EUCLIDES

vakblad voor de wiskundeleraar

februari

08 nr 4

jaargang 83

Special:

Statistiek en Kansrekening

Orgaan van de Nederlandse Vereniging van Wiskundeleraren

COLOFON

februari
08
nr 4
jaargang 83

Euclides is het orgaan van de Nederlandse Vereniging van Wiskundeleraren.
Het blad verschijnt 8 maal per verenigingsjaar.
ISSN 0165-0394

Inzendingen bijdragen

Artikelen/mededelingen naar de hoofdredacteur: Marja Bos,
Koematen 8, 7754 NV Wachtum
E-mail: redactie-euclides@nvvw.nl

Richtlijnen voor artikelen

Tekst liefst digitaal in Word aanleveren; op papier in drievoud. Illustraties, foto's en formules separaat op papier aanleveren: genummerd, scherp contrast.
Zie voor nadere aanwijzingen:
www.nvvw.nl/euclricht.html

Realisatie

Ontwerp en vormgeving, fotografie, drukwerk en mailingservices
De Kleuver bedrijfscommunicatie b.v.
Veenendaal, www.de-kleuver.nl

Nederlandse Vereniging van Wiskundeleraren

Website: www.nvvw.nl

Voorzitter

Marian Kollenveld,
Leeuwendaallaan 43, 2281 GK Rijswijk
Tel. (070) 390 63 78
E-mail: m.kollenveld@nvvw.nl

Secretaris

Wim Kuipers,
Waalstraat 8, 8052 AE Hattem
Tel. (038) 444 70 17
E-mail: w.kuipers@nvvw.nl

Ledenadministratie

Elly van Bemmel-Hendriks,
De Schalm 19, 8251 LB Dronten
Tel. (0321) 31 25 43
E-mail: ledenadministratie@nvvw.nl

Lidmaatschap

Het lidmaatschap van de NVvW is inclusief Euclides.
De contributie per verenigingsjaar bedraagt voor

- leden: € 52,50
- leden, maar dan zonder Euclides: € 35,00
- studentleden: € 26,50
- gepensioneerden: € 35,00
- leden van de VVWL: € 35,00

Bijdrage WwF (jaarlijks): € 2,50
Betaling per acceptgiro. Nieuwe leden dienen zich op te geven bij de ledenadministratie.
Opzeggingen moeten plaatsvinden vóór 1 juli.

Abonnementen niet-leden

Abonnementen gelden steeds vanaf het eerstvolgende nummer.
Niet-leden: € 55,00
Instituten en scholen: € 140,00
Losse nummers zijn op aanvraag leverbaar: € 17,50
Betaling per acceptgiro.

Advertenties en bijsluiters

De Kleuver bedrijfscommunicatie bv:
t.a.v. Ada Valkenburg
Kerkewijk 63, 3901 EC Veenendaal
Tel. (0318) 555 075
E-mail: a.valkenburg@de-kleuver.nl

Kort vooraf [Marja Bos]

Een leeswijzer bij de special 'Statistiek en Kansrekening'

Special

Meten is weten, gissen is missen. Statistiek en kansrekening, is dat eigenlijk wel wiskunde? Toeval en onzekerheid, afgezet tegen de 'zekerheden' waar de 'wis'-kunde zich normaliter op beroemt? Wie het weet, mag het zeggen - maar het ziet ernaar uit dat in brede kring de opvatting inmiddels is geaccepteerd dat dit vakgebied óók tot de wiskunde behoort. En sinds Gödel, ca. 1930, staan er wel méér zekerheden in de wiskunde op losse schroeven, nietwaar?

Hoe dan ook, dit buitenbeentje van de schoolwiskunde met z'n vele en uiteenlopende toepassingen vormt het thema van dit extra dikke nummer van Euclides. Het is overigens niet de eerste Euclides-special rond dit onderwerp. Ook in de jaargangen 1955/56, 1973/74 en 1981/1982 werden er nummers met dit thema uitgebracht, al stond in de titel toen de K van kansrekening of de W van waarschijnlijkheidsrekening vóór de S van statistiek. Jazeker, de kansrekening vormt wel degelijk het wiskundige fundament van de statistiek, maar er waren en zijn ook andere insteken en accenten mogelijk, zoals duidelijk wordt uit een aantal bijdragen in dit nummer.

Onderwijs

Wim Kleijne bijt in deze special het spits af. Hij laat u zien hoe het vakgebied 'statistiek en kansrekening' in de recente geschiedenis van het Nederlandse wiskundeonderwijs uiteindelijk een plek kreeg. En dat ging allemaal niet zonder slag of stoot.

Bert Nijdam geeft in zijn artikel een beeld van de probleemgeoriënteerde aanpak zoals die ontworpen is voor het domein 'Handelen bij Onzekerheid' van de nieuwe examenprogramma's vwo A/C en havo A vanaf 2011. Het gebruik van grote datasets speelt hierbij een belangrijke rol. Carel van de Giessen illustreert die nieuwe benadering met zijn vertaling van een artikel van Clifford Konold over explorative data-analyse.

Henk Tijms houdt een pleidooi voor een Bayesiaanse aanpak van de kansrekening in het vwo: relevante wiskunde binnen een zinvolle context, aan de hand van een mooie en relatief eenvoudige formule, de regel van Bayes.

Joke Verbeek heeft een praktische tip aan wiskundesecties: stel een kanskoffer samen!

Rob van Oord inventariseert in welke deeltjes van de bekende Zebra-reeks onderdelen uit statistiek en kansrekening aan bod komen. Hij beschrijft bovendien hoe hij zijn leerlingen deze boekjes laat gebruiken om het keuzeonderwerp in het vwo vorm te geven.

Maar vergis u niet, aandacht voor S&K is er al in het basisonderwijs! Maike den Houting laat dat zien in haar bijdrage, voorzien van veel diagrammen die rechtstreeks uit reken/wiskundemethoden voor het basisonderwijs gehaald zijn.

Om werknemers in industriële kwaliteitscontrole te scholen komen we statistiekonderwijs ook in allerlei bedrijfscursussen tegen. Arthur Bakker beschrijft samen met zijn Engelse collega's in het kader van een onderzoek naar technisch-wiskundige geletterdheid de totstandkoming van trainingsmateriaal 'statistische procescontrole' voor werknemers in de auto-industrie.

Bij vakoverstijgende projecten waarin ook het schoolvak wiskunde betrokken wordt, zien we dat het heel vaak om het onderdeel statistiek gaat. Voorbeelden hiervan worden in dit nummer beschreven door Marjanne de Nijs (een project samen met natuur- en scheikunde, biologie, Engels en Nederlands), Klaas Mars en Dédé de Haan (met economie), Cees de Hoog en Bert van der Windt (met maatschappijleer en Nederlands), en Klaske Blom en Peter Tilman (met lichamelijke opvoeding).

Maatschappelijk relevant

Zeer actueel is de oproep van 80 hoogleraren kansrekening/statistiek en medische wetenschappen tot herziening van de strafzaak tegen Lucia de B. Eén van deze statistici, Ronald Meester, praat u bij over de kanstheoretische verwikkelingen in deze zaak.

Ook andere auteurs laten in deze special zien waar statistiek en kansrekening een rol spelen in het maatschappelijk debat - al is die term misschien wel wat zwaarbeladen voor de ophef rond de loting van het Europees Kampioenschap Voetbal, waaraan Rob Bosch u in één van zijn bijdragen laat rekenen. Datzelfde voetbal is onderwerp van een bijdrage van sporteconoom Ruud Koning, die het zogeheten 'thuisvoordeel' aan een nader onderzoek onderwerpt.

Ernstiger lijkt het te worden als volgens een serieuze medische publicatie bijziendheid verband houdt met de geboortemaand… Wetenschapsjournalist Hans van Maanen stelt ondeugdelijke statistische argumenten van dit onderzoek aan de kaak, en illustreert op die manier weer eens hoe wij ons gemakkelijk door dubieuze statistiek kunnen laten misleiden, als we niet opletten. U weet het, Disraeli schijnt ooit eens gezegd te hebben: 'There are three kinds of lies: lies, damned lies and statistics.' Niet zo gek, dus, om ook daarom in het onderwijs veel aandacht te besteden aan dit algemeen vormende onderwerp met z'n grote maatschappelijke relevantie…

'Het weer' is altijd goed voor aangename gesprekstof. Kees Kok en Daan Vogelezang geven zicht op de wijze waarop de weersverwachting op basis van kansrekening tot stand komt.

Marcel Croon licht een op het eerste oog verwarrend fenomeen toe aan de hand van een voorbeeld uit de criminologie. Kun je met statistiek alles bewijzen?!

Een tweede bijdrage van Ronald Meester handelt over het modelleren van epidemieën zoals de varkenspest. Kan de kansrekening een rol spelen in het beteugelen van zo'n epidemie?

Emiel van Berkum besteedt in zijn stuk over proefopzetten aandacht aan het verzamelen van waarnemingen en de organisatie van experimenten om op 'nette' wijze conclusies te kunnen trekken. Hij doet dat aan de hand van een industriële toepassing.

En uiteraard ontkomen we in een special over kansrekening niet aan het gokken: Ben van der Genugten, hoogleraar kansrekening annex kansspeldeskundige, verklapt hoe u kunt winnen bij een spelletje poker.

Geschiedenis

De kansrekening heeft zijn wortels in het dobbel- en het kaartspel. Daarover heeft u ongetwijfeld wel eens wat gelezen, en daarom hebben we voor wat betreft de historische insteek van dit nummer de aandacht verlegd naar de iets minder bekende geschiedenis van de statistiek. Vaak wordt statistiek als een bijproduct gezien van de kansrekening, en ten dele klopt dat - denk dan vooral aan de mathematische statistiek. Voor de beschrijvende statistiek ligt het toch wat anders. Ida Stamhuis laat zien hoe de wiskundige benadering nauwelijks een poot aan de grond kreeg in de statistiek (staat-kunde!) zoals die in de 19e eeuw ten behoeve van de overheid door juristen werd opgezet.

Niet alleen lezen; ook *doen*

Natuurlijk staan er in dit speciale nummer niet alleen artikelen *over* statistiek en kansrekening, maar kunt u ook zelf aan de slag *met* aardige wiskundige probleempjes op dit terrein. Zie bijvoorbeeld de uitnodigende bijdragen van Rob Bosch, Jan van de Craats, Frits Göbel en Dick Klingens.

Cadeautje

Bijgesloten bij dit nummer vindt u een cadeautje: 'Het Ei van Columbus: rekenpuzzels & breinkrakers'. Dit aardige boekje is uitgegeven door onze zusterclub, de Nederlandse Vereniging tot Ontwikkeling van het Reken/Wiskunde Onderwijs (NVORWO) ter gelegenheid van haar 25-jarig jubileum. Leden van beide verenigingen krijgen dit boekje cadeau, om de plannen tot intensievere samenwerking te onderstrepen.

Tot slot

De redactie dankt alle medewerkers aan dit nummer voor hun bijdragen, en wenst u als lezer veel plezier en inspiratie!

Inhoud

Aan dit nummer werkten verder mee: Peter Boelens, Marjan Doijer, Jan Meerhof en Jan Smit.

Rectificatie Euclides 83-3

Bij het artikel 'Steunpunt TU/e: Wiskunde in Wetenschap' van Hans Sterk, verschenen in Euclides 83-3 (december 2007, pag. 105 e.v.), werd per abuis Ernst Lambeck niet als mede-auteur vermeld.
De redactie biedt hem hiervoor haar excuses aan.

De introductie van Statistiek & Kans-rekening als schoolvak

[Wim Kleijne]

Inleiding

In 'Honderd jaar wiskunde-onderwijs' [1] heeft Bert Zwaneveld een globaal overzicht en een globale karakteristiek gegeven van de ontwikkelingsgeschiedenis van het vak Statistiek & Kansrekening (S&K) als schoolvak in het voortgezet onderwijs. Graag sluit ik mij in dit artikel aan bij zijn overzicht, waarin hij een drietal perioden onderscheidt:

- De periode vóór 1968: de tijd van gymnasium (α en β), hbs (A en B), (m)ulo, ambachtsschool, huishoudschool enz. De wiskundevakken algebra, meetkunde, stereometrie, beschrijvende meetkunde, goniometrie werden onderwezen en geëxamineerd.
- Van 1968 tot 1988: de periode die begon met de mammoetwet waarin onder andere de schooltypen gymnasium, havo en mavo werden onderscheiden. Het was de periode van wiskunde I en wiskunde II, van de verzamelingen, van vectorrekening, van analyse en ook van het begin van S&K in de scholen.
- Vanaf 1988: de periode van wiskunde A en wiskunde B.

In dit stuk zal ik de feitengeschiedenis niet herhalen, maar zal ik proberen te achterhalen hoe onze beroepsgroep, dus die van de wiskundeleraren, indertijd reageerde op de introductie van S&K in het voortgezet onderwijs. Dat betekent dat ik me in hoofdzaak zal beperken tot de eerste van de genoemde perioden. We zullen hierbij getuige zijn van een heuse stammenstrijd tussen verschillende groepen, van oplopende emoties en van felle discussies over scherpe controverses. En dat alles binnen ons mooie vak, de wiskunde!

De jaren vijftig

We moeten zeker een halve eeuw teruggaan om een glimp te kunnen opvangen van het nieuwe schoolvak S&K. Ons lijfblad *Euclides* vormt voor deze speurtocht een mooie bron.

Het was in 1954 (uitgerekend het jaar waarin ik zelf als leerling tot de eerste klas van de toenmalige hbs werd toegelaten) dat een rapport verscheen van een leerplancommissie van WIMECOS ('Vereniging van leraren in de WIskunde, de MEchanica en de COSmografie aan Hogere Burgerscholen en Lycea', een van de voorgangers van de huidige Nederlandse Vereniging van Wiskundeleraren). In dit rapport werden voorstellen gedaan voor nieuwe wiskundeprogramma's op hbs en gymnasium. Voor het eerst in de onderwijshistorie werd nu ook 'statistiek' genoemd.

Het is instructief te lezen welke criteria de commissie aanlegde voor de keuze van de onderwerpen/wiskundevakken [2]:

a. de betekenis van de toepassingen die van de leerstof gemaakt worden hetzij in leervakken van de middelbare school, hetzij op gebieden waarmee men de leerlingen in verband met het algemene onderwijsdoel in contact wenst te brengen;

b. de mate van onmisbaarheid van de leerstof bij voortzetting van de studie;

c. de mate van geschiktheid van de leerstof om bij een juiste didactische behandeling bij te dragen tot de beoogde wiskundige vorming en tot denktraining op het gebied van de wiskunde en de hiermee structureel verwante vakken;

d. de betekenis van de leerstof uit cultureel-historisch oogpunt.

Op de keper beschouwd is het heel modern wat hier staat. Het zijn in wezen dezelfde eisen die ook nu aan de keuze van leerstof ten grondslag liggen. Tot en met de 'samenhang-gedachte', waarvan wij vaak denken daarmee iets nieuws te pakken te hebben. Achterom kijken en staan op de schouders van onze voorgangers is dus helemaal zo gek nog niet.

Welnu, de 'statistiek' is één van de vakken die volgens de commissie aan de criteria voldoet. In de erop volgende toelichting spreekt de commissie van

"het ingrijpende besluit (…) de statistiek als nieuw leervak op het programma te plaatsen."

Opmerkelijk is hierin het bijvoeglijke 'ingrijpende'. Dit woord werd namelijk niet gebruikt bij andere voor de school nieuwe vakken, zoals de differentiaal- en integraalrekening. Was er dan met het vak 'statistiek' iets bijzonders aan de hand? Natuurlijk wijst de commissie op het maatschappelijk belang van statistiek en geeft ze vele voorbeelden: van actuariaat tot kwantitatieve biologie, psychologie en sociologie. Ze spreekt tevens over de moeite die aankomende studenten in deze vakken hebben met statistiek. Maar dat alles rechtvaardigt nog niet het gebruik van de term 'ingrijpend'. En al helemaal niet als we het voorgestelde programma lezen, dat in onze ogen eigenlijk niets bijzonders is (op pag. 71 van genoemde *Euclides*):

STATISTIEK
Frequentieverdeling; histogram; gemiddelde; mediaan; standaarddeviatie; somfrequentie.
Permutaties en combinaties; het binomium van Newton.
Elementaire kansrekening; aanschouwelijke behandeling van de normale kromme; steekproeven.

N.b. We zien hier ook het woord 'kansrekening' opduiken. Het gaat dus inderdaad om S&K, statistiek én kansrekening.

Op zoek naar het 'ingrijpende'

Wat was er nu zo ingrijpend aan het besluit het vak S&K voor te stellen? Om dit op het spoor te komen zullen we eens 'luisteren' naar enkele reacties op het voorstel. Want van een behoorlijk aantal reacties is een nauwkeurige documentatie bewaard gebleven, zelfs met naam en toenaam, namelijk in het verslag van de Wimecos-vergadering


EUCLIDES

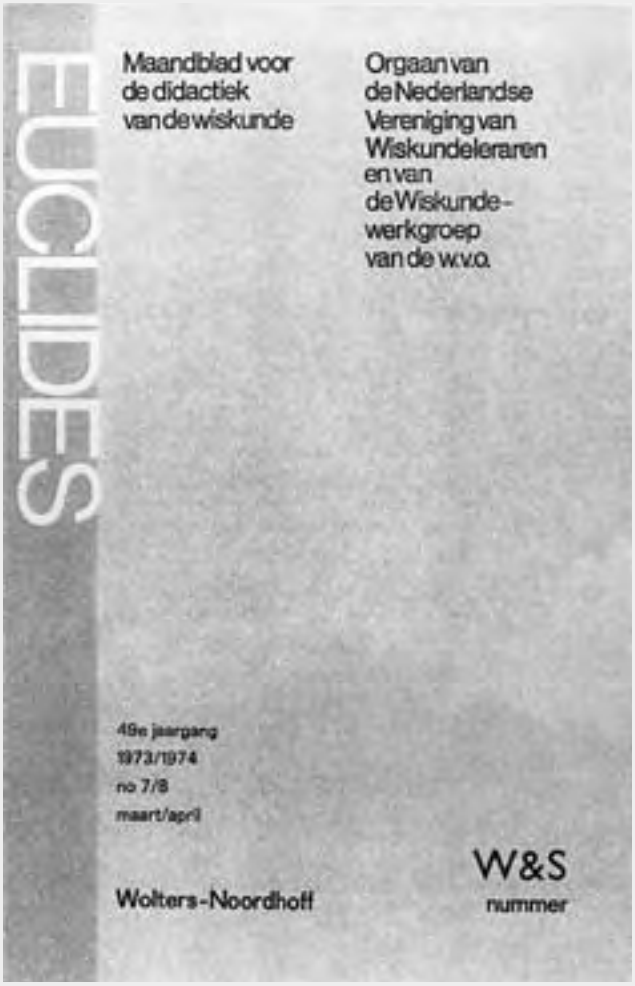
TIJDSCHRIFT VOOR DE DIDACTIEK DER EXACTE VAKKEN ONDER LEIDING VAN Dr H. MOOY EN Dr H. STREEFKERK, Dr JOH. H. WANSINK VOOR WIMECOS EN J. WILLEMSE VOOR LIWENAGEL

MET MEDEWERKING VAN

41e JAARGANG 1965/66

V-VI

P. NOORDHOFF N.V. GRONINGEN



over de voorstellen (in dezelfde *Euclides*; pag. 177). Uit de genotuleerde reacties noteer ik de volgende (de namen laat ik weg):

- Het belang van de statistiek ontgaat hem geheel.
- (…) de statistiek zal eerst wel moeilijkheden bij het onderwijs geven. (…) Overigens is hij voor de invoering ervan (…)
- Zijn hoofdbezwaar gaat tegen de statistiek. Hoewel hij van deze 'gymnastiek' niet veel weet, kan hij wel zeggen: het is geen wiskunde. Voor het doctoraal examen noch voor K5 wordt het gevraagd.
- (…) de meeste collega's weten niets van statistiek. (…) Hij is sterk tegen de statistiek.
- (…) wijst er op, dat men (…) , toch de wenselijkheid van statistiek-onderwijs kan inzien.
- Het programma statistiek is z.i. niet eenvoudig.
- Ook hem ligt de statistiek zwaar op de maag.
- Van de statistiek weet spr. niets af. Is al statistisch vastgelegd welk percentage van de leerlingen dit vak later aan de universiteit nodig heeft?
- (…) deelt mede hoe ook hij tot voor kort niets van statistiek afwist. Uit de proef die er onder leiding van dr Bunt genomen is, is gebleken, dat dit vak door leerlingen best begrepen kan worden. De leerlingen waren door de stof geboeid en de resultaten waren zeer bevredigend.

Hoewel het gaat om een beperkt aantal reacties, geven ze toch enigszins weer hoe men tegen de statistiek aankeek. (Te) kort kunnen deze samengevat worden door:
- We weten er zelf weinig of niets van en bovendien is het geen wiskunde.
En om hier een positieve ontwikkelingsdraai aan te kunnen geven, waren visionaire en doortastende collega's nodig. Met ere noem ik hier Wansink, Vredenduin en Bunt, maar er waren er natuurlijk meer. Aan het doortastende optreden van Wansink, aan het op deskundige wijze pareren van negatieve reacties door hem en anderen is het te danken dat de vergadering zich uiteindelijk achter het voorstel schaarde. Maar toen begon het pas.

Een weg met vallen en opstaan

Bunt schreef zijn baanbrekende werk voor de statistiek op scholen [3] en vergaarde door

Vraag	*Mogelijke antwoorden*
Wat is statistiek	1. *Een bepaald soort straatverlichting*
	2. *Een ander woord voor dictatuur*
	3. *Een tabel met cijfers*
	4. *Een bepaald soort journalistiek*
	5. *Een onderdeel der toegepaste wiskunde*
	6. *Een goocheltruc*
	7. *Een bureau in Den Haag*
	8. *Een onderdeel van het leger*
	9. *Een administratieve methode*
	10. ....
Waarvoor dient statistiek?	1. *Voor het vermaak van het publiek*
	2. *Voor het verkrijgen van overzicht over massale gegevens*
	3. *Om na te gaan of een geneesmiddel werkt*
	4. *Als hulpwetenschap op vrijwel ieder gebied*
	5. *Nergens voor*
	6. *Voor snelle verplaatsing van zwaar materiaal*
	7. *Voor bevolkingsadministratie*
	8. *Voor de verkeersveiligheid*
	9. *Om reclame te maken*
	10. .....

het (facultatieve) gebruik ervan een massa waardevolle evaluatiegegevens in de vorm van gebruik(er)servaringen. Toch was de ingeslagen weg een moeizame, want velen bleven de S&K een vreemde eend in de wiskundebijt vinden. Vanaf de gesprekken waarin het leerplanvoorstel van 1954 werd voorbereid (dus nog vóórdat het boek van Bunt verscheen), kwamen er telkens van verschillende kanten tegenbewegingen op, zoals in 1955, toen gesteld werd: [4]

'aan de statistiek zal in het Nederlandse onderwijs een grotere plaats ingeruimd moeten worden. De H.B.S. lijkt me daarvoor niet de aangewezen plaats om vaktechnische redenen, maar ook omdat het programma al rijkelijk is overbelast. Verder … maakt het vele taaie rekenwerk dat ik in de practijk van de statistiek daaraan steeds verbonden ontmoette, dit vak al haast niet onbruikbaar als schoolvak?'

'Vaktechnische redenen': zou hiermee dan weer bedoeld worden dat S&K niet tot de wiskunde gerekend kan worden?

Ook (zelfs) Wijdenes getuigt, zij het subtiel, van grote scepsis [5]:

Statistiek
Dat moeten we nog afwachten, wat daarvan terecht komt; ik vrees van nog geen kwart van wat het leerplan noemt en ophemelt. Men kan cijfers groeperen, diagrammen maken over volksgezondheid, huisvesting (…) *en nog veel meer. We lezen: "Het is ongetwijfeld een ongewenste situatie, wanneer degene, die dergelijke beschouwingen leest* (…) *zich geen oordeel kan vormen* (…)*". Hoe schoon gezegd, maar* (…) *daarvan komt niets terecht.*

Er werd behoorlijk zwaar geschut in stelling gebracht om de wiskundewereld te overtuigen van het belang van S&K. Zelfs werd de vakantiecursus in 1955 geheel gewijd aan het leerplanvoorstel, waarbij grote nadruk op de S&K werd gelegd. Zozeer dat *Euclides* er een dubbelnummer aan wijdde. [6] Kennisvergroting en het wegnemen van angst voor het onbekende speelden hierbij een grote rol. Bijvoorbeeld op pagina 233 van dit dubbelnummer, waar Hemelrijk inging op 'Wat is en waarvoor dient de statistiek'. Hij opent ludiek, maar met een serieuze ondertoon:

Een fundamentele kwestie

Bij dit alles speelde nog iets anders mee, dat mijns inziens aan de basis lag van de vloed van negatieve reacties. In de verdediging van het vak S&K werd telkens gewezen op de samenhang met andere vakken en op het 'nut' van het vak als 'toepassingsgebied'. In de lijst van Hemelrijk kwam dit, tussen de regels door, eigenlijk ook naar voren. Het betreft hier het destijds zeer gevoelige onderscheid tussen 'zuivere' en 'toegepaste' wiskunde. Vrijwel iedere wiskundeleraar van toen voelde zich een 'zuiver wiskundige'. Vanuit deze positie werd heel vaak misprijzend neergekeken op toegepast wiskundigen. Naar mijn mening ligt hier de werkelijke grond van de weerzin die velen hadden tegen S&K. Zij wensten zich eigenlijk niet 'in te laten' met zo'n toepassingsgebied. Seidel verwoordde dit mooi tijdens bovengenoemde vakantiecursus. Zie het dubbelnummer [6] op pagina 253/4:

Een erkenning van de toegepaste wiskunde bij middelbaar en hoger onderwijs kan helpen om de onlustgevoelens, waarmee sommige zuiveren de toepassingen beschouwen, de ondergrond, die toch vaak uit onkunde bestaat, te ontnemen. Maar niet alleen de verhoudingen binnen de wiskunde kunnen op deze wijze gezuiverd worden, ook wordt hiermee materiaal aangevoerd om het gat tussen de wiskunde en haar buurwetenschappen te dichten.

Alle pogingen en goede bedoelingen ten spijt was de tegendruk te groot en heeft het leerplanvoorstel het niet gehaald. Het werd in 1957 ingetrokken.

Vanaf de jaren '60

Ondanks dit negatieve resultaat is de bovengenoemde korte periode in de jaren vijftig van de vorige eeuw van grote betekenis geweest. In feite werd in deze jaren een gevecht geleverd met als inzet het slechten van de kunstmatige scheiding tussen de zuivere en de toegepaste wiskunde. Dit onderscheid was in de eerste helft van de 20e eeuw ontstaan. In de vele eeuwen die er aan vooraf gingen was de beoefening van de wiskunde altijd gericht geweest op zowel de praktijk als op de theorievorming. Beide gingen hand in hand. In het begin van de vorige eeuw ging men echter een scherp onderscheid maken tussen beide facetten, waarbij de beoefenaren van de zuivere wiskunde zich gingen beschouwen als de 'echte' wiskundigen, zich ver verheven voelend boven de gewone, platvloerse, toegepast wiskundigen. De problematische situatie van de introductie van de S&K was naar mijn mening dan ook eigenlijk een strijd om erkenning van gelijkwaardigheid, een strijd voor terugkeer naar de eenheid van weleer. We zijn hierin getuige van een soort van sociale strijd die halverwege de 20e eeuw binnen ons vak heeft gewoed. Ondanks de 'nederlaag' zette de strijd zich kennelijk als het ware ondergronds voort. Want toen in oktober 1968 de commissie Hemelrijk zich uitsprak over de wenselijkheid van het invoeren van S&K in het voortgezet onderwijs, was er opvallend weinig tegenstand. [7] De commissie kwam in wezen met dezelfde argumenten als in de jaren daarvoor naar voren waren gebracht. Maar nu verwoord met behulp van 'model' en 'modelleren'. Voorgesteld werd om de

S&K aangepast aan een drietal niveaus in te voeren:
- beschrijvende statistiek;
- elementaire statistiek, zonder waarschijnlijkheidsrekening, maar verdergaand dan beschrijvende statistiek alleen;
- statistiek voortbouwende op waarschijnlijkheidsrekening.

De commissie sloot de ogen niet voor de problemen die op de loer lagen:
- de gecompliceerde statistische gedachtegangen die niet volledig begrepen worden, kunnen tot gevolg hebben dat 'de reeds omvangrijke groep statistische goochelaars, die geen flauw benul hebben wat zij uitvoeren, alleen maar wordt uitgebreid';
- docenten zullen een instelling tegenover de statistische problematiek moeten ontwikkelen die duidelijk verschilt van die bij de andere onderdelen van de wiskunde.

Met dit alles in het achterhoofd werd de invoering van S&K voorgesteld.
Na een brede instemming met dit voorstel werd tot invoering besloten en daarmee gingen we de tweede periode in (1968-1988), waarbij de statistiek als officieel onderdeel van de wiskunde in het leerplan en in de examenprogramma's was opgenomen.

Een later dubbelnummer van Euclides [8] was opnieuw gewijd aan S&K. Het aardige was dat in dit nummer ook Wansink door Freudenthal in het zonnetje gezet werd ter gelegenheid van zijn tachtigste verjaardag – Wansink, die midden in de strijd van de 50er jaren stond en nu nog mocht meemaken dat zijn 'nederlaag' van toen omgezet was in een succes.

Een nieuwe periode

De introductie van S&K als onderdeel van de wiskunde in het voortgezet onderwijs markeert in feite het begin van een nieuwe periode. Een periode waarin de 'blik naar buiten' niet meer als een Fremdkörper werd ervaren, maar als een integraal deel van de wiskundebeoefening. Het onderscheid tussen zuivere en toegepaste wiskunde verdween naar de achtergrond. Daarmee waren we dan weer op weg naar de situatie die, om een bekende uitspraak te parafraseren, gekarakteriseerd kan worden als:
'er bestaat geen toegepaste wiskunde, wél bestaan er toepassingen van de wiskunde'. [9]

Sindsdien hebben we gewerkt aan de ontwikkeling van de opbouw, de vormgeving en de didactiek van dit onderdeel van de wiskunde. Opvallend is dat deze facetten momenteel niet alleen in ons land, maar wereldwijd in onderzoek zijn, getuige bijvoorbeeld de voorgenomen brede studie op dit terrein van de International Commission on Mathematical Instruction. [10]
Het zou wel eens kunnen zijn, dat we daarmee opnieuw in een overgangsfase zijn aangekomen. Een overgang naar een situatie waarin het onderwijs in de S&K wellicht op een andere leest geschoeid zal worden. Misschien wel op grond van ontwikkelingen naar aanleiding van de volgende vragen en ideeën:
- zouden we moeten uitgaan van een axiomatische opbouw?
- zouden we moeten uitgaan van dataverzamelingen?
- moeten we alles baseren op een intuïtief kansbegrip?
- zouden we kunnen uitgaan van verdelingsproblemen à la Huygens?
- …

Centraal echter in dit alles staat de vraag waar het in ons onderwijs altijd om draait: wat is het beste voor de jongere van vandaag met het oog op zijn/haar toekomst?
De tijd zal het leren.

Noten

[1] Fred Goffree (red.) (2000): *Honderd jaar wiskunde-onderwijs*. NVvW. Hoofdstuk 19, Bert Zwaneveld: *Kansrekening en statistiek*
[2] *Euclides*, 30e jaargang, 1954/55, nr. IV; pag. 152.
[3] L.N.H. Bunt (1956): *Statistiek voor het vwo*. Groningen: J.B. Wolters N.V.
[4] *Euclides*, 31e jaargang, 1955/56, nr. I; pag. 26.
[5] *Euclides*, 31e jaargang, 1955/56, nr. III; pag. 138.
[6] *Euclides*, 31e jaargang, 1955/56, nr. V en VI.
[7] Commissie Modernisering Leerplan Wiskunde (CMLW): *Rapport over de wenselijkheid en mogelijkheid van het invoeren van statistiek in het onderwijs voor M.A.V.O., H.A.V.O. en V.W.O.*; oktober 1968.
[8] *Euclides*, 49e jaargang, 1973/4, nr. 7/8.
[9] Aan Louis Pasteur wordt de uitspraak toegeschreven: '*er bestaat geen toegepaste natuurwetenschap, wél bestaan er toepassingen van de natuurwetenschap*'.
[10] *Euclides*, 82e jaargang, 2006/7, nr. 4; pag. 139: *ICMI-onderzoek*.

Over de auteur

Drs. Wim Kleijne (1942) was wiskundeleraar, rector van een lyceum en (coördinerend) inspecteur van het voortgezet onderwijs. Erelid van de NVvW. Nu met pensioen, maar nog werkzaam onder andere als algemeen voorzitter van de staatsexamencommissie vmbo-havo-vwo.
E-mailadres: *wim.kleijne@xs4all.nl*

Sabotage

HOE VAKOVERSTIJGEND LEREN KAN LEIDEN TOT HET VINDEN VAN DE DADER

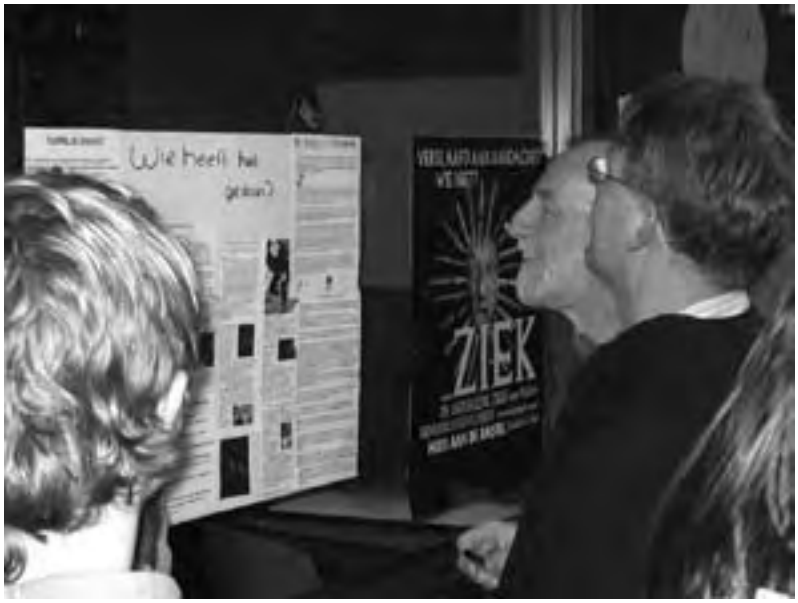
[Marjanne de Nijs]

Inspiratie

Het begon allemaal met een collega die aan de koffietafel enthousiast vertelde over 'verhalend ontwerpen'[1]. Het is een leermethode die in eerste instantie geschreven is voor het basisonderwijs. Leerlingen worden betrokken in een verhaal of activiteit waarbinnen ze worden aangespoord om kennis en vaardigheden te ontwikkelen.
Het gaf ons de inspiratie om te starten met een vakoverstijgend project in klas 2 havo/vwo. We hebben het *de leeronderneming* genoemd[2].
De basis van de leeronderneming is een 'crime scene', namelijk het practicumlokaal. Er is hier door een eindexamenleerling sabotage gepleegd om zijn eigen experimenten met alcohol en kleurstoffen te verdoezelen. Hij/zij heeft dat gedaan door alle proefopstellingen door elkaar te gooien. Voor de sabotage waren vijf leerlingen bezig met hun profielwerkstuk en allemaal hadden ze hun eigen proefopstelling. Eén van hen moet de dader zijn. Tijdens het project gaan leerlingen op zoek naar de dader en zijn/haar motief.

Onderwijs

Aan het begin van de leeronderneming zijn een aantal doelstellingen bepaald. Het belangrijkste is dat de leerlingen kennis en vaardigheden ontwikkelen in een vakoverstijgend project. Vakoverstijgend betekent hier dat de vakken wis-, natuur- en scheikunde, biologie, Engels en Nederlands in gelijke mate betrokken zijn en regelmatig in elkaar overlopen. Elke betrokken docent heeft voor zijn eigen vak bepaald welk onderdeel specifiek in het project aanwezig moet zijn en op welke wijze dat getoetst wordt. Algemene vaardigheden die ingezet worden, zijn met name samenwerken, gebruik van ICT en presenteren.
Samenwerken[3] vanwege het feit dat leerlingen op basis van hun individuele leerstijlen worden ingedeeld in vaste projectgroepen, maar ook regelmatig moeten overleggen in wat wij noemen expertgroepen. De vaste projectgroepen worden gecreëerd uit leerlingen met verschillende leerstijlen, de expertgroepen worden willekeurig ingedeeld.



De jury bekijkt de gemaakte posters

ICT speelt een centrale rol bij de leeronderneming. Er wordt in lokalen gewerkt met een digitaal schoolbord zodat rechtstreeks contact gezocht kan worden met internet. Ook is er een speciale werkruimte gemaakt op de elektronische leeromgeving (elo) waarmee wij op school werken. In deze 'digitale ruimte' is een bibliotheek met allerlei informatie over zaken die de leerlingen tijdens het project tegenkomen. Ook wordt er via deze leeromgeving gecommuniceerd tussen de leerlingen en de vijf hoofdpersonen; er kan bijvoorbeeld gechat of gemaild worden met een hoofdpersoon. Verslagen moeten digitaal worden ingeleverd en docenten geven hierop hun commentaar door er een stukje tekst bij te zetten. Leerlingen kunnen dan direct zien of ze klaar zijn of dat er nog wat verbeterd moet worden. Aan het eind van het project moet elk groepje door middel van PowerPoint en poster-presentatie de dader aanwijzen met zijn/haar motief.

Het project is gebaseerd op probleemgestuurd leren. Hierdoor willen we leerlingen uitdagen om zelf met oplossingen te komen of delen van het verhaal zelf in te vullen. Dat maakt het een dynamisch project dat deels afhankelijk is van de leerlingen en de betrokken docenten.
Voor de begeleidende docenten is het daarom minder van belang dat ze expertise hebben van hun niet-eigen vak; er wordt met name enthousiasme gevraagd voor het probleemgestuurd leren. Een docent hoeft niet met kant-en-klare antwoorden te komen als het gaat om zaken die buiten zijn vak vallen. Hij kan samen met leerlingen op zoek gaan naar een oplossing, bijvoorbeeld via internet of de elektronische leeromgeving. Ook was een vaste technische onderwijsassistent bij het project aanwezig als extra vraagbaak. Uiteraard is, aan het einde van elke dag, de uitwisseling van ervaringen tussen de vakdocenten onderling van groot belang; zo kan het project immers eventueel tussentijds bijgestuurd worden.

Start

Voordat de leeronderneming begint, zijn alle betrokken leerlingen en hun ouders ingelicht over een aankomend project. Ze zijn geïnformeerd over de tijdsduur, 25 lesuren verdeeld over 5 weken, en het moment van afsluiting.

Leerlingen komen in eerste instantie gewoon hun klaslokaal binnen en de les start zoals ze gewend zijn. Opeens komt er een docent scheikunde binnenrennen die zeer overstuur lijkt te zijn, en het verhaal begint…
In het praktijklokaal heerst grote chaos, alle proefopstellingen waar vijf leerlingen van 5H aan gewerkt hebben liggen door elkaar. De scheikundedocent vermoedt dat één van de vijf eindexamenleerlingen sabotage heeft gepleegd en vraagt de klas om hulp. Hij looft een prijs uit om uit te zoeken wie de dader is en wat zijn/haar motief is.

De rode draad van de leeronderneming is het zoeken naar de dader en het motief, alles wat ze doen staat hiermee in verband. De hoofdpersonen (potentiële daders) worden gespeeld door echte leerlingen die voor dit project een andere identiteit aannemen. Deze hoofdpersonen spelen dat ze gedurende het project hun profielwerkstuk moeten afmaken en vragen de klas daarvoor regelmatig om hulp of geven opdrachten. Vervolgens worden de groepjes beloond met aanwijzingen en 'gouden tips'.

Betrokken vakken

Het project duurt 5 weken en elke week is één les van de betrokken vakken gewijd aan de leeronderneming. Het vak Engels wordt gebruikt bij de communicatie met één van de hoofdpersonen, een uitwisselingsstudent uit Schotland. Het vak Nederlands is ondersteunend bij de verslaglegging die van leerlingen wordt gevraagd en bij het betoog waarmee uiteindelijk het project afgesloten wordt. De bètavakken worden tijdens het project gecombineerd. Leerlingen werken bijvoorbeeld aan een leugendetector; ze zoeken uit hoe ze lichamelijke reacties bij het vertellen van onwaarheden kunnen meten en voeren testen uit. Met behulp van rozenblaadjes maken ze een indicator en bepalen de zuurgraad van alcohol. Voor het vak wiskunde is gekozen om een statistisch onderzoek te integreren in dit project, ter vervanging van een aantal opgaven over gegevensverwerking. Doelstelling is dat ze zich bewust worden van zaken die van belang zijn bij een statistisch onderzoek, gegevens op de juiste wijze kunnen verzamelen, verwerken en presenteren. Ze leren gebruik maken van VU-Stat voor het invoeren van gegevens en het maken van diverse tabellen, kruisdiagrammen en cirkeldiagrammen.

Statistiek

In de derde week wordt in de klas een brief bezorgd. Deze blijkt van één van de hoofdpersonen te zijn. Deze 5H-leerling is bezig met zijn profielwerkstuk en moet uitzoeken welke kleur-smaak-combinatie over het algemeen favoriet is. Hij stelt onder andere de volgende vragen:

- Welke kleur of smaak wordt het meest gekozen?
- Is er een verband tussen de gekozen kleur en gekozen smaak?
- Maakt het uit wat de leeftijd is van de ondervraagde bij de keuze?

De docent gaat dan met de klas aan de gang; bij voorkeur worden alleen maar vragen gesteld. Hoe zouden we er achter kunnen komen, wat hebben we dan nodig, wie gaat wat doen enz…? Er komt wat structuur in de ideeën en leerlingen krijgen tips via de elektronische leeromgeving met daarop een module die speciaal ingericht is voor dit project.
Uiteindelijk gaat de klas hiermee in expertgroepjes aan de gang:

- Een groep maakt lolly's aan de hand van een recept met verschillende kleur-smaak-combinaties.
- Een expertgroepje doet theoretisch onderzoek naar het uitvoeren van een statistisch onderzoek door te zoeken op de elektronische leeromgeving en op internet.
- Een andere groep gaat in de mediatheek VU-Stat verkennen aan de hand van werkbladen.

Als de leerlingen door hebben hoe gegevens verzameld kunnen worden en VU-Stat ingedeeld is, gaan alle groepjes samenwerken. Het invulblad waarmee de gegevens verzameld worden, moet tenslotte overeenkomen met de aangemaakte velden in VU-Stat. In eerste instantie ontstaat hier chaos: leerlingen willen samenwerken, maar zien nog niet in hoeverre ze van elkaar afhankelijk zijn. De velden die in VU-Stat ingevoerd moeten worden zijn bepaald door de kleuren en smaken die door de lollymakers zijn gekozen. Het verzamelen van gegevens moet heel precies gebeuren en de groep die daar mee bezig is geweest, moet duidelijk naar de VU-Statters aangeven met welke data ze van plan zijn te komen. Hier is een rol voor de docent weggelegd om het proces te begeleiden en voor de groep wordt de problematiek van een goed statistisch onderzoek dan ook duidelijk.
Een docent kan met name vragen stellen over de manier van interviewen, de gekozen populatie en de keuze van kleur-smaak-combinaties. Voor er werkelijk data verzameld gaan worden is het ook van belang dat er al nagedacht wordt over de eindpresentatie. Kan er wel antwoord gegeven worden op de vragen die aan het begin van de week gesteld zijn met deze data?

Er wordt uiteindelijk een statistisch onderzoek uitgevoerd binnen school. Tijdens pauzes wordt aan leerlingen en docenten gevraagd van de lolly's te proeven en hun ideale kleur-smaak-combinatie te bepalen. Daarnaast worden zaken als leeftijd en jongen/meisje vastgelegd op een invulblad. De bladen gaan naar de VU-Stat-groep die alles invoert en ook de gegevens op verschillende manieren verwerkt.
Tenslotte keren de leerlingen weer terug in hun eigen projectgroep. Deze bestaat dan uit één VU-Stat-leerling, een lollymaker, een interviewer en een leerling die uitgezocht heeft hoe een statistisch onderzoek in zijn werk gaat. Deze groep gaat met behulp van staafdiagrammen, frequentietabellen en kruistabellen de informatie presenteren, en er wordt een verslag gemaakt met antwoord op de gestelde vragen en een eindconclusie. Het verslag leveren ze in bij de betreffende hoofdpersoon; deze heeft tenslotte om het onderzoek gevraagd. Dat doen ze op onze elektronische leeromgeving waarbij ze het verslag via internet in de map doen van de betrokken 5H-leerling. De docent kan aan hun documenten een stukje tekst hangen met commentaar. Dit geeft ze de kans om voordat er definitief beoordeeld wordt nog aanpassingen te doen.

Terugblikken

Bij het nabespreken van de verslagen ***(zie pag. 142)*** komen nog een paar interessante zaken boven tafel. Aangezien de groepen onderling erg afhankelijk zijn van elkaars input, is het samenwerkingsaspect van belang voor een goed verslag. Reflecteren daarop leidt tot leermomenten waarmee ze in de bovenbouw hun voordeel kunnen doen. Ze zien dat het leuk is om samen te werken met een vriendje omdat het gezellig is. Toch kan het effectiever zijn om een verslag te maken met een klasgenoot die over eigenschappen beschikt die bij jezelf wat minder ontwikkeld zijn. Het is goed om dat te benoemen bij het nabespreken.
Daarnaast wordt in het enthousiasme vaak vergeten dat het allemaal gestart is met een

brief van een 5H-leerling die wel graag antwoord wil hebben op zijn/haar vragen en deze dus uit de gepresenteerde tabellen e.d. moet kunnen aflezen. Het is voor leerlingen leerzaam dat dit wordt teruggekoppeld en kennis te nemen van conclusies die getrokken worden uit de gepresenteerde gegevens die leuk klinken, maar inhoudelijk onjuist zijn:
De populairste smaak is kers, de populairste kleur bij de grootste doelgroep is rood. De rode kerslolly is dus de populairste lolly.
Een docent heeft dan even wat overtuigingskracht nodig om leerlingen uit te leggen dat de conclusie op basis van deze redenering niet klopt. Dat is dan even een teleurstelling.

Conclusies

Leerlingen geven in hun verslag aan dat ze het project toch wel lastig vinden. Ze worden geconfronteerd met veel nieuwe dingen zoals VU-Stat en het op de juiste wijze verzamelen van gegevens. Ook moet er goed samengewerkt worden om tot een juist resultaat te komen. Daardoor ervaren leerlingen het project wel als uitdagend en kunnen ze echt trots zijn op hun eindproduct.

Als we het resultaat van het project voor het vak wiskunde vergelijken met het resultaat na 5 reguliere lessen, dan valt het niet tegen. Leerlingen kunnen nu werken met VU-Stat en hebben 'gevoeld' wat het is om een statistisch onderzoek uit te voeren. Het op de juiste wijze interpreteren van een vraag, vertalen naar een statistisch onderzoek en vervolgens antwoord geven op de vraag door je gegevens op de juiste wijze te verwerken en presenteren - dit alles sluit aan bij een aantal kerndoelen voor wiskunde die gedefinieerd zijn voor de basisvorming.

Als een groep op de juiste wijze tot de conclusie komt dat smaak en kleur gekoppeld zijn, zoals geel-banaan, rood-kers enz., dan kan hierop verder ingegaan worden. Hiermee kan een link gelegd worden naar het vak biologie, en dat was ook de bedoeling van dit vakoverstijgende project.

Dit experiment is voor ons als deelnemende docenten erg leerzaam geweest. Het zelf ontwikkelen van een verhaallijn inclusief duidelijk plot was een flinke klus. Afhankelijk van de secties die (op vrijwillige basis) mee wilden doen, moest elk vak een plek en duidelijk op kerndoelen afgestemde inhoud krijgen. Het vakoverstijgende aspect is daardoor lastig implementeren, maar uiteindelijk is het wel gelukt om op 'projectdagen' de vakken te combineren.

Wat betreft de beoordeling is dat niet gelukt; docenten hebben dus het voor hun eigen vak geproduceerde werk beoordeeld. Dat betekent bijvoorbeeld dat het niet wordt meegenomen in het cijfer als in een betoog inhoudelijke zaken met betrekking tot de bètavakken niet correct zijn weergegeven. Het verslag dat ingeleverd wordt voor statistiek zou uiteraard een tweede beoordeling kunnen krijgen voor Nederlands. Als leerlingen een verslag schrijven over verandering van lichaamstemperatuur en hartslag bij het werken met een leugendetector, dan is het interessant als niet alleen de natuurkundedocent hiervoor een cijfer geeft. Dat is een aspect dat verbeterd kan worden.

Ook was het project een behoorlijke aanslag op het rooster. We voerden dit project uit in twee klassen, terwijl de andere tien gewoon regulier les kregen. Gelukkig is dat met medewerking van behulpzame roostermakers en medecollega's op te lossen, maar het was puzzelen. Dit project leent zich beter voor een activiteitenweek waarbij leerlingen zich echt helemaal kunnen onderdompelen in het verhaal en aan het einde van de week de misdaad opgelost hebben. Docenten kunnen zich er dan ook helemaal vrij voor maken en dat komt dan de onderlinge communicatie en werkdruk ten goede; nu moest het echt overal tussendoor.

Toch is al het werk niet voor niets; het enthousiasme van leerlingen die bezig zijn om de daders te vinden maakt alles goed. Ze voeren experimenten uit, zijn creatief met oplossingen, overleggen verdachte zaken in groepjes met elkaar en hebben chat-sessies met de hoofdpersonen. Omdat onze hoofdpersonen ook echte leerlingen zijn die rondlopen, hebben die het een paar weken lastig, maar ze genieten ook erg van alle aandacht. De vragen en conclusies waarmee tweedeklassers komen, bieden je een blik in hun belevingswereld en geven je de kans om je onderwijs daarop aan te sluiten.

Afsluiting

Op de laatste dag van de leeronderneming maken alle leerlingen een poster en een PowerPoint-presentatie van hun bevindingen. Als team moeten ze dan beslissen wie ze aanwijzen als dader en wat zijn/haar motief is.
Elke groep houdt een betoog van vijf minuten over hun conclusie. Dit doen ze in de collegezaal waar hun ouders bij zitten, de vijf hoofdpersonen en een jury. De jury bestaat uit de betrokken scheikundedocent en het hoofd van de eindexamencommissie, de laatste met pruik en in toga.
De betogen zijn heel divers en leerlingen zijn zeer gedreven de dader aan te wijzen; ze koppelen de informatie en experimenten aan elkaar om hun conclusie te onderbouwen. Na de betogen geeft het hoofd van de eindexamencommissie een samenvatting en meldt dat de scheikundedocent een val opgezet heeft met gekleurde drankjes. De dader is daar ingetrapt en die kunnen we dus herkennen aan een blauwe tong. Elke hoofdpersoon wordt vervolgens gevraagd om zijn tong uit te steken. Inderdaad blijkt slechts één van de hoofdpersonen een blauwe tong te hebben en deze dader wordt vervolgens spraakmakend afgevoerd.
Voor het beste betoog is er een prijs en de leerlingen worden door de jury uitgebreid bedankt voor hun hulp bij het zoeken naar de dader. Samen met hun ouders krijgen ze nog een gekleurd (alcoholvrij) drankje ter afsluiting.

Noten

[1] E. Vos, P. Dekkers, E. Reehorst (2003): *Verhalend Ontwerpen.* Groningen: Wolters-Noordhoff; ISBN 90-01-2037-2.
[2] Het enthousiasme van Peter v.d. Berg (vakdocent biologie) is de aanzet geweest voor het project, dankzij hem en Suzan v.d. Helm (vakdocent natuurkunde/scheikunde) heeft de leeronderneming de vorm die hier beschreven is.
[3] S. Ebben, S. Ettekoven, J. van Rooijen (1997): *Samenwerkend leren.* Groningen: Wolters-Noordhoff; ISBN 90-01-30750-7.

Over de auteur

Marjanne de Nijs was tot voor kort eerstegraads docent wiskunde op het Alfrink College te Zoetermeer, een havo/vwo-school met ca. 1800 leerlingen. Vanaf 1 januari 2008 is ze werkzaam op de lerarenopleiding van de Hogeschool Utrecht.
E-mailadres: *nijs0471@planet.nl*

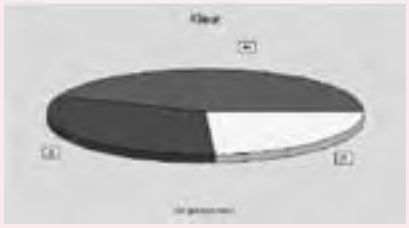
Leerlingmateriaal met hun eigen tekst

De teksten van de leerlingen bij de diagrammen staan tussen [en].

Het aantal ondervraagde jongens en meisjes van de 100 ondervraagden

[Hieruit blijkt dat er meer meisjes zijn ondervraagd dan jongens. Het kan dus zo zijn dat meisjes lolly's lekkerder vinden dan jongens.]

De gekozen kleur van de ondervraagden



[De rode kleur = Rode lolly, de groene kleur = Gele lolly, en de gele kleur = Blauwe lolly. Hieruit blijkt dat de rode lolly het meest gekozen is, opgevolgd door de gele lolly en als laatste de blauwe lolly.]

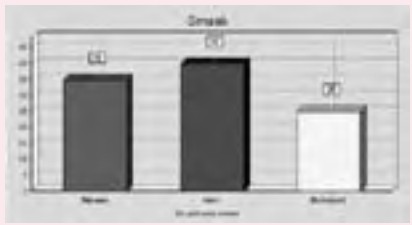
De leeftijd-kleur verhouding

[Hieruit blijkt dat er verschil is bij de voorkeur van kleur tussen de leeftijdsgroepen, bij 10-19 is de rode lolly het populairste, bij 20-29 ook, Bij 30-39 is de blauwe lolly het populairste, Bij 40-49 is de gele lolly het populairste en bij 50-59 was er maar 1 ondervraagde en die vond de blauwe lolly het lekkerste. Ook blijkt hieruit dat 10-19 de grootste doelgroep is die is ondervraagd.]

Jongen-Meisje-Kleur verhouding

[Hieruit blijkt dat er ook een verschil is tussen de jongens en meisjes bij de kleurvoorkeur. Dit verschil is echter wel kleiner. De jongens vinden de Gele lolly het lekkerst maar snel gevolgd door de rode lolly, de blauwe lolly is niet zo populair bij de jongens. Bij de meisjes is de rode lolly duidelijk de populairste lolly, daarna is blauwe lolly het populairste en tot slot de gele lolly.]

De gekozen smaak van de ondervraagde



[Hieruit blijkt dat de kers de meest gekozen smaak is met 40 keer, gevolgd door de banaan met 35 keer, en tot slot de bosvrucht met 25 keer.]

Evaluaties van twee groepen over het statistisch onderzoek

1.

EINDCONCLUSIE: De populairste smaak is kers, de populairste kleur bij de grootste doelgroep is rood. De rode kers lolly is dus de populairste lolly.

Evaluatie van de hele groep: We vonden het onderzoek best wel leuk want het is wel interessant om te weten wat de populairste lolly is. Het onderzoek ging goed want we werkten goed samen. Onze lolly's waren ook goed gelukt. We hadden op tijd de benodigde gegevens om het in VU-Stat te zetten doordat we op school veel leerlingen hadden ondervraagd en veel leerlingen reageerden positief op het onderzoek. Iedereen heeft zijn opdracht op tijd af kunnen krijgen dus dat is ook erg mooi.

We zijn erg tevreden over het onderzoek en we hebben allemaal ons best er voor gedaan. We hebben er best veel van geleerd want eerst konden we niet zulke diagrammen maken als we nu hebben gemaakt. We vinden het programma VU-Stat ook erg handig en zullen het misschien later ook eens gebruiken.
We hopen dat we een goed statistisch onderzoek hebben samengesteld.

2.

Als groep was het niet makkelijk dit onderzoek te doen. Onze lolly's waren mislukt en daarom moesten we samen met andere groepjes meelopen. We waren daar erg blij mee, want niet iedereen had dit voor ons gedaan. Het onderzoek zelf was erg leuk om te doen. VU-Stat is nieuw voor ons en het was een nieuwe ervaring. We moesten leren hoe we al onze gegevens in konden voeren en er een goede grafiek van konden maken. Dit hebben we met drie mensen uit ons groepje geleerd. De andere drie hebben hun eigen opdracht goed gedaan, zodat wij verder konden met die van ons. De samenwerking was redelijk, twee mensen hebben mij geholpen met mijn VU-Stat opdracht. Ik ben erg blij met deze hulp want zonder was het nog erg moeilijk geworden om het op tijd in te leveren. De samenwerking tussen de expertgroepjes was een stuk beter. Als ik vastliep stond er altijd wel iemand klaar om het even uit te leggen of om te zeggen hoe het wel moest.
We zijn erg tevreden met het resultaat. We hebben hard ons best gedaan op deze opdracht, maar zijn wel blij het voorbij is omdat het toch erg veel tijd kost en niet de makkelijkste opdracht is.

Probleemgeoriënteerde statistiek en kansrekening binnen wiskunde A/C

[Bert Nijdam]

Samenvatting
Een belangrijke vernieuwing in de statistiek en kansrekening in het gewijzigde programma wiskunde A en C, die de werkgroep SKACA (Statistiek & Kansrekening binnen wiskunde vwoA, vwoC en havoA) voorstaat, is een probleemgeoriënteerde aanpak. Als voorbeeld bespreken we in dit artikel de interessante statistische vraag of jongens meer ruimtelijk inzicht hebben dan meisjes. Hoe onderzoek je dit? Hoe kan een eventueel verschil in ruimtelijk inzicht het best worden uitgebeeld en hoe in een maat worden uitgedrukt? Ook de interpretatie van een verschil en een kritische blik hierop komen aan de orde. Is een optredend verschil significant en relevant?

Vernieuwing programma statistiek en kansrekening

Op instigatie van cTWO heeft de werkgroep SKACA zich gebogen over het onderdeel Statistiek en Kansrekening (S&K) voor de nieuwe examenprogramma's wiskunde A en C vanaf 2011. In augustus 2007 zijn op basis hiervan nieuwe eindtermen voor de S&K opgesteld en is dit in het rapport van de werkgroep gemotiveerd en voorzien van een voorbeeld-leerlijn (zie Conceptrapport SKACA onder wiskunde A of C via *www.ctwo.nl*).
De veranderingen die de werkgroep voorstaat, betreffen vooral de opbouw. Kort samengevat: de S&K zal worden aangeboden via een probleemgerichte aanpak vanuit interessante probleemgebieden: het vergelijken van groepen, verbanden leggen tussen verschijnselen, het voorspellen van zo'n verschijnsel en het nemen van een optimale beslissing in onzekere situaties. Sommige problemen worden gedurende de leerjaren herhaald en uitgediept (concentrische aanpak). Zo komt bijvoorbeeld het toetsen van een hypothese in klas 4 vwo impliciet aan de orde bij een significant verschil, wordt dit in klas 5 geformaliseerd in een toetsingsprocedure en wordt hieraan in klas 6 de kans op een juiste beslissing toegevoegd.
Verdere aanbevelingen:

- de leerstof, ook de kansrekening, staat in dienst van de oplossing van het gestelde probleem;
- aanbrengen van een onderzoekshouding via de empirische onderzoekscyclus;
- ter verhoging van de motivatie zo mogelijk werken met eigen data dan wel met een bestaande dataset met ruime mogelijkheden;
- waar mogelijk gebruik maken van ICT; zelf rekenen alleen ten behoeve van begrip.

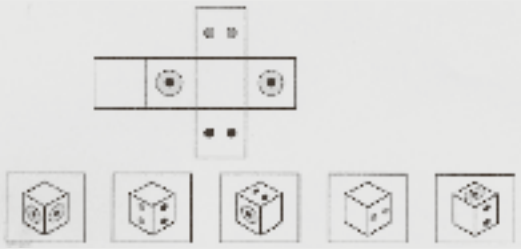
Of het voorgestelde programma inderdaad de hogere motivatie bij de leerlingen teweeg brengt en het een verwachte hogere beklijving heeft, moet in experimenten worden onderzocht. Op dit moment wordt gewerkt aan nieuwe lessen, die in het voorjaar op drie scholen in klas 4 worden gedraaid. Na verwerking van de ervaringen zullen in het nieuwe schooljaar meer scholen ermee gaan proefdraaien. Bij een goed verloop kan er in september 2009 algemeen mee worden begonnen in klas 4 vwo.
Het volgende geeft een voorbeeldopzet van een serie lessen over een concrete probleemstelling.

Aanzwengelen van een probleem

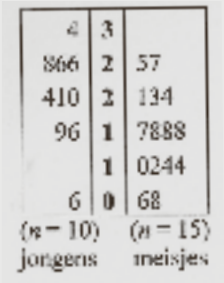
Tijdens onze vakanties met de auto naar het zuiden reed ik meestal en zat mijn vrouw naast me. Zij had dan de wegenkaart voor zich liggen en af en toe vroeg ik de weg. Nu liep dat dikwijls niet helemaal in goede banen en dan greep ik de kaart en raadpleegde deze zelf. Het liefst onder het rijden, maar dat vond zij (terecht) niet goed. Gelukkig ben ik nu in het rijke bezit van een routeplanner en is het hele probleem van de baan. Maar wat mij opviel is dat het beperkte kaartlezen van mijn vrouw ook door mijn vrienden met hun vrouwen werd ervaren. Je vraagt je af of dit toeval is of dat vrouwen in het algemeen slechter kunnen kaartlezen dan mannen. Wel viel het mij op dat mijn dochter prima met de kaarten overweg kan, alhoewel mijn zoons er toch weer handiger in zijn.
Ik wil het probleem toch nog iets algemener stellen. In de tijd uit mijn leraarschap wiskunde (nog in de jaren '60) gaf ik stereometrie en daarin viel op dat veel meisjes moeite hadden zich een goede voorstelling te vormen van de platgetekende ruimtefiguren. Eigenlijk zou ik wel willen stellen dat meisjes over minder ruimtelijk inzicht beschikken dan jongens. Toch is er reden tot twijfel. Het verschil uit vroeger tijden is nu misschien wel opgeheven vanwege de meer gelijke opvoeding van jongens en meisjes.

Meting van ruimtelijk inzicht

Door een probleem te stellen over ruimtelijk inzicht hebben we het ons niet gemakkelijk gemaakt. Zouden we hebben gesteld dat jongens in het algemeen langer zijn dan meisjes, dan hadden we lengtes van proefpersonen makkelijk op de gebruikelijke manier in cm kunnen meten. Maar ruimtelijk inzicht is een abstracte variabele, die zich niet direct laat meten [1] en waarover we slechts flauwe noties hebben.
Het al of niet kunnen kaartlezen en lezen van stereometrische figuren geven indicaties voor ruimtelijk inzicht, maar dat betreft

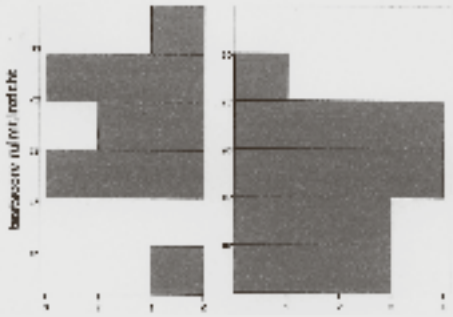


figuur 1 Voorbeelditem uit de test ruimtelijk inzicht



4	3	
866	2	57
410	2	134
96	1	7888
	1	0244
6	0	68
(n = 10)		(n = 15)
jongens		meisjes

figuur 2 Tweezijdig stamgram



figuur 3 Tweezijdig histogram

maar enkele aspecten ervan. De oplossing is om met experts een hele serie indicatoren te verzamelen en die samen op te nemen in een test voor ruimtelijk inzicht. Zoeken op internet naar zo'n test levert diverse mogelijkheden. Via *www.iq-test.nl/oefen-ruimtelijkinzicht* vonden we als voorbeeld de vraag ***uit figuur 1***, die deel uitmaakt van zo'n test. Een beetje serieuze test moet wel 30 tot 60 van dit soort items bevatten. Kies je zo'n test, dan moet je je nog wel twee dingen afvragen. Ten eerste, hoe *betrouwbaar* is de test? De mate van betrouwbaarheid is hoger naarmate het toeval er een geringere rol in speelt. Ten tweede, hoe *valide* is de test? Hiermee geef je aan in hoeverre de meting ook inderdaad het bedoelde ruimtelijk inzicht betreft. Deze zaken laten we even voor wat ze zijn en we beschouwen nu de testscore op een bepaalde test van 40 items als de bedoelde score op ruimtelijk inzicht. Deze test laten we de leerlingen in klas 4 afleggen; we bepalen per leerling als score het aantal vragen goed en voeren die scores in de computer in met een statistisch verwerkingspakket. Meest ideaal is om alle antwoorden per leerling afzonderlijk in te voeren en per leerling nog enkele kenmerken toe te voegen zoals sekse, leeftijd, profiel, wiskundecijfer in klas 3. Hiermee kunnen later vragen over eventuele beïnvloeding worden onderzocht. De leraar die het zelfonderzoek te omslachtig vindt, kan desgewenst gebruik maken van de verzamelde data bij een proefschool. Nu gaan we slechts in op de vraag of het ruimtelijk inzicht bij jongens hoger ligt dan bij meisjes.

Visualiseren van de scores

Bij de verzamelde gegevens willen we vervolgens zien of ons vermoeden tot uiting komt in de scores. Daartoe gaan we de scores van alle meisjes en jongens passend uitbeelden. Omdat de scores in principe kunnen lopen van 0 tot 40, dus bestaande uit twee cijfers, is een tweezijdig steel-blad-diagram (stamgram) erg geschikt (met het tiental van een score in de stam [2] en haar laatste cijfer ernaast).

In figuur 2 is een (fictief) resultaat van tien jongens en vijftien meisjes afgedrukt. Uit dit diagram zie je dat de scores van de jongens als geheel iets hoger aan de stam geplaatst zijn dan die van de meisjes. Het verwachte onderscheid blijkt hier dus aanwezig.

In figuur 3 zijn de scoreverdelingen uitgebeeld door twee ruggelings tegen elkaar geplaatste histogrammen, met als klassen 5-9, 10-14, 15-19, etc. Dit plaatje geeft vrijwel dezelfde informatie als het tweezijdig stamgram, alleen zijn de individuele scores er niet meer zichtbaar.

In figuur 4 is van beide groepen een boxplot getekend. Hierbij is het scoregebied van de middelste 50% scores door een box (reep) aangegeven, het bereik van de 25% laagste en de 25% hoogste scores door een lijn (bezemsteel). Twintig jaar geleden al gaf ik het de sprookjesachtige naam *repesteel*, maar internationaal gebruikt men de naam boxplot. De repesteel geeft de verdeling nog globaler aan dan het histogram, maar daarin ligt nu juist de kracht van de statistiek. Je ontdoet de data van details, maar je wint aan globale inzichtelijkheid. Aan de boxplots is duidelijk te zien dat de jongensgroep hoger scoort dan de meisjesgroep. Over de grafieken in de afbeeldingen 2, 3 en 4 wil ik nog opmerken, dat ik de scores op de test verticaal heb uitgezet. Dit is bewust gedaan. In ons probleem zeggen we namelijk dat de testscores bij meisjes en jongens verschillen, dus dat de testscores afhangen van de sekse. Zo zien we de testscores dus als de afhankelijke variabele Y en is sekse de onafhankelijke variabele X.

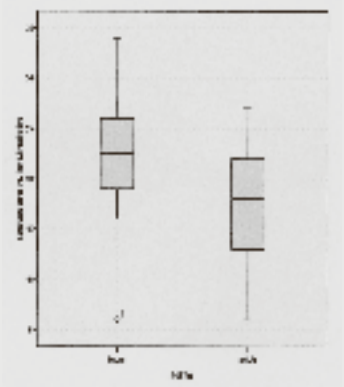
Kwantificeren van een verschil

Aan repestelen *zie* je fraai hoe groot het verschil tussen twee verdelingen is, maar dit zicht is niet objectief. Daarvoor heb je een *maat* voor het verschil nodig, bijvoorbeeld met getalwaarden tussen 0 en 1, waarbij 0 geen enkel verschil aangeeft en 1 het grootst mogelijke verschil. Net zoals er verschillende maten voor centrum of spreiding bestaan, zijn er ook diverse maten om een verschil tussen twee verdelingen in uit te drukken, elk met hun voor- en nadelen. In ons voorbeeld komen in aanmerking:

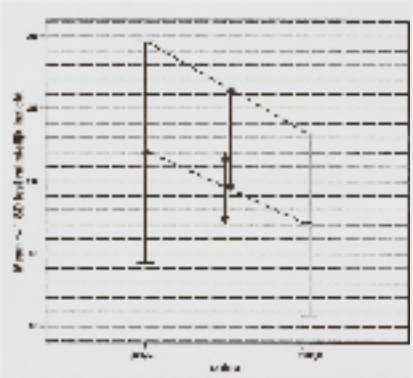
- Δ = maximale verschil tussen beide cumulatieve frequentieverdelingen, en:
- d = effectgrootte.

Δ ligt tussen 0 en 1 in, de meest gebruikte maat d kan boven 1 uitkomen.

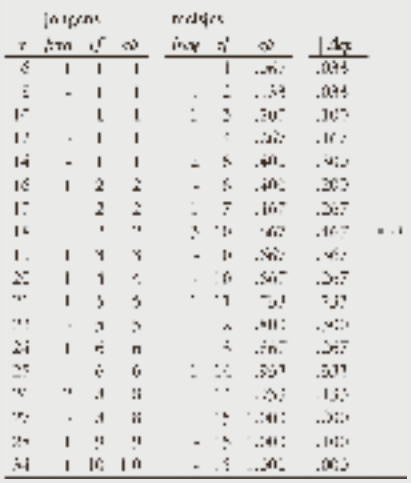
De effectgrootte d kun je definiëren als het absolute verschil tussen beide gemiddelden, gerelateerd aan de spreiding binnen de groepen en wel door het verschil te delen door het gemiddelde van de beide standaardafwijkingen. Het is ***in figuur 5*** met een 'error-bar' uitgebeeld. Van beide verdelingen zijn de gemiddelden door een



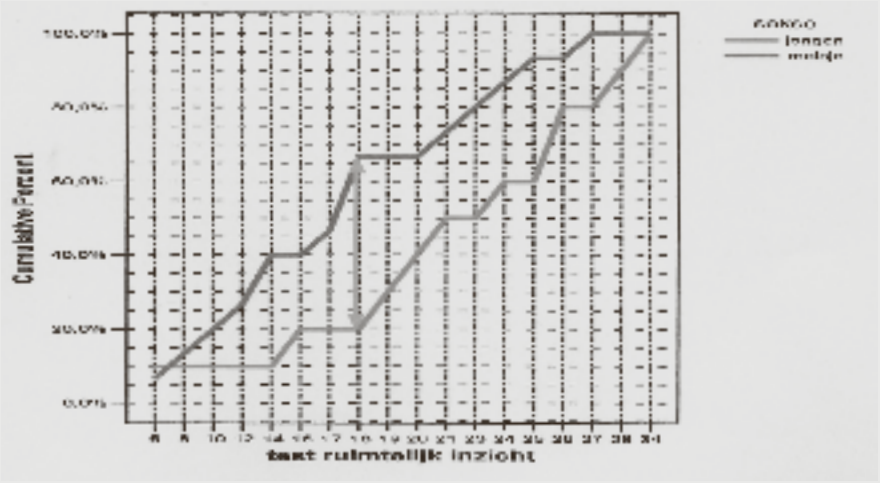
figuur 4 Boxplots (repestelen)



figuur 5 'Error-bar' met verschil tussen gemiddelden en de gemiddelde standaardafwijking



figuur 6 Frequentietabel met berekening verschilmaat Δ



figuur 7 Kwantielgrafieken (cumulatieve proportiepolygonen) ter illustratie van verschilmaat Δ

open punt uitgezet. Hieronder en hierboven zijn stelen (bars) ter lengte van één standaardafwijking geplaatst. Tussen beide bars in hebben we het verschil in gemiddelden uitgezet met een tweezijdige pijl en de gemiddelde standaardafwijking met een verticale tweezijdige steel. De effectgrootte is de verhouding tussen de pijllengte en de steellengte.

Berekening met de calculator of VU-Stat levert $\bar{y}_j = 22, \sigma_j = 7,22$ en $\bar{y}_m = 17, \sigma_m = 6,12$, dus $d = \frac{5}{6,67} = 0,75$. Deze waarde komt ongeveer overeen met de *d* voor het lengteverschil tussen 13- en 14-jarige meisjes en dat wordt over het algemeen gezien (zie Cohen) als een tamelijk groot verschil. Doorgaans worden *d*-waarden kleiner dan .4 in de sociaal-economische en gezondheidswetenschappen pas gering genoemd, bij *d*-waarden groter dan .8 spreekt men van een sterk effect [3].

Of we in ons voorbeeld de effectgrootte *d* inderdaad moeten gebruiken, valt nog te bezien. Die maat maakt immers gebruik van gemiddelden en standaardafwijkingen en die hebben zin, als de meetschaal het begrip 'ruimtelijk inzicht' op intervalniveau meet, dus als gelijke afstanden op de schaal ook gelijke verschillen betreffen bij het abstracte begrip ruimtelijk inzicht. Of dit zo is, hangt af van de gebruikte items in de test. Heb je bijvoorbeeld tien items die je met een bepaald inzicht niet kunt oplossen, maar met een beetje meer inzicht wel, dan scoren twee personen net onder en boven dat inzicht opeens tien punten meer, terwijl hun inzicht slechts weinig verschilt. Om een interval meetniveau te bereiken moeten de items gelijkmatig spreiden in moeilijkheid. Zit de test zo niet in elkaar, dan is *d* niet verantwoord en moet je een verschilmaat gebruiken die ook voor ordinaal meetniveau toepasselijk is en dat is Δ.

De bepaling van Δ is voor ons voorbeeld uitgewerkt via de tabel ***in figuur 6***. Van beide verdelingen is eerst de kolom met frequenties en met cumulatieve frequenties *cf* gemaakt, waarvan de laatste via deling door *n* is omgezet in cumulatieve proporties *cp*. Daarna is een kolom gemaakt voor de absolute verschillen tussen de cumulatieve proporties en de hoogste waarde .467 is onze Δ-waarde. $\Delta = \max_y |cp_j(y) - cp_m(y)|$ is het maximale verschil tussen beide cumulatieve verdelingen. De berekende waarde .467 ligt boven .30 en dat wordt meestal gezien als een groot verschil. Merk op dat in de bepaling van Δ geen gebruik wordt gemaakt van de scores zelf, maar alleen maar van de volgorde van de scores. Er is slechts een ordinaal meetniveau nodig.

In figuur 7 zijn de 'kwantielgrafieken' van beide verdelingen op een ordinale schaal getekend en is Δ door de tweezijdige pijl uitgebeeld als de grootste verticale afstand tussen beide grafieken.

