STAtOR

Strategie bij het bordspel Mens-erger-je-niet

The beautiful mathematics of the card game SET

Darts; Het voorspellen van de uitkomst van profwedstrijden

De sleutel tot succes; Het simuleren van Formule 1-racestrategieën

Mastermind voor BloodMatching

Hoe een statistische fout het ammoniakdebat op een dwaalspoor zet

Eerste Kamerverkiezing; De puzzelaar bepaalt de laatste zetels!

periodiek van de VVSOR jaargang 20, nummer 2, juni 2019

STAtOR

Jaargang **20**, nummer 2, juni 2019

STAtOR is een uitgave van de Vereniging voor Statistiek en Operations Research (VVSOR). *STAtOR* wil leden, bedrijven en overige geïnteresseerden op de hoogte houden van ontwikkelingen en nieuws over toepassingen van statistiek en operations research. Verschijnt 4 keer per jaar.

Redactie

Joaquim Gromicho (hoofdredacteur), Annelieke Baller, Ana Isabel Barros, Joep Burger, Kristiaan Glorie, Caroline Jagtenberg, Guus Luijben (eindredacteur), Richard Starmans, Gerrit Stemerdink (eindredacteur), Vanessa Torres van Grinsven en Sanne Willems. Vaste medewerkers: Johan van Leeuwaarden, John Poppelaars, Gerard Sierksma en Henk Tijms.

Kopij en reacties richten aan

Prof. dr. J.A.S. Gromicho (hoofdredacteur), Faculteit der Economische Wetenschappen en Bedrijfskunde, afdeling Econometrie, Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam, mobiel 06 55886747, j.a.dossantos.gromicho@vu.nl

Bestuur van de VVSOR

Voorzitter: prof. dr. Fred van Eeuwijk, db@vvsor.nl; Secretaris: dr. Sophie Swinkels, db@vvsor.nl; Penningmeester: dr. Ad Ridder, db@vvsor.nl; Algemeen: Nikky van Buuren MSc, webmaster@vvsor.nl
Voorzitters van de secties: prof. dr. ir. Mark van de Wiel (Biometrical Section); prof. dr. Albert Wagelmans (Section for Operations Research); dr. Eduard Belitser (Section Mathematical Statistics); prof. dr. Casper Albers (Social Sciences Section); dr. Michel van de Velden (Economics Section); Jonas Haslbeck MSc (Young Statisticians) Sanne Willems MSc (Section Statistics Communication).

Leden- en abonnementenadministratie van de VVSOR

VVSOR, Postbus 1058, 3860 BB Nijkerk, telefoon 033 2473408, admin@vvsor.nl. Raadpleeg onze website www.vvsor.nl over hoe u lid kunt worden van de VVSOR of een abonnement kunt nemen op *STAtOR*.

Voor advertenties

M. van Hootegem, hootegem@xs4all.nl
STAtOR verschijnt in maart, juni, oktober en december.

Ontwerp en opmaak

Pharos, Nijmegen

Uitgever

ISSN 1567-3383

EEN ZOMERZOTHEID

De boog kan niet altijd gespannen zijn, om maar eens een gemeenplaats te gebruiken. De zomer is begonnen met zijn lange avonden, dat vraagt om aangepaste activiteiten. En wat beter te doen dan in zo'n tijd dan genieten van sport en spel. De mens speelt nu eenmaal graag, het begrip Homo Ludens is niet voor niets ontstaan. Daarom vindt u in dit zomernummer een veelheid aan artikelen over de toepassingen van ons vakgebied op sport en spel. Had u ooit verwacht dat Mens-erger-je-niet wetenschappelijk geanalyseerd kan worden, net als darts, Formule 1-races en andere vormen van sport en spel? U leest het u in deze *STAtOR*!

Een geheel ander en zeer serieus spel is de puzzel die gevormd wordt door de verkiezing van de leden van de Eerste Kamer. Jacob Jan Paulus beschrijft, net als vier jaar geleden, de vele intriges die hier mee gepaard kunnen gaan. De oplossing zal bekend zijn als dit nummer verschijnt, maar dat maakt het artikel niet minder interessant.

Nóg meer spel: Joost van Sambeeck en zijn co-auteurs vergelijken het matchen van bloedtypes voor transfusies met het spelen van Mastermind.

Maar niet alles gaat over sport en spel. Paul Goedhart heeft het over het ammoniakdebat dat door een statistische fout op een dwaalspoor terecht is gekomen en Matt Briggs en Jaap Hanekamp geven daar commentaar op. De redactie van *STAtOR* hecht er aan te vermelden dat zij geen partij kiest in de soms zeer verhitte discussie die op veel andere plaatsen al over dit onderwerp is gevoerd. Wij hebben ervoor gekozen deze artikelen op te nemen om onze lezers zélf tot een oordeel te laten komen.

En natuurlijk is een nummer van dit blad niet compleet zonder bijdragen van onze columnisten, u zult zien dat bij hen ook hier en daar sport en spel een rol speelt.

Wij wensen u veel sportief speels leesplezier!

INHOUD

STRATEGIE BIJ HET BORDSPEL MENS-ERGER-JE-NIET

Bij het bordspel Mens-erger-je-niet kunnen verschillende spelstrategieën worden toepgepast. We formuleren de gebruikelijke speltactieken als drie klassieke strategieën. Ook tonen we aan dat deze strategieën sterk verbeterd kunnen worden door een nieuwe strategie die vrij eenvoudig kan worden toegepast. Dit geldt voor alle gebruikelijke varianten van bordformaat en het aantal spelers.

Ben van der Genugten & Marcel Das

De huidige literatuur over strategie bij het populaire bordspel Mens-erger-je-niet is vrijwel beperkt tot kwalitatieve suggesties over een verstandige manier van spelen (zie bijvoorbeeld Wikipedia, 2019). In dit artikel bespreken we de winstkansen bij eenvoudig toepasbare spelstrategieën. In onze *working paper* (Van der Genugten & Das, 2018) worden meer ingewikkelde spelstrategieën besproken.

Eerst geven we de precieze spelregels van de meest gebruikelijke versie van het spel. We formuleren drie eenvoudige spelstrategieën die de kwalitatieve suggesties voor een manier van spelen formaliseren. Vervolgens beschrijven we een nieuwe spelstrategie die ook vrij eenvoudig kan worden toegepast. Daarna analyseren we deze vier strategieën. We tonen aan dat deze nieuwe strategie de drie eenvoudige strategieën sterk domineert. Ten slotte geven we suggesties voor verder onderzoek.

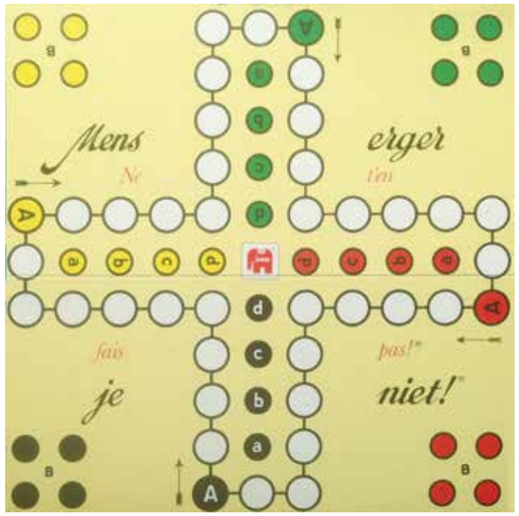
Spelregels

Het spel wordt gespeeld op de voorzijde (figuur 1) of op de achterzijde (figuur 2) van het bord afhankelijk van het aantal spelers M: voor M = 2 en 4 de voorzijde en voor M = 3 en 6 de achterzijde.

Iedere speler heeft zijn eigen kleur met P = 4 pionnen, een startgebied voor de pionnen (aan de randen van het bord) en thuisvelden voor de pionnen in het midden van het bord. Het circuit van witte velden moet door alle pionnen kloksgewijs worden afgelegd (zie de pijlen in figuren 1 en 2). De lengte C van het circuit is C = 40 aan de voorzijde en C = 48 aan de achterzijde. De startgebieden van de spelers zijn symmetrisch op het bord geplaatst.

Voor aanvang van het spel plaatsen de spelers hun pionnen in het startgebied (positie 0). Beurtelings verplaatsen zij hun pionnen afhankelijk van het aantal geworpen ogen met een dobbelsteen met D = 6 ogen, te beginnen met hun eigen startveld A (positie 1) en na het circuit volledig afgelegd te hebben (positie C) te eindigen op een eigen thuisveld (posities C+1, ..., C+P). De speler die het eerst al zijn pionnen op de thuisvelden heeft is winnaar. De volgorde van de spelers is rechtsom op het bord waarbij de speler met het startgebied linksonder het eerst aan de beurt is. (In de praktijk wordt meestal om de posities van de startgebieden geloot.)

De precieze omschrijving van de spelregels voor het verplaatsen van de pionnen afhankelijk van de bordsituatie is als volgt. Laat in een zekere bordsituatie de beurt aan een bepaalde speler zijn. Neem aan dat hij met de dobbelsteen d ogen gooit. Als d = D behoudt hij zijn beurt en als d < D dan gaat de beurt naar de volgende speler (na verplaatsing van de pion). Of en hoe hij met zijn worp een pion kan verplaatsen hangt af van de bordsituatie.



Figuur 1. Voorzijde van het bord

Figuur 2. Achterzijde van het bord

Als hij een pion op zijn startveld 1 heeft moet hij deze d velden verplaatsen tenzij op de eindpositie d+1 al een eigen pion staat; in dat geval gaat de pion op het startveld weer naar het startgebied (positie 0). Als het startveld leeg is en d < D dan mag hij naar keuze één van zijn pionnen op het circuit of op een thuisveld verplaatsen over d velden mits op de eindpositie geen eigen pion staat. Hierbij mag in de thuisvelden ook vanaf het verste thuisveld teruggeteld worden maar uiteindelijk moet de pion wel voorwaarts zijn verplaatst. (Een spelregelvariant is dat niet teruggeteld mag worden). Als het startveld leeg is en d = D dan geldt precies hetzelfde tenzij het startgebied niet leeg is. In dat geval moet een pion uit het startgebied op het startveld geplaatst worden en is dit de eindpositie.

Bij de verplaatsing kan de pion andere pionnen passeren. Als op de eindpositie van de pion een pion van een tegenstander staat, dan wordt deze pion verwijderd (geslagen) en teruggeplaatst in het startgebied van diens spelerskleur.

In zijn eerste beurt plaatst een speler een pion uit zijn startgebied op zijn startveld net alsof hij D ogen gegooid heeft. (Een spelvariant is dat geen uitzondering voor de eerste beurt gemaakt wordt en gewacht moet worden op een worp d = D.)

Het spel dankt zijn naam Mens-erger-je-niet aan de ergernis die het slaan van eigen pionnen door die van tegenstanders opwekt. Deze pionnen moeten immers weer helemaal vanaf het startgebied beginnen. Dit wordt enigszins gecompenseerd door de vreugde die het slaan van pionnen van de tegenstanders door eigen pionnen opwekt.

Spelstrategieën

Zodra een speler bij zijn worp twee of meer pionnen kan verzetten, moet hij een keuze maken welke pion hij verplaatst. Uiteraard wil hij die pion verplaatsen welke de kans op zijn winst zo groot mogelijk maakt. Maar in vrijwel alle bordsituaties is helemaal niet duidelijk welke dit is. Er is een aantal overwegingen te geven dat hierbij een rol speelt. Zo kan hij proberen

- de verst gevorderde pion zo snel mogelijk op een thuisveld te krijgen;
- indien mogelijk de pion van een tegenstander slaan;
- de tegenstander te verwarren door willekeurig een pion te kiezen;
- zijn pionnen zo veel mogelijk buiten het 'slaan bereik' van vijandelijke pionnen te houden.

In de laatste fase van het spel beschikt een speler die achterstaat (dus meer pionnen op de startvelden en op het circuit heeft) doorgaans over meer keuzemogelijkheden dan een speler die voorstaat omdat hij meer pionnen in de thuisvelden heeft.

We formuleren eerst drie eenvoudige, veel gebruikte strategieën die bovengenoemde aspecten in enigerlei vorm concretiseren.

F(ar)strat

De verst gevorderde pion wordt verzet. Daardoor wordt getracht een pion zo snel mogelijk in een thuisveld te krijgen. Er worden geen beurten verspild aan het zetten van andere pionnen om die van tegenstanders te slaan of om eigen pionnen te beveiligen voor het slaan door tegenstanders. Dit is een zeer eenvoudige strategie en wordt veelvuldig gebruikt.

H(it)strat

Als Fstrat, maar voorrang krijgt nu het slaan van een pion van een tegenstander. Als er met meerdere pionnen geslagen kan worden, dan wordt die pion genomen welke de verst gevorderde pion van een tegenstander slaat. In het uitzonderlijke geval dat meerdere pionnen dit kunnen bewerkstelligen, wordt de verst gevorderde pion van de eerstvolgende tegenstander genomen.

Deze strategie schenkt dus geen aandacht aan het beveiligen van de eigen pionnen. Ook deze strategie is relatief eenvoudig en schenkt het genoegen door slaan tegenstanders te ergeren.

R(andom)strat

De keuze van de pion wordt random gemaakt. Dit is de meest simpele strategie die met geen enkele van voornoemde overwegingen rekening houdt. Deze strategie van 'lukraak' spelen wordt nauwelijks gebruikt omdat dit als dom spelen beschouwd wordt.

T(hreat)strat

Het aspect 'pionnen zo veel mogelijk buiten het slaan bereik van vijandelijke pionnen te houden' komt aan bod bij de constructiemethode van de nieuw te introduceren strategie T(hreat)strat. Hiertoe formuleren we eerst precies wat een *directe* dreiging van een pion is. Een pion I van de ene speler wordt direct bedreigd door een pion II van een andere speler als deze bij diens eerstvolgende worp direct geslagen kan worden. Dit is het geval als

- voor de onderlinge afstand van II naar I geldt < D,
- of ook = D als de andere speler geen pionnen meer in het startgebied heeft,
- of ook als de pion I op het startveld van de andere speler staat en diens startgebied niet leeg is.

De strategie Tstrat is gebaseerd op verschilscores. De verschilscore TVScore van een pion die resulteert in het verzetten van deze pion naar een nieuwe positie is gelijk aan het aantal directe dreigingen van de eigen pionnen op de pionnen van de tegenstanders verminderd met het aantal directe dreigingen die de pionnen van de tegenstanders op de eigen pionnen hebben; dit nog met S = 2 vermeerderd als een pion van de tegenstander geslagen wordt of als een (veilig) thuisveld bereikt wordt. De beslissing volgens Tstrat is de verste pion te verzetten met de maximale TVScore.

(Na een meer specifieke analyse van spelstrategieën met verschillende waarden van het startgebied en thuisvelden is gekozen voor de waarde S = 2 die de strategiebepaling op basis van verschilscores zeer eenvoudig maakt.)

Voorbeeld (Tstrat)

In dit voorbeeld geven we de bordsituatie aan met de relatieve posities van de pionnen van de spelers op het bord ten opzichte van hun startveld. Beschouw voor M = 3 spelers de bordsituatie op de achterzijde van het bord (C = 48) in tabel 1. In figuur 3 is deze bordsituatie gevisualiseerd. Laat hierbij speler m = 3 aan de beurt zijn met een worp van d = 3 ogen. Hij moet dan een keuze maken tussen het verzetten van pion p = 3 en p = 4.

Figuur 3. Bordsituatie behorende bij spelpositie in tabel 1

SPELPOSITIE (C = 48)	m=1 (geel)	m=2 (blauw)	m=3 (groen)
p = 1	0	20	0
p = 2	0	21	0
p = 3	3	22	10
p = 4	52	43	14

Tabel 1. Spelposities voor speler m = 3 en een worp met de dobbelsteen van d = 3 ogen

Na het verzetten van pion 3 (door speler 3) bedreigt deze geen vijandelijke pion en wordt ook niet bedreigd; pion 4 van speler 3 bedreigt pion 3 van speler 1 en wordt zelf niet bedreigd. De TVScore van pion 3 is dus gelijk aan 0 (pion 3) + 1 (pion 4) = 1. Na het verzetten van pion 4 (door speler 3) blijft pion 3 van speler 3 door drie pionnen van speler 2 bedreigd en pion 3 bedreigt zelf niets; pion 4 bedreigt nu pion 3 van speler 1 en wordt zelf bedreigd door pion 2 van speler 1. Deze speler heeft immers nog pionnen in zijn startgebied en bij een worp van d=6 ogen slaat pion 2 van speler 1 pion 4 van speler 3. De TVScore van pion 4 is dus gelijk aan -3 (pion 3) + 0 (pion 4) = -3. Pion 3 geeft dus de maximale TVScore en dus is de beslissing van speler 3 het verzetten van pion 3.

