements: Precision, error, and truth* (2nd ed.). John Wiley & Sons.

Bell, S. (1999). *A beginner's guide to uncertainty of measurement.* National Physical Laboratory.

Berendsen, H. J. C. (1997). *Goed meten met fouten: Een introductie in de verwerking en presentatie van meetresultaten en de discussie van meetfouten.* Bibliotheek der RU Groningen.

Berendsen, H. J. C. (2011). *A student's guide to data and error analysis.* Cambridge University Press.

Bevington, P., & Robinson, K. D. (2002). *Data reduction and error analysis for the physical sciences* (3rd ed.). McGraw-Hill.

Drosg, M. (2007). *Dealing with uncertainties: A guide to error analysis.* Springer.

Fornasini, P. (2008). *The uncertainty in physical measurements: An introduction to data analysis in the physics laboratory.* Springer.

Grabe, M. (2014). *Measurement uncertainties in science and technology* (2nd ed.). Springer.

Hughes, I., & Hase, T. (2010). *Measurements and their uncertainties: A practical guide to modern error analysis.* Oxford University Press.

International Organization for Standardization. (2008). *Guide to the expression of uncertainty in measurement (JCGM 100:2008 GUM 1995 with minor corrections).* ISO.

Kaner, C., Bach, J., & Pettichord, B. (2001). *Lessons learned in software testing: A context-driven approach.* Wiley.

Kenney, J. F., & Keeping, E. S. (1951). *Mathematics of statistics, part two* (2nd ed.). Van Nostrand.

Kirkup, L., & Frenkel, R. B. (2006). *An introduction to uncertainty in measurement: Using the GUM (Guide to the expression of uncertainty in measurement).* Cambridge University Press.

Lyons, L. (1991). *A practical guide to data analysis for physical science students.* Cambridge University Press.

Merrin, J. (2017). *Introduction to error analysis: The science of measurements, uncertainties, and data analysis*. CreateSpace Independent Publishing Platform.

Moore, D. S. (2009). *The basic practice of statistics* (5th ed.). W.H. Freeman and Company.

Morris, A. S., & Langari, R. (2020). *Measurement and instrumentation: Theory and application* (3rd ed.). Academic Press.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression: A second course in statistics*. Addison-Wesley.

Peralta, M. (2012). *Propagation of errors: How to mathematically predict measurement errors*. CreateSpace Independent Publishing Platform.

Rabinovich, S. G. (2005). *Evaluating measurement accuracy: A practical approach*. Springer.

Ratcliffe, C., & Ratcliffe, B. (2016). *Doubt-free uncertainty in measurement: An introduction for engineers and students*. Springer.

Squires, G. L. (2001). *Practical physics* (4th ed.). Cambridge University Press.

Taylor, R. J. (1997). *An introduction to error analysis: The study of uncertainties in physical measurements* (2nd ed.). University Science Books.

H. Hoe bewijs je de correctie van Bessel?

Wiskundeboeken zeggen vaak iets over 'vrijheidsgraden' als ze uitleggen waarom je de formule voor de steekproefvariantie niet deelt door het aantal metingen, maar door eentje minder.

Als je drie keer gooit met die vijfzijdige dobbelsteen, en je krijgt 5, 5 en 2, dan is het gemiddelde 12/3 = 4. Gooide je 5, 3 en 4, dan was dat ook zo, omdat het totaal 12 is en je drie worpen hebt.

Als je het gemiddelde kent en de eerste twee worpen ook, is er dan nog iets 'vrij' aan die laatste? Als je met z'n drieën € 12 moet betalen en de andere twee geven € 3 en € 4 of € 5 en € 2, heb jij dan nog een keuze? Het totaal moet € 12 zijn, en de andere twee gaven samen € 7. Als je het totaal en/of gemiddelde kent van de drie, dan valt er niks meer te raden of kiezen als je de andere twee al weet. Het aantal *vrijheidsgraden* is drie. Als je de drie waarden kent, ligt het gemiddelde vast. Als je het gemiddelde en twee waarden weet, ligt die derde vast. In formulevorm:

$$x_3 = 3\bar{x} - x_2 - x_1$$

Algemener:

$$x_N = N\bar{x} - \sum_{i-1}^{N-1} x_i$$

Je zou dus kunnen zeggen dat die laatste waarde vastligt. Welke gemiddelde je ook kiest: geef mij alle andere waarden en ik vertel je hoe groot de laatste waarde moet zijn. Als de eerste twee 13 en 42 zijn, en het gemiddelde is 333, dan kun je nog steeds een waarde vinden die ervoor zorgt dat dat klopt.