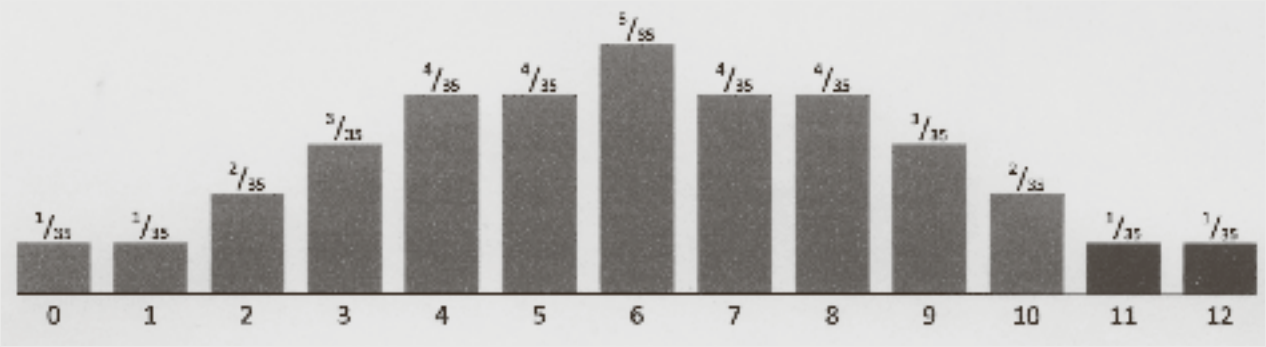
Dat het verschil via Δ als groot wordt aangemerkt en volgens *d* slechts als vrij groot, is niet verwonderlijk. Er is globaal wel degelijk een groot verschil, maar bij de jongens is er een lage uitschieter, die het gemiddelde relatief laag en de standaardafwijking hoog doet uitvallen en dan is het verschil in gemiddelde met dat van de meisjes minder groot. Zo'n uitschieter is naast het te lage meetniveau ook een bezwaar tegen het gebruik van de verschilmaat *d*.

Is het verschil significant?

In onze onderzoeksgroep is dus een verwacht verschil gevonden, een groot verschil zelfs. Toch is het nog niet verantwoord nu te roepen dat meisjes minder ruimtelijk inzicht hebben dan jongens. Mogelijk vormen onze 25 vwo-4-leerlingen geen goede afspiegeling van de gehele vwo-4-populatie. Dit kan een systematische afwijking zijn, omdat de school van de leerlingen geen doorsneeschool is, maar de afwijkingen kunnen ook toevallig zijn. Het toeval zit enerzijds in een (ongelukkige) keuze van de 25 betrokken leerlingen en anderzijds in de scoring op de test. De test voor ruimtelijk inzicht bestaat uit 40 meerkeuzevragen en bij sommige vragen zal zeker naar het antwoord worden geraden. En mogelijk had een aantal meisjes in de groep een offday en scoorden zij toevallig onder hun niveau.

We gaan eerst in op het controleren op toevallige fouten. Is ons resultaat mogelijk nog aan toeval te wijten of is het verschil meer dan toevallig, *significant*?

Volgens het toetsingsprincipe moeten we in de veronderstelling dat er geen systematisch verschil tussen beide groeperingen aanwezig is, nagaan hoe groot de kans is op minimaal een verschil als gerealiseerd in de onderzochte steekproef. Is die kans kleiner dan of gelijk aan een gekozen drempelwaarde α van (meestal) 5%, dan is het resultaat significant. Om de gevraagde kans te bepalen zijn er voor twee groepen diverse mogelijkheden, afhankelijk van welke informatie uit de steekproeven je benut, bijvoorbeeld de Z- of de T-toets voor twee groepen, de ANOVA-toets of de Mann-Whitney-U-toets. Van deze toetsen komen de eerste twee wel in aanmerking voor behandeling in het vwo, maar hier zitten qua voorwaarden en techniek zoveel haken en ogen aan, dat ik er niet enthousiast over ben. Dat geldt wel voor de *U-toets*, die let op het verschil in ligging, in heel veel gevallen bruikbaar is, waarvan het principe een begrijpelijke combinatorische oplossing betreft en die ook praktisch goed uitvoerbaar is. We zullen het principe aan de hand van een sterk vereenvoudigd voorbeeld eerst uit de doeken doen en de toets daarna



figuur 8 Kansverdeling van U bij steekproeven van omvang 3 en 4

toepassen op ons voorbeeld met de 25 leerlingen.
Als minivoorbeeld kiezen we een steekproefje van de zeven kinderen uit één gezin. Daarin halen de drie meisjes op de test voor ruimtelijk inzicht de scores 11, 14 en 17, en de vier jongens de scores 16, 21, 23 en 24. Omdat de test het ruimtelijk inzicht waarschijnlijk niet meet op interval- en wel op ordinaal meetniveau, willen we alleen de volgorde van de scores benutten. Geven we een jongen aan met een *j* en een meisje met een *m*, dan is de volgorde van jongens en meisjes op ruimtelijk inzicht van laag naar hoog: *m-m-j-m-j-j-j*. Bijna alle meisjes komen eerst, daarna de jongens. Alleen de volgorde *m-m-m-j-j-j-j* is nog extremer. Hoe groot is de kans op zo'n resultaat als jongens en meisjes eigenlijk niet verschillen op ruimtelijk inzicht, dus als de plaatsing van jongens en meisjes puur toeval is? In dat geval is elke volgorde van drie meisjes en vier jongens even waarschijnlijk. Het aantal mogelijk volgorden is gelijk aan het aantal keuzes van drie meisjes uit zeven plaatsen en dat aantal is gelijk aan $\binom{7}{3}=35$. De kans op de verkregen volgorde *m-m-j-m-j-j-j* of de extremere volgorde *m-m-m-j-j-j-j* is dan $\frac{2}{35}=5{,}7\%$. Omdat deze kans groter is dan 5%, noemen we het steekproefresultaat *niet* significant.
In het zo toegepaste volgordeprobleem zou je als volgorde hebben kunnen krijgen *m-j-m-j-m-j-j*, ook een volgorde met de meisjes voornamelijk eerst. De vraag is dan of bijvoorbeeld de volgorde *m-m-j-j-j-m-j* extremer is of niet. Je kunt deze vraag aanpakken door te tellen hoeveel meisjes aan elke jongen voorafgaan en die aantallen op te tellen. Bij de volgorde *m-j-m-j-m-j-j* zijn dit er 1 + 2 + 3 + 3 = 9 en bij de volgorde *m-m-j-j-j-m-j* zijn het er 2 + 2 + 2 + 3 = 9, dus evenveel. Het zo berekende somaantal is de grootheid *u* van de Mann-Whitney-U-toets. In ons minivoorbeeldje is de gerealiseerde *u*-waarde gelijk aan 2 + 3 + 3 + 3 = 11 en de extreemst mogelijke *u*-waarde is 3 + 3 + 3 + 3 = 4 × 3 = 12.
Door nu systematisch alle 35 volgorden uit te schrijven en per volgorde de *u*-waarde erachter te vermelden kun je tellen hoeveel volgorden leiden tot een bepaalde *u*-waarde. Bijvoorbeeld tot *u* = 9 leiden drie volgorden: *m-m-j-j-j-m-j*, *m-j-m-j-m-j-j*, *j-m-m-m-j-j-j* (2+2+2+3, 1+2+3+3, 0+3+3+3), tot *u* = 6 vijf volgorden (1+1+2+2, 1+1+1+3, 0+1+2+3, 0+2+2+2, 0+0+3+3). Omdat alle volgorden even waarschijnlijk zijn (bij evenveel ruimtelijk inzicht voor meisjes en jongens), is de kans op een *u*-waarde gelijk aan het aantal volgorden bij de *u*, gedeeld door het totaal aantal volgorden 35. De gehele kansverdeling van de grootheid *U* ziet er dan uit als ***in figuur 8*** is getekend.
Van de kansverdeling van *U* hebben we in ons gestelde probleem alleen de rechter overschrijdingskans van de gerealiseerde waarde *u* = 11 nodig, en die is gelijk aan het zwart weergegeven deel $^2/_{35}$.
Voor grotere groepen is het niet meer zo gemakkelijk de gehele kansverdeling van *U* op te stellen of benodigde overschrijdingskansen te bepalen. Gelukkig blijkt de kansverdeling van *U* dan erg goed te worden benaderd door een normale kansverdeling en via die benadering kan een overschrijdingskans worden bepaald. Daarbij heb je dan de *karakteristieken* van de *U*-verdeling nodig en daarvoor bestaan de volgende formules:
- verwachting (gemiddelde) $E_U=\frac{1}{2}n_1n_2$, in ons voorbeeldje $E_U=\frac{1}{2}\cdot 3\cdot 4=6$ (zie figuur);
- standaardafwijking $S_U=\sqrt{n_1n_2(n_1+n_2+1)/12}$, in het voorbeeld $S_U=\sqrt{3\cdot 4\cdot(3+4+1)/12}=2{,}83$.

Wellicht vallen deze formules wat rauw op uw dak, maar in het onderwijs kun je ze voor het voorbeeld op elementaire wijze laten berekenen en dan laten controleren met de formule. De formule voor de standaardafwijking is te lastig om via inductie te bewijzen, maar de formule voor de verwachting is eenvoudig plausibel te maken. In een situatie waarin jongens en meisjes niet verschillen in ruimtelijk inzicht, verwacht je over de volgorde van jongens en meisjes een symmetrische ligging en dat leidt tot de *u*-waarde $6=\frac{1}{2}n_1n_2$. Ook in de figuur zie je dat de kansverdeling van *U* symmetrisch rond de waarde 6 ligt.
Zou je de overschrijdingskans van *u* = 11 in ons minivoorbeeldje bepalen via de normale benadering, dan zou je de kans 5,6% krijgen, terwijl de kans eigenlijk 5,7% is. Je ziet dat de normale benadering al bij geringe groepsomvangen een heel aardige kansschatting oplevert. En de benadering blijkt beter te worden naarmate de groepsomvangen groter zijn. We gaan haar toepassen in ons voorbeeld met de tien jongens en vijftien meisjes.
De gerealiseerde *u*-waarde bepalen we vanuit de frequentieverdelingen in figuur 6. We tellen per jongen hoeveel meisjes een lagere score hebben en tellen daarna al die bijdragen op. De jongen met score 6 heeft geen enkel meisje onder zich, maar wel een meisje met een gelijke score 6. Die tellen we ½ eronder (en half erboven, maar daar letten we niet op). De jongen met score 16 heeft 6 meisjes onder zich en de twee jongens met scores 19 en 20 elk 10. De jongen met score 21 heeft 10½ en de jongen met score 24 12½ meisje onder zich.
Tenslotte tel je onder de twee jongens met score 26 elk 14 meisjes onder de twee jongens met scores 28 en 34 elk alle 15 meisjes. Zo vind je als *u*-waarde:

$u = ½ + 6 + 10 \cdot 2 + 10½ + 12½ + 14 \cdot 2 + 15 \cdot 2 = 107½$

De kansverdeling van U heeft bij een verondersteld gelijk ruimtelijk inzicht bij jongens en bij meisjes als:

- verwachtingswaarde $E_U = \frac{1}{2} \cdot 10 \cdot 15 = 75$, en als
- standaardafwijking $S_U = \sqrt{10 \cdot 15 \cdot (10+15+1)/12} = 18{,}03$.

De rechter overschrijdingskans van 107,5 kun je met VU-Stat of de grafische rekenmachine vinden via de normale verdeling met verwachting 75 en standaardafwijking 18,03 en als interval dat tussen 107 en oneindig. Je laat het interval lopen vanaf 107 en niet vanaf 107,5 omdat je bij de discrete U-verdeling met u-stapjes van 1 op de gerealiseerde waarde 107,5 de discontinuïteitscorrectie van -½ moet toepassen. De gevraagde kans is gelijk aan 3,8%. Deze kans is kleiner dan $\alpha = 5\%$, en daarom mag je het verschil tussen jongens en meisjes *significant* noemen!

De conclusie uit ons onderzoek met de fictieve data is, dat je mag aannemen dat jongens in 4-vwo tegenwoordig meer ruimtelijk inzicht bezitten dan meisjes, althans op de onderzochte vwo-school. Het verschil is tussen de onderzochte groepen zo groot, dat het verschil ook bij de tamelijk kleine groepen dus al significant is.

Over de U-grootheid merken we nog twee bijzonderheden op.

De eerste is dat het in principe niet uitmaakt of je bij de telling uitgaat van de meisjes of van de jongens en of je telkens telt hoeveel van de andere groep eronder of erboven liggen. Zo ontstaan er weliswaar twee verschillende u-waarden, maar die liggen symmetrisch in de U-verdeling en dan zijn de bijbehorende linker- respectievelijk rechter-overschrijdingskans gelijk.

De tweede opmerking gaat over het voorkomen van gelijke scores. Als je alle scores op volgorde sorteert, moet je gelijke scores eigenlijk op elkaar stapelen; ze vormen dan een 'knoop' in de volgorde. Het gevolg van deze knopen voor de kansverdeling van U is, dat de standaardafwijking S_U er een tikkeltje kleiner door wordt. Er bestaat een hierop aangepaste formule voor S_U, maar deze knopencorrectie zet pas zoden aan de dijk als er hele vette knopen zijn. Vergeet daarom de knopencorrectie maar, dan toets je soms iets conservatief: je noemt een verschil minder gauw significant.

Alternatieve verklaringen en relevantie

In ons voorbeeld gaven de twee groepen een groot verschil op ruimtelijk inzicht te zien, wat ook als significant is aangemerkt. Mag je nu dus aannemen dat meisjes in 4-vwo minder ruimtelijk inzicht hebben dan jongens? Dit zou je mogen doen als de meisjes en jongens een aselecte steekproef zouden vormen uit alle 4-vwo-ers. Je spreekt dan van *statistisch generaliseren*. Maar onze groepen komen van één vwo-school, dus die vormen geen aselecte keuze uit alle 4-vwo-ers. Toch zou je het resultaat mogelijk nog *inhoudelijk* naar alle 4-vwo-ers mogen *generaliseren*, als je genoeg indicaties hebt dat de onderzochte groep representatief is voor alle 4-vwo-ers, dus qua achtergronden identiek. Je kunt dit nooit compleet nagaan, maar wel ten aanzien van enkele belangrijke aspecten zoals milieu, vrijetijdsbesteding, etniciteit of religie.

Heb je een gevonden verschil significant bevonden, dan geeft een verschilmaat belangrijke informatie over de grootte van het verschil. Soms evenwel is een klein verschil al relevant, bijvoorbeeld als het om twee geneesmiddelen gaat die van levensbelang zijn voor een patiënt, soms is een groot verschil pas relevant. In ons probleem over ruimtelijk inzicht zul je voor het aantrekken voor een baan waarvoor ruimtelijk inzicht vereist wordt, niet zo gauw een meisje kiezen als het verschil tussen meisjes en jongens groot is; bij een klein verschil maakt het je niet zoveel uit. Wat in een bepaald probleem wel of niet relevant is, hangt dus af van het belang.

Tot slot

De voorgestelde aanpak bevat diverse grondbeginselen van de statistiek en is elementair in haar aanpak. Het kan naar mijn idee al deels in 4-vwo worden uitgevoerd, met enkele onderdelen later (zoals de normale benadering van de U-verdeling). De hier voorgestelde normale benadering zou ik zelfs wel als eerste kennismaking met de normale verdeling willen zien en niet via een of andere realistische populatieverdeling. Een echte normale populatieverdeling en zelfs een goede benadering ervan is zeldzaam! Dat ligt anders voor kansverdelingen. De kansverdeling van U of van een steekproefgemiddelde, of van een verschil tussen twee gemiddelden, wordt al voor niet al te kleine steekproeven goed normaal benaderd, zelfs als de populatieverdelingen nogal afwijken van een normale vorm.

Of een en ander ook echt zal lukken zullen de experimenten moeten uitwijzen die vanaf januari 2008 door enkele docenten worden gestart.

Verantwoording

Dit artikel werd geschreven met instemming van de werkgroep SKACA. Bij deze dank ik de werkgroep en met name Gerard Koolstra voor de opbouwende commentaren.

Noten

[1] Bij probleemgericht onderwijs zullen veelal dergelijke abstracte begrippen voorkomen, zoals magnetisme, aantrekkingskracht of leiderschap. Het meten ervan is een onderdeel dat van algemene aard is en in diverse schoolvakken aan de orde moet komen. Het optreden van systematische en toevallige meetfouten kan het beste in de statistiek aan de orde worden gesteld.

[2] Voor een genuanceerder beeld zijn per tiental twee takken met bladen 0-4 en 5-9 onderscheiden.

[3] De naam effectgrootte is ontleend aan experimenteel onderzoek, waarbij men de experimentele groep na toediening van een stimulus vergelijkt met een gelijkwaardige controlegroep zonder die stimulus. Het verschil in resultaat wordt dan gezien als het effect van de stimulus. De effectgrootte is een belangrijke maat voor de grootte van een verschil, en het aardige ervan is dat zo ook het nut blijkt van de maten gemiddelde en standaardafwijking. Omdat de maten gemiddelde en standaardafwijking voor één groep niet zo heel veel zin hebben, lijkt het me verstandig de standaardafwijking pas in te voeren op het moment dat je de effectgrootte *d* introduceert.

Literatuur

- J. Cohen: *Statistical power analysis for the behavioral sciences*. Mahwah (NJ, USA): Lawrence Erlbaum.
- E.L. Lehman: *Nonparametrics: Statistical methods Based on Ranks*. San Francisco (USA): Holden-Day.
- A.D. Nijdam: *Statistiek in onderzoek*, deel 1 en 2. Groningen: Wolters-Noordhoff.

Over de auteur Bert Nijdam

1964-1972: leraar wiskunde havo/vwo/kweekschool;
1971-1975: medewerker IOWO; hoofdauteur van *Kansrekening & statistiek voor het vwo*;
1975-2006: docent statistiek bij sociale wetenschappen UU; auteur universitaire leerboeken statistiek;
2007: lid werkgroep SKACA en ontwerper lesmateriaal S&K.
E-mailadres: *A.D.Nijdam@uu.nl*

De kanskoffer

EEN IDEE VOOR DE SECTIE WISKUNDE

[Joke Verbeek]

Inhoud kanskoffer

1 roulettespel
2 dubbel kaartspel in houder
3 rad van avontuur in doos
4 kanstol van pim-pam-pet
5 zakje bingo/lottogetallen in doos
6 houder met 5 dobbelstenen
7 doos met gekleurde dobbelstenen
8 25 pionnen + dobbelsteen
9 zakje met 20 balletjes (zwart, blauw, wit, geel)
10 mastermind
11 doosje met 50 punaises
12 5 munten in etui
13 doosje paperclips
14 bierviltjes
15 fruitautomaat in doos
16 doosje pokerstenen
17 yahtzee kaartjes
18 lottoballenpen

In het deksel

a sheets met boomdiagrammen
b verwisselbare schijven voor de kanstol
c boekje casino met spelregels
d gebruiksaanwijzing bij de koffer
e opdrachtkaarten voor bij de fruitautomaat
f floppydisk met vu-kans
g kaartjes (rood en blauw)
h ganzenbord

Maatschappelijke vorming

Kansrekenen is een onderwerp dat leerlingen wel aanspreekt: wie waagt er niet eens een gokje? Dit onderwerp leent zich er ook voor een en ander te verlevendigen met voorbeeldmateriaal. Al dit materiaal kan worden ondergebracht in een 'kanskoffertje'. Het sectielid dat toe is aan het onderwerp *kans* haalt het koffertje op en gebruikt alles wat van zijn of haar gading is.
Omgaan met kansen is voor elke leerling belangrijk, het hoort bij de maatschappelijke vorming en staat als zodanig in de preambule van de examenprogramma's. Het is van belang dat leerlingen leren omgaan met kansen omdat zij daar in hun leven veel mee te maken krijgen. En dan gaat het niet alleen om het wel of niet aanschaffen van allerlei loten, maar ook over het afsluiten van verzekeringen en het nemen van financiële risico's. Het is goed om daar met leerlingen uitgebreid over van gedachten te wisselen en zo bij te dragen aan hun ontwikkeling tot gecijferd mens, dat kansen kan beredeneren en afwegen en zo beslissingen kan nemen die gebaseerd zijn op ratio en niet enkel op gevoel.

Het kanskoffertje kan als demonstratiemateriaal worden gebruikt in lessen rond het onderwerp kans of het kan worden ingezet als materiaal dat door leerlingen te gebruiken is.

De samenstelling van het koffertje

1. Kijk in elk deel van je wiskundemethode welke materialen in de opgaven of voorbeelden over kansrekenen gebruikt worden. Stop die allemaal in het koffertje, dus ook punaises, munten, speelkaarten en dobbelstenen. Ook elk spel dat genoemd wordt hoort thuis in de koffer.
2. Vul de koffer aan met eigen vondsten. Denk aan een lotto- of toto-formulier, een kraslot, een bingospel, ganzenbord enzovoorts.
3. Ga actief op zoek (in de speelgoedwinkel of de spellenwinkel) naar een roulettespel, een rad van fortuin of dergelijke attributen.
4. Bij elk casino zijn spelregelboekjes te verkrijgen, evenals promotiemateriaal. Dit is goed bruikbaar als extra materiaal.
5. Maak bij het extra materiaal opdrachten voor elk niveau en plastificeer die. Zorg ook voor antwoordvellen. Denk ook aan internetopdrachten.
6. Het koffertje is ook in te zetten bij leerlingen die een sector- of profielwerkstuk willen maken met 'kans' als (deel)onderwerp. Zorg daarom ook voor open opdrachten.
   Tip: Misschien kunnen leerlingen zelf het koffertje samenstellen, als deel van hun sector- of profielwerkstuk.
7. Maak een inhoudsopgave.
8. Zorg voor een verwijzing naar opgaven in de boeken.
9. Maak een inpakwijzer.

Afspraken

- Iedere docent gebruikt het koffertje naar eigen inzicht, maar zorgt er wel voor dat het weer compleet wordt teruggezet.
- Voor de bovenbouw: een hoofdstuk *Kansrekenen* kan prima worden afgesloten met een klassenavond waarbij een croupier van het casino wordt ingehuurd om een avond met kansen te verzorgen. Overleg met de mentor.
- Voor de onderbouw: een (klassen) spelletjesavond of middag is een prima afsluiting van het onderwerp *Kans*. Overleg met de mentor.

Over de auteur
Joke Verbeek is redacteur van *Euclides* en docent wiskunde op het Arentheem College in Arnhem.
E-mailadres: *jokeverbeek@chello.nl*

MATRIX

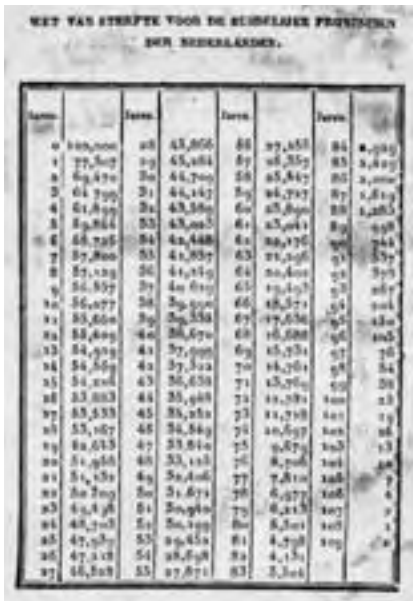
…DAAR MAAKT U ECHTE WISKUNDE VAN. Wiskunde is overal. Maar hoe maakt u het zichtbaar voor al uw leerlingen? Hoe maakt u ze nieuwsgierig? Met de wiskundemethode Matrix van Malmberg kiest u voor een unieke aanpak. Het uitgangspunt: relevante wiskunde. Matrix leert de leerlingen om het vak te snappen. Aan de hand van herkenbare situaties raakt elke leerling verwonderd. Met als gevolg: extra motivatie. Leerlingen worden uitgedaagd om zelf aan de slag te gaan en denken meer en dieper na. Bovendien heeft u als docent alle ruimte om op uw eigen manier ‘de klik’ met de leerlingen te maken. Zo kunt u flexibel inspelen op uw eigen wensen én de verschillende leerniveaus van uw leerlingen. En u kunt moeiteloos overschakelen van het boek op ePack en weer terug. Kortom, u maakt échte wiskunde van wat bij uw leerlingen leeft. Meer weten? Vraag nu een beoordelingsexemplaar van Matrix aan. Bel 073 628 8766. **VOOR JE HET WEET, HEB JE HET DOOR. MATRIX**

MALMBERG

Miscommunicatie in de Nederlandse 19e-eeuwse statistiek

[Ida Stamhuis]

figuur 1 Rehuel Lobatto (1797-1866)



figuur 2 Wet van Sterfte voor de Zuidelijke Provinciën der Nederlanden (Jaarboekje van Lobatto over 1828)
Van de 100.000 kinderen die worden geboren zijn er na 1 jaar nog 77.507 over, enz. Deze wet kan worden gebruikt om sterftekansen te berekenen, of de gemiddelde leeftijd, of de mediane leeftijd, of de gemiddelde levensverwachting van een 13-jarige enz.

Twee werelden

In de negentiende eeuw was er in Nederland een kloof tussen twee vormen van statistiek die niet te overbruggen was. De verschillende soorten statistiek vertegenwoordigden twee essentieel verschillende intellectuele tradities: die van de exacte wetenschappen en die van de letteren. In de jaren vijftig van de twintigste eeuw trok de Engelse natuurkundige en literator C.P. Snow de aandacht met zijn geschrift *The Two Worlds* waarin hij de vinger legde op de zere plek van de miscommunicatie tussen de werelden van de exacte wetenschappen en van de letteren. Hij behoorde tot beide werelden en had ervaren dat binnen de groep van de exacte wetenschappers anders wordt gecommuniceerd dan binnen de groep van de literatoren. Bovendien konden representanten van deze twee groepen elkaar in het algemeen niet begrijpen. In dit artikel zal ik een vergelijkbare onoverbrugbare kloof tussen twee manieren van het bedrijven van statistiek laten zien. Maar eerst worden ze geïntroduceerd: de statistiek van de wiskundige Rehuel Lobatto (1797-1866) en die van de juristen-statistici. De laatste vorm van statistiek zou tot in de twintigste eeuw de overheidsstatistiek domineren. Een meer wiskundige benadering van de statistiek werd pas na de Tweede Wereldoorlog gemeengoed.

Politieke rekenkunde

Lobatto (***figuur 1***) bouwde voort op een Engels-Nederlandse traditie die *Politieke Rekenkunde* wordt genoemd. De politieke rekenkunde ontstond in de zeventiende eeuw in het kader van het idee dat maatschappelijke onderwerpen konden worden onderworpen aan een kwantitatieve analyse. Men begon met het verzamelen van populatiegegevens. Hiertussen bleken allerlei wetmatigheden te bestaan, zoals een constante verhouding tussen sterftecijfers en de totale omvang van de bevolking.

Er ontstond een niet-universitaire traditie waarin getracht werd om op basis van bevolkingscijfers conclusies te trekken over de welvaart en de macht van een land. Omdat sterftecijfers vaak wel bekend waren, kon men met behulp daarvan de omvang van de bevolking schatten. Op basis hiervan had men een idee hoe groot het leger was dat zo nodig op de been kon worden gebracht. Een belangrijke politieke rekenkundige was de Engelse William Petty (1623-1687). In 1690 verscheen zijn boek *Political Arithmetick.*
De grondlegger van de politieke rekenkunde in Nederland was de politicus Johan de Witt (1625-1672) met zijn publicatie *Waerdije van Lijf-renten naer proportie van Los-renten* uit 1671, waarin hij op basis van een door hem aangenomen sterftetafel (een voorbeeld van een sterftetafel is opgenomen ***als figuur 2***) de premie bepaalde voor een zogenaamde lijfrente, waarbij de overheid een bedrag van iemand leende en een rente aan deze persoon betaalde totdat het lijf overleed waarop deze lijfrente was afgesloten.
Het is hier niet de plaats om op Lobatto's leven in te gaan. Voor ons is interessant dat hij niet alleen een wiskundige was (in 1842 werd hij hoogleraar in de wiskunde) maar dat hij zich een wiskundige voelde in hart en nieren. Dit komt tot uitdrukking in de wijze hoe hij in 1828 de geboorte van een tweeling aan een wiskundige vriend aankondigde: 'Ik ben blij u te mogen vertellen, dat mijn vrouw is bevallen van een jongen en een meisje. Ik had beslist niet kunnen voorzien, dat ik een vergelijking van de tweede graad zou oplossen, hetgeen me bijzonder geraakt heeft. Ik hoop een kleine wiskundige te hebben voortgebracht die me eens met mijn "besognes" zal kunnen helpen.' Terzijde: welke van de twee baby's zal hij op het oog hebben gehad met die laatste opmerking?

Nu terug naar de statistiek. Eén van de eerste voor ons interessante activiteiten van Lobatto was zijn jaarboekje, dat hij vanaf 1826 'op Last van Z.M. den Koning', dus namens de overheid mocht uitgeven. Lobatto formuleerde het tweeledige doel van het *Jaarboekje* als 'de verspreiding van algemeen nuttige kundigheden onder alle klassen van beschaafde ingezetenen, als mede de bekendmaking van belangrijke opgaven, meer bijzonderlijk dit land betreffende (…), zoo als, onder anderen, die nopens de bevolking, huwelijken en sterfgevallen (…).'
Elk jaar bevatte het *Jaarboekje* een sectie *Statistiek*. De sectie zag er globaal als volgt uit: tabellen met bevolkingsaantallen per stad en per provincie, geboorte- en sterfteaantallen per geslacht, per maand en per provincie, cijfers van echtelijke en buitenechtelijke geboorten, en aantallen huwelijken. Met behulp van deze populatiestatistieken zocht Lobatto naar wetmatigheden tussen deze getallen. Hij vond onder andere een constante verhouding tussen geboorte- en bevolkingsaantallen (1:27 in 1826), en tussen de geboorteaantallen van jongens en meisjes (1:0,95 in 1826). Hij vond dat hoge geboorte- en sterftecijfers samengaan zowel in de tijd (hoogste geboorte- en sterftecijfers in het voorjaar, de laagste in de zomer), als per regio (het hoogst in de provincies die het dichtst bevolkt waren en aan het water lagen). Lobatto nam in de sectie statistiek van zijn *Jaarboekje* ook elk jaar een sterftetafel op (***zie figuur 2***). Incidenteel werden ook andere gegevens opgenomen zoals over het lager onderwijs en over de bevolking van strafgevangenissen.
Verder werd af en toe een door hemzelf geschreven artikel in het statistische deel van het *Jaarboekje* afgedrukt: in 1833 over weduwenfondsen en in 1836 over maatschappijen voor levensverzekering. Hij was een deskundige op dit gebied, want hij adviseerde de overheid over levensverzekeringskwesties en had er in 1830 een boek over gepubliceerd.
In het jaarboekje van 1829 schreef hij een artikel over een mogelijke toepassing van kanstheorie op statistische gegevens en later, in 1860, schreef hij in het tijdschrift van het Nederlandse wiskundige genootschap (het huidige Koninklijke Wiskundige Genootschap) over hetzelfde onderwerp. Lobatto verwachtte dat de waarschijnlijkheidsrekening de statistiek kon bevruchten, met name als het ging om de betrouwbaarheid van statistische gegevens. Zijn artikelen gingen over de nauwkeurigheid van een gemiddelde en mondden uit in het toepassen van de daarin behandelde theorie op de populatieverhoudingen. Hij bepaalde deze uit een aantal jaarlijkse tellingen. Hij nam het gemiddelde van de verhouding tussen, bijvoorbeeld, geboorte- en bevolkingsaantallen over een tiental jaren en vroeg zich af hoe nauwkeurig dit gemiddelde de 'werkelijke' verhouding benaderde. De werkelijke verhouding zou dan de vaste waarde van die verhouding binnen de gehele populatie gedurende dat tiental jaren moeten zijn. Het artikel van 1860 bevatte een wiskundige afleiding van de kansverdeling van een gemiddelde uit de kansverdelingen van de afzonderlijke waarnemingen, speciaal in het geval van de normale verdeling (***zie figuur 3***). Vanuit hedendaags gezichtspunt, waarin een precies onderscheid wordt gemaakt tussen steekproeven en populaties, valt op de benadering van Lobatto wel het een en ander af te dingen, maar voor toen deed hij interessant werk, waar later op kon worden voortgebouwd. De keuze van de inhoud van de sectie Statistiek in Lobatto's *Jaarboekje* sloot meer aan bij zijn wiskundige aanleg en belangstelling dan bij de statistiek zoals die aan de juridische faculteiten van de universiteiten werd onderwezen. Opmerkenswaard is dat hij deze sectie 'Statistiek' en niet 'Politieke Rekenkunde' noemde. Het concept 'statistiek' was nog niet uitgekristalliseerd. Het kon blijkbaar worden opgevat als iets dat kwantitatief was, dat bevolkingsgegevens betrof en waaraan kanstheorie een relevante bijdrage kon leveren. Maar het betekende aanvankelijk iets heel anders, zoals aan de statistiekcolleges van Kluit te zien is.

Statistiek voor juristen

Vanaf 1802 werden er in Nederland colleges in de statistiek gegeven, eerst alleen in Leiden door Adriaan Kluit (1735-1806) (***figuur 4***), maar later ook aan andere universiteiten. In Leiden werd in 1815 bepaald, dat colleges in de statistiek verplicht waren voor alle studenten in de juridische wetenschappen.
Vanuit het perspectief van de hedendaagse statistiek is het ronduit verbazend dat de rechtenfaculteit als de aangewezen faculteit werd beschouwd voor dergelijke colleges. Wat was daarvan de reden? Evenals nu kwamen juristen toen vaak in overheidsfuncties terecht en moesten zij overheidsbeleid ontwerpen en uitvoeren. In Duitsland was het idee ontstaan dat aan de basis van overheidsbeleid feitelijke kennis van een staat ten grondslag moet liggen. Men was daar gaan nadenken over hoe deze kennis eruit zou moeten zien en hoe deze moest worden gesystematiseerd, zodat ook verschillende staatkundige eenheden met elkaar zouden kunnen worden vergeleken.

figuur 3 Bladzijde met wiskunde uit het artikel van Lobatto uit 1860; 'Over de Waarschijnlijkheid van Gemiddelde Uitkomsten uit een Groot Aantal Waarnemingen'

figuur 4 Adriaan Kluit (1735-1807)

figuur 5 Titelblad 'Leerboek der Statistiek of Staats-kunde' uit 1806. Uit deze titel blijkt dat onder statistiek kennis over een staat werd verstaan. Dit boek is waarschijnlijk in de eerste helft van de negentiende eeuw bij de Leidse colleges statistiek gebruikt.

figuur 6 Titelblad 'Handleiding tot het Statistisch Onderzoek' door S. Vissering, hoogleraar statistiek te Leiden van 1850-1879

Daaraan was in het achttiende-eeuwse Duitsland, bestaande uit vele staten en staatjes, grote behoefte. Statistiek was dus aanvankelijk een systematische beschrijving van een staat, 'staat-kunde', en die betekenis zit ook nog in het woord *statistiek* (***zie figuur 5***)

Wat behandelde de eerste hoogleraar in de statistiek in Nederland, Adriaan Kluit, in zijn colleges? Wij zijn daarvan heel nauwkeurig op de hoogte, omdat collegedictaten van enkele van zijn studenten in de Universiteitsbibliotheek van Leiden worden bewaard. Een collegedictaat van toen gaf vrijwel letterlijk weer wat de hoogleraar op zijn college onderwees. De titels van de collegedictaten zijn onder meer *Statistiek van Nederland* en *Voorlezingen over de Staathuishoudkunde der Nederlanden.*
Statistiek en staathuishoudkunde waren beide vakgebieden in ontwikkeling en men maakte er aanvankelijk geen onderscheid tussen.
De statistiek bevatte volgens Kluit de kennis die nodig zou zijn om de 'waare kragten' ofwel de macht en de welvaart van een land te leren kennen. Kluit onderscheidde twee aspecten aan een staat, namelijk het *natuurlijke* en het *zedelijke*. Onder de natuurlijke gesteldheid van een land verstond hij haar ligging, grenzen en naburen, grootte en klimaat. Bij het zedelijke behandelde hij allereerst de bewoners van het land, onder andere in staatkundig opzicht, maar ook hun aantal en hun karakter, evenals hun middelen van bestaan en hun welstand. Deze kwamen aan de orde aan de hand van een bespreking van visserij, akkerbouw, veeteelt, industrie, koophandel en geldzaken. Als tweede onderdeel van het zedelijke beschouwde Kluit de regering.
Kluit maakte bij zijn bespreking soms gebruik van getallen, maar meestal niet. De reden daarvoor zal zijn geweest dat er over een bepaald onderwerp geen getalsmatige gegevens beschikbaar waren, of dat het zich niet leende voor een weergave in getallen. Voor Kluit speelde het getal geen belangrijke rol. In de statistiek ging het hem er in de eerste plaats om een indruk te krijgen hoe de feitelijke toestand van een staat was; op welke manier dat gebeurde kwam op de tweede plaats.

De behandeling van het *bevolkingsaantal* is het meest kwantitatieve gedeelte van Kluits statistiek. Hij maakte daarvoor gebruik van de kennis uit de traditie van de politieke rekenkunde. Kluit onderscheidde meerdere methoden om dat te doen, bijvoorbeeld het tellen van de strijdbare mannen en dit aantal met 4 of 5 te vermenigvuldigen.

Verder was er de methode gebaseerd op 'de theorie der probabiliteit van 't leven'. Deze methode berustte erop dat er een constante verhouding bestond tussen het aantal inwoners en het jaarlijks aantal gestorvenen. In de steden was de sterfte hoger; voor de grote steden zou de verhouding tussen het jaarlijkse sterfteaantal en het bevolkingsaantal tussen 1 op 24 en 1 op 28 zijn, en op het platteland tussen 1 op 40 en 1 op 42. Kluit bracht het bevolkingsaantal met andere gegevens in verband. Hij vond het bijzonder belangrijk om te weten of een land voldoende bevolkt was. De aantallen vertoonden tussen landen grote verschillen, voor Zweden 230 per vierkante mijl, Portugal 1325 en Nederland leverde het hoogste aantal: 4800. Voor deze verschillen voerde Kluit een aantal verklaringen aan. Zo oefende volgens hem de godsdienst invloed uit op de omvang van de bevolking. In Nederland had de religie positief gewerkt. Hetzelfde gold, dankzij de traditionele vrijheid, voor de regeringsvorm. Het tegengestelde gold voor het hoge niveau van de belastingen.

Na Kluit ontstonden er al spoedig twee vakken: *statistiek* en *staathuishoudkunde*. In de statistiek werd met behulp van zowel kwantitatieve als kwalitatieve informatie over een systematische beschrijving van een staat nagedacht; de staathuishoudkunde gebruikte de resultaten van de statistiek om inzicht te krijgen in economische wetmatigheden (***zie figuur 6***). Een flink aantal juristen kreeg onderwijs in deze vakken en raakte van het belang overtuigd. Hieruit zou in 1857 de *Vereeniging voor Statistiek* voortkomen.

Reactie op elkaar

Lobatto was met zijn statistische bezigheden uniek zowel in statistisch als in wiskundig Nederland. De wiskundigen waren voorals nog niet in deze toegepaste wiskunde geïnteresseerd. De Nederlandse statistici die na 1857 de *Vereeniging voor de Statistiek* vormden waren op een geheel andere wijze met statistiek bezig. De *doelstellingen* van beide soorten statistici vertoonden grote overeenkomsten: een groeiende behoefte aan een meer abstract en formeel overheidsbestuur, gebaseerd op duidelijke en correcte informatie. De *middelen* om deze doelen te bereiken waren vooralsnog verschillend. Lobatto baseerde zich bijna alleen op (kwantitatieve) populatiegegevens, terwijl de statistici vonden dat alle aspecten van een staat in een statistische beschrijving meegenomen moesten worden, kwantitatief zowel als kwalitatief. In dit spanningsveld groeide een nieuw vakgebied, dat van de

statistiek, waarbij er in de eerste helft van de negentiende eeuw nog geen overeenstemming was wat dit precies inhield.
De manier waarop deze verschillende benaderingen van statistiek op elkaar reageerden kan laten zien hoe verschillend beider benadering van de statistiek was en hoe groot de kloof tussen beide manieren van denken.
Een manier om de reactie op elkaar te weten te komen, is na te gaan hoe de sectie statistiek in het *Jaarboekje* van Lobatto werd ontvangen door de juristenstatistici en wat de reactie van Lobatto daar dan weer op was.

Opvallend is dan eerst dat verschillende tijdschriften die meer of minder gewijd waren aan de statistiek op de wijze van de juristenstatistici, Lobatto's jaarboekje niet bespraken. Ik heb alleen een bespreking gevonden in *De Vriend des Vaderlands.* In 1838, dertien jaar na de eerste jaargang van Lobatto's jaarboekje, heeft de jurist J. Ackersdijck, hoogleraar in de statistiek te Utrecht, het statistische gedeelte van dit jaarboekje uitgebreid besproken. Wat zou zijn commentaar op Lobatto's statistiek zijn?
Eerst maakte hij een opmerking over het belang van de statistiek in het algemeen: 'Het geldt niet ijdele bespiegelingen of berekeningen; maar de kennis van werkelijke waarheden van hoog aanbelang, welke menigvuldig ten grondslag dienen moeten voor maatregelen, die op het welzijn der burgers onvermijdelijken invloed hebben; zoodat, wanneer men uit gebrek van kennis dier daadzaken mistast de gevolgen altijd noodlottig zijn.' Voor Ackersdijck ging het in de statistiek om feiten die van belang waren voor het staatsbestuur.
De inhoud van de sectie statistiek van het *Jaarboekje* beschreef hij als volgt: 'De meeste der mededeelingen betreffen de bevolking.' Hij besprak elk onderwerp en elke tabel minutieus. Hij onderstreepte ook het belang van een sterftetafel voor levensverzekeringen. Populatiestatistieken waren voor Ackersdijck zeker de moeite waard, maar hij schreef ook: 'Hiermee afstappende van het onderwerp der bevolking, kunnen wij den wensch niet onderdrukken, dat omtrent vele andere onderwerpen voor de statistiek van ons land mededeelingen in het jaarboekje mogten gevonden worden.' Hij dacht daarbij aan statistieken over onderwijsinstellingen, over in- en uitvoer, en aan mededelingen omtrent de rechtspleging.
Een andere juristenstatisticus schreef in 1850 dat hij het een groot gemis vond dat Lobatto's jaarboekje zich beperkte tot bevolkingsgegevens en niets vermeldde over zaken als landbouw, handel, nijverheid, financiën en koloniën. Hij sprak de hoop uit dat de overheid een uitvoeriger statistisch jaarboekje zou gaan uitgeven, hetgeen vanaf 1851 ook is gebeurd. Hierop zou Lobatto hebben gereageerd met de volgende zinsnede: 'De criticus heeft gelijk, wat de zaak betreft, maar ik mag geenerlei statistiek meedelen, dan waarop men staat maken kan.' Deze kritiek zou ook hebben bijgedragen aan het beëindigen van het Jaarboekje, aangezien 'nieuwe inzigten over de inrigting van dit Jaarboekje (…) hem de taak uit handen namen.' Lobatto zelf schreef een aantal jaar later dat zijn jaarboekje sinds 1850 niet meer verschenen was hoofdzakelijk omdat er een statistisch jaarboekje door het statistisch bureau van het ministerie van binnenlandse zaken werd uitgegeven. Uit archiefonderzoek blijkt dat dit bureau, dat in 1848 was opgericht, al in 1849 plannen had opgesteld om een statistisch jaarboekje uit te geven. De minister onthief toen Lobatto van zijn taak om het jaarboekje samen te stellen.

Conclusie

Uit het voorgaande volgt dat de juristenstatistici bepaalden welke onderwerpen tot de statistiek behoorden. Lobatto bestreed dat niet; hij kwam alleen op voor de zorgvuldigheid waarmee kwantificeren plaats behoorde te vinden en vond dat op dat moment voldoende reden om zich bijna geheel tot populatiegetallen te beperken. Met kwalitatieve statistiek hield hij zich al helemaal niet bezig. Indirect beschuldigde hij daarmee de juristenstatistici van onvoldoende zorgvuldigheid.
De manier waarop Lobatto met statistiek bezig was en de wijze van de juristenstatistici was te verschillend om in een nieuwe geïntegreerde vorm van statistiek te resulteren. Op zijn minst zouden ze zich vertrouwd moeten maken met elkaars denken. Dat deden ze niet. Dat ze in twee verschillende werelden leefden en dat de kloof hiertussen niet te overbruggen viel, is gebleken uit hun reactie op elkaar. Hun twee werelden doen denken aan de *Two Worlds* van C.P. Snow.
Zoals besproken, moest in 1849 het jaarboekje van Lobatto het afleggen tegen een jaarboekje uitgegeven door een statistisch overheidsbureau. Lobatto's rol was uitgespeeld en de juristenstatistici, later georganiseerd in de *Vereeniging voor de Statistiek,* maakten de dienst uit. Dit zou tot gevolg hebben dat een meer wiskundige benadering voorlopig geen kans zou krijgen in de overheidsstatistiek. Dat gebeurde pas veel later, in de twintigste eeuw. Het ontstaan van een *Vereniging voor de Statistiek,* waarin de kansrekening centraal zou staan, moest wachten tot 1945.

Literatuur

Voor dit artikel is relevant de bundel:
- Paul M.M. Klep, Ida H. Stamhuis (2002): *The Statistical Mind in a Pre-Statistical Era. The Netherlands 1750-1850*. Amsterdam: Aksant. Met name hoofdstuk 3: Ida H. Stamhuis: *An Unbridgeable Gap between Two Approaches to Statistics*; pp. 71-98.

Er is gebruik gemaakt van drie artikelen van mijn hand verschenen in *STAtOR*, periodiek van de VVS, en wel:
- *Vereeniging voor de Statistiek (VVS), een Gezelschap van Juristen*. In: *STAtOR* 3 (2002/2); pp. 13-17.
- *Rehuel Lobatto (1797-1866), Eenling als Statisticus, Eenling als Wiskundige*. In *STAtOR* 4 (2004/ 1); pp. 15-20.
- *Statistiek in de Negentiende Eeuw. Een Onoverbrugbare Kloof*. In: *STAtOR* 4 (2004/ 4); pp. 14-19.

Over de auteur

Ida H. Stamhuis doceert wetenschapsgeschiedenis aan de Vrije Universiteit Amsterdam.
E-mailadres: *ih.stamhuis@few.vu.nl*

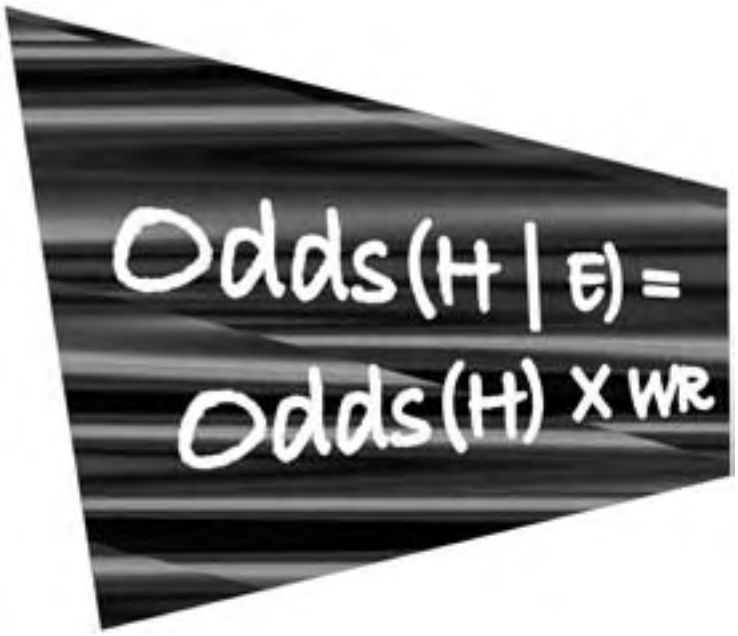
Bayesiaanse kansrekening

EEN VERFRISSEND EN LEERZAAM ALTERNATIEF

[Henk Tijms]

Inleiding

Bayesiaanse kansrekening verenigt eenvoud met praktische toepasbaarheid en is een tak van de kansrekening die zich richt op beslissingssituaties met onzekerheid die niet voor herhaling vatbaar zijn. Voorbeelden van dit soort problemen zijn medische beslissingen bij de behandeling van een patiënt of beslissingen in de rechtszaal over schuld of onschuld van een verdachte. Grote bedrijven - met name farmaceutische - gebruiken steeds meer Bayesiaanse methoden in plaats van traditionele statistische methoden bij het op de markt brengen van nieuwe producten of bij het nemen van investeringsbeslissingen op basis van onderzoeksdata.
In het voortgezet onderwijs zou men er verstandig aan doen om - naast de klassieke kansrekening over herhaalbare kansexperimenten - de nodige aandacht te besteden aan Bayesiaanse kansrekening, en het accent minder te leggen op de statistiek. Het is aardig om aan te halen wat Freudenthal al over kansrekening en statistiek schreef in de jaren zeventig op de bladzijden 610 en 613 van zijn boek *Mathematics as an Educational Task*: '… pervade all mathematics by probability at an early stage- as soon as the children get to know about fractions …' en '… I simply do not understand the philosophy of those who propose to teach high school children a lot of statistical techniques …'. Woorden die nog niets aan zeggingskracht verloren hebben. Zeker nu de praktische bruikbaarheid van significantietests uit de klassieke statistiek steeds meer ter discussie staan en deze tests meer en meer vervangen worden door Bayesiaanse methoden (zie [7] en [11]).
Leerlingen met wiskunde A of C, voor wie wiskunde meestal niet de eerste keus is, zullen met Bayesiaanse kansrekening het nut en de kracht van de wiskunde meer gaan waarderen. Een basiskennis van Bayesiaans denken zal in het latere leven van veel leerlingen, met name van hen die later in de medische wereld of de wereld van de rechtspraak hun beroep uitoefenen, van pas komen. Medische en juridische



blunders zoals in de recente zaak van Sally Clark, die in Engeland ten onrechte tot levenslang veroordeeld werd vanwege de wiegendood van twee haar opeenvolgend geboren baby's, zouden hiermee voorkomen kunnen zijn. In Nederland zou de zaak Lucia de B., de vermeende seriemoordenares, ook anders gelopen zijn als bij de rechters en advocaten enig inzicht in de Bayesiaanse kansrekening aanwezig was geweest. [2]

De formule van Bayes

Het begrip conditionele kans is essentieel in de (Bayesiaanse) kansrekening. Het is een intuïtief begrip dat voor de meeste leerlingen direct duidelijk is zonder een theoretische beschouwing. Bijvoorbeeld, bijna iedereen redeneert als volgt wanneer gevraagd wordt naar de kans op twee azen bij het random trekken van twee kaarten uit een grondig geschud pak van 52 kaarten. De kans op een aas met de eerste kaart is $\frac{4}{52}$. Gegeven dat één aas uit het pak kaarten is, dan is de kans op een aas met de tweede kaart gelijk aan $\frac{3}{51}$. De gevraagde kans is dus $\frac{4}{52}\times\frac{3}{51}$. Dit is een toepassing van de basisregel:
$P(A \text{ en } B) = P(A) \cdot P(B \mid A)$
In woorden, de kans dat zowel gebeurtenis A als gebeurtenis B optreedt (A en B), wordt uitgerekend door de kans dat gebeurtenis A optreedt te vermenigvuldigen met de conditionele kans dat gebeurtenis B optreedt gegeven de informatie dat gebeurtenis A is opgetreden ($B \mid A$). Vele aardige sommetjes zijn voorhanden om deze basisregel of de equivalente formule:
$P(B|A) = \dfrac{P(A \text{ en } B)}{P(A)}$ te oefenen.

Een leuke illustratie is de volgende. Iemand heeft buiten jouw gezichtsveld twee dobbelstenen geworpen en vertelt je dat één van de dobbelstenen een zes gaf. Wat is de kans dat de andere dobbelsteen ook een zes geeft? Het antwoord wordt berekend als $\frac{\frac{1}{36}}{\frac{11}{36}} = \frac{1}{11}$. Het antwoord is dus niet $\frac{1}{6}$, zoals menigeen verwacht zou hebben. Zou echter één van de dobbelstenen op de grond gevallen zijn en je zou gezien hebben dat die steen een zes gaf, dan was het antwoord wél $\frac{1}{6}$ geweest. Dit voorbeeld illustreert fraai een basisles uit de kansrekening: kansen veranderen

als extra informatie ter beschikking komt. Uitgaande van de regel voor voorwaardelijke kans is het een kleine stap naar de *Regel van Bayes*. Deze regel kent verschillende formuleringen. De meest aansprekende is de formulering in termen van 'odds', en niet:
$P(H|E) = P(E|H)\frac{P(H)}{P(E)}$ met
$P(E) = P(E|H)\cdot P(H) + P(E|\neg H)\cdot P(\neg H)$,
de wet van totale waarschijnlijkheid [1].

Klassieke denkfout

Beschouw de volgende situatie:
0,8% van vrouwen van begin 40 die aan een routineonderzoek meedoen, hebben borstkanker. 90% van de vrouwen met borstkanker testen positief, terwijl van de vrouwen zonder borstkanker 7,5% een positieve testuitslag krijgt. Mevrouw Peterson in de genoemde leeftijdsgroep krijgt bij een routineonderzoek een positieve testuitslag.
Wat is de kans dat ze werkelijk borstkanker heeft?
Deze vraag werd in een Amerikaans onderzoek aan een groot aantal medici voorgelegd. Veel van de medici gaven een verkeerd antwoord en schatten de kans in de orde van 75%. Het correcte antwoord is 8,8%! Blijkbaar dachten veel medici dat de kans op borstkanker gegeven een positieve uitslag in de buurt moet liggen van de kans op een positieve uitslag gegeven borstkanker. Dit fenomeen staat in de literatuur bekend als de 'prosecutor's fallacy' naar de aanklager die de kans op onschuld van een verdachte met dezelfde zeldzame lichaamskenmerken als de dader verwart met de (uiterst kleine) kans dat een willekeurig gekozen persoon deze kenmerken heeft. Een interessante bevinding in het Amerikaanse onderzoek was ook dat het percentage foute antwoorden veel lager lag als het probleem in de volgende aangepaste vorm werd gepresenteerd:
Gemiddeld 80 van 10.000 vrouwen van begin 40 die aan een routineonderzoek meedoen hebben borstkanker. Van elke 80 vrouwen met borstkanker krijgen 72 vrouwen een positieve testuitslag, terwijl van de 9920 vrouwen zonder borstkanker 744 vrouwen een positieve testuitslag krijgen.
Als 10.000 vrouwen van begin 40 een routineonderzoek ondergaan, welk percentage van de vrouwen met een positieve testuitslag heeft dan werkelijk borstkanker?
De berekening $\frac{72}{72+744}\times 100\% = 8{,}8\%$ werd wel correct gemaakt door de meeste doktoren. Gigerenzer bepleit in zijn boek *Calculated Risks* de aanpak om probleemstellingen zoals bovenstaande in termen van 'natuurlijke frequenties' te presenteren. Deze 'houtje-touwtje'-aanpak werkt ook heel goed in het beroemde *driedeurenprobleem* (ook wel het *Monty Hall probleem*) [3]: denk in dat je 3000 keer aan de quiz meedoet, dan sta je gemiddeld 2000 keer voor de verkeerde deur en veranderen van deur geeft dus in gemiddeld 2000 van de 3000 gevallen de hoofdprijs. Het blijft echter een wat adhoc aanpak die niet altijd bruikbaar is. Wat dat betreft is werken met kansbomen een betere en systematischer aanpak. Op deze aanpak gaan we niet in, maar wel op Bayes in odds-vorm.

Bayes in odds-vorm

Het begrip 'odds' wordt veelvuldig gebruikt in de wereld van het gokken. Het is een andere manier om kansen weer te geven. De odds van een hypothese is de kans dat de hypothese waar is, gedeeld door de kans dat de hypothese niet waar is. Een hypothese die waar is met kans p, heeft dus odds $\frac{p}{1-p}$. Omgekeerd betekent odds a een kans $\frac{a}{1+a}$. Voor een hypothese H en evidentie E voor de hypothese, luidt de regel van Bayes in odds-vorm:

$$\frac{P(H|E)}{P(\neg H|E)} = \frac{P(H)}{P(\neg H)}\cdot\frac{P(E|H)}{P(E|\neg H)}$$

of in woorden: *posterior odds* = *prior odds* × *waarschijnlijkheidsratio*.

In deze formule representeert de *prior* kans $P(H)$ de waarde van de kans dat de hypothese H ('mevrouw Peterson heeft borstkanker') waar is *voordat* de evidentie E ('de testuitslag van mevrouw Peterson is positief') bekend is, en representeert de *posterior* kans $P(H \mid E)$ de waarde van de kans dat de hypothese H waar is *nadat* evidentie E bekend is. Analoog $P(\neg H)$ en $P(\neg H \mid E)$.
De kans $P(E \mid H)$ geeft de kans op de evidentie E als de hypothese H waar is. Analoog $P(E \mid \neg H)$.
Het quotiënt $P(H) / P(\neg H)$ heet de *prior odds* van de hypothese H en het quotiënt $P(H \mid E) / P(\neg H \mid E)$ is de *posterior odds* van de hypothese.
Het quotiënt $P(E \mid H) / P(E \mid \neg H)$ heet de *waarschijnlijkheidsratio* vanuit de evidentie E. Deze ratio is een maat voor hoe sterk de evidentie is. Als de waarschijnlijkheidsratio hoger dan 1 is, dan spreekt de evidentie in het voordeel van de hypothese en zijn de posterior odds hoger dan de prior odds. Belangrijk is dat men de precieze waarden van de teller en noemer in de waarschijnlijkheidsratio niet hoeft te weten, de verhouding volstaat.
In een rechtszaak waar de evidentie een DNA-spoor is, is het een forensische expert die een waarde aan de waarschijnlijkheidsratio toekent, maar is het de rechtbank die de prior odds een waarde geeft.

De afleiding van de Bayes-formule in odds-vorm is simpel. De formule volgt direct door de relaties:
$P(H|E) = P(E|H)\cdot\frac{P(H)}{P(E)}$ en $P(\neg H|E) = P(E|\neg H)\cdot\frac{P(\neg H)}{P(E)}$
op elkaar te delen.
De intuïtieve formule $P(A \text{ en } B) = P(A) \cdot P(B \mid A)$ volstaat dus om de belangrijke formule van Bayes in odds-vorm af te leiden; meer voorkennis is niet nodig. Dit gevoegd bij de vele interessante en leerzame toepassingen van de Bayes-formule maakt dat de Bayesiaanse kansrekening een bij uitstek geschikt onderwerp is voor het vwo-onderwijs. De Bayes-formule heeft een ereplaats in het boek *25 Big Ideas* (Oneworld Publishers, 2005) van de wetenschapsschrijver Robert Matthews.

Laten we de formule nu toepassen om voor het eerder gegeven voorbeeld de aangepaste waarde van de kans te berekenen dat mevrouw Peterson kanker heeft gegeven dat de routinetest positief uitvalt. De gegevens zijn:
$P(H) = 0{,}008$, $P(\neg H) = 0{,}992$, $P(E|H) = 0{,}9$, $P(E|\neg H) = 0{,}075$
Invullen geeft de posterior odds:
$$\frac{P(H|E)}{P(\neg H|E)} = \frac{0{,}008}{0{,}992}\cdot\frac{0{,}9}{0{,}075} = 0{,}09677$$
Oftewel, de kans dat mevrouw Peterson werkelijk de ziekte heeft bij een positieve testuitslag, is $\frac{0{,}09677}{1+0{,}09677} = 0{,}088$.

Inzicht met Bayes

In de media horen we regelmatig over achteraf onterechte veroordelingen die in feite alleen gebaseerd waren op een bekentenis van de verdachte. De regel van Bayes in odds-vorm geeft inzicht in de onderliggende fout.
Laat $H = S$ de gebeurtenis zijn dat de verdachte schuldig is en laat $E = B$ de gebeurtenis zijn van bekentenis van schuld. Wil de bekentenis bijdragen tot de veroordeling, dan moet gelden dat $P(S \mid B)$ groter is dan de prior kans $P(S)$. Uit de regel van Bayes volgt met simpele algebra dat $P(S \mid B) > P(S)$ alleen dan geldt als $P(B \mid S)$ groter is dan $P(B \mid \neg S)$ (voor $0 < p, q < 1$ is

$\frac{p}{1-p} > \frac{q}{1-q}$ alleen dan als $p > q$). Oftewel, P($S \mid B$) groter dan $P(S)$ geldt alleen dan als de bekentenis met grotere kans van een schuldige afkomstig is dan van een onschuldige. In werkelijkheid is het zeer discutabel of dit altijd het geval is. De regel van Bayes onderstreept dat een veroordeling nooit op een bekentenis zonder verder bewijs mag berusten.

Oefenopgaven

Tot slot geven wij een aantal instructieve opgaven voor de leerling. Tussen haakjes staat het antwoord in termen van odds (de kans $\frac{a}{a+1}$ correspondeert met odds a). De opgaven vereisen het gebruik van een formule maar bovenal denkwerk over de hypothese H en de evidentie E. Relevante wiskunde binnen een zinnige context!

1. Een langgezocht scheepswrak ligt in een bepaald gebied met een kans geschat op 40%. Een nader onderzoek zou met een kans van 90% tot het wrak leiden als het wrak inderdaad in het gebied zou zijn. Het onderzoek wordt uitgevoerd maar leidt niet tot het wrak.
   Wat is de kans dat het wrak zich toch in het betreffende gebied bevindt? (odds 1/15)
2. In een bepaalde streek regent het gemiddeld één keer in de tien dagen. Voor de dagen dat het regent is in 85% van de gevallen regen voorspeld, terwijl voor dagen dat het niet regent in 25% van de gevallen regen is voorspeld. Voor morgen is regen voorspeld. Wat is de kans dat het morgen werkelijk zal regenen? (odds 17/45)
3. In een stad zijn twee taxibedrijven, de 'Gele Rijders' en de 'Witte Rijders', waarbij 85% van de taxi's van het eerste bedrijf zijn en 15% van het tweede bedrijf. Op een regenachtige avond is een taxi na een aanrijding doorgereden. Een getuige van het ongeluk meent de taxi geïdentificeerd te hebben als een taxi van de 'Witte Rijders'. Als de rechtbank de betrouwbaarheid van de waarneming van de getuige test onder vergelijkbare weersomstandigheden, dan blijkt dat in 80% van de gevallen de getuige de kleur van de taxi juist identificeert en in 20% van de gevallen verkeerd.
   Wat is de kans dat de doorgereden taxi inderdaad van de 'Witte Rijders' is? (odds 12/17)
4. In een bepaalde regio worden regelmatig alcoholcontroles uitgevoerd. Een aangehouden automobilist wordt eerst aan een blaasproeftest onderworpen. Als deze test leidt tot een positieve reactie, wordt de automobilist meegenomen voor een bloedproef. De blaasproeftest geeft bij dronkenschap in 90% van de gevallen een positieve reactie, terwijl bij géén dronkenschap in 5% van de gevallen een positieve reactie ontstaat. Momenteel wordt een automobilist alleen voor een blaasproeftest aangehouden bij verdacht verkeersgedrag. Het voorstel is nu om willekeurig automobilisten aan te houden voor een blaasproeftest. Van alle weggebruikers in de betreffende regio rijdt gemiddeld 1 op de 25 onder invloed.
   Wat is de kans dat een willekeurig aangehouden automobilist bij wie de blaasproeftest positief uitvalt, onnodig aan een bloedproef onderworpen wordt? (odds 3/4)
5. Van tweelingen is 30% eeneiig (noodzakelijk van hetzelfde geslacht) en 70% twee-eiig. Elvis Presley, the king of rock 'n roll, had een tweelingbroer die bij zijn geboorte overleed.
   Wat is achteraf de kans dat Elvis de helft van een eeneiige tweeling was? (odds 6/7)
6. Een atleet wordt uitgeloot voor de dopingcontrole. De dopingtest geeft een betrouwbare uitslag in 90% van de gevallen. Geschat wordt dat 5% van de atleten doping gebruikt. De testuitslag wijst op dopinggebruik bij de uitgelote atleet.
   Wat is de kans dat de atleet werkelijk doping heeft gebruikt? (odds 9/19)
7. Een moord is gepleegd. De dader is één van de twee personen X en Y. Beide personen zijn op de vlucht en in eerste instantie even waarschijnlijk als dader. Nader onderzoek onthult dat de dader bloedgroep A heeft. Tien procent van de bevolking heeft deze bloedgroep. Persoon X blijkt bloedgroep A te hebben, maar de bloedgroep van Y is onbekend. Wat is de kans dat Y de dader is? (odds 1/10)

Noten (red.)

[1] $\neg H$ moet worden gelezen als 'niet-H'.
[2] Zie ook het artikel van Ronald Meester op pagina 160.
[3] Zie ook het artikel van Jan van de Craats op pagina 171 in dit nummer.

Referenties

[4] A.P. Dawid (2005): *Probability and Proof*. On-line manuscript (Word-document), University College London (*http://tinyurl.com/tz85o*).
[5] H. Freudenthal (1973): *Mathematics as an Educational Task*. Dordrecht: Reidel Pubishers.
[6] G. Gigerenzer (2002): *Calculated Risks*. New York: Simon and Schuster.
[7] S.N. Goodman (1999): *Toward evidence-based medical statistics*. 1: The p value fallacy. 2: The Bayes factor. In: *Annals of Internal Medicine*, Vol. 130, pp. 995-1013.
[8] H. Joyce (2002): *Beyond reasonable doubt*. In: *Plus Magazine*, Issue 21 (*http://plus.maths.org*).
[9] D.V. Lindley (2006): *Understanding Uncertainty*. Chichester: Wiley.
[10] R.A.J. Matthews (1997): *Shock-horror statistics*. In: *Teaching Statistics*, Vol. 18, pp. 9-11.
[11] A. O'Hagan, B.R. Luce (2003): *A primer on Bayesian Statistics*. University of Sheffield: CHEBS (*www.shef.ac.uk/chebs*; zie Downloads).
[12] M. Sjerps (2004): *Forensische statistiek*. In: *Nieuw Archief voor Wiskunde*, Vol. 5, pp. 106-111.
[13] H.C. Tijms (2007): *Understanding Probability: Chance Rules in Everyday Life*. Cambridge: Cambridge University Press (2e editie).
[14] W.A. Wagenaar (1988): *The proper seat / a Bayesian discussion of the position of expert witnesses*. In: *Law and Human Behavior*, Vol. 12, pp. 499-510.

Over de auteur

Prof. dr. Henk Tijms is hoogleraar operationele research aan de afdeling econometrie van de Vrije Universiteit. Hij is auteur van verschillende leerboeken op het gebied van de kansrekening en operationele research.
E-mailadres: *tijms@feweb.vu.nl*

Rev. Thomas Bayes, 1702-1761

De loting voor het EK Voetbal en de 'groep des doods'

[Rob Bosch]

Aanbeveling aan Marco van Basten

Zo af en toe spelen de kansrekening en statistiek een rol in het publieke debat. Het proces tegen en de daarop volgende veroordeling van Lucia de B. is hiervan een prominent voorbeeld. Hierbij ging het om de mogelijk ondeugdelijke statistische argumenten in het bewijs tegen de verdachte. Op een ander, minder dramatisch terrein heeft de kansrekening de gemoederen van voetballiefhebbers een aantal weken bezig-gehouden. Het betreft hier de loting voor het Europees kampioenschap voetbal dat in juni 2008 in Zwitserland en Oostenrijk wordt gehouden. Bij de vooruitblik op de loting kwam men op grond van kanstheoretische argumenten tot de aanbeveling aan de Nederlandse bondscoach Marco van Basten om de laatste kwalificatiewedstrijd (tegen Wit-Rusland) maar te verliezen (Nederland was al zeker van plaatsing voor het toernooi). Een dergelijke uitslag zou de kansen op een gunstige loting vergroten. In kranten, voetbaltijdschriften, televisieprogramma's en op het internet kreeg dit argument uitvoerige aandacht [1]. Zie het onderstaande bericht op *www.EK-WK.nl*:

Marco van Basten

Is winnen wel zo handig?
21-11-2007, 11:13

Er staat vanavond tegen Wit-Rusland niks meer op het spel, of toch wel? Als Oranje vanavond verliest zou dat weleens gunstige gevolgen voor de EK-loting kunnen hebben. Als het Nederlands elftal vanavond wint of gelijkspeelt bij Wit-Rusland, kan dat weleens voor een nieuwe 'groep des doods' gaan zorgen.
Tijdens de loting worden er namelijk 4 verschillende potten gemaakt, waarbij landen uit dezelfde pot niet tegen elkaar kunnen loten. De organisatoren Zwitserland en Oostenrijk én regerend Europees Kampioen Griekenland zijn zeker van een plek in de eerste pot. Oranje komt ook in deze pot, tenzij het verliest van Wit-Rusland en Duitsland wint van Wales. Dan zijn het de Duitsers die in Pot 1 terecht komen.

Weer een 'Poule des doods'?

De echte Europese toplanden zullen echter in Pot 2 en 3 te vinden zijn. We denken dan aan Italië, Duitsland, Frankrijk, Spanje, Tsjechië én Portugal en Engeland (de laatste 2 zijn nog niet zeker van kwalificatie).
Zo zou Oranje (als het in Pot 1 terecht komt) dus zomaar kunnen loten tegen wereldkampioen Italië, angstgegner Portugal en bijvoorbeeld Roemenië. Niet echt een ideale loting. Komt Oranje in Pot 2 dan is er dus 75% kans dat ze tegen Zwitserland, Oostenrijk of Griekenland komen… op het eerste gezicht mindere tegenstanders.

Het argument is duidelijk; als je in pot 1 terecht komt, tref je in de groepsfase niet één van de drie zwak geachte tegenstanders. Wanneer je echter in één van de andere potten terecht komt, dan is de kans groot (75%) dat je tegen één van deze matige tegenstanders speelt.

Groep des doods

Het gaat wat Nederland betreft goed(?) - we verliezen - maar doordat Duitsland gelijk speelt tegen Wales, komen we toch in pot 1 terecht en het schrikbeeld van een lastige groep doemt op. De voor ons land ongunstige geachte indeling wordt na deze uitslagen als volgt [2] :

pot 1 = Zwitserland, Oostenrijk, Nederland, Griekenland
pot 2 = Frankrijk, Polen, Rusland, Turkije
pot 3 = Duitsland, Portugal, Roemenië, Spanje
pot 4 = Italië, Kroatië, Zweden, Tsjechië

De procedure bij de loting is in het eerste internetbericht al aan de orde geweest, maar nog even een korte herhaling. De landen uit pot 1 (geplaatste landen) worden in de bovenstaande volgorde verdeeld over de groepen A, B, C en D. Daarna trekt men uit pot 2 een land dat in groep A terecht komt. Vervolgens trekt men een tweede land uit pot 2 dat in groep B wordt geplaatst. Uit de twee overgebleven landen in de pot trekt men het land voor groep C, waarna het resterende land in groep D terecht komt. Deze procedure herhaalt men met de potten 3 en 4. Bij deze procedure is het dus mogelijk dat in groep C, aangevoerd door het sterkst geachte land uit pot 1, Nederland, de sterkste landen uit de potten 2, 3 en 4 terecht komen. We hebben dan een zware groep of een 'groep des doods'. Zo'n groep lijkt niet erg waarschijnlijk, maar de voetbalanalisten van het internet krijgen gelijk (zie weer *www.EK-WK.nl*):

Zware loting voor Oranje EK 2008
02-12-2007, 10:22
Zojuist heeft in Luzern (Zwitserland) de loting plaats gevonden voor het EK Voetbal 2008 in Oostenrijk en Zwitserland. Nederland is samen met de landen Frankrijk, Italië en Roemenië ingedeeld in poule C. We kunnen voorzichtig spreken van een 'groep des doods'.

Jawel, de loting van 2 december 2007 geeft weer een 'groep des doods'. We zijn ingedeeld bij wereldkampioen Italië, Frankrijk, de nummer 2 van het afgelopen wereldkampioenschap, en het sterke Roemenië (winnaar van onze kwalificatiepoule). *Weer* een 'groep des doods', want in 2006 bij het wereldkampioenschap in Duitsland werd Nederland ook al ingedeeld in een poule die door iedereen als de sterkste groep werd gekwalificeerd [4]. Voor Nederland heeft een loting bij een belangrijk toernooi dus veel weg van een noodlot.

Loting Europees Kampioenschap Voetbal 2008

	Poule A	Poule B	**Poule C**	Poule D
pot 1	Zwitserland	Oostenrijk	**Nederland**	Griekenland
pot 2	Turkije	Polen	**Frankrijk**	Rusland
pot 3	Portugal	Duitsland	**Roemenië**	Spanje
pot 4	Tsjechië	Kroatië	**Italië**	Zweden

Kansen

Als wiskundige wil je graag weten hoe groot de kans is dat door loting zo'n 'groep des doods' gevormd wordt. Wellicht dat je met die kennis enig kwantitatief licht kunt laten schijnen op de analyses van de voetbalkenners. De eerste moeilijkheid die je daarbij tegenkomt, is dat het begrip 'groep des doods' niet precies gedefinieerd is en dat je voor een wiskundige analyse daarmee dus niet uit de voeten kunt. Laten we daarom vooraf proberen te definiëren wat we met dat begrip bedoelen. De meest voor de hand liggende definitie is de volgende:
een 'groep des doods' is een groep die bestaat uit het sterkste land uit pot 1, het sterkste land uit pot 2, enz.; kortom, de op basis van de lotingsprocedure sterkst mogelijke poule, waarbij de sterkte van de landen wordt bepaald op basis van een bepaald criterium, bijvoorbeeld de wereldranglijst van de FIFA [3].
Als we deze definitie hanteren, dan is door loting slechts één 'groep des doods' mogelijk, namelijk de volgende [5]:

groep des doods	
pot 1	Nederland
pot 2	Frankrijk
pot 3	Spanje
pot 4	Italië

In het licht van de gegeven (voor de hand liggende) definitie is er bij de loting op 2 december j.l. dus geen 'groep des doods' tot stand gekomen. Als we deze 'wiskundig' gedefinieerde groep vergelijken met de opmerkingen in het eerste internetbericht, dan zien we dat alle landen uit deze groep in het commentaar op het internet genoemd zijn als zware tegenstanders. Blijkbaar is onze definitie zo gek nog niet. De kans dat deze 'groep des doods' door loting tot stand komt, is eenvoudig uit te rekenen. Het sterkste land uit pot 1, Nederland, wordt niet geloot maar direct als groepshoofd geplaatst - in dit geval als aanvoerder van groep C. De kans dat het sterkste land uit pot 2, Frankrijk, in deze groep wordt ingedeeld is uiteraard ¼. Voor de potten 3 en 4 geldt voor Spanje en Italië dezelfde kans, zodat de kans op de zware samenstelling gelijk is aan:

$\frac{1}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{64}$

We kennen nu de kans, maar wat valt er over deze kans van ongeveer 1,56% te zeggen. Enerzijds is er niet iets buitengewoon onwaarschijnlijks gebeurd. Anderzijds zal er gemiddeld bij 64 toernooien slechts één keer een 'groep des doods' optreden, en bij 64 kampioenschappen spreken we dan toch over 256 jaar. Je kunt je tegen deze achtergrond de opwinding van voetbaljournalisten wel voorstellen als er zo'n groep wordt samengesteld.