Winstkansen voor alle borden en spelers

Voor de spelanalyse is een speciaal MATLAB-computerprogramma geschreven, geschikt voor allerlei waarden van M, P, C, D en strategiecombinaties (ook voor afwijkende startveldposities en bij coalitievorming van spelers). Winstkansen zijn hiermee bepaald door simulatie

Nr	StratComb	M = 2	M = 3	M = 4	M = 6
1	F(F) (F)F	51 49	34 33	25 25	17 16
2	H(F) (F)H	61 59	41 41	31 31	19 18
3	R(F) (F)R	62 60	51 50	38 37	22 21
4	T(F) (F)T	86 85	79 78	61 60	44 44
5	F(H) (H)F	41 39	26 27	20 19	15 15
6	H(H) (H)H	50 50	34 33	25 25	16 17
7	R(H) (H)R	47 46	33 32	30 29	18 17
8	T(H) (H)T	81 80	75 74	60 59	45 45
9	F(R) (R)F	40 38	17 17	15 14	11 11
10	H(R) (R)H	54 53	27 26	19 19	14 13
11	R(R) (R)R	51 49	34 33	25 25	17 17
12	T(R) (R)T	82 81	72 71	52 51	42 41
13	F(T) (T)F	15 14	05 04	05 05	04 04
14	H(T) (T)H	20 19	06 06	07 07	05 05
15	R(T) (T)R	19 18	08 08	09 09	06 06
16	T(T) (T)T	51 49	34 33	26 25	17 16

Tabel 2. Winstkansen van een strategie tegen tegenstanders met allen dezelfde strategie (in %)

met een simulatierun van 100.000 spelen, leidend tot een nauwkeurigheid van ± 0,5%. Het verschil van twee kansen heeft dus steeds een nauwkeurigheid van ± 1%.

Tabel 2 geeft voor een speler met een bepaalde strategie zijn winstkansen tegen tegenstanders die allemaal op hun beurt dezelfde strategie voeren. De tabel bevat 16 maal twee strategiecombinaties. De kansen zijn gegeven voor de eerste en laatste positie van de speler op de volgende manier. Bijvoorbeeld, voor Nr = 2 hebben we de strategiecombinaties H(F) en (F)H. De strategiecombinatie H(F) geeft de eerste speler de strategie H(strat) en alle andere spelers de strategie F(strat). De combinatie (F)H geeft juist het tegengestelde: de laatste speler speelt strategie H en alle andere spelers spelen strategie F. De winstkans hangt af van het totaal aantal spelers M. Zo is de winstkans van de speler met strategie H voor M = 2 gelijk aan 61% voor de combinatie H(F) en 59% voor de combinatie (F)H; voor M = 6 is deze kans 19% voor H(F) en 18% voor (F)H (zie tabel 2, Nr = 2).

Tabel 2 toont duidelijk de toenemende kwaliteit van strategieën in de volgorde F, H, R, T. Dit geldt voor alle varianten van het aantal spelers. Bijvoorbeeld, voor Nr = 1 - 4 nemen de winstkansen tegen (F) toe in deze volgorde. Hetzelfde geldt voor de andere verzamelingen van de 4 combinaties 5 - 8, 9 - 12 en 13 - 16. Enkele uitzonderingen zijn er voor de strategieën H en R waarbij soms H een iets grotere winstkans heeft dan R.

De nieuwe strategie T domineert de anderen vrij sterk. Een verrassing is de slechte kwaliteit van de vaak gespeelde strategie F. De toevalstrategie R doet het zelfs beter en heeft min of meer dezelfde kwaliteit als strategie H. In de meeste gevallen is de winstkans van een speler die start iets hoger dan die waarbij hij de laatste beurt heeft. Bijvoorbeeld, voor Nr = 4 and M = 2 is de winstkans 86% bij T(F) en 85% bij (F)T.

Uit tabel 2 volgt de som van de winstkansen van de tegenstanders eenvoudig door de individuele winstkans af te trekken van 100%. Echter, deling van deze som door het aantal tegenstanders geeft niet altijd een goede benadering van de individuele winstkansen van de tegenstanders. De reden is dat de onderlinge positie van de tegenstanders een belangrijke rol speelt. Er zijn talrijke

Nr	StrC	Kansen (%)
1	FFFF	25 25 25 25
2	FFFH	17 15 27 31
3	FFFR	20 21 22 37
4	FFFT	14 13 13 60
5	HHHF	34 24 22 19
6	HHHH	25 25 25 25
7	HHHR	29 22 20 29
8	HHHT	08 12 11 59
9	RRRF	28 29 29 14
10	RRRH	25 27 29 19
11	RRRR	25 25 25 25
12	RRRT	17 16 16 51
13	TTTF	32 31 32 05
14	TTTH	31 30 32 07
15	TTTR	31 30 30 09
16	TTTT	25 25 25 25

Tabel 3. Winstkansen (in %) voor de verschillende spelers bij enkele strategiecombinaties en M = 4 spelers

verschillende volgordes van de spelers en daarom beperken we ons tot wat eenvoudige voorbeelden.

Tabel 3 geeft een illustratie voor M = 4 spelers. De strategiecombinaties Nr = 1 – 16 zijn dezelfde als in tabel 2 maar nu alleen met de laatste positie voor de speler met de enkele (in de meeste gevallen afwijkende) strategie. Dus StrC = FFFF in tabel 3 komt overeen met StratComb = F(F) in tabel 2, StrC = FFFH komt overeen met StratComb = F(H). De winstkansen voor de vierde speler komen overeen met de getoonde winstkansen in tabel 2, tweede kolom bij M = 4.

We zien dat de kansen van strategiecombinaties die de strategie H bevatten sterk afhangen van de volgorde. Voor de andere strategiecombinaties verschillen de winstkansen niet veel of zijn zelfs aan elkaar gelijk.

Slotconclusies

In dit artikel hebben we laten zien dat de gebruikelijke eenvoudige strategieën Fstrat, Hstrat en Rstrat sterk gedomineerd worden door de nieuwe strategie Tstrat, die ook eenvoudig is toe te passen. Verrassend is dat de vaak gespeelde strategie Fstrat slecht presteert tegen de andere strategieën. De toevalstrategie Rstrat doet het zelfs veel beter en heeft een vergelijkbare kwaliteit met de strategie Hstrat.

In het *working paper* (Van der Genugten & Das, 2018) worden ook andere bordvarianten beschouwd en strategieconstructies die gebaseerd zijn op meer ingewikkelde en dus minder toepasbare scores van spelsituaties. Ook worden speelduren vermeld. Voor verder onderzoek is het interessant na te gaan of de strategie Tstrat nog verbeterd kan worden door strategieën die ook eenvoudig in de praktijk zijn toe te passen.

Literatuur

Genugten, B., van der, & Das, M. (2018). *Strategies for the board game 'Man, don't get upset', the Dutch variant of Ludo*. Working Paper, Tilburg University.

Wikipedia (2019). *Ludo* (*board game*). https://en.wikipedia.org/wiki/Ludo_(board_game) (26-04-2019)

Ben van der Genugten is emeritus hoogleraar Waarschijnlijkheidsrekening en Statistiek, Tilburg University.
E-mail: b.vandergenugten@chello.nl

Marcel Das is directeur van CentERdata en hoogleraar Econometrie en Dataverzameling, Tilburg University.
E-mail: das@uvt.nl

THE BEAUTIFUL MATHEMATICS OF THE CARD GAME SET

The card game SET was invented in 1974 by Marsha Falco, a population geneticist at Cambridge university. When explaining the combinatorics of genes to veterinarians, she used cards with symbols to visually represent expressions of various genes. She quickly realised that combining these symbols could be made into a great game. Apart from being a fun game to play, SET connects to more serious mathematics with applications in combinatorics and computer science. Here, we will give a glimpse of the rich combinatorics behind the game and its relation to recent research on the *cap set problem*.

Dion Gijswijt

The cards in SET are characterised by four attributes, each with three possible values: Shape (diamond, oval, or peanut), Number (one, two, or three), Colour (red, green, or purple), and Shading (open, striped, or filled). There are 81 cards in total: one card for each combination of attributes. Three cards form a SET if for each of the four attributes the cards either have three different values, or they all have the same value. In Figure 1 two examples of SETs are shown. In the SET on the left, the three cards are the same for the attributes Shape and Number, and are all different for the attributes Colour and Shading. In the SET on the right, the three cards are all different for each of the four attributes.

Figure 1. Two examples of a SET. In the first, two attributes are equal and two are all-different. In the second, the four attributes are all-different

The game begins by placing twelve cards face up on the table. The first player to spot a SET shouts out 'SET!' and points out the found SET. Assuming the player correctly identified a SET, she may take the three cards. The empty spots are filled with new cards. If no SET can be formed using the cards on the table, three new cards are added. The game ends when all cards are used up and no SETs can be formed from the remaining cards on the table. The player who has collected the most SETs is declared the winner.

Puzzle 1

Can you find the six SETs among the twelve cards in figure 2?

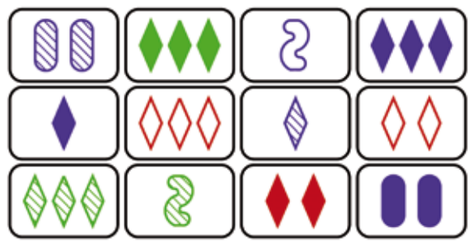


Figure 2. Find the six SETs

Finite geometry

The mathematical structure of SET becomes apparent when we encode the cards in the following way. Let $\mathbb{F} = \{0,1,2\}$ be the field of three elements.[1] For each of the four attributes, we encode the three possible values by the elements from $\mathbb{F}$. For instance, let's encode them as follows:

Colour:	red=0,	green=1,	purple=2
Shape:	diamond=0,	oval=1,	peanut=2
Number:	one=0,	two=1,	three=2
Shading:	open=0,	striped=1,	full=2

This way, the leftmost card in Figure 1 becomes the vector (0,1,2,0). The other two cards in that set become (2,1,2,1) and (1,1,2,2). The 81-card deck is thus identified with the four-dimensional space $\mathbb{F}^4$.

Surprisingly, the SETs now have a natural algebraic interpretation. Indeed, three different 'cards' $x,y,z \in \mathbb{F}^4$ form a SET if and only if

$$x + y + z = 0 \qquad (1)$$

You may want to check that the three cards we just encoded indeed sum to zero: (0,1,2,0) + (2,1,2,1) + (1,1,2,2) = (0,0,0,0) modulo 3. SETs can also be interpreted geometrically. Three cards x,y,z form a SET if and only if they lie on a common line in $\mathbb{F}^4$.

A simple consequence of (1) is that for any two cards x,y there is a unique third card that completes the SET, namely $z = -x - y$. It follows that the total number of SETs in the game equals $\frac{1}{3}\binom{81}{2} = 1080$.

Although there are 4 attributes in SET, the game is easily generalised to n attributes. For instance, the case $n = 3$ can be modelled by using only the 27 red cards and the case $n = 5$ can be modelled by introducing an additional attribute Size (small, medium, and large) to get (a somewhat impractical) game with 243 cards. In the general situation, cards correspond to vectors of length n (i.e. elements of $\mathbb{F}^n$) and three different cards x, y, and z form a SET if and only if they satisfy (1). We will come back to the higher dimensional setting later.

Puzzle 2

A magic square is a 3×3 array of SET cards in which all rows, columns, and diagonals form a SET. Let x, y, and z be three cards that do not form a SET. Show that we can always (uniquely) complete the following magic square:

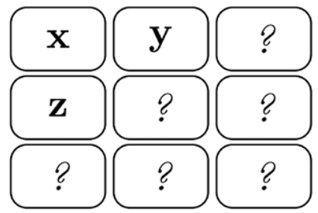


Figure 3. The resulting set of 9 cards is a 2-dimensional plane inside $\mathbb{F}^4$

How many SETs are there within the initial 12 cards? It is quite possible that there is no SET at all. The expected number of SETs is equal to $\frac{1}{79}\binom{12}{3} \approx 2{,}78$ since choosing three distinct cards uniformly at random gives a SET with probability 1/79. The maximum possible number of SETs in 12 cards is harder to determine. A wonderful exposition of this problem was presented by N.G. de Bruijn (2002) in *Nieuw Archief voor Wiskunde*. It turns out that the maximum number of SETs is equal to 14 and the optimal configuration is unique (up to affine transformations).

At the opposite end, we may ask for the number of cards that remain at the end of the game. This number can be equal to 0, 6, 9, 12, 15, or 18. According to computer simulations (see McMahon, Gordon, H. Gordon & R. Gordon), 6 or 9 remaining cards is most likely (46,8% and 44,5% of the time, respectively). The situation with 3 cards is missing from the list. This is not a typo: there is a simple mathematical reason why this situation cannot occur!

Puzzle 3

A game of SET cannot end with only three cards remaining on the table. Why not?

Caps

A collection of twelve cards does not necessarily contain a SET. In fact, the largest number of cards without a SET is equal to 20 as was shown by Pellegrino in 1971 (Pellegrino, 1971; three years before SET was invented!). The unique (up to affine transformations) configuration of 20 cards without a SET is depicted in Figure 4.

A subset of $\mathbb{F}^n$ that does not contain three points on a line is called a *cap* or a *cap set*. This corresponds to a collection of cards in n-attribute SET of which no three cards form a SET. The maximum size of a cap in $\mathbb{F}^n$ is denoted $a(n)$. So for the normal game of SET we have $a(4) = 20$. Finding upper and lower bounds on $a(n)$ (and more generally for caps in affine and projective spaces over finite fields) is one of the central questions in finite geometry.

Puzzle 4

Show that $a(2) = 4$

In dimension 3, there is a cap of size 9 This can be seen from Figure 4 by considering the 9 dots in the top three 3×3 squares. In fact, this is optimal: $a(3) = 9$. The proof is a short counting argument. We provide it here for the interested reader, but it can be safely skipped.

Proof. Suppose that $a(3) > 9$. Then, there is a cap $C \subseteq \mathbb{F}^3$ of size 10. If we partition $\mathbb{F}^3$ into three parallel planes, then each plane contains at most $a(2) = 4$ elements from C. The 10 elements of C are therefore distributed over the three planes according to one of the two partitions $10 = 4 + 4 + 2$ and $10 = 4 + 3 + 3$.

Let H_0 be a plane that contains 2 or 3 elements from C. Say it contains a, b, and possibly an element c. Let l be the line through a and b. There are three planes H_1, H_2, H_3 that intersect H_0 in l. Together, H_0, H_1, H_2, H_3 contain all elements of $\mathbb{F}^3$. (Figure 5)

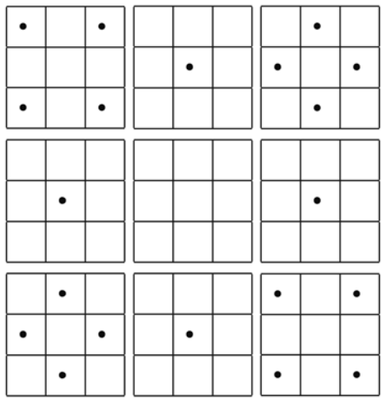
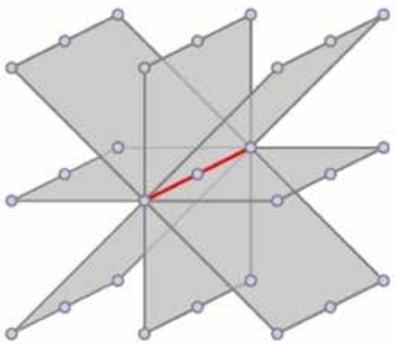


Figure 4. The four-dimensional space $\mathbb{F}^3$ is depicted as a 3×3 grid of 3×3 squares. The squareswith dots represent the 20 vectors that form the Pellegrino cap



Figuur 5. In $\mathbb{F}^3$, there are four planes through a line. Together, they cover the whole space

Since H_1, H_2, and H_3 contain both a and b, they each contain at most 2 other elements from C. Also, H_0 contains at most 1 other element of C. This gives a total of $2 + (2 + 2 + 2 + 1) = 9$ elements in C, a contradiction!

Pellegrino's result that $a(4) = 20$ is quite a bit harder to prove[2], but a short argument based on clever counting can be found in Davis & Maclagan (2003). Also for dimensions 5 and 6 the maximum size of a cap is known, see Table 1. For dimensions $n \geq 7$, the exact value of $a(n)$ is not known, although there exist lower bound (via constructions) and upper bounds (using Fourier analysis).