$$x_3 = 3\bar{x} - x_2 - x_1 = 3 \cdot 333 - 42 - 13 = 944$$

Die laatste waarde zorgt ervoor dat de som van alle afwijkingen van het gemiddelde gelijk is aan 0. Dat is per definitie zo.

Die eerste twee waarden kunnen vrij variëren. Die derde niet. Was de eerste waarde 121 geweest en de tweede 400, dan kwamen we met een gemiddelde van 33 op een andere derde waarde uit.

$$x_3 = 3\bar{x} - x_2 - x_1 = 3 \cdot 333 - 121 - 400 = 478$$

Is dit nou een 'bewijs'?

Nee. Dat waren die simulaties ook al niet.

Het probleem van het schatten van een variantie uit een steekproef is dat je niet werkt met het gemiddelde van de verdeling, maar met een schatting daarvan. En die schatting is een beetje fout. Hoe groot is het verschil tussen de echte en de geschatte? Dat kun je in een formule zetten.

$$\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2-\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2$$

Dat kun je uitschrijven als:

$$\frac{1}{N}\sum_{i=1}^{N}({x_i}^2-2x_i\mu+\mu^2)-\frac{1}{N}\sum_{i=1}^{N}({x_i}^2-2x_i\bar{x}+\bar{x}^2)$$

De termen onder beide somtekens worden opgeteld voor precies dezelfde waarden van i, dus mogen we er ook één grote som van maken.

$$\frac{1}{N}\sum_{i=1}^{N}\left(({x_i}^2-2x_i\mu+\mu^2)-({x_i}^2-2x_i\bar{x}+\bar{x}^2)\right)$$

Dat kan simpeler:

$$\frac{1}{N}\sum_{i=1}^{N}(\mu^2-\bar{x}^2+2x_i(\bar{x}-\mu))$$

Van de drie termen zijn er twee waar geen indexjes i in zitten. Die mag je dus ook vóór het somteken zetten. (Waarom zou je immers eerst N keer hetzelfde getal optellen en dan door N delen?)

$$\mu^2-\bar{x}^2+\frac{1}{N}\sum_{i=1}^{N}2x_i(\bar{x}-\mu)$$

We kunnen nóg meer voor het somteken halen: alles waar de waarden $2x_i$ mee worden vermenigvuldigd kun je naar links halen en dan die som zelf daarmee vermenigvuldigen.

$$\mu^2-\bar{x}^2+2(\bar{x}-\mu)\frac{1}{N}\sum_{i=1}^{N}x_i$$

Nu pas wordt duidelijk waarom we al die moeite deden: oplettende lezers zien nu namelijk iets heel bekends staan. Als je de som van alle x_i deelt door hun aantal (N dus) heb je het gewoon over het *steekproefgemiddelde*. Dat is $\bar{x}$.

We kunnen het nu korter schrijven en zijn van die stomme som af.

$$\mu^2 - \bar{x}^2 + 2(\bar{x} - \mu)\bar{x}$$

Haakjes, daar houden we ook al niet van. Iets langer, maar wel haakloos:

$$\mu^2 - \bar{x}^2 + 2\bar{x}^2 - 2\mu\bar{x}$$

Daar zit twee keer iets met $\bar{x}$ in, dus dat kan sowieso korter.

$$\mu^2 - 2\mu\bar{x} + \bar{x}^2$$

Weet je, soms vind ik haakjes opeens toch wel leuk. Ik weet namelijk dat

$$(a - b)^2 = a^2 - 2ab + b^2$$

En dat het stuk links van het isgelijkteken toch wel lekker kort is. Vandaar:

$$(\mu - \bar{x})^2$$

Tsjonge! Het verschil tussen werken met het echte gemiddelde en het steekproefgemiddelde is nu wel heel simpel opgeschreven. Weten we nu hoe groot dat verschil is? Nee. Want de ene steekproef is nu eenmaal de andere niet, ook niet als de verdelingen hetzelfde zijn.