Nog een groep des doods

Zoals we boven gezien hebben is er met de gegeven definitie maar één 'groep des doods' mogelijk. Maar onze definitie is wellicht wat te beperkt. Enkele landen zijn, ook op basis van de wereldranglijst, min of meer gelijkwaardig. Als we bij de loting waren ingedeeld bij Frankrijk, Duitsland en Tsjechië, dan zou men waarschijnlijk ook voorzichtig van een 'groep des doods' hebben gesproken (alle landen worden in het internetbericht genoemd als sterke tegenstanders). Temeer daar zowel Duitsland als Tsjechië hoger staan op de wereldranglijst dan Roemenië. We kiezen daarom voor een pragmatische definitie van zo'n groep. De sterke landen uit de verschillende potten zijn:

sterke landen in de pot	
pot 1	Nederland
pot 2	Frankrijk
pot 3	Portugal, Duitsland, Spanje
pot 4	Italië, Kroatië, Tsjechië

We spreken *nu* van een 'groep des doods' als er uit de verschillende potten een land getrokken wordt dat voorkomt in de boven-

staande rijtjes. Een mogelijke zware poule is nu bijvoorbeeld:

groep des doods	
pot 1	Nederland
pot 2	Frankrijk
pot 3	Spanje
pot 4	Tsjechië

Deze poule zou door de journalisten waarschijnlijk ook voorzichtig een 'groep des doods' worden genoemd. We berekenen de kans op een 'groep des doods' bij de pragmatische definitie.
Nederland is weer groepshoofd in groep C. De kans dat uit pot 2 Frankrijk in groep C komt is ¼. Voor pot 3 en 4 zijn de kansen op een sterke tegenstander gelijk aan ¾. De kans dat Nederland deel uitmaakt van een 'groep des doods' is nu gelijk aan:
$\frac{1}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{9}{64}$
Met een kans van 14% kunnen we nauwelijks meer spreken van een verrassing. De angst van de voetbalkenners dat bij plaatsing in pot 1 voor Nederland een zware groep dreigt, lijkt dan ook gerechtvaardigd.

Het alternatief

Laten we nog even nagaan hoe de zaken ervoor zouden staan als niet Nederland, maar Duitsland als groepshoofd van groep C geplaatst zou zijn. Nederland zou in dit geval de plaats van Duitsland in pot 3 innemen. We komen dan alleen in een 'groep des doods' als we in groep C terecht komen. De kans hierop is ¼. De kans dat we in die groep Frankrijk ontmoeten, is ¼. De kans dat Italië, Tsjechië of Kroatië aan de groep wordt toegevoegd, is ¾. De kans op een 'groep des doods' is in dit geval dus:
$\frac{1}{4}\times\frac{1}{4}\times\frac{3}{4}=\frac{3}{64}$
De kans dat Nederland bij de loting in een sterke groep terecht komt, is in het eerste geval dus drie keer zo groot als bij de indeling van Nederland in pot 3. Om een sterke groep zoveel mogelijk te vermijden is het dus gunstig om in pot 3 te komen. De aanbeveling van de voetbaljournalisten om in Wit-Rusland te verliezen, is vanuit sportief oogpunt misschien niet fraai, maar vanuit strategische overwegingen lijkt het zo gek nog niet.

Alles in één pot

De lotingsprocedure waarbij de zestien deelnemende landen worden verdeeld over 4 potten, wordt mede gemotiveerd door het vermijden van zeer zwakke en zeer sterke groepen. Dit verlangen wordt echter enigszins doorkruist door de overweging de organiserende landen en de huidige kampioen als eerste in een groep te plaatsen. Stel dat je de groepen samenstelt door een pot met 16 landen te maken en dan steeds vier landen te trekken die samen een groep vormen, dan wordt de loting veel spannender, want alles is nu mogelijk. De kans op een 'groep des doods' lijkt nu wel (aanzienlijk) groter te worden. Immers, de sterke landen uit de potten 2, 3 en 4 kunnen bij deze procedure in dezelfde groep terecht komen. Hoe groot is bij deze procedure de kans dat Nederland in een 'groep des doods' terecht komt?
Alle groepen {Nederland, L_2, L_3, L_4} zijn uiteraard even waarschijnlijk. Behalve Nederland zijn er nog 7 sterke landen. Er zijn dus $\binom{7}{3}$ sterke groepen met Nederland. In het totaal kunnen we met Nederland $\binom{15}{3}$ groepen vormen. De kans op een supergroep is bij deze procedure dus gelijk aan:
$\frac{\binom{7}{3}}{\binom{15}{3}}=\frac{1}{13}\approx 0{,}077$
Deze lotingsprocedure zou voor Nederland dus gunstiger zijn dan de lotingsprocedure waarbij Nederland in pot 1 geplaatst is, maar minder gunstig dan de variant met Nederland in pot 3. De vorming van een supergroep is, zoals te verwachten, mede afhankelijk van de lotingsprocedure. De kans dat er bij de loting met alle landen in één pot een 'groep des doods' (al dan niet met Nederland) gevormd wordt, is een pittige opgave. Die laten we dan ook graag aan de lezer en zijn of haar leerlingen over.

Tot slot

We hebben gezien dat we, voordat we kansen kunnen uitrekenen, de begrippen helder moeten formuleren. Anders gezegd, we moeten een helder stochastisch model maken. In de media komen we vaak kansuitspraken over vaag gedefinieerde gebeurtenissen tegen. Dergelijke uitspraken zijn dan oncontroleerbaar. Kortom, het bepalen van kansen begint met het opstellen van een goed kansmodel.

We hebben in dit stukje twee definities gegeven van een 'groep des doods'. Hierbij horen twee modellen die verschillende kansen geven voor het ontstaan van zo'n groep. Er zijn buiten deze modellen nog zeker andere mogelijk.
Bijvoorbeeld, is een groep met drie sterke landen en één zwak land voor de sterke landen niet veel lastiger dan een groep met vier sterke landen? We zouden een dergelijke groep ook een 'groep des doods' kunnen noemen. Vanuit een andere optiek kunnen we de kansen op een 'groep des doods' uitrekenen voor verschillende lotingsprocedures. Is er een redelijke procedure te bedenken waarbij de kans op een supergroep heel klein is? De bijbehorende kanssommen moeten voor een aantal leerlingen zeker te doen zijn.
Wellicht is het WK in Zuid-Afrika in 2010 een aardig kansrekenproject voor leerlingen.

Noten

[1] Van Basten verklaarde niet gevoelig te zijn voor deze kanstheoretische overwegingen en zeker op winst te willen spelen in de laatste kwalificatiewedstrijd tegen Wit-Rusland.

[2] In het internetartikel gaat men uit van een enigszins andere indeling. Dat artikel is echter geschreven voordat de definitieve rangschikking in de kwalificatiegroepen bekend was. Voor het in het genoemde bericht gegeven argument maakt dit kleine verschil echter niet uit.

[3] FIFA = La Fédération Internationale de Football Association = de wereldvoetbalbond.

[4] In de *Nieuwe Wiskrant* (jrg. 25, nr. 4, juni 2006) schreef Hans van Maanen een bijdrage over eenzelfde thema bij het WK2006, getiteld 'Nederland in de poule des verderfs', waarin hij de kansen op het overleven in zo'n poule bespreekt. Het artikel is digitaal beschikbaar via: *www.vanmaanen.org/hans/artikelen/*

[5] Sterkte op basis van de FIFA-wereldranglijst 17 december 2007 (zie *http://fr.fifa.com*).

Over de auteur

Rob Bosch is redacteur van *Euclides* en universitair hoofddocent wiskunde aan de Nederlandse Defensie Academie te Breda.
E-mailadres: *robbosch@live.nl*

Lucia de B. en de statistiek

[Ronald Meester]

Inleiding

De rechtszaken tegen de verpleegster Lucia de B., die veroordeeld werd voor de vermeende moord op een aantal van haar patiënten, hebben de afgelopen jaren heel wat stof doen opwaaien. Een van de bijzondere aspecten in deze casus was het gebruik van statistiek en kansrekening in de bewijsvoering. In eerste instantie werd Lucia de B. een verdachte doordat zich tijdens haar diensten onevenredig veel incidenten [1] leken te hebben voorgedaan, en de vraag drong zich op of het optreden van zoveel incidenten aan het 'toeval' kon worden toegeschreven. De officier van justitie vroeg een statisticus, Henk Elffers, om op deze vraag een antwoord te geven, en het antwoord van Elffers was ondubbelzinnig: als alle incidenten gewoon toevallig zouden plaatsvinden, zo zei hij, dan zou de kans dat Lucia de B. minstens zoveel incidenten zou meemaken als feitelijk was gebeurd, kleiner zijn dan 1 op de 342 miljoen. Een onwaarschijnlijk klein getal, en dit getal inspireerde Henk Elffers om te concluderen: 'Als statisticus heb ik de vraag: "Kan dat toeval zijn?", beantwoord met een ongeclausuleerd "Nee".' [2]

De berekening, en vooral ook de interpretatie van Elffers, werden vrijwel onmiddellijk door vele wiskundigen, onder wie ikzelf, scherp bekritiseerd. In hoger beroep heb ik daarop als getuige-deskundige betoogd dat het getal dat Elffers leverde, weinig of geen betekenis heeft, en dat men in het algemeen buitengewoon voorzichtig dient te zijn met dit soort berekeningen. De rechters waren tijdens dit verhoor *not amused* over wat ik te vertellen had. Het werd mij snel duidelijk dat zij de getuigenis van Elffers wensten te aanvaarden, omdat hij een eenduidig getal leverde, namelijk de al eerder genoemde 1 op de 342 miljoen. Mijn opvatting dat een eenduidig getal als dit helemaal niet zinvol te leveren was, en dat er sowieso ernstige problemen waren in Elffers' methode, werd niet gewaardeerd.

Elffers' methode

Laten we nu eerst eens kijken wat Elffers eigenlijk gedaan had. De data die Elffers gebruikte, waren als volgt. De verdachte had in de periode waar het om ging, in twee ziekenhuizen gewerkt: in het Juliana Kinderziekenhuis (JKZ) en in het Rode Kruisziekenhuis (RKZ) alwaar ze op twee verschillende afdelingen diensten had gedraaid. In onderstaande tabel vinden we de gegevens over de diensten in de bestudeerde periode.

ziekenhuis (en nummer afdeling)	**JKZ**	**RKZ-41**	**RKZ-42**
totaal aantal diensten	1029	336	339
aantal diensten Lucia de B.	142	1	58
totaal aantal incidenten	8	5	14
aantal incidenten in diensten Lucia de B.	8	1	5

Op het eerste gezicht lijkt dit inderdaad wel wat te veel van het goede. Zo vonden bijvoorbeeld alle 8 incidenten in het JKZ plaats tijdens een dienst van Lucia de B. Onder de aannames dat:

1. er een vaste kans p bestaat dat gedurende een dienst een incident optreedt, en een kans
2. $1-p$ dat er geen incident optreedt, en
3. incidenten onafhankelijk van elkaar optreden,

berekende Elffers eerst de *conditionele kans* dat alle incidenten in het JKZ gedurende Lucia's diensten plaatsvonden, *gegeven* dat er in totaal 8 incidenten plaatsvinden. Deze kans is eenvoudig uit te rekenen. Stel in het algemeen dat het totaal aantal diensten n is, en dat Lucia de B. er daarvan r voor haar rekening neemt. Laat nu A_k de gebeurtenis zijn dat er precies k incidenten plaatsvinden, en laat L_x de gebeurtenis zijn dat Lucia de B. tijdens haar diensten x incidenten ziet. We berekenen dan:

$$P(L_x \mid A_k) = \frac{P(L_x \cap A_k)}{P(A_k)} = \frac{\binom{r}{x}p^x(1-p)^{r-x}\cdot\binom{n-r}{k-x}p^{k-x}(1-p)^{n-r-k+x}}{\binom{n}{k}p^k(1-p)^{n-k}} = \frac{\binom{r}{x}\binom{n-r}{k-x}}{\binom{n}{k}}$$

In het concrete geval waarnaar we hier kijken is $n = 1029$, $r = 142$, $k = 8$, en tenslotte $x = 8$. De (conditionele) kans is dan gelijk aan $\binom{142}{8}/\binom{1029}{8} \approx 1{,}1\times10^{-7}$.

De kansverdeling die we hebben gevonden, staat bekend als de *hypergeometrische verdeling*. Elffers bracht hiermee het probleem terug tot een berekening die je ook in termen van vazen en ballen kunt formuleren. Merk op dat de uitkomst van de berekening helemaal niet afhangt van de onbekende p. Volgens Elffers is dat maar goed ook, omdat deze onbekende p volgens hem niet te bepalen of te schatten is.

Bovenstaande berekening kun je voor elke afdeling doen waar Lucia de B. werkte. Bij het RKZ wordt dan de kans berekend dat de verdachte *minstens* 1, respectievelijk minstens 5 incidenten zou zien tijdens haar diensten. [3]

Echter, nog steeds volgens Elffers, deze berekening is niet helemaal eerlijk voor de verdachte, omdat de berekening wordt uitgevoerd juist omdát de verdachte zoveel incidenten had meegemaakt. Vergelijk dit maar met de winnaar van een loterij: als je achteraf gaat uitrekenen wat de kans is dat juist deze persoon wint, dan zul je zien dat die kans heel klein is, maar natuurlijk maakt dit feit de winnaar niet tot een verdacht persoon. Het is daarom beter, aldus Elffers, om in het JKZ, het ziekenhuis waar de verdachte opviel, de kans uit te rekenen dat *één van de verpleegkundigen* alle 8 incidenten meemaakt. Ruwweg (maar niet helemaal) komt dat neer op het vermenigvuldigen van de gevonden kans met het aantal verpleegkundigen, namelijk 27. Voor

het RKZ is deze correctie niet nodig, omdat de verdachte toen al geïdentificeerd was in het JKZ.
Tenslotte komt Elffers tot zijn getal van 1 op de 342 miljoen door de uitkomsten van deze drie berekeningen met elkaar te vermenigvuldigen; *met* correctie in het JKZ, en *zonder* correctie in het RKZ.

Kritiek op de gang van zaken

Mijn kritiek kan onderverdeeld worden in een aantal verschillende categorieën. In deze bijdrage maak ik het volgende onderscheid.

1. De manier waarop de gegevens tot stand kwamen.
2. De keuze van het model van Elffers.
3. De berekeningen van Elffers.
4. De interpretatie van de cijfers van Elffers door het hof.

In deze zaak is weinig goed gegaan vanuit statistisch perspectief, en ik zal nu deze vier punten één voor één langslopen, en kort weergeven wat er mis is gegaan.

1. De manier waarop de gegevens tot stand kwamen

In het rapport van de Commissie Evaluatie Afgesloten Strafzaken onder leiding van J.W.M. Grimbergen is duidelijk te lezen wat er op dit punt is misgegaan. Uit het rapport komt overduidelijk naar voren dat Lucia de B. in een veel te vroeg stadium als enige verdachte werd aangemerkt. Dit had tot gevolg dat incidenten waarvan al duidelijk was dat Lucia de B. er niets mee te maken had, als niet-relevant terzijde werden geschoven, en dus in bovengenoemde data helemaal niet voorkomen.
Het is duidelijk dat dit een onacceptabele gang van zaken is. Voordat het statistisch onderzoek überhaupt zou kunnen beginnen, is er al een fout begaan die niet meer gecorrigeerd kan worden: de gegevens waarop de statisticus zich moet baseren, zijn eenzijdig vastgesteld, en zijn dus in feite incorrect. Als alle incidenten waar Lucia de B. op voorhand geen invloed op kan hebben weggelaten worden uit de berekeningen, dan is het duidelijk dat dit sterk ten nadele van haar is. Als bijvoorbeeld op het JKZ niet 8 maar 16 incidenten zouden hebben plaatsgevonden (ik noem maar wat), dan zou Lucia de B. niet alle, maar slechts de helft van alle incidenten hebben meegemaakt. Dit zou van grote invloed zijn op de uitkomst van de berekeningen. Ik wijs hier met grote nadruk op het feit dat deze rechtszaak in die zin bijzonder was, dat helemaal niet vaststond dat er überhaupt een moord was gepleegd. Dat gegeven maakt dat we buitengewoon voorzichtig moeten zijn met de data.

2. De keuze van het model van Elffers

Elffers heeft een wel heel eenvoudig model gekozen, en de vraag is of dit eenvoudige model de realiteit wel voldoende kan beschrijven.
Allereerst valt op dat hij de kans op een incident per dienst gelijk neemt voor alle diensten. Je kunt je afvragen of dat reëel is. Wellicht is er een verschil tussen bijvoorbeeld dag- en nachtdiensten, of is er verschil in de beleving van patiënten bij verschillende verpleegkundigen. Deze nuance is in het model van Elffers niet mee te nemen.
Ten tweede beperkt het model van Elffers zich tot de afdelingen waar de verdachte werkte. Deze keuze is niet goed of fout, maar het is wel een keuze. De vraagstelling was of de verdachte een 'te groot' aantal incidenten mee had gemaakt, en in deze vraag komt het woord 'afdeling' niet voor. Als je andere afdelingen of zelfs andere ziekenhuizen mee zou nemen, dan zou het model misschien te groot en te gecompliceerd worden, maar dat wil niet zeggen dat andere afdelingen of ziekenhuizen daardoor irrelevant zijn. Immers, ook als Lucia de B. iets overkomen is dat een heel kleine kans heeft, dan nog zal een dergelijke gebeurtenis 'ergens in Nederland' ook wel eens voorkomen. Het is een beetje te vergelijken met de winnaar van de loterij waar ik hierboven al over schreef. Net zoals het onredelijk is om uit te rekenen wat de kans is dat de feitelijke winnaar wint, is het ook onredelijk om zonder verder commentaar de berekeningen te beperken tot juist die afdeling waar iets bijzonders is gebeurd.
Maar er is nog iets belangrijkers. Door te conditioneren op het aantal incidenten dat plaatsgevonden heeft, valt de onbekende p weg uit zijn berekeningen, en zoals ik al schreef, is dat volgens Elffers maar goed ook, want over die onbekende p is volgens hem niet veel te zeggen. Welnu, of dat waar is, is op zijn minst onduidelijk, maar zelfs als hij hierin gelijk heeft, speelt die p op de achtergrond een belangrijke rol. Laat ik dit proberen duidelijk te maken met het volgende gedachtenexperiment.
Stel je even voor dat in een lange periode voordat Lucia op de afdeling kwam werken, er gemiddeld twee maal zoveel incidenten plaatsvonden als gedurende de tijd waarin Lucia werkzaam was op de afdeling. Dit zou natuurlijk relevante informatie zijn, want het aantal sterfgevallen gedurende de tijd dat Lucia er werkte, zou dan lager zijn dan normaal, en dan zou het toch wel enigszins merkwaardig zijn dat Lucia de B. als verdachte aangemerkt zou worden (ik roep in herinnering dat we helemaal niet weten of er überhaupt een moord gepleegd is). Ook deze overwegingen kunnen in het model van Elffers niet meegenomen worden, en dus is het heel goed mogelijk dat belangrijke informatie gemist wordt. In het eerder genoemde rapport van de commissie Grimbergen wordt op dit punt dan ook terecht forse kritiek geleverd op de gang van zaken.
In een gezamenlijke publicatie met Van Lambalgen, Gill en Collins [4] geef ik met een berekening aan wat er zou kunnen gebeuren als je een model maakt waarin deze overweging wel een plaats krijgt. Kern van dat model is, dat het aantal incidenten dat een bepaalde verpleegkundige meemaakt, een Poisson-verdeling krijgt met een parameter μ die eventueel van verpleegkundige tot verpleegkundige kan variëren. We krijgen dan dat het aantal incidenten X van een verpleegkundige die r diensten draait, de kansverdeling:

$$P(X=k)=e^{-\mu r}\frac{(\mu r)^k}{k!}\quad (k=0,1,\ldots)$$

heeft. [5] De hypothese dat Lucia de B. geen relatie heeft tot de incidenten, zou dan geformuleerd kunnen worden door te zeggen dat elke verpleegkundige dezelfde μ heeft; deze hypothese geven we aan met H_d. De officier van justitie zou dan de hypothese kunnen formuleren dat de μ_L die bij Lucia de B. hoort, anders is dan de 'normale' μ van de andere verpleegkundigen. Deze hypothese geven we aan met H_p. Als we nu E schrijven voor de gebeurtenis dat verpleegkundige i precies k_i incidenten meemaakt, Lucia de B. nummer j krijgt, en het aantal verpleegkundigen I is, dan geldt dat:

$$P(E\mid H_d)=\prod_{i=1}^{I} e^{-\mu r_i}\frac{(\mu r_i)^{k_i}}{k_i!}$$

en

$$P(E\mid H_p)=\frac{e^{-\mu_L r_j}(\mu_L r_j)^{k_j}}{k_j!}\cdot\prod_{i=1,\,i\neq j}^{I} e^{-\mu r_i}\frac{(\mu r_i)^{k_i}}{k_i!}$$

De zogenoemde 'likelihood ratio' wordt dan gegeven door:

$$LR=\frac{P(E\mid H_p)}{P(E\mid H_d)}=e^{(\mu-\mu_L)r_j}\left(\frac{\mu_L}{\mu}\right)^{k_j}$$

Hoe hoger dit getal, hoe relatief meer waarschijnlijk de gebeurtenissen zijn onder de aanname H_p van de officier van justitie, dan onder de aanname H_d van de verdediging. De formule is niet altijd even makkelijk te interpreteren, maar als nu bijvoorbeeld uit data voorafgaand aan de bestudeerde periode, zou blijken dat $\mu > \mu_L$, dan zien we dat dit getal kleiner (en dus ontlastender voor Lucia de B.) wordt naarmate het aantal incidenten k_j dat zij meemaakt, groter wordt. Zo *tegenintuïtief* kan het kennelijk zijn.

3. De berekeningen van Elffers

Hoeveel kritiek er op de keuze van het model ook te geven is, zodra het model eenmaal vastgesteld is, zouden berekeningen binnen dat model eenduidig moeten kunnen zijn. Maar zelfs dat is niet het geval bij de berekeningen van Elffers. Ik wil hier twee punten in Elffers' berekening nader belichten.

Elffers zegt zich in zijn berekening alleen te willen baseren op dienstrooster-data, dat wil zeggen, de diensten van verdachte en andere verpleegkundigen, en de opgetreden incidenten. Het enige dat telt, is het aantal diensten dat door de verdachte gedraaid wordt, samen met het totaal aantal gedraaide diensten. Het aantal verpleegkundigen op de afdeling zou dus totaal irrelevant moeten zijn. Immers, de vraag die Elffers gesteld werd is of er reden is om te denken dat zich tijdens de diensten van de verdachte iets opvallends voorgedaan heeft. De enige gegevens die daarvoor nodig zijn, zijn gegevens omtrent de incidenten die zich tijdens, of juist buiten de diensten van de verdachte hebben afgespeeld; het aantal verpleegkundigen waarover de resterende diensten verdeeld worden is totaal irrelevant. We voeren nu een klein gedachtenexperiment uit. Als de diensten buiten de verdachte niet onder 26 resterende verpleegkundigen verdeeld zouden zijn maar onder 99 (ik noem maar een willekeurig getal), dan zou er voor Lucia de B. eigenlijk niets mogen veranderen. Maar als er 100 verpleegkundigen zouden zijn, dan zou de correctie van Elffers een vermenigvuldiging met 100, en niet met 27 moeten zijn. Natuurlijk, in de realiteit waren er 27 verpleegkundigen, en geen 100, maar dit gedachtenexperiment maakt duidelijk dat de feitelijkheid van de 27 verpleegkundigen geen enkele rol kan en mag spelen in de feitelijke berekening. Echter, in de berekening van Elffers speelt dit getal *wel* een rol.

Een tweede gedachtenexperiment geeft aan dat ook de spreiding van de diensten over de ziekenhuizen van belang is. Als de verdachte al haar diensten op één afdeling zou hebben gedraaid (in plaats van op drie), waarbij het totaal aantal diensten, het aantal door de verdachte gedraaide diensten, het totaal aantal incidenten en het aantal incidenten tijdens een dienst van de verdachte gelijk zouden zijn aan die van de reële casus, dan zou er een totaal andere kans uit komen, een veel grotere kans wel te verstaan. Dus het feit dat de verdachte haar diensten spreidt over meerdere afdelingen, is nadelig voor haar, en het lijkt duidelijk dat er iets mis moet zijn. De achtergrond van dit probleem is dat de vermenigvuldiging die Elffers toepast, onjuist is. Je mag de getallen van de drie afdelingen niet met elkaar vermenigvuldigen, en dit is als volgt eenvoudig in te zien.

Stel dat een bepaalde verpleegkundige op 10 afdelingen werkt, en dat ze op elke afdeling een 'normaal' aantal incidenten meemaakt. Laten we, om precies te zijn, zeggen dat de kans dat ze minimaal zoveel incidenten meemaakt als feitelijk gebeurt, op elke afdeling gelijk is aan ½. Volgens Elffers zou dan de kans dat ze op de 10 afdelingen bij elkaar zoveel incidenten meemaakt als feitelijk gebeurd is, gelijk moeten zijn aan $(½)^{10}$. Dit is een zo kleine kans dat dit de verpleegkundige verdacht maakt… Onzin, natuurlijk! Maar het is wel wat Elffers in zijn berekeningen doet.

4. De interpretatie van de cijfers van Elffers door het hof

In haar vonnis van 24 maart 2003 nam de rechtbank het getal van Elffers over, maar uit de motivering bij het vonnis blijkt herhaaldelijk dat de rechters de betekenis van het getal 1 op de 342 miljoen niet begrepen hebben. We geven een drietal voorbeelden.

Onder punt 7 van de bewijsoverwegingen schrijft de rechtbank het volgende:

In zijn rapport van 29 mei 2002 komt dr. H. Elffers tot de conclusie dat de kans dat een verpleegkundige bij toeval net zoveel incidenten (…) *zou meemaken als verdachte 1 op de 342 miljoen bedraagt.*

Onder punt 8 vinden we:

Uit zijn rapport van 29 mei 2002 blijkt verder dat dr. H. Elffers tevens de navolgende deelkansen heeft berekend:

a. De kans dat bij toeval een van de 27 verpleegkundigen bij 142 uit 1029 diensten betrokken zou zijn bij 8 incidenten (…) *is kleiner dan 1 op 300.000.*

b. De kans dat verdachte bij toeval (…).

Onder punt 11 vinden we tenslotte:

De rechtbank is van oordeel dat uit de door dr. H. Elffers uitgevoerde waarschijnlijkheidsberekeningen, zoals neergelegd in zijn rapport van 29 mei 2002, volgt dat het uitermate onwaarschijnlijk moet worden geacht dat de verdachte de in de tenlastelegging genoemde incidenten bij toeval heeft meegemaakt. Deze berekeningen geven derhalve aan dat het in hoge mate waarschijnlijk is dat er een verband bestaat tussen het werkzaam zijn van de verdachte en het zich voordoen van bedoelde incidenten.

Ik citeer zo uitgebreid uit de bewijsoverwegingen om te laten zien dat de rechtbank systematisch formuleringen hanteert zoals *kans* (…) *bij toeval*. Elffers beoogt de kans te berekenen dat iets gebeurt als we aannemen dat incidenten gewoon per toeval en zonder samenhang plaatsvinden. De rechtbank meent echter ten onrechte dat Elffers de kans op toeval uitrekent. Deze fout wordt vaak gemaakt, en staat bekend als de *prosecutor's fallacy*. Kort gezegd komt de fout hierop neer dat $P(A \mid B)$ en $P(B \mid A)$ verwisseld worden, terwijl dit doorgaans toch echt twee geheel verschillende kansen zijn. Dit is dus niet alleen maar een kwestie van verkeerde woorden, zoals Elffers ter zitting van het hof wilde doen geloven. De interpretatie van de rechtbank is fundamenteel onjuist, en suggereert een veel sterkere bewijskracht van de getallen dan deze werkelijk hebben. Bovendien is de conclusie die de rechtbank onder punt 11 trekt ('deze berekeningen geven derhalve aan…') ook nog eens fout, zelfs als men de eerdergenoemde onjuiste interpretatie van de getallen voor lief neemt.

Conclusie

Is de conclusie dat we dan maar helemaal moeten afzien van het gebruik van kansrekening en statistiek in het strafrecht, gezien de mate van subjectiviteit die hierbij een rol speelt?

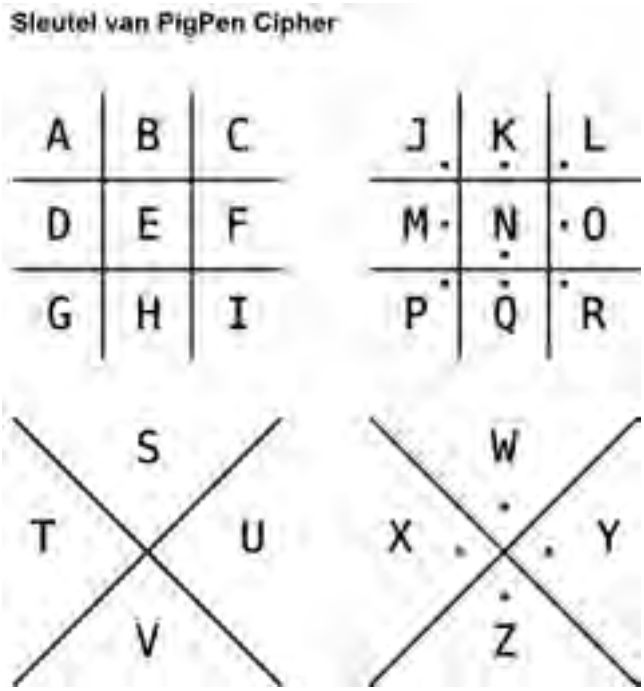
Nee, dat is niet de conclusie die ik trek, en dat brengt me ook bij de rol die de statisticus als expert zou moeten spelen in een proces. Een statisticus staat in een proces ten dienste van rechter, officier van justitie en/of verdediging. Zijn taak is om ervoor te zorgen dat gegevens die zich lenen voor statistische analyse, op de juiste wijze geïnterpreteerd worden. Dat betekent niet dat er slechts op één manier mee om te gaan is, dat heb ik betoogd. Juiste interpretatie betekent aan de ene kant dat de rechters inzicht moeten krijgen in de betekenis van gegevens. Is het gek dat bij een zuivere munt 65 van de 100 keer 'kop' gegooid wordt? Ja, dat is gek, en het is de taak van de statisticus om dat uit te leggen. Maar ook, en misschien wel vooral, is het belangrijk dat rechters gaan begrijpen dat getallen een beperkte betekenis hebben. Een getal is nooit alleen maar een getal, maar altijd het gevolg van keuzes. Uiteindelijk moet de rechter de getallen zelf

beoordelen, en ik kan me zo voorstellen dat hij of zij daar de hulp van een goede statisticus heel goed kan gebruiken. De statisticus helpt, kort gezegd, bij het interpreteren en begrijpen van de data.
Toen ik, tijdens de zitting in hoger beroep, de rechter duidelijk trachtte te maken hoe statistiek werkt, en ik vertelde over het maken van berekeningen onder bepaalde hypotheses, werd ik onderbroken: 'Mijnheer Meester, hier in de rechtszaal houden we ons niet bezig met hypothetische situaties, maar met de realiteit.'
Dat bedoel ik nou.

Noten

[1] Wat een incident precies is, is voor mijn bijdrage hier niet van belang. Het komt ongeveer neer op de noodzaak tot reanimatie, ongeacht de afloop ervan.
[2] In: *STAtOR* 5(2), juni 2004, p. 11.
[3] Deze 'minstens' was in het JKZ niet nodig, omdat daarbij alle incidenten al tijdens Lucia de B.'s diensten plaatsvonden.
[4] In: *Law, Probability & Risk* 2007.
[5] De keuze van een Poisson-verdeling ligt enigszins voor de hand, vanwege de aanname van onafhankelijkheid tussen verschillende diensten.

Over de auteur
Ronald Meester is hoogleraar kansrekening aan de Vrije Universiteit Amsterdam. Reacties naar diens e-mailadres (*rmeester@few.vu.nl*).



Vrij, Nederland!

Laatst las ik in mijn dagblad tot mijn schrik
Dat die verrekte hitsige Fransozen
Gemiddeld toch nog vaker minnekozen
Dan Nederlandse mannen zoals ik.

Dus klamp ik voortaan alle vrouwen aan:
"Kom, help mij om de Fransman te verslaan!"

Gepubliceerd door *Driek van Wissen* op 30-11-2002 op *www.gedichten.nl*

(*... Fransen vrijen gemiddeld 167 keer per jaar en zijn daarmee koplopers in Europa, de Nederlanders doen het 158 keer per jaar ...*)

Vergrijzing

WISKUNDE SCHOLEN PRIJS 2007, AFLEVERING 1

[Dédé de Haan en Klaas Mars]

In week 25 van 2007 was het zover: in die week werden de prijzen van de Wiskunde Scholen Prijs 2007 uitgereikt. Op woensdag 20 juni in Amstelveen en Hoorn, op vrijdag 22 juni in Rotterdam.
In drie afleveringen zult u kennismaken met de drie prijswinnende projecten. Deze aflevering gaat over het project Vergrijzing van de Gereformeerde Scholengemeenschap Randstad in Rotterdam.

PO Wiskunde A en Economie **Havo 4**

Vergrijzing

Groepsgrootte:	Presentatievorm:	aantal SLU:
4	Verslag min 6 A4 + interview + stage	15

Wordt de AOW onbetaalbaar?

Tijdens de laatste verkiezingen van afgelopen november was één van de grote strijdpunten de betaalbaarheid van de AOW.
De komende veertig jaar zal de Nederlandse bevolking sterk vergrijzen. Het aantal AOW-ers zal stijgen van 10 naar misschien wel 30% en het aantal werkenden dat de AOW moet opbrengen zal sterk dalen.
Tijdens de verkiezingen kwamen de politici met allerlei oplossingen om de AOW betaalbaar te houden. Bv. verhogen van de pensioenleeftijd naar 67 jaar, afschaffen van de VUT, i.p.v. een 36-urige werkweek naar een 40-urige werkweek, de ouderen meer belasting laten betalen, enz..

De bedoeling van deze PO is om in kaart te brengen hoe groot en wat het probleem nu eigenlijk is. Moeten wij ons zorgen gaan maken en welke oplossingen zijn goed. Hoe denken ouderen en hoe denken jongeren over dit probleem. Zijn jongeren in de toekomst nog wel bereid om deze zware lasten te dragen?
Je gaat als groepje een statistisch onderzoek verrichten om te zien hoe ouderen en jongeren over dit probleem denken en welke oplossing zij bespreekbaar vinden. Stel eerst een hypothese op. Deze hypothese ga je testen in het statistisch onderzoek.

Als economisch deskundige ga je de paragraaf schrijven over de betaalbaarheid van de vergrijzing. Daarin moet je duidelijk onderbouwd aangeven hoe het volgens jou verder moet met de verzorgingsstaat. Zorg dat je het verhaal onderbouwt met tabellen, grafieken, rekenvoorbeelden e.d.
Ieder lid van de groep levert apart een logboek in, zodat zichtbaar wordt wat je hebt gedaan en hoe lang je er mee bezig geweest bent.

figuur 1

Jaap zegt:
het volgende is: wat zijn de nadelen van de Clash tussen de generaties.
Jan zegt:
wat bedoelen ze daar nu weer mee?
Petra zegt:
Woordenboek: Clash betekent botsing van meningen die meestal tot een breuk leidt.
Jaap zegt:
Ik denk dat jongeren later meer moeten betalen voor de AOW omdat er dan veel meer ouderen zijn die AOW hebben en dat ze daarom er veel nadelen aan hebben omdat ze dan veel moeten betalen voor die ouwtjes
Jan zegt:
volgens mij niet
Petra zegt:
ik denk dat de jongeren moeten betalen en daardoor wordt de generatiekloof groter en groter wat tot een clash leidt
Ilze zegt:
dat denk ik ook
Jaap zegt:
ok
Jan zegt:
ok

figuur 2

Vakoverstijgend project

De opdracht behelst een onderzoek naar het probleem van de vergrijzing in Nederland. Het project is een samenwerking van de vakken wiskunde A en economie, uitgevoerd in havo-4.
In ***figuur 1*** staat het voorblad van de praktische opdracht, waarin de bedoeling van de opdracht geschetst wordt.
Binnen de school heeft dit project een voorbeeldfunctie: per profiel en per jaarlaag moet er een vakoverstijgend project komen dat meer samenhang tussen de vakken aanbrengt, vanuit een praktische situatie.
De opdracht die hier besproken wordt zal in 2008 ook in vwo-4 gebruikt gaan worden.

Originele aspecten

De opdracht werd gemaakt rond verkiezingstijd: niet lang na de laatste Tweede Kamer verkiezingen in november 2006. In alle debatten speelde het probleem van de vergrijzing een belangrijke rol, waarmee vergrijzing een bijzonder actueel onderwerp was. Wie herinnert zich niet de 'Bosbelasting' (de fiscalisering van de AOW), en alle reacties daarop?
Daarnaast zaten er twee andere originele aspecten aan de opdracht: de leerlingen moesten een chat-sessie over vergrijzing voeren op MSN. Deze chat-sessie moest als bijlage aan het in te leveren werkstuk toegevoegd worden. Een voorbeeld van een gedeelte van een chat-sessie kunt u zien in ***figuur 2***.
Verplicht onderdeel van de opdracht was tevens een maatschappelijke stage bij een verzorgingstehuis, inclusief een interview met een personeelslid of bewoner.

De oriëntatie

Het project werd uitgevoerd in groepen van vier leerlingen. De groepen werden samengesteld op grond van kwaliteiten: er moest in ieder geval een 'wiskundige' in, maar ook een 'creatieveling'.
De oriëntatie op het onderwerp vond gedeeltelijk in de wiskundeles plaats, maar ook thuis, zoals te lezen is in ***figuur 3***.

Het statistisch onderzoek

De wiskundige activiteiten zaten voornamelijk in het statistisch onderzoek en de verwerking ervan. De leerlingen moesten meerdere hypothesen opstellen en door middel van een enquête testen. Hierbij kwamen allerlei problemen aan de orde:

- Wat voor soort vragen stel je? Zijn ze open, of gesloten, en wat doe je dan met de antwoorden?
- Hoe voorkom je dat de vragen suggestief worden?
- Wat is een aselecte steekproef?
- Hoe verwerk je je gegevens in Excel?

Het maken van het onderzoeksplan alleen duurde al zeker vijf lessen!

In ***figuur 4*** staat een voorbeeld van een enquête, in ***figuur 5*** een voorbeeld van verwerkte resultaten van de enquête.
Het onderzoek draagt op een natuurlijk manier bij aan meer kennis over het onderwerp. De wiskunde geeft hier duidelijk meerwaarde aan het project.
De opdracht rond het onderzoek en de overige opdrachten zijn te lezen in ***figuur 6***.

Reacties van leerlingen

Zonder uitzondering is er met veel animo door de leerlingen gewerkt aan dit project. Dit uitte zich in het zonder pauze doorwerken in de mediatheek, zonder pauze doorwerken bij het afnemen van de enquête, en langer doorwerken dan gebruikelijk.
De leerlingen waren erg enthousiast over de maatschappelijke stage. Een aantal leerlingen heeft zelfs vervolg-afspraken gemaakt en later nog een vrije zaterdag vrijwilligerswerk gedaan in een verzorgingstehuis.

Enkele reacties van leerlingen zoals die in de verschillende verslagen gemeld zijn:

- *Het was een groot feest voor de ouderen, ik heb van deze dag ongelooflijk genoten. Meneer Mars en meneer De Ruiter: gun dit volgende groepen ook!*
- *Eindelijk een project dat leuk is en waar je wat van leert.*

Dat in dit project via deze stage duidelijk andere vaardigheden aan bod kwamen, blijkt ook uit de reactie van een leerling die op school slecht functioneert door ziekte en problemen: *'Ze vonden dat ik feeling had voor werken met ouderen en boden me vakantiewerk aan.'* Een mooie opsteker voor deze leerling!

De prijsuitreiking

Op vrijdag 22 juni 2007 was het zo ver; de prijsuitreiking zou plaatsvinden. Heel havo-4 was, onder leiding van wiskundedocent Klaas Mars, in de docentenkamer verzameld. Er was zelfs taart geregeld!

a. Ga naar *www.beeldengeluid.nl*. Tik in de archieven bij de zoekmachine "vergrijzing" in en bekijk de hieronder beschreven uitzending.

Titel: Netwerk
Netzender: Nederland 1
Zendgemachtigde: KRO
Uitzenddatum: 2004-10-10
Beschrijving: DOOR VERGRIJZING DREIGEND GENERATIECONFLICT OVER FINANCIERING PENSIOENEN EN AOW

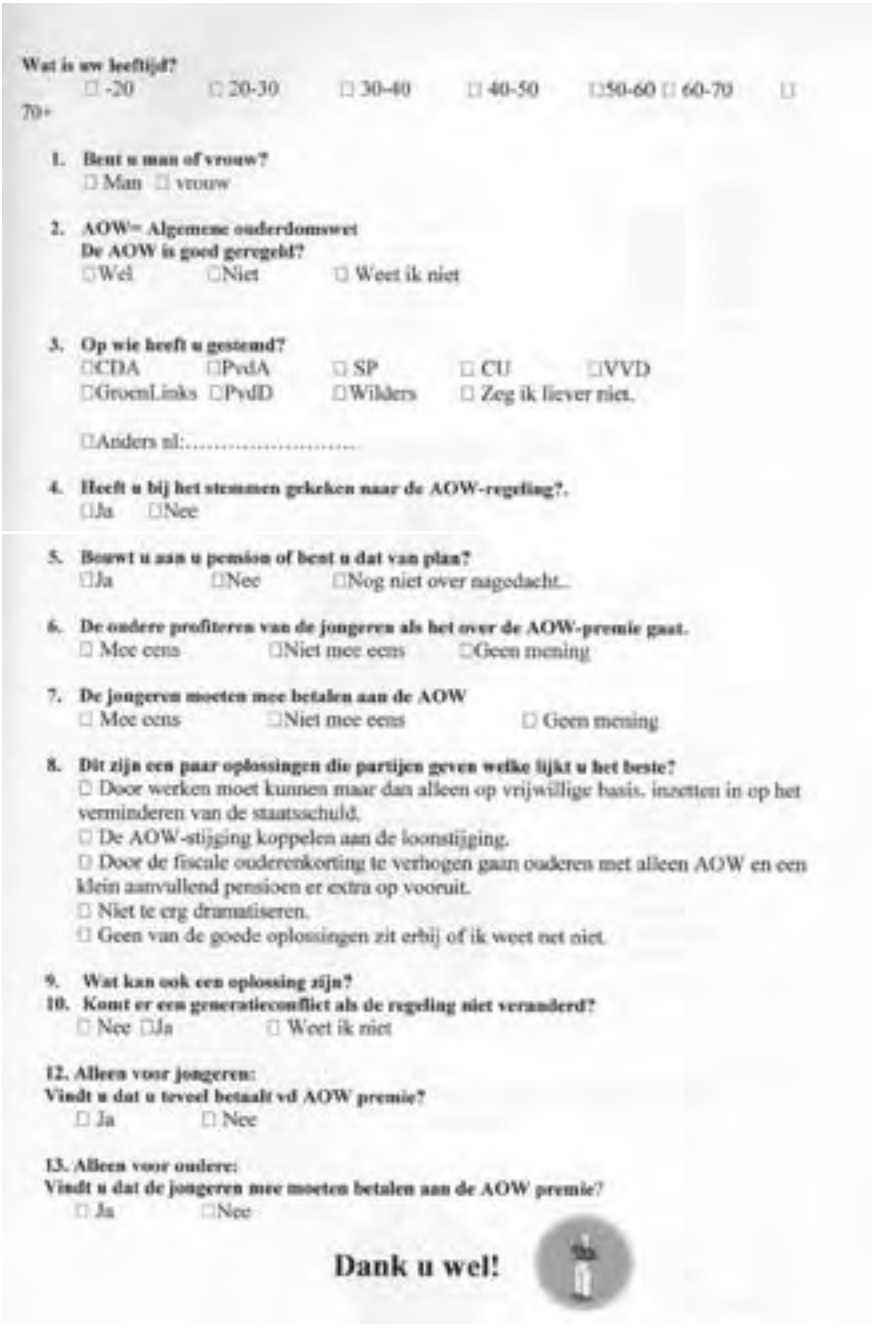
Verzamel naar aanleiding van deze uitzending materiaal om de onderzoeksvragen die je in de oriënterende wiskundeles hebt geformuleerd te kunnen beantwoorden. Je hoeft alleen maar aantekeningen te maken. Op school werkt één lid van de groep het verder uit.

b. Chatsessie.
Deze opdracht voer je thuis uit. Je gaat een chatsessie houden op MSN over vergrijzing. Aan de orde moeten komen:

- Verklaring van de titel, babyboomers, demografische situatie
- Clash tussen generaties, nadelen voor jongeren
- Verschil tussen AOW en pensioen
- Vergrijzing als economisch probleem
- Eventuele oplossingen

Je print de chatdiscussie uit en doet dit als bijlage bij je werkstuk. Woensdag of donderdag werkt iemand van je groepje de discussie uit voor het verslag.

figuur 3

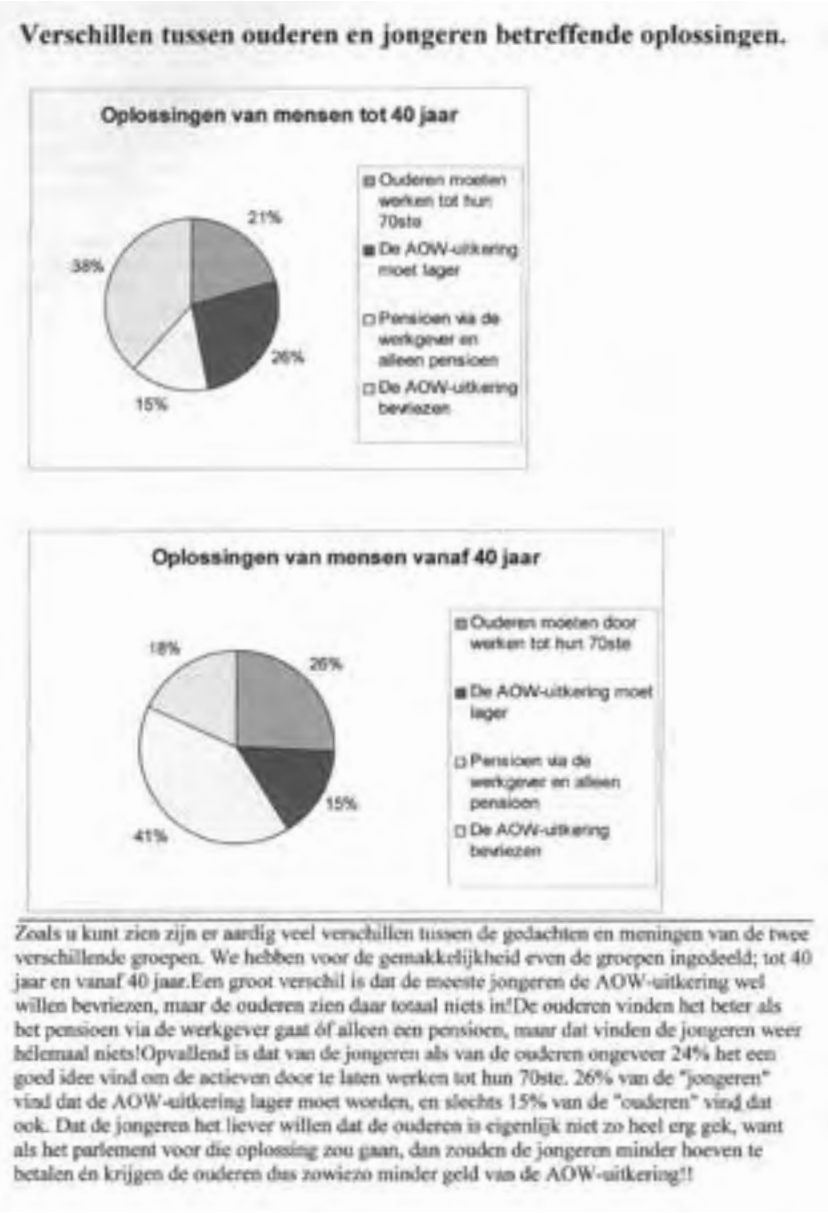


Wat is uw leeftijd?
☐ -20 ☐ 20-30 ☐ 30-40 ☐ 40-50 ☐ 50-60 ☐ 60-70 ☐ 70+

1. **Bent u man of vrouw?**
   ☐ Man ☐ vrouw
2. **AOW= Algemene ouderdomswet**
   De AOW is goed geregeld?
   ☐ Wel ☐ Niet ☐ Weet ik niet
3. **Op wie heeft u gestemd?**
   ☐ CDA ☐ PvdA ☐ SP ☐ CU ☐ VVD
   ☐ GroenLinks ☐ PvdD ☐ Wilders ☐ Zeg ik liever niet.
   ☐ Anders nl:.....
4. **Heeft u bij het stemmen gekeken naar de AOW-regeling?.**
   ☐ Ja ☐ Nee
5. **Bouwt u aan u pensioen of bent u dat van plan?**
   ☐ Ja ☐ Nee ☐ Nog niet over nagedacht...
6. **De ondere profiteren van de jongeren als het over de AOW-premie gaat.**
   ☐ Mee eens ☐ Niet mee eens ☐ Geen mening
7. **De jongeren moeten mee betalen aan de AOW**
   ☐ Mee eens ☐ Niet mee eens ☐ Geen mening
8. **Dit zijn een paar oplossingen die partijen geven welke lijkt u het beste?**
   ☐ Door werken moet kunnen maar dan alleen op vrijwillige basis. inzetten in op het verminderen van de staatsschuld.
   ☐ De AOW-stijging koppelen aan de loonstijging.
   ☐ Door de fiscale ouderenkorting te verhogen gaan ouderen met alleen AOW en een klein aanvullend pensioen er extra op vooruit.
   ☐ Niet te erg dramatiseren.
   ☐ Geen van de goede oplossingen zit erbij of ik weet net niet.
9. **Wat kan ook een oplossing zijn?**
10. **Komt er een generatieconflict als de regeling niet veranderd?**
    ☐ Nee ☐ Ja ☐ Weet ik niet
12. **Alleen voor jongeren:**
    Vindt u dat u teveel betaalt vd AOW premie?
    ☐ Ja ☐ Nee
13. **Alleen voor ondere:**
    Vindt u dat de jongeren mee moeten betalen aan de AOW premie?
    ☐ Ja ☐ Nee

Dank u wel!

figuur 4



figuur 5

figuur 6

De sfeer was opgewonden. Na een introductie over het doel van de Wiskunde Scholen Prijs en wat statistieken over het aantal deelnemers en de verschillende categorieën, kwam het belangrijkste deel: het voorlezen van het juryrapport en het overhandigen van de prijs.

Het juryoordeel

Volgens het juryrapport:
'Een mooi, leerzaam bovenbouwproject met maatschappelijke relevantie. Het project is op natuurlijke wijze vakoverstijgend. Het wiskundige onderzoek voegt echt iets toe aan de uitwerking van het thema; allerlei statistische vaardigheden komen aan bod. Het spreekt de jury aan dat het een open opdracht is – er is genoeg ruimte voor een eigen opzet. Het project is betekenisvol en rijk: diverse media worden gebruikt, waaronder een chat-sessie en een maatschappelijke stage, zodat allerlei vaardigheden aan de orde komen. Zonder twijfel de hoofdprijs waardig!'

Overdraagbaarheid

De beide ontwerpers van de opdracht denken dat het project zonder meer overdraagbaar is naar andere scholen – in ieder geval kan het gebruik van verschillende media (chat-sessie, maatschappelijk stage) een eye-opener zijn voor collega's.

Napraten met docenten en leerlingen

De docenten Mars (wiskunde) en De Ruiter (economie) zijn erg verheugd over het project, het enthousiasme en de betrokkenheid van de leerlingen, en het winnen van de prijs. Maar die prijs, dat vinden ze maar bijzaak: het gaat ze om het stimuleren van de leerlingen. Voor Klaas Mars is dat niet iets wat hij met dit project nu voor het eerst doet. Vorig jaar won hij óók de Wiskunde Scholen Prijs, met een ander project ('Armoede in Nederland'). Daarnaast is de school ook op andere terreinen bijzonder actief: als deelnemer aan het Universum-project, waarbij de school subsidie krijgt om te stimuleren dat vooral meisjes kiezen voor wiskunde B. Ook is Klaas Mars ambassadeur van het SALVO-project (Samenwerkend Activerend Leren V.O.), waarbij meer samenhang tussen de exacte vakken aangebracht wordt, waardoor de motivatie van de leerlingen toeneemt en hun resultaten vooruitgaan.

Het is in dit licht niet zo verbazingwekkend dat ook de secties economie en wiskunde A elkaar op het havo gevonden hebben om samen een praktische opdracht te ontwikkelen.
Ook de leerlingen zijn zonder uitzondering vol lof over dit project, en vinden dat het terecht een prijs heeft verdiend.
Kortom: een duidelijk voorbeeld van een 'good practice'.

Informatie
Wie meer over dit project wil weten, kan contact opnemen met Klaas Mars *(e-mail: mrk@gsr.nl).*
Meer informatie over de Wiskunde Scholen Prijs is te vinden op *www.wiskundescholenprijs.nl.*
U kunt zich nog (vrijblijvend) aanmelden voor de Wiskunde Scholen Prijs 2008 tot en met 29 februari 2008, via *scholenprijs@fi.uu.nl.*

De winnaars (links achteraan Geert de Ruiter, rechts vooraan Klaas Mars)

Taart in Rotterdam bij de prijsuitreiking

Noot
Met dank aan economiedocent Geert de Ruiter.

Over de auteurs
Dédé de Haan is werkzaam bij het Freudenthal Institute for Science and Mathematics Education (Universiteit Utrecht) en organisator van de Wiskunde Scholen Prijs.
E-mailadres: *d.dehaan@fi.uu.nl*
Klaas Mars is werkzaam als wiskundedocent bij de Gereformeerde Scholengemeenschap Randstad in Rotterdam.
E-mailadres: *mrk@gsr.nl*

De Wiskunde Scholen Prijs
De prijs wordt jaarlijks uitgereikt sinds 2002. Het is een stimuleringsprijs, bedoeld om scholen uit te nodigen met hun sterke punten op het gebied van wiskundeonderwijs naar buiten te treden, en op die manier goede initiatieven zichtbaar te maken en het imago van wiskunde te verbeteren.
Scholen kunnen strijden in drie categorieën: vmbo, onderbouw havo/vwo en bovenbouw havo/vwo.
Ieder project dat ingezonden wordt (en dat kan van alles zijn, van vakoverstijgend tot verdiepend) wordt beoordeeld door een deskundige jury. In iedere categorie valt € 1000,00 te winnen.
De Wiskunde Scholen Prijs is voortgekomen uit het WisKids-project. Dit was een gezamenlijk initiatief van het Wiskundig Genootschap, de Nederlandse Vereniging voor Wiskundeleraren en de Nederlandse Vereniging tot Ontwikkeling van het Reken-Wiskunde Onderwijs.

Bijziendheid en geboortemaand

[Hans van Maanen]

Season of Birth, Natural Light, and Myopia

Yossi Mandel, MD, MHA,[1,2] Itamar Grotto, MD, MPH,[1,3] Ran El-Yaniv, PhD,[4] Michael Bellin, MD, MA,[5] Eran Imanli, MD, MHA,[1,6] Uri Polat, PhD,[7] Elisha Barnea, MD[7]

Purpose: To investigate the possible roles of season of birth and perinatal duration of daylight hours (photoperiod) in the development of myopia.
Design: Retrospective, population-based, epidemiological study.
Participants: A total of 276 911 adolescents (157 663 male, 119 248 female) 16 to 22 years old. All were Israeli-born conscripts to the Israeli Defense Forces who were examined during the 5-year period 2000 through 2004.
Methods: Noncycloplegic refraction was determined by autorefractometer and validated by qualified optometrists. Myopia, defined on the basis of right eye spherical equivalence, was classified as mild (−0.75 to −2.99 diopters [D]), moderate (−3.0 to −5.99 D), or severe (−6.0 D or worse). The photoperiod was recorded from astronomical tables and classified into 4 categories. Using multivariable logistic regression models, we calculated odds ratios (ORs) for several risk factors of myopia including season of birth.
Main Outcome Measure: The OR for photoperiod categories as risk factors for myopia.
Results: Overall prevalences of mild, moderate, and severe myopia were 18.8%, 8.7%, and 2.4%, respectively. There were seasonal variations in moderate and severe myopia according to birth month, with prevalence highest for June/July births and lowest for December/January. On multivariate logistic regression, the ORs of photoperiod categories for moderate and severe myopia were highly significant and demonstrated a dose–response pattern. Odds ratios for severe myopia were highest for the shortest versus the longest photoperiods (1.24; 95% confidence interval, 1.15–1.33; P<0.001). Mild myopia was not associated with season of birth or perinatal light exposure. Other risk factors were gender (1.14 for female), education level (1.32 for age above 12), and father's origin (1.31 for Eastern vs. Israeli origin).
Conclusion: Myopia in this population is associated with birth during summer months. The exact associating mechanism is not known but might be related to exposure to natural light during the early perinatal period. Ophthalmology 2007;xx:xxx © 2007 by the American Academy of Ophthalmology.

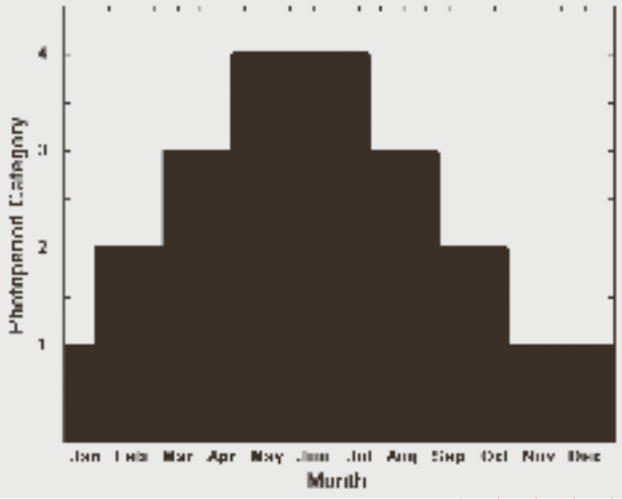
Wetenschapsnieuwtje?

Het is geen wetenschap en het is geen journalistiek, dus het zal wel wetenschapsjournalistiek wezen. Voor een 'grappig' onderzoeksresultaatje gaan de kolommen immers altijd open. We nemen een min of meer willekeurig, dat wil in dit geval zeggen een zo recent mogelijk, voorbeeld. De uiterste inleverdatum voor dit artikel was 1 september 2007, op 29 augustus 2007 werden de lezers van de doordeweekse wetenschapspagina van *NRC Handelsblad* verrast met de eenkolommer 'Kind geboren in juni of juli wordt later vaker bijziend'. De eerste zinnen: 'Kinderen die in juni of juli geboren worden hebben een grotere kans om later zwaar bijziend te worden dan kinderen die in december of januari ter wereld komen. In Israël scheelt het wel 24 procent. Dat bericht de Israëlische oogheelkundige Yossi Mandel in *Ophthalmology* (augustus) na een onderzoek onder 278.000 dienstplichtigen van 16 tot 22 jaar oud.' Het nieuwtje stond eerder op de website van 'Eurekalert' (*www.eurekalert.org*), de bron waaraan alle wetenschapsjournalisten zich laven, dus het bericht gonsde de hele wereld over. Het ziet er niet naar uit dat ook maar één van hen de moeite heeft genomen het wetenschappelijke artikel er even bij te pakken en op zijn merites te beoordelen.

Het tijdschrift *Ophthalmology*, uitgegeven door de American Academy of Ophthalmology, is een van de best aangeschreven oogheelkundige bladen. Het artikel van Mandel is echter niet in het augustusnummer van het blad te vinden. Wel blijkt het op het internet beschikbaar te zijn gesteld. 'Season of birth, natural light, and myopia' heet het, en Mandel was, toen hij het onderzoek uitvoerde, als oogarts werkzaam aan het Edith Wolfson Medical Center in Holon.

Fotoperiode

De conclusie van het artikel laat aan duidelijkheid niets te wensen over, al moeten de auteurs de verklaring schuldig blijven: 'Bijziendheid in deze bevolking houdt verband met geboorte tijdens de zomermaanden. Het exacte mechanisme achter het verband is onbekend, maar zou gerelateerd kunnen zijn met de blootstelling aan daglicht gedurende de vroege perinatale periode.' Dus doordat kinderen in de zomer rond hun geboorte meer daglicht krijgen, worden ze misschien wel sterker bijziend. Voor het onderzoek is Mandel niet over een nacht ijs gegaan. Hij verzamelde de gegevens van maar liefst 276.911 (afgerond 277.000) adolescenten die zich tussen 1999 en 2005 in Israël voor de dienstplicht meldden. Om redenen die hij niet nader toelicht, bekeek Mandel alleen het rechteroog, en noteerde of dat licht (tussen -0,75 en -2,99), matig (tussen -3,0 en -5,99) of ernstig (meer dan -6,0) bijziend was. Al evenmin vertelt Mandel waarom hij zoveel tussenliggende waarden in de gegevens kan weggooien, en waarom hij deze indeling in drie categorieën kiest.
Vervolgens moet de 'fotoperiode', dus de hoeveelheid daglicht, worden gemeten, en ook hier is over nagedacht. 'Daglichttijden werden berekend uit astronomische tabellen. Voor elke dag van het jaar namen wij het gemiddelde van de fotoperiode van de volgende 30 dagen. Deze gemiddelden werden verder gegroepeerd in 4 categorieën (90-91 dagen elk) met fotoperioden van



figuur 1 Mandels eerste grafiek: 'Fotoperiodecategorieën naar geboortemaand'

10,1-10,8 uur, 10,81-12,2 uur, 12,21-13,57 uur en 13,58-14,23 uur.'
Behulpzaam levert Mandel er een grafiek bij - een grafiek die stellig kan meedingen naar de prijs voor de grootste inkt-dataverhouding; ***zie figuur 1***.
Wat Mandel hier dus eigenlijk gedaan heeft, is het herindelen van de seizoenen. In de winter is er weinig daglicht (categorie 1), in het begin van de lente en het eind van de herfst wat meer (2), aan het eind van de lente en het begin van de herfst nog wat meer (3), en in de zomer het meest. Waarbij opgemerkt moet worden dat Mandels 'zomer', dus categorie 4, op grond van een reconstructie met behulp van astronomische tabellen ongeveer moet lopen van 27 april tot 27 juli. Een simpele tabel zou veel vragen en rekenwerk hebben voorkomen.
Belangrijker is, dat Mandel hier weer een schat aan gegevens weggooit. Waarom maar vier categorieën, en niet acht of twaalf of 183? En vooral, waarom zo ingewikkeld?
In de conclusie heeft Mandel het over de 'perinatale periode', maar in feite kijkt hij naar de post-natale periode: de eerste dertig dagen na de geboorte.
Waarom dertig? Waarom niet zestig, of waarom niet juist naar de tweede maand na de geboorte gekeken? Mandel zegt er niets over, oogheelkundige overwegingen lijken geen rol te hebben gespeeld. En hoeveel uren komt een baby de eerste maand buiten? Is er eigenlijk wel een verband tussen de hoeveelheid daglicht die beschikbaar is en de hoeveelheid daglicht die een kind op zijn bolletje krijgt? Had Mandel niet eerst moeten aantonen dat kinderen in hun eerste maand 's zomers meer licht ontvangen dan 's winters? Nu lijkt het, met alle respect zoals dat heet, toch meer op astrologie - het op niets af relateren van hemelverschijnselen met de menselijke lotgevallen - dan op wetenschap.
Maar misschien is de statistiek ijzersterk.

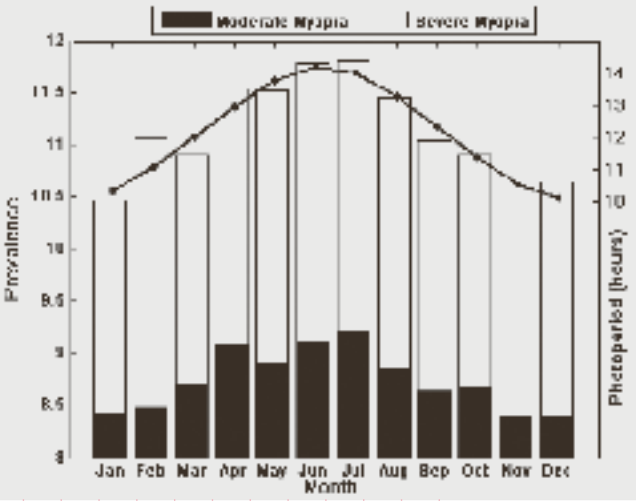
Lichtcategorie	mild bijziend	matig bijziend	ernstig bijziend
2	1,00 (0,98–1,03)	1,03 (0,99–1,06)	1,09 (1,02–1,17)
3	1,03 (1,00–1,05)	1,06 (1,02–1,11)	1,11 (1,03–1,19)
4	1,03 (1,00–1,06)	1,08 (1,04–1,13)	1,24 (1,16–1,33)

tabel 1 Odds ratio's van verschillende maten van bijziendheid naar hoeveelheid licht na de geboorte vergeleken met het donkerste jaargetijde

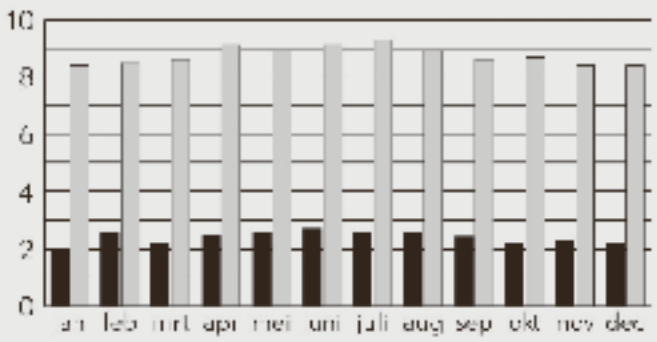
Als bewijs voor zijn gelijk levert Mandel een tabel, en wederom een grafiek. In de tabel (***zie tabel 1***) worden de daglichtcategorieën afgezet tegen de bijziendheid.
De eerste verrassing is, dat hieruit in het geheel niet blijkt dat bijziendheid gerelateerd is aan daglicht - voor milde bijziendheid zeker niet. En aangezien dat de grootste categorie is, zou het best kunnen dat bijziendheid over het geheel genomen niet gerelateerd is aan daglichtcategorie. Mandel gaat meteen over op analyse van de subgroepen, maar geeft geen totaaloverzicht.
Alleen 'matige' (-3 tot -6) en 'ernstige' (-6 en hoger) bijziendheid staan in verband met Mandels fotocategorieën, en dan nog niet heel overtuigend. Voor matige bijziendheid zijn de odds (dus de verhouding bijziend op niet-bijziend) voor de zomer 1,08 keer zo groot als voor de winter. Voor ernstige bijziendheid is die odds ratio 1,24.
Waarbij niet vergeten moet worden dat de statistiek wordt toegepast op een onvoorstelbaar grote populatie (van een steekproef is eigenlijk geen sprake, wat het gebruik van statistiek en betrouwbaarheidsintervallen al wat merkwaardig maakt). Met zulke grote groepen is het namelijk betrekkelijk eenvoudig een significant resultaat te boeken; bij een steekproef van bijvoorbeeld 10 000 mensen is een onbetekenend IQ-verschil van 0,2 punten al statistisch significant.

Seizoenseffect?

Mandel heeft nog een troef om zijn slag binnen te halen. '***Figuur 2*** toont de prevalentie van matige en ernstige bijziendheid naar geboortemaand.' Merk op dat hij de *y*-as laat beginnen bij 8 procent, dat doet hij om 'het seizoenseffect beter grafisch te laten uitkomen', zo zegt hij in het bijschrift.
Vreemd is allereerst dat Mandel hier opeens weer een indeling in twaalf geboortemaanden geeft, in plaats van zijn zo zorgvuldig berekende categorieën. En hoewel Mandel steeds al zijn analyses op matige en ernstige bijziendheid apart uitvoert, voegt hij ze hier samen: de witte balkjes, de ernstige bijziendheid, heeft hij bovenop de zwarte gestapeld. Dat zal het seizoenseffect ongetwijfeld beter laten uitkomen, maar het ontneemt nogal het zicht op de invloed van de geboortemaand op ernstige bijziendheid. Die moet nu met liniaal en potlood gereconstrueerd worden - en dat levert een nieuwe grafiek (hier ***figuur 3***), en een nieuw inzicht op.
Uit de reconstructie blijkt immers, dat februari er, qua ernstige bijziendheid, ook wel mag wezen. Juni scoort inderdaad hoog, met 2,7 procent, maar naast mei, juli en augustus komt ook de donkere februarimaand tot 2,6 procent ernstig bijzienden.
Omdat absolute aantallen ontbreken, is het onmogelijk een exacte statistische toets te doen, maar de zwarte balkjes ogen niet alsof ze daartegen bestand zijn.
De derde creatieve vondst is natuurlijk om in figuur 2 ook meteen de fotoperiode in te tekenen, met de as rechts. In het bijschrift licht Mandel toe dat die lijn 'de maandelijkse gemiddelde fotoperiode' voorstelt, en door een wonderbaarlijk toeval is die *y*-as precies zo uitgevallen dat de golfvorm samenvalt met het verloop van de balkjes.



figuur 2 Mandels tweede grafiek: 'Geboorteseizoen en prevalentie van bijziendheid'



figuur 3 Geboortemaand en prevalentie van bijziendheid (zwart is ernstig, grijs is matig bijziend)

Waarbij de aantekening past dat de 'maandelijkse gemiddelde fotoperiode' iets anders is dan de 'fotoperiode' over de komende dertig dagen - die Mandel juist zo netjes had uitgerekend. Het scheelt, uiteraard, een dag of veertien, en het zou zomaar kunnen dat ook daardoor de seizoenseffecten wat beter uitkomen.

24 procent

Rest de vraag: om hoeveel mensen gaat het nu? Wat betekent die odds ratio van 1,24 hier? Betekent het werkelijk: 'In Israël scheelt het 24 procent', zoals *NRC Handelsblad* schreef? Wat is het werkelijke effect van die vier uren extra daglicht in periode 4?
Er is nog een derde figuur, maar die wordt gelukkig gedubbeld door een tabel waaruit de cijfers ook zijn af te lezen. Elders in het artikel valt zelfs het totale aantal adolescenten in elke fotocategorie te vinden (waarbij Mandel overigens, verrassenderwijs, een standaardafwijking geeft: het aantal kinderen geboren tijdens een fotoperiode van de eerste categorie is volgens hem 70 358 met een standaardafwijking van 25,41).

Lichtcategorie	totaal	% matig bijziend	% sterk bijziend
1	70.358	8,5	2,2
2	69.045	8,6	2,4
3	69.559	8,9	2,4
4	67.949	9,0	2,7

tabel 2 Aantal deelnemers en percentage bijziendheid naar lichtcategorie

Afijn. Het scheelt, als we deze tabel aanhouden, in bijziendheid de helft van een procent - als 1000 kinderen in de donkere maanden waren geboren in plaats van in de lichte, zouden er 5 minder matig bijziend zijn en 5 minder ernstig bijziend. Maar '1 op de 200' klinkt toch wat minder dan 'het scheelt 24 procent'. Als alle 276.911 kinderen in dit onderzoek tussen 27 april en 27 juli waren geboren, zou dat Israël in vijf jaar tijd 517 ernstig bijzienden hebben gescheeld. Nog steeds aannemende, natuurlijk, dat er een oogheelkundige redenering zit achter de ingewikkelde berekening van de gemiddelde lichtblootstelling in de eerste dertig dagen en dat deze niet, na een paar meer vanzelfsprekende berekeningen, de enige was die een statistisch presentabel resultaat gaf. Het zou allemaal een stuk simpeler zijn als Mandel gewoon de geboortedata direct had gekoppeld aan de dioptrieën.

Correlatie of causaliteit

Bekend is dat ook geslacht, opleiding en genetische factoren van belang zijn in de ontwikkeling van bijziendheid. Dat blijkt ook uit het onderzoek van Mandel en collega's. Het blijkt zelfs heel goed. Terwijl bij matige bijziendheid de odds ratio van winter tegenover zomer uitkomt op 1,08, is die voor vrouw tegenover man maar liefst 1,18. Het geslacht speelt dus een veel belangrijker rol in de ontwikkeling van de bijziendheid dan de eerste dertig dagen licht.
Hetzelfde geldt voor de herkomst. Als de vader van het kind niet in Israël is geboren, is de odds dat het bijziend wordt 1,4 keer zo groot als wanneer de vader wel in Israël is geboren.
Nog meer scheelt de opleiding. Bekend is dat bolleboosjes vaker bijziend zijn — het heeft te maken met veel lezen, of in ieder geval veel kijken naar dingen dicht bij de ogen. Als de odds op matige bijziendheid van een kind dat minder dan twaalf jaar onderwijs heeft genoten weer op 1 wordt gesteld, dan is de odds voor een kind dat meer dan twaalf jaar in de klas heeft gezeten 1,72. Ook opleiding speelt dus een veel grotere rol dan de theoretische blootstelling aan daglicht.

Wie niet wil dat zijn kind bijziend wordt, moet niet het licht, maar de school mijden.

Waarbij we, ten slotte, voor het gemak maar voorbijgaan aan de sprong die in het artikel wordt gemaakt van correlatie en causaliteit. Als er de eerste dertig dagen van een kinderleven veel licht is, heeft dat kind iets meer kans bijziend te worden. Maar komt dat door dat licht?
De onderzoekers nemen min of meer vanzelfsprekend aan dat er sprake is van een oorzakelijk verband. Ze gaan zelfs serieus in op de vraag, of dat verband niet vertroebeld zou kunnen worden doordat hoger-opgeleide, bijziende echtparen liever kinderen in de zomer krijgen, waardoor er meer bijziende kinderen in de zomer worden geboren. Nee, dat is niet zo, menen zij, het ligt aan de melatonine-huishouding - melatonine is een krachtig hormoon waarvan de productie mede afhangt van het dag-en-nachtritme.
Het zou kunnen, maar er is natuurlijk nog wel meer dat in de zomer anders is dan in de winter. De temperatuur bijvoorbeeld. Of misschien is het niet de lengte van de dag, maar de zonshoogte, of de hoeveelheid ultraviolet licht. Misschien laten domme ouders hun jonggeborenen op donkere winteravonden eerder uit hun handen vallen, zodat die op hun hoofdje terechtkomen en later minder intelligent worden en minder bijziend worden.

Filterfunctie

Misschien is het allemaal wel onzin. Een onzinnig idee, slecht uitgevoerd, misleidend gepresenteerd.
Wie een wetenschappelijk artikel schrijft, krijgt dat altijd gepubliceerd. Misschien niet in een topblad, maar er zijn genoeg wetenschappelijke bladen die om kopij zitten te springen en het review-proces niet zo nauw nemen. Zelfs topbladen als *Nature*, *Science* en *Ophthalmology* laten nog wel eens een steekje vallen. Dat iets in een wetenschappelijk tijdschrift staat, is dus geen kwaliteitsgarantie. (Of dat vroeger anders was, zullen we hier in het midden laten.)
Als die filterfunctie er niet (meer) is, valt die taak toe aan wetenschapsjournalisten. Zoals een economisch redacteur een jaarverslag doorpluist en de jaarrede van de directeur kritisch tegen het licht houdt, zoals een parlementair redacteur doorvraagt bij kamerleden en niet klakkeloos de conclusie van elk rapport overneemt, zo moeten wetenschapsredacteuren een vinding op haar merites beoordelen voordat ze erover schrijven.
Een bericht in de krant zetten zonder het oorspronkelijke artikel te lezen, is een van de zeven hoofdzonden van de wetenschapsjournalistiek. Over de andere zes zullen we hier maar zwijgen.

Over de auteur

Hans van Maanen is zelfstandig wetenschapsjournalist. Hij schrijft regelmatig voor *de Volkskrant*, *Quest*, *Akademienieuws* en *AMC Magazine*. Ook schreef hij verschillende populair-wetenschappelijke boeken, zoals *Een gezonde geest*, *FC Algebra* en de *Encyclopedie van misvattingen*. In 2007 ontving Hans van Maanen twee prijzen voor wetenschapsjournalistiek: de Van Walreeprijs van de KNAW en de Eurekaprijs van NWO.
E-mailadres: *hans@vanmaanen.org*

Alwéér die drie deuren?