Challenge 1

Determine (with or without computer assistance) the number $a(7)$.

n	1	2	3	4	5	6	7
a(n)	2	4	9	20	45	112	?

Table 1. Known values of $a(n)$. Sequence *A090245* in the OEIS

The cap set problem

How does the maximum cap size $a(n)$ grow when we let $n \to \infty$? Consider all 2^n cards with the property that every coordinate is equal to 0 or 1. Three of these cards cannot form a SET, so we have a cap of size 2^n. Since the total number of cards is 3^n, we find that $2^n \leq a(n) \leq 3^n$. A little analysis shows that there exists a number σ such that $a(n)$ grows roughly as σ^n. To be precise, $\sigma = \lim_{n\to\infty} a(n)^{1/n}$. The number σ is called the *asymptotic solidity* and $2 \leq \sigma \leq 3$ because of the remarks above.

The *cap set problem* asks whether $\sigma = 3$, see for example Terence Tao's blog post (Tao, 2007). We can find lower bounds on σ by constructing large caps. If there is a cap of size M in the n-attribute SET, then we obtain the lower bound $\sigma \geq M^{1/n}$. For example, the Pellegrino cap gives the lower bound $\sigma \geq 20^{1/4} \approx 2{,}1147$. The current record is $\sigma \geq 2{,}217389$ obtained by Edel (2004) by constructing a very large cap in $\mathbb{F}^{480}$ (a variant of SET with 480 attributes!). On the other hand, better and better *upper bounds* have also been found. Until recently, the best upper bound was $a(n) \leq 3^n / n^{1+\epsilon}$ for a very very small $\epsilon > 0$. This would suggest that maybe $\sigma = 3$.

However, in 2016 Jordan Ellenberg and myself unexpectedly solved the cap set problem by showing that $\sigma \leq 2.75511$ (Ellenberg & Gijswijt, 2017). Our proof uses the so-called polynomial method, and is based on earlier work by Croot, Lev and Pach (2017). An interesting feature of the proof is that it is very short and elementary. In fact, the proof is only about 2 pages! The details can be found in Ellenberg and Gijwijt (2017).

Although the cap set problem seems like an innocent question about a simple card game, it has connections to many other areas of mathematics. For instance, the result that $\sigma < 3$ also resolves other combinatorial conjectures such as the *Erdős-Szemerédi sunflower conjecture*. This in turn refutes a conjecture by Coppersmith and Winograd about possible ways of obtaining *Fast Matrix Multiplication*: a scheme for multiplying dense $n \times n$ matrices asymptotically much faster than the best algorithms currently known. For more about these connections, see for instance (Kalai, 2017).

Notes

1. This just means that we compute modulo 3. For instance, $1 + 2 = 0$ and $2 \cdot 2 = 1$.
2. Donald Knuth has written a computer program to count the number of caps in of each size using the group of affine transformations to reduce the search space. See Knuth (n.d.).

References

Bateman, M., & Katz. N. (2012). New bounds on cap sets. *Journal of the American Mathematical Soci*ety, 25(2), 585–613.

Bruijn, N. G. de. (2002). Set! *Nieuw Archief voor Wiskunde, 3*, 320–325.

Croot, E., Lev, V. F., & Pach, P. P. (2017). Progression-free sets in are exponentially small. *Annals of Mathematics 185*(1), 331–337.

Davis, B., & Maclagan, D. (2003). The card game SET. *The Mathematical Intelligencer, 25*(3), 33–40.

Edel, Y. (2004). Extensions of generalized product caps. *Designs, Codes and Crypt, 31*(1), 5–14.

Ellenberg, J. & Gijswijt, D. (2017). On large subsets of with no 3-term arithmetic progression. *Annals of Mathematics 185*(1), 339–243.

Kalai, G. (2017). *Polymath 10.* https://gilkalai.wordpress.com/category/polymath10/

Knuth, D. (n.d.). *SETSET-ALL.* http://www-cs-faculty.stanford.edu/~uno/programs.html

McMahon, L., Gordon, G., Gordon, H., & Gordon, R. (2016). *The Joy of SET.* Princeton University Press.

Pellegrino, G. (1971). Sul massimo ordine delle calotte in $S_{4,3}$. *Matematiche (Catania), 25*(10).

Tao, T. (2007). *Open question: best bounds for cap sets.* https://terrytao.wordpress.com/2007/02/23/open-question-best-bounds-for-cap-sets/

Dion Gijswijt is universitair docent wiskunde aan de Technische Universiteit Delft.
E-mail: dion.gijswijt@gmail.com

Bij darten wordt vaak het gemiddelde en het checkout percentage van spelers getoond. Op basis daarvan hebben Steffen Liebscher en Thomas Kirschstein een model ontwikkeld om te voorspellen wie een wedstrijd gaat winnen. Je kunt die voorspelling verbeteren door een andere definitie te kiezen voor het gemiddelde en het checkout percentage dan gebruikelijk. Maar kan het model ook de gokmarkt verslaan?

DARTS het voorspellen van de uitkomst van profwedstrijden

Miriam Loois

Darts is de laatste jaren steeds vaker op tv te zien. Vroeger hadden 'we' vooral Raymond van Barneveld, tegenwoordig zijn er steeds meer Nederlandse toppers. Michael van Gerwen is al jaren de nummer een van de wereld. Bij darts gooi je om de beurt drie pijlen op een dartbord. In figuur 1 is te zien hoe een dartbord eruit ziet en hoeveel punten je krijgt afhankelijk van waar je gooit. Een dartbord bestaat uit 20 delen, en een enkele en dubbele Bullseye. De dubbele of binnenste Bullseye is 50 punten waard, de enkele 25. Voor de zwarte en witte vakjes geldt dat je de score tussen 1 en 20 krijgt. De buitenste gekleurde ring verdubbelt het aantal punten, de binnenste gekleurde ring verdriedubbelt het aantal punten. Je krijgt dus de meeste punten als je de triple 20 raakt.

Spelregels

Het doel bij darts is om zo snel mogelijk een score van 501 te halen (terugtellend van 501 tot 0), dan heb je één leg gewonnen. Je moet altijd eindigen met een darts in de buitenste ring, een 'dubbel'. In theorie is het mogelijk om in 9 darts uit te gooien, maar dat gebeurt bijna nooit. Op televisie is dat sinds 1984 53 keer voorgekomen. Je kunt een 9-darter gooien door bijvoorbeeld zeven keer triple 20 te gooien, één keer triple 19, en te eindigen met dubbel 12.

Het commentaar bij darts kan wiskundigen veel stof tot nadenken geven. Zo komen er uitspraken voorbij als 'je kunt een wedstrijd niet winnen in de eerste *leg*, maar wel verliezen', en 'hij staat verdedigend te darten'. Qua statistieken wordt er bijna altijd gekeken naar het gemiddelde en het checkout percentage:

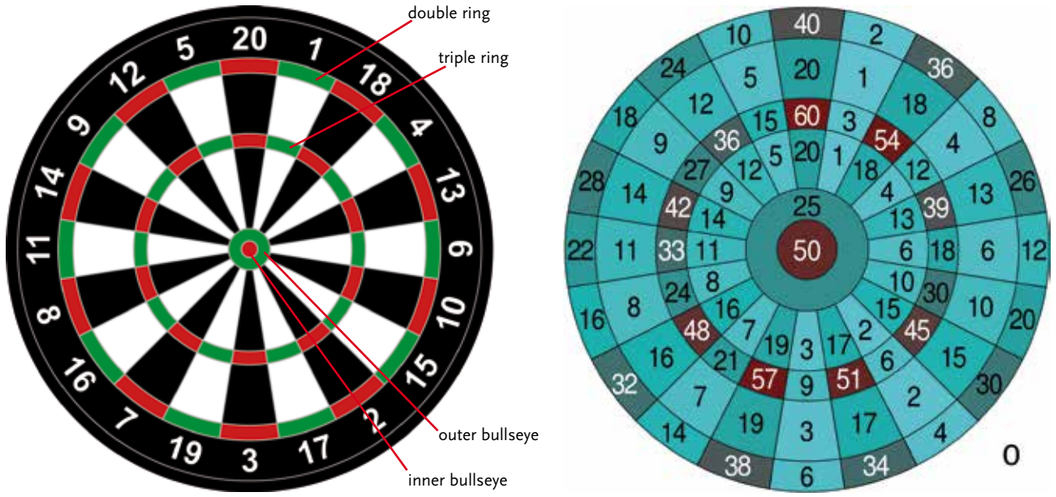
Het gemiddelde
Het gemiddelde wordt berekend door de gemiddelde score van alle gegooide pijlen te berekenen, en te vermenigvuldigen met drie. Topdarters hebben vaak een gemiddelde tussen de 90 en (ruim) 100.

Het checkout percentage
Het checkout percentage wordt berekend door naar alle mogelijkheden te kijken om de leg met één pijl uit te gooien. Dit kan als de overgebleven score ofwel 50 is (uit te gooien met dubbele Bull), of een even getal tussen 2 en 40. Het checkout percentage is het percentage succesvolle pogingen. Als een darter dus 40 over heeft, eerst twee pijlen mist, en dan dubbel 20 gooit, geldt dit als één succesvolle poging uit drie.

Bij darts wisselt het format van een wedstrijd vaak. De meeste wedstrijden zijn van de vorm 'best of *x* legs'. Bij 'best of 7' is de speler die als eerste 4 legs heeft gewonnen de winnaar. Naarmate het toernooi vordert, neemt het aantal te winnen legs toe. Vaak is het genoeg om met één leg verschil te winnen. De spelers mogen om en om starten met gooien in een leg, dus in zo'n format is de startende darter in het voordeel. Soms, zoals bij het WK, is het format 'best of *x* sets'. Een set wordt dan weer gewonnen door de speler die als eerste 3 legs heeft. Er zijn nog meer varianten, bijvoorbeeld een waarin de eerste pijl altijd een dubbel moet zijn.

Model voor het voorspellen van de uitkomst

Steffen Liebscher en Thomas Kirschstein van de Martin-Luther-Universität Halle-Wittenberg hebben een model ontwikkeld om de uitkomst van een dartwedstrijd te voorspellen. Het model is geschat op ruim 800 wedstrijden uit 2016 op professioneel niveau. Alleen wedstrijden van het format 'best of *x* legs' zijn meegenomen. In hun model hangt de kans dat de beginnende speler *a* wint van speler *b* af van het verschil in gemiddelde en checkout percentage tussen de twee spelers, en een constante die



Figuur 1. (links) Een dartbord; (rechts) De score die je krijgt afhankelijk van waar je gooit. Bron: https://en.wikipedia.org/wiki/Darts

het voordeel van het starten met gooien aangeeft:

$$\phi^{-1}(P_{ab}) = c + \alpha(A\nu_a - A\nu_b) + \beta(CO_a - CO_b)$$

Waarbij:
P_{ab} = Kans dat speler *a* (beginnende speler) wint van speler *b*
$\phi(x)$ = Cumulatieve standaard normale verdeling
$A\nu_i$ = Gemiddelde van speler *i*
CO_i = Checkout-percentage speler *i* (tussen 0 en 100)

Het model is geschat op de helft van de data (de trainingsset), de andere helft is gebruikt om het model te valideren (validatieset). Ook is de McKelvey & Zavoina's R^2 berekend. Dit is een zogenaamde 'pseudo R^2', een alternatief voor de R^2 die gebruikt wordt bij lineaire regressie. Er zijn verschillende pseudo R^2-maten, die gebruikt kunnen worden voor bijvoorbeeld logistische regressie, GLM's, of een probit model zoals hier.

Als het model geschat wordt met de gebruikelijke definities voor het gemiddelde en checkout-percentage levert dit een McKelvey & Zavoina's R^2 op van 60%. De drie geschatte parameters zijn te vinden in tabel 1. Laten we als voorbeeld een wedstrijd nemen tussen Michael van Gerwen en Vincent van der Voort. In de trainingsset heeft van Gerwen een gemiddelde en checkout-percentage van 102,7 en 46,2% en van der Voort 92,6 en 40,4% . Als van Gerwen start, is de kans dat hij de wedstrijd wint volgens het model dus:

$$\phi\left((0{,}076 + 0{,}135 * (102{,}7 - 92{,}6) + 0{,}025 * (46{,}2 - 40{,}4)\right) = \phi(1{,}58) = 94{,}3\%.$$

Het model kan 72% van de uitslagen in de validatieset goed voorspellen. Ruim meer dan de 50% als je een muntje op zou gooien om te bepalen wie er gaat winnen. Hier is echter wel een kanttekening bij te plaatsen. Om te voorspellen worden de gemiddeldes en checkout-percentages uit de gehele datset gebruikt (en de parameters uit de trainingsset). En die bevatten ook het gemiddelde en checkout-percentage van de te voorspellen wedstrijd. Als iemand vaak in de dataset voorkomt is dat niet zo erg. Maar als iemand bijvoorbeeld maar één keer in de validatieset voorkomt, is het een beetje alsof je bij voetbal voorspelt dat degene die de meeste doelpunten scoort wint. Daarom zullen we, als we verderop het model gaan vergelijken met de quotes op goksites, gebruik maken van de gemiddeldes en checkout-percentages uit de trainingsset om te voorspellen.

Om het model te verbeteren hebben Steffen en Thomas een andere definitie gekozen voor het checkout-percentage. Ze hebben niet naar de mogelijkheid om te finishen per pijl gekeken, maar per beurt van drie pijlen. Als een speler voor hij begint aan een reeks van 3 pijlen op een finish staat (een resterende score tussen 2 en 158, of 160, 161, 164, 167, of 170), telt dit als één mogelijkheid. Het maakt dan niet uit of hij uitgooit in één, twee of drie pijlen voor zijn succespercentage. Indien deze definitie wordt gehanteerd, stijgt de McKelvey & Zavoina's naar 67%. Ook stijgt het percentage goed voorspelde wedstrijden in de validatieset licht.

Verbeterd model

Maar het is mogelijk om het model nog verder te verbeteren, door ook een andere definitie voor het gemiddelde te kiezen. Het gemiddelde wordt nu gebaseerd op alle pijlen. Op het moment dat iemand kan finishen door bijvoorbeeld dubbel 8 te gooien, zal dit zijn gemiddelde laten dalen, ook al gooit hij precies waar hij wil gooien. Daarom is in een alternatieve definitie alleen het gemiddelde meegenomen van de pijlen indien de resterende score hoger is dan 180. Hierdoor stijgt de McKelvey & Zavoina's R^2 nog iets verder naar 70%. Op basis van deze nieuwe definities heeft van Gerwen een gemiddelde van 111,6 en een checkout-percentage van 48,1%. Interessant is dat bij de nieuwe definities een stijging van het gemiddelde van 10 ongeveer evenveel oplevert als een stijging van het checkout-percentage met 10%. In de andere modellen telt het gemiddelde relatief (veel) zwaarder.

	Constante (voordeel startende speler)	α (gewicht verschil gemiddelde)	β (gewicht verschil checkout-percentage)	McKelvey & Zavoina's *R2*
Gebruikelijke definities	0,076 (0,069)	0,135 (0,013)	0,025 (0,007)	60%
Aangepaste definitie checkout-percentage	0,049 (0,070)	0,101 (0,015)	0,062 (0,012)	67%
Aangepaste definitie gemiddelde en checkout-percentage	0,056 (0,071)	0,079 (0,011)	0,080 (0,011)	70%

Tabel 1. Geschatte parameters (en standaarddeviatie) bij verschillende definities van gemiddelde en checkout-percentage

Kun je dit model gebruiken om de gokmarkt te verslaan? Op de site www.oddsportal.com is terug te vinden hoeveel je op goksites kon winnen als je geld inzette op een dartswedstrijd. Neem bijvoorbeeld de wedstrijd tussen Michael van Gerwen en Garry Anderson tijdens de Champions League of Darts in 2016. Als je van tevoren één euro had ingezet op van Gerwen, en hij zou winnen, kreeg je € 1,34 uitbetaald, 34 cent winst. Bij Anderson had je € 3,23 gekregen, € 2,23 winst. Op basis van die quotes is het mogelijk om uit te rekenen wat de winkansen zijn die de markt inschat. In dit geval is de geschatte kans dat van Gerwen wint

$$\frac{1}{1{,}34} \Big/ \left(\frac{1}{1{,}34} + \frac{1}{3{,}23}\right) = 70{,}7\%.$$

Om te testen of het model de markt kan verslaan vergelijken we hier de voorspelde kans dat de uiteindelijke winnaar zal winnen. De voorspelling is dan correct als die kans groter dan 50% is. Op deze manier zijn 264 wedstrijden in de validatieset vergeleken. Het was niet mogelijk om alle wedstrijden te vergelijken, omdat voor sommige wedstrijden geen quote beschikbaar was, of omdat een speler niet in de trainingsset zat, en er daarom geen statistieken van hem beschikbaar waren. In figuur 2 is de voorspelde kans van de goksites vergeleken met de kans op basis van het statistische model. De markt had het 182 keer goed (68,9%), het model 176 keer (66,7%). Het model verslaat dus nog niet de markt, maar komt wel in de buurt.