Maar wacht eens even… Kunnen we hier niet de verwachtingswaare van bepalen? Ja. En als je het niet erg vindt, draai ik die twee variabelen dan eerst even om, voordat ik die rare dubbele E ervoor zet.

$$\mathbb{E}((\bar{x} - \mu)^2)$$

De verwachtingswaarde van het verschil tussen het steekproefgemiddelde en het gemiddelde van de verdeling. Het klinkt als te mooi om waar te zijn, maar dat is niet anders dan de variantie van het steekproefgemiddelde. En laten we daarvan nou precies weten hoe groot die is.

$$\mathbb{V}(\bar{x}) = \frac{\sigma^2}{N}$$

We weten nu dus dit: als we rekenen met een formule die deelt door N en uitgaat van het ‘verkeerde’ gemiddelde, namelijk het steekproefgemiddelde in plaats van het ‘echte’ gemiddelde, dan is de verwachtingswaarde die we vinden voor deze schatter ${s_N}^2$

$$\mathbb{E}({s_N}^2) = \sigma^2 - \frac{\sigma^2}{N} = \sigma^2\left(1 - \frac{1}{N}\right) = \sigma^2\left(\frac{N-1}{N}\right)$$

Op die regel met drie istekens staat vier keer hetzelfde. De laatste notatie is het makkelijkste. Je zou namelijk niet met deze ${s_N}^2$willen rekenen, maar iets wat precies een factor groter is om van die fout af te komen. Dan moet je dus delen door die laatste term tussen die ronde haken. Oftewel: vermenigvuldigen met het omgekeerde.

We zouden dus het liefst rekenen met

$$s^2 = \left(\frac{N}{N-1}\right){s_N}^2$$

En die ${s_N}^2$ hadden we zelfd bedacht: dat was namelijk de schatting van de variantie alsof je gewoon door N mocht delen. Aan het einde van dit hoofdstuk krijgen we dus toch weer zo'n stom somteken terug.

$$s^2 = \left(\frac{N}{N-1}\right)\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

Goed nieuws: hier kan nog een N worden weggedeeld.

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

Daar hebben we hem!

Of zoals wiskundigen in een klassieke taal zeggen: *q.e.d.*

> Dat *q.e.d.* is een afkorting van *quod erat demonstrandum*, een Latijnse term die in het Nederlands kan worden vertaald met "hetgeen bewezen moest worden". Engelsen zeggen voor de grap ook wel: *Quiet Easy Done*. Volgens Wikipedia zijn er ook Nederlanders die *wak* in hun bewijs schrijven en dat staat dan voor 'waarheid als een koe'. Je moet wel wiskundige zijn om daar de humor van in te zien. (Die laatste stelling is óók een waarheid als een koe. En bij een slechte grap roep je 'boe'.)

I. In één pagina: het bewijs van “die √N”

Wanneer een variabele normaal verdeeld is met een onbekend gemiddelde μ en standaardafwijking σ, kunnen we bewijzen dat het steekproefgemiddelde $\overline{X}$ ook normaal verdeeld is met verwachting μ en standaardafwijking

$$\frac{\sigma}{\sqrt{N}}$$

Stap 1: Bepaal het gemiddelde van $\overline{X}$

Omdat elke X_i dezelfde verwachting $\mathbb{E}[X_i] = \mu$ heeft, kunnen we de verwachting $\mathbb{E}$ van $\overline{X}$ als volgt berekenen:

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[X_i] = \frac{1}{N}\cdot N\cdot\mu = \mu$$