[Jan van de Craats]

Het *driedeurenprobleem* is zo langzamerhand een klassieker geworden. Je ziet het op allerlei plaatsen opduiken, en telkens leidt het weer tot heftige discussies. Ook in de klas zal het zijn uitwerking niet missen. Leuk voor een spannend project kansrekening, vooral als je de simpele oplossing hieronder pas achteraf geeft, nadat alle stofwolken zijn opgetrokken.

Het probleem

In zijn eenvoudigste vorm luidt het probleem als volgt. Stel dat je aan een televisiequiz meedoet en alle vragen goed hebt beantwoord. De quizmaster brengt je bij drie gesloten deuren. Achter een van die deuren staat de hoofdprijs: een splinternieuwe auto. Achter de beide andere deuren staat niets. De quizmaster, die vertelt dat hij weet waar de auto staat, vraagt je één deur aan te wijzen.
'Weet je het zeker?' vraagt hij vervolgens, en opent een van de andere deuren. Daarachter geen auto. 'Ik weet het goed met je gemaakt' zegt hij. 'Je mag nog wisselen. Als je de juiste deur kiest, is de auto voor jou!' Wissel je of niet?

Pech of geluk

Wat moet je doen? Waar de auto staat, weet je niet. Als je wisselt, heb je geluk of pech, en als je niet wisselt ook. Het blijft een gok. Heb je het idee dat de quizmaster kwaadaardig is en je alleen maar die keuze aanbiedt omdat hij weet dat je de eerste keer goed gekozen hebt, dan kun je maar beter bij je eerste keuze blijven. Maar denk je dat hij je juist wil helpen omdat hij weet dat je de verkeerde keuze hebt gemaakt, dan moet je natuurlijk wisselen.

Het kan ook zo zijn - en dat wordt bij besprekingen van het driedeurenprobleem meestal stilzwijgend verondersteld - dat de quizmaster neutraal is, dat wil zeggen dat hij de opdracht heeft om na de eerste keuze van de kandidaat altijd een deur te openen waarachter niets staat en dat hij de kandidaat daarna ook altijd de mogelijkheid moet geven om te wisselen. Onder die aannames kun je de hele gang van zaken als een *kansexperiment* definiëren dat je in principe net zo vaak kunt herhalen als je wilt. Je kunt het ook op de computer of op je grafische rekenmachine simuleren. Dat op zichzelf is al een leerzaam project: het dwingt je om alle veronderstellingen expliciet te maken en het geheel netjes te programmeren. Voer vervolgens een lange serie simulaties uit waarbij je elke keer vaststelt of wisselen je de auto oplevert of niet.

Wisselen loont!

Maar het moet natuurlijk ook mogelijk zijn te *beredeneren* hoe de kansen liggen. Dat is niet zo eenvoudig, zoals de geschiedenis van het driedeurenprobleem leert: talloos veel artikelen zijn er al over geschreven. Zelfs professionele wiskundigen zijn ermee de fout ingegaan.
Die foutieve redeneringen leiden meestal tot de conclusie dat wisselen niets uitmaakt. Groot is de verrassing als bij een simulatie blijkt dat daar niets van klopt. Toch is een correcte redenering, met de juiste conclusie, niet moeilijk te volgen.
Mijn favoriete uitleg gaat als volgt.

Noem de deuren A, B en C. Zonder beperking van de algemeenheid kun je aannemen dat de auto achter deur A staat, en dat jouw eerste keuze volledig door het toeval wordt bepaald. Je kiest dus met gelijke kansen deur A, B of C. Heb je A gekozen, dan krijg je de prijs als je niet wisselt. Heb je echter B of C gekozen, dan krijg je de auto als je wél wisselt. In twee van de drie gevallen levert wisselen dus de auto op. Bij een neutrale quizmaster is wisselen dus verstandig: je verdubbelt daarmee je winstkans.

Een ander probleem?

Hoe komt het dat zo veel mensen deze simpele oplossing niet vinden, of zelfs niet willen aanvaarden, totdat ze door een simulatie van de juistheid ervan worden overtuigd? Dat is meer een kwestie van psychologie dan van wiskunde. Het verhaal van de quizmaster die deuren opent zet je blijkbaar op het verkeerde been. Het is een afleidingsmanoeuvre, die je in verwarring probeert te brengen. Eigenlijk geeft dat openen van een deur waar niets achter staat helemaal geen extra informatie: van tevoren weet je immers al dat van de twee deuren die je niet gekozen hebt, er altijd minstens één geen auto verbergt. Je zou het kansexperiment daarom ook zo kunnen inrichten, dat de quizmaster helemaal geen deur opent, maar eenvoudig zegt dat je bij wisselen de *beide* andere deuren mag openen. Er

is maar één auto, dus die grootmoedigheid kost hem niets. En het driedeurenprobleem wordt er niet wezenlijk anders door.

Maar als je het probleem zo formuleert, snapt iedereen meteen dat het (nog steeds bij een neutrale quizmaster) voordelig is om te wisselen. En dat je je winstkans daarmee verdubbelt. Het driedeurenprobleem is nu *triviaal* geworden: zelfs de meest verstokte wiskunde-analfabeet zal er geen moeite meer mee hebben. Je hebt dan ook geen computersimulaties nodig om jezelf te overtuigen: alles spreekt vanzelf. Het enige dat misschien niet direct vanzelf spreekt, is dat het hier in wezen echt om hetzelfde probleem gaat. Daarover moet je toch wel even nadenken.

Voor de lessen kansrekening is het leuk om het driedeurenprobleem na afloop in verband te brengen met de binomiale verdeling. Bij n uitvoeringen van het bijbehorende kansexperiment met als strategie dat je altijd wisselt, is de succeskans per keer gelijk aan $p_w = \frac{2}{3}$. De kansvariabele die het totale aantal successen telt, is dus binomiaal verdeeld met parameters n en p_w. Neem je als strategie juist om nooit te wisselen, dan is het totale aantal successen in een serie van n uitvoeringen van het experiment binomiaal verdeeld met parameters n en $1 - p_w = \frac{1}{3}$. Degenen die het driedeurenprobleem op de computer of op de grafische rekenmachine hebben gesimuleerd, hebben dus eigenlijk de binomiale verdeling gesimuleerd!

Meer dan drie deuren

Vaak wordt ook gevraagd hoe het zit als er meer dan drie deuren zijn, bijvoorbeeld vier (en nog steeds maar één auto). Stel dan dat de quizmaster nadat je je eerste keuze gemaakt hebt, een of twee lege deuren opent en je vervolgens weer de kans geeft te wisselen. Ook dan is wisselen lucratief. Opende hij één deur, dan vergroot je je kans door te wisselen van een kwart naar een half. Opende hij twee deuren, dan is je winstkans bij wisselen zelfs driekwart. De triviale variant van het *vierdeurenprobleem* is nu dat je bij wisselen twee, respectievelijk drie van de nog niet gekozen deuren mag openen.
Bij meer dan vier deuren gelden soortgelijke redeneringen. Toch moet je in de praktijk voorzichtig zijn. Wie garandeert je dat de quizmaster inderdaad neutraal is?

Over de auteur
Prof.dr. Jan van de Craats is hoogleraar wiskunde aan de Universiteit van Amsterdam en aan de Open Universiteit.
E-mailadres: *craats@science.uva.nl*

Verschenen / De Archimedes-codex

Uit de flaptekst – Verloren gewaande tekeningen en stukken tekst van Archimedes zijn teruggevonden onder de bladzijden van een gebedenboek uit de dertiende eeuw. Deze verborgen tekst, die nu wordt ontcijferd door handschriftkundigen, laat zien dat Archimedes 2200 jaar geleden meer wist van wis- en natuurkunde dan Isaac Newton in de zeventiende eeuw. Hij kende de waarde van pi, ontwikkelde de theorie van het soortelijk gewicht en zette de eerste stappen op weg naar de wiskundige analyse. Alles wat we van Archimedes weten is gebaseerd op drie manuscripten, waarvan er twee verloren zijn gegaan. Het derde is een zogenaamde palimpsest, een codex waarvan de tekst is weggeschraapt om het perkament opnieuw te gebruiken, in dit geval als gebedenboek. William Noel, een handschriftkundige, en Reviel Netz, een wiskundig historicus, vertellen het verhaal van het gebedenboek van 1229 tot nu, laten zien hoe Archimedes' waardevolle tekst tevoorschijn kwam en leggen uit waarom de tekst zo belangrijk is. In hun boek combineren ze een adembenemend plot, verloren boeken en wetenschappelijke speurzin, en doen zo een aantal ontdekkingen die de geschiedenis van wetenschap en wiskunde voor altijd zullen veranderen. De uitgave van *De Archimedes-codex* is daarmee een belangrijke gebeurtenis voor de wetenschappelijke wereld en een fascinerende leeservaring voor iedereen die houdt van spannende, waargebeurde verhalen.

De website *www.archimedespalimpsest.org* bevat een schat aan informatie over de Archimedes-codex voor wetenschappers en leken.

Ondertitel: De geheimen van een opzienbarende palimpsest ontsluierd
Auteurs: Reviel Netz, William Noel
Vertaling: Boukje Verheij
Uitgever: Athenaeum-Polak & Van Gennep, Amsterdam (2007)
ISBN 978 90 253 6322 2
Prijs: € 19,95 (320 pag.)

Aankondiging / Statistiekwedstrijd

Het Internationale Statistiek Geletterdheid Project (ISLP) van de Internationale Vereniging voor Statistiekonderwijs (IASE) organiseert een wedstrijd statistiek voor leerlingen van 10-18 jaar.

De einddatum voor registratie is 28 februari 2008.

Meer informatie over werking en deelname is te vinden op:
www.stat.auckland.ac.nz/~iase/islp/competition
Oefenmateriaal is eveneens beschikbaar. Aan deze wedstrijd kan in de eigen taal worden deelgenomen.

Registratieformulieren in andere talen dan in het Engels kunnen verkregen worden via een e-mailbericht aan de ISLP director, Juana Sanchez (jsanchez@stat.ucla.edu).

Er wordt uitgekeken naar deelnemers uit België en Nederland!

Kruis of munt en de gulden snede

[Rob Bosch]

We gooien een aantal keren met een munt en constateren dat in deze serie worpen geen enkele keer twee maal achtereen kruis gegooid is. Vertelt deze wetenschap ons iets over de eerste worp? Met andere woorden, hoe groot is met deze wetenschap de kans op munt of kruis bij de eerste worp? Maakt daarbij de lengte van de serie nog verschil? We vragen hier de kans op munt bij de eerste worp onder de voorwaarde dat er in de reeks geen dubbel kruis - twee maal kruis in opeenvolgende worpen - is opgetreden. Deze voorwaardelijke kans schrijven we als $P(A \mid B)$, waarbij A de gebeurtenis is dat *bij de eerste worp munt boven* komt en B de gebeurtenis is dat *in de reeks geen dubbel kruis* optreedt. Voor de voorwaardelijke kans geldt:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad (1)$$

Om deze kans te kunnen uitrekenen moeten we weten hoeveel series er zijn zonder dubbel kruis (B) en vervolgens hoeveel van deze series met munt beginnen ($A \cap B$) [1]. Voor kleine reeksen staan de mogelijke rijtjes van kruis en munt in de onderstaande tabel waarbij n het aantal worpen in een serie is.

worpen	rijtjes zonder dubbel kruis				| B |	| A ∩ B |
$n = 1$	**m**	k			2	1
$n = 2$	**mm**	**mk**	km		3	2
$n = 3$	**mmm**	kmm	**mkm**	**mmk**		
	kmk				5	3
$n = 4$	**mmmm**	kmmm	**mkmm**	**mmkm**		
	mmmk	kmkm	kmmk	**mkmk**	8	5
$n = 5$	**mmmmm**	kmmmm	**mkmmm**	**mmkmm**		
	mmmkm	**mmmmk**	kmkmm	kmmkm		
	kmmmk	**mkmkm**	**mkmmk**	**mmkmk**		
	kmkmk				13	8

Met behulp van deze tabel kunnen we de kans uitrekenen voor een serie van 5 worpen. Het totaal aantal mogelijke series van 5 worpen is uiteraard $2^5 = 32$, alle met gelijke kans $\frac{1}{32}$. Voor $P(B)$ vinden we $\frac{13}{32}$, terwijl de kans $P(A \cap B)$ gelijk is aan $\frac{8}{32}$, zodat

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{8}{13}$$

De lezer heeft in de tabel ongetwijfeld een bekend patroon ontdekt; jawel *Fibonacci*.

Zowel het aantal rijtjes zonder dubbel kruis als de rijtjes die met munt beginnen, lijken Fibonacci-getallen te zijn. Dat is inderdaad het geval zoals het onderstaande argument aantoont.

Laat $R(n)$ het aantal rijtjes zonder dubbel kruis zijn bij een serie van n worpen. Een dergelijk rijtje eindigt met kruis of munt. Als zo'n rijtje op munt eindigt, dan vormen de eerste $(n-1)$ worpen een rijtje zonder dubbel kruis. Het aantal van zulke rijtjes is $R(n-1)$.

Als de serie van n worpen op kruis eindigt, dan moet de voorlaatste worp munt zijn. De laatste twee worpen liggen dus vast. In dit geval vormen de eerste $(n-2)$ worpen een rijtje zonder dubbel kruis. Daarvan zijn er $R(n-2)$. Voor het aantal rijtjes $R(n)$ geldt dus:

$R(n) = R(n-1) + R(n-2)$

Dit is de bekende Fibonacci-relatie. Uit de tabel lezen we de beginvoorwaarden $R(1) = 2$ en $R(2) = 3$ af.

De Fibonacci-rij $F(n)$ is 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89…

Voor de getallen $R(n)$ geldt $R(1) = F(3)$, $R(2) = F(4)$ enz. Kortom:

$R(n) = F(n+2)$

Het aantal rijtjes $M(n)$ van n worpen zonder kruis dat met munt begint, is nu niet moeilijk meer. Als het rijtje met munt begint, dan vormen de worpen 2 tot en met n een rijtje zonder dubbel kruis van lengte $(n-1)$. Voor $M(n)$ vinden we:

$M(n) = R(n-1) = F(n+1)$

Voor de (voorwaardelijke) kans P dat een serie van n worpen met munt begint als we weten dat we in die serie geen enkele keer tweemaal achtereen kruis hebben gegooid, geldt dat:

$$P = \frac{F(n+1)}{F(n+2)}$$

In de onderstaande tabel zijn de kansen weergegeven voor enkele waarden van n.

n	P	≈
1	$\frac{1}{2}$	0,5000
2	$\frac{2}{3}$	0,6667
3	$\frac{3}{5}$	0,6000
4	$\frac{5}{8}$	0,6250
5	$\frac{8}{13}$	0,6154
10	$\frac{89}{144}$	0,6181
14	$\frac{610}{987}$	0,6180

We zien dat vanaf $n = 5$ de kans nauwelijks meer verandert. De kans convergeert blijkbaar naar een getal in de buurt van 0,6180. De ratio's van opeenvolgende Fibonacci-getallen convergeren naar de *gulden snede* φ. Dat wil zeggen: $\lim_{n\to\infty} \frac{F(n+1)}{F(n)} = \varphi$.

Voor onze kans vinden we dus:

$$\lim_{n\to\infty} \frac{F(n+1)}{F(n+2)} = \frac{1}{\varphi} \qquad (2)$$

Voor φ geldt: $\varphi = \frac{1+\sqrt{5}}{2} \approx 1{,}6180339887$

Daaruit volgt: $\frac{1}{\varphi} = \frac{\sqrt{5}-1}{2} \approx 0{,}6180339887$

Omdat $\frac{1}{\varphi} = \varphi - 1$ is, hadden we (2) ook kunnen schrijven als:

$$\lim_{n\to\infty} \frac{F(n+1)}{F(n+2)} = \varphi - 1 \qquad (3)$$

De afloop van het sommetje was van te voren moeilijk direct te zien. Maar ja, de Fibonacci-getallen en de gulden snede duiken binnen en buiten de wiskunde op de meest onverwachte plaatsen op.

Noot (red.)

[1] De gebeurtenissen A en B treden tegelijk op. Spreek uit als: A én B.

Over de auteur

Rob Bosch is redacteur van *Euclides* en universitair hoofddocent wiskunde aan de Nederlandse Defensie Academie te Breda. E-mailadres: *robbosch@live.nl*

VU-Statistiek en gemeentepolitiek

[Cees de Hoog en Bert van der Windt]

In 's-Gravenzande op een vmbo-locatie houden leerlingen zich gedurende de maanden april en mei bezig met een vakoverstijgend project. Het doel van dit project is tweeledig: ten eerste willen we dat leerlingen enig inzicht krijgen in de plaatselijke politiek en ten tweede moeten leerlingen 25 mensen enquêteren. Zij hebben hiervoor zelf een enquête gemaakt en verwerken de gegevens met behulp van het computerprogramma VU-Statistiek.

figuur 1 Klas 3 in de raadzaal van 's-Gravenzande

figuur 2 De verwerking van de enquête

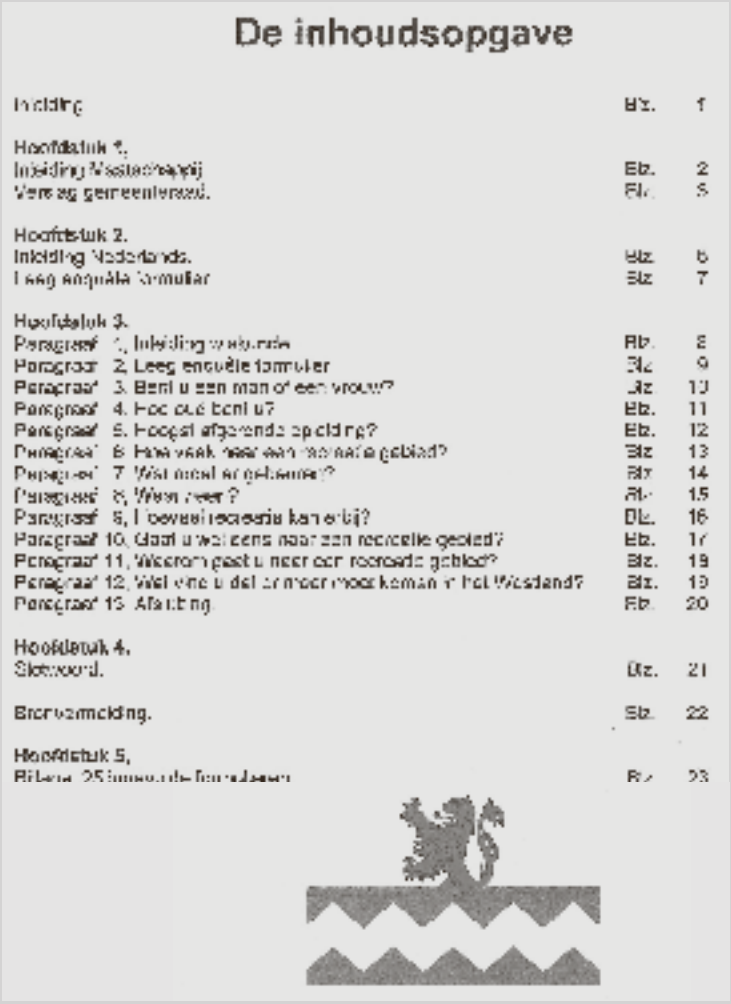
Twee praktijkopdrachten

Voorheen bezochten elk jaar alle leerlingen van 3 vmbo-tl de raadzaal in 's-Gravenzande (Gemeente Westland). In een aantal lessen maatschappijleer werd een debat voorbereid. De leerlingen verdiepten zich in de lokale politiek en lokale vraagstukken. Zij bestudeerden verschillende lokale kranten en de site van de gemeente. Tijdens het debat zaten zij letterlijk op de stoel van een raadslid. Ook stelden de leerlingen via een e-mail een vraag over het bestudeerde onderwerp aan het desbetreffende raadslid. Veel raadsleden namen hun maatschappelijke verantwoordelijkheid en beantwoordden de vragen serieus. Bij de debatten was een aantal raadsleden betrokken. Zowel de politici als de leerlingen zagen de debatten als zinvol, verrassend en inspirerend.

Bij wiskunde wordt al zeven jaar een werkstuk gemaakt met behulp van het veelgebruikte computerprogramma VU-Statistiek, een statistisch verwerkingsprogramma met heel veel mogelijkheden. In bijna elk wiskundeschoolboek wordt optioneel gebruik gemaakt van dit programma; zeker de moeite waard om er eens mee te werken. In de eerste jaren was de keuze van het onderwerp voor het werkstuk geheel vrij. Maar na veel werkstukken over mobiele telefoons, scooters, computers, auto's en muziek hebben we de onderwerpen veranderd. Twee jaar hebben we het onderwerp laten sporen met de sector waarin men examen ging doen. De laatste jaren hebben we onderwerpen met meer diepgang gekozen, onderwerpen die meer maatschappelijk gerelateerd zijn, zoals 'leven met kanker', 'leven met een handicap' en 'leven na de dood'. Wij verzekeren u dat gesprekken met leerlingen over dit soort onderwerpen een nieuwe dimensie toevoegt aan het beroep wiskundeleraar.

Drie vakken, één werkstuk

Bovenstaande praktijkopdrachten hebben we vorig schooljaar gecombineerd tot één project. De docenten maatschappijleer zochten een manier om het project 'lokale politiek' uit te breiden, de docenten wiskunde zochten interessante onderwerpen voor het maken van een werkstuk met behulp van VU-Statistiek en de docenten Nederlands wilden graag meedoen omdat het maken van een goede enquête een vaardigheid is die zij de leerlingen willen bijbrengen.
Na een korte brainstormsessie besloten de begeleidende docenten een lijst met onderwerpen te maken. De onderwerpen moesten een lokaal, politiek tintje hebben en maatschappelijk relevant zijn. Bij het vak maatschappijleer spreekt men van een maatschappelijk probleem als (1) er veel mensen bij betrokken zijn, (2) er verschillende meningen over bestaan en (3) de politiek zich ermee bemoeit. Binnen een paar dagen stonden 60 titels op papier. Een aantal voorbeelden: hokken, megadisco, verkeersdrempels, maatschappelijke stage, het nieuwe winkelcentrum in 's-Gravenzande, afschaffen van het zwemonderwijs in 's-Gravenzande, overlast hangjongeren, straatvervuiling, openingstijden van de winkels, veiligheid op straat, drankgebruik onder jongeren, recreatiemogelijkheden. We besloten de leerlingen een onderwerp toe te kennen, omdat een paar onderwerpen (hangjongeren, drankgebruik, hokken, megadisco) erg populair zijn en andere thema's wellicht niet kunnen rekenen op de warme belangstelling van onze leerlingen (het nieuwe winkelcentrum, maatschappelijke stage, afschaffen van schoolzwemmen).



De inhoudsopgave

Inleiding Blz. 1

Hoofdstuk 1.
Inleiding Westerschelpij Blz. 2
Verslag gemeenteraad. Blz. 5

Hoofdstuk 2.
Inleiding Nederlands. Blz. 6
Leeg enquête formulier Blz. 7

Hoofdstuk 3.
Paragraaf 1. Inleiding onderzoek Blz. 8
Paragraaf 2. Leeg enquête formulier Blz. 9
Paragraaf 3. Bent u een man of een vrouw? Blz. 10
Paragraaf 4. Hoe oud bent u? Blz. 11
Paragraaf 5. Hoogst afgeronde opleiding? Blz. 12
Paragraaf 6. Hoe vaak naar een recreatie gebied? Blz. 13
Paragraaf 7. Wat moet er gebeuren? Blz. 14
Paragraaf 8. Waar naar ? Blz. 15
Paragraaf 9. Bevoegdheden kamerlid? Blz. 16
Paragraaf 10. Gaat u wel eens naar een recreatie gebied? Blz. 17
Paragraaf 11. Waarom gaat u naar een recreatie gebied? Blz. 19
Paragraaf 12. Wat vind u dat er moet komen in het Westland? Blz. 19
Paragraaf 13. Afsluiting Blz. 20

Hoofdstuk 4.
Slotwoord. Blz. 21

Bronvermelding. Blz. 22

Hoofdstuk 5.
Bijlage 25 [illegible] Blz. 23

De inleiding.

[illegible]

1. Hoe vaak per jaar gaat u naar een bos of ander natuurgebied in het Westland?
   Bij deze vraag wil ik achter komen of de mensen nog wel vaak naar een natuur gebied gaan. Zo kan je zien of de mensen er nog wel behoefte aan hebben.

2. Als er te weinig natuur en recreatie in het Westland komt. Wat vindt u dat er dan moet gebeuren?
   [illegible]

3. Waar gaat u per jaar wel een keer of 3 heen?
   Of deze vraag wo. Ik duidelijk maken waar het meeste behoefte naar is. Deze antwoorden kunnen worden [illegible] te kunnen zien, dit bij je kaart leggen. Hier kan je de recreatie [illegible] van de bijv. te kunnen zien waar het vaak open dan naar een [illegible] gaat.

4. Hoeveel recreatie gebieden kunnen er in het Westland bij?
   Hier wou ik komen naar hoeveel mensen gebieden er nog bij kunnen. Is er teveel of is er nog te weinig recreatie in het Westland. Oftewel is het goed gesteld met de recreatie in het Westland of niet?

6. Gaat u wel eens (1 of 2 keer per maand) naar een recreatie gebied?
   Bij deze vraag wou ik kijken waarom de mensen naar een recreatie gebied gaan of juist niet. Hier kan je ook weer een [illegible] de mensen echt behoefte aan recreatie hebben.

8. Als u naar een natuur- of recreatiegebied gaat, waarom gaat u daar dan heen?
   Deze vraag, ik kan beter [illegible] ik wou. Ik ben benieuwd de mensen naar een recreatie gebied gaan. Ik ga ervan uit dat het nog een beetje dit gaat of niet. Bij deze vraag wordt er weer een [illegible] gekeken. Waarom ga je er heen of wat doe je er voor.

7. De recreatie begint te verdwijnen, wat vindt u dat er meer in het Westland moet komen?
   Deze vraag zegt veel over de hoofdvraag. Iedereen weet dat er dingen in het Westland gaan verdwijnen zoals de [illegible]. Bij deze vraag ga ik kijken wat de mensen nog graag willen hebben in het Westland.

Verloop

De docente maatschappijleer bereidde het debat in de raadzaal voor. Iedereen kreeg een rol als raadslid. De leerlingen kregen de opdracht het toegewezen onderwerp als een maatschappelijk probleem te beschrijven. Verder formuleerden ze een hoofdvraag en stelden over hun onderwerp een vraag aan een raadslid. Voorbeelden van hoofdvragen:

- Hoe is het drankgebruik van jongeren in het Westland?
- Hoe is het gesteld met natuur en recreatie in het Westland?
- Hoe veilig is het Westland?
- Hoe denken mensen over zorg en welzijn in het Westland? en
- Heeft het Westland behoefte aan een megadisco?

Half april volgden er in de raadzaal vijf interessante debatten. De raadsleden waren verheugd over de betrokkenheid van de leerlingen en de leerlingen vonden het best spannend een politiek debat te voeren. Dit alles resulteerde in goede verslagen van het debat.

In dezelfde week kregen de leerlingen bij het vak Nederlands de opdracht een enquête te maken. Naast drie verplichte vragen moesten ze zeven deelvragen over hun onderwerp formuleren. De zelfverzonnen vragen moesten uiteindelijk leiden tot een antwoord op de hoofdvraag. De eerste drie verplichte vragen gaan over het geslacht, de leeftijd en de hoogst afgeronde opleiding van de geënquêteerde. De eisen die aan de vragen gesteld werden, zijn mede bepaald door het programma VU-Statistiek. Er zijn drie soorten vragen mogelijk: een open vraag waarbij het antwoord een getal moet zijn, een meerkeuzevraag (de respondent mag één antwoord aangeven) en een multipuntvraag (de geënquêteerde mag meer dan één antwoord aankruisen). Deze drie soorten vragen moesten allemaal voorkomen. De leerlingen formuleerden de vragen en de antwoordmogelijkheden en mailden hun ontwerp naar hun docenten Nederlands. Deze beoordeelden het enquêteformulier (vragen, lay-out, antwoordmogelijkheden), corrigeerden vervolgens de vragen en gaven adviezen. Nadat de enquête aangepast was, moest deze worden doorgestuurd naar de docent wiskunde. Deze bekeek de vragen door een 'VU-Statistiek-bril'. Ook dit gebeurde geheel per e-mail. Daarna probeerden de leerlingen de vragenlijst uit. Als het nodig was werd de vragenlijst nog aan-

gepast. Elke leerling liet daarna de vragenlijst door 25 mensen invullen, 5 personen uit elk van de volgende leeftijdsgroepen: 10-20 jaar, 20-30 jaar, 30-40 jaar en 50 jaar en ouder. Bovendien moest de verdeling man/vrouw ongeveer gelijk zijn.

Het werkstuk

Inmiddels was het mei. Het wiskundeboek voor klas 3 was uit en het was tijd voor het maken van het werkstuk. Eerst werkten de leerlingen een lessenserie door over het programma VU-Statistiek. Ze leerden cirkel- en staafdiagrammen maken. Ze leerden werken met klassengrenzen, procenten en verschillende manieren om een diagram op te maken. Verder werd aandacht besteed aan het splitsen van diagrammen. Het splitsen gebeurde op grond van leeftijd, opleiding en geslacht.
Er werd ook aandacht besteed aan het trekken van conclusies. Vervolgens bekeken we de opties dataplot, kruistabel, puntenwolk, centrummaten en boxplot. De leerlingen leerden het exporteren van diagrammen en tabellen naar Word. De lessenserie ging vanzelf over in het maken van het werkstuk. In de laatste opdrachten gingen de leerlingen aan de slag met het inbrengen en labelen van de verplichte vragen. De geschreven lessenserie hield op bij de opdracht, de zelf gemaakte vragen en antwoordmogelijkheden in de computer te zetten.
Nadat de tien vragen ingevoerd en gelabeld waren, brachten de leerlingen de antwoorden van de 25 enquêteformulieren in. Natuurlijk is een totaal van 25 respondenten niet genoeg, maar met 4000 formulieren is de werkwijze hetzelfde, en daar gaat het om.
Toen alle vragen en antwoorden in de computer stonden, kon men aan het werkstuk beginnen. De opdracht was bij elke vraag ten minste één diagram of tabel te maken. Het moesten plaatjes zijn waarover iets te vertellen viel. Leerlingen vonden het moeilijk in te zien dat een conclusie iets anders is dan het vertalen van het plaatje in woorden. Als leerlingen de verwerking vaak splitsten, konden er interessante verwerkingen ontstaan. Als een leerling de verwerking splitst op man/vrouw, dan kan de leerling zien of mannen er anders over denken dan vrouwen. De praktijk leerde dat leerlingen liever tijd stopten in het mooi opmaken van het diagram, dan in het schrijven van goede conclusies.

Aan de verwerking worden eisen gesteld

In het werkstuk moesten de volgende verwerkingen voorkomen: een gesplitst cirkeldiagram; een gesplitst staafdiagram, procenten, centrummaten, frequentietabel met procenten, cumulatief diagram. Naast deze verplichte verwerkingen waren er nog andere interessante mogelijkheden. Kruistabel, puntenwolk en boxplot leidden soms tot heel interessante plaatjes met goede conclusies. Vaak was hulp van ons daarbij wel nodig. Maar dat hoort in een leerproces. Sommige leerlingen moest duidelijk gemaakt worden dat ze geen tien verschillende plaatjes, maar tien *goede* plaatjes moesten maken.
Nadat de kern van het werkstuk klaar was, begon voor sommige leerlingen het moeilijkste deel. Er moesten een inleiding, een afsluiting voor het wiskundedeel en een slotwoord over het gehele project geschreven worden. Vooral verwoorden wat ze geleerd hadden op wiskundig gebied, maar zeker ook op het gebied van hun eigen vaardigheden, was lastig. Sommige leerlingen werkten voor het eerst met paginanummering, anderen leerden een paginanummer te onderdrukken. Een aantal leerlingen werkte voor het eerst met 'tab' en 'ctrl-enter' om op een nieuwe pagina te beginnen (in plaats van een pagina vol met enters). Het schrijven van een samenvatting van het werkstuk, een antwoord op een hoofdvraag geven en het terugkijken op het eigen leerproces (reflectie) was voor de meeste leerlingen nieuw.

Voordelen van dit project

De betrokken docenten constateerden een aantal duidelijke voordelen. Zo schrijft elke leerling een uniek werkstuk. Er is geen enkel onderdeel gekopieerd. Van de eerste tot de laatste letter is het door de leerling zelf bedacht (geen 'ctrl-C 'gevolgd door 'ctrl-V'). Elke leerling in de klas schrijft over een ander onderwerp. Men kan elkaar op technisch gebied assisteren, maar het trekken van conclusies kan alleen gedaan worden door de leerling zelf. Je kunt elk jaar de onderwerpen veranderen. De kans op 'geleende' kennis is klein. Samenwerking met collega's uit andere vakgebieden is een verrijking en de leerlingen maken in totaal minder werkstukken. Het contact met de leerling is door het gebruik van e-mail laagdrempelig. Dit werkstuk wordt geheel op school gemaakt (als de leerling doorwerkt). Bijkomend voordeel aan het eind van het jaar: geen 'huiswerk' voor wiskunde. Kennis wordt functioneel en de opdracht is zowel praktisch als theoretisch van aard.

Praktische zaken

De periode april/mei is gekozen omdat we dan makkelijker kunnen beschikken over de lokalen met computers. Bij ons op school is het mogelijk de mediatheek te reserveren voor alle wiskundelessen in de derde klassen. Het werkstuk is opgenomen in het PTA (programma van toetsing en afsluiting) van klas 3 voor wiskunde, Nederlands en maatschappijleer en telt dus mee voor het examen. Na de contacten via e-mail kost het doorwerken van de lessenserie over VU-Statistiek en het maken van het werkstuk zo'n tien wiskundelessen, ongeveer de benodigde tijd voor een hoofdstuk. Wel moeten we opmerken dat het maken en afnemen van de enquête op tijd moet gebeuren. Wij doen dit terwijl wij het laatste hoofdstuk uit het boek behandelen.

Tot slot

Terugkijkend kunnen wij concluderen dat de aanpak van een statistisch onderzoek in combinatie met twee andere vakken zeer geslaagd is geweest. Statistiek leent zich bijzonder goed voor vakoverstijgende projecten en zou een mogelijk onderdeel kunnen vormen voor een sectorwerkstuk in klas 4. Leerlingen zeggen zelf ook trots te zijn op het geleverde werk en geven aan dat er werkelijk iets geleerd is in plaats van

Hoofdvraag en deelvragen:
De hoofdvraag die ik heb bedacht, is: *Hoe bekend is Stip Westland in het Westland?* De deelvragen (die ik heb verwerkt in paragrafen en grafieken / tabellen) die ik heb bedacht om dit werkstuk te kunnen maken en de antwoorden daarop zijn:

- *Welk geslacht hebben de meeste mensen die ik heb geënquêteerd?*
  De meeste mensen die ik heb geënquêteerd, zijn vrouw, namelijk 52 % (13 vrouwen).
- *Welke leeftijd hebben de meeste mensen die ik heb geënquêteerd?*
  De meeste mensen die ondervraagd zijn, zijn tussen de 20 en 39 jaar.
- *Hebben de mensen die ik heb geënquêteerd een goede opleiding?*
  De meeste mensen hebben een goede opleiding.
- *Wat denken mensen dat Stip Westland is?*
  Het merendeel (mannen en vrouwen samen) denkt dat Stip Westland 'Een loket voor hulp bij uw woning, de zorg en welzijn' is (aantal enquêtes: 25; aantal mannen en vrouwen die kiezen voor het antwoord 'Een loket voor hulp bij uw woning, de zorg en welzijn': 13; 25-13 = 12).
- *Vinden de mensen die ik heb geënquêteerd Stip Westland goed bereikbaar?*
  De meeste van de ondervraagden vinden Stip Westland niet zo goed bereikbaar.
- *Waar moet Stip Westland volgens de mensen die ik heb geënquêteerd meer aandacht aan besteden?*
  In totaal hebben de meeste mensen geen mening over dit onderwerp.
- *Hoeveel procent van de bevolking maakt volgens Westlanders gebruik van Stip Westland?*
  De meeste mensen denken dat maar weinig procent van de bevolking gebruik maakt van Stip Westland.
- *Wat vinden de meeste mensen die ik heb geënquêteerd van de openingstijden van Stip Westland?*
  De meeste mensen vinden de openingstijden niet goed. Ze willen zowel 's avonds als 's middags dat Stip Westland ook open is.
- *Waarom hebben de meeste mensen die ik heb geënquêteerd gebruik gemaakt van Stip Westland, en hebben ze wel eens gebruik gemaakt van Stip Westland?*
  De meeste mensen die ik heb ondervraagd hebben geen gebruik gemaakt van Stip Westland. Van de mensen die ik heb geënquêteerd die wel gebruik gemaakt hebben van Stip Westland, hebben de meeste mensen gebruik gemaakt van de huishoudelijke hulp.
- *Welk cijfer geven de mensen die ik heb geënquêteerd gemiddeld aan Stip Westland?*
  De meeste mensen die ik heb ondervraagd geven een 6.4 als cijfer aan Stip Westland ((6,2 + 6,5) : 2 = 6,35 = 6,4).

een werkstuk te produceren met behulp van knippen en plakken van internet.

Wij vinden het jammer dat in de komende jaren het onderwerp statistiek niet meer getoetst wordt in het centrale examen. Wij betreuren dit zeer, omdat in de vervolgopleidingen méér vmbo-leerlingen te maken krijgen met cirkeldiagrammen, procenten en beelddiagrammen, dan met de stelling van Pythagoras, goniometrie of gelijkvormigheid. Een deel van de geletterdheid in Nederland bestaat toch uit het kunnen lezen en interpreteren van tabellen en diagrammen?

Over de auteurs

Bert van der Windt en Cees de Hoog werken bij de scholengroep I.S.W. (Locatie Sweelincklaan in 's-Gravenzande, Gemeente Westland). De locatie kent een brede onderbouw vmbo-bl, -kl, -tl, een havo-startklas en een bovenbouw vmbo-tl. Bert en Cees werken voornamelijk in de bovenbouw, maar zijn in de onderbouw actief betrokken bij het ontwikkelen van practicumlessen en projecten (in lengte variërend van één tot tien lessen).

Bert van der Windt werkt nu zo'n 3 jaar in het vmbo. Daarvóór heeft hij 10 jaar gewerkt in de onderbouw van havo en vwo. De overstap naar het vmbo was niet eenvoudig, maar voor zijn didactische en pedagogische ontwikkeling is het zeer goed geweest. Naast het lesgeven in wiskunde ligt zijn interesse in het ontwikkelen van vakoverstijgende projecten.
E-mailadres: *biw@caiway.nl*

Cees de Hoog werkt 30 jaar in het vmbo. Naast lesgeven houdt hij zich bezig met het ontwikkelen en toepassen van practicumopdrachten en spelletjes. Hij heeft in de afgelopen jaren de ICT-Testbeelden bij de methode *Moderne wiskunde* geschreven en is nu bij deze methode betrokken als auteur.
E-mailadres: *ceesdehoog@kabelfoon.nl*

Voetbalpoule

Ons voetbalelftal speelt bepaald niet groots.
Vandaar dat het er straks ook niks van bakt,
Al had de loting anders uitgepakt:
Voor ons is elke poule een poule des doods.

En het gemodder voor het eigen doel
Maakt deze poule ook tot een modderpoule.

Gepubliceerd door *Driek van Wissen* op 8-12-2007 op *www.gedichten.nl*

(*...Bij de loting voor het EK-voetbal is Nederland ingedeeld bij Italië, Frankrijk en Roemenië ...*)

Een meetkundige (on)waarschijnlijkheid

[Dick Klingens]

figuur 1

figuur 2

figuur 3

figuur 4

figuur 5

0.

Gegeven is een vaste cirkel met straal r. Hoe groot is de kans dat de lengte van een *willekeurig* getrokken koorde in die cirkel kleiner is dan r?

We zullen laten zien, dat de gevraagde kans afhangt van de manier waarop de koorde getekend wordt.

1.

Zij AB de koorde. We kiezen eerst het punt A op de cirkel en vervolgens het punt B. De keuze van A is niet aan voorwaarden gebonden, maar bij de keuze van B maken we de - overigens niet onlogische - afspraak dat, bij een verdeling van de cirkel in n gelijke delen, de kans dat B in zo'n deel terecht komt, voor alle delen gelijk is.

Zie figuur 1. We kiezen $n = 6$ en maken een verdeling van de cirkel waarbij het punt A één van de deelpunten is.

Voor $AB < r$ zijn er nu 2 gunstige mogelijkheden, immers B moet op de boog 2-1-6 gekozen worden.

De gevraagde kans is dus $P(AB < r) = \frac{2}{6} = \frac{1}{3}$.

In het algemene geval is de kans dat B op een boog met lengte s terecht komt gelijk aan $\frac{s}{2\pi r}$.

We kunnen een en ander ook in termen van hoeken formuleren. We kiezen het punt A weer willekeurig op de cirkel en daarna tekenen we een koorde AB die met de raaklijn t in A aan de cirkel een hoek maakt tussen 0° en 180° (***zie figuur 2***).

Voor zo'n hoek die kleiner is dan 30° of groter dan 150°, is de lengte van de koorde kleiner dan die van de straal van de cirkel.

Voor de gezocht kans $P(AB < r)$ vinden we daarmee ook:

$P(AB < r) = \frac{30+30}{180} = \frac{1}{3}$

2.

We kiezen het punt A nu niet *op* de cirkel, maar *erbinnen*. En we eisen dat A dan het midden is van een koorde BC die kleiner is dan r (***zie figuur 3***).

De afstand d van het middelpunt M van de cirkel tot een zijde van een in die cirkel ingeschreven regelmatige zeshoek is gelijk aan:

$d = r\cos 30° = \frac{1}{2}r\sqrt{3}$

waarmee d dan de lengte is van de straal van de ingeschreven cirkel van die zeshoek.

Kiezen we A buiten die incirkel, dan is BC kleiner dan r.

De kans kan in dit geval worden bepaald via de oppervlaktes van de beide cirkels:

$$P(BC < r) = \frac{opp(\text{ring})}{opp(\text{grootste cirkel})} = \frac{\pi r^2 - \pi \cdot (\frac{1}{2}r\sqrt{3})^2}{\pi r^2} = 1 - \frac{3}{4} = \frac{1}{4}$$

3.

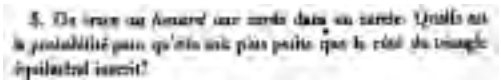
De getekende koorde is evenwijdig aan een zekere middellijn PQ van de cirkel. Het punt A ligt dan op de middellijn m die loodrecht op PQ staat (***zie figuur 4***). Zoals we bij punt 2 gezien hebben, is de afstand van M tot een zijde van een ingeschreven regelmatige zeshoek van de cirkel gelijk aan $d = \frac{1}{2}r\sqrt{3}$. De lijn m snijdt twee zijden van deze zeshoek in de punten P' en Q'.

Gunstig voor A is in dit geval een positie *buiten* het lijnstuk $P'Q'$, maar *binnen* de cirkel, zodat, en nu op basis van lengtes van lijnstukken:

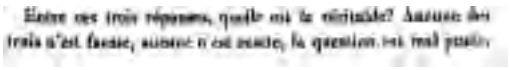
$$P(BC < r) = \frac{2r - 2\cdot\frac{1}{2}r\sqrt{3}}{2r} = 1 - \frac{1}{2}\sqrt{3} \ (\approx 0{,}13)$$

We vinden hiermee drie *juiste*, maar *verschillende* antwoorden op de in punt 0 gestelde vraag. De vraag die dán rest is natuurlijk hoe dat komt…

Tot slot



Hetgeen in de voorgaande paragrafen beschreven is, staat bekend als de *Paradox van Bertrand*. Joseph Louis Bertrand (1822-1890, Frankrijk) stelt in 1889 in zijn boek *Calcul des probabilités* het probleem in een iets andere vorm (zie [1; pp. 4-5] en ook ***figuur 5***):
Hoe groot is de kans dat de lengte van een willekeurige koorde van een cirkel kleiner is dan de lengte van een zijde van een in die cirkel ingeschreven gelijkzijdige driehoek?
Bertrand lost het door hem gestelde het probleem evenwel op door de kans te berekenen - hij rekent trouwens bijna niet - dat de koorde *langer* is dan de zijde van de driehoek en geeft daarbij als antwoorden: $\frac{1}{3}$, $\frac{1}{2}$ en $\frac{1}{4}$.



De lezer wordt aangemoedigd Bertrands antwoorden te verifiëren (wellicht worden daarbij dan nog andere mogelijke antwoorden gevonden…).
En, leerlingen confronteren met deze paradox mag natuurlijk ook!

Literatuur

[1] Joseph Bertrand (1889): *Calcul des probabilités*. Paris: Gauthier-Villars et Fils.
Het boek is digitaal beschikbaar via:
http://gallica.bnf.fr/ark:/12148/bpt6k99602b/

[2] En uiteraard via *Google* met als zoekwoorden: *bertrand paradox –russell* (let wel, er staat een minteken voor *russell* om te voorkomen dat een andere paradox gevonden wordt dan de bedoelde).

Over de auteur

Dick Klingens is als wiskundeleraar verbonden aan het Krimpenerwaard College te Krimpen aan den IJssel. Daarnaast is hij eindredacteur van *Euclides*.
E-mailadres: *dklingens@pandd.nl*

De boekenkast viel om

Over kansrekening en statistiek zijn natuurlijk heel wat studieboeken verschenen, maar ook op het populair-wetenschappelijke vlak is er het nodige te vinden. Hierbij een lijstje van wat aardige titels, zonder welke pretentie of samenhang dan ook. Sommige nog steeds rechtstreeks verkrijgbaar, andere misschien alleen tweedehands of in de ramsj.

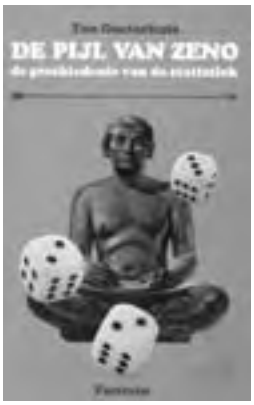
📖 Larry Gonick, Woollcott Smith: *Het Stripverhaal van de Statistiek*. Utrecht: Epsilon Uitgaven (nr. 32, 1996).
'Dit boek behandelt de kernzaken van de moderne statistiek […] uitgelegd met behulp van eenvoudige, heldere, en jawel, leuke illustraties.'

📖 Darrell Huff: *How to lie with statistics*. London: Pelican Books (1954).
Een grappige klassieker over elementaire statistiek en statistische misleiding: '… the crooks already know these tricks; honest men must learn them in self-defence.'

📖 J.S. Cramer, J. Hemelrijk: *Statistiek eenvoudig*. Amsterdam: Uitgeverij Nieuwezijds (1998).
'Zonder in te gaan op de techniek of formules, wordt de gedachtegang [van de moderne statistiek; red.] uitgelegd en aan de hand van actuele voorbeelden gedemonstreerd.'

📖 Peter Sprent: *Risico's*. Haarlem: Aramith Uitgevers (Nederlandse vertaling, 1990).
'Of we nu op paarden wedden, erop gokken dat de trein te laat is, of de gezondheid van onze kleinkinderen bedreigen door spuitbussen te gebruiken, risico's kunnen we op geen enkele manier vermijden. Maar hoewel de meeste mensen voor de hand liggende gevaren uit de weg gaan, variëren onze reacties sterk, omdat zij gewoonlijk door emoties bepaald worden. Duizenden mensen zijn doodsbang om te vliegen, hoewel dagelijks meer voetgangers sterven dan vliegtuigpassagiers.'

📖 J.H. van der Vlis, E.R. Heemstra: *Geschiedenis van kansrekening en statistiek*. Utrecht: Matrijs / Rijswijk: Pandata (1989).
Het boek behandelt de ontwikkeling van de statistiek aan de hand van beschrijvingen van belangrijke personen uit het vak, van hun werk, en van de invloed die zij uitoefenden. Ook technische aspecten van nieuwe methoden en ideeën worden beschreven, en met voorbeelden toegelicht.



📖 Ton Oosterhuis: *De Pijl van Zeno; de geschiedenis van de statistiek*. Baarn: Uitgeverij De Fontein (1991).
'In dit boekje wordt de geschiedenis van de statistiek geplaatst tegen de achtergrond van de maatschappelijke ontwikkelingen, de denkwerelden van het verleden, de problemen waar de wetenschappers mee worstelden.'

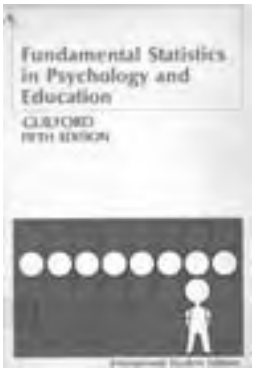
📖 James Burke: *Hoe groot is de kans dat…; risico's en kansen in het alledaagse leven*. Den Haag: BZZTôH (1992).
Inhoud:
- Hoe groot is de kans dat mijn vliegtuig neerstort? (1 op 600.000, even groot als de kans om door de bliksem te worden getroffen.)
- Welke honden bijten het meest? (Duitse Herders)
- Maak je minder kans een wedstrijd te winnen na seks op de avond ervoor? (Nee)

Deze en vele andere vragen worden behandeld in dit boekje. De auteur goochelt met cijfers en neemt een loopje met de normale wijze waarop cijfers en statistieken gebruikt worden. Daarbij komt hij tot verbluffende en soms absurde conclusies.

📖 John Allen Paulos: *Ongecijferdheid*. Amsterdam: Uitgeverij Bert Bakker (1989).
'Mijn benadering is steeds licht wiskundig en maakt gebruik van enkele elementaire begrippen uit de waarschijnlijkheidsrekening en de statistiek die, hoewel diepzinnig, niet meer van de lezer verlangen dan gezond verstand en wat rekenwerk.'

📖 Hans van Maanen: *Zoete koek & speculatie*. Amsterdam: Uitgeverij Contact (2004).
Medische onderzoeksresultaten krijgen in de media veel aandacht. Wetenschapsjournalist Van Maanen kijkt met kritische blik naar de gehanteerde onderzoeksmethoden, en stelt met name onjuist gebruik van elementaire statistiek aan de kaak. 'Kaneel is zelfs niet verdacht, koffie is altijd de gebeten hond, en dat proefschrift over stress en verkoudheid is al helemaal dubieus. Enkele langere artikelen, onder meer over de abominabele voorlichting bij het bevolkingsonderzoek naar borstkanker, over de Franse paradox en natuurlijk over homeopathie, completeren deze nieuwe bundel over de "rafelranden van de wetenschap".'

📖 Henk Tijms: *Spelen met kansen.* Utrecht: Epsilon Uitgaven (nr. 43, 2e herziene en vermeerderde druk, 2002).
'Dit boek laat zien hoe rijk aan toepassingen de kansrekening is. Het bevat een groot aantal voorbeelden, waarvan vele op het gebied van casinospelen en loterijen. Zo wordt de lezer spelenderwijs vertrouwd gemaakt met de grondbegrippen van de kansrekening.'



📖 J.P. Guilford, B. Fruchter (1897): *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill (5e editie, herdruk 1973).
Bevat toegepaste statistiek op de terreinen van de psychologie en het onderwijs, bedoeld voor studenten in die vakken. Zoals de auteurs zelf schrijven, is het boek meer dan een 'kookboek'. Er is veel aandacht voor toepassingen van de verschillende statistische technieken.

📖 Bert Nijdam e.a.: *Statistiek en Kansrekening voor het V.W.O.* Utrecht: Instituut voor Ontwikkeling van het Wiskunde Onderwijs, IOWO (1973).
In 1976 werd de statistiek definitief opgenomen in het examenprogramma Wiskunde I, zodat in 1974 alle leerlingen van het vijfde leerjaar les in statistiek kregen. Het boek is geschreven op basis van het experimentele (paarse) boekje *Waarschijnlijkheidsrekening en Statistiek voor het V.W.O.* van H. Freudenthal, B. Nijdam en J.J. Wouters, dat gebruikt is bij experimenten voorafgaande aan de invoering van het nieuwe vak en bij bijscholingscursussen voor leraren.
In het boek staat de statistiek centraal, terwijl alleen die delen van de kansrekening zijn opgenomen die voor een goed begrip van de statistiek noodzakelijk zijn.

📖 Hans Freudenthal (1957): *Waarschijnlijkheid en statistiek*. Haarlem: De Erven F. Bohn N.V.; deel 57 in de tweede reeks van de Volksuniversiteits Bibliotheek (3e druk, 1966).
Het boek is een verzameling van negen 'short stories', die - volgens de schrijver - fundamenteel zijn en interessant, maar het is geen roman. Het geheel leest ook niet als een roman (toch wel wat wiskunde). Binnen de tekst vinden we in elke paragraaf, meestal gekoppeld aan een voorbeeld, een aantal opgaven (de eerste: "De kans, om met drie dobbelstenen meer dan 10 te werpen. Reken die uit, maar doe het handig!").

Diagrammen in het basisonderwijs

[Maike den Houting]

Deeltijdstudenten in een kort onderzoek naar gebruik en inzet van diagrammen in de verschillende basisschool-rekenmethodes en de verschillende groepen van de bovenbouw

Gele en roze balletjes

Mijn jongste zoon Merijn kreeg van Sinterklaas in groep 2 een dartboard van klittenband met twee bijbehorende balletjes, een roze en een gele. Niet elk kind had dezelfde kleuren, sommigen hadden twee roze of juist twee gele balletjes gekregen. 'Mam', zei hij bij thuiskomst, 'er zijn eigenlijk vier soorten cadeautjes.' Ik vroeg wat hij bedoelde, dus hij legde me uit: 'Je kunt twee roze balletjes krijgen, twee gele, een gele en een roze of een roze en een gele.' Verbaasd keek ik hem aan en prees hem vervolgens de hemel in. Ik werkte in die periode als wiskundedocent op een middelbare school en was juist bezig met combinatoriek. Misschien kwam het daardoor wel, dat de zaken thuis me meer opvielen. Korte tijd later had mijn oudste zoon Joran (op dat moment 9 jaar) namelijk net zo'n briljante ingeving toen hij met de elektrische gitaar van zijn vader om zijn schouder riep: 'De gitaar kan in 500 verschillende standen.' Inderdaad zaten er twee knopjes op die van 1 tot 10 gingen en een knopje met 1 tot 5. Het zette me aan het nadenken over de manier waarop ik met zoiets als wegendiagrammen bezig was met die 13/14-jarigen op de middelbare school. Nu ik als docent vakdidactiek rekenen op de pabo werkzaam ben, heb ik me meer kunnen verdiepen in wat basisschoolkinderen aan wiskunde doen. Voor die tijd vond ik de rekenboeken van de basisschool nogal onbegrijpelijk. Onderwerpen dwars door elkaar en nergens theorieblokken. Hoe leren deze kinderen dan de belangrijke concepten?

Wiskundetaal

Kerndoel 23 van het rekenonderwijs op de basisschool is dat kinderen wiskunde-taal leren gebruiken. Een nogal globale omschrijving van wat een kind moet kunnen aan het eind van groep 8. Wat houdt het gebruiken van wiskundetaal in en op welk niveau moeten kinderen die taal beheersen? Dat ligt niet vast. Redeneren over combinaties, het interpreteren en maken van diagrammen en tabellen zijn slechts een aantal voorbeelden van wiskundetaal. De SLO heeft geprobeerd dit kerndoel wat handen en voeten te geven door het beschrijven van meer concrete tussendoelen en leerlijnen, maar ook dat zijn slechts handreikingen voor de leerkracht [1].
Welke taal stelt de SLO voor als het gaat om diagrammen? En in welke groep? Ik vond het volgende voorstel:
Groep 3/4: Taal voor het uitdrukken of benoemen van verbanden (bijv. het staafdiagram om gegevens overzichtelijk weer te geven).
Groep 5/6: Modellen, schema's en grafieken voor het uitdrukken van verdelingen (bijv. tabel, cirkelgrafiek, staafgrafiek) en voor het uitdrukken van verbanden / verloop (bijv. dubbele getallenlijn, tabellen, lijngrafiek).
Groep 7/8: Modellen, schema's en grafieken voor het uitdrukken van verbanden van tijd en afstand, groei, en andere tijdgebonden zaken (bijv. lijngrafiek, staafgrafiek: histogram).
Het blijft echter een vertaling. Elke realistische rekenmethode in het basisonderwijs (*Alles Telt*, *De Wereld in Getallen*, *Pluspunt*, *Rekenrijk*, *Talrijk*, *Wis & Reken*) gaat er op een eigen wijze mee aan de slag. Er is dus geen uniforme leerlijn diagrammen en dat maakt het meteen lastig om u daarvan in dit artikel een beeld te geven. Niet elke methode doet bijvoorbeeld hetzelfde aan kansrekening of combinatoriek. (***In figuur 1*** wel een aantal voorbeelden uit dit domein.) Niet elke methode heeft assen met negatieve getallen, niet elke methode gaat even ver als het gaat om het zelf tekenen van een cirkeldiagram.

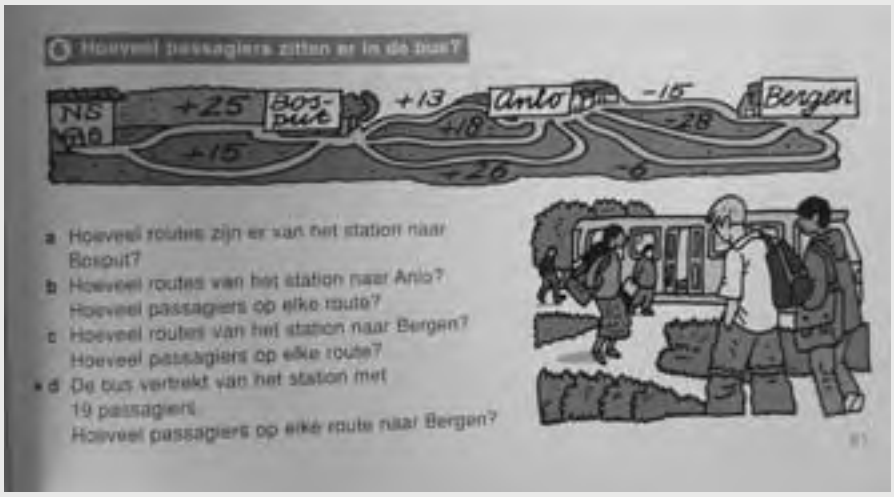
Ondanks de onderlinge verschillen neem ik u mee op reis vanaf groep 1 tot en met groep 8. Ik richt me daarbij voornamelijk op diagrammen. Mijn doel is u te laten zien dat kinderen niet zonder bagage naar het voortgezet onderwijs vertrekken. De bagage mag dan divers zijn, zij is zeker groot genoeg om rekening mee te mogen houden in het voortgezet onderwijs. Bovendien valt er van de aanpak op de basisscholen misschien nog iets te leren. We zullen zien!

Groep 1-2

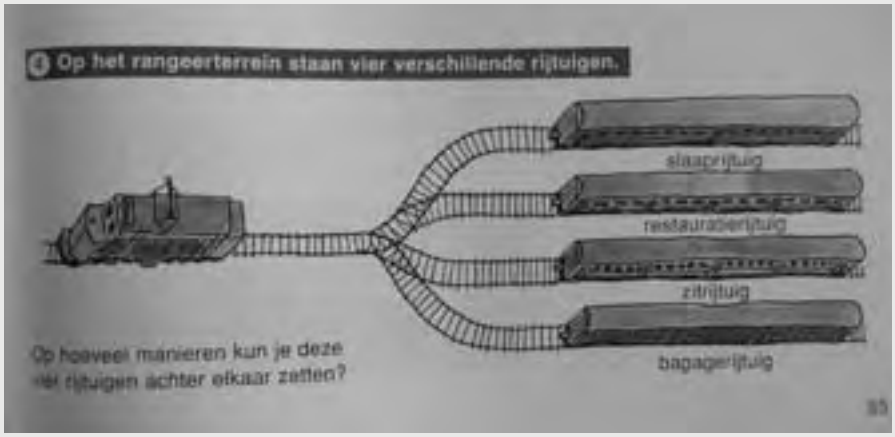
Ik hoor u denken, groep 1-2 en diagrammen of kans? Voor leerkrachten in groep 1 en 2 zijn er geen methodeboeken beschikbaar, maar er zijn wel veel boeken met talloze lesideeën voor reken/wiskunde-activiteiten. Bij onze kleuters begint juist het gevoel voor kans en het kennismaken met grafieken. Bij het spelen van een dobbelsteenspel met een pion over een recht scorespoor start het al met gevoel voor wie er gaat winnen. Een heel jonge kleuter denkt vaak dat hij de meeste kans maakt om het eerst bij het eindpunt te komen, al ligt hij nog zo ver achter zijn medespeler(s). Het kind voelt zich het centrum van de wereld waar alles om draait. Gaandeweg beginnen kinderen in te zien dat hun kans op een overwinning kleiner wordt naarmate de ander verder voorligt. Ze merken dat ze kunnen verliezen. Ook knikkerspelletjes op het schoolplein zijn een eerste aanzet tot het inschatten van kansen. Er komt daardoor steeds meer tactiek in hun knikkerspel.

Niet alleen gevoel voor kans ontwikkelt zich in de kleuterleeftijd, ook leren kinderen dan al gegevens in een tabel te lezen. Denk aan ontwikkelingsmateriaal waarbij in kolommen en rijen moet worden gewerkt. Daarnaast maken kinderen spelenderwijs kennis met diagrammen. Wanneer ze bijvoorbeeld praten over hun lievelingskleur, dan kan een leerkracht dat samen met de kinderen vormgeven in een staafdiagram. Ieder kind pakt een vouwblaadje in zijn of haar lievelingskleur en legt dat op de

figuur 1



De Wereld in Getallen, groep 5



De Wereld in Getallen, groep 6

Rekenrijk, groep 8

grond neer midden in de kring. Dezelfde kleuren worden naast elkaar gelegd en zo ontstaan er staven. Deze staven worden weer met elkaar vergeleken, waardoor er op de grond bijna een echt staafdiagram is ontstaan. Welke kleur is het vaakst de lievelingskleur? Mooi daarbij is soms nog de afwezigheid van het conservatiebegrip. Conservatiebegrip betekent dat het kind doorziet dat iets evenveel blijft, ook al verandert de vorm. Zonder conservatiebegrip zijn vijf dennenappels meer dan vijf eikeltjes, is een liter water in een brede bak minder dan in een smalle hoge vaas en zijn vijf blauwe vouwblaadjes die verder uit elkaar liggen en zo een langere staaf vormen, meer dan de zes rode vouwblaadjes. Daar speelt een leerkracht natuurlijk op in door ze de blaadjes 1 op 1 met elkaar te laten corresponderen. Een belangrijk concept dat een leerkracht voor ogen heeft, is het inzicht dat elk vouwblaadje staat voor één kind. In tegenstelling tot staafdiagrammen in de hogere groepen en op de middel-bare school, is in een staafdiagram dat op deze wijze is ontstaan elke eenheid nog afzonderlijk te herkennen.
Andere voorbeelden zijn het vastleggen van de groei van een plantje met behulp van stroken, of staafdiagrammen maken van het favoriete broodbeleg of van de schoenmaten van alle kinderen in de klas (***zie figuur 2***).

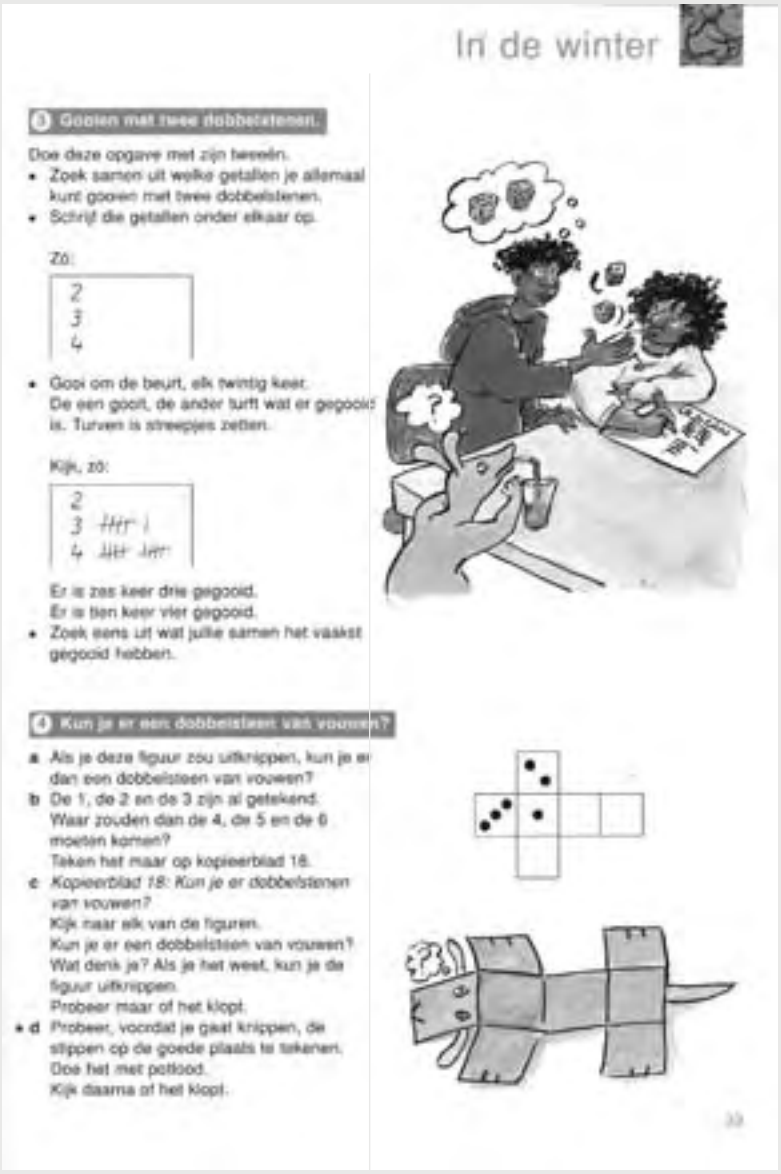
Groep 3-5

Vanaf groep 3 krijgen kinderen les uit een rekenmethode. In deze periode leren de kinderen de turfstreepjes kennen en leren ze tabellen te gebruiken en in te vullen. Voorbeelden van tabellen zijn verhoudingstabellen, somtabellen en informatietabellen zoals een tabel met kledingmaten of de maandkalender. Een aantal methoden laat kinderen experimenteren met kanstollen, het gooien van kop of munt, de uitkomsten van twee dobbelstenen (***zie figuur 3***).
De kinderen leren daarnaast voornamelijk eenvoudige staafdiagrammen te interpreteren en te tekenen. De getallen worden gaandeweg groter, met als gevolg dat één hokje niet altijd meer 1 voorstelt maar een groter aantal kan weergeven. ***In figuur 4*** een aantal voorbeelden van diagrammen uit groep 3 tot en met 5 waarin deze ontwikkeling te zien is.

figuur 2

Tijdens het bezoek aan de kinderboerderij hebben de kinderen van groep 1 en 2 allerlei soorten dieren geteld. Bij terugkomst op school hebben ze voor elk geteld dier een gekleurde plakker op een groot vel geplakt, geordend naar soort dier. Toen alle plakkers erop zaten, zijn we erover gaan praten: waarvan waren er het meest, het minst, evenveel enz. Van de vogels waren er zoveel dat we er een papier bovenaan bij moesten maken. Er waren ook problemen: de geiten waren lastig te tellen omdat die steeds wegliepen. Ook in de paddenpoel troffen we het niet, want er was geen pad/kikker te zien of te horen.'(foto: Olga Meyns).

figuur 3



In de winter

3 Gooien met twee dobbelstenen.

Doe deze opgave met zijn tweeën.

- Zoek samen uit welke getallen je allemaal kunt gooien met twee dobbelstenen.
- Schrijf die getallen onder elkaar op.

Zo:

2
3
4

- Gooi om de beurt, elk twintig keer. De een gooit, de ander turft wat er gegooid is. Turven is streepjes zetten.

Kijk, zo:

2
3 𝍸 |
4 𝍸 𝍸

Er is zes keer drie gegooid.
Er is tien keer vier gegooid.

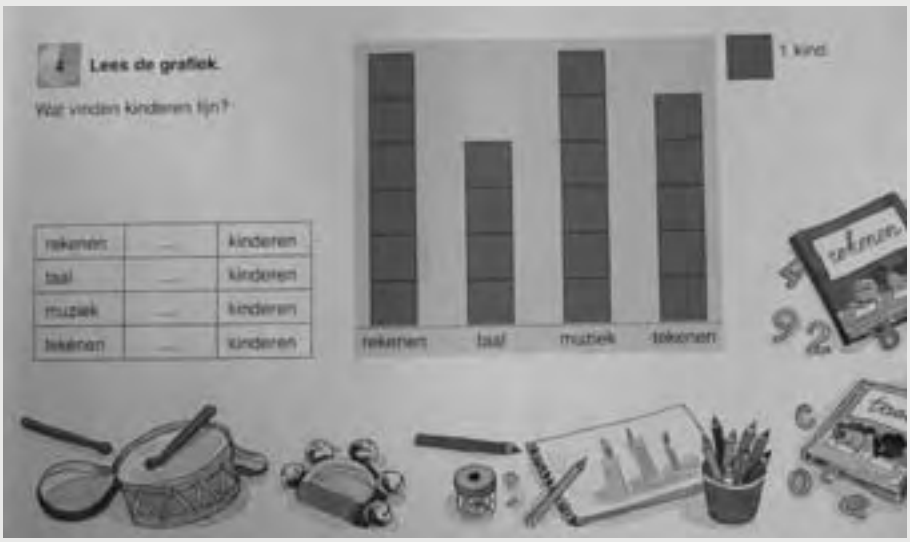
- Zoek eens uit wat jullie samen het vaakst gegooid hebben.

4 Kun je er een dobbelsteen van vouwen?

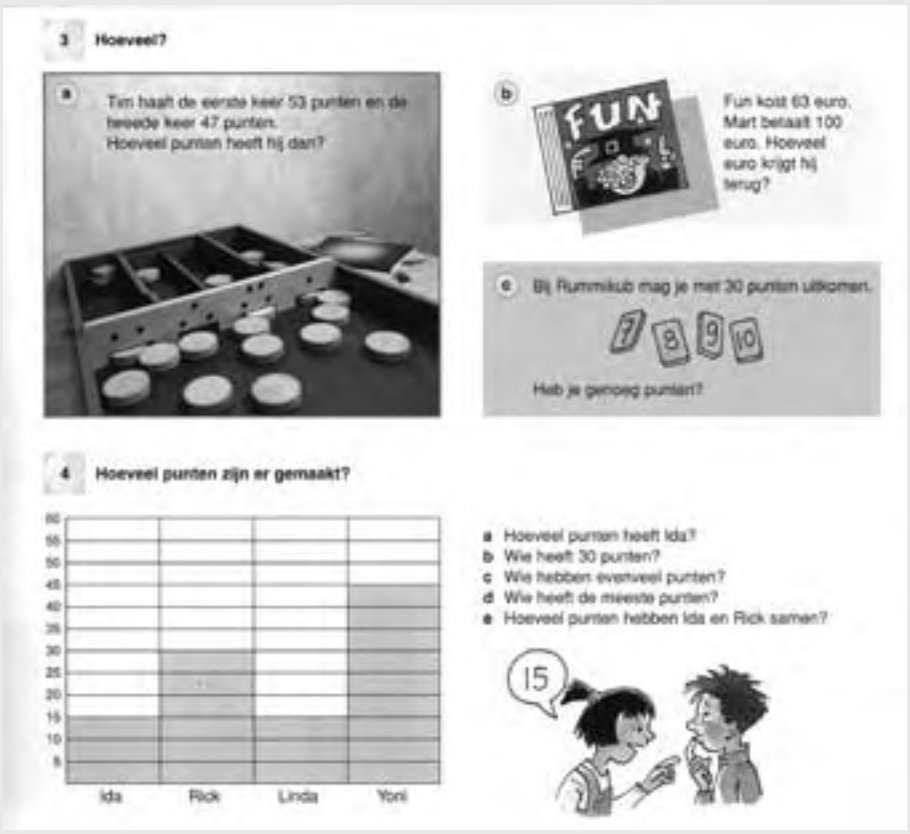
a Als je deze figuur zou uitknippen, kun je er dan een dobbelsteen van vouwen?
b De 1, de 2 en de 3 zijn al getekend. Waar zouden dan de 4, de 5 en de 6 moeten komen?
Teken het maar op kopieerblad 18.
c *Kopieerblad 18: Kun je er dobbelstenen van vouwen?*
Kijk naar elk van de figuren.
Kun je er een dobbelsteen van vouwen?
Wat denk je? Als je het weet, kun je de figuur uitknippen.
Probeer maar of het klopt.
• d Probeer, voordat je gaat knippen, de stippen op de goede plaats te tekenen.
Doe het met potlood.
Kijk daarna of het klopt.

figuur 4

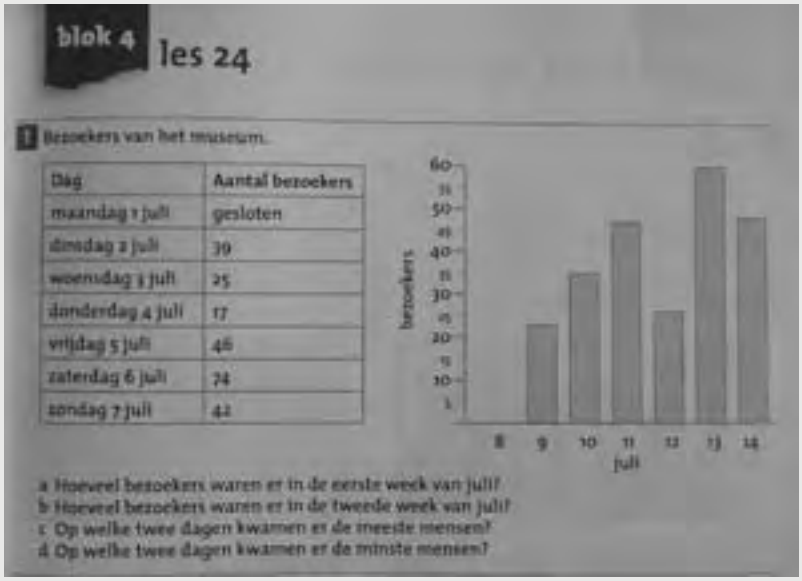
Wis en Reken, groep 3



Pluspunt, groep 4



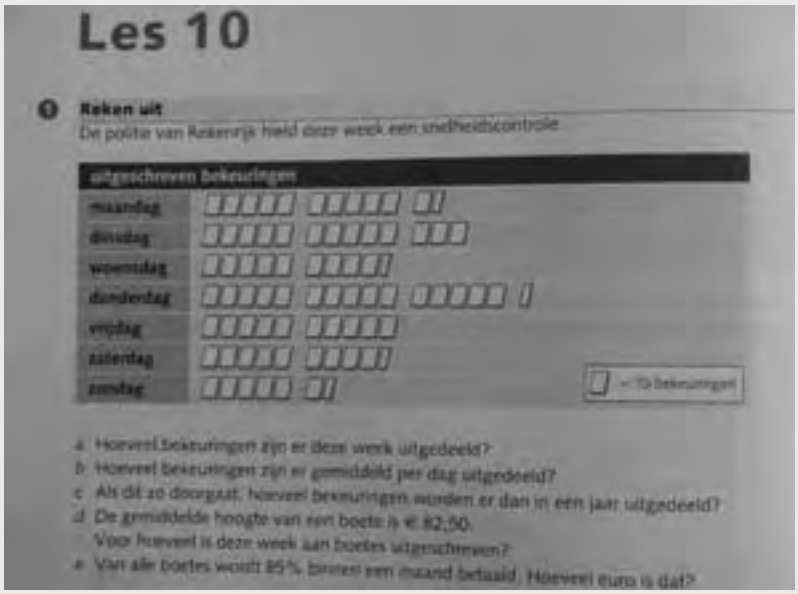
Pluspunt, groep 5



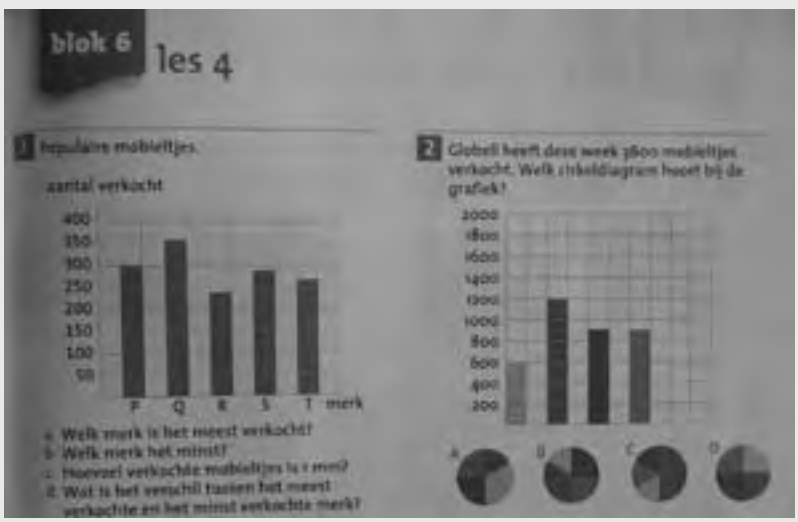
Alles telt, groep 5

---

figuur 5



Beelddiagram
Rekenrijk, groep 6



Staaf- en cirkeldiagram
Alles telt, groep 6

Groep 6-8

Vanaf groep 6 vind je naast het staafdiagram meer verschillende soorten diagrammen en ook frequenter. Het gaat daarbij om diagrammen om te ordenen, hoeveelheden te vergelijken, ontwikkelingen in beeld te brengen, kansen te berekenen of verdelingen te tonen (***zie figuur 5***).