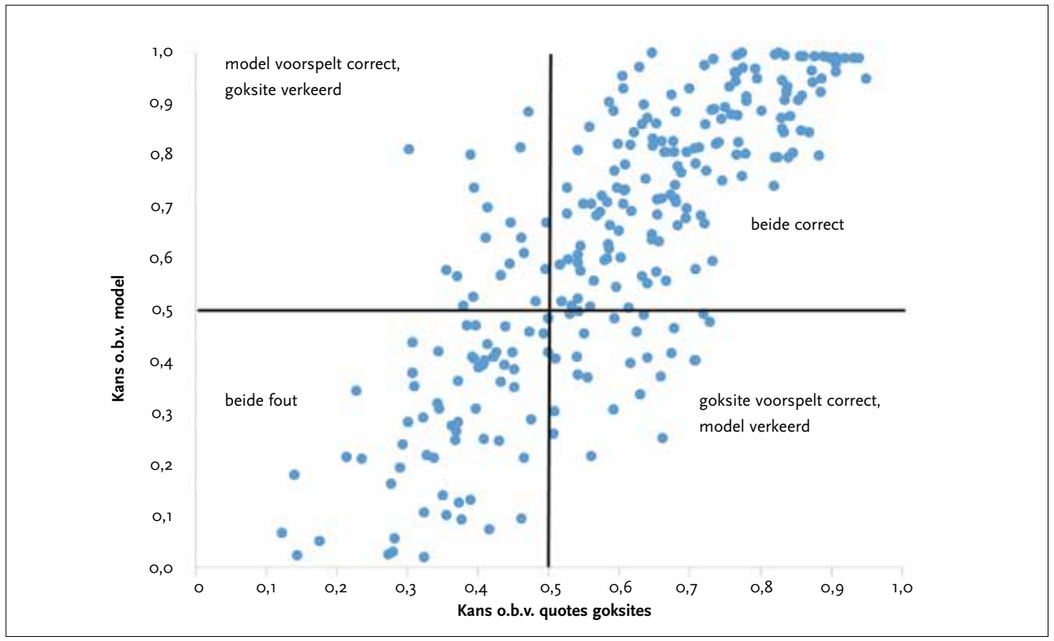
Kan het model verbeterd worden zodat de markt wel verslagen kan worden? Er is nog winst te behalen, door het model bijvoorbeeld op meer data te schatten. Ook kan er gekeken worden hoe het gemiddelde en checkout-percentage zo goed mogelijk geschat kunnen worden, bijvoorbeeld via een gewogen gemiddelde van afgelopen wedstrijden. En wellicht is er dan een (kleine) marge te behalen door statistisch verantwoord te gokken.

Many thanks to Thomas Kirschstein for providing the data he used for his article, providing R-code to scrape results and answering many questions.

Literatuur

Data wedstrijdverloop: http://live.dartsdata.com/
Data quotes goksites: http://www.oddsportal.com/darts/results/
Liebscher, S., & Kirschstein, T. (2017). Predicting the outcome of professional darts tournaments. *International Journal of Performance Analysis in Sport*, *17*(5), 666–683.
Loois, M. (2018). De financiële wiskunde achter het gokken op sportwedstrijden. *De actuaris, 25*(3).

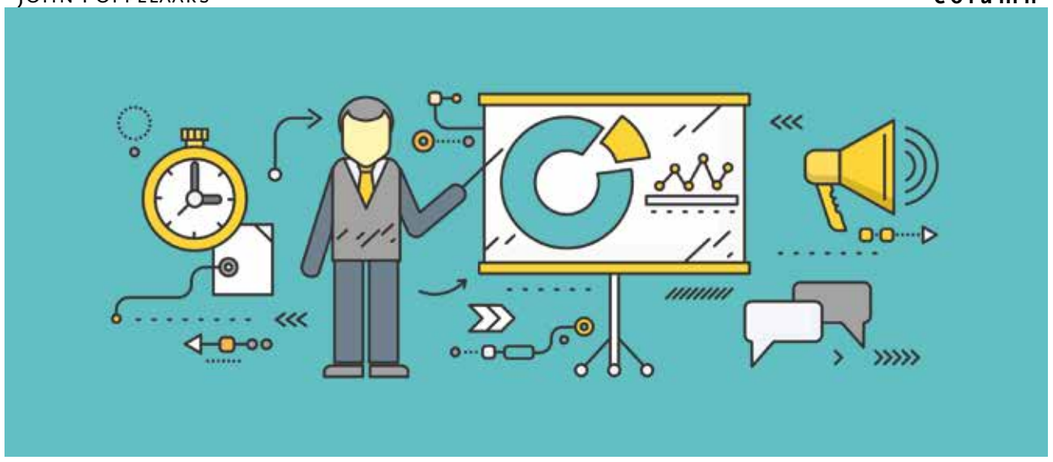
Miriam Loois is docent Toegepaste Wiskunde aan de Hogeschool van Amsterdam.
E-mail: miriamloois@gmail.com



Figuur 2. Van tevoren geschatte kans dat uiteindelijke winnaar de wedstrijd zal winnen (goksites versus model)

John Poppelaars

column



VERDER DAN VOORSPELLEN

Met enige regelmaat geef ik lezingen over het gebruik van data en modellen in besluitvorming. Laatst voor de vereniging van brancheorganisaties voor mediabedrijven en uitgeverijen, de Mediafederatie[1]. Om me te verdiepen in de uitdagingen van de uitgeefbranche had ik ter voorbereiding het boek *The Bestseller Code*[2] van Jodie Archer en Matthew Jockers gelezen. Archer & Jockers hebben een voorspelmethode ontwikkeld die op basis van eigenschappen als thema, plot, stijl, karakters en het vocabulaire van een manuscript voorspelt of het een bestseller wordt. Het is grappig om te lezen dat uit hun analyse volgt dat *Fifty Shades of Grey* en *Harry Potter* erg vergelijkbare boeken zijn. Je zou toch anders verwachten! Na mijn lezing kwam een aantal van de aanwezige uitgevers bij me langs om meer details over deze magische formule voor bestsellers te weten te komen.

De reactie van de uitgevers op de *Bestseller Code* is een mooi voorbeeld van hoe de aandacht voor en de daadwerkelijke toepassing van data-analyse en voorspelmodellen aan het toenemen is. Steeds meer organisaties starten data labs of stellen teams van data scientists samen om data te verzamelen en te analyseren, op zoek naar de 'bestseller-code' patronen waarmee de uitdagingen van die organisaties kunnen worden aangepakt. Het precieze aantal organisaties dat actief *analytics* methoden gebruikt is lastig te bepalen. Forbes[3] schat in dat 53% van de bedrijven bij hun besluitvorming *analytics* toepast. De verwachting is dat dit percentage in de komende jaren verder zal toenemen. Daarbij concentreren de meeste organisaties zich nu op het ontwikkelen van verklarende modellen of voorspelmodellen (*root cause* of *predictive analytics*), modelgedreven besluitvorming (*prescriptive analytics*) is maar nog beperkt aanwezig. Slechts 11% van grote en middelgrote bedrijven gebruikt een vorm van *prescriptive analytics* in besluitvorming volgens Gartner[4] in hun laatste marktscan. Daarmee laten, in mijn optiek, veel bedrijven de kans liggen om hun prestaties substantieel te verbeteren. Bovendien laten ze veel van de waarde die de voorspelmodellen kunnen genereren onbenut. Een geanonimiseerd voorbeeld laat zien waarom.

De operations manager van een productiebedrijf wil af van rigide onderhoudsschema's voor de machines, ze leiden tot onnodig productiecapaciteitsverlies en hoge onderhoudskosten. Het is de ambitie van het bedrijf om het onderhoud kosteneffectief te maken door over te stappen op conditiegebaseerd onderhoud. Met de data van sensoren in de machines kunnen conditievoorspelmodellen worden gemaakt die de toekomstige machineconditie kunnen voorspellen waardoor onderhoud kan worden ingepland wanneer dat ook echt nodig is. De verwachting is dat dit tot een substantiële kostenbesparing zal leiden. Samen met het data science team van het bedrijf ontwikkelt de onderhoudsdienst met succes conditievoorspelmodellen die met hoge betrouwbaarheid de toekomstige conditie van de machines kunnen voorspellen. Toch zijn de modellen niet voldoende om het onderhoud beter te plannen. Ze leveren alleen een verwachte conditie van de machines op en geven geen antwoord op de vraag of en welk onderhoud zou moeten worden uitgevoerd. Dat is juist de vraag waar de operations manager in geïnteresseerd is. Ondanks dat het inzicht in de (toekomstige) conditie van de machines is gegroeid, wordt de operations manager dus slechts beperkt geholpen in het nemen van onderhoudsbesluiten. De waarde van de voorspellingen is daarom slechts beperkt, er is meer nodig dan een voorspelling.

Onderhoudsbeslissingen zijn complex en gaan verder dan op basis van de verwachte conditie vaststellen welk onderhoud moet worden ingepland. Ter illustratie, als een machine uit productie wordt genomen moet worden bepaald hoe om te gaan met het verlies aan productiecapaciteit. Kan *fulfilment* van de klantvraag worden uitgesteld zonder schade aan de klantrelatie? Of moet een alternatief worden gezocht, bijvoorbeeld door elders capaciteit in te kopen of door de overige machines extra in te zetten? Wellicht kunnen producten op voorraad worden geproduceerd zodat de klant niets merkt van de beperkte productiecapaciteit, echter kan de extra voorraad dan wel worden opgeslagen in het magazijn of is extra externe opslagcapaciteit nodig? Hoeveel extra manuren zijn nodig voor de extra productie op de overige machines? Wegen de extra kosten op tegen de marge op de verkochte producten? Is onderhoud van de machine wel de beste beslissing of is vervangen van de machine misschien een betere optie?

De bovenstaande vragen zijn slechts een subset van de mogelijke vragen van de operations manager. Deze vragen kunnen niet beantwoord worden met een conditievoorspelmodel maar vragen om een beslissingsondersteunend model waarin onderhoud en inzet van machines en daarmee de bediening van klanten integraal kan worden afgewogen. Een dergelijk beslissingsondersteunend model levert de operations manager expliciete adviezen over hoe de machines in te zetten en wanneer en hoe onderhoud uit te voeren. Het biedt bovendien de mogelijkheid de invloed van aannames op de besluitvorming alsook de robuustheid van de planning voor veranderende omstandigheden, zoals vraagfluctuaties en transport- of opslagkostenwijzigingen, te toetsen.

Duidelijk is dat de waarde van beslissingsondersteunende modellen in het bovenstaande voorbeeld groter zal zijn dan die van de voorspelmodellen. Hoe komt het dan dat veel organisaties nog niet voorbijgaan aan voorspellen en de stap naar model gedreven besluitvorming maken? Het is mijn ervaring dat bij veel bedrijven simpelweg de kennis ontbreekt. Ook wordt de ontwikkeling van beslissingsondersteunende modellen (*aka prescriptive analytics*) als complex gezien of wordt ingeschat dat het niet tot een goede *return on investment* zal leiden. Die inschatting is mijn optiek onterecht, wellicht maakt onbekend onbemind? De INFORMS Edelman-competitie[5] laat ieder jaar weer zien dat verbluffende resultaten kunnen worden geboekt met beslissingsondersteunende modellen. Diezelfde competitie laat ook zien dat het niet vanzelf gaat om goede modellen te ontwikkelen en in te bedden in besluitvorming. Echter organisaties die zich verdiepen in de mogelijkheden en bereid zijn te investeren in het verbreden van hun *analytics* competenties kunnen net als de deelnemers aan de Edelman-competitie voorbijgaan aan het voorspellen en waarde creëren met model gedreven besluitvorming.

Noten

1. Business uit Data, Mediafederatie https://mediafederatie.nl/agenda/business-uit-data/
2. The Bestseller code, http://www.archerjockers.com/
3. 53% Of Companies Are Adopting Big Data Analytics, Forbes Dec 24, 2017, https://www.forbes.com/sites/louiscolumbus/2017/12/24/53-of-companies-are-adopting-big-data-analytics/
4. Forecast Snapshot: Prescriptive Analytics Software, Worldwide, 2019, https://www.gartner.com/document/3899065
5. INFORMS Edelman Award https://www.informs.org/Recognizing-Excellence/INFORMS-Prizes/Franz-Edelman-Award

John Poppelaars is Practice Leader Advanced Analytics bij BearingPoint, Amsterdam.
E-mail: john.poppelaars@bearingpoint.com

De sleutel tot succes

HET SIMULEREN VAN FORMULE 1 RACESTRATEGIEËN

Dat het kiezen van de juiste racestrategie in de Formule 1 een cruciale rol kan spelen, werd duidelijk tijdens het wereldkampioenschap in 2010. Met wel vier coureurs in de strijd om de titel was het seizoen ongekend spannend. Fernando Alonso had de beste papieren, maar liep door een strategische fout van zijn team de wereldtitel mis. Ferrari dacht namelijk dat een vroege pitstop de sleutel tot de winst zou zijn, maar achteraf bleek dat Alonso beter langer door had kunnen rijden. Sebastian Vettel, die wel langer doorreed, won uiteindelijk de race en staat nu te boek als de jongste wereldkampioen ooit in de Formule 1.

Claudia Sulsters

Om een fout in de racestrategie en de daarbij horende gevolgen zoveel mogelijk te voorkomen, gebruiken Formule 1-teams simulatiesoftware om de race vooraf te simuleren met verschillende variabelen. Voor deze simulatiemodellen worden grote hoeveelheden data gebruikt die gegenereerd worden door de sensoren aan boord van de Formule 1-auto's. Dit artikel* beschrijft een simulatiemodel op basis van *open source data*, dat kan worden gebruikt om de optimale racestrategie voor een Formule 1-team te bepalen. Het simulatiemodel beschrijft onder andere de invloed van het brandstofverbruik en de bandenslijtage op de rondetijden. Daarnaast imiteert het de meeste gebeurtenissen op het circuit, zoals de chaos na de start, pitstops, inhaalacties, en het uitvallen van coureurs door een crash of een technisch probleem.

Ontwerp van het simulatiemodel

Elke simulatie begint met het bepalen van de coureurs die de race niet zullen finishen. Zij raken bijvoorbeeld betrokken bij een crash of krijgen een technisch probleem. Voor elke coureur kiezen we willekeurig een racestrategie. Vervolgens verzamelen alle coureurs zich op de *starting grid*, de lichten gaan uit en de race gaat van start. Vlak na de start is het vaak chaos: alle coureurs proberen één of meer plaatsen te winnen en crashes komen vaak voor op dit moment in de race. De volgende stap is het simuleren van een rondetijd voor elke coureur en elke ronde van de race. Deze rondetijd wordt gesimuleerd met behulp van de kwalificatietijd en parameters die de invloed van het brandstofverbruik en de bandenslijtage op de rondetijd beschrijven. Inhaalacties worden toegevoegd door te kijken naar het verschil in (cumulatieve) rondetijd tussen opeenvolgende coureurs. Net als bij een echte Formule 1-race is een inhaalactie niet altijd succesvol. Dit hangt namelijk af van het verschil in rondetijd, waar de inhaalactie wordt ingezet (in een bocht of op een recht stuk) en hoe agressief de andere coureur verdedigt. Aan het einde van elke ronde simuleren we eventuele pitstops. Het simuleren van rondetijden herhaalt zich totdat de zwart-wit geblokte vlag wordt gezwaaid en we weten welke coureur als eerste over de streep is gekomen.