Stap 2: Bepaal de variantie van $\overline{X}$

Aangezien de X_i ’s onafhankelijk en identiek verdeeld zijn, geldt dat de variantie van hun som de som van hun varianties is:

$$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) = \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{Var}(X_i) = \frac{1}{N^2}\cdot N\cdot\sigma^2 = \frac{\sigma^2}{N}$$

Stap 3: Bepaal de verdeling van $\overline{X}$

Omdat de X_i ’s normaal verdeeld zijn, is elke lineaire combinatie van normaal verdeelde variabelen ook normaal verdeeld. Het steekproefgemiddelde $\overline{X}$ is een lineaire combinatie van de normaal verdeelde X_i ’s.

Conclusie

- Het gemiddelde van $\overline{X}$ is gelijk aan μ.
- De variantie van $\overline{X}$ is gelijk aan $\frac{\sigma^2}{N}$, de standaardafwijking is $\frac{\sigma}{\sqrt{N}}$.
- $\overline{X}$ is zelf normaal verdeeld.

J. Ja, dáááág!

Het is je misschien opgevallen: Meten & Weten staat vol met regels. Over afronden. Over significantie. Over onzekerheden. Over hoe je dingen moet schrijven, noteren, formuleren, presenteren, controleren, dubbelchecken en interpreteren. Regels, die er niet voor niets zijn.

Maar in de echte wereld zeg je tegen vrienden: "Ik zit twee kilo boven m'n streefgewicht", en niet "mijn lichaamsmassa wijkt af met (2,0 ± 0,5) kg van mijn doelwaarde." Dat klinkt alsof je een geavanceerd protocol voor hamstervoeding hebt opgesteld. Je zegt gewoon: "Het is hier ruim dertig graden", ook al zou de drammerige meteoroloog in je hoofd willen toevoegen dat "de plaatselijke temperatuur op dit moment tussen de 303 en 308 kelvin ligt". En je biertje? Dat is gewoon nul-punt-nul, of alcoholvrij. Zelfs als er op het etiket "0,05 ± 0,02 % vol" staat. (Het flesje zit voller trouwens.)

Over procenten gesproken: volgens de internationale norm — NBN EN ISO 80000-1:2013 — hoort er een spatie vóór te staan: 25 %. Want het is een eenheid, net als kg of m. De Stijlgids van de Europese Commissie bevestigt dat, en ook in formele Nederlandse teksten zie je netjes 10 %, 90 %, enzovoort. En wat doe ik in Meten & Weten? Meestal zet ik die spatie wél. Soms met tegenzin, omdat ik als auteur en docent het goede voorbeeld hoor te geven. Maar soms doe ik het zonder. Gewoon 25%, 90%, enzovoort. Fout, dus? Misschien. Maar "een kwart" is informeel en "25%" is dat ook.

> Op de website van Onze Taal keren ze het om. *Andere tekens die bij voorkeur aan het getal vast worden geschreven: het procentteken en het promilleteken: 13%, 10‰ (overigens wordt hierbij in wetenschappelijke teksten vaak wel een spatie gezet).* Daar lijken ze te zeggen: het hoort niet, maar in wetenschappelijke teksten doen ze het tóch.

Ik hoop dat je bij het lezen van dit boek niet struikelt over een spatie die er soms niet staat omdat ik het informeel vond waar het daar niet was.

Soms, heel soms, doe je iets gewoon bewust anders dan het hoort. Ik stop heel vaak voor rood licht, maar als er helemaal niemand aankomt of in de buurt is en ik de laatste trein zou missen door een halve minuut te wachten, rij ik toch maar verder. (Na heel goed kijken!)

Wat ook niet hoort, is een slotwoord maken van een bijlage. Maar ik had er gewoon lol in om de bijlagen A t/m J titels te geven waarvan de eerste letters A t/m J zijn. Veel mensen vinden dat niet grappig. Ik wel, en het is mijn boek.

Dus ja, soms zeg ik tegen een norm of regel: "Ja, dáááág!"

De laatste pagina is leeg.

Bijna dan...