Een aantal voorbeelden:

- het lijndiagram, ook met meerdere lijnen in één diagram (bijvoorbeeld tijd-afstandgrafieken of het verloop van de temperatuur);
- het cirkeldiagram / het sectordiagram (in samenhang met breuken, procenten, verhoudingen);
- het samengestelde staafdiagram met absolute hoeveelheden, in een enkele methode ook met relatieve groottes;
- het strokendiagram;
- het beelddiagram;
- het boomdiagram.

Het zelf tekenen van de verschillende diagrammen krijgt in de ene methode meer aandacht dan in de andere. Boomdiagrammen komen niet in elke methode voor, wat betekent dat dit voor kinderen in het voortgezet onderwijs een nieuw soort diagram kan zijn.

Kinderen verkennen de nieuwe diagrammen en leren ze interpreteren en zelf te maken. Ze leren dat de gegevens in tabellen preciezer zijn, maar dat diagrammen een duidelijker beeld laten zien van de gegevens. Er worden in de klas discussies gevoerd over welk soort diagram geschikt is in welke situatie. Als je een week lang het aantal fietsers telt dat tussen 11:00 uur en 11:15 uur langs je school rijdt, zet je dit dan in een lijndiagram met rechte lijnstukken, een lijndiagram met een vloeiende lijn, een staafdiagram of een cirkeldiagram? Ook eigen constructies hebben een belangrijke plek in het basisonderwijs. Daar is ruimte voor.

De diagrammen worden gaandeweg natuurlijk steeds complexer (***zie figuur 6***). Een aantal aspecten die de diagrammen complexer maken:

- Meerdere lijngrafieken in één plaatje.
- Lijngrafieken optellen of het verschil bepalen.
- Scheurlijnen of zaagtanden.
- Grotere getallen. Langs de verticale as staat bijvoorbeeld '× 1000'.
- Assen met negatieve getallen.
- Gevarieerder uiterlijk van de diagrammen. Onder meer diagrammen uit de krant of uit andere bronnen. Een voorbeeld is de groeigrafiek uit het boekje van het consultatiebureau. Daarin staan lengte en gewicht in één diagram, met daarbij boven- en ondergrenzen.
- Er wordt naar steeds ingewikkelder interpretaties en conclusies gevraagd. Kinderen moeten bijvoorbeeld de snelheid halen uit een tijd-afstand-grafiek of meerdere gegevens in één diagram met elkaar combineren. Ook trekken kinderen conclusies op basis van informatie uit meerdere diagrammen.

Diagrammen komen overigens niet alleen voor bij het onderdeel rekenen, maar ook bij de zaakvakken. De Cito- eindtoets toetst naast rekenen en taal specifiek het onderdeel informatieverwerking.

Na groep 8

En dan op naar de middelbare school! Wat kun je als wiskundedocent verwachten? Kinderen hebben geleerd om gegevens uit diagrammen te lezen en ook kritisch te bekijken. Maar hoe ver ze daarin gegaan zijn, blijft afhankelijk van de methode en de leerkracht die ze hadden (werden ze door hun leerkracht veel geprikkeld met zaken uit de actualiteit bijvoorbeeld). Een mooi voorbeeld waarbij kinderen eind groep 7 werken aan een artikel in de krant, vindt u in een doorkijkje van de SLO; zie [2]. Voor een wiskundedocent dus best goed om eerst af te tasten wat zijn leerlingen al aan bagage hebben voor ze met een misschien erg eenvoudig diagram uit het wiskundeboek komen. Gooi eens een pittig diagram in de groep en kijk wat er gebeurt. Uitdagen is immers een grotere motivatie voor kinderen.

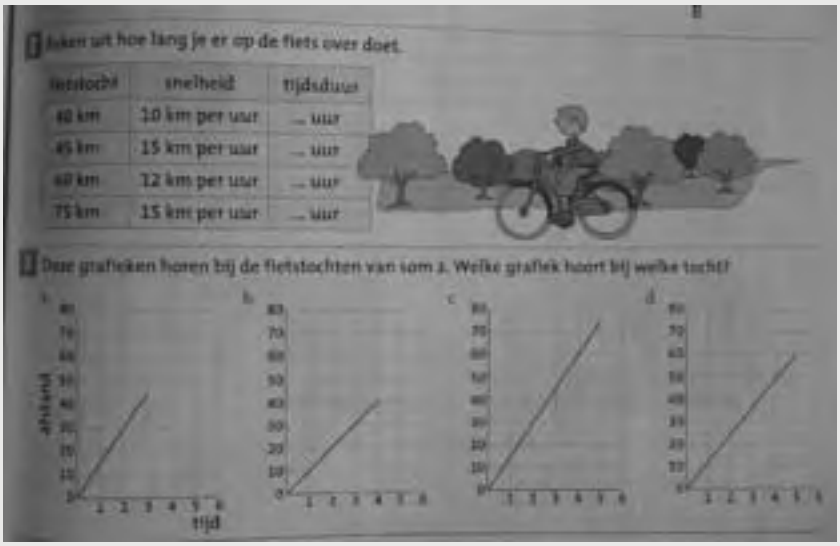
Op de pabo

Een rekenmethode zien we niet als iets waar je elke dag twee bladzijden uit moet maken. Dat zou niet stroken met de gedachte van het realistisch rekenonderwijs, hoe realistisch deze rekenmethoden ook zijn. Tijdens de module 'bovenbouw rekenen' laten we de studenten nadenken hoe je kinderen kunt prikkelen door aanvullende vragen te stellen bij het boek, maar vooral door het loslaten van het boek. Wiskunde is immers in de wereld om ons heen. Actualiteiten en krantenartikelen zijn daarbij een mooi hulpmiddel. Deze worden door veel leerkrachten voornamelijk ingezet in hun taallessen, bijvoorbeeld door kinderen een muurkrant te laten maken of door een klassikaal gesprek te voeren naar aanleiding van één of meer krantenberichten (de zogenaamde 'krantenkring'). Maar zulke actualiteiten en artikelen zijn ook uiterst geschikt als aanzet tot een rekenprobleem of tot het doen van onderzoek. De sturing in een methode van de basisschool is minder groot dan die in een boek van de middelbare school, maar desalniettemin worden niet altijd belangrijke mentale processen bij de leerling aangewakkerd. Dat is de taak van de leerkracht. Een goede leerkracht zal de kinderen uitdagen met goede, rijke, uitdagende problemen en die vind je niet vaak in een boek. Dat zijn actuele, betekenisvolle, open problemen die op allerlei niveaus kunnen worden aangepakt. Dat proberen we de studenten deze module mee te geven en bij te brengen. Zij zullen problemen mee moeten nemen naar de klas en ook spontane situaties in de klas moeten kunnen uitbuiten om er wiskunde mee te gaan doen. De leerkracht heeft dus een heel belangrijke rol in het organiseren van wiskundige activiteiten naast de methode.

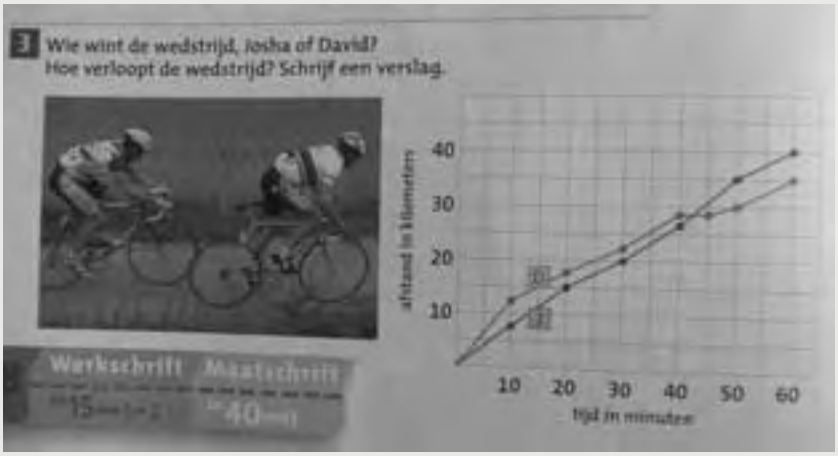
Om onze pabo-studenten eens goed los te weken van de rekenmethode geven we ze de opdracht de kinderen in de stageklas een onderzoek te laten uitvoeren. De duur van dit project is minimaal vier dagdelen en in die tijd stellen de kinderen onderzoeksvragen op, stellen ze een hypothese op, voeren ze een onderzoek uit en verwerken ze de resultaten van dat onderzoek in diagrammen en tabellen. Het opstellen van de onderzoeksvragen is iets dat we oefenen met onze studenten. Wat voor een soort vragen leent zich voor een onderzoek? Hoe krijg je goede vragen uit de kinderen? Hoe laat je ze de verantwoordelijkheid nemen over hun aanpak van de probleemstelling? Studenten gaan vaak wel te werk vanuit een thema, zoals verkeer of voeding. Maar het is de bedoeling dat de vragen echt uit de kinderen zelf komen en die authenticiteit verhoogt de intrinsieke motivatie om het probleem aan te pakken. Zo zie ik vragen als: Zijn duurdere etenswaren lekkerder? Welke sport doen ze in mijn klas? Hoe gaan we de speelplaats inrichten en hoeveel geld is daarvoor nodig? Dat maakt het voor de leerkracht nogal spannend, want in plaats van degene te zijn met het antwoord, heb je ineens de rol van coach en ben je zelf net zo nieuwsgierig naar de uitkomsten.

Het onderzoek is een vakoverstijgend project, waar ook het vak Nederlands aan is gekoppeld (hoe maak je een goed verslag, welke vragen zet ik in mijn enquête, hoe geef ik het eindproduct vorm als dit een muurkrant of poster wordt, hoe houd ik een interview, enz). Daarnaast moeten de studenten coöperatieve werkvormen inzetten. Op de opleiding leren zij vele werkvormen kennen en toepassen. Ook in deze module.

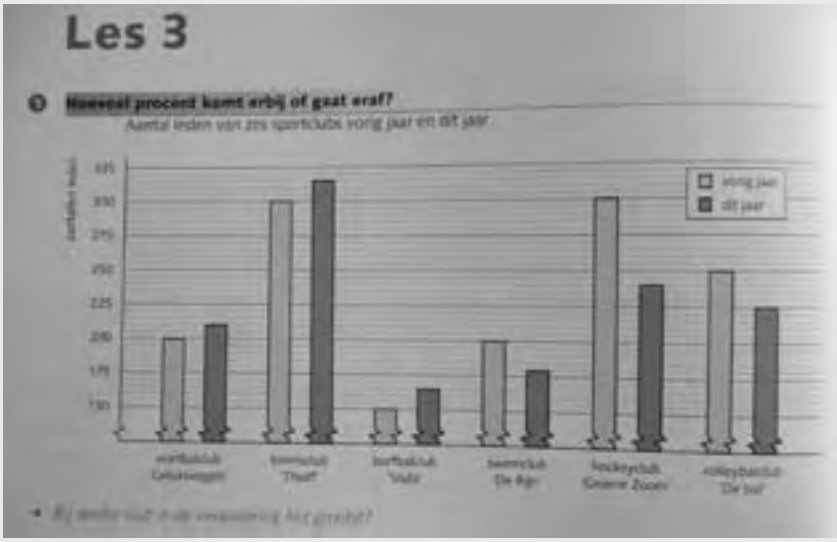
figuur 6



Alles telt, groep 6



Meerdere lijnen in een grafiek.
Alles telt, groep 7



Scheurlijnen
Rekenrijk, groep 8

TI-*nspire*™ TECHNOLOGIE

Een nieuwe visie vanuit meerdere wiskundige invalshoeken

Elke leerling leert op een andere manier.
De een begrijpt vergelijkingen vlot, de ander grafieken. De nieuwe TI-Nspire™ technologie voor Wiskunde en Exact is geschikt voor verschillende individuele manieren van leren. Lesmateriaal wordt gepresenteerd en onderzocht naar de voorkeur van de individuele leerling. Leerlingen kunnen daardoor wiskundige relaties en verbanden veel gemakkelijker waarnemen.

Als rekenmachine en als software voor de computer beschikbaar.

TI-Nspire™ TECHNOLOGIE
Voor een beter begrip van de wiskunde.

www.education.ti.com/nederland

1 ALGEBRA

2 GRAFIEKEN/ MEETKUNDE

3 LIJSTEN/ SPREADSHEETS

4 TEKSTVERWERKEN

VIERDYNAMISCH GEKOPPELDE OMGEVINGEN, TE BEWAREN IN ÉÉN DOCUMENT

Nu tijdelijk
TI-Nspire™ bundel
(handheld +software)
voor slechts € 101,75 ! *
tel 020 - 58 29 490

* exclusief € 9 verzendkosten

Texas Instruments

Uw expertise. Onze technologie. Succes voor de leerling.

Vanuit het vakgebied rekenen vul ik de gereedschapskoffer van de pabo-student met verschillende soorten diagrammen en tabellen. Ik leer ze kritisch te kijken naar de functionaliteit van elk soort diagram. Zij moeten de groepjes kinderen in de bovenbouw tijdens het project daarin immers gedegen kunnen begeleiden. Niet door te zeggen welk soort diagram het meest geschikt is, maar door kritische vragen te stellen die de kinderen laten nadenken over hun keuzes.
We richten ons ook op zaken als misleiding in diagrammen, onder andere door de keuze van de as-indeling of de titel. Begrippen als modus, mediaan en gemiddelde passeren als manieren om gegevens te presenteren. Ik vraag ze niet alleen om deze te berekenen, maar vooral om na te denken wanneer nu welke maat een goede duiding is van een reeks gegevens. Een schoenverkoper heeft werkelijk niets aan de gemiddelde maat schoenen wanneer hij weer gaat inkopen.

Tot slot

De wiskundetaal van mijn eigen kinderen en de minder gestructureerde manier van leren omgaan met wiskundetaal op de basisschool maken me bewust van de vaak nogal rechtlijnige aanpak van diagrammen in het voortgezet onderwijs. Al zijn ook daar grote verschillen tussen methoden onderling. Vaak zie je dat het wiskundeboek de leerling erg bij de hand neemt, bijvoorbeeld in de vorm van een gestructureerd stappenplan of gewoonweg door heel expliciet naar een bepaald type diagram te vragen.

Sommige rekenmethoden van de basisschool lijken steeds meer op wiskundeboeken van de middelbare school door te kiezen voor een meer rechtlijnige aanpak. Waarschijnlijk omdat er naar aansluiting wordt gezocht? Toch blijft de sturing op de basisschool veel minder groot. Er is minder sturing ten aanzien van hoe het diagram gebruikt wordt, de leerling moet vooral zelf afwegen welke manier het beste bij zijn probleem of gegevens past.

Via vaste stappenplannen werken is wel handig, maar het lijkt me nuttig voor het verder vormen van de manier van denken van leerlingen om ze niet altijd klakkeloos zo'n trucje in te laten zetten. Ik merk bij mijn studenten vaak een hang naar eigen maniertjes van oplossen. Bijvoorbeeld 'delen door een breuk is het vermenigvuldigen met het omgekeerde' of 'om van cm^2 naar m^2 te gaan moeten er 4 nullen af'. Zo hebben ze het geleerd, vertellen ze me dan, zo hebben ze het altijd gedaan, zo weten ze in ieder geval zeker dat er een geldig antwoord uitkomt. Maar ze kunnen mij niet meer uitleggen waaróm het op die manier een goed antwoord oplevert. Ik laat ze ervaren dat sommige vaste procedures ook heel onhandig kunnen zijn. Studenten die bij het vermenigvuldigen van breuken altijd alles in teller/noemer gaan schrijven geef ik bijvoorbeeld de opdracht om $6\times315\frac{1}{6}$ uit te rekenen. Dat resulteert dan na veel gereken in $\frac{11346}{6}$ en dat moet dan nog worden vereenvoudigd. Studenten overtuigen elkaar van handiger manieren.

Af en toe een goede discussie in de klas op gang brengen, het laten afwijken van de vaste oplossingsroute of het laten presenteren van eigen onderzoeksgegevens maakt dat een leerling pas echt moet nadenken over het hoe en waarom. Het soort diagram open laten in de vraagstelling zal een leerling tot denken zetten over de meest geschikte presentatie en kan na een les prachtig voer voor een discussie zijn. Dan pas zie je of ze daadwerkelijk inzicht hebben in het geheel en of er niet alleen routinematig iets wordt gemaakt of opgelost. Samen met de leerlingen praten over hun eigen constructies en over handige strategieën brengt leerlingen weer op een hoger niveau. Als leerkracht heb je immers bepaalde doelen voor ogen die je door middel van die leergesprekken kunt bereiken.

Ik heb in mijn voormalige baan als wiskundedocent weinig ruimte gegeven aan kinderen voor eigen constructies. Het boek was daarbij erg bepalend. Zoals het wiskundeboek het wilde, zo deden de kinderen. En ik ging daar in mee. In de theorieblokken kunnen die methodes erg voorschrijvend zijn en daarmee de kinderen het denken ontnemen.
Op de basisschool staan er geen theorieblokken in het boek waar de kinderen uit werken. Zij worden daar dus ook niet door afgeleid of gestuurd. Wel is er een vaak zeer uitgebreide docentenhandleiding. Daarin vindt de basisschoolleerkracht algemene richtlijnen, maar ook didactische adviezen en tips op opgaveniveau. Bijvoorbeeld op welke manier hij kan inspelen op de verschillende aanpakken die in de klas zijn gebruikt bij het maken van de opgave en voor welke van die aanpakken hij de kinderen door middel van reflectie zou moeten motiveren. Sommige leerkrachten in het basisonderwijs varen helaas nog altijd op hun eigen trukendoos en openen die voor de kinderen van de klas. Maar dat is absoluut niet de bedoeling.
Misschien is het mooi wanneer er minder voorschrijvende theorieblokken in de wiskundemethodes zouden staan en de docent meer didactische adviezen in een aparte handleiding vindt. Dan zou er minder sturing zijn in de aanpak en is er daarmee meer ruimte voor eigen constructies. Dat hele oorspronkelijke denken van jonge kinderen, zoals over die gele en roze balletjes, wil je immers de ruimte blijven geven en zeker niet verloren laten gaan.

Noten

[1] De uitwerking van de SLO vindt u op *http://tule.slo.nl/RekenenWiskunde/F-L23.html*.

[2] Een mooi voorbeeld waarbij kinderen eind groep 7 werken aan een artikel in de krant vindt u op *http://tule.slo.nl/RekenenWiskunde/F-L23-Gr78-Doorkijkje.html*.

Over de auteur

Maike den Houting is sinds augustus 2005 werkzaam als docent vakdidactiek rekenen/wiskunde aan de Saxion Hogescholen te Deventer (pabo). Daarnaast is zij medeauteur van de nieuwe Wageningse Methode. Tot augustus 2005 was zij werkzaam als docent wiskunde in het vmbo en de onderbouw havo/vwo.
E-mailadres: *m.i.j.denhouting@saxion.nl*

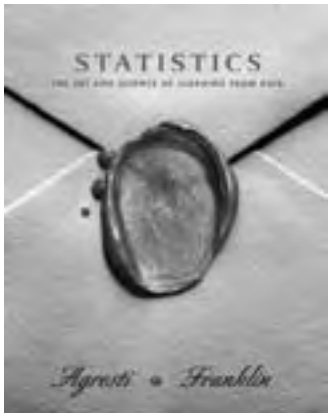
De Yule-Simpson paradox

OF WAAROM SOMS DE INDRUK ONTSTAAT DAT JE MET STATISTIEK OM HET EVEN WAT KUNT BEWIJZEN

[Marcel Croon]

Vaak hoor je beweren in discussies waarin argumenten door statistische cijfers worden ondersteund, dat je met statistiek alles kunt bewijzen. Het is niet aan mij om hier in dit artikel dit type van tegenargument volledig te ontkrachten, maar graag wil ik een bepaald op het eerste oog paradoxaal fenomeen bespreken dat misschien wel mede ten grondslag ligt aan dit misverstand over de toepasbaarheid van statistische methoden.

In statistische analyses op verschillende inhoudelijke gebieden (sociologie, psychologie, criminologie, economie, medische wetenschappen, etc.) krijgt men vaak te maken met een fenomeen dat voor een leek moeilijk te begrijpen valt. Dit fenomeen staat bekend als Simpsons paradox naar de naam van de Engelse statisticus die hierover in 1951 een kort artikel schreef in the Journal of the Royal Statistical Society. Aangezien in 1903 een andere Engelse statisticus, Yule, al op het fenomeen had gewezen, wordt er ook wel terecht met de term *Yule-Simpson paradox* naar verwezen.



Een voorbeeld

Ik kan deze paradox het makkelijkst toelichten aan de hand van een voorbeeld dat ik ontleend heb aan een criminologisch onderzoek waarin op basis van cijfers verzameld in Florida tussen 1976 en 1980 het verband werd nagegaan tussen de huidskleur van een verdachte (blank versus zwart) en het veroordelen tot de doodstraf. Je zou bijvoorbeeld kunnen verwachten dat zwarte mensen in Amerika sneller tot de doodstraf worden veroordeeld dan blanke mensen. Onderstaande tabel beschrijft hoe de kans dat de doodstraf wordt uitgesproken verandert met de huidskleur van de verdachte. Ik heb deze gegevens ontleend aan het uitstekende leerboek voor elementaire toegepaste statistiek dat Alan Agresti en Christine Franklin in 2007 publiceerden. Deze auteurs refereren hierbij zelf naar een artikel van Radelet en Pierce dat in 1991 in de *Florida Law Review* verscheen.

Huidskleur verdachte	Doodstraf?			
	Ja	Nee	totaal	% Ja
Blank	39	308	347	11,2
Zwart	32	345	377	8,8
totaal	71	653	724	9,8

tabel 1 De relatie tussen de huidskleur van de verdachte en de rechterlijke uitspraak in de totale groep

Uit ***tabel 1*** kan men opmaken dat globaal genomen blanke verdachten een groter risico (11,2 %) op de doodstraf lopen dan zwarte verdachten, waarvan slechts 8,8 % tot de doodstraf wordt veroordeeld.
Bij het opstellen van deze tabel is echter een belangrijke derde variabele over het hoofd gezien: de huidskleur van het slachtoffer. Ook voor deze derde variabele maken we slechts een onderscheid tussen blank en zwart zodat we kunnen volstaan met een vergelijking van de resultaten voor twee deelgroepen. De eerste deelgroep betreft de misdaden waarbij het slachtoffer blank was; in de tweede deelgroep waren de slachtoffers zwart.

Tabel 2 beschrijft nu de relatie tussen de huidskleur van de verdachte en het al dan niet uitspreken van de doodstraf voor de misdaden met een blank slachtoffer.

Huidskleur verdachte	Doodstraf?			
	Ja	Nee	totaal	% Ja
Blank	39	279	318	12,3
Zwart	29	121	150	19,3
totaal	68	400	468	14,5

tabel 2 De relatie tussen de huidskleur van de verdachte en de rechterlijke uitspraak voor misdaden met een blank slachtoffer

Voor misdaden met een zwart slachtoffer geeft ***tabel 3*** de observaties.

Huidskleur verdachte	Doodstraf?			
	Ja	Nee	totaal	% Ja
Blank	0	29	29	0,0
Zwart	3	224	227	1,3
totaal	3	253	256	1,2

tabel 3 De relatie tussen de huidskleur van de verdachte en de rechterlijke uitspraak voor misdaden met een zwart slachtoffer

Als we naar de cijfers in de laatste twee tabellen kijken, dan stellen we iets heel vreemds vast: als we de oorspronkelijke tabel opsplitsen door rekening te houden met de huidskleur van het slachtoffer, blijken zwarte verdachten een grotere kans op de doodstraf te lopen dan blanke verdachten. Bij blanke slachtoffers lopen zwarte verdachten 7 % (= 19,3 – 12,3) meer kans op veroordeling tot de doodstraf dan blanke verdachten; bij zwarte slachtoffers is dat verschil weliswaar kleiner (1,3 %) maar de kans is nog steeds in het nadeel van de zwarte verdachten.

Dit voorbeeld toont aan waar het in de Yule-Simpson paradox om gaat: de richting van het verband tussen twee variabelen slaat om als we een groep opsplitsen in een aantal deelgroepen, met andere woorden, als we rekening houden met een derde variabele.

Hoe verklaren we deze paradox?

Om een verklaring voor ons getallenvoorbeeld op te stellen moeten we eerst de samenhang tussen de huidskleur van het slachtoffer en het verdict nagaan. Informatie hierover vinden we in ***tabel 4***.

Huidskleur slachtoffer	Doodstraf?			
	Ja	Nee	totaal	% Ja
Blank	68	400	468	14,5
Zwart	3	253	256	1,2
totaal	71	653	724	9,8

tabel 4 Relatie tussen de huidskleur van het slachtoffer en de rechterlijke uitspraak

Op basis van deze tabel mogen we concluderen dat de kans dat een doodstraf wordt uitgesproken, veel groter is als het slachtoffer blank is dan wanneer hij zwart is.

Vervolgens kijken we in ***tabel 5*** ook naar de samenhang (in tabelvorm) tussen de huidskleur van de verdachte en huidskleur van het slachtoffer.

Huidskleur verdachte	Huidskleur slachtoffer			
	Blank	Zwart	totaal	% Blank
Blank	318	29	347	91,6
Zwart	150	227	377	39,8
totaal	468	256	724	64,6

tabel 5 Relatie tussen de huidskleur van de verdachte en de huidskleur van het slachtoffer

Uit deze tabel maken we op dat blanke verdachten meestal betrokken zijn bij misdaden met een blank slachtoffer; zwarte verdachten zijn meestal betrokken bij misdaden met een zwart slachtoffer.
Het beeld in de totale groep dat zwarte verdachten minder snel tot de doodstraf worden veroordeeld, is het gevolg van twee andere 'causale' relaties: (1) zwarte verdachten zijn vaker betrokken bij misdaden met een zwart slachtoffer, en (2) in misdaden met een zwart slachtoffer wordt minder snel de doodstraf uitgesproken. Hierdoor komt het dat in de totale groep – wanneer we geen rekening houden met de huidskleur van het slachtoffer – zwarte verdachten minder snel de doodstraf krijgen. Als we echter de totale groep opsplitsen volgens de huidskleur van het slachtoffer, dan zien we dat zwarte verdachten in elke deelgroep vaker de doodstraf krijgen.
De huidskleur van het slachtoffer is een wat met een Engelse term *confounder* genoemd wordt. In het Nederlands spreken we van een *storende variabele* die het verband tussen de twee kernvariabelen beïnvloedt. Om het ware directe verband tussen die twee kernvariabelen te onderzoeken, moet de storende variabele constant worden gehouden. In ons voorbeeld moeten we dan de relatie tussen de huidskleur van de dader en de aard van het vonnis onderzoeken in de twee deelgroepen op basis van de huidskleur van het slachtoffer.

Wat is er vanuit wiskundig perspectief aan de hand?

Een meer wiskundige verklaring van de Simpson paradox gaat uit van de vaststelling dat uit de twee ongelijkheden $\frac{a}{b} > \frac{c}{d}$ **én** $\frac{e}{f} > \frac{g}{h}$ *niet* automatisch volgt dat $\frac{a+e}{b+f} > \frac{c+g}{d+h}$
Let op dat we hierboven niet de breuken zelf bij elkaar hebben opgeteld, maar wel rechtstreeks de overeenkomstige tellers en noemers. Een eenvoudig getallenvoorbeeld kan als illustratie dienen: ofschoon $\frac{4}{1} > \frac{18}{11}$ **én** $\frac{13}{14} > \frac{5}{9}$, hebben we toch $\frac{4+13}{1+14} < \frac{18+5}{11+9}$
Om aan te geven dat deze observatie relevant is voor ons voorbeeld, definiëren we drie variabelen A, B, en C die elk slechts twee waarden kunnen aannemen. De variabele A correspondeert met de huidskleur van het slachtoffer en we stellen $A = 1$ indien het slachtoffer blank is en $A = 2$ indien zwart. De variabele B stelt de huidskleur van de verdachte voor met dezelfde conventie als voor A: $B = 1$ indien blank en $B = 2$ indien zwart. Tenslotte stelt varia-

bele C het verdict voor: $C = 1$ indien doodstraf en $C = 2$ indien geen doodstraf. De kansen dat de doodstraf wordt uitgesproken indien slachtoffer en verdachte een bepaalde huidskleur hebben, kunnen aan de hand van vier voorwaardelijke kansen worden bepaald. Onderstaande tabel bevat hier een overzicht van.

Kans op doodstraf	Slachtoffer	Verdachte
$p(C = 1 \mid A = 1, B = 1)$	Blank	Blank
$p(C = 1 \mid A = 1, B = 2)$	Blank	Zwart
$p(C = 1 \mid A = 2, B = 1)$	Zwart	Blank
$p(C = 1 \mid A = 2, B = 2)$	Zwart	Zwart

In onze twee deeltabellen constateerden we nu dat de kans op de doodstraf groter was voor zwarte dan voor blanke mensen als we bovendien controleerden voor de huidskleur van het slachtoffer. Hiermee corresponderen twee ongelijkheden tussen de vier voorwaardelijke kansen:

$p(C = 1 \mid A = 1, B = 2) > p(C = 1 \mid A = 1, B = 1)$, en
$p(C = 1 \mid A = 2, B = 2) > p(C = 1 \mid A = 2, B = 2)$

Voor voorwaardelijke kansen in termen van willekeurige gebeurtenissen E en F geldt per definitie:

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

zodat we voor onze twee ongelijkheden kunnen schrijven:

$$\frac{p(C=1, A=1, B=2)}{p(A=1, B=2)} > \frac{p(C=1, A=1, B=1)}{p(A=1, B=1)} \quad \text{(ongelijkheid 1)}$$

en

$$\frac{p(C=1, A=2, B=2)}{p(A=2, B=2)} > \frac{p(C=1, A=2, B=1)}{p(A=2, B=1)} \quad \text{(ongelijkheid 2)}$$

Als we nu de overeenkomstige tellers en noemers uit de twee ongelijkheden optellen dan krijgen we aan de linkerkant:

$$\frac{p(C=1, A=1, B=2) + p(C=1, A=2, B=2)}{p(A=1, B=2) + p(A=2, B=2)} = \frac{p(C=1, B=2)}{p(B=2)} = p(C=1 \mid B=2)$$

en aan de rechterkant:

$$\frac{p(C=1, A=1, B=1) + p(C=1, A=2, B=1)}{p(A=1, B=1) + p(A=2, B=1)} = \frac{p(C=1, B=1)}{p(B=1)} = p(C=1 \mid B=1)$$

De voorwaardelijke kansen $p(C = 1 \mid B = 1)$ en $p(C = 1 \mid B = 2)$ stellen nu de kansen voor dat een blanke c.q. zwarte verdachte tot de doodstraf wordt veroordeeld zonder dat rekening wordt gehouden met de huidskleur van het slachtoffer. Deze kansen vinden we voor ons voorbeeld terug in ***tabel 1*** hierboven.
Gelet op de merkwaardige ombuiging van het ongelijkheidsteken tussen breuken die kan optreden als we overeenkomstige tellers en noemers optellen, kunnen we nu concluderen dat uit de ongelijkheden (1) en (2) *niet* noodzakelijk moet volgen dat:
$p(C = 1 \mid B = 1) > p(C = 1 \mid B = 2)$
En inderdaad: in onze eerste tabel hebben we precies de omgekeerde ongelijkheidsrelatie geconstateerd. Vanuit een zuiver wiskundig standpunt bekeken is de Yule-Simpson paradox helemaal geen paradox, maar iets dat met heel elementaire wiskundige inzichten verklaard kan worden.

Literatuur

- A. Agresti, C. Franklin (2007): *Statistics. The Art and Science of Learning from Data.* Upper Saddle River (New Jersey): Pearson Prentice Hall.
- M. Radelet, G. Pierce (1991): *Choosing who will die: Race and the death penalty in Florida*. In: *Florida Law Review*, 43; pp. 1- 34.

Over de auteur

Marcel Croon is hoofddocent aan het departement Methoden en Technieken van Onderzoek van de Faculteit Sociale Wetenschappen van de Universiteit van Tilburg. Hij verzorgt onderwijs en doet onderzoek op het gebied van de toegepaste statistiek voor de sociale wetenschappen. E-mailadres: *m.a.croon@uvt.nl*

Statistiek en geheimschriften

[Rob Bosch]

figuur 1 Een met Pigpen Cipher gecodeerde tekst

Inleiding

Door de eeuwen heen hebben mensen hun geschreven boodschappen gecodeerd om ze zo te beschermen tegen onderschepping. Een tekst werd daarvoor op een bepaalde wijze getransformeerd in een voor de 'tegenstanders' onbegrijpelijk bericht. Alleen diegene die het geheim van de transformatie kende, werd in staat geacht de oorspronkelijke boodschap te achterhalen. Maar geheimschriften zijn uitdagende puzzels, en rivalen gingen al snel de intellectuele uitdaging aan om deze geheime codes te ontcijferen. Was een code eenmaal gekraakt, dan ontstond de noodzaak om een nieuw en beter geheimschrift te ontwikkelen. Zo begon een voortdurende wedloop tussen cryptografen, de bedenkers van geheime codes, en crypto-analysten, de brekers ervan. We bespreken in dit artikel een geheimschrift dat gedurende lange tijd in het verleden als onbreekbaar werd beschouwd, totdat een *statistische* aanval buitengewoon effectief bleek om de code te kraken.

We gebruiken in het vervolg de term *klare tekst* voor de oorspronkelijke tekst. De versleutelde tekst noemen we de *cijfertekst*. De termen *versleutelen*, *versluieren* en *vercijferen* worden als synoniemen gebruikt. De methode om een tekst naar een, voor tegenstanders, onbegrijpelijke tekst te transformeren heet *vercijferalgoritme*. Bij een dergelijk vercijferalgoritme hoort een *sleutel* die - als het goed is - alleen bij de verzender en de beoogde ontvanger van het bericht bekend is. Ter illustratie: als vercijferalgoritme kunnen we afspreken de letters in de klare tekst een aantal plaatsen in het alfabet op te schuiven. De bijbehorende sleutel is dan het aantal plaatsen van de verschuiving. Ten slotte, een bepaalde wijze van vercijfering noemen we vaak kortweg een *cijfer*. Het bovengenoemde verschuivingsalgoritme heet bijvoorbeeld een *Caesar Cijfer*.

Substitutiecijfer

Een eenvoudige manier om een tekst te vercijferen is het vervangen van de letters uit de klare tekst door andere letters uit het alfabet of door onbegrijpelijke symbolen. We noemen een dergelijke vercijfering een *substitutiecijfer*. Van deze wijze van vercijfering zijn in de geschiedenis van het geheimschrift een groot aantal voorbeelden bekend. Zo gebruikte Julius Caesar een substitutiecijfer waarbij de letters uit de klare tekst drie plaatsen werden verschoven. De letter a in de klare tekst wordt de letter D in de cijfertekst, de letter b schrijven we in de cijfertekst als E enz. Hierbij wordt de z gevolgd door de letter a, waardoor een y in de klare tekst vervangen wordt door de letter B. In het vervolg schrijven we de klare tekst steeds in kleine letters en de cijfertekst in hoofdletters.

De klare tekst:

Julius Caesar

wordt na verschuiving (*substitutie*):

MXOLXV FDHVDU

Hoe onbegrijpelijk de cijfertekst ook moge lijken, als we eenmaal het vercijferalgoritme (verschuiving) kennen, dan is de ontcijfering een fluitje van een cent. We proberen gewoon alle 26 mogelijke verschuivingen en wachten totdat er een zinnige boodschap tevoorschijn komt. Hierbij komt een zwakte van dit type geheimschrift aan het licht. Het aantal sleutels om de cijfertekst te openen is slechts 26. Dat geeft de ontcijferaar de mogelijkheid om alle sleutels (de 26 mogelijke verschuivingen) te proberen om zo - in korte tijd - de oorspronkelijke tekst te kunnen lezen. Maar we kunnen het natuurlijk beter, veel beter doen dan Julius Caesar. Als we de letters uit de klare tekst op een willekeurige wijze vervangen door andere letters of symbolen, krijgen we een vercijferalgoritme met een gigantisch aantal sleutels.

We geven een voorbeeld van een substitutiecijfer met een symbolisch alfabet(= een verzameling van 26 tekens). Als alfabet kiezen we de volgende symbolen: ⊕, ⊗, ⊖, ⊘, ..., ⊜

Ieder symbool stelt hierin een letter van ons alfabet voor. In het symbolisch alfabet luidt een vercijferde tekst als volgt:

⊕⊗⊖⊕⊘⊕⊜⊕⊗⊖⊕

Als we niet beschikken over de sleutel (welk symbool stelt welke letter voor), dan valt het niet mee dit abracadabra te ontcijferen. Voor de beoogde ontvanger van het bericht, die over de onderstaande sleutel beschikt, is het ontcijferen echter een simpel klusje.

sleutel

a	→	⊕
b	→	⊗
c	→	⊘
d	→	⊜
...	→	...
r	→	⊖
...	→	...

De sleutel laat zien dat we de letter a in de klare tekst vervangen door ⊕, de letter b door ⊗, enz. De sleutel s van het substitutiecijfer is dus een bi-jectie van ons (letter)alfabet A naar het alfabet van symbolen S:

$s: A \rightarrow S$

Het aantal bi-jecties van A naar S is gelijk aan 26!. Immers, voor a hebben we de keuze uit 26 symbolen, voor b hebben we daarna nog de keuze uit 25 symbolen, …

Het totaal aantal mogelijke koppelingen wordt dan $26 \cdot 25 \cdot 24 \cdots 2 \cdot 1 = 26!$ Het aantal sleutels van het substitutiecijfer is dus $26! \approx 4 \cdot 10^{26}$. Het één voor één proberen van de sleutels is zelfs voor een krachtige computer onbegonnen werk. (De lezer kan zelf nagaan hoe lang een miljard computers die elk een miljard mogelijkheden per seconde kunnen nagaan, er over zouden doen om alle sleutels te proberen.) Er waren in de tijd dat geheimschriften vooral gebaseerd waren op substitutiecijfers, zeker geen monniken genoeg om dit werk te doen. Het lijkt dan ook een bijna onmogelijke opgave om een geheimschrift te kraken dat gebaseerd is op een substitutie.
Als we de sleutel meedelen aan diegene voor wie ons bericht bestemd is, dan kan deze persoon de klare tekst in een handomdraai weer terughalen door de symbolen volgens de opgegeven sleutel weer te vertalen naar de oorspronkelijke letters; met andere woorden, door de inverse functie s^{-1} toe te passen op de cijfertekst.

Geheimschriften waarin letters door symbolen zijn vervangen, ogen altijd wat mysterieus. Een bekend voorbeeld van zo'n geheimzinnige code is het zogenaamde *Pigpen Cipher*, het geheimschrift van de vrijmetselaars, waarin symbolen als ⊏, ∨ en ◘ voorkomen (***zie figuur 1***). Deze code werd overigens ook gebruikt door de federale troepen tijdens de Amerikaanse burgeroorlog. In plaats van door mysterieuze symbolen kunnen we de letters in de klare tekst ook gewoon door andere letters uit het alfabet vervangen. De boodscha p die met symbolen vercijferd werd tot:

zou met de sleutel:

sleutel		
a	→	X
b	→	R
c	→	P
d	→	U
…	→	…
r	→	Z
…	→	…

worden geschreven als
XRZXPXUXRZX
De sleutel s is in dit geval een bi-jectie van A naar A:
$s: A \to A$
Anders gezegd, de sleutel is een permutatie van de letters van het alfabet. In plaats van de letter a door het symbool ⊕ te vervangen, schrijven we in de cijfertekst voor de letter a de letter (het symbool) X. *Welk* symbool we ter vervanging van de letters a, b, c, … kiezen maakt voor het substitutiecijfer niet uit. Met andere woorden, een geheimschrift waarbij we de letters van ons alfabet door symbolen vervangen, is gelijkwaardig of isomorf met een geheimschrift waarin we een lettersubstitutie toepassen, ook al lijken de door verschillende substituties verkregen geheimschriften uiterlijk helemaal niet op elkaar.
Bij de substitutie van de letters in de klare tekst door symbolen vervangen we iedere a in de klare tekst door het zelfde symbool. Hetzelfde geldt voor de overige 25 letters. We gaan, bij wijze van spreken, van het ene alfabet over op een ander alfabet dat mogelijk uit andere tekens bestaat. We gebruiken voor de vercijfering dus een ander, gelijkwaardig, alfabet. Zo'n substitutiecijfer noemen we daarom een *monoalfabetisch* substitutiecijfer. Ter vergelijking, we zouden bijvoorbeeld kunnen afspreken dat we de letter a de ene keer vervangen door het symbool ⊕ en een andere keer door ◘. Als we zoiets doen voor iedere letter, dan gebruiken we als het ware twee vercijferalfabetten. Dat maakt de ontcijfering een stukje lastiger. Een dergelijk substitutiecijfer noemen we een *polyalfabetisch* substitutiecijfer. In het vervolg houden we ons maar bij het eenvoudiger monoalfabet.

Een cijfertekst

We vercijferen een boodschap met een monoalfabetisch substitutiecijfer, waarbij we de letters in de klare tekst door andere letters vervangen. Dit levert het onderstaande vercijferde bericht op. Zoals eerder opgemerkt: ter onderscheiding van de klare tekst en de cijfertekst, schrijven we de cijfertekst in hoofdletters. Om de lengte van de woorden in de tekst niet prijs te geven, hebben we de cijfertekst in blokjes van vier letters verdeeld. Dit is bovendien handig voor de verzending van de boodschap en maakt het ontcijferen voor de ontvanger wat overzichtelijker.

DDRO	DBDW	XIQB	EWAY	THYO	DZHI	DDET
WIGV	DDRX	GRGH	CAHZ	DYWI	QBIJ	ZIYW
YJYW	DQWN	ADEF	HRXD	YZDB	JCVL	HRDD
RAED	SJDR	YWDH	RHCM	IDDD	RLGJ	TWOU
GETD	RODF	EHHF	YTDP	DXDY	BGTD	UDEF
VWRB	DYZW	NPGR	TDEO	GDTZ	WNCH	RODY
DFIY	DR					

Uiteraard nodigen we de lezer uit het bovenstaande bericht te ontcijferen. U weet inmiddels dat er 26! sleutels zijn. Als u deze allemaal gaat proberen, duurt het even, maar misschien is het geluk u welgezind en heeft u slechts 1% van de tijd nodig die het zou kosten om ze allemaal te proberen.
Het proberen van alle mogelijke sleutels heeft door het gigantische aantal geen enkele zin. We moeten om het cryptogram te kraken iets beters verzinnen. Een even eenvoudige als succesvolle aanpak bestaat uit een *statistische analyse* van de cijfertekst.
Het idee om via statistische analyse een bericht te ontcijferen dat versleuteld is met een monoalfabetisch substitutiecijfer werd rond 800 ontwikkeld in Arabië. Arabische geleerden bestudeerden de *hadith*, die de dagelijkse uitspraken van de profeet Mohammed bevat. Om vast te stellen of uitspraken daadwerkelijk aan de profeet konden worden toegeschreven, bestudeerden zij de zinsbouw en de afzonderlijke woorden in de teksten, om zo na te gaan of de teksten pasten in het taalpatroon van de profeet. Hierbij ontdekten zij dat in de bestudeerde teksten niet alleen bepaalde woorden meer voorkwamen dan andere, maar dat dit ook op het niveau van de afzonderlijke letters gold. Zo komen in het Arabisch de letter a en de letter l aanzienlijk vaker voor dan de andere letters. Op basis van een groot aantal teksten kon men nu de relatieve letterfrequenties bepalen. Hoewel het niet goed mogelijk is deze ontdekking toe te schrijven aan een bepaalde geleerde, heeft de verhandeling van al-Kindi (*Een manuscript over het ontcijferen van cryptografische berichten*) waarschijnlijk een grote invloed gehad op de Arabische crypto-analysten. Het belangrijkste idee uit dit werk, vertaald naar het Nederlands, is dat in onze taal de letters e en n aanzienlijk vaker voorkomen dan de andere letters. Als we nu de letters in een klare tekst vervangen door andere letters of symbolen dan zullen in de cijfertekst de symbolen die voor e en n staan het meeste voorkomen. Dit geeft een simpele mogelijkheid om op zijn minst twee symbolen of letters in de cijfertekst te identificeren. Deze *frequentieanalyse* kan worden uitgebreid tot de andere letters. In de onderstaande tabel vinden we de relatieve frequentie van de letters in het Nederlands. (Als we vermoeden dat de klare tekst in het Engels geschreven is, dan moeten we natuurlijk een frequentietabel van de letters in de Engelse taal gebruiken.) We gebruiken een tabel met relatieve frequenties omdat de absolute frequenties uiteraard afhangen van de lengte van de tekst. In de tabel is de letter ij opgevat als de twee afzonderlijke letters i en j.

relatieve frequentie van de letters in het Nederlands
bron: Genootschap Onze Taal (onderzoek PWT, 1985)

letter	%	letter	%	letter	%	letter	%
a	7,5	h	2,4	o	6,1	v	2,9
b	1,6	i	6,5	p	1,6	w	1,5
c	1,2	j	1,5	q	0,0	x	0,0
d	5,9	k	2,3	r	6,4	y	0,0
e	18,9	l	3,6	s	3,7	z	1,4
f	0,8	m	2,2	t	6,8		
g	3,4	n	10,0	u	2,0		

de meest voorkomende letters zijn e, n a, t, i, r

We zien, zoals we met onze kennis van de Nederlandse taal ook wel weten, dat de letters e en n het meest voorkomen en dat we de letters x, y en q zelden aantreffen in een tekst. In een lange tekst zullen we de bovenstaande verdeling ongeveer terugvinden. Als we iedere letter in zo'n tekst door een andere letter of een symbool vervangen, dan zal dezelfde frequentieverdeling optreden voor de vervangende letters of symbolen. Omgekeerd geldt dit ook, als we een frequentietabel opstellen van de letters of symbolen in een cijfertekst, dan zal deze een beeld laten zien dat ongeveer gelijk is aan de verdeling in de bovenstaande tabel. Een waarschuwing is hier echter op zijn plaats. De frequentieverdeling van de letters is gebaseerd op het turven van de letters in een zeer groot aantal uiteenlopende teksten (boeken, krantenartikelen etc.). Met name in een kort bericht kan de frequentieverdeling van de letters afwijken van de 'natuurlijke' frequenties in de tabel. Een zinnetje als 'Bob en Bea blijven bij de bovenburen' leidt tot een frequentie van de letter b in de tekst die aanzienlijk groter is dan de natuurlijke frequentie van deze letter. De frequenties in de tabel moeten dan ook worden opgevat als gemiddelden.
Een - min of meer arbitraire - klasse-indeling van de letters geeft een aardig beeld van het voorkomen van de diverse letters:

Nederlandse tekst

rel. frequentie			letters
10,0	$<\cdots$		e, n
6,5	$<\cdots\leq$	10,0	a, i, t
4,5	$<\cdots\leq$	6,5	d, o, r
2,5	$<\cdots\leq$	4,5	l, g, s, v
1,0	$<\cdots\leq$	2,5	b, c, h, j, k, m, p, u, w,z
	$\cdots\leq$	1,0	f, q, x, y

Een letter met een relatieve frequentie van 1,5% in een tekst zal waarschijnlijk één van de letters op de rij b, c, …, z zijn. Mogelijk is het een l of een f, maar de kans dat een dergelijke letter een d, o of r voorstelt is bijzonder klein. Met de bovenstaande tabellen kunnen we nu de aanval openen op de gegeven cijfertekst.

De ontcijfering

We turven het voorkomen van de letters in de cijfertekst en verwerken vervolgens de gegevens in een tabel met relatieve frequenties. Deze tabel geeft het volgende beeld:

relatieve frequentie van de letters in de cijfertekst

letter	%	letter	%	letter	%	letter	%
A	2,3	H	6,3	O	3,4	V	1,1
B	3,4	I	4,6	P	1,1	W	6,9
C	2,3	J	2,9	Q	1,7	X	2,3
D	20,1	K	0,0	R	9,8	Y	9,8
E	4,6	L	1,1	S	0,0	Z	3,4
F	2,9	M	0,6	T	4,6		
G	4,6	N	1,7	U	1,1		

de meest voorkomende letters zijn D, Y, R, W, H

De letters D, R en Y met respectievelijk relatieve frequenties 20,1, 9,8 en 9,8 springen eruit.

figuur 2 Enigma-codeermachine, door Duitsland gebruikt in de Tweede Wereldoorlog

We vermoeden, of eigenlijk weten we bijna zeker, dat de D de letter e voorstelt en dat de R of Y de letter n vervangt. Voor de keuze tussen R of Y onderzoeken we de cijfertekst op combinaties van twee en drie letters. Dergelijke combinaties worden respectievelijk *bigrammen* en *trigrammen* genoemd. Er zijn ook frequentieverdelingen voor bi- en trigrammen, die we voor de ontcijfering kunnen gebruiken. Voor een monoalfabetisch substitutie-ijfer hebben we in het algemeen aan de letterfrequenties genoeg. We zien dat de combinaties DR en DDR vaak in de cijfertekst voorkomen. Hieruit leiden we af dat R waarschijnlijk staat voor de letter n. Immers, 'en' en 'een' zijn veel voorkomende combinaties in het Nederlands. We hebben op basis van de frequentieverdeling nu twee letters geïdentificeerd. Een tipje van de sluier is daarmee opgelicht. Als we in de cijfertekst de D en R door respectievelijk de letters e en n vervangen, dan ziet het raadsel er nog als volgt uit:

eenO	eBeW	XIQB	EWAY	THYO	eZHI	eeET
WIGV	een X	GRGH	CAHZ	eYWI	QBIJ	ZIYW
YJYW	eQWN	AeEF	HnXD	YZeB	JCVL	Hnee
nAEe	SJen	YWeH	nHCM	Ieee	nLGJ	TWOU
GETe	nODF	EHHF	YTeP	eXeY	BGTe	UeEF
VWnB	eYZW	NPGn	TeEO	GeTZ	WNCH	nOeY
eFIY	en					

Een klasse-indeling van de letters in de cijfertekst helpt ons bij het bepalen van overige letters:

Cijfertekst

rel. frequentie			letters
10,0	$<\cdots$		D
6,5	$<\cdots\leq$	10,0	R, Y, W, H
4,5	$<\cdots\leq$	6,5	E, G, H, I, T
2,5	$<\cdots\leq$	4,5	B, F, J, O, Z
1,0	$<\cdots\leq$	2,5	A, C, L, N, P, Q, U, V
	$\cdots\leq$	1,0	K, M, S

De letters Y, W en H komen in de cijfertekst veel voor. Vergelijking met de natuurlijke frequenties laat zien dat deze letters waarschijnlijk een a, i, of t voorstellen. We kunnen verschillende mogelijkheden proberen maar het gaat sneller als we cijfertekst er nog even bijpakken. Het bigram Hn komt een aantal keren voor. Op de derde rij zelfs aan het eind van een woord. We kunnen dus uitsluiten dat H staat voor r of t. Het bigram HH doet ons besluiten dat H de letter a voorstelt. Blijft voor de Y en W nog over i of t. Passen we deze letters in de bigrammen YW en WY alsmede in het trigram YWY dan lijkt Y = t en W = i de verstandigste keuze. We kijken even waartoe dit leidt:

eenO	eBei	XIQB	EiAt	TatO	eZaI	eeET
iIGV	eenX	GnGa	CAaZ	etiI	QBIJ	ZIti
tJti	eQiN	AeEF	anXe	tZeB	JCVL	anee
nAEe	SJen	tiea	naCM	Ieee	nLGJ	TiOU
GETe	nOeF	EaaF	tTeP	eXet	BGTe	UeEF
VinB	etZi	NPGn	TeEO	GeTZ	iNCa	nOet
eFIt	en					

We zijn al een aardig eind op weg. Vanaf hier kunnen we de frequentietabellen gebruiken om nog meer letters te vinden, maar er zijn inmiddels meerdere wegen die naar Rome leiden. De redundantie in de taal stelt ons in staat woorden te herkennen ondanks een foute spelling of het ontbreken van letters. Denk hierbij bijvoorbeeld aan de televisiequiz *Twee voor Twaalf*, waarin het teams lukt om woorden te raden waarvan slechts 5 van de 12 letters gegeven zijn. We kiezen deze weg maar de lezer kan natuurlijk zijn eigen idee volgen. Op de vierde regel komt een woord voor dat 'gespeld' is als AeeSJentie … Hierin zijn 6 van de 10 letters bekend, dat moet dus wel lukken. Bovendien vertelt de frequentietabel ons dat de S een zeldzaam voorkomende letter is, dus waarschijnlijk f, q, x of y. Ten slotte gebruiken we nog een belangrijk wapen om in te breken. Vaak weten we globaal waar een tekst over gaat. Er zijn dan een aantal woorden die meer passen bij de tekst dan andere. Een vercijferd weerbericht bevat waarschijnlijk woorden als wind(kracht), graden, regen etc, terwijl een aanvalsplan woorden voor data en tijd zal bevatten. We kunnen dan zoeken op de structuur van die woorden. Deze methode was een niet onbelangrijk wapen in de cryptoanalyse van de Enigma-code in de Tweede Wereldoorlog (***zie figuur 2***). Het cryptogram AEeSJentie ontcijferen we met deze kennis als het woord frequentie. Als dit klopt, dan hebben we gevonden dat A = f, E = r, S = q en J = u. Dit blijkt redelijk overeen te komen met de frequentietabel. Invullen van deze letters in de cijfertekst geeft:

eenO	eBei	XIQB	rift	TatO	eZaI	eerT
iIGV	eenX	GnGa	CfaZ	etiI	QBIJ	ZIti
tJti	eQiN	ferF	anXe	tZeB	JCVL	anee
nfre	quen	tiea	naCM	Ieee	nLGu	TiOU
GrTe	nOeF	raaF	tTeP	eXet	BGTe	UerF
VinB	etZi	NPGn	TerO	GeTZ	iNCa	nOet
eFIt	en					

We gaan ervan uit dat de lezer vanaf hier zijn weg wel zal weten te vinden.

Einde van het substitutiecijfer

Een substitutiecijfer heeft een gigantisch aantal sleutels maar, zoals we boven hebben gezien, dat betekent geenszins dat de code moeilijk te kraken is. Met een statistische analyse is de klus binnen een half uurtje geklaard (met een geschikt computerprogramma zelfs binnen een minuut). De ontdekking van deze methode betekende uiteraard het einde van het monoalfabetisch substitutiecijfer. De crypto-analysten hadden de slag gewonnen, waarna de cryptografen op zoek moesten naar een beter cijfer. Een nieuw cijfer zou uiteraard de natuurlijke frequenties van de letters en combinatie van letters moeten omzeilen. Een aantal nieuwe geheimschriften, die de natuurlijk frequenties van de letters niet meer prijsgaven, werd ontwikkeld. De kern van deze codes is dat een bepaalde letter in de klare tekst niet steeds door hetzelfde symbool of letter wordt vervangen. Deze nieuwe geheimschriften waren lange tijd erg succesvol. Maar uiteindelijk bleken ook deze cijfers niet bestand tegen een slimme statistische aanval. Bij de ontcijfering van codes heeft de statistiek vaak een belangrijke rol gespeeld. Zo bleek in de Tweede Wereldoorlog de statistische analyse een effectief wapen voor de ontcijfering van Duitse en Japanse codes.

Spannende statistiek

Bij de ontcijfering van het monoalfabetisch geheimschrift kwamen we een aantal eenvoudige statistische technieken tegen. Allereerst het *turven* van de diverse letters in de cijfertekst. Daarna het verwerken van

deze gegevens in een *relatieve frequentietabel*. Voor een goed overzicht hebben we vervolgens een *klasse-indeling* gemaakt. Een vergelijking van *frequentieverdelingen* leidde daarna tot de ontcijfering van de cijfertekst.
Als we in een schoolboek het eerste hoofdstuk over statistiek opslaan, dan komen we daarin de boven cursief geschreven woorden tegen. De ontcijfering van een geheimschrift is een mooi en vooral spannend voorbeeld van de toepassing van de elementaire statistiek. In de opgaven in de schoolboeken wordt de leerling gevraagd (relatieve) frequentieverdelingen op te stellen van verzamelingen van meestal oninteressante data. Het doel is dan een overzichtelijke presentatie van de gepresenteerde data. Bij de ontcijfering van het substitutiecijfer is het opstellen en vergelijken van frequentieverdelingen echter een belangrijk middel om een op eerste gezicht moeilijk probleem op te lossen. Het nut van frequentietabellen komt hierbij mijns inziens beter naar voren dan bij veel opgaven in de schoolboeken.

Tot slot

De lezer die het geheimschrift heeft ontcijferd, vond de volgende sleutel voor het substitutiecijfer:

sleutel

a	→	H	h	→	B	o	→	G	v	→	L
b	→	Z	i	→	W	p	→	V	w	→	U
c	→	Q	j	→	N	q	→	S	x	→	K
d	→	T	k	→	F	r	→	E	y	→	M
e	→	D	l	→	X	s	→	I	z	→	P
f	→	A	m	→	X	t	→	Y			
g	→	O	n	→	R	u	→	J			

Literatuur

- Robert Edward Lewand (2000): *Cryptological Mathematics*. Washington: The Mathematical Association of America.
- Simon Singh: *Code, de wedloop tussen makers en brekers van geheime codes en cijferschrift*. Amsterdam: Arbeiderspers (1999).
- Martin Gardner (1972): *Codes, Ciphers and Secret Writing*. New York: Dover Publications.

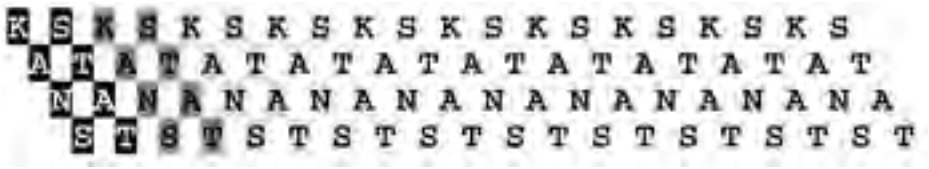
Over de auteur

Rob Bosch is redacteur van *Euclides* en universitair hoofddocent wiskunde aan de Nederlandse Defensie Academie te Breda.
E-mailadres: *robbosch@live.nl*

Hoe krijg je met een zebra een kans van slagen?

STATISTIEK EN KANSREKENING IN DE ZEBRAREEKS

[Rob van Oord]



Ik vind dat je als docent je leerlingen de mogelijkheid moet bieden zelfstandig eens wat dieper in te gaan op een toepassing van de wiskunde. Ook voor de nabije toekomst blijft het mogelijk om dit te blijven doen: in het nieuwe vak wiskunde D. Leerlingen kunnen individueel, in groepjes of klassikaal werken aan keuzeonderwerpen [1].
Door de NVvW is een reeks boekjes met keuzeonderwerpen ontwikkeld met toepassingen van de wiskunde, de zogenoemde Zebrareeks; de naam die is verzonnen door Jan Breeman (†1996) voor het deel in het programma dat in de oorspronkelijke plannen van de tweede fase - toen nog niet ingevuld - als een schuin gestreept blok werd aangegeven.
In dit artikel schets ik in het kort op welke manier mijn leerlingen zebraboekjes gebruiken [2]. Vervolgens bespreek ik wat er aan statistiek en kansrekening te vinden is in deze boekjes. Daarbij vertel ik ook welke opdrachten mijn leerlingen in de regel uitvoeren.

De plaats van het keuzeonderwerp in het PTA

Ik heb het keuzeonderwerp, verplicht deel van het examenprogramma, voor 5% in het PTA van wiskunde B1 en B12 opgenomen als onderdeel in de zesde klas. Ik weet dat er ook veel collega's zijn die dit verwerken in de Praktische Opdracht. Ik heb ervoor gekozen om hiervoor de leerlingen uit een van de boekjes van de Zebrareeks te laten werken. Zij moeten dit boekje eerst zelfstandig doornemen, vervolgens in overleg met mij een eindopdracht kiezen en daarover een presentatie voorbereiden.
Bij ons op school zijn er in de examenklas vier toetsweken. In drie daarvan geef ik een schriftelijke toets over de gewone stof. In één termijn plan ik geen toets, maar reserveer ik die termijn als het ware voor het werken aan het keuzeonderwerp.
Omdat het laten houden van 30 presentaties in de toetsweek onwenselijk en ook bijna onmogelijk in te roosteren is, spreid ik de presentaties over de lesweken in januari en februari. Ik gebruik op die manier een van de wekelijkse lessen voor de presentaties (3 per les) waarbij de (dit jaar halve) klas aanwezig moet zijn. Omdat de groep nu erg groot is, splits ik de klas dit jaar in tweeën en werk ik zelf 10 weken lang een extra uur in het rooster om alle presentaties te kunnen afwerken.
Mijn collega die wiskunde A12 geeft, laat leerlingen in tweetallen een boekje bestuderen en presenteren. De leerlingen met wiskunde A12 kiezen vaker een boekje met A-onderwerpen en vinden het moeilijk om zo'n onderwerp te begrijpen. Ook voor deze leerlingen telt het cijfer voor de presentatie voor 5% mee in het dossiercijfer.

Vaardigheden

Bij de beoordeling let ik vooral op de algemene vaardigheden uit Domein A1 van het examenprogramma [3].

- De leerlingen leren doelgericht de informatie uit het gekozen boekje te selecteren;
- ze leren te beoordelen welke informatie ze later nodig hebben voor de eindopdracht en de presentatie;
- ze leren bij het doorwerken van het boekje te reflecteren op hun eigen belangstelling en motivatie;
- ze leren waar de behandelde wiskunde toepassingen heeft in studie- en/of beroepssituaties;
- ze leren een verband te leggen tussen hun eigen kennis, vaardigheden en belangstelling en de praktijk in een studie en/of beroep;
- ze leren mondeling te communiceren over een wiskundig onderwerp, eventueel met gebruikmaking van (moderne) hulpmiddelen.

Ik vind deze manier van werken voor het keuzeonderwerp een goede voorbereiding op het vervolgonderwijs.
De boekjes zijn zo geschreven dat ze zelfstandig kunnen worden doorgewerkt. Van verschillende vragen en opdrachten staan

opdrachten staan er geen antwoorden in; die zijn geschikt om te gebruiken voor de verwerking. Ik vind het geen bezwaar dat ik zelf een boekje (nog) niet heb bestudeerd; het is dan leuk om te zien wat de leerling er uit haalt. Door de jaren heen weet ik op die manier ook van een aantal boekjes waar het over gaat. Ik vind niet dat de leerling de stof uit zijn hoofd moet kennen, in die zin dat hij er een tentamen over zou moeten kunnen doen. Bij de presentatie blijkt of hij de stof begrepen heeft en er zinnige dingen over kan zeggen. De presentatie moet gericht zijn op de medeleerlingen. Ik vind dat het verhaal ook andere leerlingen enthousiast moet maken voor het onderwerp. Ook moet blijken of er verdieping van de stof heeft plaatsgevonden bij het onderzoeken en uitwerken van de eindopdracht. Eigen inbreng en creativiteit spelen daarbij een rol.
Het cijfer hangt voor het grootste deel af van de presentatie. Criteria zijn: opbouw, enthousiasme, wiskundige inhoud (moet er in zitten, maar hoeft niet van wetenschappelijk niveau te zijn), uitwerking van de eindopdracht, gebruik van hulpmiddelen. Ik maak aantekeningen en probeer een onduidelijk of zwak punt te vinden. Na de presentatie is er tijd voor vragen. Ik stel dan altijd een vraag die over de wiskundige inhoud gaat, op grond van mijn aantekeningen. De leerling levert meteen zijn schrift met gemaakte opgaven en het logboek in. Voor het vaststellen van het definitieve cijfer gebruik ik de informatie die ik hieruit krijg. Het eindcijfer kan maximaal 1 punt afwijken van het cijfer van de presentatie.

Wat moeten de leerlingen doen?

- Eind oktober: keuze van een onderwerp/zebraboekje en intekenen op een datum voor de presentatie;
- november tot februari: 8 tot 12 slu voor het doorwerken van het boekje; antwoorden moeten in een schrift; logboek bijhouden met werkzaamheden;
- januari, februari: 5 tot 8 slu voor de eindopdracht (in overleg vastgesteld) en voorbereiding van de presentatie; uitvoering opdracht in schrift; logboek bijhouden;
- januari, februari: presentatie van max. 15 minuten aan de klas.

Op school heb ik van elk boekje vier exemplaren liggen (van sommige zelfs zes), door de jaren heen aangeschaft van de sectiepot. In oktober leg ik ze allemaal op tafeltjes en vertel ik kort wat over de inhoud.
Ze krijgen een lijst mee met alle beschikbare titels en een week de tijd om een keuze te maken. Inmiddels heb ik dan een lijst met data waarop de presentaties gehouden moeten worden. Degenen die er snel van af willen zijn, tekenen vroeg in, en de leerlingen met uitstelgedrag willen zo laat mogelijk. Voordeel van dit intekenen is dat de beter gemotiveerde leerlingen eerst komen, en dat heeft een positieve invloed op het niveau van de rest - want je kunt niet 'afgaan' voor de hele klas. Ik vind dat de leerlingen wel zelf het boekje dat ze gaan bestuderen, moeten aanschaffen. Het staat daarom ook op de boekenlijst.
Ik weet inmiddels dat de leerlingen in 8 tot 12 slu het boekje zelf kunnen doorwerken. Een enkele leerling komt tussendoor wel eens met een vraag over de inhoud. Ik probeer hem of haar dan met wat suggesties verder te helpen. Uitleg van de stof geef ik niet. Ik vind dat ze zelf keuzes moeten maken hoe ver ze gaan met het snappen van de inhoud. Wanneer het hele boekje is doorgewerkt, wordt in overleg met mij een eindopdracht gekozen. Dit doe ik omdat er soms meer dan één leerling met hetzelfde boekje aan de slag gaat; ik zorg er dan voor dat de eindopdrachten voor de leerlingen verschillend zijn. Het is ook een controle wat er zoal gedaan is.
Tot slot moet elke leerling een presentatie geven waarin kort verteld wordt over de inhoud van het boekje en wat uitgebreider over de eindopdracht. Voor het maken van de eindopdracht en het voorbereiden van de presentatie schat ik 5 tot 8 slu. De ervaring leert dat steeds meer leerlingen een goede PowerPoint-presentatie in elkaar weten te zetten en dat de presentaties beslist de moeite van het aanzien en -horen waard zijn.

Inspiratie door eigen keuze

Ik probeer mijn leerlingen duidelijk te maken dat ze vooral een boekje moeten kiezen over een onderwerp dat ze zelf interessant vinden. Dit kan zijn in verband met hun vervolgstudie, maar ook omdat ze nieuwsgierig zijn naar het onderwerp. Veel leerlingen willen wel eens iets weten van de geschiedenis van de wiskunde, de Babylonische wiskunde of over meetkundige mysteries als Perspectief, Buckyballs, Fractals of Zeepvliezen. Als ze goed gemotiveerd zijn voor hun keuze, dan levert dat ook wat moois op. Misschien zijn ze wel geboeid voor hun leven, en zullen ze zich vaker laten inspireren door het gekozen item.
Leerlingen die verder gaan in de (dier-) geneeskunde, raad ik het boekje *Kattenaids en Statistiek* aan. Leerlingen met een economische vervolgopleiding beveel ik het boekje over de Poisson-verdeling aan. Voor de leerlingen met belangstelling in de maatschappij is het boekje over Verkiezingen een aanrader. Voor de minder exacte leerlingen is het boekje over Schatten niet al te moeilijk en toch interessant. Degenen met interesse voor computers en programmeren geef ik *Rekenen met kansen* in overweging.

Kritische noot

Collega's zullen zich afvragen of de boekjes niet te moeilijk zijn. Snappen de leerlingen de uitgelegde stof wel echt? Hoe is het niveau van de wiskunde in de spreekbeurt? Hoe kun je het cijfer hard maken met van te voren vastgestelde criteria? Zijn de cijfers vaak niet te hoog, vergeleken met de gewone wiskundetoetsen? In het volgende stukje ga ik hier nader op in.
Ik vind dat de leerling vooral moet proberen te begrijpen waar het boekje over gaat. Er komt geen tentamen waarin ze het geleerde uit hun hoofd moeten kunnen ophoesten of toepassen. Ze moeten het op een begrijpelijke manier kunnen uitleggen aan de klas. Omdat ze niet willen afgaan voor de groep, doen ze hun uiterste best om het zo goed mogelijk te begrijpen en uit te leggen. Ze mogen de hulp van de PowerPoint gebruiken bij de presentatie of de formules op het bord schrijven. Ook mogen ze een stencil gebruiken met gegevens of formules of een tekening (in perspectief). Ik kopieer dan een en ander voordat de les begint voor de hele klas. Wie van de leerling in een presentatie verwacht dat er een soort college komt met wiskundige uiteenzettingen, komt bedrogen uit. Het begrijpelijk reproduceren van wat er in het boekje staat, vind ik al moeilijk genoeg. Als ze dat in hun presentatie laten zien, ben ik al dik tevreden.
En nu het cijfer. Natuurlijk heb je geen harde cijfers. Je kent de leerlingen al jaren. Je weet hoe ze werken, wie altijd alles op tijd af heeft, wie slim is maar slordig, wie er echt veel tijd aan besteedt. Vanaf het moment dat je de boekjes uitdeelt, houd je een beetje in de gaten hoe ze er mee bezig zijn. Er zijn altijd leerlingen die meteen aan de slag gaan. Van hen hoor je in de les hoe ze het vinden, wat ze moeilijk vinden, waar ze wat over zouden willen weten. Er zijn ook leerlingen waarvan

Uit het logboek van Ellis Heijboer, maart 2007

14-03 (1½ uur) - Vandaag de opdrachten van 3.3 gemaakt. Dit was moeilijker dan ik dacht. Daarom heb ik advies aan Shirley en Roos gevraagd. Die hebben namelijk hetzelfde boekje en hadden de presentatie al eerder dan ik. Dankzij hun uitleg begreep ik het nu wel en kon ik verder met de opgaven. Omdat er van deze opdrachten geen antwoorden achterin stonden, was dit echt een test of ik het allemaal wel begreep. Voor zover ik weet is alles goed gelukt.

15-03 (1½ uur) - Vandaag begonnen aan het Verjaardagsprobleem. Dit is hoofdstuk 4. Ik vond dit trouwens ook erg interessant. Je kijkt echt naar vraagstellingen die veel dichter bij mij liggen dan zoiets als Pruissen. Ook heb ik het bijna-verjaardagsprobleem doorgenomen en na het oefenen van de voorbeelden snap ik hoe het in elkaar zit.

16-03 (1 uur) - Omdat ik de uitleg goed begreep, ging het maken van de opdrachten soepel en vlot. Ik heb er geen problemen mee gehad om deze sommen te maken. Je past ze in feite toe op de formule van Poisson.

18-03 (¾ uur) - Eindopdracht De Lotto a en b gemaakt. Het is een praktisch stuk aangezien men maar al te graag wil weten wat hun winstkans is bij de lotto. Ook werden er tips gegeven om de kans op een hogere prijs te verbeteren. Al met al vond ik dit gedeelte erg interessant.

20-03 (3¼ uur) - Begonnen met het maken van de PowerPoint presentatie. Aangezien ik hier nog niet zo veel ervaring mee heb, was het behoorlijk tijdrovend. Ook vond ik het moeilijk om te bedenken wat ik aan de klas zou kunnen vertellen en om het zo begrijpelijk mogelijk te maken. Ik heb besloten om van elk onderdeel een korte uitleg te geven en een voorbeeld + uitleg. Dit vond ik het meest overzichtelijk.

21-03 (1¾ uur)- Vandaag heb ik geoefend met mijn presentatie voor mijn ouders. Ook heb ik nog enkele punten toegevoegd die ik was vergeten. Ik ben ervan overtuigd dat ik me genoeg in het onderwerp heb verdiept om het begrijpelijk uit te leggen aan de klas.

je weet dat ze pas in het laatste weekend het hele boekje doen, en dan nog gauw de eindopdracht. Het is een kwestie van regelmatig rondlopen door de klas en vragen stellen: of ze al begonnen zijn, of ze het leuk vinden, of ze al een idee hebben waarover ze de eindopdracht willen gaan doen. In het cijfer weeg ik intuïtief al deze dingen ook mee, en de klas zal dit ook aanvoelen als ik een cijfer geef. De echt goede leerling die er veel aan gedaan heeft en een 9 krijgt, is ook bekend bij de medeleerlingen. Ook degene die eigenlijk te laat is begonnen en een 6½ of 7 krijgt zal er vrede mee hebben. Zelf heb ik me alleen een aantal malen verbaasd over leerlingen met een geweldige presentatie die ik niet verwacht had. Maar ik geef dan toch het hoge cijfer. Een enkele keer ben ik teruggekomen op een lager cijfer dan door de leerling verwacht. Meestal begin je bij de eerste presentaties niet te hoog te cijferen.
Ik vind dat je als docent wiskunde je moet bekwamen in het geven van een cijfer voor werkstukken en presentaties. Dit kan alleen door het te doen. Je moet er een keer mee beginnen. En na een paar jaar ervaring kom je tot de ontdekking hoe je een eerlijk cijfer kunt geven voor dit soort moeilijk te beoordelen werk. Als houvast neem je de schriften mee naar huis en houd je waar mogelijk de tussentijdse voortgang in de gaten.