Did Not Finish

Het komt bijna nooit voor dat alle coureurs de finish halen. Het uitvallen van coureurs wordt gesimuleerd met behulp van een *Did Not Finish*-kans (DNF-kans). Deze kans wordt per coureur bepaald aan de hand van data over het wel of niet finishen in vorige races van dat seizoen. We gebruiken hiervoor de Bayesiaanse methode. Het aantal keer dat een coureur is uitgevallen in n races kan gemodelleerd worden als een binomiale verdeling met parameters n en p, met p de DNF-kans. Voor elke race zijn er namelijk twee mogelijke uitkomsten: wel of niet finishen. Het doel van de Bayesiaanse methode is om de priorverdeling van de DNF-kans te updaten aan de hand van nieuwe data. We kiezen de bètaverdeling als prior omdat deze verdeling twee belangrijke eigenschappen heeft: ten eerste is de bètaverdeling gedefinieerd op het interval (0,1), net als onze DNF-kans, en ten tweede is de bètaverdeling de zogeheten *conjugate prior* van de *Bernoulli likelihood*. Dit laatste betekent dat de posterior-

verdeling ook een bètaverdeling is. De parameters van de priorverdeling schatten we op basis van alle data van alle coureurs. Voor een individuele coureur updaten we deze prior vervolgens met data over het niet finishen van de betreffende coureur. Zo heeft Daniel Ricciardo, die relatief weinig uitviel, bijvoorbeeld een DNF-kans van 0,12 terwijl Daniil Kvyat, een coureur die veel vaker uitviel, een DNF-kans van 0,24 heeft.

Chaos bij de start

De start van een F1-race is vaak de beste kans voor een coureur om posities te winnen in het veld. Het simulatiemodel gebruikt de empirische verdeling van het aantal posities dat een coureur bij de start in vorige races heeft gewonnen of verloren om de chaos bij de start te simuleren. In deze empirische verdelingen wordt echter maar aan een beperkt aantal posities kansmassa toegekend omdat ze zijn gebaseerd op een kleine steekproef. Om deze reden zijn de empirische verdelingen niet geschikt om een aantal gewonnen of verloren posities uit te simuleren voor het simulatiemodel. De oplossing is om een mix van normaalverdelingen te gebruiken, waarbij de verwachting van elk van deze normaalverdelingen een positie met kansmassa is. Op deze manier ontstaat een *smooth* empirische verdeling waarmee we wel het aantal gewonnen of verloren posities kunnen simuleren.

Rondetijden

Nu bekend is welke coureurs ongeschonden de eerste bocht zijn doorgekomen, simuleren we een rondetijd voor elke coureur met behulp van een regressiemodel. Dit model beschrijft de invloed van het brandstofverbruik en de bandenslijtage op de rondetijd. Het is namelijk bekend dat de rondetijden dalen naarmate de race vordert, omdat Formule 1-auto's lichter worden door het brandstofverbruik. Daarnaast heeft de slijtage van de banden invloed op de rondetijd. Wanneer een coureur net is gewisseld naar nieuwe banden nemen de rondetijden iets toe, omdat de banden eerst moeten worden opgewarmd. Als de banden de optimale temperatuur hebben bereikt, kan de coureur een paar snelle rondes rijden. Wanneer de coureur echter te lang doorrijdt op de banden nemen de rondetijden weer toe als gevolg van de bandenslijtage. We nemen dus aan dat er een lineair verband bestaat tussen de rondetijd en het brandstofverbruik en een kwadratisch (convex) verband tussen de rondetijd en de bandenslijtage. Een rondetijd bestaat dan uit een basis rondetijd, een correctie voor het brandstofverbruik, een correctie voor de bandenslijtage en willekeurige fluctuaties:

(rondetijd) = (basisrondetijd) + (brandstofcorrectie) + (bandenslijtage) + (willekeurige fluctuaties).

Voor elke coureur schatten we vervolgens het effect van het brandstofverbruik en de bandenslijtage voor elke type band op de rondetijd met behulp van een regressiemodel. De parameters van het regressiemodel worden geschat met behulp van de kleinste kwadratenschatter. Als basisrondetijd gebruiken we de kwalificatietijd van de coureur. De geschatte parameters kunnen vervolgens worden gebruikt om nieuwe rondetijden te simuleren.

Inhalen

Tot nu toe hebben we interacties tussen coureurs, zoals inhaalacties, genegeerd. Inhaalacties zijn aan het simulatiemodel toegevoegd met behulp van het 'inhaalmodel'. Stel je twee coureurs voor die vlak achter elkaar rijden: als de achterste coureur sneller is dan de voorste coureur kan hij proberen in te halen. Of deze inhaalactie succesvol is hangt af van onder andere de relatieve snelheid tussen de coureurs, de afstand tussen de coureurs, waar op het circuit de inhaalactie wordt ingezet en hoe agressief de voorste coureur zijn positie verdedigt. Voor alle opeenvolgende paren van coureurs op het circuit berekenen we het verschil in cumulatieve rondetijd als $\delta_{jk} = C_k - C_j$, waarbij coureur j voor coureur k ligt. Als dit verschil negatief is en kleiner dan een bepaalde 'inhaaldrempel', dan betekent dit dat coureur k snel genoeg is om coureur j in te halen. We simuleren aan de hand van de 'inhaalkans' of de inhaalactie succesvol is of niet. Als de inhaalactie succesvol is, wisselen de coureurs van positie. Als de inhaalactie niet succesvol is, dan heeft de coureur de volgende ronde mogelijk opnieuw een kans om in te halen.

Het vergelijken van pitstopstrategieën

We kunnen het simulatiemodel gebruiken om de optimale pitstopstrategie voor Mercedes te bepalen tijdens de Grand Prix van Japan in 2016. Mercedes reed destijds met het rijdersduo Lewis Hamilton en Nico Rosberg. Volgens de reglementen van Pirelli, de bandenleverancier van de Formule 1, waren er drie soorten banden beschikbaar: de zachte band, de medium band en de harde band. Daarnaast was de top 10 volgens de reglementen verplicht om de race op de zacht band te starten en moesten alle coureurs verplicht tenminste één pitstop te maken en de harde band gebruiken. Deze restricties zorgden ervoor dat niet alle combinaties van bandensoorten mogelijk waren. Ten slotte verwachtte Pirelli dat een twee-stop strategie optimaal zou zijn, dus nemen we aan dat elke coureur twee keer een pitstop maakt.

De verzameling van mogelijke bandenstrategieën voor Lewis Hamilton en Nico Rosberg bestaat dus uit:
- Zacht – medium – hard
- Zacht – hard – medium
- Zacht – hard – hard

Daarnaast gebruiken we de gemiddelde duur van een pitstop van Mercedes tijdens de Grand Prix van Japan in 2015 als pitstopduur.

Op basis van data weten we wat de werkelijke pitstopstrategie was van Lewis Hamilton en Nico Rosberg tijdens de Grand Prix van Japan in 2016. Lewis Hamilton wisselde zijn banden in rondje 13 van de zachte band naar de harde band en in ronde 33 opnieuw naar de harde band. Nico Rosberg gebruikte dezelfde bandenstrategie, maar zijn pitstops vonden plaats in ronde 12 en ronde 29. We gebruiken het simulatiemodel om drie alternatieve strategieën te simuleren en te vergelijken met de huidige strategie:
- Strategie 1: dezelfde bandenstrategie, maar de pitstops vinden eerder plaats, namelijk in ronde 10 en ronde 28.
- Strategie 2: bandenstrategie zacht – hard – zacht met pitstops in ronde 12 en ronde 29.
- Strategie 3: bandenstrategie zacht – hard – medium met pitstops in ronde 12 en ronde 29.

We simuleren elke combinatie van strategieën 1000 keer en berekenen het gemiddelde van de som van de posities van Nico Rosberg en Lewis Hamilton. De combinatie van strategieën die resulteert in het laagste gemiddelde is de optimale strategie voor Mercedes. Volgens tabel 1 is dat de werkelijke strategie voor Lewis Hamilton in combinatie met strategie 2 voor Nico Rosberg. Met behulp van statistische testen kunnen we bepalen dat deze combinatie van strategieën significant beter is dan de werkelijke strategie. De conclusie is dus dat Mercedes de juiste strategie had gekozen voor Lewis Hamilton, maar voor Nico Rosberg beter strategie 2 had kunnen kiezen.

* Dit artikel is gebaseerd op het artikel 'Simulating Formula One Racestrategies', winnaar van de Business Analytics Best Paper Award 2018. Het volledige artikel is te vinden op: https://beta.vu.nl/nl/Images/werkstuk-sulsters_tcm235-877826.pdf.

Claudia Sulsters is recentelijk afgestudeerd van de masteropleiding Business Analytics aan de Vrije Universiteit en werkt momenteel als data scientist bij Data Science Lab. in Amsterdam.
Email: claudia.maria@live.nl
LinkedIn: https://www.linkedin.com/in/claudiasulsters/

STRATEGIE		GEMIDDELDE VAN DE SOM VAN DE POSITIES	P-VALUE T.O.V WERKELIJKE STRATEGIE
Lewis Hamilton	*Nico Rosberg*		
Werkelijk	Werkelijk	5,02	-
Werkelijk	Strategie 2	4,59	$1{,}003 \times 10^{-9}$
Werkelijke	Strategie 3	4,91	$2{,}445 \times 10^{-5}$

Tabel 1. Het vergelijken van racestrategieën

STRAVA-BILLEN EN BOLLETJESTRUIEN

In de zomer kijk ik voor mijn doen tamelijk veel tv. Voetbal ook wel, maar dat duurt me toch te vaak te lang. Wielrennen? Nog saaier. Drie weken hard fietsen door Frankrijk? Geen bal aan. Vooral de vlakke etappes zijn pas echt *boring*; die bekijk ik alleen in-samenvatting. Een kopgroep van acht man krijgt de ganse dag 12 minuten voorsprong of zo. Met het algoritme van een paar gesjeesde econometristen wordt vervolgens berekend dat het peloton kop-over-kop met 63 kilometer per uur die kopgroep 100 meter voor de meet inrekent en dat de sprinter met de dikste billen de etappe wint. Althans, dat moet er wel bijgezegd worden, mits hij onderweg heeft weten te ontsnappen uit een kluwen van kromme wielen en gebutste lijven. Als de winnaar allang binnen is, zoemt de camera in op zijn gehavende shirt en broek en het langst op zijn bebloede verrafelde bil.

Maar de bergen dan? Eigenlijk ook niks. Eerst erop en dan eraf, want voor andersom heb je twee bergen nodig en ik zie ze nooit aan de voet ervan omdraaien en dezelfde weg terugknarren, in welk geval trouwens één berg zou volstaan. En voor de doordenkende lezer, ook geldt meestal $a+b \neq b+a$, vanwege de asymmetrische vorm der toppen, terwijl $a+b = b+a$ *iff* de hellingen aan beider zijden der top hetzelfde zijn.

Maar allee, de Mont Ventoux, de Tourmalet, de Alpe d'Huez en nog zo'n paar, zijn toch van de buitencategorie! Op de hellingen van die reuzen zoek ik inderdaad het 'puntje' van mijn stoel. Omhoog gaat het met de snelheid van een rennende ontblote dikbuik en naar beneden met naad-op-stang en kop-in-stuur. Borst bloot omhoog en extra-shirtje-aan omlaag. Omhoog worden de verschillen gemaakt, omlaag vallen de klappen en soms doden. Topsport en sensatie, ga er maar voor zitten.

Ik hou van wielrennen. Voor mij de pure sport. Zelf bezit ik zo'n supersonische carbonfiets met Strava op mijn smartding. Strava? Ook de toerfietser is in de greep van het gezellige gemedia van prestaties en getallen. Een ritje door de polder volbracht, betekent *kudo*'s van je *volgers* en bij de bedwinging van een bergreus in een nieuw pr word je zelfs met loftuitingen beloond. Tijden, hoogtemeters en hartslagen, alles staat op Strava, Daar doe je het voor. Geen kudo of loftuit op Strava ontvangen: rit niet gefietst. Pure sport is ook genadeloos.

Strava herbergt trouwens nog smeriger valkuilen. Ikzelf heb dat aan den bille ondervonden. Zo meldde Strava enige tijd alweer geleden dat één mijn allerbeste fietskameraden op het Strava-segment in mijn 'achtertuin', de Onlanden bij Groningen, nummer één stond in het lijstje (sic!) van mijn volgers. Ik stond slechts tweede in mijn eigen 'achtertuin'! Toen de zuidenwind blies ben ik rustig tegen de straffe wind in naar de start van 'mijn' segment gefietst. Starten, gassen, op volle snelheid met kop bijna op Triathlonstuur. Drie-en-zeventig was ik. In de laatste bocht naar links ging het mis. Gruis op het rafelig asfalt. Van knie tot schouder, bil ontveld (de linker dus) en elleboogbot in 't zicht. De huidrafels heeft de dienstdoende arts van de dokterspost in Leek bijgeknipt en na het hechtwerk mocht ik de volgende dag in het vliegtuig bij de nooduitgang zitten op weg naar het congres over *computer science* en sport. Twee-en-veertig km per uur of zo reed ik; staat op Strava, nog steeds.

Kan het gekker? Ja dat kan. De directie van de Ronde van Italië kwam in 2017 met een bizar idee: een klassement voor de snelste dalers (*Trouw*, 2017). Het idee werd meteen weggehoond. Snel dalen stimuleren? Idioterie. Een beter idee werd toen gek genoeg niet geopperd door de wielrenopperhoofden, namelijk een speciaal klassement voor snelste klimmers.

Natuurlijk, een klassement voor klimmers bestaat al, sinds 1933 zelfs, en de aanvoerder ervan mag zich sinds 1975 hullen in de 'bolletjestrui'. Die bolletjes zijn trouwens niet bol, maar dat snapt u wel. Het is een begerenswaardig tricot, wit met een boel rood ingekleurde cirkeltjes met een diameter van 5 cm of zo, en goed voor veel publiciteit bovendien. Voor dat bergklassement krijgen renners punten. Een buitencategorieberg levert 25, 20, 16, 14, 12, 10, 8, 6, 4 en 2 punten voor de eerste tot en met de tiende renner die de top passeert. Het onderlinge tijdsverschil speelt daarbij geen enkele rol, de daadwerkelijke klimtijd ook niet. Voor tien minuten eerder boven krijg je dezelfde beloning als met tien seconden. Alleen de volgorde van de passage op de col telt dus.

En dat is toch wel gek.

De winnaar van de bolletjestrui is lang niet altijd de beste klimmer van die ronde. Neem Laurent Jalabert, die de trui in 2001 en 2002 won – hij begon ooit als sprinter. Of neem Richard Virenque die de trui zeven keer won in de Tour de France, een record. Maar is Virenque daarmee de beste Tour de France-klimmer aller tijden? Nou nee, zou ik zeggen. Jalabert en Virenque wonnen hun truien vooral door vroeg en vaak mee te gaan in lange ontsnappingen in etappes met de bergen aan het einde ervan. Zo sprokkelden ze hun punten bij elkaar door met grote voorsprong aan de voet van de berg te beginnen. Zo deden Anthony Charteau in 2010 en Thomas Voeckler in 2012 het ook. De bolletjestrui gaat dan dus niet naar de beste klimmer, maar naar de slimste. Daar heb ik zeker respect voor en dat ooit voor die opzet is gekozen snap ik ook. Het was vroeger nu eenmaal niet mogelijk om van alle renners op elke berg de klimtijden bij te houden. Maar in het huidige Strava-tijdsgewricht is dat geen probleem meer.

Waarom die technieken dan niet gebruiken en de klimmer die echt het snelst is belonen? Via internet en tv kunnen de beelden en de statistieken van klimtijden *real time* worden getoond. Zo krijgen we spannende wedstrijden binnen de grote wedstrijd met voor mij als enig nadeel het sneller slijten van de stoelpuntbekleding.

Kortom, ik pleit voor een trui met gele bolletjes voor de snelste klimmer. Die trui kan de huidige bolletjestrui vervangen. Deze suggestie zou ik willen doen aan het adres van Christian Prudhomme, de directeur van de Société du Tour de France. Maar wanneer hij erg aan die huidige trui hecht, kunnen ze ook naast elkaar bestaan: een 'gewone' bolletjestrui voor de puntensprokkelaars op de bergtoppen en een gele bolletjestrui voor de man die echt het snelste klimt. Prudhomme zou de wielerliefhebbers onder ons een groot plezier doen. En wat Strava in mijn achtertuin betreft, ik raad je aan een slijtvaste (tegenwoordig te koop) en zeker geen grote broek aan te trekken.

P.S. Een opmerkzame lezeres van een eerdere versie van deze column attendeerde mij op het feit dat de letters a en b hierboven niet gedefinieerd (sommigen noemen het gedeclareerd) zijn. Deze lezeres heeft volkomen gelijk. Toch heb ik die definities weggelaten. Mocht u daar problemen mee hebben, dan zit ik daar even niet mee: ik kijk tv. En nog iets, de lezeres wees mij ook op het vrouwenfietsen, dat dat ook bestaat en dat ook daar valpartijen voorkomen met op dezelfde plekken als bij de mannen opengereten kleding. Oh.

Literatuur

Heuvelman, D., Schoonderwalt, F. van, & Sierksma, G. (2006). *Tour de France Top 100; De honderd beste tourrenners aller tijden.* Z.pl.: Tirion Sport.

Sierksma, G. (2017). Waarom geen gele bolletjestrui voor de echt snelste klimmer? *Trouw*, 15 juli 2017.

Gerard Sierksma is emeritus hoogleraar Kwantitatieve Logistiek en Sportstatistiek aan de Rijksuniversiteit Groningen.
E-mail: g.sierksma@rug.nl

MASTERMIND VOOR BLOODMATCHING

Mastermind was en is een leuk denkspel uit de jaren 80 van de vorige eeuw. Anno 2019 is het state of the art en bloedserieus binnen het vakgebied van de transfusiegeneeskunde. Dit artikel beschrijft hoe met OR (Dynamisch Programmeren) en selectieve matching de kans op het ontstaan van antistoffen bij bloedtransfusies kan worden geminimaliseerd. Tevens wordt een optimale volgorde van antigenen bepaald waarin op geschikt bloed dient te worden gezocht.

Joost van Sambeeck, Ellen van der Schoot, Nico van Dijk, Henk Schonewille & Mart Janssen

Een bloedtransfusie is een veilige, veel voorkomende en soms levensreddende medische behandeling, waarbij rode bloedcellen afkomstig van een donor worden ingebracht in de bloedbaan van een transfusieontvanger die deze nodig heeft vanwege bijvoorbeeld een chirurgische ingreep (orgaantransplantatie, hartchirurgie) of een hematologische aandoening (leukemie, sikkelcelanemie, thalassemie). Hierbij is het van belang dat de *bloedgroepen* van de donor en transfusieontvanger zoveel mogelijk overeenkomen, of anders gezegd, dat hun bloedgroepen *matchen*.

De bloedgroep van een individu wordt bepaald door de aan- of afwezigheid van *antigenen* op de celwand van rode bloedcellen. Indien een bepaald antigeen aanwezig is op de rode bloedcellen van een donor, maar afwezig op de rode bloedcellen van de transfusieontvanger, dan kan het immuunsysteem van de transfusieontvanger overgaan tot het aanmaken van *antistoffen* tegen dit lichaamsvreemde antigeen. Deze antistoffen kunnen vervolgens problemen veroorzaken bij een volgende bloedtransfusie of, in het geval van een vrouwelijke ontvanger, een (toekomstige) zwangerschap. Een reactie van het immuunsysteem kan echter voorkómen worden door uit de beschikbare voorraad –indien mogelijk– rode bloedcellen te selecteren waarbij het desbetreffende antigeen ook ontbreekt.

Over het algemeen is de ABO-bloedgroep classificatie, evenals de Rhesus-D factor welbekend, welke een drietal antigenen betreft (A, B, D). Binnen de transfusiegeneeskunde onderscheidt men echter meer dan 15 verschillende, klinisch relevante, bloedgroepantigenen. Idealiter zou men bij een bloedtransfusie met al deze antigenen rekening willen houden. Indien er onbeperkte hoeveelheden rode bloedceleenheden voorradig zouden zijn, is dat in principe geen probleem. In de praktijk zijn voorraden echter eindig (variërend van 60 tot 250 eenheden voor ziekenhuizen en ruwweg 1000 eenheden voor een regionaal distributiecentrum). Daarom moet er een selectie worden gemaakt van antigenen waar *wel* op wordt gematcht en antigenen waar *niet* op wordt gematcht. Daarbij mag verondersteld worden dat zowel bij donors als ontvangers een volledige typering van de antigenen heeft plaatsgevonden. Een tweetal onzekerheidsaspecten spelen een rol:

- de verdeling van bloedgroepen in de populatie;
- het vermogen van een antigeen om antistofvorming op te wekken, indien er niet op dit antigeen gematcht wordt (de *immunogeniciteit* van een antigeen).

Uitgiftestrategieën

Figuur 1 laat een situatie zien, waarbij we een transfusieontvanger op 15 antigenen proberen te matchen. Hierbij is het van belang dat

- alle antigenen die *afwezig* (wit) zijn op de rode bloedcellen van de transfusieontvanger ook *afwezig* zijn op de rode bloedcellen van de donor,
- alle antigenen die *aanwezig* (oranje) zijn op de rode bloedcellen van de transfusieontvanger ook *aanwezig* zijn op de rode bloedcellen van de donor,

een zogeheten antigeen identieke matchingstrategie. In dit voorbeeld is een identieke match op alle 15 antigenen niet mogelijk. Daarom zullen er antigenen geschrapt moeten worden (van rechts naar links), totdat er een identieke match is gevonden.

In principe is het matchingsprobleem bij een gespecificeerde vraag en een gespecificeerde voorraad eenvoudig. Echter is de volgorde waarin antigenen worden geschrapt vooralsnog willekeurig gekozen. Bovendien zijn zowel de bloedgroepen van de gevraagde eenheden als de bloedgroepen van de beschikbare eenheden stochastisch. Een eerste vraag die gesteld kan worden, is:

1. Hoe kan de kans op het voorkómen van antistofvorming zo groot mogelijk worden gemaakt?

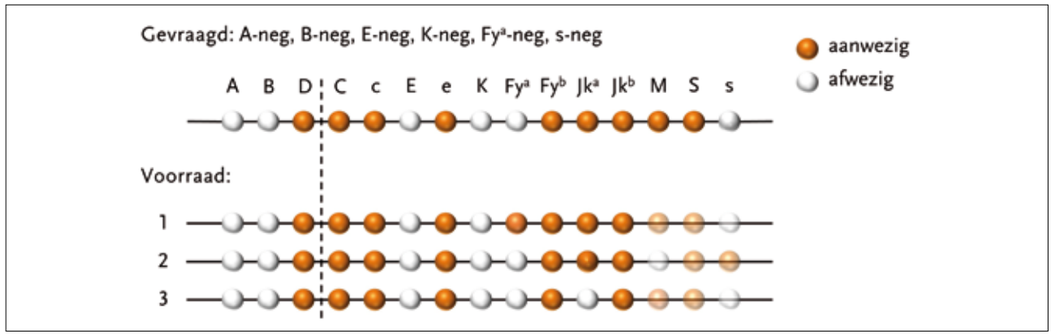
Dit artikel beoogt bovenstaande vraag te beantwoorden. Bekend is dat wanneer transfusieontvangers uitsluitend op antigenen A, B en D worden gematcht, 88% van de antistoffen gericht zijn tegen antigenen C, c, E, e, K, Fy^a, Fy^b, Jk^a, Jk^b, M, S, s (Evers et al., 2016). Uitgaande van deze set van 15 antigenen bedraagt het aantal mogelijk te matchen bloedgroepen ruim 32.000 (= 2^{15}). In een voorraad van 60 tot 1000 eenheden kunnen deze bloedgroepen niet allemaal aanwezig zijn. Er zal dus een aantal antigenen geschrapt moeten worden met als doel de kans op antistofvorming zo klein mogelijk te houden. Dit schrappen roept een tweede vraag op, te weten:

2. In welke volgorde moeten antigenen worden geschrapt?

Voor het 12-tal eerdergenoemde antigenen (naast A, B en D) zijn er al 12!, ongeveer een half miljard, mogelijke volgordes denkbaar. Dit is waar OR uitkomst kan bieden!

Dynamisch programmeren

Beide vraagstellingen kunnen worden geabstraheerd als een langste pad probleem startend in {A, B, D} en eindi-



Figuur 1. Voorbeeld van een antigeen identieke uitgiftestrategie. Antigenen worden geschrapt (van rechts naar links), totdat er een antigeen identieke match is gevonden (eenheid 2)

gend met de genoemde 15 antigenen, zoals weergegeven in figuur 2. Een matchingsstrategie met $l+3$ antigenen is gedefinieerd als een verzameling

$$M = \{A, B, D\} \cup M_l,$$

met $M_l = \{ \mu_1, ..., \mu_l \}$. De maximale proportie antistoffen die voorkómen kan worden door op antigenen A, B, D, en $\mu_1, ..., \mu_l$ te matchen (wiskundig: $V(M)$), wordt iteratief met de volgende DP-vergelijking berekend:

$$\underbrace{V(M)}_{l+3 \text{ antigenen}} = \max_{i=1,...,l} \{ \underbrace{r(M \setminus \mu_i \to M)}_{1 \text{ antigeen}} + \underbrace{V(M \setminus \mu_i)}_{l-1+3 \text{ antigenen}} \}$$

waarbij $r(M \setminus \mu_i \to M)$ de één-staps opbrengst weergeeft en $V(\{A, B, D\}) = 0$. Deze één-staps opbrengst berust op de verdeling van bloedgroepen in de populatie, zoals cryptisch weergegeven door de formule:

$$r(M / \mu_i \to M) = p(\mu_i) \sum_{\varphi} \underbrace{w(\varphi,\mu_i)}_{\substack{\text{ontvanger} \\ \text{bloedgroep } \varphi}} \cdot \underbrace{L(\varphi, M, n, k)}_{\substack{\text{voorraad (donor)} \\ \text{bloedgroep } \varphi}},$$

met

- $p(\mu_i)$: de proportie antistoffen gericht tegen antigeen μ_i,
- $w(\varphi, \mu_i)$: de kans dat een transfusieontvanger bloedgroep φ heeft, gegeven dat het antigeen μ_i niet aanwezig is op zijn/haar rode bloedcellen,
- $L(\varphi, M, n, k)$: de kans dat er k eenheden in voorraad zijn met bloedgroep φ, als matchingstrategie M wordt toegepast en de voorraadomvang gelijk is aan n.

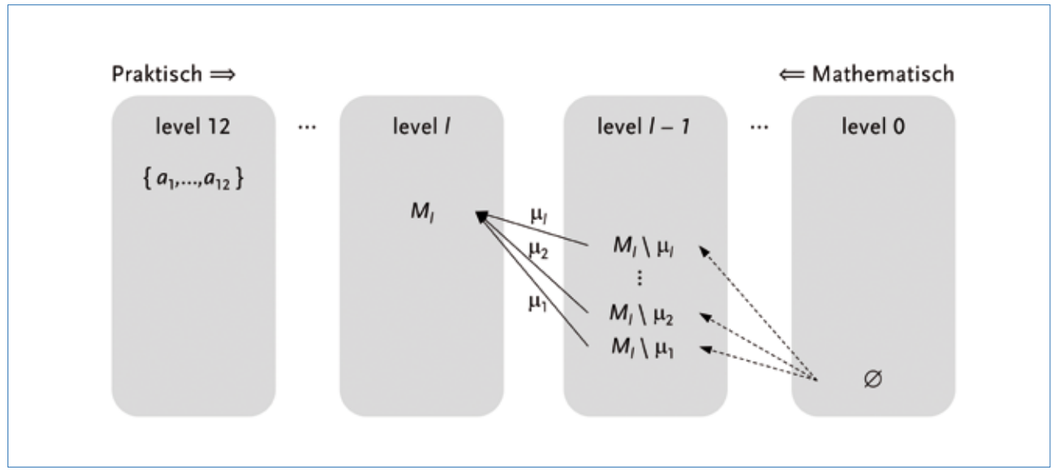
Merk op dat bovenstaande DP-vergelijking niet simpelweg neerkomt op het toevoegen van één enkel antigeen aan een eerdere optimale matchingsstrategie (met $l-1+3$ antigenen). De volgorde waarin antigenen worden opgenomen (een pad) kan namelijk compleet veranderen.

Er zijn ruwweg een half miljard paden om van level 0 naar level 12 te komen. De tijd die het zou kosten om de lengte van elk pad apart te berekenen (en vervolgens het langste pad te selecteren) zou, naar schatting, neerkomen op 17 dagen. Met behulp van dynamisch programmeren wordt exact hetzelfde pad in slechts één seconde gevonden. De berekening van tabel 1 reduceert hiermee van bijna één jaar tot minder dan één minuut.

Resultaten

Tabel 1 geeft de resultaten weer van een uitgiftestrategie die de huidige transfusiepraktijk het best representeert. Voor een klein ziekenhuis ($n = 60$) kan 46% van de antistoffen voorkómen worden, wanneer een transfusieontvanger twee rode bloedceleenheden krijgt toegediend. Dit percentage stijgt naarmate er meer eenheden in voorraad zijn en daalt naarmate er meer eenheden gevraagd worden. Indien vooraf al een vermoeden bestaat dat een bloedtransfusie nodig zal zijn, kunnen ziekenhuizen de juiste rode bloedceleenheden het best bij een distributiecentrum bestellen. Immers; hoe groter de voorraad, des te meer antigenen gematcht kunnen worden.

Tevens is een optimale volgorde bepaald, waarin antigenen geschrapt dienen te worden, totdat een match gevonden wordt. Hoewel deze volgorde afhankelijk is van de voorraadomvang (n) en het aantal gevraagde eenheden (k), blijken ze toch sterk overeen te komen (tabel 2).



Figuur 2. Transitiediagram: 4096 toestanden (matchingsstrategieën) en 24.576 lijnen (beslissingen). Het representeert en ontbindt alle mogelijke volgordes waarin twaalf antigenen kunnen worden opgenomen (Mathematisch) of geschrapt (Praktisch). Tevens illustreert het hoe de maximale proportie antistoffen die voorkómen kan worden voor een matchingsstrategie $M = \{A, B, D\} \cup M_l$ wordt berekend

n	$k = 1$	$k = 2$	$k = 3$	$k = 5$	$k = 10$
60	0,65	0,46	0,32	0,13	0,00
120	0,78	0,63	0,53	0,38	0,11
250	0,88	0,78	0,70	0,59	0,38
1000	0,97	0,94	0,91	0,85	0,74

Tabel 1. Maximale proportie antistoffen die voorkómen kan worden door op 15 antigenen te matchen, waarbij n de voorraadomvang en k het aantal gevraagde eenheden weergeeft

Conclusie

Technologische ontwikkeling in de transfusiegeneeskunde gaat gepaard met nieuwe logistieke uitdagingen. Een voorbeeld hiervan is in dit artikel besproken: het optimaal matchen van bloeddonors en transfusieontvangers. Dit toont de toepasbaarheid van klassieke en moderne OR.

Literatuur

Evers, D., Middelburg, R.A., de Haas, M., Zalpuri, S., de Vooght, K.M.K., van de Kerkhof, D., Visser, O., Péquériaux, N.C., Hudig, F., Schonewille, H., Zwaginga, J.J., & van der Bom, J.G. (2016). Red-blood-cell alloimmunisation in relation to antigens' exposure and their immunogenicity: a cohort study. *Lancet Haematology*, *3*(6), e284–e292.