Natuurlijk zullen er nauwelijks onvoldoendes vallen en behoorlijk wat 'hoge' cijfers. Maar dat mag ook best als je ziet dat ze er hun best voor gedaan hebben. De toetsen met analyse zorgen vanzelf voor een reëel eindcijfer van het schoolexamen voor het vak. Ik zit desondanks vaak nog onder het cijfer voor het centraal examen. Maar ik ga dan ook voor goed wiskundeonderwijs.

Boekjes met statistiek en kansrekening

Statistiek en kansrekening zijn te vinden in de Zebradeeltjes 1, 3, 5, 8, 17 en 23. Van elk deeltje vertel ik kort iets over de benodigde voorkennis, de inhoud en mogelijke eindopdrachten.

1 Kattenaids en Statistiek (*J. van den Broek, P. Kop*)
Voorkennis: Toetsen van hypothesen, dus geschikt aan het eind van klas 6-vwo.
Dit boekje gaat over toetsing van hypothesen. Eerst wordt besproken wanneer een binomiale stochast kan worden benaderd door een normaal verdeelde stochast. De verwachtingswaarde van de steekproeffractie wordt gekoppeld aan de standaardfout. Ten slotte wordt uitgelegd hoe je twee populatiefracties kunt vergelijken en kunt toetsen op significantie.

In een van de eindopdrachten laat een leerling zien hoe je kattenaids van twee groepen asielkatten, de ingebrachte zwerfkatten en de van huis ingebrachte katten, op significantie kunt toetsen. Een andere leerling laat zien of het voorkomen van infectie aan de urinewegen bij mensen met een lage dan wel hoge immuunstatus significant verschilt. Nog een leerling onderzocht hoe met behulp van logistische regressie het voorkomen van kattenaids gerelateerd wordt aan meerdere kenmerken. Hoe logistische regressie werkt heeft ze zelf op internet moeten opzoeken. Weer een ander meisje onderzocht verschillende manieren van het voeren van varkens. Ze heeft daarvoor zelf gegevens bij de universiteit van Wageningen opgevraagd. Stuk voor stuk leerzame presentaties, waarbij het mij vooral duidelijk werd dat de leerlingen goed begrepen hadden wat ze over hypothesetoetsing geleerd hadden. Het boekje sluit goed aan bij de stof die ze in de les hebben gehad.
Aantal slu (genoteerd): -, 22, 14½, 19½.
Behaalde cijfers: 7½, 8, 8, 8.

3 Schatten, hoe doe je dat? (*W. Kremers, J. Smit*)
Voorkennis: kansverdeling, verwachtingswaarde, betrouwbaarheid, dus geschikt aan het eind van klas 6-vwo.

Dit boekje gaat over het bepalen van een schatting van de grootte van een populatie op grond van een steekproef. Aan de orde komen de begrippen gemiddelde, permutaties, combinaties, hypergeometrische verdeling, binomiale verdeling, verwachtingswaarde, standaardafwijking, variantie en betrouwbaarheid.
De eindopdrachten gaan over het schatten van de omvang van een populatie dieren aan de hand van gemerkte steekproeven. Het aantal gemerkte dieren in een tweede vangst wordt vergeleken met het aantal dieren dat bij de eerste vangst gemerkt werd, ook wel de vang–merk–terugvang–methode genoemd. Er kan slechts een uitspraak gedaan worden over bijvoorbeeld het 95%-betrouwbaarheidsinterval waarbinnen de schatting ligt. Bij dit boekje staan (helaas) wel alle antwoorden achterin.
Je zou de leerlingen kunnen opdragen zelf een onderzoek te laten doen volgens deze methode.
Aantal slu (genoteerd): 13½, -, -, 20.
Cijfers: 7, 7½, 8, 8½.

5 Poisson, de Pruisen en de Lotto (*F. Heierman, R. Nobel, H. Tijms*)
Voorkennis: binomiale verdeling, $\binom{n}{k}$-formules, dus geschikt vanaf eind 5-vwo.
Dit boekje gaat over de Poisson-verdeling. Als de stochast X het aantal successen is bij een *zeer groot* aantal (n) onafhankelijke experimenten met een *zeer kleine* succeskans (p), dan is $P(X=k)=e^{-\lambda}\cdot\frac{\lambda^k}{k!}$. Hierin is $\lambda = n \cdot p$ de verwachtingswaarde van de binomiaal verdeelde stochast X. De Poisson-verdeling kreeg bekendheid door een klassieke studie van L. von Bortkiewicz naar het aantal soldaten dat gedood werd door een klap met een paardenhoef.
In hoofdstuk 2 wordt een theoretische afleiding gegeven van de formule. In een volgend hoofdstuk worden enkele historische gegevens vergeleken met de theorie.
In het daarop volgend hoofdstuk zie je hoe de oplossing van het verjaardagsprobleem [zie ook het artikel van Rob Bosch op pag. 223 (red.)] zeer goed benaderd kan worden met de Poisson-verdeling. Dit is de kans dat bij een groep van m personen er minstens twee op dezelfde dag jarig zijn. Je kunt de kans exact uitrekenen, bijvoorbeeld voor (m =) 23 personen. Het benaderen van die kans gaat zo. Bij 23 personen zijn er als het ware $\binom{23}{2}=253$ onafhankelijke deelexperimenten, alle met kans $\frac{1}{365}$, dus $\lambda = 253\cdot\frac{1}{365}=\frac{253}{365}\approx 0{,}69315$
Stochast X is het aantal keer dat er twee mensen op dezelfde dag jarig zijn.
De kans dat er dus minstens twee op dezelfde dag jarig zijn is gelijk aan:
$1-P(X=0)=1-e^{-0{,}69315}\cdot\frac{0^0}{0!}\approx 1-0{,}499999\approx 0{,}5$
Uiteraard is het interessant om te berekenen hoe groot de kans is dat de jackpot valt bij de lotto. Een leerling bestudeerde daarvoor de gegevens van de Duitse lotto. Zij kwam theoretisch op 355 winnaars tegen 339 in werkelijkheid.
In 1995 was er een recordaantal winnaars van de jackpot: 133 van de 69,8 miljoen ingevulde lijstjes, alle met de getallen 7–17–23–32–38–42. Naar aanleiding hiervan liet een leerlinge 30 klasgenoten ieder een zestal getallen opschrijven uit 1 t/m 49. Dit om te kijken welke getallen het meest populair zijn. Door haar is een gedetailleerd logboek ingeleverd (***zie kader***). Dit geeft een beeld hoe leerlingen het werken aan het zebraboekje ervaren.
Aantal slu (genoteerd): 17, -, 12, 19.
Cijfers: 7½, 9, 9, 9.

8 Verkiezingen, een web van paradoxen (*A. van Deemen, E. van der Hout, P. Kop, H.C.M. de Swart*)
Voorkennis: procenten, tabellen, geschikt vanaf begin 5-vwo.
In dit boekje wordt aan de hand van enkele voorbeelden uitgelegd dat verschillende kiesregels tot geheel verschillende uitslagen kunnen leiden. Het is geen moeilijke wiskunde. Je kunt je afvragen of de inhoud valt onder het kopje statistiek. Je moet goed kunnen (op)tellen en gegevens overzichtelijk kunnen noteren.
Het interessante van het verhaal is dat je afhankelijk van de gekozen kiesregel de uitslag kunt beïnvloeden. Door juist anders te stemmen dan je zou willen kun je toch je zin krijgen.
In de vorige twee jaargangen van *Euclides* stonden regelmatig stukjes van Rob Bosch over paradoxale situaties bij stemmingen [rubriek *(Wis)Kundig Kiezen* (red.)]. Ik gaf een van mijn leerlingen de opdracht om een artikel van Rob Bosch te gebruiken bij haar eindopdracht [4]. Zij deed ook opdracht 4.1 waarbij onderzocht moet worden in hoeverre de uitslag van de Tweede Kamerverkiezingen in 1982 anders zou zijn geweest wanneer andere kiessystemen waren gebruikt. Een ander meisje moest uitzoeken of een ander kiesmechanisme een andere uitslag van de Tweede Kamerverkiezingen van 2007 zou geven. Hier kwam ze niet uit. Dit heeft te maken met het feit dat je in Nederland op één naam stemt en geen voorkeurvolgorde kiest. Vervolgens deed ze de eindopdrachten 4.2, 4.2 en 4.3, waarin onderzocht wordt hoe in de Veiligheidsraad van de VN wordt gestemd. Weer een ander meisje onderzocht de verschillen tussen het Nederlandse en Britse kiessysteem.
Aantal slu: 14¼, 12½, 8½.
Cijfers: 8, 8½, 9.

17 Christiaan Huygens (*R. Vermij, H. van Dijk, C. Reus*)
Voorkennis: rijen, lineair verband, elementair kansbegrip, Σ-teken, dus geschikt vanaf begin 5-vwo.
Het boekje over Huygens gaat vooral over zijn levensweg van onderzoeker tot internationaal erkend geleerde.
De vrije val, botsingen en het slingeruurwerk spelen daarbij een grote rol. Op pag. 25 t/m 29 wordt aandacht besteed aan het boekje dat hij schreef over kansrekening: *Van Rekeningh in Spelen van Geluck* [5]. Daarin worden 'voorstellen' (opgaven) beschreven over onder meer dobbelsteenspelletjes.
Er zijn nog geen leerlingen geweest die dit boekje hebben gekozen.

23 Experimenteren met kansen; simulatie met de grafische rekenmachine (*H. Pfaltzgraff*)
Voorkennis: elementaire kansrekening, binomiale verdeling, verwachtingswaarde, normale verdeling, integralen, dus geschikt aan het eind van klas 6-vwo.
In het boekje worden allerlei simulaties behandeld die op de grafische rekenmachine (TI 83Plus) kunnen worden uitgevoerd. Het behandelt hoe je met eenvoudige programmaatjes leuke resultaten kunt boeken. Bijvoorbeeld bij het probleem van de spelshow met de drie deuren, lootjes trekken voor Sinterklaas, benadering van π door te gooien met dart-pijlen, het meningsverschil tussen Pascal en Chevalier de Méré over de kans op minstens één zes gooien met 4 dobbelstenen.
Vorig jaar heeft een leerling dit boekje gekozen. Hij heeft enkele van de eindopdrachten die in het boekje staan gemaakt. Via een viewscreen van de TI kon de klas meekijken naar de resultaten.
Aantal slu: 12.
Cijfer: 7.

Noten

[1] *Domein Ga* resp. *Fb: Keuze-onderwerpen*. Dit domein omvat een of meer keuze-onderwerpen. De onderwerpen worden gekozen door de school. De onderwerpen kunnen, indien de school daarvoor kiest, voor elke kandidaat verschillend zijn. De totale studielast van de keuze-onderwerpen is 40 uur.

[2] Ik schreef al eerder over de tweede fase en het zebrablok: *De tweede fase en het zebrablok*; in: *Euclides* 78(3), pp. 86-89.

[3] *Domein A: Vaardigheden (oude tweede fase) / Subdomein A1: Informatievaardigheden*
1. De kandidaat kan, mede met behulp van ICT, informatie verwerven, selecteren, verwerken, beoordelen en presenteren.
Domein A: Vaardigheden (vernieuwde tweede fase) / Subdomein A2: Algemene vaardigheden
1. Informatievaardigheden - De kandidaat kan doelgericht informatie zoeken, beoordelen, selecteren en verwerken.
2. Communiceren - De kandidaat kan adequaat schriftelijk, mondeling en digitaal in het publieke domein communiceren over onderwerpen uit de wiskunde.
3. Reflecteren op leren - De kandidaat kan bij het verwerven van vakkennis en vakvaardigheden reflecteren op eigen belangstelling, motivatie en leerproces.
4. Studie en beroep - De kandidaat kan toepassingen en effecten van wiskunde en natuurwetenschappen in verschillende studie- en beroepssituaties herkennen en benoemen en een verband leggen tussen de praktijk van deze studies en beroepen en de eigen kennis, vaardigheden en belangstelling.

[4] Zie *Euclides* 82(3), pp. 103-104: *De Condorcet Paradox* (Rob Bosch).

[5] Huygens' *Van Rekeningh in Spelen van Geluck* is ook afzonderlijk uitgebracht bij Epsilon. Het boekje werd vertaald en toegelicht door Wim Kleijne.
Epsilon-uitgave nr. 40, Utrecht 1998, ISBN 90-5041-047-2.

Illustraties

Jeannette van der Kleij

Over de auteur

Rob van Oord is sinds 1974 docent wiskunde aan het Coenecoopcollege te Waddinxveen. Hij is op verschillende plaatsen betrokken bij het wiskundeonderwijs, onder meer als mede-auteur (samen met H. Melissen) van Zebraboekje 11(*Schuiven met auto's, munten en bollen*), als voorzitter van de werkgroep havo-vwo van de NVvW, als lid van de commissie *Wiskunde B-dag* en als lid van de kerngroep wiskunde D voor de module *Wiskunde in Wetenschap* bij de TU Delft.
E-mailadres: *robvanoord@tiscali.nl*

Kansen in de weersverwachting

[Kees Kok en Daan Vogelezang]

Wat kun je verwachten als de weerman of -vrouw 's avonds op tv aankondigt dat er de volgende dag een 'grote kans op neerslag' is? Neem je dan een paraplu mee of een regenpak? Ga je met de bus in plaats van op de fiets naar school? Allemaal vragen die je je zou kunnen stellen bij zo'n bericht. Wat je je ook kunt afvragen: Is die kans voor mij persoonlijk, en geldt dat dan voor De Bilt of voor heel Nederland, voor de hele dag morgen of alleen een deel daarvan? Een dergelijke mate van detail wordt niet vaak in het weerbericht gegeven. Kijk je daarentegen in het weerbericht in de krant, dan staan daar vaak tabellen die in percentages een kans op neerslag aangeven (30% bijvoorbeeld). Toch blijft dan de vraag staan hoeveel neerslag er zal vallen: 1 mm, 10 mm? En als de kans 90% is, valt er dan ook meer regen dan bij een kans van 30%?
Om wat meer duidelijkheid te scheppen rondom deze vragen zullen we twee methoden beschrijven die op het KNMI gebruikt worden om kansverwachtingen te maken, een fysische en een statistische, maar ook wat je kunt doen met een kansverwachting!

Inleiding

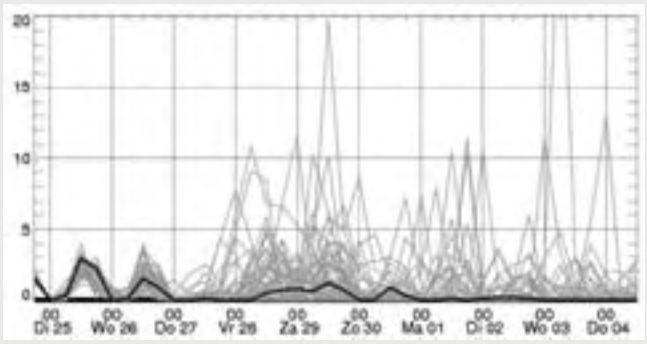
Vanaf het allereerste moment dat er een weersverwachting werd gemaakt, waren meteorologen zich ervan bewust dat al te stellige uitspraken niet waargemaakt konden worden. Hun onzekerheid over de verwachting werd uitgedrukt in kansen. Ook toen in het midden van de vorige eeuw de eerste numerieke weermodellen ontwikkeld werden, bleef er een grote mate van onzekerheid. In deze modellen zijn de natuurwetten vastgelegd die de stroming, de warmte- en de vochtbalans en hun evolutie in de tijd beschrijven. De atmosfeer wordt hierbij opgedeeld in boxen die in de beginjaren typisch in de orde van enkele honderden kilometers lang en breed en ca. 1 km dik waren. Deze noodgedwongen grove oplossing van de atmosfeer leverde naar hedendaagse maatstaven zeer gebrekkige verwachtingen op. Een techniek om deze verwachtingen te verbeteren is *statistische nabewerking*. Bovendien stelde statistische nabewerking ons in staat om de onzekerheid in de verwachting op een objectieve manier te kwantificeren, uitgedrukt in de vorm van kansen. De techniek die hiervoor de afgelopen decennia gebruikt wordt is MOS.

MOS

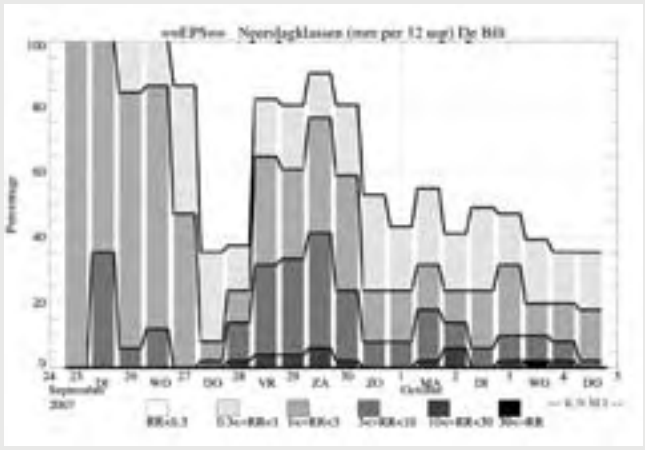
MOS staat voor *Model Output Statistics*. Bij MOS wordt een statistische relatie gezocht tussen een opgetreden gebeurtenis (bijvoorbeeld, meer dan 0,3 mm neerslag in 12 uur tijd in De Bilt) en verklarende variabelen uit een weermodel (modelverwachtingen). MOS kan ook gebruikt worden voor weerelementen die niet door het weermodel berekend worden. Mist bijvoorbeeld wordt niet door weermodellen uitgerekend, maar de kans op mist kan wel bepaald worden met behulp van verwachte temperatuur, vochtigheid, luchtdruk en windsnelheid. De gebeurtenis die we bekijken, is dus binair (treedt wel of niet op). Om de kans op optreden te verklaren met de weers variabelen uit het model gebruiken we de methode van *logistische regressie* (***zie Kader 1 op pag. 206***). De statistische relatie moet bepaald worden op een voldoende grote dataset, meestal zo'n 300 dagelijkse weerberekeningen. Altijd wordt getest of de statistische relatie ook goed werkt op een onafhankelijk deel van de dataset, een deel dat niet is gebruikt bij de afleiding van de relatie.

EPS

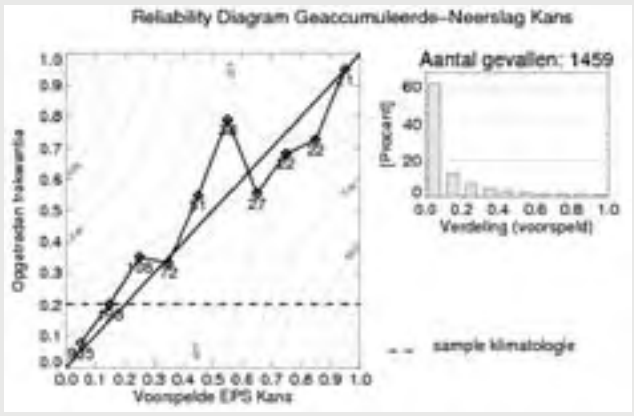
In de loop der jaren nam de rekenkracht van computers explosief toe en werden (en worden) de modellen steeds fijnmaziger (dat wil zeggen de boxen steeds kleiner) en de weersverwachtingen (inclusief de statistisch bewerkte) steeds beter. Maar al in de zestiger jaren van de vorige eeuw begon men zich te realiseren dat de atmosfeer zich nooit exact zal laten voorspellen. Dit komt door het feit dat verstoringen in de atmosfeer, hoe klein ook, soms de neiging hebben te groeien. Dit wordt ook wel het *butterfly effect* genoemd: als een vlinder besluit een bepaalde kant op te vliegen, dan wordt de luchtbeweging op die plek verstoord en dit kan onder bepaalde condities en na een bepaalde termijn bijvoorbeeld een depressie tot gevolg hebben in een totaal ander deel van de wereld (Lorenz, 1979). En aangezien we de vliegintenties van individuele 'vlinders' nooit zullen kennen, is een belangrijke consequentie van deze intrinsieke 'gevoeligheid' van de atmosfeer dat deze niet deterministisch voorspelbaar is. De mate waarin verstoringen in de atmosfeer groeien, hangt o.a. sterk af van het weer in de omgeving van die verstoring. Van deze eigenschap van de atmosfeer is gebruik gemaakt bij de ontwikkeling (rond 1985) van een nieuwe methode om kansverwachtingen te maken: *ensemble verwachtingen*. Het beste en meest gebruikte ensemble-systeem is het EPS, *Ensemble Prediction System*, afkomstig van het Europese Centrum voor Middellange Termijn Weersverwachtingen (ECWMF; zie ook *www.ecmwf.int*).
Hoe gaat EPS in zijn werk? De bovenbeschreven gevoeligheid van de atmosfeer houdt ook in dat voor het maken van verwachtingen met een weermodel de uitgangssituatie erg goed bekend moet zijn. Met allerlei waarnemingen afkomstig van satellieten, weerballonnen, buienradars, grondstations, schepen etc. wordt hiervan zo goed mogelijk een schatting gemaakt, maar een exact beeld tot op de kleinste schalen over de hele aardbol zal nooit lukken. Naast deze 'best mogelijke' schatting wordt ook een schatting gemaakt van de onzekerheid in die schatting. Deze onzekerheid wordt gebruikt om verschillende alternatieve uitgangstoestanden te berekenen die, ieder voor zich, ook prima bij de waarnemingen gepast zouden kunnen hebben. In het EPS worden er 50 van deze zgn. verstoorde begintoestanden berekend die allemaal een even grote waarschijnlijkheid hebben. Met alle (51) begintoestanden wordt een berekening van het weer gemaakt tot 10 (tegenwoordig zelfs 15) dagen vooruit. Wat resulteert noemen we het ensemble van weersverwachtingen, waarbij elk ensemblelid evenveel kans heeft om uit te komen, namelijk (afgerond) 2%.



figuur 1 Neerslagverwachting met EPS. X-as: datum/tijd; Y-as: neerslag in millimeter



figuur 2 Neerslagverwachting met EPS – Stacked Bar diagram. X-as: datum/tijd; Y-as: kans op neerslag



figuur 3 Betrouwbaarheidsdiagram: EPS neerslag > 20mm [+48 : +144] Waterschap Friesland, berekend over de periode dec. 1999-jan. 2004
Linker diagram: de opgetreden frequentie als functie van de verwachte kans, in intervallen van 10% breedte (0-10,11-20 etc.). De nummers in de grafiek geven het aantal gevallen in een interval weer. De gestreepte lijn geeft het opgetreden percentage weer (20%).
Rechter diagram: de verdeling van de verwachte kansen.

In welke mate de berekeningen in de loop van de verwachting divergeren, verschilt van dag tot dag: de ene keer blijven de verwachtingen erg op elkaar lijken en is de atmosfeer blijkbaar goed voorspelbaar, de andere keer lopen de verwachtingen snel uiteen en is de voorspelbaarheid laag.
De stap naar een kansverwachting voor een bepaalde plek of gebied en voor een bepaalde tijd is nu gemakkelijk te maken. Als voorbeeld nemen we een neerslagverwachting voor De Bilt (***zie figuur 1***). De dunne grijze lijnen in figuur 1 tonen de 50 verstoorde leden van het ensemble, de dikke lijn geeft de niet verstoorde berekening weer. Al na een paar dagen ontstaan er verschillen in de verwachte neerslag zowel in hoeveelheid alsook qua timing. Om de informatie kwantitatief handzamer te maken hebben we de neerslagverwachtingen gecategoriseerd in klassen en opgeteld per 12 uursperiode (***zie figuur 2***). Op elk moment in de toekomst kan dan bepaald worden hoeveel ensembleleden er bijvoorbeeld meer dan 0,3 mm (in De Bilt in 12 uur tijd) berekenen. Dat percentage is bijvoorbeeld voor zaterdag 29 september 90%. Tegelijkertijd is de kans op meer dan 3 mm 40%. Op deze wijze is voor een locatie een kans op een bepaalde hoeveelheid neerslag te bepalen.
Uiteraard zijn de (kans)verwachtingen die uit EPS komen, ook weer statistisch na te bewerken om de kwaliteit en de bruikbaarheid te verhogen. Ook is nabewerking (MOS) nodig voor grootheden die niet direct uit EPS komen (zoals bijvoorbeeld de kans op mist of op onweer). Tenslotte kan de meteoroloog de verwachting voordat die naar de afnemer gaat, nog bijsturen (zie hieronder).

Verifiëren van kansverwachtingen

Van een deterministische weersverwachting, bijvoorbeeld 'er valt morgen meer dan 1 mm regen in De Bilt', is goed te bepalen of die is uitgekomen of niet. Bij een kansverwachting 'de kans dat er morgen meer dan 1 mm regen in De Bilt valt is 30 procent' is dat minder makkelijk, maar wel mogelijk. Van één enkele verwachting kun je namelijk niet zeggen of hij goed of fout is. Maar neem je alle verwachtingen bij elkaar die aangaven dat de kans 30% was, dan kijken we in de metingen achteraf of dit in 3 van de 10 gevallen ook is uitgekomen. Hetzelfde doen we voor andere kansen (0%, 10%, 20% etc. tot 100%). De uitkomsten hiervan worden in een grafiek gezet (***zie figuur 3***). In het ideale geval liggen de uitkomsten op de diagonale lijn; in dat geval noemen we de verwachting *reliable* (betrouwbaar). Je weet dan dat als

Kader 1 – Logistische regressie

Logistische regressie is een regressiemethode waarbij de data gefit worden aan een logistische kromme (Brelsford & Jones, 1967) en niet (zoals bij lineaire regressie) aan een rechte lijn.

X = predictor, Y = predictand (voorspelde parameter, bijvoorbeeld kans op neerslag)

Bij kansverwachtingen hebben we meestal te maken met een twee-klassenprobleem: het wel of niet overschrijden van een drempel (bijvoorbeeld hoeveelheid neerslag in een bepaalde periode op een bepaalde plaats). De kans P dat de drempel wordt overschreden, wordt hierbij uitgedrukt als:

$$P = \frac{1}{1+e^{fx}}$$

met $fx = a_0 + a_1 X_1 + a_2 X_2 + \cdots$

De onafhankelijke variabelen (predictoren) X_i worden in principe geselecteerd via de zogenaamde *forward stepwise* selectiemethode. Daarbij worden predictoren op volgorde van significantie aan de vergelijking toegevoegd, totdat niet meer aan een te specificeren significantiecriterium wordt voldaan. De regressiecoëfficiënten a_i worden bepaald met de *maximum likelihood* methode, een iteratieve methode die het product van alle berekende kansen op de opgetreden klassen in de afhankelijke dataset maximaliseert.

het tabelletje in de krant 20% zegt, het 1 op de 5 keer optreedt en 4 keer niet! Je moet er hierbij wel voor waken dat waarnemingen en verwachtingen hetzelfde bedoelen: verwacht je voor heel Nederland, dan moet je ook waarnemingen van heel Nederland gebruiken, verwacht je voor één plek, dan moet je ook daartegen verifiëren.
Overigens is 'betrouwbaarheid' nodig, maar nog niet genoeg. Je wilt graag dat de kansverwachting zo 'scherp' mogelijk is, waarmee wordt bedoeld dat je het liefst kansen richting 0% of 100% wilt. Op dat moment bied je de gebruiker ervan namelijk meer zekerheid dan wanneer je hem een 50/50-verwachting biedt.

De meteoroloog

MOS en EPS zijn automatische systemen die zonder tussenkomst van de mens getallen produceren. De meteoroloog heeft daarmee vergeleken extra informatie, bijvoorbeeld over recente waarnemingen en lokale omstandigheden. Deze informatie gebruikt hij/zij om de kansen bij te stellen (bijvoorbeeld door ze te 'vertalen' naar lokale omstandigheden) en vooral ook door meer 'scherpte' (zekerheid) toe te voegen en tegelijkertijd zo 'betrouwbaar' mogelijk te zijn. Dat is geen makkelijke taak.
Het KNMI is bij wet verplicht de burgers en de maatschappij te waarschuwen voor gevaarlijk weer. Dat gebeurt o.a. door middel van de 'weeralarmen'. Een weeralarm wordt uitgegeven als de meteoroloog ervan overtuigd is dat er een kans van minstens 90% is dat het gevaarlijke weer zich zal aandienen tussen nu en 12 uur later. Je ziet dat er dan nog steeds 10% kans kan zijn dat het niet uitkomt. Maximaal 1 op de 10 keer zal een weeralarm dus onterecht blijken te zijn uitgegeven.

Hoe gebruiken we kansverwachtingen: cost/loss

Hebben we eenmaal een 'scherpe' en 'betrouwbare' kansverwachting, dan rest nog de vraag: Wat doen we er nu mee? Beslissen dus! Op dat moment kom je op een individueel niveau uit: de een zal bij een 50% neerslagkans zijn paraplu meenemen, de ander al bij 25% of pas bij 75%. Deze beslissing hangt grotendeels af van de gevolgen die iemand zal ondervinden als het uiteindelijk wel regent. Hier komt de zogeheten *cost/loss-analyse* om de hoek kijken.
Een bekend voorbeeld hiervan is gladheidsbestrijding (zout strooien) bij kans op nachtvorst. Boven een bepaalde verwachte kans is het economisch nuttig om preventief te strooien. De kosten (*cost*) die daarmee gepaard gaan, wegen dan op tegen de kosten die je ondervindt indien er niet gestrooid zou zijn en er tal van ongelukken, files etc. plaatsvinden (*loss*). Onder die bepaalde kans kun je beter niet strooien. Wél strooien heeft tot gevolg dat uiteindelijk de preventiekosten hoger worden dan de eventuele schade.
Je kunt je ook voorstellen dat een fruitteler dezelfde nachtvorstkans als een wegbeheerder krijgt. De eerste zal overwegen of hij zijn fruitbomen zal beschermen (door besproeien) of het risico nemen dat de oogst verloren gaat. Beiden gebruiken dezelfde verwachting maar zullen er anders mee omgaan. Bij zo'n cost/loss-analyse is

Kader 2 – Het KNMI

Het KNMI (Koninklijk Nederlands Meteorologisch Instituut) is opgericht in 1854 door prof. dr. C.H.D. Buys Ballot en maakt onderdeel uit van het Ministerie van Verkeer en Waterstaat.

Christophorus Henricus Didericus Buys Ballot (1817-1890)

De hoofdvestiging is in De Bilt en er werken in totaal ca. 500 mensen. Het KNMI is hét nationale instituut voor weer, klimaat en seismologie. Er wordt onder meer seismologisch, meteorologisch en klimaatonderzoek verricht. Vanuit de centrale weerkamer in De Bilt verzorgt het KNMI de weersverwachtingen en waarschuwingen voor het algemene publiek, de luchtvaart en de maritieme sector. Sinds 1999 ontplooit het KNMI geen commerciële activiteiten meer. Presentatie van weersverwachtingen via radio, televisie en kranten doet het KNMI sindsdien niet meer. Informatievoorziening via Internet is vervolgens voor het KNMI een belangrijke schakel met de buitenwereld geworden. Meer informatie over de KNMI-organisatie is te vinden via de KNMI-website (*www.knmi.nl*).

het dus cruciaal dat de kansverwachtingen betrouwbaar zijn.

Tot slot

Gezien de intrinsieke onvoorspelbaarheid van het weer is het een goede zaak om de onzekerheid zo veel mogelijk mee te nemen in een weersverwachting. Een voor de hand liggende manier om dat te doen is in de vorm van kansen. Dit bevat veel meer informatie voor de gebruiker dan 'een best mogelijke schatting' van het verwachte weer. De gevoeligheid voor het betreffende weer en daarmee de afweging hoe te reageren, is voor iedere gebruiker namelijk verschillend.
Maar om kwantitatief om te gaan met kansen is het van groot belang om te weten wat de kans precies voorstelt. Dus een antwoord op de vragen: wat, waar en wanneer. Als dat bekend is en de kans is betrouwbaar, dan kan iedere gebruiker voor zich bepalen welke beslissing het meest profijtelijk is.

Literatuur

- E.N. Lorenz: *Predictability: Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?* Address at the annual meeting of the American Association for the advancement of science in Washington (29 december 1972).
- W.M. Brelsford, R.H. Jones (1967): *Estimating probabilities*. In: *Monthly Weather Review*, Vol. 95, pp. 570-576.

Interessante websites

- *www.knmi.nl*
- *www.ecmwf.int*

Over de auteurs

Daan Vogelezang is afgestudeerd in de Meteorologie en Oceanografie aan de Universiteit van Utrecht en is sinds 1993 werkzaam bij het KNMI. De laatste jaren werkt hij als onderzoeksmedewerker in de statistische werkgroep die onderdeel is van de afdeling Weer-Onderzoek. Hij houdt zich met name bezig met gegevensverwerking en onderzoek in relatie tot kansverwachtingen.
E-mailadres: *daan.vogelezang@knmi.nl*
Kees Kok heeft Toegepaste Wiskunde gestudeerd aan de UvA en werkt sinds 1981 op het KNMI. Hij is nu als onderzoeker werkzaam in dezelfde groep als Daan. Hij houdt zich vooral bezig met statistische postprocessing technieken, met name kansverwachtingen en verificatie.
E-mailadres: *kees.kok@knmi.nl*

Data-analyse met behulp van educatieve software

[Clifford Konold / vertaling en bewerking Carel van de Giessen]

In augustus 2007 zijn nieuwe eindtermen voor Statistiek en Kansrekening havo/vwo voorgesteld. In deze voorstellen is het vertrekpunt voor de statistiek het analyseren van data. Voor het nemen van verantwoorde beslissingen is het theoretisch kader van de kansrekening nodig. Interessante probleemgebieden zijn: het vergelijken van groepen, verbanden leggen tussen verschijnselen, het voorspellen van zulke verschijnselen en het nemen van een optimale beslissing in onzekere situaties. Om met deze aanpak vertrouwd te raken zullen leerlingen experimenten doen.
Bijgaande tekst is een vertaling van een artikel van C. Konold. Het oorspronkelijke artikel uit 1990, later ook verschenen in het Duitse blad *Computer und Unterricht* (1995), is met instemming van de auteur vertaald, ingekort en bewerkt voor de huidige Nederlandse situatie. In de tekst staan tussen haakjes toelichtingen die op de Nederlandse situatie slaan. De originele dataset is met VU-Statistiek in het Nederlands omgezet. De vertaler is van mening dat het artikel enkele belangrijke aspecten van de data-analyse op school duidelijk maakt.

Vooraf

Weinig mensen zullen ontkennen dat door de computer het gebruik van statistiek in de praktijk is veranderd. Nog minder mensen zullen ontkennen dat tot op heden de computer weinig invloed heeft gehad op het onderwijs in de statistiek (althans wat het vmbo/havo/vwo betreft). Velen zijn echter van mening dat bij een eerste inleiding in de statistiek de computer geen belangrijke rol mag spelen. Het bezwaar is dat volgens sommigen een inleiding in de statistiek tegelijk met een kennismaking van software onnodig gecompliceerd wordt. Aan veel bezwaren kan tegemoet worden gekomen door verstandig inzetten van educatieve software.

Educatieve software voor data-analyse

Dit artikel gaat over het gebruik van software bij het analyseren van data door leerlingen die nog geen ervaring met data-analyse hebben. De software maakt het mogelijk data te bekijken en te 'bevragen', waardoor leerlingen snel uitkomen bij de essentiële aspecten van data-analyse: Hoe stel je interessante vragen en hoe kom je tot plausibele antwoorden? Educatieve software (zoals VU-Statistiek, Tinkerplots, Fathom, Station) biedt leerlingen een eenvoudige toegang tot een exploratieve analyse van data.

Doelen van data-analyse en educatieve software

Er bestaan verschillende meningen over wat we over statistiek of data-analyse moeten onderwijzen. Sta je voor de keus software voor het onderwijs te gebruiken dan moet je een helder beeld hebben van wat je wilt bereiken. Hier is dat om leerlingen vlot in staat te stellen een aantal samenhangende vragen te stellen met als doel een 'coherent verhaal' bij een dataset te maken. Daarom bevatten de data liefst meerdere variabelen waar wat mee gedaan kan worden. Data-analyse wordt vaak als een interactief en iteratief proces gezien waarin op grond van een vraagstelling relevante data verzameld en onderzocht worden. Vervolgens kan de vraag opnieuw geformuleerd en toegespitst worden, waarna nieuwe data bekeken worden, enzovoort. Educatieve software geeft ondersteuning in het herkennen van patronen in data (de structuur), trends en verschillen, en helpt bij het doordenken van data.

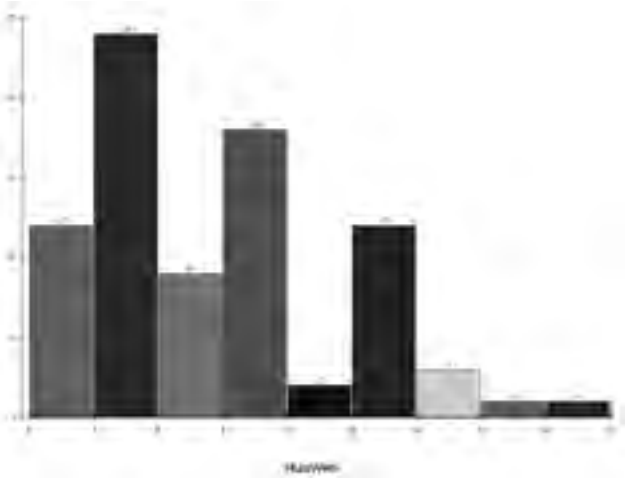
Eenvoud

De software moet geen overdaad aan soorten grafieken bieden. Daar hebben we twee redenen voor. De duidelijkste is omdat de software dan makkelijker te gebruiken is. Hoe minder typen mogelijk zijn, des te minder gecompliceerd het is. Ook valt te denken aan software die weliswaar veel opties biedt maar die afgestemd kan worden op de behoefte van de beginnende leerlingen.
Een belangrijk argument om het aantal grafieken beperkt te houden is dat er tijd nodig is om een grafiek te leren lezen. Voor veel mensen is het lezen van een histogram een tweede natuur geworden. Een expert kan met één blik op een histogram zowel typerende als atypische kenmerken vaststellen en die informatie gebruiken bij verder onderzoek. Voor veel beginners is daarentegen een histogram altijd nog een wirwar van informatie. Ze weten niet waar ze op moeten letten en zien het ongewone niet. Het beperken van het aantal soorten grafieken en diagrammen geeft de leerlingen de mogelijkheid aan elke soort zoveel tijd te besteden dat ze voldoende ervaring opdoen, zodat dit gereedschap een onbewuste uitbreiding van hun gewone waarnemingssysteem wordt.

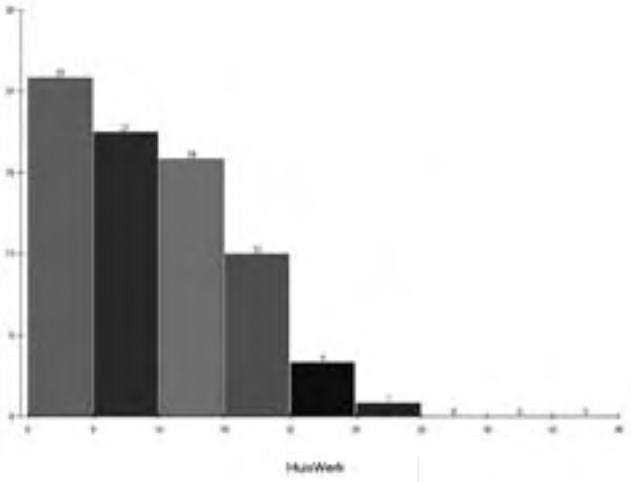
Grafieken die iets vertellen

Een goede schrijver weet aan elk woord betekenis te geven. Cleveland (1993) vraagt iets dergelijks voor het visualiseren van data: 'Wij denken graag dat we relevante informatie opnemen als we veel zien. Het resultaat van een visualisering moet echter louter en alleen afgemeten worden aan hoeveel we over het onderzochte fenomeen te weten komen.' Omdat we bij educatieve software het aantal grafieken willen beperken is het van belang dat de gekozen grafieken 'iets te zeggen hebben'. Zulke grafieken zitten niet vol irrelevante details - wat Tufte (1983) 'chartjunk' heeft genoemd - maar bieden een gelegenheid om relevante kenmerken en relaties in de data vast te stellen. Voor

figuur 1



figuur 2



figuur 3

het weergeven van niet-numerieke data gebruiken we frequentietabellen en staafdiagrammen, voor numerieke data histogrammen, boxplots en spreidingsdiagrammen. Deze worden niet alleen gebruikt vanwege de goede gebruiksmogelijkheden maar ook omdat ze zo vaak voorkomen.
Overigens vinden we dit laatste argument niet doorslaggevend. Cirkeldiagrammen worden in de massamedia heel veel gebruikt, maar ze zijn niet zo geschikt om relevante patronen in de data aan te geven of vergelijkingen te maken. Volgens Tufte is er maar één ding slechter dan een cirkeldiagram: meer cirkeldiagrammen. Omgekeerd zijn er ook grafieken die buiten het terrein van de data-analyse niet veel voorkomen maar toch heel geschikt zijn voor het weergeven van centrum en spreiding en voor het vergelijken van groepen. Dat kunnen redenen zijn om zulke grafieken toch op te nemen in de software.

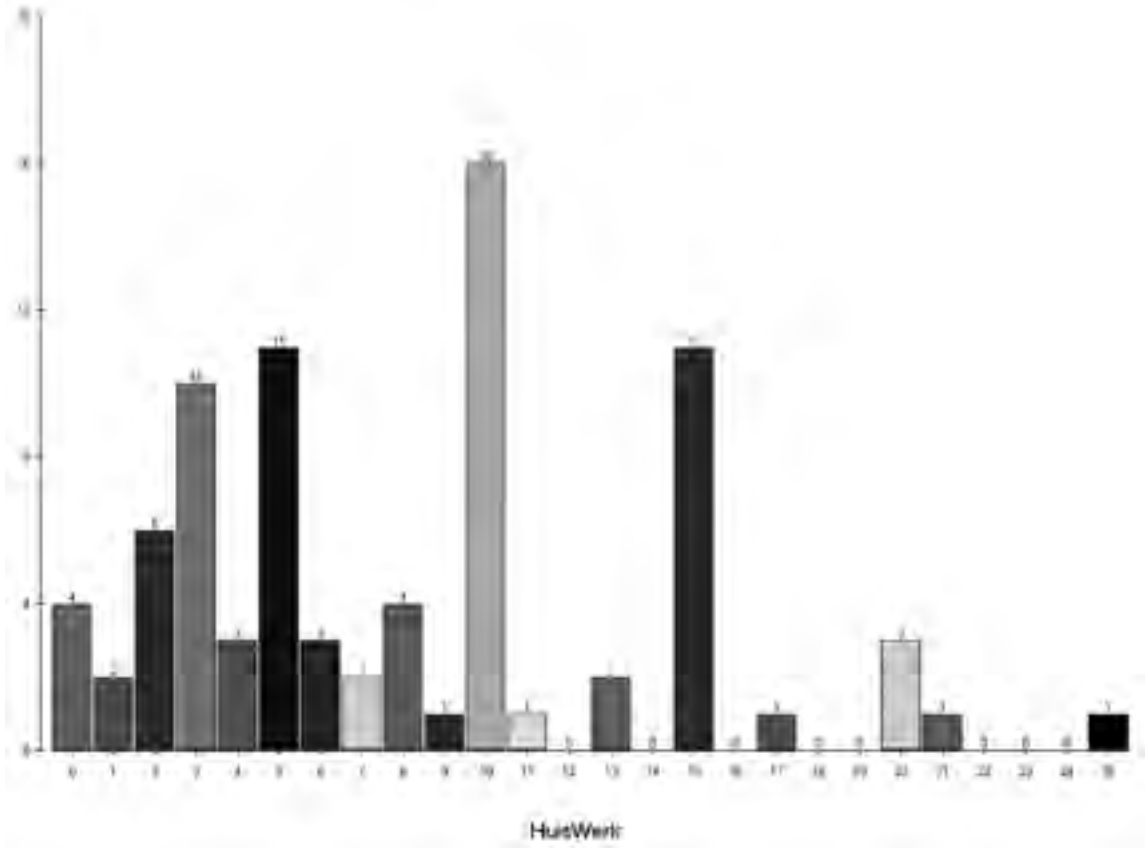
Om te laten zien hoe leerlingen met educatieve software omgaan gebruiken we een dataset van een enquête uit 1990 onder 82 leerlingen in twee steden in Massachusetts in de Verenigde Staten: de kleine stad Amherst waar een universiteit is gevestigd, en de industriestad Holyoke. De anonieme dataset bevat informatie over onder meer sekse, leeftijd, gezinsgrootte, burgerlijke staat van de ouders, godsdienst, schoolcijfers, opleidingsniveau van de ouders. Voor de scores in een datatabel ***zie figuur 1***. Daarin zijn 12 records (van de 82) en 9 variabelen (van de 32) te zien.
Dit is een van de datasets die we in een bepaald schooljaar van de Holyoke Highschool gebruikten. Vragen die de leerlingen op grond van deze data zouden kunnen stellen zijn:

- Moeten meisjes eerder thuis zijn dan jongens als ze uitgaan (stappen)?
- Hangen de opvattingen van een persoon over abortus samen met de godsdienstige overtuiging?
- Voorspelt de plaats in de rij kinderen-van-een-gezin leidinggevende kwaliteiten?
- Hebben kinderen van alleenstaande ouders slechtere schoolresultaten dan kinderen met twee ouders?
- Bestaat er een verband tussen sekse en uurloon van een leerling?
- Heeft een baantje negatieve gevolgen voor de schoolprestaties?

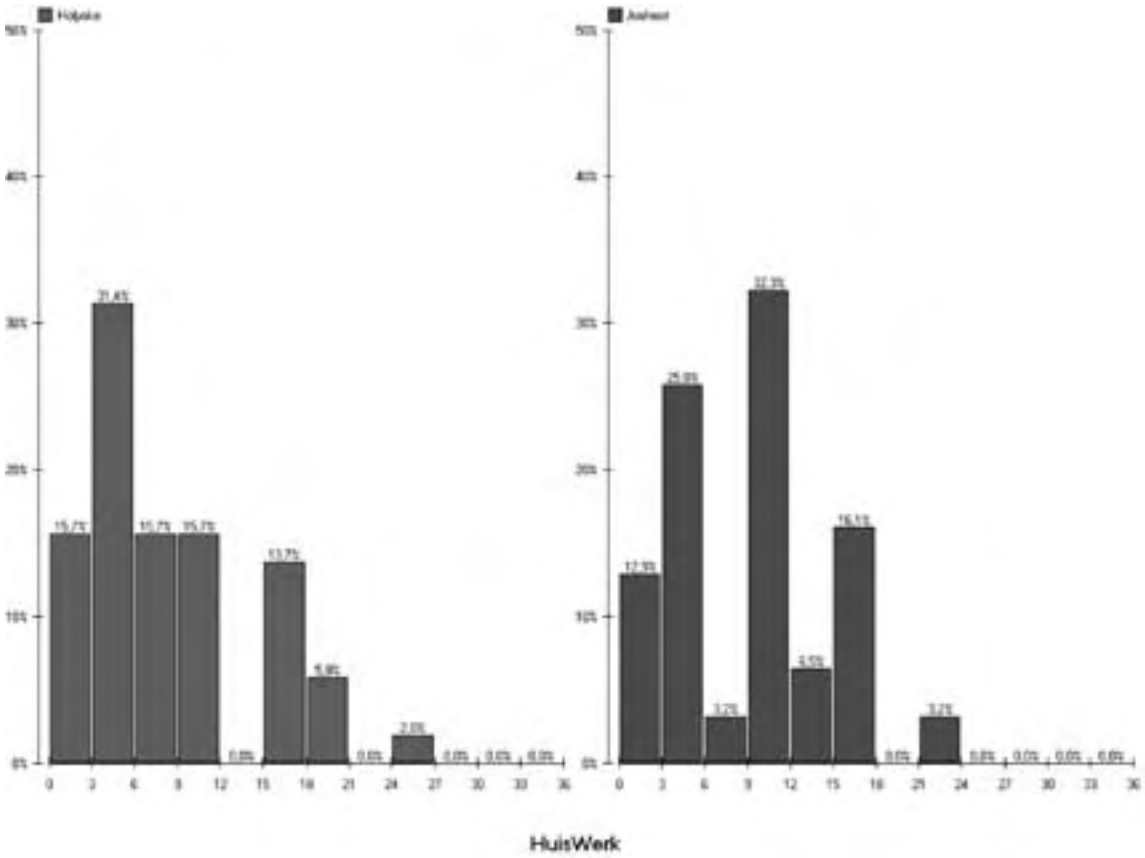
Over deze laatste vraag maken veel ouders zich zorgen. Weliswaar worden bij een vraag als deze slechts weinig variabelen bekeken, maar al gauw worden de andere variabelen in het onderzoek betrokken. Dat is een van de voordelen van het gebruiken van datasets met veel variabelen: ze dagen de leerlingen uit hun vermoedens over mogelijke verklaringen voor de in de data waargenomen trends te formuleren en te testen.

Histogrammen geven een globaal beeld van data

Laten we beginnen de laatste vraag te onderzoeken door het histogram te bekijken van de tijd die leerlingen per week aan huiswerk besteden. Een histogram/staafdiagram (***zie figuur 2***) hoor je eenvoudigweg te krijgen door een optie of knop voor staafdiagram te kiezen en dan de variabele *HuisWerk* te selecteren, zonder dat de gebruiker vooraf verschillende parameters, zoals intervalbreedte, hoeft op te geven.
We zijn van mening dat leerlingen niet gevraagd kan worden om al te beslissen over zaken waarvan ze de gevolgen nog niet kennen. Veel educatieve software kiest daarom de instelling waarbij de data het best getoond worden. Als het diagram er eenmaal staat, heeft de leerling een overzicht en de mogelijkheid om te besluiten of en hoe het diagram veranderd zou moeten worden om meer van de data te laten zien. Door de intervalbreedte aan te passen kan de gebruiker zien hoe de vorm van de verdeling eruitziet bij een grove of een fijne klassenindeling.
De pieken in de data van het *HuisWerk* verdwijnen ***in figuur 3***, de klassenbreedte is veranderd van 3 in 5, waardoor de vorm van de verdeling beter zichtbaar wordt. Kijkend naar steeds hogere aantallen uren



figuur 4



figuur 5

huiswerk zien we steeds minder leerlingen. Een klassenbreedte van 1 laat meer details zien (***zie figuur 4***). Misschien is de ene verdeling die van leerlingen uit Holyoke met pieken in de buurt van 5 en dan afnemend, en de andere van leerlingen uit Amherst met een piek bij 10 en dan afnemend. Als dat inderdaad zo is, zou dat passen bij plaatselijke stereotypen van leerlingen aan beide scholen.

We hopen dat dit het soort veronderstellingen zijn die leerlingen maken als ze de diagrammen bekijken. Het is een van de redenen dat leerlingen data zouden moeten analyseren waarover ze al iets weten. Achtergrondkennis verschaft namelijk een basis om interessante vragen te stellen en ontdekkingen te interpreteren. De software moet het makkelijk maken zulke hypothesen te onderzoeken.
Leerlingen vormen deelgroepen van een variabele (bijvoorbeeld *HuisWerk*) voor elke waarde van een andere variabele (bijvoorbeeld School met waarden Holyoke en Amherst). Door de verdeling van één variabele te vergelijken met die van een andere variabele, kunnen ze verbanden tussen variabelen ontdekken.
In de histogrammen van ***figuur 5*** is de variabele *HuisWerk* gegroepeerd op de variabele *School*. Het resultaat is een tweetal histogrammen, één voor leerlingen uit Holyoke en één voor leerlingen uit Amherst. Merk op dat de verdelingen zijn gelabeld met de waarden van de variabele *School*.
Omdat de aantallen leerlingen in de enquête verschillen, zijn de frequenties niet absoluut maar relatief weergegeven. De optie om frequenties zowel absoluut als relatief weer te geven maakt het gemakkelijker de twee verdelingen op het oog te vergelijken. Misschien zijn histogrammen echter niet de beste keus om deze vraag te onderzoeken, want we krijgen meer informatie uit de details van de histogrammen dan we eigenlijk willen.

Verdelingen vergelijken aan de hand van boxplots

In figuur 6 staan boxplots van dezelfde data als van de histogrammen in ***figuur 5***. Er is een mediaan van 10 te zien voor de leerlingen uit Amherst, ter vergelijking een waarde 6 voor de leerlingen uit Holyoke. Het kruisje boven 25 in het boxplot van Holyoke is een 'uitschieter', een waarde die zo ver buiten de bulk data ligt dat speciale aandacht gerechtvaardigd is. Merk op hoe makkelijk het is om de twee boxplots te vergelijken als ze beide boven een gemeenschappelijke as liggen.

Wat tot nu toe is gedaan, kan beschouwd worden als een voorlopig onderzoek om vertrouwd te raken met de verdeling op afzonderlijke variabelen alvorens in te gaan op de vraag naar verbanden er tussen.
Terug naar de oorspronkelijke vraag om die iets te verfijnen: 'Besteden leerlingen met een baantje minder tijd aan hun schoolwerk dan leerlingen die geen baan hebben?' ***Figuur 7*** is een boxplot van *HuisWerk* gegroepeerd op de variabele *Baan*. Verrassend is dat de 56 leerlingen met een baantje ('ja') een hogere mediaan voor huiswerk hebben dan de 26 die geen baantje hebben ('nee') hebben.

Verhalen vertellen in plaats van bevindingen rapporteren

Het mag verleidelijk zijn te stoppen met nadenken in de veronderstelling dat de vraag beantwoord is. Er zijn echter vele verklaringen mogelijk voor een waargenomen verschil en vele interpretaties. Allereerst zou er een toevallig verschil kunnen bestaan bij het samenstellen van de steekproef. Bij poker kun je altijd wel vijf opeenvolgende kaarten van eenzelfde kleur krijgen, zelfs wanneer de stapel goed is geschud (ongeveer twee op de duizend keer). Iemand die onbekend is met poker en in je hand een 'flush' schoppen ziet, zou kunnen denken dat de hele stapel uit schoppen bestond. Je kunt de onjuistheid hiervan aantonen door de hele stapel te laten zien en uitleggen dat wat er gebeurd is niet erg vaak gebeurt, tenminste niet onder eerlijke spelers. Evenzo, als je kijkt naar de totale populatie van elke school, zou je leerlingen kunnen vinden die dezelfde tijd voor huiswerk opgaven, of ze nu in Holyoke of Amherst woonden en of ze een baan hadden of niet. De vraag is hoe moeilijk het is om een steekproef van de school-'stapels' te trekken en een resultaat te krijgen als we deden. Deze vraag wordt aangeduid met inferentiële statistiek en kan met behulp van random simulatieprocessen worden onderzocht. We zullen deze 'kansverklaring' hier niet onderzoeken, maar het is van belang dat leerlingen hieraan denken als ze data analyseren.

Laten we aannemen dat de verschillen in huiswerktijd in feite karakteristiek zijn voor alle leerlingen in Holoyoke en Amherst. Dan is er nog een groot aantal verklaringen mogelijk voor deze verschillen. Eén van de uitdagingen van het onderwijzen van data-analyse is om leerlingen verder te krijgen dan het rapporteren van bevindingen: het maken van 'verhalen' en testen hoe plausibel die zijn.
Verhalen beschrijven mogelijke verklaringen voor onze waarnemingen en verhalen over

figuur 6

figuur 7

data doen er zeer toe. Hierna volgen enkele simpele verhalen die kunnen verklaren dat leerlingen met een baan de neiging hebben meer te studeren. Deze verhalen kunnen gevolgd worden door het onderzoeken van verbanden in de dataset.

- In Amherst zijn meer leerlingen van plan naar de universiteit te gaan, dus studeren ze harder en werken ze om geld voor hun studie te sparen.
- Sommige leerlingen zijn meer gemotiveerd dan andere en hebben daarom vermoedelijk een baan en zijn ijverig op school.
- Leerlingen die een baan hebben zijn ouder dan leerlingen zonder baan en als leerlingen in hogere klassen zitten krijgen ze meer huiswerk.

Als datasets veel records en veel variabelen bevatten, kunnen verklaringen die leerlingen geven vaak 'getest' worden. Omdat in dit geval de dataset de leeftijd van leerlingen, de school, het plan om te gaan studeren en een score van de motivatie bevat, kunnen leerlingen de bovengenoemde mogelijkheden onderzoeken. Het voert te ver alle genoemde verklaringen hier te onderzoeken.

Samenvatting

Het doel van dit artikel is om te laten zien hoe educatieve software voor data-analyse leerlingen verder kan helpen. Door middel van interessante en steeds moeilijker opdrachten leren ze datasets te onderzoeken en uit brokjes informatie verhalen te maken die de brokjes samenbinden tot een begrijpelijk en overtuigend geheel. Wij geloven dat de computer geen panacee is, en dat het belangrijk is om leerlingen verschillende grafische en statistische vaardigheden ook zonder computer te leren. Er bestaan namelijk genoeg activiteiten die beter zonder computer gedaan kunnen worden. Onze leerlingen tekenen hun eerste boxplot van een bescheiden aantal data met de hand. Maar als de leerlingen de karakteristieke

cyclus van exploratieve data-analyse (de data weergeven in een passende tabel en/of grafiek; de data samenvatten in passende kentallen; nagaan of verwachtingen tot uiting komen in de data, en of deze relevant en significant zijn) gaan doorlopen willen we ze achter de computer hebben. In dit artikel hebben we aangegeven dat leerlingen zelfs met eenvoudige software problemen ondervinden bij data-analyse, en die vragen om nadere doordenking en beproeving.

Noot

Dit artikel is eerder verschenen (in het Duits) in *Computer und Unterricht*, 17, (1995), pp. 42-49.
Vertaling en bewerking: Carel van de Giessen, met dank aan Arthur Bakker.

Literatuur

- W.S. Cleveland (1993): *Visualizing data.* Summit (NJ, USA): Hobart Press.
- E.R. Tufte (1983): *The visual display of quantitative information.* Cheshire (CT, USA): Graphics Press.
- Anne van Streun, Carel van de Giessen (2007): *Een vernieuwd statistiekprogramma. Deel 1: Statistiek leren met "Data-analyse".* In: *Euclides*, 82(5), pp. 176-179.
- Anne van Streun, Carel van de Giessen (2007): *Een vernieuwd statistiekprogramma. Deel 2: Data-analyse, een mogelijke opzet.* In: *Euclides*, 82(6), pp. 217-221.

Over de auteur

Clifford Konold (1949) is werkzaam aan het Scientific Reasoning Research Institute, University of Massachusetts, Amherst USA. Konolds onderzoeksaandacht gaat uit naar de wijze waarop kinderen en volwassenen redeneren over toeval; hij past zijn onderzoeksresultaten toe op het ontwerp van onderwijs en educatieve software. Zo heeft hij onder meer het educatieve statistiekprogramma TinkerPlots ontwikkeld.
E-mailadres: *konold@srri.umass.edu*

Over de vertaler/bewerker

Carel van de Giessen was wiskundeleraar en auteur. Hij is lid van cTWO en van de werkgroep die voorstellen heeft gedaan voor het nieuwe domein *Handelen bij onzekerheid* voor het nieuwe programma wiskunde A en C.
E-mailadres: *carelvdg@planet.nl*

Kansrekening

[Rob Bosch]

Op het eerste gezicht is er niets opwindends aan het volgende sommetje.
Op tafel staan 6 vazen met 5 witte ballen en 1 zwarte bal. We trekken uit iedere vaas 1 bal. Hoe groot is de kans dat we geen enkele zwarte bal trekken?
De kans dat we uit vaas 1 geen zwarte bal trekken is uiteraard $\frac{5}{6}$. Omdat de trekkingen uit de verschillende vazen onafhankelijk zijn, is de kans dat we geen enkele zwarte bal trekken gelijk aan $(\frac{5}{6})^6 \approx 0{,}3349$.
Inderdaad een eenvoudige en weinig opwindende opgave. Een simpele variatie op dit sommetje is:
Op tafel staan 10 vazen met 9 witte ballen en 1 zwarte bal. We trekken uit iedere vaas 1 bal. Hoe groot is de kans dat we geen enkele zwarte bal trekken?
We vinden op dezelfde manier als hierboven dat de kans gelijk is aan $(\frac{9}{10})^{10} \approx 0{,}3387$
De twee kansen zijn, wellicht enigszins verrassend, nagenoeg gelijk. Het maakt blijkbaar niet zoveel uit hoeveel vazen we op tafel zetten als we in iedere vaas maar een aantal ballen (waaronder 1 zwarte) stoppen dat gelijk is aan het aantal vazen. Hoe zit dit?

We nemen nu n vazen met $(n-1)$ witte ballen en 1 zwarte bal. We trekken uit iedere vaas 1 bal. Hoe groot is de kans dat we geen enkele zwarte bal trekken?
De kans dat we uit een vaas niet de zwarte bal trekken is $\frac{n-1}{n}$. De kans dat we uit geen enkele vaas de zwarte bal trekken, is dus: $(\frac{n-1}{n})^n$, anders geschreven: $(1-\frac{1}{n})^n$
Maar deze uitdrukking kennen we in de gedaante:

$$\lim_{n\to\infty}(1-\tfrac{1}{n})^n = (1+\tfrac{-1}{n})^n = e^{-1} = \tfrac{1}{e}$$

Voor een groot aantal vazen is de gevraagde kans dus ongeveer gelijk aan $\frac{1}{e}$.
Ter illustratie enkele numerieke waarden:

n	P
2	0,2500
6	0,3349
10	0,3387
20	0,3585
40	0,3632
100	0,3660
	0,3679 (=1/*e*)

De lezer heeft de kans $\frac{1}{e}$ wellicht vaker gezien. Ja, bij de Sinterklaasloterij, waarbij we de kans willen weten dat niemand zijn eigen lootje trekt; deze kans blijkt bij een grote groep Sinterklaasvierders ook gelijk te zijn aan $\frac{1}{e}$.[1]

Een opmerking. Bij de limiet $\lim_{n\to\infty}(1-\frac{1}{n})^n$ treffen we nogal eens de volgende onjuiste redenering aan:

$\lim_{n\to\infty}(1-\frac{1}{n}) = 1$ en $1^n = 1$, dus
$\lim_{n\to\infty}(1-\frac{1}{n})^n = 1$

Het kanssommetje toont aardig aan waar de schoen wringt: weliswaar wordt de kans op een zwarte bal per trekking heel klein, maar daar staat een groot aantal vazen tegenover waaruit we moeten trekken, waardoor de kans op een zwarte bal toch niet 0 wordt.

Bovenstaande opgave heeft een duidelijk relatie met de Poisson-verdeling. De lezer kan dat nagaan in het vijfde Zebraboekje: *Poisson, de Pruisen en de Lotto*.[2]

Noten (red.)

[1] Zie ook:
Hans Blom: *Sint en de letter e*. In: *Euclides* 75(5), 2001; pp. 212-213.

[2] Henk Tijms, Frank Heierman, Rein Nobel (2000): *Poisson, de Pruisen en de Lotto*. Utrecht: Epsilon Uitgaven.

Over de auteur

Rob Bosch is redacteur van *Euclides* en universitair hoofddocent wiskunde aan de Nederlandse Defensie Academie te Breda.
E-mailadres: *robbosch@live.nl*

Vergelijkend atletiek-onderzoek in de tweede klas

EEN PROJECT LO & STATISTIEK

[Peter Tilman en Klaske Blom]

Inleiding

Sinds een paar jaar streven we er op 't Hooghe Landt naar om het aantal verschillende docentengezichten in de onderbouwklassen te beperken. Docenten die wel aan deze klassen lesgeven, geven meerdere vakken in de vorm van leergebieden. Deze docenten zien de klas dus vaker dan wanneer ze één vak zouden geven. We proberen hiermee de overgang met de basisschool te verkleinen. Bovendien, en belangrijker, denken we op deze manier meer diepgang in de stof te kunnen creëren. Een voorbeeld hiervan is de docent in de onderbouw die zowel aardrijkskunde als geschiedenis geeft in het leergebied M&M, mens en maatschappij. Hij kan veel beter dwarsverbanden aanbrengen tussen de 'oude' vakken omdat hij weet wat op welk moment behandeld wordt. In de derde klas halen we de leergebieden weer uit elkaar om de aansluiting met de bovenbouw beter te laten verlopen; in de bovenbouw zijn namelijk geen leergebieden, dus moeten leerlingen wennen aan de gescheiden vakken.
Naast *diepgang* willen we leerlingen in de leergebieden laten ervaren dat vakken die ogenschijnlijk ver uit elkaar liggen, ook *samenhang* kunnen vertonen. Er is voor gekozen om projecten te ontwikkelen waarbij twee onverwachte vakken samen optrekken om de dwarsverbanden aan te brengen. Voorbeelden hiervan zijn wiskunde en lichamelijke opvoeding met een project statistiek; natuur & gezondheid en beeldende vorming met een project 'uitgestorven dierrassen'; het leergebied economie en scheikunde is nog in ontwikkeling. Tijdens de gekoppelde vakken werken leerlingen aan eenzelfde opdracht. Het is de bedoeling dat het eindwerkstuk dat als afsluiting van een project gemaakt wordt, meerwaarde heeft boven 'werken uit het boek'.

Wij zijn geen 'nieuwe leren' school, maar op kleine schaal gebruiken we de didactische aanpak die dit leren kenmerkt. In dit geval is dat: eigen onderzoek doen om te kunnen werken met eigen data. Het project dat we hier beschrijven is een samenwerkingsverband tussen lo (lichamelijke opvoeding) en wiskunde in het tweede leerjaar, zowel voor de vmbo-t-, als de havo- en de vwo-klassen. Het idee erachter is dat leerlingen tijdens een aantal atletiekonderdelen van hun lolessen metingen doen, die ze tijdens de wiskundeles statistisch verantwoord verwerken om er vervolgens conclusies uit te kunnen trekken. Het project vervangt een hoofdstuk statistiek. De achterliggende hoop is dat de leerlingen gemotiveerder zijn voor het onderwerp statistische verwerking als ze met echte en eigen data werken, en dat ze beter het nut van een goede dataverwerking gaan inzien, omdat de data uit hun eigen wereld komen. Door gegevens manipulerend te verwerken en onware conclusies te trekken kun je opeens de sportheld worden die je nooit was…!

Het project

Gedurende 7 weken hebben leerlingen 4 blokken lo en 3 blokken wiskunde. (Een blok bij ons is een les van 2 keer 45 minuten.) De lo-docent organiseert de atletiekopdrachten: stoten, lopen, driesprong, en het bijbehorende meetwerk en het verzamelen van de data. Leerlingen krijgen hiervoor een verzamelformulier dat ter plekke tijdens de gymles wordt ingevuld. Elke leerling noteert zijn of haar eigen gegevens voor zichzelf en vervolgens op een verzamelstaat, zodat in de wiskundelessen ook gewerkt kan worden met de resultaten van de hele groep. De lo-docent zorgt ervoor dat de formulieren bij zijn wiskundecollega terechtkomen. Tijdens de daaropvolgende wiskundelessen worden de gegevens zowel individueel als klassikaal verwerkt. De bedenkers van het project (een lo- en een wiskundedocent) hebben een werkboekje gemaakt; ***zie Kader 1*** op pag. 215. In dit werkboekje valt ook te zien welke statistiekonderdelen aan bod komen. U ziet als voorbeeld het onderdeel 'Stoten' ***in Kader 2***.
Er is een verschil in aanpak in de klassikale lessen tussen de diverse klassen. In de vmbo-t-klassen wordt de theorie van en aanpak bij dit statistiekproject klassikaal besproken, in de havo-klassen wordt de theorie en uitleg opgehangen aan de opdrachten voor het project en in de vwo-klassen worden leerlingen geacht zelf uit te zoeken hoe ze bijvoorbeeld een mediaan moeten berekenen. Ze mogen daarvoor hun wiskundeboek gebruiken, maar ook elke andere goede bron. Leerlingen werken in groepen van 3 à 4 personen en in hun eindverslag moet elke leerling vooral inzoomen op zijn/haar eigen prestaties in vergelijking met het klassengemiddelde. De samenwerkingsopdrachten zijn uitdrukkelijk gericht op de samenwerking waarbij iedere leerling een eigen rol vervult, zoals voorzitter, secretaris, materiaalverzamelaar of wiskundig brein.

Tijdens het eerste projectjaar bleek er te veel overlap in de wiskundeopdrachten die bij de verschillende atletiekonderdelen uitgevoerd moesten worden: herhaaldelijk snelheden uitrekenen, staaf-, cirkel-, steelblad-diagrammen tekenen, gemiddelde / modus / mediaan berekenen op grond van de eigen metingen. Er werden vragen gesteld zoals: 'Maak een histogram bij de hardloopgegevens', 'Maak een histogram bij de driesprong gegevens', 'Teken een cirkeldiagram bij de snelheden van de kogel'. De herhaling bleek niet nodig en te saai; bovendien waren deze opdrachten zo tijdrovend dat leerlingen niet aan verdieping toekwa-

men. In het tweede projectjaar is dit aangepast. Leerlingen werken nu niet alleen met hun eigen groepsgegevens, maar ook met de gegevens van de hele klas of met een deelgroep, zoals de gegevens van alle jongens of meisjes. Het accent in de opdrachten is verschoven naar het vergelijken van de diverse uitkomsten en vervolgens het trekken van relevante conclusies.

Vragen

Een aantal vragen komt op bij bestudering van het project.
Hoe ging het met de tijdwaarneming tijdens de lo-lessen? Hoe secuur waren de metingen? Hoe waren leerlingen ermee bezig? Hoe precies konden ze zijn? Waren er ook leerlingen die veel liever alle tijd gesport hadden?
En hoe hielden de gymdocenten de leerlingen gemotiveerd? Want dat het project winst oplevert voor de wiskundelessen lijkt op voorhand evident. Maar waar zit de winst voor lo? En hoe ging het tijdens de wiskundelessen? Sjoemelden leerlingen met hun gegevens? Was het anoniem? En zijn de beoogde doelstellingen over vergrote motivatie voor en betrokkenheid bij dit hoofdstuk statistiek gehaald door te werken met echte, eigen data?
We gaan het vragen aan de betrokken collega's.

Ervaringen

Leerlingen zijn enthousiast over het project omdat het bijzonder leuk en verrassend is om met eigen gegevens te werken tijdens de wiskundelessen. Uitschieters naar boven en beneden in de metingen worden gewoon genoteerd. Metingen bijstellen of weglaten is helemaal niet aan de orde. Bij lo heerst een open sfeer waarin leerlingen gewend zijn naar elkaars prestaties te kijken.
Als wij, beide auteurs, aan onze eigen gymlessen denken, herinneren we ons toch vooral de stiekeme en ook gemene opmerkingen die we laatdunkend over elkaar konden maken, zoals over een zoveelste mislukte sprong van P., over een te dikke medeleerling die niet mee kon rennen, over degene die altijd de hordes omliep.
Maar de lo-collega's schetsen een ander beeld: leerlingen weten heel goed van elkaar wat ze wel en niet kunnen, en accepteren het als gegeven feit. Ze zijn wel competitief ingesteld, weten heel goed wie de beste en de snelste is, zijn gefocust op tijden en afstanden, maar zonder oordelen: hier is een andere generatie die veel makkelijker met hun lijf omgaat, waar te dikke meiden ook gewoon met naveltruitjes lopen; het heeft absoluut goede kanten! De metingen worden dan ook eerlijk genoteerd, en in het tweede projectjaar ook op grote verzamelstaten ingevuld, zodat dat in de wiskundelessen niet meer hoeft te gebeuren.
Tot nu toe zat de winst vooral in het werken met eigen data binnen de wiskundelessen. Leerlingen waren daardoor bijzonder gemotiveerd om de diverse statistiek-onderdelen te leren en goed uit te voeren. De vorm die ze voor de uitwerking van hun metingen kozen, was vrij voor de hand liggend: een verslag of een poster, omdat vooral de gegevens wiskundig goed verwerkt moesten worden. Iets eigens toevoegen aan het geheel, behalve de eigen data, deden de leerlingen niet.
De wiskundecollega's hebben kritiek op hun lo-collega's: ze vinden dat hun collega's dit project er 'te veel bij doen'. Het lijkt alsof er niets veranderd is in de gymlessen; er wordt tijdens een 'gewone' les ook even atletiek gedaan om aan de verplichtingen van het project te voldoen, maar het feit dat de gegevens verwerkt gaan worden bij wiskunde, krijgt nauwelijks aandacht. Dit lijkt een gemiste kans. In de gymlessen zou toch ook aandacht kunnen komen voor verdieping in de sportachtergronden? Te denken valt aan onderwerpen zoals het verbeteren van sportprestaties, hartslag of (sport)voeding. Het zou voor leerlingen dan nog interessanter worden om bij het trekken van hun conclusies tijdens de wiskundelessen deze informatie te verwerken en toe te passen op hun eigen uitwerkingen van de metingen. Het project zou er meer cachet van krijgen, aldus enkele betrokkenen. De lo-collega's onderschreven de kritiek voor een deel, en na het tweede projectjaar is het project inmiddels volledig op de schop. Tijdens een studiedag zijn er nieuwe plannen gemaakt waarin het projectmatige karakter veel meer tot zijn recht gaat komen.

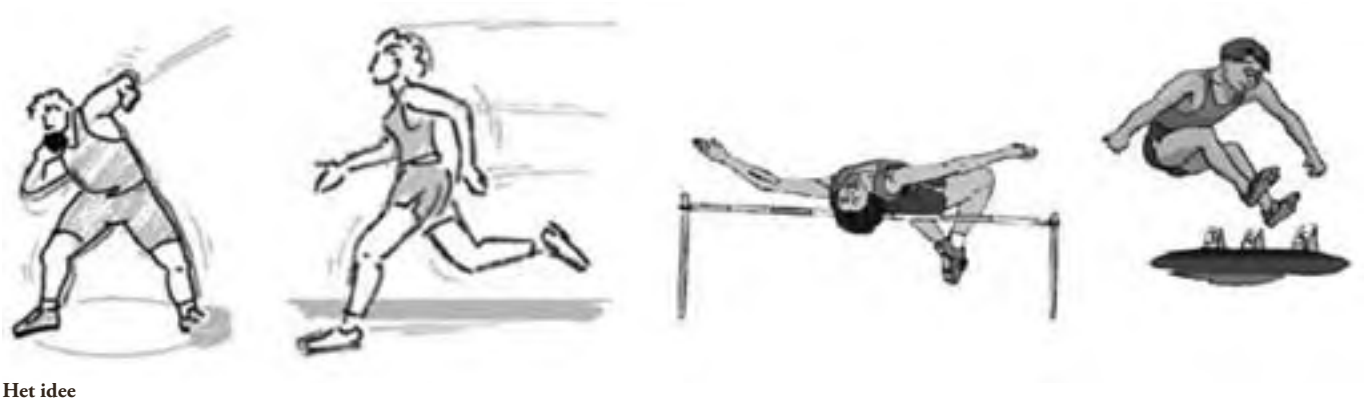
Het vervolg

Het project is inmiddels helemaal herzien. Er is gekozen voor een totaal andere opzet, waarbij de wiskunde veel meer impliciet (maar wel degelijk) aanwezig is. De opdracht wordt veel opener. Leerlingen gaan aan het werk aan de hand van de hoofdvraag: *Maak een plan om je conditie in zeven weken tijd te verbeteren.* De bedoeling is dat leerlingen 'echt' een eigen onderzoek gaan doen. Collega's zijn tevreden omdat er nu een plan ligt waarin werkelijk samengewerkt en samen geleerd moet worden, zowel door collega's als door leerlingen. Wie weet berichten we u over enige tijd weer over onze ervaringen.