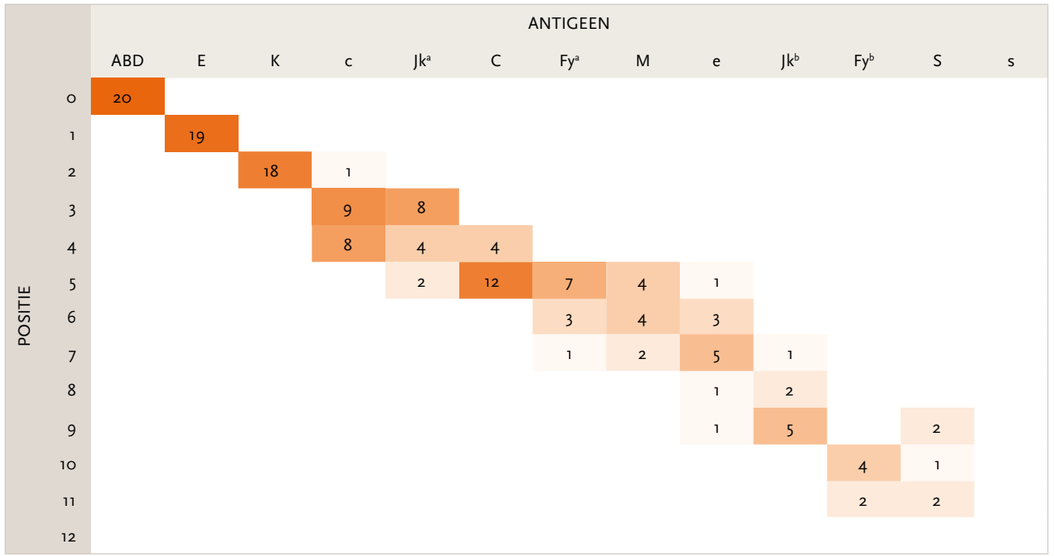
Joost van Sambeeck is promovendus bij Sanquin Research en verbonden aan de onderzoeksgroepen SOR (Stochastic Operations Research) en CHOIR (Center for Healthcare Operations Improvement & Research) van de Universiteit Twente.
E-mail: j.vansambeeck@sanquin.nl

Ellen van der Schoot is hoofd van de afdeling Experimentele Immunohematologie bij Sanquin Research en hoogleraar aan de Universiteit van Amsterdam.
E-mail: e.vanderschoot@sanquin.nl

Nico van Dijk is verbonden aan de onderzoekgroepen SOR en CHOIR aan de Universiteit Twente.
E-mail: n.m.vandijk@utwente.nl

Henk Schonewille is onderzoeker bij Sanquin Research.
E-mail: h.schonewille@sanquin.nl

Mart Janssen is hoofd van de afdeling Transfusion Technology Assessment bij Sanquin Research.
E-mail: m.janssen@sanquin.nl



Tabel 2. Verdeling van de optimale positie per antigeen binnen een matchingsstrategie. De oranje-schaal correspondeert met de opgegeven frequentie

Het artikel van Paul Goedhart is een reactie op de heranalyse van Hanekamp et al. van een aantal WUR-experimenten om de ammoniakemissie van verschillende bemestingstechnieken te schatten. De discussie over de wetenschappelijke onderbouwing van het ammoniakbeleid is eerder gevoerd in het vakblad *Soil Use and Management*. Het Rathenau Instituut spreekt van een hardnekkige controverse. Het conflict heeft ook een politieke kant, doordat sommige veehouders zich beperkt voelen in hun ondernemerschap door milieuregels van de overheid. Goedharts artikel gaat met name over het berekenen van de onzekerheidsmarges en de zijns inziens incorrecte statistische analyse van Hanekamp et al. Om te voorkomen dat het een preek voor eigen parochie wordt heeft de *STAtOR*-redactie Hanekamp om een reactie gevraagd, zodat de lezer zelf een mening kan vormen.

HOE EEN STATISTISCHE FOUT HET AMMONIAKDEBAT OP EEN DWAALSPOOR ZET

Paul Goedhart

Ammoniak (NH_3) is slecht voor de natuur, dat is onbetwist. In Nederland wordt door de agrarische sector veel ammoniak uitgestoten vooral door het bemesten van landbouwgrond en door verliezen uit de stal. Sinds 1985 is er overheidsbeleid om deze uitstoot te reduceren. Zo is het verplicht om emissiearme bemestingstechnieken te gebruiken zoals bijvoorbeeld de zodenbemester in plaats van mest bovengronds uit te rijden. Alle overheidsmaatregelen tezamen hebben volgens het rekenmodel NEMA geleid tot een reductie van de ammoniakemissie met 64% (zie www.rivm.nl/ammoniak/emissies). Het meetmodel OPS van het Rijksinstituut voor Volksgezondheid en Milieu (RIVM), dat gebruik maakt van gemeten ammoniakconcentraties in de lucht, geeft echter aan dat de reductie kleiner is. Het verschil in de uitkomst van het rekenmodel NEMA en het meetmodel OPS wordt het ammoniakgat genoemd. Sommigen claimen op grond hiervan dat het overheidsbeleid niet werkt en dat het bijvoorbeeld onzin is om bovengronds uitrijden van mest te verbieden.

Het rapport *Ammoniak in Nederland* van chemicus Jaap Hanekamp (University College Roosevelt, Middelburg), wetenschapsjournalist Marcel Crok en de Amerikaanse wiskundige Matt Briggs wordt op 24 januari 2017 aangeboden aan de Tweede Kamerleden Elbert Dijkgraaf (SGP) en Jaco Geurts (CDA). In dit rapport (Hanekamp et al., 2017a) constateren zij: 'een opeenstapeling van rekenkundige, modelmatige, en argumentatieve tekortkomingen. Men kan om die redenen niet concluderen dat het ammoniakbeleid wetenschappelijk goed onderbouwd zou zijn, zoals de overheid beweert. Integendeel.' Naast kritiek op de door het RIVM uitgevoerde trendanalyse op gemeten ammoniakconcentraties in de lucht, zijn de auteurs zeer kritisch over het emissieonderzoek van Wageningen Universiteit en Research (WUR). Dit onderzoek richt zich mede op de vergelijking van verschillende bemestingstechnieken zoals bovengronds uitrijden en de zodenbemester. Uit het WUR-onderzoek blijkt dat emissiearme technieken de uitstoot van ammoniak flink verminderen. Hanekamp et al. heranalyseren acht WUR-experimenten om de ammoniakemissie van een specifieke techniek in te schatten. Op basis van hun heranalyse concluderen zij dat de onzekerheidsmarges van de geschatte emissies 'fors

Bovengronds uitrijden van mest

Zodenbemester aan het werk

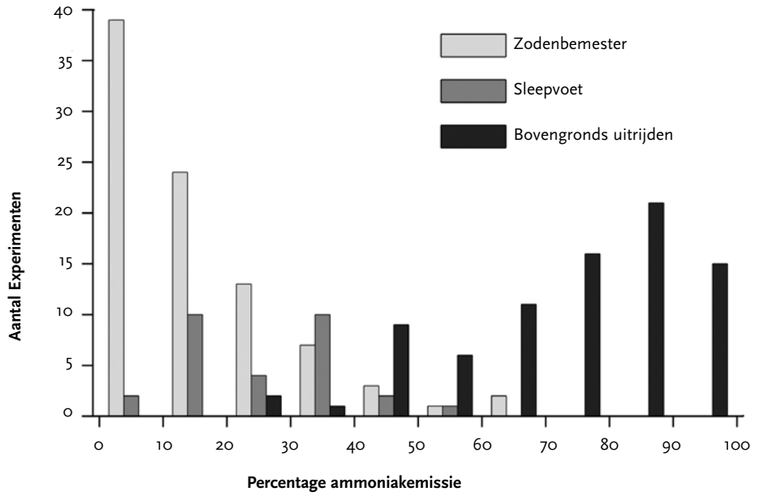
bleken' en tevens dat 'het niet onwaarschijnlijk (is) dat de verschillende bemestingstechnieken dichter bij elkaar liggen wat betreft ammoniakemissies dan tot op heden is gerapporteerd en toegepast in het beleid'.

Samen met de kritiek op de trendanalyse van het RIVM plaatsen ze grote vraagtekens bij het door de overheid gevoerde ammoniakbeleid. Nu zijn onzekerheden de *core business* van elke statisticus, en daarom ben ik nagegaan of de kritiek van Hanekamp et al. op het WUR-onderzoek hout snijdt.

Ammoniakconcentratie en windprofiel

Een veel gebruikte experimentele methode om de ammoniakemissie van een bemestingsmethode te bepalen is de zogenaamde *Integrated Horizontal Flux* methode. Aan het begin van een experiment wordt eenmalig mest uitgereden op een veelal cirkelvormige plot van ongeveer 40 m diameter. In het centrum van de plot wordt op diverse hoogtes zowel de ammoniakconcentratie als de windsnelheid gemeten (figuur 1). Deze metingen worden gebruikt om een ammoniakconcentratieprofiel $c(h)$ en een windprofiel $u(h)$ te schatten, beide als functie van de hoogte h. De ammoniakemissie wordt vervolgens gebaseerd op de hoeveelheid ammoniak die de wind wegblaast: $\int_{h_o}^{h_p} u(h)\, c(h)\, dh$. Het integratie-interval loopt van de hoogte h_o waarop de windsnelheid nul is, tot de hoogte h_p waarop de ammoniakconcentratie gelijk is aan de achtergrondconcentratie (die ook wordt gemeten). Na een correctie voor de achtergrondconcentratie en een normalisatie

Figuur 1. Links een meetmast met windsnelheidsmeters op verschillende hoogtes, en rechts een zogenaamde Impinger oftewel een flesje met zuur om ammoniak af te vangen



Figuur 2. Histogram van de geschatte percentages ammoniakemissie voor drie technieken om mest op landbouwgrond aan te brengen; de zodenbemester en de sleepvoet zijn emissie-arme technieken

wordt uiteindelijk een percentage emissie van de oorspronkelijke hoeveelheid ammonium in de mest verkregen. Deze procedure wordt gevolgd voor verschillende opeenvolgende perioden na toediening van de mest. De som van de berekende percentages over de opeenvolgende perioden is de zogenaamde emissiefactor. Deze geeft aan welk aandeel van de ammonium in de mest verloren gaat via vervluchtiging.

Het is gebruikelijk om voor het concentratie- en windprofiel een lineaire relatie te nemen met als verklarende variabele de logaritme van de hoogte h: $c(h) = -A\ln(h) + B$ en $u(h) = D\ln(h) + E$. De parameters A, B, D en E worden uit de metingen geschat met lineaire regressie. Met deze profielen kan de integraal expliciet berekend worden:

$$\int_{h_0}^{h_p} u(h)\,c(h)\,dh = [-AD(h\ln^2 h - 2h\ln h + 2h) + (BD - AE)(h\ln h - h) + BE\,h]_{h_0}^{h_p}$$

De integraal, en dus de emissie, is een functie $F(A,B,D,E)$ van de vier regressieparameters. Hanekamp et al. berekenen een 95%-betrouwbaarheidsinterval voor de uitkomst door het invullen van 95%-betrouwbaarheidsgrenzen voor de individuele parameters zelf in $F(A,B,D,E)$. Gebruikmakend van het superscript + voor een rechter betrouwbaarheidsgrens berekenen zij $F(A^+,B^+,D^+,E^+)$ voor de bovengrens in plaats van $F^+(A,B,D,E)$. Deze fout verergeren zij door het sommeren van de zo verkregen betrouwbaarheidsgrenzen over de opeenvolgende meetperioden. Oftewel, met bijvoorbeeld drie meetperioden, berekenen zij $(F_1^+ + F_2^+ + F_3^+)^+$ in plaats van het correcte $(F_1 + F_2 + F_3)^+$.

Het zo berekende interval wordt door hen gepresenteerd als een 95%-interval. Een basale fout.

Het volgende voorbeeld illustreert dat het optellen van betrouwbaarheidsgrenzen een veel hogere overdekkingskans geeft. Veronderstel dat we twee steekproefgemiddelden $\bar{x}_1$ en $\bar{x}_2$ hebben, beide gebaseerd op n waarnemingen, uit normale verdelingen met gemiddelden respectievelijk μ_1 en μ_2 en bekende variantie σ^2. De éénzijdige α-rechter betrouwbaarheidsgrens voor μ_i is dan $\bar{x}_i + z_\alpha\sigma/\sqrt{n}$, met $\Phi(z_\alpha) = \alpha$, en de som van de bovengrenzen is $\bar{x}_1 + \bar{x}_2 + 2\,z_\alpha\sigma/\sqrt{n}$. Dit is echter géén betrouwbaarheidsgrens voor $\mu_1 + \mu_2$. Immers, uit $\bar{x}_1 + \bar{x}_2 \sim N(\mu_1 + \mu_2, 2\,\sigma^2/n)$ volgt onmiddellijk dat de correcte, en lagere, bovengrens gelijk is aan $\bar{x}_1 + \bar{x}_2 + \sqrt{2}z_\alpha\sigma/\sqrt{n}$, nu met een factor $\sqrt{2}$ in plaats van 2. De consequentie is dat de intervallen van Hanekamp et al. een veel grotere overdekkingskans hebben dan de door hen geclaimde 95%. De intervallen zijn veel te breed en dit verklaart de door hen geclaimde forse onzekerheden. In bovenstaand voorbeeld heeft het foutieve interval met $\alpha = 95\%$ overigens een overdekking van 99%. In het algemene geval, met K steekproeven, heeft het foutieve éénzijdige interval een overdekking van $\Phi(z_\alpha\sqrt{K})$.

De volgende analogie kan wellicht nog duidelijker maken dat het optellen van betrouwbaarheidsgrenzen niet een gelijke maar een grotere overdekkingskans geeft. Als ik met één dobbelsteen gooi, dan is de kans 5/6 = 0,83 dat ik een getal ≤ 5 gooi. Hanekamp et al. zouden dan redeneren dat gooien met twee dobbelstenen een gelijkblijvende kans 5/6 geeft op een uitkomst ≤ 10 (5 + 5), waarbij dus de predictiegrens 5 die hoort bij één dobbelsteen, verdubbeld is. Het is eenvoudig in te zien dat deze kans groter is, en wel gelijk aan 33/36 = 0,92. Het probleem wordt pregnanter als we met meer dobbelstenen gooien. Bijvoorbeeld met 6 dobbelstenen is de kans niet 0,83 maar 0,99 op een uitkomst ≤ 30 (6 × 5).

De forse onzekerheden van Hanekamp et al. zijn dus gebaseerd op een statistische fout. Daarmee vervalt een groot deel van hun kritiek op het WUR-onderzoek.

Hoe het verder ging?

Op het rapport van Hanekamp et al. volgen vele berichten op sociale media. Dit culmineert in een rondetafelgesprek met de vaste Kamercommissie voor Economische Zaken op 22 februari 2017, waarbij Hanekamp cum suis én vertegenwoordigers van het RIVM en de WUR aanwezig zijn. Hierin worden harde verwijten gemaakt. Hanekamp spreekt over 'de afwezigheid van … wetenschappelijke tegenspraak … en de ondermijnende consequenties daarvan'. Vogelaar, voorzitter van het Mesdag Zuivelfonds, beticht de onderzoekers van 'misleiding van de Kamer' en 'een NH_3-anarchie in onze democratie'. De statistische fout is vervolgens tijdens een bijeenkomst met Hanekamp en Crok in Wageningen op 6 maart 2017 door mij naar voren gebracht. Dat weerhoudt de auteurs van het rapport er niet van om hun bevindingen in ongewijzigde vorm wetenschappelijk te publiceren (Hanekamp et al., 2017b). In een reactie daarop kwantificeer ik de onzekerheid met behulp van de parametrische bootstrap (Goedhart & Huijsmans, 2017). De betrouwbaarheidsintervallen zijn dan een factor drie minder breed. Hanekamp et al. zijn niet onder de indruk. Zij schrijven dat hun en onze methode '… only differ in the manner of estimating this uncertainty. Both approaches have strengths and shortcomings' (Briggs et al., 2017). Op 8 februari 2019 informeert Hanekamp leden van de Tweede Kamer over zijn bezwaren tegen het Landelijk Meetnet Luchtkwaliteit van het RIVM (Rotgers & Hanekamp, 2018). Tijdens deze briefing zegt Hanekamp over nieuwe data die hij van de WUR heeft ontvangen: '… de onzekerheden … die zijn zo groot dat de vraag rijst of al de experimenten die er ooit voor gedaan zijn überhaupt wel datgene hebben opgeleverd wat we hadden willen weten. … De eerste bevindingen lijken toch heel anders dan wat de WUR altijd heeft gezegd daarover.' Mogelijk is opnieuw de gewraakte fout gemaakt.

Hanekamp et al. richten hun pijlen vooral op de onzekerheid in de meetmethode voor ammoniakemissie. Waar ze aan voorbij gaan, is dat er 199 emissie-experimenten zijn gedaan, geheel conform de heilige graal van de statistiek: herhalen, herhalen en nog eens herhalen. Een eenvoudige statistische analyse op de emissie percentages van figuur 2 geeft aan dat er zeer significante

én grote verschillen zijn tussen de bemestingstechnieken (Goedhart & Huijsmans, 2017) en dat de emissiearme technieken tot een veel lagere ammoniakemissie leiden. Dit rechtvaardigt de conclusie dat de door Hanekamp et al. gezaaide twijfel over het emissiebeleid van de overheid gebaseerd is op een totaal verkeerd gebruik van statistische methoden.

Als statisticus krijg je niet vaak de kans om rechtstreeks bij te dragen aan het publieke en politieke debat. Het is evident dat statistische inbreng hier cruciaal is.