Over de auteurs

Peter Tilman is wiskundedocent en vakgroepcoördinator onderbouw in Amersfoort aan het Meridaan College, vestiging 't Hooghe Landt.
E-mailadres: *tmw.tilman@meridiaan-hl.nl*
Klaske Blom is redacteur van *Euclides* en wiskundedocent in Amersfoort aan het Meridaan College, vestiging 't Hooghe Landt.
E-mailadres: *kablom@tiscali.nl*

Kader 1 – Project wiskunde en LO 2006-2007 voor klas 2T, 2H & 2V
Docentenboekje voor lo- en wiskundedocenten



Het idee

In de lo-lessen doen de leerlingen metingen op allerlei sportgebied. Dit kan gaan over afstanden, hoogtes, tijden etc. In de wiskundelessen gaan ze vervolgens deze gegevens op een statistische manier verwerken. Dit gebeurt zowel individueel als klassikaal. (…)

Hoe gaat het praktisch in zijn werk

- Wiskunde en lo worden om en om gegeven. Het kan zijn dat een klas met lo begint, maar het kan ook zijn dat ze met wiskunde begint. Docenten werken in duo's: een lo- en een wiskundedocent geven beiden les aan twee dezelfde klassen (in dit voorbeeld A en B). De ene week heeft de lo-docent klas A en de wiskundedocent klas B; de volgende week draait dit om. Elke docent heeft per week een blokuur van 2 keer 45 minuten tot zijn beschikking. Iedere klas heeft 4 blokuren lo en 3 blokuren wiskunde (in het ideale geval) of andersom. In een schema staan alle benodigde gegevens.
- In de lo-les zorgen de leerlingen voor het verzamelen van de resultaten. Voor iedere leerlingen is daarvoor een verzamelformulier. Als de gegevens verzameld zijn, gaan de leerlingenformulieren weer naar de lo-docent. Deze geeft ze aan de wiskundecollega (uitwisseling via postvak).
- In de wiskundeles verwerken de leerlingen de gegevens (zie hieronder voor ons programma). Een probleem hierbij is dat als een klas als eerste les in het project een wiskundeles heeft, er nog geen gegevens beschikbaar zijn. De volgorde van de onderdelen is als volgt:

Start	**Les 1**	**Les 2**	**Les 3**	**Les 4**	**Les 5**	**Les 6**	**Les 7**
wi	Klas A	Klas B	Klas A	Klas B	Klas A	Klas B	Klas A
verwerken gegevens:	*hoogspringen* (*)	*stoten*	*stoten*	*lopen*	*lopen*	*driesprong*	*driesprong*
lo	Klas B	Klas A	Klas B	Klas A	Klas B	Klas A	Klas B
onderdeel:	*stoten*	*stoten*	*lopen*	*lopen*	*driesprong*	*driesprong*	*hoogspringen*

(*) met geschatte gegevens

Kader 2 – Stoten

lo	**wiskunde**
Lln. meten afstand van geworpen kogel, basketbal en medicinbal. Lln. meten de tijd op tussen loslaten bal en moment dat bal op grond komt.	Afstanden van de drie verschillende ballen Snelheid berekenen van de drie verschillende ballen
Bovenstaande zaken worden op scoreformulier ingevuld.	*individueel* - afstand van de drie ballen als staafdiagram; - snelheid van de drie ballen als staafdiagram; - conclusies trekken uit de diagrammen a.d.h.v. gesloten opdrachten zoals: "Bereken de snelheid van de drie verschillende ballen en maak hier vervolgens een histogram van."
	klassikaal Uitwisselen van de gegevens. Leerlingen zitten in groepjes. Een groepslid zorgt ervoor dat ieder groepje de gegevens van alle klasgenoten krijgt. Hierbij verschil maken tussen jongen en meisje. - snelheden van de klas als klassenindeling (3 ballen); - gemiddelde, modus, mediaan van de klas als geheel bepalen voor afstand en snelheid (3 ballen); - gemiddelde, modus, mediaan van jongens en meisjes bepalen voor afstand en snelheid (3 ballen); - conclusies trekken (zoals wie heeft het goed gedaan?, verschil tussen de ballen, verschil jongens/meisjes); - uitwisselen van de conclusies.

Kansrekening en statistiek bij de verspreiding van besmettelijke ziektes

[Ronald Meester]

Inleiding

Van tijd tot tijd worden we in Nederland opgeschrikt door het uitbreken van een besmettelijke ziekte onder dieren. Zo'n tien jaar geleden bijvoorbeeld zorgde een uitbraak van de varkenspest voor gruwelijke beelden op de televisie, en nog niet zo lang geleden was er de dreiging van mond- en klauwzeer uit Engeland. Bij elke (dreigende) uitbraak van een dergelijke ziekte moeten er maatregelen getroffen worden om verdere verspreiding te voorkomen. Hierbij kun je denken aan vaccinatie, een vervoersverbod, of het zogenaamde 'ruimen' van bedrijven, een eufemisme voor het doden van alle dieren op een bepaald bedrijf.

De keuze voor bepaalde (combinaties van) methodes wordt gemaakt aan de hand van een aantal overwegingen. Bij deze overwegingen spelen vooral economische motieven een rol, want het gaat eigenlijk altijd over geld. Daarnaast is het natuurlijk ook belangrijk dat een bepaalde voorgestelde methode ook het beoogde effect zal hebben. Het is duidelijk dat sommige van deze overwegingen met elkaar conflicteren: alle dieren vaccineren zou een bepaalde uitbraak waarschijnlijk kunnen stoppen, maar deze maatregel kan bijvoorbeeld te duur bevonden worden.

Maar hoe kunnen we voorspellen of een bepaalde maatregel wel het beoogde effect zal hebben? Bij het beantwoorden van deze vraag spelen zaken als ervaring natuurlijk een grote rol, maar ook de wiskunde - en in het bijzonder de kansrekening en statistiek - kan hier een bijdrage leveren. Wiskundigen kunnen proberen om de uitbraak van een epidemie te modelleren, en kunnen vervolgens proberen om aan de hand van dat model een voorspelling te doen over het verdere verloop van een epidemie. Als het model aangeeft dat de epidemie verder uit de hand zal lopen, dan moeten er mogelijk maatregelen getroffen worden. Het is zelfs mogelijk om met behulp van een wiskundig model na te gaan welke maatregelen het meest effectief zouden kunnen zijn.

Natuurlijk is het modelleren van een gecompliceerd proces als een epidemie een hachelijke zaak, en iedereen die zich hiermee bezighoudt zal erkennen dat een model de werkelijkheid maar zeer beperkt kan weergeven. De precisie die modellen in bijvoorbeeld de natuurkunde bereiken is hier zeker niet te verwachten, maar ondanks dat kan een verstandig model ons wel degelijk inzicht geven over de vraag wat een goede maatregel zou kunnen zijn.

De keuze van een model is niet eenvoudig. Om te beginnen kunnen we kiezen tussen een deterministisch of een stochastisch model. In een deterministisch model speelt toeval geen rol, terwijl dat in een stochastisch model wel zo is. Als vuistregel kun je denken aan een deterministisch model als de epidemie al een tijdje op streek is - het gaat dan over grote aantallen en de toevallige fluctuaties worden dan als het ware uitgemiddeld - terwijl een stochastisch model meer op zijn plaats lijkt aan het begin van een epidemie. Maar ook als je het type model hebt bepaald zijn er vaak nog vele manieren om het model op te zetten. Dit noopt tot enige bescheidenheid, want het model is altijd gebaseerd op je eigen keuzes.

In dit artikel wil ik een voorbeeld geven van een eenvoudig model waarin kansrekening en statistiek een voorname rol spelen. Kansrekening en statistiek omvatten zo veel meer dan vazen met ballen, en in de praktijk is de toepassing ervan vaak een stuk weerbarstiger dan je misschien denkt. Ik wil ook laten zien hoe een toegepast probleem kan leiden tot interessante theoretische vragen - vragen waaraan we zonder deze toepassing nooit gedacht zouden hebben. Dit illustreert de interactie tussen theorie en praktijk die ik zelf als uitermate inspirerend ervaar. Het voorbeeld komt uit de recente wiskundige onderzoekspraktijk (zie [1]).

Het model

Stel dat een besmettelijke ziekte ontstaat op een bepaald bedrijf; we noemen dat bedrijf dan *besmet*. Om het verloop van de ziekte te beschrijven kiezen we als tijdseenheid een week. In zo'n week kunnen twee dingen gebeuren. Het kan zijn dat de ziekte op het bedrijf ontdekt wordt; in dat geval wordt in dit model het bedrijf geruimd, en besmet het geen andere bedrijven. Als de ziekte op het bedrijf echter niet ontdekt wordt, dan zal het een aantal andere bedrijven in de omgeving besmetten. Om concreet te zijn stellen we dat de kans dat er in dat geval precies k bedrijven in een week besmet worden door het oorspronkelijke bedrijf, gelijk is aan:

$$p_k = e^{-\lambda}\frac{\lambda^k}{k!}$$

voor $k = 0, 1, 2, \ldots$ Dit is de zogenoemde *Poisson-verdeling* die in de kansrekening een voorname rol speelt. Deze verdeling is in dit model heel natuurlijk: als je eist dat het proces geen geheugen heeft (dat wil zeggen dat toekomstige ontwikkelingen alleen mogen afhangen van de huidige situatie), dan wordt deze verdeling je als het ware opgedrongen. Het bewijs hiervan geef ik niet, maar is niet eens zo moeilijk.

De parameter λ kunnen we kiezen om ervoor te zorgen dat het model zo realistisch mogelijk wordt; daarover zo meer. Als het aantal besmette bedrijven op deze manier wordt gekozen, dan is het niet zo moeilijk

om uit te rekenen dat er gemiddeld gesproken in een week λ bedrijven besmet worden door het oorspronkelijk besmette bedrijf. Dit gemiddelde wordt meestal de *verwachting* van de Poisson-verdeling genoemd.
De kans dat het bedrijf ontdekt wordt in een bepaalde week stellen we op π. Als het bedrijf ontdekt wordt (hetgeen met kans π gebeurt), dan is de epidemie na een week dus afgelopen; als het bedrijf niet ontdekt wordt (en dit gebeurt met kans $1-\pi$), dan zijn er na een week met kans p_k precies $k+1$ besmette bedrijven: het oorspronkelijke bedrijf plus de k besmette bedrijven.
Voor de tweede week doen we precies hetzelfde, maar nu voor elk van de bedrijven die na de eerste week (nog) besmet zijn. Elk van deze bedrijven besmetten op hun beurt dus weer een (toevallig) aantal bedrijven met dezelfde Poisson-kansverdeling als eerst, mits ze niet ontdekt worden. En voor elk bedrijf geldt dat de kans op ontdekking gelijk is aan π. Dit proces herhaalt zich week na week, totdat er geen besmette bedrijven meer zijn, en de epidemie over is.
Voordat we er wiskundig naar gaan kijken, wil ik opmerken dat dit model niet realistisch is als de epidemie te groot wordt. In dat geval zullen er namelijk gewoon niet genoeg bedrijven zijn die besmet kunnen worden door locale geografische uitputting. Echter, voor niet al te grote epidemieën is dit misschien helemaal geen gekke beschrijving.
Het model dat ik zojuist heb geschetst heeft twee onbekenden, namelijk λ en π. Gemiddeld gesproken zal een bepaald bedrijf verantwoordelijk zijn voor
$R = (1-\pi)(\lambda+1)$
besmette bedrijven in de daarop volgende week. Immers, allereerst moet het bedrijf niet ontdekt worden, hetgeen met kans $1-\pi$ gebeurt, en vervolgens besmet het gemiddeld λ andere bedrijven, en de extra 1 is het bedrijf zelf dat immers de volgende week ook nog gewoon besmet is.
Welnu, het is bekend - en niet zo moeilijk te bewijzen - dat voor $R<1$ de epidemie zeker uitsterft. Met andere woorden, om er voor te zorgen dat de ziekte zich niet blijft uitbreiden, is het zaak ervoor te zorgen dat $(1-\pi)(\lambda+1)<1$. Dit is echter een stuk lastiger dan het op het eerste gezicht lijkt, want hoe bepalen we hoe groot λ en π eigenlijk zijn? De grootste moeilijkheid hier is dat we niet de hele epidemie kunnen observeren. Immers, we observeren (natuurlijk) alleen maar bedrijven waarbij de ziekte ontdekt is, en die vervolgens geruimd worden. Natuurlijk geeft dit aantal een indicatie voor het werkelijke (onbekende) aantal besmette bedrijven, en de wiskundige kunst is nu om op basis van de observaties die we doen, een schatting te maken van R zodat we kunnen nagaan of de getroffen maatregelen voldoende zijn om R kleiner dan 1 te maken.
Dit is nu een mooi voorbeeld van een theoretische vraag, die naar voren kwam door een werkelijk praktisch probleem. Het proces dat we beschouwen, heet ook wel een *vertakkingsproces*, en de vraag waarnaar we nu kijken is, wat we kunnen zeggen over het (toevallige) aantal nakomelingen van een bepaald 'individu' (in ons geval is een individu een bedrijf) als we iets weten over alleen die individuen die geen nakomelingen hebben. Zonder praktijkmotivatie zouden we dit zeker niet bestudeerd hebben, terwijl de theorie mooi blijkt te zijn. In de rest van dit artikel wil ik een impressie geven van wat je er wiskundig zoal over kunt zeggen.

Het schatten van R

Laten we het aantal besmette bedrijven na n weken aangeven met X_n, en het aantal dat hiervan ontdekt wordt in de komende $(n+1)$-ste week met Z_n. Het is belangrijk je te realiseren dat we de X_n's niet kunnen waarnemen, maar de Z_n's wel. Wanneer we dus $Z_0, Z_1, Z_2, \ldots, Z_n$ hebben waargenomen, is de vraag wat we op basis van die informatie over R kunnen zeggen. Hiervoor hebben we om te beginnen de volgende stelling.

Stelling 1. *De getallen* $A_n = \dfrac{Z_n}{Z_{n-1}}$ (1)
zijn goede schattingen voor R in de zin dat als het proces niet uitsterft, we zeker weten dat de rij A_n *naar R convergeert.*

Hoewel het bewijs van deze stelling nog niet eens zo simpel is, is het niet zo moeilijk om in te zien waarom hij wel waar zal zijn. Dit gaat als volgt. Stel dat Z_{n-1} gelijk is aan z. Dan zal X_{n-1} ongeveer gelijk zijn aan $\frac{z}{\pi}$ immers, Z_{n-1} is ongeveer een fractie π van X_{n-1}. Omdat elk besmet bedrijf gemiddeld voor R besmette bedrijven zorgt in de volgende week, zal X_n ongeveer gelijk zijn aan $\frac{zR}{\pi}$, en hiervan wordt een fractie π ontdekt, hetgeen leidt tot $Z_n = \frac{zR}{\pi}\times\pi = zR$. Hieruit concluderen we dat $\frac{Z_n}{Z_{n-1}}$ wel in de buurt van R zal liggen.
Op het eerste gezicht lijkt het probleem hiermee opgelost, maar kunnen we niets beters doen? Hiertoe is het belangrijk op te merken dat bij de berekening van A_n alleen maar informatie van de weken $n-1$ en n gebruikt wordt! De rest van de beschikbare informatie wordt genegeerd, en dat betekent dat we - door die extra informatie verstandig te gebruiken - wellicht een betere schatting voor R kunnen doen.
In de volgende stelling wordt wel alle infor-

matie gebruikt.

Stelling 2. *De getallen* $B_n = \dfrac{\sum_{i=1}^{n} Z_i}{\sum_{j=0}^{n-1} Z_j}$ (2)

zijn goede schattingen voor R in de zin dat als het proces niet uitsterft, we zeker weten dat B_n *naar R convergeert.*

Hoewel het niet direct in de stellingen tot uitdrukking komt, is de rij B_n echt beter dan de rij A_n in de zin dat de convergentie *sneller* is, maar daar ga ik hier niet verder op in. De reden waarom Stelling 2 waar is, is eigenlijk dezelfde als voor Stelling 1; het formele bewijs is lastiger omdat de uitdrukking wat ingewikkelder is.

Het schatten van π en λ afzonderlijk

Het kan belangrijk (of alleen maar interessant) zijn om niet alleen informatie over *R* te hebben, maar ook over π en λ zelf. Als we ook de X_j's zouden kunnen waarnemen, dan zou het veel gemakkelijker zijn, vanwege het volgende resultaat.

Stelling 3. *De getallen* $C_n = \dfrac{\sum_{i=0}^{n} Z_i}{\sum_{j=0}^{n} X_j}$ (3)

zijn goede schattingen voor π *in de zin dat als het proces niet uitsterft, we zeker weten dat* C_n *naar* π *convergeert.*

Ook dit resultaat is vrij makkelijk te begrijpen: immers, de teller is het aantal ontdekkingen en de noemer het totale aantal bedrijven dat ontdekt had kunnen worden. Helaas is de stelling niet bruikbaar in de praktijk, omdat we de X_j's niet kunnen waarnemen.
Het blijkt echter dat we toch nog iets meer kunnen doen dan dit. Hiertoe moeten we ons echter wel in wat bochten wringen, en dat resulteert in de volgende stelling. Ik ben me er van bewust dat het onmogelijk duidelijk kan zijn waarom de stelling waar is.

Stelling 4. *De getallen*

$$D_n = \frac{1}{n}\sum_{i=0}^{n} (Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1} - B_n\right)^2 \quad (4)$$

zijn goede schattingen voor

$$\gamma = (1-\pi)\lambda + (1-\pi)^2 + (1-\pi)(\lambda+1)^2 \quad (5)$$

in de zin dat als het proces niet uitsterft, we zeker weten dat D_n *naar* γ *convergeert.* [a]

Ik kan in dit artikel niet echt uitleggen waarom deze stelling waar is. De formules lijken een beetje uit de lucht te vallen, maar de γ in (5) blijkt toch een mooie interpretatie te hebben als een bepaalde spreidingsmaat. De liefhebbers verwijs ik naar [1].
Wat hebben we nu bereikt? Welnu, wanneer we $Z_1, \ldots, Z_n$ observeren, dan kunnen we met behulp van (2), (4) en (5) twee vergelijkingen met twee onbekenden λ en π opstellen, en met een beetje geluk kunnen we dan λ en π bepalen. Immers, als we $Z_1, \ldots, Z_n$ kennen, dan kennen we ook B_n en D_n en dan hebben we het stelsel vergelijkingen:

$$B_n = (1-\pi)(\lambda+1)$$
$$D_n = (1-\pi)\lambda + (1-\pi)^2 + (1-\pi)(\lambda+1)^2$$

Dit alles lijkt allemaal vrij voorspoedig te verlopen, maar er zit toch nog een vervelend addertje onder het gras. Weliswaar zijn bovengenoemde stellingen allemaal juist, maar de convergentie in Stelling 4 blijkt in de praktijk bijzonder langzaam te gaan. Hierdoor kan het gebeuren dat het bovengenoemde stelsel helemaal geen oplossing heeft, of bijvoorbeeld geen oplossing voor π tussen 0 en 1; dat is ons meerdere malen overkomen. Dit betekent zeker niet dat al deze moeite zonde van de tijd is. Wat we er (ook) van leren, is dat de beschikbare gegevens soms *principieel* te weinig informatie bevatten om een betrouwbare schatting te geven van de onbekende parameters. Dit weerspiegelt de spanning tussen realiteit en theorie: zelfs als het in theorie prachtig werkt (zoals hier), dan nog is het niet altijd zeker dat we het in de praktijk ook kunnen gebruiken.
In de volgende paragraaf geven we een toepassing waarbij we ons beperken tot het schatten van *R* en γ. De liefhebber kan proberen om hieruit een schatting voor π en λ te destilleren.

Toepassing: de varkenspestepidemie van 1997

In deze paragraaf bestuderen we de al eerder genoemde varkenspestepidemie van 1997 in Nederland. Omdat gedurende verschillende periodes van die epidemie verschillende maatregelen van kracht waren, leek het ons verstandig om voor elk van die periodes een aparte schatting van *R* en van γ te maken. In de tabel valt te zien dat we vier periodes hebben onderscheiden, die respectievelijk 10, 8, 9 en 30 weken duurden; gedurende elk van die periodes veranderden de genomen maatregelen niet. Verder zien we dat in de eerste periode 101 bedrijven geruimd werden, 160 in de tweede, 107 in de derde en 51 in de laatste periode. In de laatste twee kolommen vinden we de schattingen van *R* en γ via bovenstaande theorie. (Indien gewenst kun je proberen om hieruit een schatting voor π en voor λ te destilleren. Lukt dat in alle gevallen?)
Het is aardig om op te merken dat in de eerste twee periodes geldt dat $R > 1$, terwijl de strengere maatregelen in periodes 3 en 4 er voor zorgden dat *R* onder de 1 terecht kwam; dit was nodig om de epidemie tot staan te brengen.
In principe zijn dit soort berekeningen mogelijk *tijdens* het verloop van de epidemie, en dus niet alleen na afloop. Aan de hand van de ontdekte besmette bedrijven kunnen we de schatting van *R* bepalen, en zien of deze kleiner is dan 1. Natuurlijk moet het model ook gevalideerd worden; er moet gekeken worden of bijvoorbeeld de epidemie inderdaad uitsterft als de schatting voor *R* kleiner is dan 1. Of het model dan bij een volgende gelegenheid ook weer bruikbaar is, is altijd maar weer de vraag. Wat dat betreft denk ik dat modellen als deze wel een rol kunnen spelen bij de besluitvorming, maar beleid kan en mag zeker niet alleen afhangen van wat wiskundigen er over zeggen; daarvoor zijn de modellen simpelweg niet nauwkeurig genoeg. De modellen en de berekeningen kunnen wel een bijdrage leveren wanneer het erom gaat te begrijpen welke maatregel wellicht het meest effectief is. En naast deze (enigszins bescheiden) praktische motivatie [b], levert het denken over dit soort problemen ook gewoon mooie wiskunde op.

periode	aantal weken	aantal geruimde bedrijven	*R*	γ
1	10	101	1,15	1,03
2	8	160	1,04	1,87
3	9	107	0,857	0,813
4	30	51	0,880	0,859

Noten

[a] Strikt gesproken is dit niet exact wat we bewijzen, maar het verschil met wat hier staat, is puur technisch en voor dit verhaal niet van belang.

[b] (Red.) Zie ook het centraal examen VWO Wiskunde A1 2006 (2e tijdvak), Varkenspest.

Literatuur

[1] R. Meester, P. Trapman (2006): *Estimation in branching processes with restricted observations*. In: *Advances of Applied Probability*, pp. 1098-1115.

Over de auteur

Ronald Meester is hoogleraar waarschijnlijkheidsrekening aan de Vrije Universiteit in Amsterdam.
E-mailadres: *rmeester@few.vu.nl*

Statistische proces-controle in de auto-industrie

[Arthur Bakker, Celia Hoyles, Phillip Kent en Richard Noss]

In het bedrijfsleven speelt statistiek een belangrijke rol. Reden voor Bakker en zijn Engelse collega's om daar onderzoek naar te doen en trainingsmateriaal te ontwikkelen.

Inleiding

Bedrijven gebruiken in toenemende mate statistische technieken om hun productieprocessen te verbeteren. Een veelgebruikte techniek is statistische procescontrole (SPC). In dit artikel beschrijven we deze techniek en de software die we ontwikkeld hebben voor SPC-cursussen in twee autofabrieken.

Onderzoek naar technisch-wiskundige geletterdheid

In de eerste fase van ons vierjarige onderzoek hebben we proberen te achterhalen welke wiskundige kennis werknemers nodig hebben. Wij hebben ons vooral gericht op mensen zonder academische opleiding in een aantal grote sectoren: de farmaceutische industrie, verpakkingsindustrie, auto-industrie en de financiële sector. In totaal hebben we twaalf bedrijven onderzocht. Onze interesse ging uit naar wat we 'technisch-wiskundige geletterdheid' (*techno-mathematical literacies*) noemen: functionele wiskundige kennis die specifiek is voor de situatie en de instrumenten die daarin worden gebruikt. Om een voorbeeld te geven: met de hand een statistische analyse uitvoeren vergt andere kennis dan met een grafische rekenmachine of een computerprogramma. In bedrijven worden de meeste berekeningen door software en machines uitgevoerd, waardoor werknemers wiskundige resultaten vaker moeten interpreteren dan produceren.

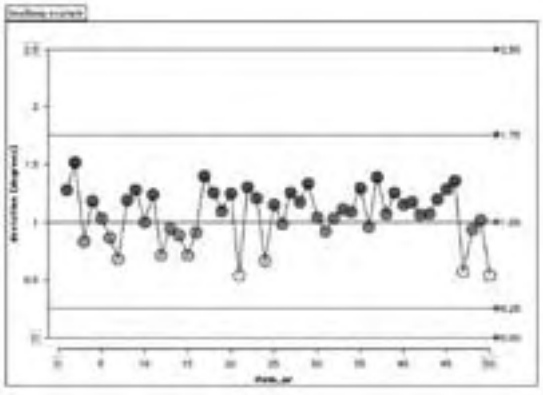
In de tweede fase van het onderzoek ontwikkelden wij cursusmateriaal voor een brede doelgroep om het gebrek aan technisch-wiskundige geletterdheid die wij in de eerste fase hadden ontdekt aan te pakken. In dit artikel beperken we ons tot SPC in de auto-industrie.

In een van de autofabrieken hebben we enkele interviews gehouden met operators, groepsleiders en managers van verschillende rangen en standen om een beeld te krijgen van de statistische verbetertechnieken die ze hanteerden. We namen deel aan een SPC-cursus, enquêteerden de dertien deelnemers en hielden met drie van hen een uitgebreid interview. Daarna hebben we in overleg met de SPC-trainers drie computertools ontwikkeld met begeleidende activiteiten die vervolgens in drie cursussen gebruikt werden. Elke cursus is geëvalueerd door middel van evaluatieformulieren en interviews.

SPC: statistische procescontrole

Laten we kort schetsen wat SPC is. Productieprocessen moeten aan strenge eisen voldoen, niet alleen voor de veiligheid van de producten, maar ook om redenen van efficiëntie en milieubelasting. Alle productieprocessen zijn echter variabel: de ene keer is de verflaag bijvoorbeeld enkele micrometers dikker dan de andere keer. Wat betekent het dan dat een proces onder controle is? Volgens de SPC-theorie moeten de meetpunten dicht bij een streefwaarde (*target*) en binnen bepaalde grenswaarden blijven.

Een voorbeeld.
De wet eist dat koplampen tussen 0° en 2,5° naar beneden en naar de berm wijzen, anders worden tegenliggers mogelijk verblind. Zulke grenswaarden heten specificatiegrenzen. Meetpunten buiten die grenzen leiden mogelijk tot boetes, afval, gevaarlijke situaties of klachten; ***zie figuur 1***.
Om te kunnen voorspellen tussen welke waarden het proces blijft, maakt men gebruik van een wetmatigheid van de normale verdeling: als het proces stabiel is en de meetpunten normaal verdeeld zijn, kunnen we ervan uitgaan dat 99,7% van de meetpunten binnen het interval $[m - 3SD, m + 3SD]$ valt. Hierbij staat m voor het gemiddelde en SD voor de standaarddevia-



figuur 1 SPC-grafiek (reconstructie in TinkerPlots) met hoeken van 50 koplampen. De specificatiegrenzen zijn 0 en 2,5 graden naar beneden; 0,5 en 2,0 graden zijn de controlewaarden. Dit proces is onder controle.

tie. Deze twee waarden heten controlewaarden (*control limits*): de *lower control limit* (*LCL*) is $m - 3SD$ en de *upper control limit* (*UCL*) is $m + 3SD$. Als we ervoor kunnen zorgen dat deze controlewaarden (0,5° en 2,0° ***in figuur 1***) voldoende ver van de specificatiegrenzen verwijderd blijven, dan is de kans dat een meetpunt door puur toeval de ruimere specificatiegrenzen overschrijdt nihil.
In de praktijk blijven productieprocessen echter lang niet altijd stabiel. Er is dus behoefte aan regels die waarschuwen dat er mogelijk een afwijking optreedt. We spreken dan van een speciale oorzaak, die moet worden opgespoord en verwijderd. SPC-experts hebben enkele regels gedefinieerd voor het opsporen van trends voordat meetpunten een controlewaarde overschrijden. Een voor de hand liggende regel is dat er waarschijnlijk iets mis is als er een meetpunt buiten de controlewaarden ligt. Maar er zijn ook subtielere regels.
Stel bijvoorbeeld dat er zeven meetpunten achter elkaar boven het gemiddelde liggen. De kans hierop is ongeveer $\left(\frac{1}{2}\right)^7 = \frac{1}{128} \approx 0{,}0078$. Dat is nogal klein als we uitgaan van random oorzaken van variatie. Veel waarschijnlijker is een speciale oorzaak die het gemiddelde heeft doen opschuiven. De verkeerde instelling van een machine, of grondstoffen van een andere leverancier bijvoorbeeld. Dergelijke kansregels helpen bij het tijdig signaleren van speciale oorzaken, maar het blijft mogelijk dat er iets mis is wat niet door een van de bekende regels aan het licht komt. Ook komt door puur toeval (ongeveer 1 op 300 keer) wel eens een rijtje van zeven meetwaarden boven het gemiddelde voor zonder dat er iets aan de hand is. Grondige kennis van het proces blijft dus een vereiste om de meetwaarden te interpreteren.

In de praktijk worden operators geacht op te letten of er trends in de data zijn die kunnen duiden op een probleem. Als er zeven meetpunten achter elkaar boven het gemiddelde worden gevonden, moeten ze daar een notitie van maken of hun baas waarschuwen. Het komt echter regelmatig voor dat ze trends en patronen niet opmerken, bijvoorbeeld doordat ze gewoon niet naar de grafieken kijken. Ook is er verwarring over het verschil tussen specificatiegrenzen en controlewaarden. Waar managers het meest over klagen, is dat operators de instellingen van machines veranderen omdat er net een paar hoge of lage waarden zijn geweest. Als die toevallig zijn, wordt het proces er bij veranderde instellingen alleen maar slechter op. Enige achtergrondkennis over SPC, in het bijzonder inzicht in random variatie, is dus beslist geen overbodige luxe in dit soort bedrijven.

Procescapaciteitsmaten

Zoals we hiervoor al schreven, moeten de controlewaarden voldoende ver van de specificatiegrenzen verwijderd blijven. Maar hoe groot moet die afstand zijn als we een minimale kans willen hebben op een overschrijding van een specificatiegrens? Een historisch gegroeide vuistregel die voor veel processen wordt gehanteerd, is dat de verhouding van de afstand tussen de specificatiegrenzen ($USL - LSL$) tot de afstand tussen de controlewaarden ($UCL - LCL = 6SD$) minstens 4 : 3 moet zijn. Bij een stabiel gemiddelde wordt dan een afwijking geaccepteerd van $4SD$ voor de afstand tussen gemiddelde en de specificatielimieten. Zelfs als het gemiddelde dan een beetje opschuift of de variatie tijdelijk toeneemt, blijft de kans op overschrijding van de specificatiegrenzen uiterst klein. Bij cruciale productie-onderdelen wordt soms zelfs $5 \cdot SD$ of $6 \cdot SD$ geëist om de overschrijdingskans nog kleiner te maken (bij $6SD$ is die ongeveer 3 op een miljoen).

Statistici hebben maten ontwikkeld die zulke informatie in getallen samenvatten: de procescapaciteitsmaten. De eenvoudigste is C_p:

$$C_p = \frac{USL - LSL}{6\sigma}$$

met *USL* = *upper specification limit*, *LSL* = *lower specification limit*, σ = *standaarddeviatie* (*populatie*).
C_p geeft een indruk van de variatie in het proces, maar houdt geen rekening met de locatie van het gemiddelde. Als het proces weinig variatie vertoont maar ver verwijderd is van de streefwaarde, dan is C_p toch nog hoog. Om die reden is er ook een maat geformuleerd die *wel* rekening houdt met de locatie van het gemiddelde, C_{pk}:

$$C_{pk} = \min(C_{pkl}, C_{pku})$$

waarbij $C_{pku} = \frac{USL - \overline{X}}{3\sigma}$, $C_{pkl} = \frac{\overline{X} - LSL}{3\sigma}$

met $\overline{X}$ = gemiddelde (populatie) en *USL*, *LSL*, σ zoals in C_p.

figuur 2 Cp = 1,5, want 6SD (5,4 cm) past 1,5 in de afstand tussen LSL en USL (8 cm)

figuur 3 Staven waarmee Cp geschat kan worden

figuur 4 Cpk = 1,17 want 3SD past 1,17 keer in de afstand tussen LSL en het gemiddelde.

figuur 5 De berekening van Cpk kan naar behoefte zichtbaar en onzichtbaar gemaakt worden

figuur 6 Grafiek van Cpk = 0

figuur 7 Berekening van Cpk = 0

In de praktijk gebruikt men schatters van de standaarddeviatie omdat de populatie-standaarddeviatie onbekend is. Voor werknemers die sinds hun zestiende geen wiskunde meer gebruikt hebben, zijn dergelijke berekeningen tijdens een training even schrikken, maar in de praktijk hoeven ze die gelukkig niet zelf uit te voeren. Wel moet iedereen weten dat C_{pk} minstens $\frac{4}{3} = 1\frac{1}{3} \approx 1{,}33$ moet zijn. Bij een kleine interviewronde bleek echter lang niet iedereen die met SPC in aanraking kwam, dit te weten.

Ontwikkeling van computertools

Het zal de lezer niet verbazen dat werknemers, van operators tot managers, of ze nu veel of weinig opleiding hadden genoten, moeite hadden met het interpreteren en berekenen van de procescapaciteitsmaten. Daarom besloten we computertools te ontwikkelen die de belangrijkste relaties aanschouwelijk zouden maken en de berekeningen meer op de achtergrond zouden plaatsen.

Figuur 2 laat een grafiek in een computertool zien die bij C_p = 1,5 hoort ($USL - LSL$ = 10 – 2 = 8 cm; $6SD$ = 7,8 – 2,4 = 5,4 cm; dus $\frac{8}{5{,}4} \approx 1{,}5$). Merk op dat het niet nodig is om het gemiddelde (grijze verticale lijn) of de target (zwarte stippellijn) af te lezen; het schatten van de verhouding van de twee staven ***in figuur 3*** is voldoende.
Figuur 4 heeft een C_{pk} van 1,17. Door *hide/show values* aan te klikken, kunnen gebruikers de achterliggende berekeningen inspecteren en hun vermoedens testen (***zie figuur 5***). Door de bolletjes en vierkantjes te verplaatsen kunnen ze de controlewaarden en specificatiegrenzen wijzigen.

Het was een interessante uitdaging om samen met de SPC-trainers de tools zo te ontwerpen dat ze aan hun en onze eisen voldeden. Om een indruk te geven van onze communicatie geven we eerst een voorbeeld van onze eenzijdige wiskundige blik, en daarna van de beperkte statistische kennis van een trainer.

Een eye-opener voor ons

De definitie van C_{pk} is het minimum van twee andere waarden (zie boven). In de eerste versie van de C_{pk}-tool lieten we deelnemers daarom eerst deze twee waarden schatten, maar de trainers vonden dit geen goede aanpak. Zij keken alleen naar de kant van het proces die risicovol was. Waarom zou je beide waarden berekenen als je meteen ziet aan welke kant van het proces de kans op overschrijdingen het grootst is? Kortom, wij gingen aanvankelijk uit van de wiskundige definitie van C_{pk}, terwijl de trainers die waarde in termen van de context interpreteerden. Op advies van de trainers hebben we er toen voor gezorgd dat de risicovolle kant van het proces donkerder gekleurd was dan de oninteressante kant, en dat gebruikers niet de C_{pk}'s van beide kanten hoefden te schatten. Dit is een van de vele voorbeelden van de door ons ervaren discrepanties tussen de formele wereld van de wiskunde en de praktijk waarin wiskunde wordt gebruikt.

Een eye-opener voor een SPC-trainer

Een van de trainers (met een diploma in SPC maar geen academische opleiding) gebruikte tijdens de training steeds de volgende vuistregel om de deelnemers een idee te geven van wat de procescapaciteitsmaten betekenden voor de praktijk. 'Als C_{pk} = 0,80', zei hij, 'dan valt 20% van de productie buiten de controlewaarden.'
Wat vindt u daarvan?
Volgens ons is dit alleen waar als de verdeling uniform is, gecentreerd rond het gemiddelde en als $0 < C_{pk} < 1$. Wij wilden de trainer niet in het openbaar tegenspreken, maar vroegen hem bij het beoordelen van de eerste versie van onze C_{pk}-tool om waarden als C_{pk} = 0 te demonstreren met de tool. Zijn eerste reactie was: dan is het hele proces buiten de controlewaarden. Maar toen hij dit uitprobeerde met de computertool, bleek zijn verwachting tot zijn verrassing niet te kloppen: C_{pk} is 0 als het gemiddelde samenvalt met een van de specificatiegrenzen (***zie figuur 6 of 7***) en dan is de helft van het proces nog steeds binnen de specificatiegrenzen. Hij vroeg ons toen deze vraag ook in het cursusmateriaal op te nemen.

SPC-trainingen

De trainers verzorgden het grootste deel van de training, wij alleen het gedeelte over de procescapaciteitsmaten met de software. Na een korte introductie maakten deelnemers opgaven als:
- Schat de C_p en C_{pk} van dit proces.
- Maak grafieken die horen bij een C_{pk} van 1, 2, -0,5 en 0.
- Verander je grafiek en laat je buurman de C_p en C_{pk} van het proces raden.

Ook was er ruimte voor discussie zodat deelnemers het geleerde aan hun praktijk konden koppelen. Zo was er een groep die protesteerde bij de aanname dat data normaal verdeeld zijn, want in hun ervaring was dit zelden het geval. De technische uitleg van de trainer over de centrale limietstelling en transformaties ging uiteraard over de meeste hoofden heen. Maar de deelnemers aan de SPC-trainingen bleken snel in staat om C_p en C_{pk} te schatten aan de hand van ons cursusmateriaal. Zo begreep de meerderheid welke grafiek bij $C_{pk} = 0$ hoorde.

Evaluatie

Uit de evaluatie blijkt dat de deelnemers de tools en bijbehorende activiteiten zeer waardeerden. Een greep uit de reacties.
Een deelnemer schreef dat hij nu eindelijk begreep waar men het over had als het in vergaderingen over C_{pk} ging:
Now when I go to meetings I know what they are talking about, whereas before I would think: 'What the hell is a C_{pk}?', whereas now I know and understand what they're talking about.
Men waardeerde het dat de tools visueel waren en niet veel wiskundige kennis vereisten:
- It is a lot more interesting to use the tools rather than just doing the algebra.
- It is faster to use the tools because it calculates everything automatically.
- The visual aid also helps to see what is happening, instead of numbers alone.
Verder rapporteerden enkelen dat hun zelfvertrouwen was toegenomen.

De trainers waren erg trots op de tools die waren ontwikkeld. Ze demonstreerden de tools op internationale bijeenkomsten en conferenties, en werden daarna steeds bestormd door aanwezigen die de tools ook wilden hebben. Dat kan, ga naar onze website (zie onder) en registreer. Onze enige voorwaarde is dat u ons vertelt wat uw ervaringen zijn.

Toepassing van het geleerde

We vroegen enkele deelnemers ook of ze de opgedane kennis hadden toegepast. Dennis gaf ons een mooi voorbeeld. Als supervisor kreeg hij te maken met het volgende probleem. Bij sommige auto's kwam er niet genoeg water uit de achterruitsproeier. Wat bleek bij nadere inspectie? Als de slang te lang was, duwden operators het uitstekende stukje gewoon terug in het dak van de auto, zich niet realiserend dat bij sommige lengtes wel eens een knik in de slang zou kunnen ontstaan. Denkend aan de SPC-training besloot Dennis de slangen van de container naar de achterruitsproeier te meten. De leverancier had de opdracht gekregen slangen van 280 cm aan te leveren, maar er bleek nogal wat variatie in de lengtes te zijn. Dennis vergeleek de frequentieverdeling van zijn data vervolgens met de C_p-tool en vertelde ons:
I look at your tool, put my numbers against yours, and drag them sideways, and just watch the waves spread into a little ripple, and then, you know, it's miles out of spec [specification].
Ook realiseerde hij zich door de vergelijking met de tool dat hij de leverancier nooit specificatiewaarden had gegeven, alleen een streefwaarde van 280 cm. Met de data in de hand onderhandelde Dennis toen met de leverancier over de specificatiewaarden die acceptabel waren voor de autofabrikant.

Meer statistiek in het onderwijs?

We waren in het onderzoek op zoek naar voorbeelden van en gebrek aan technisch-wiskundige geletterdheid. Wat ons opviel was dat een groot deel van onze voorbeelden eerder statistisch dan wiskundig van aard was.

In bedrijven waar SPC wordt gebruikt moeten werknemers, zelfs de laagopgeleide, het verschil weten tussen de streefwaarde en het feitelijke gemiddelde. Ze moeten ook weten wat een voortschrijdend gemiddelde is, wat specificatiegrenzen (die niet statistisch zijn) en controlewaarden zijn (die afgeleid zijn van de standaarddeviatie). Ze moeten weten dat er zoiets is als random variatie, wat de oorzaken van variatie zijn, en dat er patronen kunnen zijn. Een basaal begrip van verdeling (vooral de normale verdeling) is nodig, evenals enig inzicht in trends en cyclische effecten. Het denken in afhankelijke en onafhankelijke variabelen bleek ook een nuttige vaardigheid te zijn die weinig werknemers zich eigen hadden gemaakt.

De statistische kennis die werknemers nodig hebben, verschilt vaak van de kennis die op school onderwezen wordt. Dit is deels onvermijdelijk: scholen bereiden leerlingen voor op veel verschillende beroepen. Toch zijn er ook technieken die in veel werksituaties bruikbaar zijn en voor zover wij weten niet op school onderwezen worden, statistische procescontrole (SPC) bijvoorbeeld. Ons advies is daarom –zeker in het beroepsonderwijs– om meer aandacht te besteden aan bovengenoemde statistische inzichten die aan SPC en andere technieken ten grondslag liggen.

Over de auteurs

Dr. Arthur Bakker was drie jaar *research officer* in het project *Techno-mathematical Literacies in the Workplace*, dat gesubsidieerd werd door de *Economic and Social Research Council (Teaching and Learning Research Programme)*. Hij werkt sinds 2006 weer aan het Freudenthal Instituut, nu als postdoctoraal onderzoeker.
E-mailadres: *A.Bakker@fi.uu.nl*
Dr. Phillip Kent was vier jaar *research officer* in het project. Prof. Celia Hoyles en Prof. Richard Noss waren de *co-directors*.
De software is geprogrammeerd door Chand Bhinder in Flash en kan worden gedownload van *www.lkl.ac.uk/technomaths/tools/spctools.*

De kansrekening van verjaardagen

[Rob Bosch]

Op dezelfde dag jarig

Opgaven in de kansrekening, zelfs betrekkelijk eenvoudige, geven vaak een resultaat dat niet strookt met onze intuïtie. Een opgave waarvan de uitkomst door velen als verbluffend wordt ervaren, is het bekende verjaardagenprobleem:

Hoe groot moet een groep mensen minstens zijn, zodat de kans dat er twee of meer mensen op dezelfde dag jarig zijn groter is dan 1/2?

Bij deze opgave nemen we gemakshalve aan dat een jaar 365 dagen telt en dat geboortedagen gelijkmatig over het jaar verdeeld zijn. Uiteraard zijn er (dan) in een groep van 366 personen altijd twee personen op eenzelfde dag jarig, terwijl in een groepje van 20 mensen de kans niet groot lijkt dat zoiets optreedt. Immers, de 20 mogelijke verjaardagen maken nog geen 6% uit van het totaal aantal mogelijke verjaardagen. Alvorens verder te lezen kan de lezer zijn idee over het antwoord op papier zetten, om het later te vergelijken met het door berekening verkregen antwoord.

Berekening

We berekenen de kans dat minstens twee verjaardagen op eenzelfde dag vallen in een groep van k personen. Voor de berekening stellen we ons een vaas met 365 briefjes voor. Op ieder briefje is een dag van het jaar geschreven. De k personen trekken nu achtereenvolgens een briefje. De datum op een getrokken briefje vatten we op als de verjaardag van diegene die het briefje trekt. Uiteraard wordt na de trekking het briefje weer in de vaas gestopt. Omdat er bij elk van de k onafhankelijke trekkingen 365 mogelijkheden zijn, is het aantal mogelijke verdelingen van de briefjes (verjaardagen) gelijk aan: 365^k (1)

We kijken nu naar het aantal mogelijkheden waarbij *geen* briefje twee maal wordt getrokken (geen dubbele verjaardagen). De eerste persoon heeft de keuze uit 365 briefjes. Omdat dit briefje niet meer getrokken mag worden, blijven er voor de tweede persoon nog 364 mogelijkheden over. De derde persoon heeft daarna nog keuze uit 363 briefjes enz. Voor de laatste persoon resteren nog $(365 - k + 1)$ briefjes. Het totaal aantal verdelingen waarbij geen briefje dubbel voorkomt, is dus:

$365 \cdot 364 \cdot 363 \cdots (365 - k + 1)$ (2)

Uit (1) en (2) berekenen we de kans dat in de k trekkingen geen enkel briefje meer dan eenmaal voorkomt. Deze kans is:

$$\frac{365 \cdot 364 \cdot 363 \cdots (365-k+1)}{365^k} \quad (3)$$

De kans p dat in de k trekkingen één of meerdere briefjes minstens tweemaal getrokken worden, is:

$$p = 1 - \frac{365 \cdot 364 \cdot 363 \cdots (365-k+1)}{365^k} \quad (4)$$

Dit is de kans dat er in een groep van k personen minstens twee personen op dezelfde dag jarig zijn.

We kunnen nu voor verschillende waarden van k (de groepsgrootte) de kans p uitrekenen (een programma als Maple is hierbij wel erg handig) en zien voor welke waarde van k deze kans voor het eerst boven ½ uitkomt. In onderstaande tabel zijn voor enkele waarden van k de bijbehorende kansen p gegeven.

Verjaardagstabel

$k = 5$	$k = 10$	$k = 20$
0,0271	0,1170	0,4144
$k = 30$	$k = 50$	$k = 75$
0,7063	0,9704	0,9997

Uit de tabel zien we dat de gevraagde waarde tussen de 20 en 30 ligt. Verdere berekening toont:

Verjaardagstabel

$k = 22$	$k = 23$
0,47757	0,5073

Bij een groepsgrootte van 23 is de kans al groter dan ½ dat minstens twee mensen uit deze groep op dezelfde dag jarig zijn! Dit resultaat is inderdaad verbluffend: in een klas van 23 leerlingen kom je met een kans van meer dan 50% twee leerlingen tegen met dezelfde verjaardag. De verjaardagstabel laat nog meer opmerkelijke kansen zien. Bij een voetbalwedstrijd is de kans meer dan 70% dat twee spelers (inclusief reservebank) op dezelfde dag jarig zijn. Op een school met 75 personeelsleden is het zo goed als zeker (99,97%) dat twee medewerkers een gemeenschappelijke verjaardag hebben.

Intuïtie als slechte raadgever

Als je aan iemand, onbekend met het bovenstaande resultaat, vraagt een groepsgrootte te noemen waarvoor het in de opgave gestelde geldt, dan krijg je meestal een antwoord in de orde van grootte van 182. De intuïtieve redenering die daarbij waarschijnlijk gevolgd wordt, is het idee dat we een kans van 50% willen hebben en dat daarbij de helft van het aantal dagen hoort, en dus een groep van ongeveer 182 personen. Deze redenering is fout, zoals we hebben gezien. Als verklaring voor dat foute antwoord hoor je vaak zeggen dat 182 het goede antwoord is, maar op een *andere* vraag - namelijk de vraag hoe groot een groep minstens moet zijn zodat de kans minstens ½ is dat er iemand is met een verjaardag op een *bepaalde* dag. Het

intuïtieve argument lijkt inderdaad beter te passen bij deze vraag. Immers, als ik denk aan de verjaardagsbriefjes, dan wil ik een bepaald briefje trekken. Bij 365 briefjes zal het gemiddeld 182 trekkingen vergen om het bewuste briefje eruit te halen. We gaan hieronder na of dit inderdaad het geval is.

Jarig op 4 oktober

We kiezen een bepaalde dag van het jaar, zeg 4 oktober, en berekenen de kans dat er iemand in een groep van k personen op die dag jarig is. Daartoe vullen we weer een vaas met 365 briefjes en rekenen de kans uit dat in k trekkingen (*met* terugleggen) een bepaald briefje te voorschijn komt. De kans dat iemand het bewuste briefje *niet* trekt is $\frac{364}{365}$. Bij k onafhankelijke trekkingen is de kans dat het 4-oktober-briefje niet getrokken wordt gelijk aan $\left(\frac{364}{365}\right)^k$.
De kans q dat dit briefje wèl één van de getrokken briefjes is, is dus gelijk aan:

$$q = 1-\left(\frac{364}{365}\right)^k \qquad (5)$$

Dit is de kans dat iemand uit de groep op de afgesproken dag, 4 oktober, jarig is.
Om uit te vinden bij welke groepsgrootte deze kans ½ overschrijdt, maken we weer een tabelletje:

Verjaardagstabel

k = 5	k = 10	k = 20	k = 50
0,0136	0,0271	0,0534	0,1282
k = 100	k = 200	k = 250	
0,2399	0,4223	0,4964	

Uit de tabel zien we dat zelfs 250 personen niet voldoende zijn voor een kans van 1/2. Verdere berekening toont:

Verjaardagstabel

k = 252	k = 253
0,4991	0,5005

Bij 253 personen komt de kans voor het eerst boven de ½. We hebben dus een groep van minsten 253 personen nodig voor een kans van ½ op een bepaalde verjaardag binnen de groep!
Merk het enorme verschil op tussen het antwoord op de eerste en de tweede vraag. Het is ook opmerkelijk dat men in het algemeen het aantal personen in de eerste vraag schromelijk overschat, terwijl men het aantal personen in het tweede probleem juist onderschat. Onze intuïtie blijkt bij dit soort vragen geen goede raadgever.

Tot slot

De twee verjaardagopgaven zijn voorbeelden van de volgende algemene problemen.
Een vaas is gevuld met n briefjes met daarop de getallen 1, 2, 3, …, n. We trekken, met terugleggen, k briefjes uit de vaas.

a. Hoe groot is de kans dat niet alle briefjes in de trekking verschillend zijn?
b. Hoe groot is de kans dat een bepaald briefje (zeg het briefje met nummer 1) minstens 2 maal wordt getrokken?

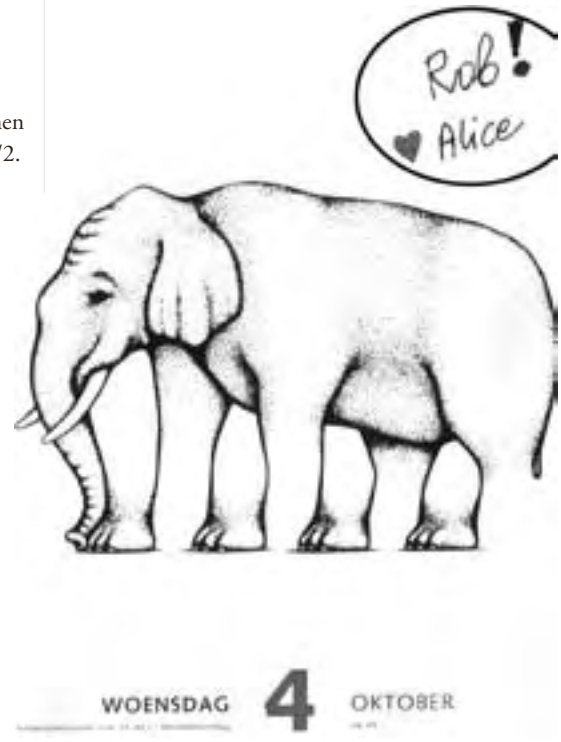
De kans onder a is gelijk aan:

$$P_a = 1-\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} \qquad (6)$$

Het antwoord op vraag b is:

$$P_b = 1-\frac{(n-1)^k}{n^k} = 1-\left(1-\tfrac{1}{n}\right)^k \qquad (7)$$

Nemen we n = 365, dan vinden we de eerder genoemde kansen terug.
Kiezen we n = 52 of n = 12, dan kunnen we met (6) de vraag beantwoorden hoe groot een groep moet zijn zodat er minstens twee mensen in dezelfde week respectievelijk dezelfde maand jarig zijn.
Substitutie van deze waarden in (7) geeft de kans dat minstens één persoon in de groep in dezelfde week respectievelijk zelfde maand jarig is als u of ik.
De lezer kan nagaan dat ook deze antwoorden verrassend zijn. Uiteraard zijn er meerdere variaties mogelijk, die we ook graag aan de lezer overlaten.



Literatuur

- William Feller: *An Introduction to Probability Theory and its Applications*. New York: Wiley (third edition, 1968)
- Henk Tijms: *Understanding Probability*. Cambridge: Cambridge University Press (second edition, 2007).

Over de auteur

Rob Bosch is redacteur van *Euclides* en universitair hoofddocent wiskunde aan de Nederlandse Defensie Academie te Breda.
E-mailadres: *robbosch@live.nl*

Nieuw: Moderne wiskunde 9

Volledig herziene editie voor vmbo

9e editie voor vmbo

- Veel praktische wiskunde
- Extra aandacht voor rekenvaardigheden
- Afwisselend en motiverend
- Ook volledig digitaal beschikbaar

Meer informatie op www.modernewiskunde.wolters.nl

Wolters-Noordhoff

Thuisvoordeel en statistiek

[Ruud Koning]

Inleiding

Thuisvoordeel is een zeer bekend fenomeen uit de sport. Sporters die thuis spelen, halen in het algemeen meer succes, en dat heeft er toe geleid dat oud-minister Winsemius er zelfs een boekje voor managers over heeft geschreven: 'Speel nooit een uitwedstrijd.' Dit roept wel de vraag op hoe groot thuisvoordeel nu eigenlijk is. Als je verder thuisvoordeel echt wilt gebruiken, is het ook nuttig te weten wat nu de oorzaken zijn van dat voordeel.

In deze bijdrage gaan we nader in op meting van het thuisvoordeel. Het leidende voorbeeld is meting van thuisvoordeel in de Nederlandse eredivisie in het seizoen 2006-2007, maar dat voorbeeld kan ook voor andere sporten (en natuurlijk andere seizoenen) worden gebruikt.

Thuisvoordeel wordt in het algemeen toegerekend aan vier verschillende factoren:

- steun van het (thuis)publiek;
- bekendheid met de locale omstandigheden;
- reistijd;
- spelregels.

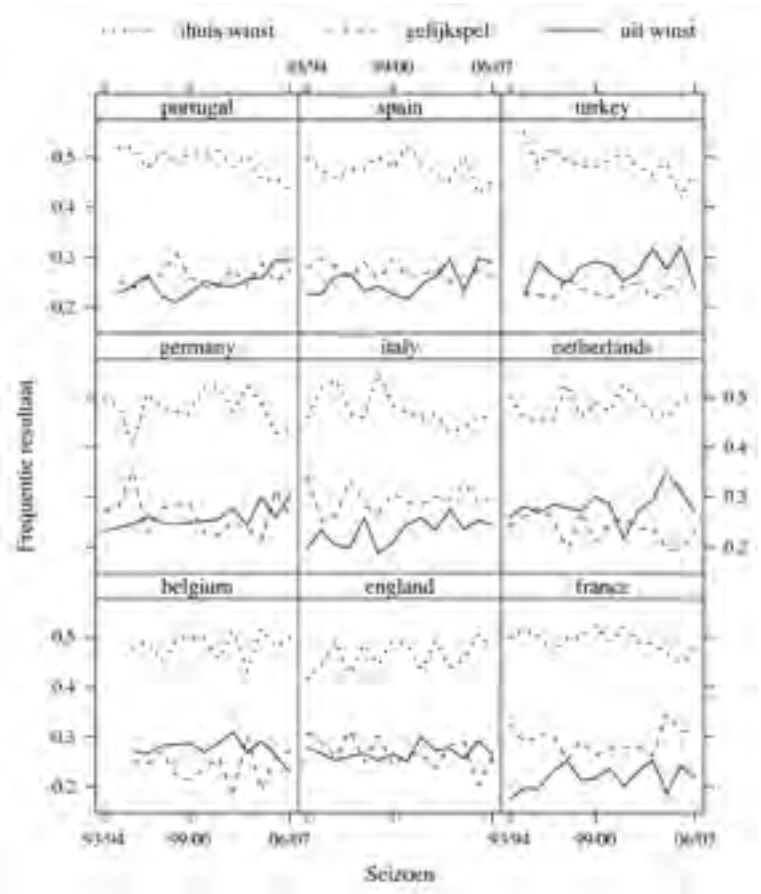
De eerste factor behoeft weinig toelichting. Een voorbeeld van de tweede factor is de bekendheid van een sporter met de belijning in zijn eigen sporthal, terwijl die in een andere sporthal anders kan zijn. Reizen kan vermoeiend zijn, dus het bezoekende team of de bezoekende sporter heeft op dit gebied een nadeel. In sommige sporten bieden de spelregels bepaalde voordelen aan het team dat thuis speelt. Zo moeten uitspelende honkbalteams in de Verenigde Staten aan slag beginnen, zodat het thuisteam steeds weet hoeveel punten het moet maken om de wedstrijd te winnen.

Een recent overzicht van thuisvoordeel in verschillende sporten wordt gegeven door Stefani (2007). Hij analyseert alleen teamsporten, en definieert thuisvoordeel als de mate waarin een team thuis beter presteert dan in uitwedstrijden. Thuisvoordeel wordt gemeten door het verschil te nemen van het percentage gewonnen wedstrijden door het thuisteam en het percentage verloren wedstrijden. Als in een competitie 100 wedstrijden worden gespeeld, waarvan 55 door het thuisteam worden gewonnen, 10 in een gelijkspel eindigen, en 35 door het thuisteam worden verloren, is het thuisvoordeel 20% (= 55% – 35%). Stefani vindt dat rugby de sport is met het grootste thuisvoordeel (25,1%), gevolgd door voetbal (21,7%), NBA basketbal (21,0%), American Football (17,5%), NHL ijshockey (9,%) en MLB honkbal (7,5%). De verschillen tussen de sporten zijn groot te noemen, waarbij opvalt dat de sporten met het grootste thuisvoordeel continu zijn: spelers beginnen en zijn in het algemeen pas uitgespeeld bij het eindsignaal. Vermoeidheid ten gevolge van reizen kan een verklaring zijn voor dat grote thuisvoordeel.

Deze aanpak om thuisvoordeel te meten heeft één groot voordeel, namelijk eenvoud, en twee nadelen: de maatstaf zegt niets over mogelijke variatie van thuisvoordeel tussen verschillende teams, en de maatstaf is ook niet goed toepasbaar in individuele sporten.

Voetbal

Thuisvoordeel in voetbal is van alle tijden, en van alle landen. Dit wordt nog eens duidelijk geïllustreerd ***in figuur 1***. Ruwweg de helft van alle wedstrijden wordt gewonnen door het thuisspelende team, een kwart eindigt in een gelijkspel, en een kwart wordt gewonnen door het team dat uit speelt. Echter, dit zijn gegevens op geaggregeerd niveau, voor een hele competitie over langere tijd. Voetballiefhebbers weten echter dat het soms spookt in de Euroborg, terwijl elders wordt geschamperd over 'Ajax-publiek'. Thuisvoordeel zal niet hetzelfde zijn voor elke club, dus een iets fijnzinniger maat is nodig.



figuur 1 Thuisvoordeel in verschillende nationale competities in Europa

Een praktische procedure om thuisvoordeel te meten voor individuele teamsporten in een volledige competitie is voorgesteld door Clarke en Norman (1995). Zij gaan ervan uit dat het resultaat in een wedstrijd is toe te dichten aan twee factoren: thuisvoordeel en kwaliteitsverschil. Beide factoren worden voor elk team geschat. Nu zit er wel een addertje onder het gras als je thuisvoordeel voor individuele teams gaat meten.
Stel dat de competitie uit drie teams A, B, en C zou bestaan. Team A is het sterkste team, en wint zowel thuis als uit van B met 2-1, en van C met 3-1. Team B wint thuis en uit van C met 2-1. Er is geen thuisvoordeel, en alle teams scoren thuis even veel als uit, en krijgen thuis even veel goals tegen als uit. De doelsaldo's van de drie teams staan vermeld in de eerste twee kolommen van ***tabel 1***.

	geen thuisvoordeel		thuisvoordeel C	
	HGD	AGD	HGD	AGD
team A	3	3	3	1
team B	0	0	0	2
team C	-3	-3	1	-3

tabel 1 Doelsaldo in thuis- en uitwedstrijden (zonder en met thuisvoordeel voor team C)

Er is geen thuisvoordeel, het doelsaldo van alle teams in thuis- en uitwedstrijden is even groot. Nu krijgt team C een thuisvoordeel van twee doelpunten, dus het speelt thuis gelijk tegen team A, en wint met 3-2 van team B. Het doelsaldo van team A in thuiswedstrijden is nu +3, en in uitwedstrijden is het +1, zoals ook in de laatste twee kolommen van ***tabel 1*** staat. Echter, het lijkt nu alsof team A *ook* een thuisvoordeel heeft, terwijl dit niet het geval is. Het verschil in doelsaldo tussen thuis- en uitwedstrijden is geen goede maatstaf voor *individueel* thuisvoordeel, aangezien het ook het gemiddelde thuisvoordeel in de gehele competitie bevat.
Clarke en Norman stellen dus een andere maatstaf van thuisvoordeel voor. Allereerst modelleren zij het doelpuntenverschil in een wedstrijd tussen teams i (thuis) en j (uit) als volgt:

$$w_{ij} = u_i - u_j + h_i + \varepsilon_{ij} \qquad (1)$$

In deze vergelijking is w_{ij} het doelpuntenverschil in een wedstrijd. Dit is positief als het thuisteam wint, nul als de wedstrijd in een gelijkspel eindigt, en negatief als het uitteam wint. Dit verschil hangt af van drie factoren: het verschil in kwaliteit tussen beide teams ($u_i - u_j$), het thuisvoordeel van team i (h_i), en toevalsfactoren (het weer, een onterechte strafschop gegeven door de scheidsrechter) die worden gemodelleerd met een stochastische storingsterm (ε_{ij}), die verwachting 0 heeft en gelijke variantie voor elke waarneming. De verwachting van het resultaat is dus $w_{ij} = u_i - u_j + h_i$, ofwel, het verwachte resultaat hangt af van het kwaliteitsverschil ($u_i - u_j$) en het thuisvoordeel van het thuisspelende team h_i. Het resultaat van een individuele wedstrijd wordt uiteraard ook nog bepaald door het toeval, ε_{ij}, maar uiteindelijk zijn we meer geïnteresseerd in de parameters u en h, die kwaliteit en thuisvoordeel gedurende een heel seizoen weergeven, dan in toeval dat soms een belangrijke rol speelt bij de bepaling van de uitslag van een individuele wedstrijd. In zekere zin is het ontwikkelen en gebruiken van thuisvoordeel ook een kwaliteit, maar dat bedoelen we niet met de parameter u_i.
u_i meet de kwaliteit van een team, als op een neutraal terrein zou worden gespeeld. Het gemiddelde van alle u's is 0, zodat een team met een positieve u beter is dan het gemiddelde team, en een team met een negatieve u is slechter. Thuisvoordeel heeft nu ook een natuurlijke interpretatie: het is het verwachte doelpuntenverschil als beide teams even goed zijn (dus als $u_i - u_j = 0$).
De restrictie dat de som van alle kwaliteitsparameters 0 is, is noodzakelijk, anders kan het model niet worden geschat. Als in vergelijking (1) alle u's met een constante worden verhoogd, verandert het verschil niet, en dus wordt de kansverdeling van de geobserveerde grootheid (het doelpuntenverschil) dan niet uniek bepaald door de parameters. Er zijn dan meer parameters dan we kunnen schatten op basis van de gegevens (het is net zo min mogelijk het gewicht van leerlingen onder te verdelen in gewicht van de armen en benen, en het gewicht van de overige lichaamsdelen, als men alleen het totaalgewicht meet). De identificerende restrictie $\sum_i u_i = 0$ lost dit probleem op. De keuze $\sum_i u_i = 0$ is aantrekkelijk, omdat de kwaliteitsparameters u_i en de parameters h_i die het thuisvoordeel meten, dan kunnen worden geïnterpreteerd als afwijkingen ten opzichte van een gemiddeld niveau. Een andere identificerende restrictie zou kunnen zijn $u_1 = 0$, dus de kwaliteit van het eerste team is 0, en de kwaliteitsparameters van alle andere teams worden gemeten ten opzichte van team 1. Echter, het is informatiever om te weten dat een bepaald team een positieve u heeft, en dus beter dan gemiddeld is, dan dat dat team beter is dan team 1.
De parameters van het model kunnen worden geschat met behulp van de methode van de kleinste kwadraten. In het voorbeeld met drie teams ziet het model er als volgt uit:

$$\begin{pmatrix} 2-1 \\ 3-1 \\ 1-2 \\ 2-1 \\ 1-3 \\ 1-2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \\ -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ -2 & -1 & 0 & 0 & 1 \\ -1 & -2 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_A \\ u_B \\ h_A \\ h_B \\ h_C \end{pmatrix} + \begin{pmatrix} \varepsilon_{AB} \\ \varepsilon_{AC} \\ \varepsilon_{BA} \\ \varepsilon_{BC} \\ \varepsilon_{CA} \\ \varepsilon_{CB} \end{pmatrix} \qquad (2)$$

Aangezien $u_C = -u_A - u_B$ hebben we voor de wedstrijd A tegen C: $u_A - u_C = 2u_A + u_B$, zoals duidelijk is uit de tweede regel van de design matrix.
De methode van de kleinste kwadraten schat de parameters door de geschatte residuen zo klein mogelijk te maken:

$$\min_{u,h} \sum \left(w_{ij} - (u_i - u_j + h_i) \right)^2 \qquad (3)$$

De som in deze vergelijking wordt genomen over alle wedstrijden. De waarden voor u_i en h_i die deze doelstellingsfunctie minimaliseren, heten de *kleinste kwadratenschatters*. Als we het model in matrixnotatie schrijven zoals in vergelijking (2), krijgen we:

$$y = X\theta + \varepsilon$$

en de kleinste kwadratenschatter is dan:

$$\hat{\theta} = (X'X)^{-1}X'y$$

De schattingsresultaten van het voorbeeld staan ***in tabel 2***.

	geen thuisvoordeel		thuisvoordeel C	
	u	h	u	h
team A	1	0	1	0
team B	0	0	0	0
team C	-1	0	-1	2

tabel 2 Berekening thuisvoordeel in het voorbeeld

In de eerste twee kolommen staan de schattingen voor het geval team C geen thuisvoordeel heeft, in de laatste twee kolommen staan de schattingen voor de parameters wanneer dat wel het geval is. Duidelijk is dat team A het team met de beste kwaliteit is, gevolgd door teams B en C. Het thuisvoordeel van team C komt naar voren in de schatting voor h_C, die toeneemt van 0 naar 2. Overigens, de parameter u_C is niet geschat, maar uitgerekend op basis van de schattingen voor u_A en u_B.
Deze aanpak om kwaliteit en thuisvoordeel van elkaar te scheiden voor individuele teams kan ook op een andere manier worden uitgevoerd. We gebruiken hierbij ***tabel 3***, waarin de gegevens van de Nederlandse eredivisie van het seizoen 2006-2007 staan vermeld.

	team	HW	HD	HL	H.f	H.a	HGD	AW	AD	AL	A.f	A.a	AGD	GD	p	h	u
1	Ajax	12	3	2	44	12	32	11	3	3	40	23	17	49	75	0,257	1,535
2	AZ	10	6	1	44	13	31	11	3	3	39	18	21	52	72	-0,085	1,774
3	ADO Den Haag	2	4	11	19	36	-17	1	4	12	21	36	-15	-32	17	-0,805	-0,184
4	Excelsior	6	3	8	27	28	-1	2	3	12	16	37	-21	-22	30	0,570	-0,594
5	Feyenoord	10	5	2	29	24	5	5	3	9	27	42	-15	-10	53	0,570	-0,260
6	FC Groningen	8	4	5	32	26	6	7	2	8	22	28	-6	0	51	0,070	0,267
7	sc Heerenveen	10	4	3	35	14	21	6	3	8	25	29	-4	17	55	0,882	0,333
8	Heracles Almelo	7	6	4	27	19	8	0	5	12	5	45	-40	-32	32	2,320	-1,747
9	NAC Breda	6	7	4	22	21	1	6	0	13	21	33	-12	-11	43	0,132	-0,069
10	N.E.C.	8	3	6	22	20	2	4	5	8	14	24	-10	-8	44	0,070	0,045
11	PSV	15	0	2	53	14	39	8	6	3	22	11	11	50	75	1,070	1,156
12	Roda JC	11	2	4	29	14	15	4	7	6	18	22	-4	11	54	0,507	0,354
13	Sparta Rotterdam	6	5	6	20	24	-4	4	2	11	20	42	-22	-26	17	0,445	-0,642
14	FC Twente	13	3	1	47	15	32	6	6	5	20	22	-2	30	66	1,445	0,413
15	FC Utrecht	11	4	2	30	11	19	2	5	10	11	33	-22	-3	48	1,882	-0,722
16	Vitesse	7	4	6	30	23	7	3	4	10	20	32	-12	-5	38	0,507	-0,000
17	RKC Waalwijk	5	5	7	19	24	-5	1	4	12	14	36	-22	-27	27	0,382	-0,639
18	Willem II	7	2	8	21	27	-6	1	5	11	10	37	-27	-33	31	0,632	-0,931

tabel 3 Berekening individueel thuisvoordeel

Er namen 18 teams deel aan deze competitie, dus elk team speelt 17 thuiswedstrijden. In die tabel staan gegevens over thuiswedstrijden (in de kolommen met een *H*), en gegevens over uitwedstrijden (kolommen met een *A*). De eerste kolom is het aantal gewonnen thuiswedstrijden (*HW*), die wordt gevolgd door het aantal gelijk geëindigde thuiswedstrijden (*HD*) en het aantal verloren thuiswedstrijden (*HL*). Vervolgens wordt voor de thuiswedstrijden het aantal doelpunten voor (*H.f*), tegen (*H.a*) en het doelsaldo in thuiswedstrijden (*HGD*) gegeven. De laatste twee kolommen geven de schattingen voor thuisvoordeel h en kwaliteit u. Die zijn als volgt berekend:

1. H is het gemiddelde thuisvoordeel voor de gehele competitie, $H = \sum_i \frac{HGD_i}{17}$. In dit geval vinden we $H = 11$.
2. Het thuisvoordeel voor elk team is $h_i = \frac{HGD_i - AGD_i - H}{16}$, dus het verschil van het doelsaldo in thuis- en uitwedstrijden, verminderd met het gemiddelde thuisvoordeel, gedeeld door 16. Er moet door 16 worden gedeeld, omdat de waarnemingen $HGD_i - AGD_i - H$ aan twee restricties voldoen: $H = \sum_i \frac{HGD_i}{17}$ en $\sum HGD_i + \sum AGD_i = 0$.
3. Kwaliteit tenslotte is dat deel van het doelsaldo in thuiswedstrijden dat niet te danken is aan thuisvoordeel: $u_i = \frac{HGD_i - (18-1)\cdot h_i}{18}$.

De resultaten van deze rekenpartij staan in de laatste twee kolommen van ***tabel 3***. Het team met de hoogste kwaliteit was AZ Alkmaar, maar hun thuisvoordeel was laag. Als elke competitiewedstrijd op neutraal terrein zou zijn gespeeld, was AZ misschien wel kampioen geworden. Aan de andere kant zien we ook dat de kwaliteit van PSV lager is dan die van Ajax, maar dat het thuisvoordeel in Eindhoven de doorslag heeft gegeven. Uit de tabel is in elk geval duidelijk dat het ene team een veel groter thuisvoordeel heeft dan het andere team.
Het model is in staat om ruim 40% van de variatie in wedstrijdresultaten te verklaren. De geschatte residuen $\hat{\varepsilon}_{ij}$ volgen inderdaad bij benadering een normale verdeling.
De Clarke-Norman methode voor het schatten van thuisvoordeel is bruikbaar voor alle sporten die in een volledige competitie worden gespeeld, zoals voetbal, hockey, waterpolo, en basketbal. Een nadeel van deze methode is dat die niet bruikbaar is om thuisvoordeel in individuele sporten te schatten.

Conclusie

Thuisvoordeel is belangrijk in sport. Goede meting van thuisvoordeel is echter moeilijk, omdat sportwedstrijden niet als experiment worden uitgevoerd. De vorm van de competitie (volledige competitie met uit- en thuiswedstrijden, of een toernooivorm waarbij niet elke speler tegen elkaar speelt) bepaalt in grote mate in hoeverre thuisvoordeel meetbaar is, en welke maatstaf handig is. In het voorbeeld is gebleken dat thuisvoordeel niet gelijk is voor elk team in de Nederlandse eredivisie in het seizoen 2006-2007. De methode die is besproken om thuisvoordeel te schatten, is goed toepasbaar in sporten waarin een volledige competitie met uit- en thuiswedstrijden wordt gespeeld. Thuisvoordeel in individuele sporten, die vaak in toernooivorm worden gespeeld, is moeilijker. Zie bijvoorbeeld Koning (2008) voor meting van thuisvoordeel in tennis. Toch is ook voor verschillende individuele sporten het belang van thuisvoordeel aangetoond.

Literatuur

- S.R. Clarke, J.M. Norman (1995): *Home ground advantage of individual clubs in English soccer*. In: *The Statistician*, Vol. 44(4), pp. 509-521.
- R.H. Koning (2008): *Home advantage in professional tennis*. Manuscript.
- R. Stefani (2007): *Measurement and interpretation of home advantage*. In: J. Albert, R.H. Koning (red.), *Statistical Thinking in Sports*, Boca Raton: Chapman & Hall/CRC; pp. 203-216.

Over de auteur

Ruud H. Koning is bijzonder hoogleraar Sporteconomie aan de Faculteit Economie en Bedrijfskunde van de Rijksuniversiteit Groningen. Hij heeft econometrie gestudeerd, en is verbonden aan de vakgroep Economics & Econometrics. Meer informatie vindt u op *www.rhkoning.com.* E-mailadres: *r.h.koning@rug.nl*

Proefopzetten

[Emiel van Berkum]

Inleiding

Statistiek houdt zich onder andere bezig met het analyseren van data. Uit data wil men dan conclusies trekken of vragen beantwoorden. De manier waarop de data (waarnemingen) verzameld worden heet een proefopzet; dit is dus de manier waarop het experiment georganiseerd wordt. Het is heel belangrijk dit goed te doen. Men wil duidelijke en goede conclusies kunnen trekken. Bovendien wil men dit zo efficiënt mogelijk doen.

De manier waarop het experiment georganiseerd is kan veel beperkingen opleveren met betrekking tot de conclusies die getrokken kunnen worden; er kunnen verkeerde conclusies worden getrokken. De Engelsman Sir Ronald Fisher was de eerste die zich dat realiseerde. Hij ging in 1919 als statisticus werken bij Rothamsted Agricultural Experimental Station om data te analyseren die in de loop der jaren verzameld waren. Hier werd veel onderzoek gedaan naar variëteiten van gewassen. Het werd hem al snel duidelijk dat veel data nutteloos waren omdat er geen goede proefopzetten gebruikt waren. Hij heeft in zijn periode in Rothamsted de basis voor *Design of Experiments* gelegd en zijn boek met die naam is een klassieker in de statistische literatuur. Veel terminologie van *Design of Experiments* is ontleend aan de landbouw en de door Fisher ontwikkelde theorie wordt nog steeds in de landbouw toegepast. Maar niet alleen daar; op alle terreinen waar statistische experimenten plaatsvinden is de theorie van proefopzetten belangrijk. Bij onderzoek naar nieuwe medicijnen is het van belang na te denken over de opzet van het experiment (denk aan het placebo-effect).

Soms zijn gepaarde waarnemingen van belang, bijvoorbeeld bij het testen van nieuwe producten. Een proefpersoon kan makkelijker twee producten met elkaar vergelijken dan ze afzonderlijk beoordelen. Overal waar subjectieve beoordeling een rol speelt kan een gepaarde opzet nuttig zijn. Een geavanceerde vorm hiervan is de zogeheten 'conjoint analysis', een statistische techniek die bij marktonderzoek gebruikt wordt en die geschikt is om te bepalen wat de kenmerken zijn die maken dat een bepaald product verkozen wordt boven een ander. Proefopzetten wordt ook steeds vaker toegepast in een industriële omgeving. In dit artikel wordt een aantal belangrijke principes bij proefopzetten genoemd. Hierbij zal een voorbeeld van een industriele toepassing leidraad zijn.

Verloting

Een van de belangrijke principes bij proefopzetten is verloting van de volgorde van waarnemen. Als men besloten heeft hoe men waarnemingen gaat doen, is het van belang om de volgorde waarin men die waarnemingen gaat doen te verloten. Een eerste reden is dat meestal aangenomen wordt dat de waarnemingen onafhankelijk zijn. Het verloten van de volgorde is daarvoor van wezenlijk belang. Een tweede reden is dat op die manier het effect van niet-bekende factoren gemodelleerd wordt als een toevallige fout.

Een eenvoudig voorbeeld van een niet goed gekozen proefopzet is het volgende. Bij het bakken van tegels is veel uitval. Men wil onderzoeken hoe de temperatuur van de oven de kwaliteit van de tegels beïnvloedt. Daarom gaat men 10 keer tegels bakken bij 370 graden en 10 keer bij 340 graden. Het is echter erg kostbaar om telkens de temperatuur van de oven opnieuw in te stellen en daarom bakt men eerst tien keer tegels bij een temperatuur van 370 graden en daarna tien keer bij 340 graden. Men zag geen verschil in de kwaliteit van de tegels. Pas na verder experimenteren kwam men erachter dat de oven een langere opwarmperiode nodig had dan gedacht. Hoewel de knop op 370 graden stond was de werkelijke temperatuur in de oven minder dan 370 graden. Toen de waarnemingen bij 340 graden aan de beurt waren, was de opwarmperiode voorbij en was de werkelijke temperatuur gelijk aan die 340 graden. Hierdoor zag men geen verschil. Als een willekeurige (verlote) volgorde was aangehouden, zou de temperatuur van 370 graden niet 'benadeeld' zijn. Indien de volgorde van alle 20 waarnemingen was verloot waren er vermoedelijk wel verschillen zichtbaar geworden.

Bij een experiment zijn er altijd bronnen van variatie. Hierbij kan een aantal soorten variatie worden onderscheiden. Allereerst is er de variatie die men wil onderzoeken: bij de tegels verwacht men dat de temperatuur van de oven invloed heeft. Deze zorgt voor variatie in de kwaliteit van de tegels. Dit is een *gewenste* bron van variatie. Er zijn ook *ongewenste* bronnen van variatie. Men heeft allereerst te maken met onnauwkeurigheid bij het meten en met variatie in het experimenteel materiaal. Daarnaast kunnen er ook andere factoren een rol spelen. Zo kan het bij het bakken van tegels verschil uitmaken wie de oven bedient. De plaats van de tegels in de oven is mogelijk van belang. Dit zijn ongewenste bronnen van variatie. Zo kunnen er ook onbekende ongewenste bronnen van variatie zijn. Deze leveren een bijdrage aan de zogeheten restterm van het model. In het algemeen formuleert men een model waarbij het te onderzoeken kenmerk enerzijds afhangt van een aantal meetbare factoren, maar anderzijds ook van een toevallige fout: de restterm in het model. Deze restterm wordt stochastisch verondersteld en vaak wordt aangenomen dat hij normaal verdeeld is. Een eenvoudig model kan bijvoorbeeld zijn:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

waarbij x_1 de temperatuur aangeeft en x_2 de plaats in de oven (bijvoorbeeld $x_2 = 0$ betekent 'onder' en $x_2 = 1$ betekent 'boven'). Door de volgorde van het doen van de waarnemingen te verloten zorgt men ervoor dat de invloed van (on)bekende ongewenste fouten gezien kan worden als een bijdrage aan de restterm in het model.

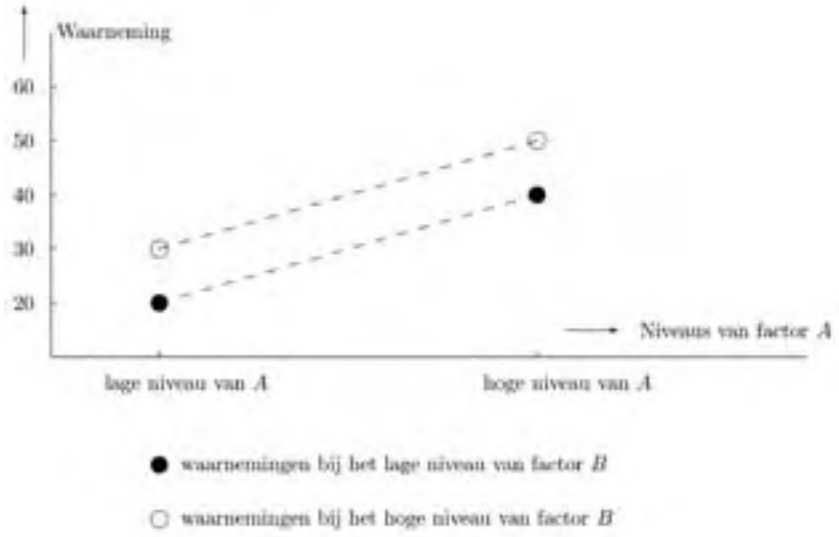
Blokvorming

De variatie die het gevolg is van bekende ongewenste factoren, kan in de analyse apart genomen worden door deze factoren in het model op te nemen. Het is van belang om daar in de proefopzet rekening mee te houden. In het voorbeeld van de tegels is genoemd dat er verschil kan zijn tussen de verschillende ploegen. Voor een goede analyse moeten dan de waarnemingen in gelijke mate over de verschillende ploegen verdeeld zijn. Dat heet een *gebalanceerde blokkenproef*. De ploegen zijn dan de blokken. Zo kan er nog een andere blokfactor zijn, bijvoorbeeld de verschillende partijen grondstof. De waarnemingen moeten dan ook in gelijke mate over de partijen verdeeld zijn. De theorie van het proefopzetten gaat onder andere over het construeren van efficiënte proefopzetten als er veel factoren een rol spelen.

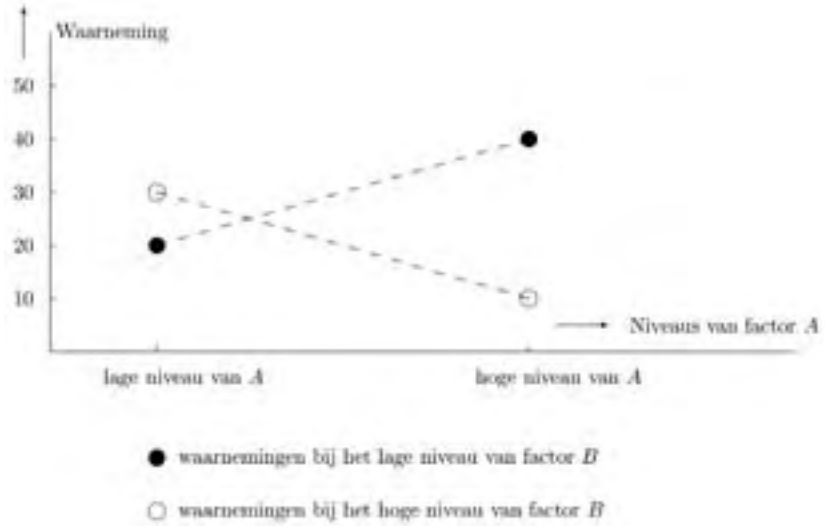
Proefopzetten die in dit kader genoemd kunnen worden zijn Latijnse vierkanten.

Partij	Operateurs				
	1	2	3	4	5
1	A	D	B	E	C
2	D	A	C	B	E
3	C	B	E	D	A
4	B	E	A	C	D
5	E	C	D	A	B

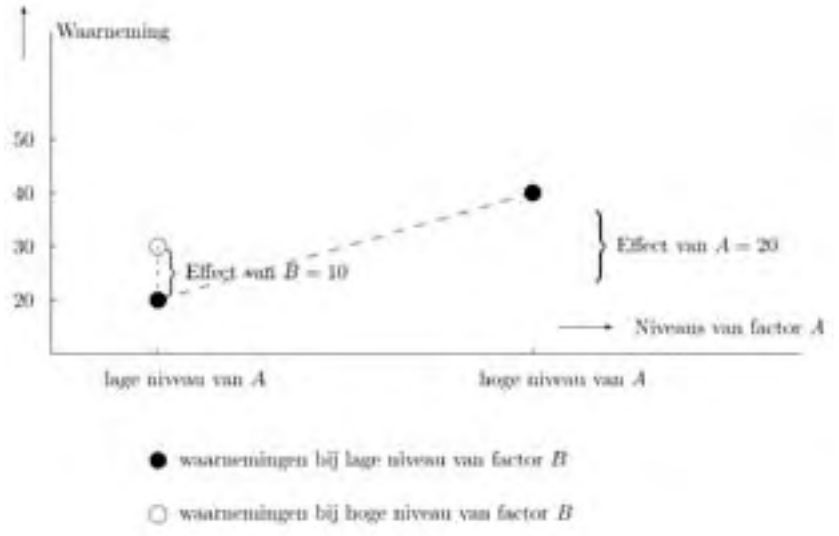
figuur 1 Latijns vierkant



figuur 2 Geen interactie



figuur 3 Interactie



figuur 4 One-factor-at-a-time opzet

Een Latijns vierkant van orde p heeft p rijen en p kolommen. In de p^2 cellen van het vierkant staan de eerste p letters van het alfabet zodanig dat iedere letter precies een keer voorkomt in elke rij en elke kolom. Een voorbeeld van zo'n proefopzet is het volgende.

Stel dat men in een experiment de invloed van de temperatuur op een kwaliteitskenmerk wil onderzoeken en dat er 5 verschillende temperaturen gekozen zijn in het experiment. Om het experiment uit te voeren zijn verschillende operators en verschillende partijen van de grondstof nodig. Deze operators en partijen vormen een ongewenste bron van variatie. Men wil de gegevens zo kunnen analyseren dat de variatie die het gevolg is van de operators en de partijen apart genomen kan worden en men goed conclusies kan trekken over de invloed van de temperatuur. Daarvoor is een Latijns vierkant handig. Met iedere rij wordt nu een partij geassocieerd (er moeten dus ook 5 partijen gekozen worden), met iedere kolom wordt een operator geassocieerd (dus ook 5 operators). De letters in het vierkant stellen de 5 temperaturen voor. ***In figuur 1*** staat dit weergeven.

Dit stelt een proefopzet voor. Als de letter B in rij 3 en kolom 2 voorkomt, betekent het dat men een waarneming gaat doen bij temperatuur B, partij 3 en operator 2. In plaats van 5 × 5 × 5 = 125 hoeft men nu slechts 25 waarnemingen te doen. Bovendien kan de variatie die het gevolg is van ongewenste bronnen van variatie in de analyse apart genomen apart genomen worden. Het meest eenvoudige voorbeeld van een blokkenproef is een experiment met gepaarde waarnemingen (zie de inleiding).

Herhaling

Herhalingen zijn nodig om iets over de mate van variatie te zeggen. Stel dat men bij een temperatuur van 340 graden één waarneming heeft en één waarneming bij 370 graden en laat de kwaliteit bij 340 graden 6,4 zijn en die bij 370 graden 6,7. Nu kan men niet zien of dit het gevolg is van toeval (bijvoorbeeld van de meetfout) of het gevolg van de temperatuursverandering. Als er meerdere waarnemingen zijn per temperatuur krijgt men een indruk van de meetfout en kan statistisch getoetst worden of het verschil van het toeval afhangt of niet.

Factoriële opzetten

Bij veel problemen kan een groot aantal factoren een rol spelen. In een industrieel probleem komt het regelmatig voor dat meer dan tien factoren mogelijk van belang zijn voor het te onderzoeken kwaliteitskenmerk. Om te onderzoeken welke factoren van belang zijn wordt dan een opzet gekozen waarbij slechts twee verschillende niveaus per factor worden gekozen. Een proefopzet die - helaas - in de praktijk nog veel gekozen wordt is een zogeheten *one-factor-at-a-time* opzet. Hierbij wordt vanuit een bestaande situatie telkens maar één factor gevarieerd. Het idee is dat men dan goed zou kunnen zien of die betreffende factor invloed heeft of niet. Voorbeeld: bij een chemisch proces wil men onderzoeken welke factoren de 'filtration rate' beïnvloeden. De te onderzoeken vier factoren zijn: temperatuur, druk, concentratie en roersnelheid. Bij de huidige instellingen wordt een waarneming gedaan. Vanuit de huidige instellingen worden dan om de beurt elk van de vier factoren gewijzigd om de invloed van elke factor afzonderlijk te kunnen meten. Zo krijgt men vijf waarnemingen. Zo'n opzet is niet efficiënt en bovenal kan hij leiden tot foute conclusies omdat men op deze manier geen interacties kan detecteren.

Interacties

We spreken van een interactie als de invloed van een factor afhangt van het niveau van een andere factor. Het volgende eenvoudige rekenvoorbeeld geeft dat aan.

We gaan uit van twee factoren A en B, elk op twee niveaus.

Factor A	Factor B	
	B_{laag}	B_{hoog}
A_{laag}	20	30
A_{hoog}	40	50

tabel 1

In tabel 1 staan de gemiddelden van de waarnemingen bij elk van de vier niveaucombinaties

Het effect van A wordt gedefinieerd als het verschil van het gemiddelde van de waarnemingen op het hoge niveau en het gemiddelde van de waarnemingen op het lage niveau. Het effect van A is dus:

$$A = \frac{40+50}{2} - \frac{20+30}{2} = 20$$

Het effect van B is:

$$B = \frac{30+50}{2} - \frac{20+40}{2} = 10$$

In figuur 2 staan de waarnemingen weergegeven. De bovenste stippellijn tussen de twee open cirkels verbindt de twee punten van de twee waarnemingen die op het hoge niveau gedaan zijn. De helling van de lijn geeft dus de invloed van factor A aan bij het hoge niveau van B. De helling van de andere lijn geeft de invloed van factor A bij het lage niveau van factor B aan.

De twee lijnen lopen hier evenwijdig. Dat betekent dat de invloed van factor A niet afhangt van het niveau dat factor B heeft. Er wordt gezegd dat factor A en factor B geen interactie hebben.

Stel nu dat één van de waarnemingen een andere waarde heeft; ***zie tabel 2***.

Factor A	Factor B	
	B_{laag}	B_{hoog}
A_{laag}	20	30
A_{hoog}	40	10

tabel 2

Het effect van A is nu dus

$$A = \frac{40+10}{2} - \frac{20+30}{2} = 0$$

Toch heeft factor A wel degelijk effect. Het effect van factor A hangt af van het niveau van factor B. ***In figuur 3*** is dit weergegeven. Hier snijden de lijnen elkaar. Er wordt gezegd dat factor A en factor B een interactie hebben.

Nu kan het gevaar van een *one-factor-at-a-time* duidelijk gemaakt worden.

Factor A	Factor B	
	B_{laag}	B_{hoog}
A_{laag}	20	30
A_{hoog}	40	

tabel 3

In een one-factor-at-a-time opzet worden slechts de waarnemingen gedaan van ***tabel 3***. ***In figuur 4*** worden de waarnemingen

weergegeven. Als men op grond hiervan een zo hoog mogelijk waarde wil bereiken, ligt het voor de hand om zowel het hoge niveau van *A* als het hoge niveau van *B* te kiezen. Als er een interactie is kan dat een verkeerde conclusie zijn.
Om de interacties te kunnen meten moet gekozen worden voor een factoriële opzet. Dat kan een volledige factoriële opzet zijn, maar dat is niet nodig. Bij een volledige factoriële opzet worden alle niveaucombinaties gekozen. In het voorbeeld van het chemisch proces zouden dan $2^4 = 16$ waarnemingen nodig zijn.

Fractionele factoriële opzetten

Bij een 2^k-proefopzet, dat wil zeggen een opzet met k factoren die elk twee niveaus hebben, is het aantal niveaucombinaties gelijk aan 2^k. In de praktijk kan het te kostbaar zijn om bij alle niveaucombinaties waarnemingen te doen. Het is niet ongebruikelijk dat het aantal factoren in een experiment 10 of nog meer is. Met de theorie van proefopzetten is het mogelijk om waarnemingen te doen bij slechts een fractie van de combinaties. Als men wil zorgen dat de proefopzet mooie eigenschappen heeft wat betreft de analyse, is het gewenst om als aantal waarnemingen een aantal te kiezen dat gelijk is aan een macht van 2, dus bijvoorbeeld de helft of een kwart van het totaal aantal combinaties. Zo'n opzet wordt een 2^{k-p}-proefopzet genoemd. Zo'n fractionele opzet is ook mogelijk als er niet zoveel factoren zijn. Ook in het geval van het chemisch proces kan men ervoor kiezen om niet alle 16 niveaucombinaties te kiezen, maar een halve fractie. Men doet dan dus 8 waarnemingen. Aan de hand van dit voorbeeld kan ook duidelijk gemaakt worden dat zo'n opzet veel efficiënter is dan een one-factor-at-a-time opzet. De one-factor-at-a-time opzet heeft 5 waarnemingen. Als men echter het effect van een factor wil meten, heeft men dus slechts een waarneming op het hoge niveau en een waarneming op het lage niveau. Dan kun je niet concluderen of het verschil een gevolg is van de meetonnauwkeurigheid of het gevolg van het effect van die factor. Daarom moet men de opzet minstens één keer herhalen. Dan telt de opzet 10 waarnemingen en heeft men toch slechts twee waarnemingen op het hoge niveau van de betreffende factor en twee op het lage niveau. De fractionele opzet met slechts 8 waarnemingen zit zo mooi in elkaar dat alle vier de waarnemingen op het hoge en alle vier op het lage niveau van de factor gebruikt kunnen worden om het effect van die factor te kunnen schatten. De fractionele opzet is dus veel efficiënter. Bovendien kan men in die opzet 3 van de 6 interacties van twee factoren schatten. Deze kunnen in de one-factor-at-a-time opzet helemaal niet geschat worden.

Conclusie

Bij een statistisch onderzoek is het van belang eerst goed na te denken over het experiment. Welke factoren kunnen van belang zijn? Hoeveel verschillende niveaus per factor moeten er gekozen worden? Bij welke niveaucombinaties moeten er waarnemingen gedaan worden? Hoeveel herhalingen moeten er gedaan worden? Te vaak wordt begonnen met waarnemingen zonder dat over deze vragen goed is nagedacht.
In het ergste geval kunnen effecten niet geschat worden die men had willen schatten en kunnen de verkeerde conclusies getrokken worden. Als het meezit heeft men slechts een inefficiënte proefopzet gekozen.

Noot

Dit artikel verscheen eerder in *STAtOR*, jaargang 6 nummer 4 (december 2005), en is voor *Euclides* licht bewerkt (zie *http://www.vvs-or.nl/stator).*

Literatuur

- G.W. Cobb (1997): *Introduction to Design and Analysis of Experiments.* New York: Springer.
- R.A. Fisher (1966): *The Design of Experiments,* 8th edition. New York: Hafner Publishing Company.
- R.L. Mason, R.F. Gunst, J.L. Hess (2003): *Statistical Design and Analysis of Experiments,* 2nd edition. Hoboken: Wiley-Interscience.
- D.C. Montgomery (2005): *Design and Analysis of Experiments,* 6th edition. New York: Wiley.
- P.M. Upperman (1974): *Statistische Methoden en Proefopzetten.* Rotterdam: Universitaire Pers.

Over de auteur

Emiel van Berkum is adjunct-directeur Service Onderwijs Wiskunde aan de Technische Universiteit Eindhoven.
E-mailadres: *e.e.m.v.berkum@tue.nl*

Deeltje 5 uit de Zebrareeks [1] is gewijd aan de Poisson-verdeling. Uit dit boekje neem ik, met andere getalletjes, de volgende toepassing over (zie [1], pag. 4, voorbeeld 1.1).

Stel dat in een boek drukfouten voorkomen met een gemiddelde van vier drukfouten per vijf bladzijden. Wat is de kans dat een gegeven bladzijde geen enkele drukfout bevat?

We nemen aan dat het aantal drukfouten per bladzijde een Poisson-verdeling volgt. Bij de Poisson-verdeling worden de kansen op $N = 0, 1, 2, \ldots$ drukfouten op een bepaalde bladzijde uitgerekend met een formule waarin het *gemiddeld* aantal drukfouten per bladzijde voorkomt. Bij *onze* opgave is dat gemiddelde gelijk aan 0,8 (4 drukfouten in 5 bladzijden). Pas op… 0,8 is een *gemiddelde* en moet niet verward worden met de *kans* op een drukfout per bladzijde. Als de stochast N het aantal drukfouten op een gegeven bladzijde aangeeft, dan berekenen we de kansen $P(N = n)$ met de volgende formule:

$$P(N=n)=\frac{\lambda^n}{n!}e^{-\lambda} \quad (n=0,1,2,\cdots) \qquad (1)$$

Hierin is λ het gemiddeld aantal drukfouten per bladzijde. In de opgave is, zoals gezegd, het gemiddeld aantal drukfouten per bladzijde λ gelijk aan 0,8. De gevraagde kans is dus gelijk aan:

$$P(N=0)=\frac{(0,8)^0}{0!}e^{-0,8}\approx 0,4493$$

Deze kans is eenvoudig uit te rekenen (opzoeken in een Poisson-tabel kan natuurlijk ook). Daarom nu een iets(?) lastiger sommetje:

Hoe groot is de kans op een *even* aantal drukfouten op een willekeurige pagina?

In de berekening die volgt, laten we λ gewoon even staan. Aan het eind van de berekening kunnen we dan $\lambda = 0,8$ substitueren om het numerieke antwoord te krijgen. Overigens, het aantal druk-

Even en oneven

[Rob Bosch]

Siméon Denis Poisson (1781-1840)

fouten op een bladzijde kan 0 zijn en aangezien 0 even is (deelbaar door 2), moeten we die kans meerekenen.
Voordat we de gevraagde kans uitrekenen, een vraag aan de lezer. Hoe groot denkt u dat deze kans ongeveer is?

Met formule (1) zien we dat de kans op een even aantal drukfouten gelijk is aan:
$$\sum_{n=0}^{\infty} P(N=2n) = \sum_{n=0}^{\infty} \frac{\lambda^{2n}}{(2n)!} e^{-\lambda} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^{2n}}{(2n)!}$$

Uitgeschreven wordt dit:
$$\sum_{n=0}^{\infty} P(N=2n) = e^{-\lambda}\left(1+\frac{\lambda^2}{2!}+\frac{\lambda^4}{4!}+\frac{\lambda^6}{6!}+\cdots\right)$$
Het rechterlid bevat de evenmachtstermen uit de ontwikkeling van e^{λ}. Deze termen krijgen we uit de ontwikkeling van e^{λ} en die van $e^{-\lambda}$:
$$e^{\lambda} = 1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\cdots \qquad (2)$$
$$e^{-\lambda} = 1-\frac{\lambda}{1!}+\frac{\lambda^2}{2!}-\frac{\lambda^3}{3!}+\cdots \qquad (3)$$
De som van (2) en (3) geeft:
$e^{\lambda}+e^{-\lambda} = 2\cdot\left(1+\frac{\lambda^2}{2!}+\frac{\lambda^4}{4!}+\frac{\lambda^6}{6!}+\cdots\right)$, zodat:
$\frac{1}{2}(e^{\lambda}+e^{-\lambda}) = 1+\frac{\lambda^2}{2!}+\frac{\lambda^4}{4!}+\frac{\lambda^6}{6!}+\cdots$
waaruit volgt:
$\frac{1}{2}e^{-\lambda}(e^{\lambda}+e^{-\lambda}) = e^{-\lambda}\left(1+\frac{\lambda^2}{2!}+\frac{\lambda^4}{4!}+\frac{\lambda^6}{6!}+\cdots\right)$

De kans op een even aantal drukfouten is dus:
$$\sum_{n=0}^{\infty} P(N=2n) = \tfrac{1}{2}e^{-\lambda}(e^{\lambda}+e^{-\lambda}) = \tfrac{1}{2}(1+e^{-2\lambda}) \qquad (4)$$
Substitutie van $\lambda = 0{,}8$ in (4) geeft het antwoord 0,6009.

Hoe goed was uw schatting van het antwoord? Had u wellicht een kans van minder dan een half verwacht?
We kunnen het antwoord nog controleren aan de hand van een Poisson-tabel voor $\lambda = 0{,}8$:

n	$P(N = n)$
0	0,4493
1	0,3595
2	0,1438
3	0,0383
4	0,0077
5	0,0012
6	0,0002

Uit deze tabel lezen we af dat:
$$\sum_{n=0}^{\infty} P(N=2n) = 0{,}4493+0{,}1438+0{,}0077+0{,}0002+\cdots \approx 0{,}6010$$
We keren ten slotte nog even terug naar de kans in (4), waar we vonden dat:
$\sum_{n=0}^{\infty} P(N=2n) = \frac{1}{2}(1+e^{-2\lambda})$ Omdat $1+e^{-2\lambda} > 1$ is voor alle λ, geldt onafhankelijk van het gemiddelde λ:
$\sum_{n=0}^{\infty} P(N=2n) > \frac{1}{2}$

Dit betekent:
Als het aantal drukfouten in een boek een Poisson-verdeling volgt, dan is de kans op een even aantal drukfouten op een willekeurige bladzijde altijd groter dan de kans op een oneven aantal drukfouten op die bladzijde!
Een resultaat om *even* bij stil te staan.

Literatuur

[1] Henk Tijms, Frank Heierman, Rein Nobel (2000): *Poisson, de Pruisen en de Lotto*. Utrecht: Epsilon Uitgaven.
[2] Frederick Mosteller (1987): *Fifty Challenging Problems in Probability*. New York: Dover Publications.

Over de auteur

Rob Bosch is redacteur van *Euclides* en universitair hoofddocent wiskunde aan de Nederlandse Defensie Academie te Breda.
E-mailadres: *robbosch@live.nl*

Pokeren: bluffen met wiskunde

[Ben van der Genugten]

Inleiding

Poker is een kaartspel van Amerikaanse origine, onder breed publiek bekend geworden door opwindende scènes in westerns. De laatste jaren is de belangstelling voor poker behoorlijk toegenomen. Een echte hype is de pokervariant *Texas Hold'em*, niet alleen doorgedrongen tot casino's maar ook te spelen op het internet.
Menigeen denkt bij poker aan allerlei psychologische trucs om tegenstanders te misleiden. Zo kan een speler *bluffen* door te bieden met een lage kaart, erop hopend dat tegenstanders met een betere kaart weggaan. Of ook juist het tegenovergestelde, hij kan *verleiden* door met een hoge kaart slechts mee te gaan om tegenstanders er toe te brengen te bieden met een mindere kaart, om dan vervolgens genadeloos toe te slaan door te over bieden. Is dit allemaal zuiver psychologie? Doel van dit verhaal is te laten zien dat een zakelijke wiskundige analyse vanuit de speltheorie tot zulke beslissingen kan leiden als onderdeel van verstandige, soms zelfs optimale strategieën.

Tweepersoonsstraightpoker

Texas Hold'em is een complex spel, te ingewikkeld om gedetailleerd de eraan ten grondslag liggende ideeën te illustreren. Daarom nemen we een eenvoudige pokervariant waarbij alleen aan het begin kaarten gedeeld worden en het spel verder alleen uit biedronden bestaat. Deze vorm staat bekend als *Straightpoker*. In de praktijk wordt dit weinig gespeeld omdat het wat saai gevonden wordt. Maar het is uitermate geschikt om ideeën aan te illustreren. We kiezen hier ook zeer eenvoudige spelregels. Om de kansberekening eenvoudig te houden delen we ook niet uit een volledig kaartspel echte pokerhanden van 5 kaarten, maar nemen we een kaartspel van 5 kaarten van dezelfde kleur met waarden T(ien) < B(oer) < V(rouw) < H(eer) < A(as). We bekijken in deze paragraaf het geval van twee deelnemers en in de volgende paragraaf dat van drie.

a) Spelverloop

We beschrijven het spelverloop in detail. Eerst storten beide spelers I en II elk een bedrag als begininzet in de pot, de *Ante*. We nemen als spelregel Ante = 1. Vervolgens krijgen ze elk een aselect getrokken kaart uit het spel van 5 kaarten. De hiernavolgende mogelijke spelverlopen zijn weergegeven ***in figuur 1*** op pag. 237. Speler I is het eerst aan de beurt (zie knoop 1 in de spelboom). Hij heeft de keuze tussen de beslissingen passen (*Pass*) en bieden (*Bet*).
Als speler I biedt, dan moet hij het biedbedrag in de pot stoppen, de Bet. We nemen als spelregel Bet = 2. Dan is speler II aan de beurt (zie knoop 7). Deze heeft hier de keuze tussen de beslissingen weggaan (*Fold*) en meegaan (*Call*). Als hij meegaat, dan moet hij ook de Bet in de pot stoppen: beide spelers hebben hier dan evenveel in gestopt. Dan laten beide spelers hun kaart zien (de *showdown*): de speler met de hoogste kaart wint de pot (eindknoop 9). Deze boekt in dit geval dus een winst van 3, en de ander een verlies van 3 (= winst van -3). Maar als speler II weggaat, dan volgt geen showdown en krijgt speler I als enig overgebleven speler de pot (eindknoop 8). Dus een winst 1 voor I (en een winst -1 voor II).
Als speler I past, dan is speler II aan de beurt (knoop 2). Hij zit dan in dezelfde situatie als speler I eerst (bij knoop 1): hij kan passen of bieden. In het geval hij biedt verloopt het spel als voorheen aangegeven: eindknoop 6 met showdown en een winst van 3 voor de winnaar, en eindknoop 5 zonder showdown met een winst van 1 voor speler II. Daarentegen, als speler II ook past, dan hebben alle spelers gepast en stopt het spel: beide spelers krijgen gewoon hun Ante uit de pot terug (eindknoop 3 met winst 0).
De naamgeving van de verschillende beslissingen bij poker is wat verwarrend. Bij *Texas Hold'em* heet passen schuiven (*Check*) en wordt met passen juist opgeven bedoeld! We doen dat niet: in de boom heeft elk type beslissing mooi een unieke beginletter.

b) Beginners onder elkaar

Hoe zullen spelers I en II het spel spelen? Welke strategieën zullen zij voeren? In dit verband is een spelstrategie van een speler een tabel die in iedere spelsituatie aangeeft welke beslissing hij neemt. De tabel van speler I moet dus voorschrijven wat hij in de knopen 1 en 4 doet, afhankelijk van zijn kaart. Een simpele overweging gebaseerd op *kansrekening* is als volgt.
Speler I in knoop 1 past als hij de lage kaart T of B heeft, omdat bij een eventuele showdown de kans dat hij wint, kleiner is dan die van speler II. Hij biedt bij een hoge kaart H of A omdat dan deze kans groter is. Bij de middelste kaart V zijn de kansen gelijk en daarom loot hij met gelijke kansen ½ tussen passen (P) en bieden (B); een dergelijke beslissing heet *gerandomiseerd*.
Voor knoop 4 zijn de overwegingen dezelfde met betrekking tot de keuze tussen weggaan (F) en meegaan (C). Een strategie die gerandomiseerde beslissingen bevat, heet *gemengd*. Een strategie met alleen niet-gerandomiseerde beslissingen (alleen kansen 0 of 1) heet *zuiver*. ***Tabel 1*** bevat de (gemengde) strategie van I en een soortgelijke strategie van II. Dit zijn simpele strategieën, uitsluitend ingegeven door kansen bij de showdown, en daarom aangeduid als 'beginner'. De tabel bevat de kansen behorend bij de beslissingen. Een kenmerk van zulke beginnersstrategieën is: altijd weggaan of passen (passief spelen) bij een lage kaart en altijd meegaan of bieden (actief spelen) bij een hoge kaart, zonder acht te slaan op beslissingen van tegenstanders.

Speler I als beginner					Speler II als beginner				
Kaart	(1) P	(1) B	(4) F	(4) C	Kaart	(2) P	(2) B	(7) F	(7) C
T	1	0	1	0	T	1	0	1	0
B	1	0	1	0	B	1	0	1	0
V	½	½	½	½	V	½	½	½	½
H	0	1	0	1	H	0	1	0	1
A	0	1	0	1	A	0	1	0	1
Winst tegen beginner II: $-0{,}05$ $(=-\frac{1}{20})$					Winst tegen beginner I: $+0{,}05$ $(=+\frac{1}{20})$				

tabel 1 Beginners onder elkaar

De gecombineerde tabel van strategieën van alle spelers heet een *strategieprofiel.* Een strategieprofiel legt de kansverdeling van de winsten van de spelers vast en dus ook de eruit volgende verwachte winsten (gemiddelde winsten op de lange duur). De berekening is eenvoudig maar bewerkelijk: sommeren over alle 5 × 4 = 20 kaartcombinaties en bij kaart V ook nog over de mogelijke uitkomsten van de gerandomiseerde beslissing. ***Tabel 1*** geeft ook deze verwachte winsten. De som van beide verwachte winsten is 0, omdat dit ook voor iedere spelrealisatie geldt. Speler I is in het nadeel want zijn verwachte winst is negatief.

c) Optimaal spel tegen beginners

Vanzelfsprekend wil speler I zich verbeteren. Hij zal tegen speler II een optimale strategie willen voeren, per definitie een strategie die zijn (verwachte) winst maximaliseert, gegeven de beginnersstrategie van II. Aangetoond kan worden dat er onder de optimale strategieën altijd een zuivere is: randomisatie bij optimaal spel tegen een gegeven beginnersstrategie is overbodig. ***Tabel 2*** geeft een optimale zuivere strategie en de bijbehorende winsten, zowel voor speler I als speler II.

Optimale speler I tegen beginner II				
Kaart	(1) P	B	(4) F	C
T	1	0	1	0
B	1	0	1	0
V	1	0	1	0
H	0	1	1	0
A	0	1	0	1
Winst van I: -0,05 ($=-\frac{1}{20}$)				

Optimale speler II tegen beginner I				
Kaart	(2) P	B	(7) F	C
T	0	1	1	0
B	0	1	1	0
V	0	1	1	0
H	0	1	1	0
A	0	1	0	1
Winst van II: +0,25 ($=+\frac{1}{4}$)				

tabel 2 Optimale strategieën tegen beginners

Vergelijk deze tabel 2 met tabel 1. De verschillen zijn opmerkelijk. De belangrijkste conclusie is dat beide spelers kunnen winnen tegen beginners als ze optimaal spelen. Maar dan moeten ze wel weten dat hun tegenstanders als beginners spelen!

d) Optimaal spel tegen elkaar

Wat kan speler I doen als hij de strategie van speler II helemaal niet kent? Als hij een bepaalde strategie kiest, dan is het ergste wat speler I kan overkomen dat speler II juist als antwoord hierop *die* strategie kiest die zijn eigen winst maximaliseert, en dus de verwachte winst van speler I minimaliseert. Daarom doet speler I er bij zijn keuze verstandig aan een strategie te kiezen die zijn minimale winst maximeert. Zo'n strategie heet een *maximinstrategie* en de bijbehorende minimale winst heet de maximinwaarde. Speler I heeft 1024 zuivere strategieën en oneindig veel gemengde. En juist het maximaliseren over ook de gemengde blijkt tot een veel hogere maximale waarde te leiden dan alleen maximaliseren over de zuivere. Dus is het zeker verstandig dit te doen. Maar dit is gemakkelijker gezegd dan gedaan. Gelukkig bestaan er allerlei computerprogramma's om zulke oplossingen met de computer te vinden. Door een maximinstrategie te spelen kan speler I zich verzekeren van een winst gelijk aan zijn maximinwaarde, wat speler II ook doet.

De overwegingen voor speler II zijn dezelfde. Door een maximinstrategie te spelen kan speler II zich verzekeren van een winst gelijk aan *zijn* maximinwaarde, wat speler I ook doet. Omdat de som van de winsten van beide spelers altijd 0 is, is dit hetzelfde als zijn verlies te kunnen beperken tot minus zijn maximinwaarde. Blijkbaar moet daarom altijd gelden: maximinwaarde (I) ≤ -maximinwaarde (II). ***Tabel 3*** geeft de maximinstrategieën van beide spelers.

Optimale speler I tegen optimale II				
Kaart	(1) P	B	(4) F	C
T	$\frac{15}{16}$	$\frac{1}{16}$	1	0
B	$\frac{15}{16}$	$\frac{1}{16}$	1	0
V	$\frac{15}{16}$	$\frac{1}{16}$	1	0
H	0	1	$\frac{1}{2}$	$\frac{1}{2}$
A	$\frac{1}{2}$	$\frac{1}{2}$	0	1
Winst van I: -0,1125 ($=-\frac{9}{80}$)				

Optimale speler II tegen optimale I				
Kaart	(2) P	B	(7) F	C
T	$\frac{3}{4}$	$\frac{1}{4}$	1	0
B	$\frac{3}{4}$	$\frac{1}{4}$	1	0
V	$\frac{3}{4}$	$\frac{1}{4}$	1	0
H	0	1	$\frac{5}{8}$	$\frac{3}{8}$
A	0	1	0	1
Winst van II: +0,3125 ($=+\frac{5}{16}$)				

tabel 3 Optimale strategieën tegen elkaar

Het meest opmerkelijke in de tabel is wellicht dat maximinwaarde (I) = -maximinwaarde (II), ofwel de winst van speler I is gelijk aan het verlies van speler II. Deze gemeenschappelijke waarde noemt men de *spelwaarde*. Speler I kan door een maximinstrategie (in termen van winst) te volgen minimaal de spelwaarde krijgen en speler II kan met zijn maximinstrategie zijn verlies altijd tot deze spelwaarde beperken. Met recht kunnen we dit strategieprofiel daarom karakteriseren als *optimaal spel van spelers tegen elkaar*. Blijkbaar is dit spel gunstig voor speler II en ongunstig voor speler I.

Karakteristiek bij deze vorm van optimaal spel is dat de bijbehorende strategieën gemengd zijn. We bekijken de vorm hiervan nader.

In knoop 1 moet speler I zelfs met de laagste kaart T niet altijd passen: gemiddeld in 1 van de 16 gevallen moet hij bieden, terwijl zeker is dat hij bij de showdown altijd zal verliezen. Hij kan alleen hopen dat speler II zal weggaan. Dit is dus *bluffen*. Het is blijkbaar een onderdeel van een wiskundig begrip *optimaliteit*: bluffen met wiskunde, en staat geheel los van bluffen met psychologie!

Het is niet verrassend dat in knoop 1 speler I met de hoge kaart K altijd moet bieden, omdat hij een grote kans heeft een eventuele showdown te winnen. Des te verrassender is het dat bij de allerhoogste kaart A hij gemiddeld in 3 van de 8 gevallen moet passen. Dit kan alleen als speler II ertoe gebracht wordt te bieden en op deze manier de pot zit te spekken. Dit is dus *verleiden*. Ook dit is dus een onderdeel van optimaliteit.

De in het voorgaande besproken begrippen optimaliteit vormen eigenlijk twee uitersten: optimaal spel bij c) met een bekende strategie van de tegenstander en bij d) met een volledig onbekende strategie van de tegenstander. In de praktijk heeft een speler meestal wel enige informatie over de strategie van zijn tegenstander of bouwt hij die in de loop van meerdere spelronden op, zodat de klasse van diens mogelijke strategieën kleiner is. Hierdoor kan hij zijn maximinwaarde vergroten en zal de bijbehorende maximinstrategie zich wijzigen. Deze zal in het algemeen nog steeds gerandomiseerde beslissingen bevatten, en die zijn vaak weer te interpreteren in termen van bluffen en verleiden.

Het feit dat er sprake is van een gemeenschappelijke waarde, geldt niet alleen voor *Straightpoker* met deze specifieke spelregels maar zeer algemeen voor alle tweepersoons-nulsomspelen.

Driepersoonsstraightpoker

De spelboom voor de generalisatie van twee- naar driepersoonsstraightpoker wordt gegeven door ***figuur 2***.
De spelboom is al behoorlijk veel groter dan voor tweepersoonsstraightpoker. Beginnersstrategieën zijn weer te baseren op kansen om met een bepaalde kaart bij een eventuele showdown te winnen. Een optimale strategie van een speler bij gegeven beginnersstrategieën van beide tegenstanders zijn ook weer te berekenen en deze kan steeds zuiver gekozen worden. Randomisatie, en in het bijzonder bluffen en verleiden, is dus volstrekt overbodig. We laten de resultaten hiervan onvermeld.
Voor het geval dat een speler de strategieën van de tegenstanders helemaal niet kent, is nu niet meer op de manier van de paragraaf *Tweepersoonsstraightpoker* goed te definiëren wat optimaal is. Natuurlijk bestaat er een (gemengde) maximinstrategie van een speler tegen de *coalitie* van zijn beide tegenstanders. Maar het idee dat deze zullen samenwerken om hem een zo'n klein mogelijke winst te gunnen, is wel wat pessimistisch en ook niet realistisch: ze willen elk voor zich hun eigen winst maximaliseren. Bovendien is het maken van onderlinge afspraken over te volgen strategieën verboden.
Maar hoe dan wel? Er bestaat in ieder geval nu niet meer een strategieprofiel dat gekarakteriseerd kan worden als optimaal tegen elkaar. Iemand die beweert dat hij optimaal speelt, zonder dat hij de strategieën van zijn tegenstander kent, verkoopt eigenlijk onzin. Toch bestaat er in dit geval wel een strategieprofiel dat doorgaans als *verstandig tegen elkaar* gekarakteriseerd mag worden: een zgn. *Nash-evenwicht*. Dit heeft de eigenschap dat voor *iedere* speler geldt: zijn strategie van het profiel is optimaal, zolang alle tegenstanders ook hun strategieën van het profiel spelen. Voor iedere speler geldt dus dat het geen zin heeft af te wijken van zijn strategie als de anderen dat ook niet doen. Zulke strategieën zijn meestal ook weer gemengd.
Tabel 4 geeft een numerieke benadering van zo'n evenwicht (alleen de kansen op de linkertak zijn vermeld, die van de rechtertak volgen hieruit):

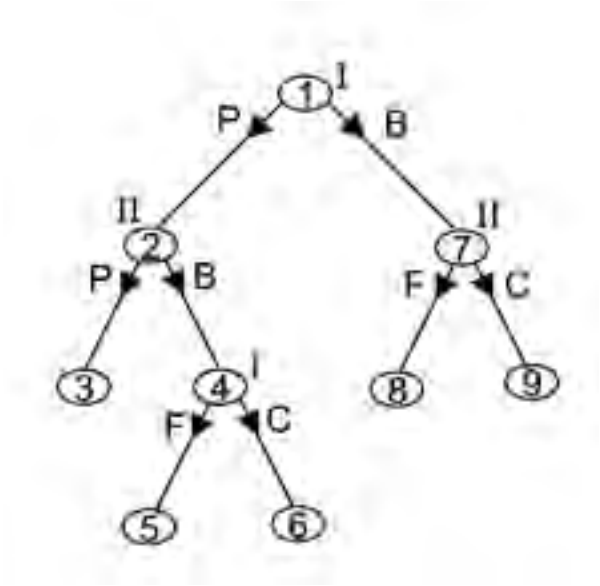
Speler I			Speler II			Speler III		
Kaart	(1) P	(5) F	Kaart	(2) P	(19) F	Kaart	(3) P	(12) F
T	0,96	1	T	0,84	1	T	0,87	1
B	0,96	1	B	0,84	1	B	0,87	1
V	0,96	1	V	0,84	1	V	0,87	1
H	0	0,11	H	0,40	0,90	H	1	0,98
A	0,74	0	A	0	0	A	1	0
Kaart	(13) F	(16) F	Kaart	(6) F	(9) F	Kaart	(20) F	(23) F
T	1	1	T	1	1	T	1	1
B	1	1	B	1	1	B	1	1
V	1	1	V	1	1	V	1	1
H	0,22	0,78	H	0,98	1	H	0,66	0
A	0	0	A	0	0	A	0	0
Winst van I: -0,15			Winst van I: -0,02			Winst van I: +0,17		

tabel 4 Nash-evenwicht bij driepersoonsstraightpoker

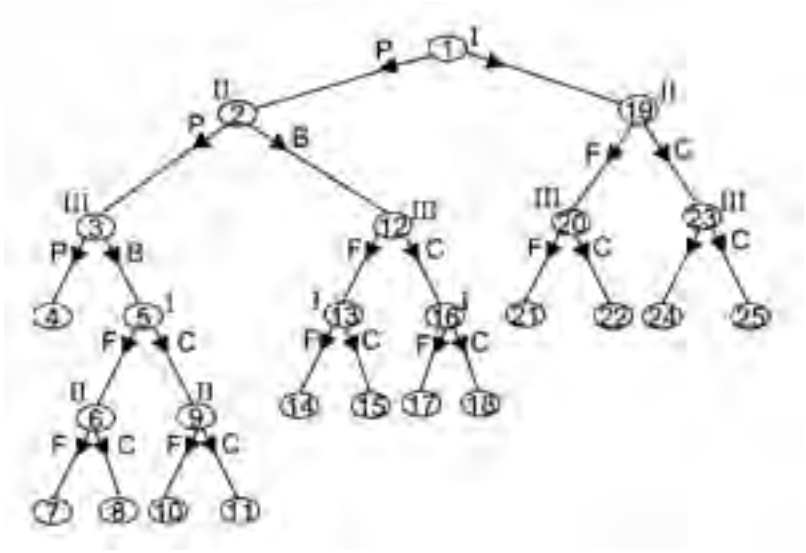
We zien opnieuw dat spelers bluffen en verleiden. Dit vormt dus weer onderdeel van verstandige spelstrategie.
Ieder spel heeft altijd een Nash-evenwicht. Voorts is optimaliteit bij tweepersoons-nulsomspelen, zoals beschreven in de paragraaf *Tweepersoonsstraightpoker* hiervan een speciaal geval. Toch lijkt het idee ervan mooier dan het is. In het algemeen kunnen er vele Nash-evenwichten zijn met allemaal verschillende spelerswinsten. Bovendien zegt het niets als spelers niet individueel maar tegelijk van hun strategieën afwijken. Het is daarom onjuist een evenwicht optimaal te noemen; hooguit zijn er vaak verdere argumenten om het verstandig spel te noemen. Daarom blijft het interessant te zien dat ook deze manier van verstandig spel leidt tot gerandomiseerde beslissingen, waaronder bluffen en verleiden.
Hoe zal dit spel nu concreet in de praktijk gespeeld worden? In dit spel is na het delen van de kaarten de berekening van de kans op de hoogste kaart in een eventuele showdown zeer eenvoudig. Veel lastiger is het inschatten van de strategieën van de tegenstanders. Ongetwijfeld is het verstandig op grond van bovengenoemde overwegingen ook gerandomiseerde beslissingen te gebruiken, maar de precieze vorm zal meestal niet erg hard te maken zijn.

Andere vormen van poker

De spelboom voor drie personen is nog juist analyseerbaar. Bij nog meer personen wordt de boom al snel onhanteerbaar. Dit treedt nog sneller op als overbieden (*raisen*) wordt toegestaan. Als gespeeld wordt met een volledig kaartspel en echte pokerhanden, dan wordt ook de kansberekening gecompliceerd. En dit gaat dan allemaal alleen nog maar over *Straightpoker*.
De complexiteit bij *Texas Hold'em* is nog vele malen groter omdat in verschillende fasen nieuwe kaarten getrokken worden, gevolgd door nieuwe biedronden. Maar ook bij deze pokervorm kan bluffen en verleiden gezien worden als onderdelen van verstandige gerandomiseerde beslissingen, nodig om ingedekt te zijn voor onbekende spelstrategieën van de tegenstanders. Er zijn de laatste jaren computerprogramma's ontwikkeld die spelers adviezen geven over de te nemen beslissingen gedurende de loop van het spel. Deze starten altijd met bepaalde strategiekeuzen van de tegenstanders en passen vervolgens deze keuzen aan in de loop van het spel op basis van genomen beslissingen. Vaak kunnen deze empirische gegevens opgeslagen worden zodat ze bij een volgende keer gebruikt kunnen worden. De algoritmen van zulke programma's verzorgen niet alleen de kansberekening (meestal door simulatie) maar ook de randomisatie. Aan deze programma's valt nog veel te verbeteren, met name met betrekking tot de randomisatie. Hier ligt de uitdaging voor de toekomst. Want zoals dit artikel laat zien, verstandig spel berust niet alleen op het goed kunnen inschatten van kansen maar ook op een goede manier van randomiseren, waaronder bluffen en verleiden.



figuur 1 Spelboom tweepersoonsstraightpoker



figuur 2 Spelboom driepersoonsstraightpoker

Noot
Met dank aan Ruud Hendrickx en de referenten voor hun commentaar.

Over de auteur
Prof.dr. Ben van der Genugten is als hoogleraar Waarschijnlijkheidsrekening en Mathematische Statistiek verbonden aan de Universiteit van Tilburg bij het departement Econometrie & OR. Naast fundamentele research heeft hij veel contractonderzoek verricht met betrekking tot de toepassing van de Wet op de Kansspelen. Doel hierbij was meestal inzicht te geven in de behendigheid van een spel ten opzichte van toeval. Bij vele processen was hij als getuige-deskundige betrokken, waaronder de pokerprocessen in de periode 1996-1998.
E-mailadres: *Ben.vdGenugten@uvt.nl*

De nieuwe grafische calculator FX-9860G SD: Even divers als Wiskunde zelf.

- Groot display met natuurlijke weergave van wiskundige formules
- Geheugen: 64 kB RAM en 1.5 MB flash ROM
- Meer dan 1000 technische en wetenschappelijke functies
- Spreadsheet functie
- Geometrische programmas (add-in)
- eActivity jr.
- PC emulatie software
- USB interface voor data overdracht

Ook beschikbaar zonder SD geheugenkaart slot als FX-9860G.

Met SD geheugenkaart slot*

* Compatible SD kaarten: TOSHIBA: SD-NA032MT, SD-NA064MT, SD-NA128MT, SD-NA256MT, SD-NA512MT, SD-FA128MT, SD-FA256MT
SanDisk: SDSDB-64-J60, SDSDB-128-J60, SDSDB-256-J60, SDSDB-512-J60, SDSDH-256-903, SDSDH-512-903

CASIO Benelux B.V. • Prof.W.H. Keesomlaan 6 • 1183 DJ Amstelveen • Tel: 020-545 10 70 • Fax: 020-545 10 89

CASIO
www.casio-europe.com

Grote griepmeting weer van start

Persbericht van 31 oktober 2007
Voor de vijfde keer in de geschiedenis wordt een grootschalige online griepmeting gedaan in Nederland en België. Onderzoekers uit beide landen kunnen zo volgen hoe de griep zich deze winter verspreidt. Op de website van de GroteGriepMeting *(www.degrotegriepmeting.nl)* kunnen deelnemers zelf hun griepverschijnselen bijhouden.

Wekelijks kunnen vrijwilligers op de website aangeven hoe grieperig ze zijn. Aan de hand van de aan- of afwezigheid van griepachtige symptomen wordt bepaald of een deelnemer ofwel verkouden of een vermoedelijke griep heeft. Alle metingen komen in een kaart van Nederland en België, die laat zien waar de griep zich manifesteert. Blauwe stippen geven de gevallen van verkoudheid aan en rode stippen de vermoedelijke griepgevallen. De meting is anoniem en de resultaten zijn vrij beschikbaar voor iedereen. De vijfde Grote Griepmeting wordt gehouden van 1 november 2007 tot 1 mei 2008.

Betrouwbaar

Dit is het vijfde jaar dat de griepmeting in Nederland en België plaatsvindt. In Portugal gaat voor de derde keer een identieke griepmeting via internet van start. De afgelopen jaren brachten bijna veertigduizend 'griepmeters' in de drie landen de griepgolf in kaart. De bedenker van de griepmeting, Carl Koppeschaar, begon in 2003 met de metingen als educatief webproject, met als doel mensen te betrekken bij wetenschappelijk onderzoek en te informeren over griep. 'Maar inmiddels verschenen al twee wetenschappelijke publicaties waaruit bleek dat het project zeer betrouwbare resultaten oplevert', aldus Koppeschaar en zijn team van wiskundigen, virologen en epidemiologen. 'Op grond daarvan vinden deze winter ook proefmetingen plaats in Zweden en in Italië. Groot-Brittannië, Duitsland en Slowakije willen volgend jaar gaan meedoen. Verder is er al belangstelling vanuit Oostenrijk, Frankrijk, Finland, Ierland, Polen, Malta, Spanje, Estland en Luxemburg. Op den duur kan dus één grote, Europese griepmeting door vrijwilligers ontstaan.'

Griepprik

Bijzondere aandacht bij de vijfde griepmeting gaat uit naar de effectiviteit van de griepprik en de bereidheid om de griepprik te nemen. Met enkele extra vragen proberen de onderzoekers inzicht te krijgen of er een verschil in bescherming is tussen mensen die de griepprik nemen op aanraden van een arts, en mensen die op eigen initiatief de griepprik halen.

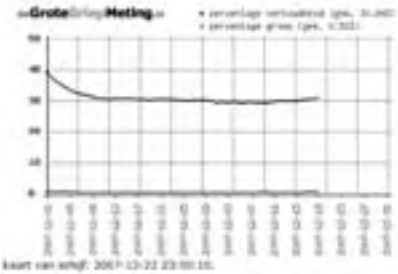
Pandemie

Echte griep, of influenza, is een infectieziekte die met name wordt gevreesd vanwege de voortdurende verandering van het virus en daarmee de dreiging van een wereldwijde pandemie. Berucht is de Spaanse Griep, die in 1918-1919 naar schatting 40 miljoen slachtoffers eiste. De afgelopen jaren maken virologen zich ernstig zorgen over de verbreiding van de vogelgriep, omdat dat virus verwant is aan het menselijk virus en zou kunnen veranderen in een variant die van mens op mens overdraagbaar is. Grieponderzoeker Sander van Noort en zijn collega's denken dat de Grote Griepmeting een belangrijk systeem is om de griepactiviteit tijdens een eventuele pandemie te kunnen blijven meten. Met name wanneer het rapporteringsysteem via de dan overbelaste huisartsen niet meer goed zou werken. 'Uit de voorgaande jaren weten we dat de griepmeting via vrijwilligers gelijke start-, piek- en einddata van de griepactiviteit meet als het rapporteringsysteem van de huisartsen', zegt Van Noort. 'Een voordeel is dat de online griepmeting de resultaten 3-4 dagen eerder heeft. Maar het belangrijkst is dat de meting een uniek platform biedt om de griepactiviteit tussen verschillende landen te vergelijken. Naarmate meer landen zich bij de meting aansluiten, kunnen we de verspreiding van griep van land tot land door Europa volgen en ook eerder waarschuwen.'

Informatie

Meer informatie waaronder actuele meetresultaten (van meer dan 23.500 Nederlandse en Belgische vrijwillige 'griepmeters') is te vinden op de website van de GroteGriepMeting (*www.degrotegriepmeting.nl*).

figuur 1 Op 23 december 2007 (om 13:00u): 457 gevallen van verkoudheid



figuur 2 Percentages griep en verkoudheid uitgezet tegen de tijd (tot 22 december 2007, 22:550u)

NVvW, NVORWO, Het Ei van Columbus en het Vierde Bartjens Rekendictee (2007)

[Jaap Vedder en Marian Kollenveld]

De leden van de NVvW treffen een boekje aan bij dit nummer van Euclides. Het heet: *Volgens Bartjens..., Het Ei van Columbus: rekenpuzzels & breinkrakers.* Onder deze titel brengt de Nederlandse Vereniging tot Ontwikkeling van het Reken/Wiskunde Onderwijs (NVORWO) ter gelegenheid van haar 25-jarig bestaan een boekje uit met uitdagende inspirerende rekenopgaven voor leerlingen en leraren. De NVvW en de NVORWO geven dit boekje aan al hun leden cadeau!

Het Ei als lustrumidee

Sinds jaren verschijnt in het vaktijdschrift van NVORWO *Volgens Bartjens…* de rubriek 'Het Ei van Columbus'. Ingegeven door de populariteit van deze rubriek bij leerlingen en leraren in de bovenbouw van de basisschool én het succes van een eerdere uitgave 'Rekenpuzzels en Breinkrakers' heeft de NVORWO de beste eieren van de afgelopen jaren verzameld in een lustrumuitgave. Deze is bedoeld voor leerlingen van 10 tot 14 jaar. Ook voor (aanstaande) leraren zijn de opgaven vast en zeker een bron van inspiratie, enerzijds om zelf mee aan de slag te gaan, anderzijds om te benutten in de klas.

De voortdurende discussie over het rekenniveau van de basisschool en over de doorlopende leerlijn van basisschool tot en met hoger onderwijs (pabo) is voor de NVORWO mede aanleiding dit boekje uit te geven. Het boekje bevat rekenpuzzels en breinkrakers die kunnen bijdragen aan het uitbreiden van de eigen rekenvaardigheid en vooral aan het leren oplossen van problemen, wat de wiskundige attitude vergroot. De opgaven boeien dan ook jong en oud.

De NVvW en de NVORWO

De NVORWO en de NVvW gaan hun contacten intensiveren. De discussie rond de doorlopende leerlijn rekenen en taal noopt daartoe. Het grote probleem bijvoorbeeld van de instroom in de pabo die negatief in het nieuws is, heeft te maken met de opbouw van ons stelsel. Steeds meer zullen alle leerlingen 'rekenen' bij moeten houden, ook als ze geen wiskunde in het pakket hebben. Met een Ei van Columbus voor alle leden van beide verenigingen onderstrepen we dat er voor onze verenigingen een fors gemeenschappelijk deel in onze agenda's zit. Daar gaan we aan werken!

Het Vierde Bartjens Rekendictee

Voor de vierde maal werd in november het Bartjens Rekendictee in Zwolle gehouden. Meester Willem Bartjens, geboren in Amsterdam, maar verhuisd naar Zwolle om daar les te geven, heeft een groot stempel gedrukt op het Nederlandse rekenonderwijs. Wie kent niet de uitdrukking: *Volgens Bartjens*?

Het nu Landelijke Rekendictee bestond uit een voorronde onder scholieren (3e klas voortgezet onderwijs) en een voorronde onder pabo-studenten. Er werden 32 scholieren afgevaardigd en negen pabo's deden mee met een team van hun vier beste rekenaars. Verder waren er vier voorrondes via internet: de Stentor, het Limburgs Dagblad, de Provinciale Zeeuwse Courant (PZC) en via Kennisnet. Totaal namen aan deze internetrondes 1234 personen deel. (Hoe toevallig is 1-2-3-4?)

De finale vond plaats in Zwolle op 23 november j.l. Totaal 135 finalisten: 120 kwamen uit de voorrondes aangevuld met 15 prominenten, onder wie twee wethouders van Zwolle en twee hoofdinspecteurs van het onderwijs.

Na een spannende barrage werd wiskundedocent Epi van Winsen uit Margraten (Limburg) uitgeroepen tot de allerbeste rekenaar van 2007. Hij was in de finale geplaatst via de voorronde van Kennisnet. Er waren schitterende prijzen voor de beste pabo-student, de beste scholier, de beste internet-rekenaar en de beste prominent.

De finale werd gepresenteerd door Marjolein Kool, hoofdredacteur van *Volgens Bartjens*…. Zij stelde ook de opgaven samen.
Het was een spannende finale met veel publiek. Volgend jaar doen waarschijnlijk nog meer regionale dagbladen mee. Alle finalisten kregen ook een exemplaar van *Het Ei* mee om vast te oefenen voor volgend jaar. Op *www.bartjens.net* staat meer over het Rekendictee.

Tot slot
Rekenen/wiskunde en problemen oplossen zijn uitdagend voor veel leerlingen tussen 10 en 14 jaar en ouder, en inspirerend voor (aanstaande) leraren. De antwoorden van de opgaven zijn te vinden op de websites van de NVORWO (*www.nvorwo.nl* en *www.volgens-bartjens.nl*). Daar zijn ook alle opgavenbladen te downloaden om te printen en in de klas te gebruiken. *Het Ei van Columbus* is te koop voor € 4,95 in de boekhandel.

Over de auteurs
Jaap Vedder is voorzitter van de NVORWO en voorzitter van de Stichting Bartjens Rekendictee.
E-mailadres: *voorzitter@nvorwo.nl*
Marian Kollenveld is voorzitter van de NVvW.
E-mailadres: *m.kollenveld@nvvw.nl*

Inloggen op de site van de NVvW

[Metha Kamminga]

Het is mogelijk om op de site van de NVvW in te loggen (rechts bovenaan) met uw e-mailadres als login en te kiezen voor 'ingelogd blijven' als u vaak op dezelfde computer werkt.
In juni 2007 is daarover een mailing geweest. Niet iedereen is echter op de hoogte van deze mogelijkheid om meer informatie te krijgen op de site en meer mogelijkheden te hebben om mee te discussiëren. De mensen die niet kunnen inloggen of hun wachtwoord zijn vergeten, kunnen een mail sturen naar < *webmasters@nvvw.nl* > (denk om het meervoud in het e-mailadres tussen < en >!).
Eenmaal ingelogd is het mogelijk uw profiel en wachtwoord aan te passen.

Vanaf nu is het overigens ook mogelijk om lid te worden via een digitaal inschrijfformulier onder Vereniging | Lidmaatschap | Aanmeldingsformulier.

Aankondiging / Het voortbestaan van de wiskunde in het HBO

Conferentie georganiseerd door de werkgroep-HBO van de NVvW

[Metha Kamminga]

Als je leerlingen toegepaste wiskunde leert, leer je ze wiskunde die niet toepasbaar is.' (Een stelling van professor Hans Freudenthal.)

De conferentie wordt gehouden op **woensdagmiddag 12 maart 2008** van 13:00 tot 17:00 uur en is bedoeld voor iedereen die te maken heeft met wiskunde in het HBO. Dus niet alleen wiskundedocenten!

De conferentie is gericht op toekomstig beleid over aansluiting, samenwerking met VO en MBO en de manier waarop wiskunde in het HBO wordt ingezet. Ook het vak wiskunde D in het HAVO krijgt aandacht. Er zijn twee rondes, beide plenair, met korte inleidingen van beleidsmakers van onder andere lerarenopleidingen, hogescholen en MBO-instellingen, het Freudenthal Instituut, het bestuur van BON (Beter Onderwijs Nederland), en, niet in de laatste plaats, ook het bestuur van de HBO-raad laat zich vertegenwoordigen.
Beide rondes worden afgesloten met een forumdiscussie. In het forum hebben zitting de inleiders aangevuld met mensen van het bedrijfsleven die een duidelijke mening hebben over wiskunde in het HBO.
Als afsluiting wordt ter plekke door alle deelnemers over een aantal vooraf opgestelde petities (richting OCW en HBO-raad) gestemd per GSM.
Op de site van de NVvW, *www.nvvw.nl*, bij Vereniging | Werkgroepen | Werkgroep-HBO, is informatie te vinden met de mogelijkheid tot digitaal aanmelden. Neem op 12 maart 2008 uw collega van de toegepaste vakken mee en iedereen die zich maar interesseert voor de wiskunde in het HBO.

Puzzel 834

Stochastische wandelingen

[Frits Göbel]

Een stoffelijk punt wandelt over de x-as. Het punt start in de oorsprong en het maakt stappen van de grootte 1, met kans ½ naar rechts en kans ½ naar links, steeds onafhankelijk van de voorgeschiedenis. We beginnen eenvoudig.

Opgave 1
Bepaal de kans dat het punt na $2n$ stappen weer in 0 is.

In de volgende opgave hebben we te maken met een aapje, dat ook in 0 begint en ook met kans ½ naar links of naar rechts springt, maar bij hem – het is namelijk een mannetje – wordt de 'staplengte' steeds 1 groter, te beginnen met lengte 1.

Opgave 2
Bepaal de kans dat het aapje na 8 stappen weer in 0 is.

In opgave 3 zien we een wat oudere aap. Zijn sprongen hebben de lengten 2 en 3, en gaan naar links of rechts. De vier verschillende mogelijkheden hebben alle de kans ¼.

Opgave 3
Bepaal de kans dat deze aap na 7 stappen weer in 0 is.

Voor de laatste opgave gaan we een dimensie hoger. Het punt beweegt nu op het vierkante rooster met (0, 0) als startpunt. Het gaat steeds met kans ¼ in een van de vier mogelijke richtingen.

Opgave 4
Wat is de kansverdeling van de afstand tot (0, 0) na 6 stappen?
Als afstand nemen we de Manhattan-afstand: de afstand van (0, 0) tot (x, y) is dan $|x| + |y|$.

Oplossingen kunt u mailen naar *a.gobel@wxs.nl* of per gewone post sturen naar F. Göbel, Schubertlaan 28, 7522 JS Enschede.
Er zijn weer maximaal 20 punten te verdienen met uw oplossing.
De deadline is 3 maart 2008.
Veel plezier!

Oplossing 832

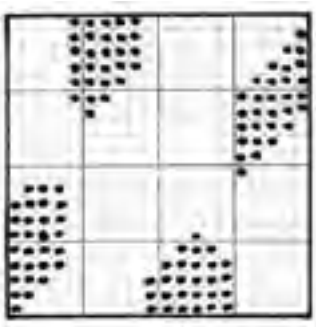
Witte en zwarte dames

Er waren 13 inzendingen, waarvan er 11 goed zijn voor 20 punten. Deze werden ingestuurd door Wim van den Camp, Jozef Hanenberg, Hans Linders (terug van weggeweest!), Kees Verhoeven, Floor van Lamoen, Ton Kool, Lieke de Rooij, Wobien Doyer, Gerhard Riphagen, Niels Wensink en Hans Klein.

Opgave 1 is door iedereen goed opgelost. Twee inzenders, Wobien Doyer en Gerhard Riphagen, bewezen dat de oplossing uniek is.
Ook *opgave 2* werd door alle inzenders tot een goed einde gebracht. Menigeen vond zelfs een veilige stelling op het bord van 7 × 7 met zeven Dames van iedere kleur, waarmee dan meteen opgave 3 is opgelost. Behalve de paren (7, 6) en (7, 7) werden ook nog de volgende paren gevonden: (8, 6) door Jozef Hanenberg en Hans Klein, (9, 6) door Wim van den Camp en Floor van Lamoen, en (10, 6) door Wobien Doyer.
Naar aanleiding van de formule voor $s(n)$ schrijft Jozef Hanenberg:
'Je zou meer punten moeten krijgen voor een stand met 6 witte en 8 zwarte Dames dan voor een stand met 6 witte en 7 zwarte Dames. Dat zou je voor elkaar kunnen krijgen door de volgende formule te gebruiken:

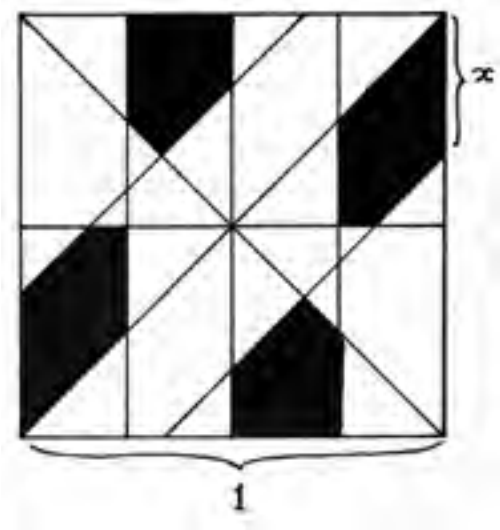
$$s(n) = \frac{2 \cdot \max(\sqrt{w \cdot z})}{n^2}$$

waarbij het maximum genomen wordt over alle veilige standen.'
Een goed idee!



figuur 1

Zoals gezegd, (7, 7) levert al een oplossing voor *opgave 3*, maar bijna niemand heeft het hierbij gelaten! Wim van den Camp stuurde (58, 56) op 20 × 20 met een score 0,28; ***zie figuur 1***. Deze hoge waarde en de figuur leidden al snel tot het idee dat mijn limiet van $2 - \sqrt{3}$ wel eens fout zou kunnen zijn. En inderdaad: dit getal kan worden verhoogd tot $\frac{7}{24}$.
Zie figuur 2; voor $x = \frac{1}{3}$ is de bezette oppervlakte maximaal. Ook Jozef Hanenberg en Kees Verhoeven vonden de waarde $\frac{7}{24}$. Een duidelijke verbetering. Ik waag me maar niet aan een nieuw vermoeden!
Ik ontving ook veel andere veilige stellingen voor $n > 7$ met een hoge s-waarde: (9, 10) voor $n = 8$ van Gerhard Riphagen. Hij en Floor van Lamoen vonden (12, 12) voor $n = 9$; (14, 15) voor $n = 10$ van Jozef Hanenberg en Kees Verhoeven, en deze laatste verbeterde bovengenoemde (58, 56) nog tot (59, 56). Floor van Lamoen tenslotte vond (576, 576) voor $n = 65$.
Al met al heel wat resultaten die beter of zelfs veel beter waren dan wat ik had!
Twee inzenders kwamen niet toe aan opgave 3 doordat zij de definitie van $s(n)$ niet doorgrondden.



figuur 2

Ladderstand
De top van de ladder ziet nu als volgt uit:
H.J. Brascamp 472
J. Meerhof 395
L. de Rooij 353
G. Riphagen 309
L.H. van den Raadt 231
H. Klein 203
N. Wensink 198
W. Doyer 196
T. Kool 119

SERVICEPAGINA

Voor overige internet-adressen zie
www.wiskundepersdienst.nl/agenda.php

Voor Wiskundeonderwijs Webwijzer zie
www.wiskundeonderwijs.nl

Publicaties van de Nederlande Vereniging van Wiskundeleraren

Zebraboekjes

1. Kattenaids en Statistiek
2. Perspectief, hoe moet je dat zien?
3. Schatten, hoe doe je dat?
4. De Gulden Snede
5. Poisson, de Pruisen en de Lotto
6. Pi
7. De laatste stelling van Fermat
8. Verkiezingen, een web van paradoxen
9. De Veelzijdigheid van Bollen
10. Fractals
11. Schuiven met auto's, munten en bollen
12. Spelen met gehelen
13. Wiskunde in de Islam
14. Grafen in de praktijk
15. De juiste toon
16. Chaos en orde
17. Christiaan Huygens
18. Zeepvliezen
19. Nullen en Enen
20. Babylonische Wiskunde
21. Geschiedenis van de niet-Euclidische meetkunde
22. Spelen en Delen
23. Experimenteren met kansen
24. Gravitatie
25. Blik op Oneindig
26. Een Koele Blik op Waarheid

Zie verder ook *www.nvvw.nl/zebrareeks.html* en/of *www.epsilon-uitgaven.nl*

Nomenclatuurrapport Tweede fase havo/vwo
Dit rapport en oude nummers van Euclides (voor zover voorradig) kunnen besteld worden bij de ledenadministratie (zie Colofon).

Wisforta – wiskunde, formules en tabellen
Formule- en tabellenboekje met formulekaarten havo en vwo, de tabellen van de binomiale en de normale verdeling, en toevalsgetallen.

Honderd jaar wiskundeonderwijs, lustrumboek van de NVvW
Het boek is met een bestelformulier te bestellen op de website van de NVvW: *www.nvvw.nl/lustrumboek2.html*
Voor overige NVvW-publicaties zie de website: *www.nvvw.nl/Publicaties2.html*

Kalender

In de kalender kunnen alle voor wiskundedocenten toegankelijke en interessante bijeenkomsten worden opgenomen. Relevante data graag zo vroeg mogelijk doorgeven aan de hoofdredacteur, het liefst via e-mail (*redactie-euclides@nvvw.nl*). Hieronder vindt u de verschijningsdata van Euclides in de lopende jaargang. Achter de verschijningsdatum is de deadline vermeld voor het inzenden van mededelingen en van de *eind*versies van geaccepteerde bijdragen; zie daarvoor echter ook *www.nvvw.nl/euclricht.html*.

nr.	verschijnt	deadline
5	11 maart	22 januari
6	22 april	4 maart
7	3 juni	8 april
8	1 juli	15 mei

dinsdag 19 februari, Amsterdam
Mastercourse: Golven als dynamische systemen
Organisatie UvA

donderdag 6 maart, Amsterdam
Mastercourse: Computerarchitectuur
Organisatie UvA

do. 13 en vr. 14 maart, Garderen
Finale Wiskunde A-lympiade
Organisatie FIsme

dinsdag 18 maart, Groningen
17e Johann Bernoulli-lezing
Organisatie RU Groningen

do. 27 en vr. 28 maart, Noordwijkerhout
Nationale Rekendagen
Organisatie FIsme

woensdag 9 april, Amsterdam
Mastercourse: Laat de Spelen beginnen! Olympische wiskunde
Organisatie UvA

vrijdag 11 april, op de scholen
Wiskunde Kangoeroe
Organisatie Stichting Wiskunde Kangoeroe

woensdag 16 april, op de scholen
De Grote Rekendag
Organisatie FIsme

zaterdag 17 mei, Utrecht
HKRWO-Symposium XIV
Organisatie HKRWO

za. 13 t/m vr. 18 juli, RAI, Amsterdam
Fifth European Congress of Mathematics
Organisatie VU, CWI en UvA

do. 24 t/m di. 29 juli, Leeuwarden
Bridges Leeuwarden (11th Bridges Conference): Mathematics, Music, Art, Architecture, Culture
Organisatie The Bridges Organisation (www.bridgesmathart.org)

MATRIX

WAT BIJ UW LEERLINGEN LEEFT...

...DAAR MAAKT U ECHTE WISKUNDE VAN. Wiskunde is overal. Maar hoe maakt u het zichtbaar voor al uw leerlingen? Hoe maakt u ze nieuwsgierig? Met de wiskundemethode Matrix van Malmberg kiest u voor een unieke aanpak. Het uitgangspunt: relevante wiskunde. Matrix leert de leerlingen om het vak te snappen. Aan de hand van herkenbare situaties raakt elke leerling verwonderd. Met als gevolg: extra motivatie. Leerlingen worden uitgedaagd om zelf aan de slag te gaan en denken meer en dieper na. Bovendien heeft u als docent alle ruimte om op uw eigen manier 'de klik' met de leerlingen te maken. Zo kunt u flexibel inspelen op uw eigen wensen én de verschillende leerniveaus van uw leerlingen. En u kunt moeiteloos overschakelen van het boek op ePack en weer terug. Kortom, u maakt échte wiskunde van wat bij uw leerlingen leeft.
Meer weten? Vraag nu een beoordelingsexemplaar van Matrix aan. Bel 073 628 8766. **VOOR JE HET WEET, HEB JE HET DOOR. MATRIX**

MALMBERG

Nieuw: Netwerk 4e editie

Compleet voor vmbo, havo en vwo onderbouw en Tweede Fase

4e editie: compleet voor vmbo, havo en vwo onderbouw en Tweede Fase

- Compacte leerlijnen
- Pragmatisch
- Voor vmbo: wiskunde met je handen
- Voor havo/vwo: puzzelen en denken
- Aandacht voor rekenvaardigheden
- Ook volledig digitaal beschikbaar

Meer informatie op www.netwerk.wolters.nl

Wolters-Noordhoff