Literatuur

Briggs, W. M., Hanekamp, J. C., & Crok M. (2017). Comment on Goedhart and Huijsmans (2017). *Soil Use and Management, 33*, 603–604. https://doi.org/10.1111/sum.12393

Goedhart, P. W., & Huijsmans, J. F. M. (2017). Accounting for uncertainties in ammonia emission from manure applied to grassland. *Soil Use and Management, 33,* 595–602. https://doi.org/10.1111/sum.12381

Hanekamp, J. C., Crok, M., & Briggs, M. (2017a). *Ammoniak in Nederland; Enkele kritische wetenschappelijke kanttekeningen. V-focusrapport*. Wageningen. https://bit.ly/2KoHRSg

Hanekamp, J. C., Briggs, W. M., & Crok, M. (2017b). A volatile discourse; Reviewing aspects of ammonia emissions, models and atmospheric concentrations in The Netherlands. *Soil Use and Management, 33*, 276–287. https://doi.org/10.1111/sum.12354

Rotgers, G. R., & Hanekamp, J. C. (2018). *Ammoniak in Nederland. Een noordoostelijke spelbreker*. Leeuwarden: Stichting Mesdag Zuivelfonds. https://bit.ly/2HXXOGp

Paul W. Goedhart is sinds 1983 statisticus bij Biometris, Wageningen Universiteit en Research.
E-mail: paul.goedhart@wur.nl

Response to Paul Goedhart

We are grateful to Goedhart for helping to highlight the over-certainty that was ubiquitous in the discourse of atmospheric ammonia before 2017. It is good to be part of the change that woke people up to the idea that we were too sure of what was happening! Even now it is difficult to imagine estimates taken directly from the Ryden and McNeil (R&M) model were assumed to be, or were used as if they were, without error or uncertainty. We provided illustrations thereof in our report and paper Goedhart refers to. There are three main sources of the over-certainty and error in the R&M model.

The first arises in the internal parts of the R&M model, which uses a statistical technique to approximate atmospheric ammonia by height. That technique assumed certain mathematical parameters (from a regression) were fixed. We initially suggested one way of adding uncertainty to these parameters, and Goedhart later suggested another. The differences between these techniques, given the prior over-certainties, and the other difficulties of the R&M model, were not especially large or important. Indeed, as we will soon show in upcoming work, the differences are trivial when other problems with the model are accounted for.

Now since all probability, and therefore all probability models, is conditional on the assumptions made, the goodness of the uncertainty method in the mathematical parameters depends on how well the assumptions made accord with reality. Goedhart used a common technique, found in all statistical toolboxes, tied to the idea that the R&M model parameters "had" correlations. Most researchers would use this technique as a kind of default, and Goedhart's tutorial on this is adequately done. We used a different technique that did not put much interest in the correlations between parameters; our method instead turned an eye towards the predictive ability of the R&M model to better reproduce atmospheric ammonia. That is, we were not trying to adhere to frequentist theory, but instead to physical reality. We wanted predictions, and predictive uncertainties, matching with observations. In the limited samples of data we had at the time, the method we used did the job asked of it.

Second came the other problems with the R&M model, which show our initial attempts did not go far enough. We will leave a full discussion of these shortcomings to our new paper. Here we can say that the R&M model often 'breaks' when fed observations that are not perfectly in line with the crude statistical assumption R&M made (about ammonia by height). Absurd, i.e. completely non-physical, values are the result. And this is so regardless which uncertainty-in-the-parameters technique is used. That this has been missed so far in the literature is rather stunning.

Goedhart was right on the money when he said we ought to 'herhalen, herhalen en nog eens herhalen' (repeat, repeat, and repeat) our analysis. We did, obviously, take his advice, and subsequently discovered these much larger flaws.

The third and last problem are the observations themselves, produced by experiments that were, as we will show, not the cleanest, to say the least. This often leads to over-estimates of atmospheric ammonia.

Goedhart is right once again that 'statistische inbreng hier cruciaal is' to the ammonia debate. But since everybody knows how easy it is to manipulate statistics, or how confusing they can be to interpret, what's far more important is physics. We want a model that makes skillful and accurate predictions of reality. Everybody can understand good forecasts. With the fixes we suggest in our upcoming paper, we are well on our way to delivering these.

Matt Briggs and Jaap C. Hanekamp

Matt Briggs is an independent researcher.
E-mail: matt@wmbriggs.com

Jaap C. Hanekamp is a faculty member of the University College Roosevelt, Middelburg, the Netherlands and an adjunct faculty member of the University of Massachusetts Amherst School of Public Health and Health Sciences, Amherst, MA, USA. E-mail: j.hanekamp@ucr.nl

Vrouwen hebben invloed op de omvang van *STAtOR*!

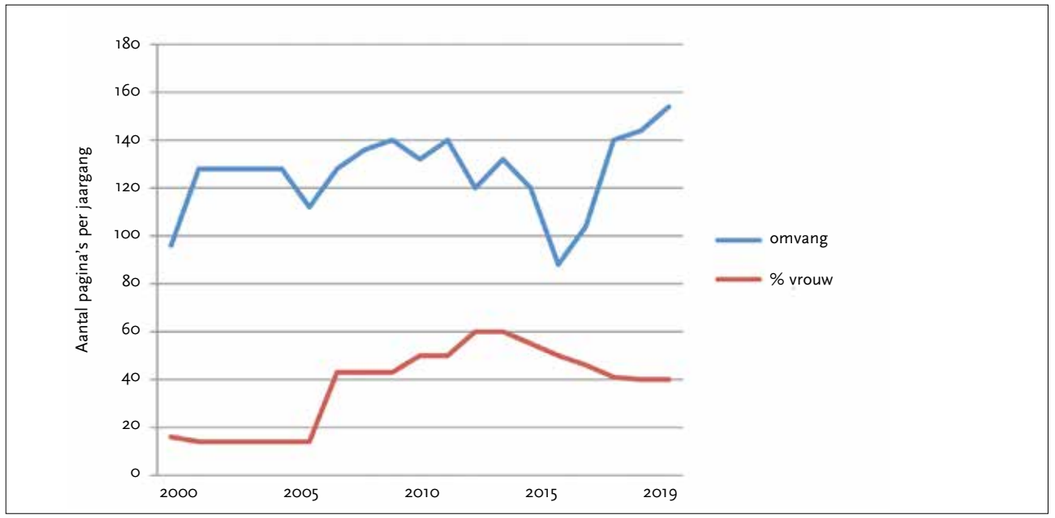
De vakantie staat voor de deur en dat betekent een periode van ontspannen. Ik heb daar steeds een beetje moeite mee gehad, bij mij was er nooit zo'n strikte scheiding tussen werk en vrije tijd. Mijn werk heb ik altijd als ontspannen ervaren en tijdens vakanties zag ik altijd wel iets dat mijn cijfergerichte geest bezig hield. Ik had een soort lijfspreuk, vrij naar Hieronymus van Alphen, 'mijn werken is spelen, mijn spelen is werken'. Mijn vrouw had als commentaar daarop dat ze hoopte dat ik niet zou overwerken.

We kennen allemaal onderzoeken waarbij correlatie en causaliteit door elkaar worden gehaald. Soms uit onkunde, soms met opzet. In de jaren 80 van de vorige eeuw leverde Rothamsted in de UK bij het computerprogramma Genstat enkele setjes testdata waarmee men kon controleren of de installatie op het lokale computersysteem goed was verlopen. Een van die datasets bestond uit het aantal waargenomen ooievaars en het aantal geboortes in het Verenigd Koninkrijk gedurende een reeks van jaren. Uiteraard was er een sterke correlatie, anders zou de befaamde Britse humor gefaald hebben. Na een uitgebreide congresborrel waarbij dit ter sprake kwam beweerde een Australiër dat hij in zijn land eens op zoek zou gaan naar de correlatie tussen het aantal schapen en het aantal patiënten met een venerische ziekte ...

Met cijfers spelen is leuk. Daarom heb ik de 20e jaargang van *STAtOR* aangegrepen om alle afgelopen jaargangen eens in cijfers om te zetten. Het kost even tijd, maar dan heb je ook wat. In een spreadsheet heb ik het aantal pagina's per nummer gezet, samen met het aantal mannelijke en vrouwelijke redactieleden. In de beste traditie van flutonderzoeken heb ik een grafiek gemaakt om te zien of er een verband is tussen deze gegevens. De verrassende conclusie is dat zodra het aandeel vrouwen in de redactie stijgt de omvang van ons blad sterk gaat variëren! (Zie figuur 1) Een boreaal politicus zal hier ongetwijfeld iets van vinden, ik beperk me slechts tot de kale waarnemingen.

Terug naar het ontspannen, dat gun ik al onze lezers. Ik nodig u uit om verder te spelen met deze data (de dataset vindt u op de VVSOR-webpagina's). U kunt er van alles aan toevoegen, van de CBS-prijsindex tot het aantal tropische dagen, het aantal bezoekers van dierentuinen of het aantal veroordelingen voor rijden onder invloed. Gebruik uw inzicht en uw fantasie en verras ons met opzienbarende correlaties! Uiteraard worden de fraaiste verklaringen voor de omvang van *STAtOR* in ons blad gepubliceerd.

Gerrit Stemerdink is eindredacteur van *STAtOR*.
E-mail: gjstemerdink@hotmail.com



Figuur 1. Het aandeel vrouwen in de *STAtOR*-redactie en de omvang van *STAtOR* van 2000 tot 2019

EERSTE KAMERVERKIEZING

De puzzelaar bepaalt de laatste zetels!

Het was spannend dit jaar … Welke partij pakt met een paar strategische stemmen een extra zetel in de Eerste Kamer? Hoe kan een strategische stem een zetel verschil maken? Wat gebeurde er in het verleden en wat zou er in mei bij de verkiezing kunnen gebeuren? Reden genoeg voor een analytische terug- en vooruitblik. *(Ten tijde van verschijning van dit artikel heeft de Eerste Kamerverkiezing reeds plaatsgevonden.)*

Jacob Jan Paulus

De Eerste Kamer wordt gekozen via een getrapte verkiezing, de burgers kiezen Provinciale Statenleden die vervolgens de Eerste Kamerleden kiezen. In maart waren de verkiezingen voor de Provinciale Staten, en de gekozen Statenleden stemmen in mei voor de Eerste Kamer. De Statenleden zijn in principe vrij in hun keuze en kunnen dus ook op een andere partij stemmen dan waarvoor ze zelf in de Provinciale Staten zijn gekozen. Dat biedt de mogelijkheid om de zetelverdeling in de Eerste Kamer te beïnvloeden. In de landelijke media wordt daarom elke 4 jaar bij deze verkiezing gespeculeerd over strategisch stemgedrag.[1] En de historie leert dat dit ook gebeurt.

Hoe worden de senatoren van de Eerste Kamer gekozen?

Elk lid van de Provinciale Staten mag één stem uitbrengen, waarbij de stemmen worden gewogen omdat het inwoneraantal en aantal Statenleden per provincie verschilt. Deze weging, de zogenaamde *stemwaarde,* wordt bepaald door het aantal inwoners van de provincie te delen door het aantal Statenleden in die provincie maal 100. Zie figuur 1. Zo telt in 2019 een stem van een Statenlid uit Zuid-Holland voor 668 en is daarmee ongeveer 6.8 keer zo zwaar als die van een Statenlid uit Zeeland, waarvan de waarde 98 is. Dit jaar mogen Bonaire, Sint-Eustatius en Saba voor het eerst meestemmen, al is het met een hele kleine waarde.

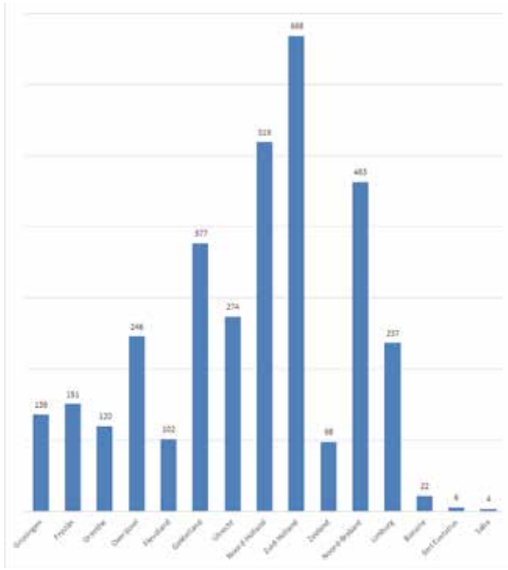
De totale waarde van de stemmen uitgebracht op een bepaalde partij is het behaalde stemcijfer voor die partij. De stemcijfers bepalen vervolgens de zetelverdeling in de Eerste Kamer. De Kieswet[2] schrijft hier het volgende over:

> Artikel U 8: Zoveel maal als de kiesdeler is begrepen in het stemcijfer van een lijst wordt aan die lijst een zetel toegewezen.
>
> Artikel U 9: De overblijvende zetels, die restzetels worden genoemd, worden achtereenvolgens toegewezen aan de lijsten die na toewijzing van de zetels het grootste gemiddelde aantal stemmen per toegewezen zetel hebben. Indien gemiddelden gelijk zijn, beslist zo nodig het lot.

De analyses worden echter een stuk eenvoudiger als je je realiseert dat deze procedure er in feite op neer komt dat de kiesdeler wordt verlaagd net zo lang tot er in totaal precies 75 zetels worden toegewezen.

Wat gebeurde er in 2011?

Via de Kiesraad[3] en Databank Verkiezingsuitslagen[4] zijn de 2011 verkiezingsuitslagen van de Provinciale Staten en Eerste Kamer te verkrijgen, en kan gemakkelijk worden achterhaald hoe er per provincie door de Statenleden is gestemd. De uitslag van de Provinciale Statenverkiezing geeft het aantal stemmen dat elke partij per provincie kan uitbrengen, en de uitslag van de Eerste Kamerverkiezing geeft het aantal stemmen dat elke partij per provincie heeft ontvangen. Deze vergelijking laat zien dat naast een ongeldige stem van D66 in Noord-Holland, door minstens 13 Statenleden niet op de eigen partij is gestemd!



Figuur 1. Stemwaarde Eerste Kamerverkiezingen 2019

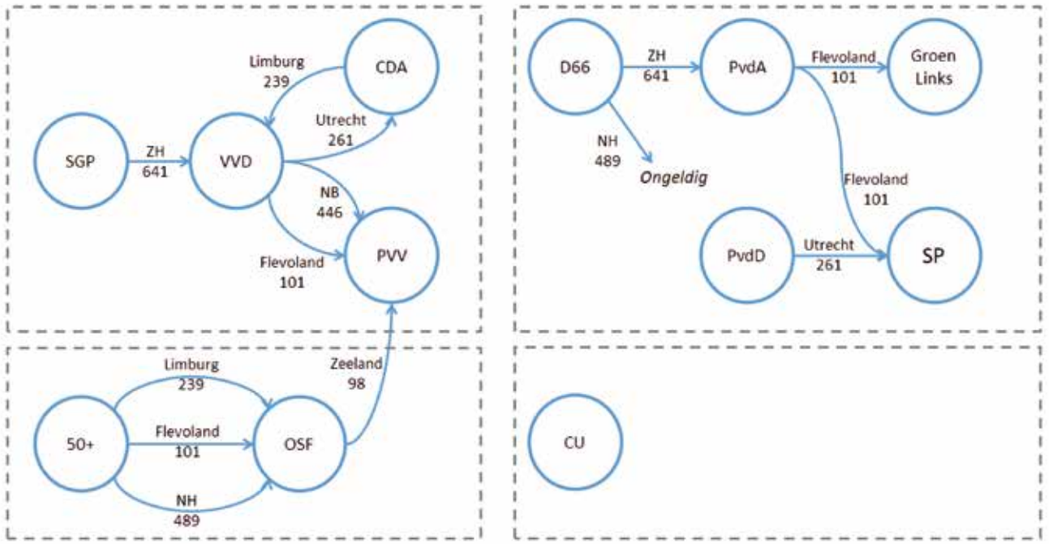
Neem de provincie Utrecht als voorbeeld. Bij de Provinciale Statenverkiezingen haalt het CDA daar 6 zetels, maar krijgt bij de verkiezing van de Eerste Kamer van de Utrechtse Statenleden 7 stemmen. Omgekeerd haalt de VVD in Utrecht 11 zetels in de Provinciale Staten maar krijgt voor de Eerste Kamer maar 10 stemmen uit Utrecht. Het is aannemelijk dat een VVD-Statenlid zijn stem aan het CDA heeft gegeven. Omdat er in Utrecht ook een stem van de PvdD naar de SP gaat, hadden we dit ook anders kunnen verklaren (PvdD naar CDA en VVD naar SP) maar dat is politiek gezien niet logisch.

In figuur 2 zijn alle zo te achterhalen verschuivingen samengevat. Iedere pijl representeert een verschoven stem, waarbij de richting aangeeft van welke naar welke partij de stem is gegaan, en staat bij de pijl de betreffende provincie en de waarde van de stem. De lokale partijen zijn landelijk verenigd in de Onafhankelijke SenaatsFractie (OSF).

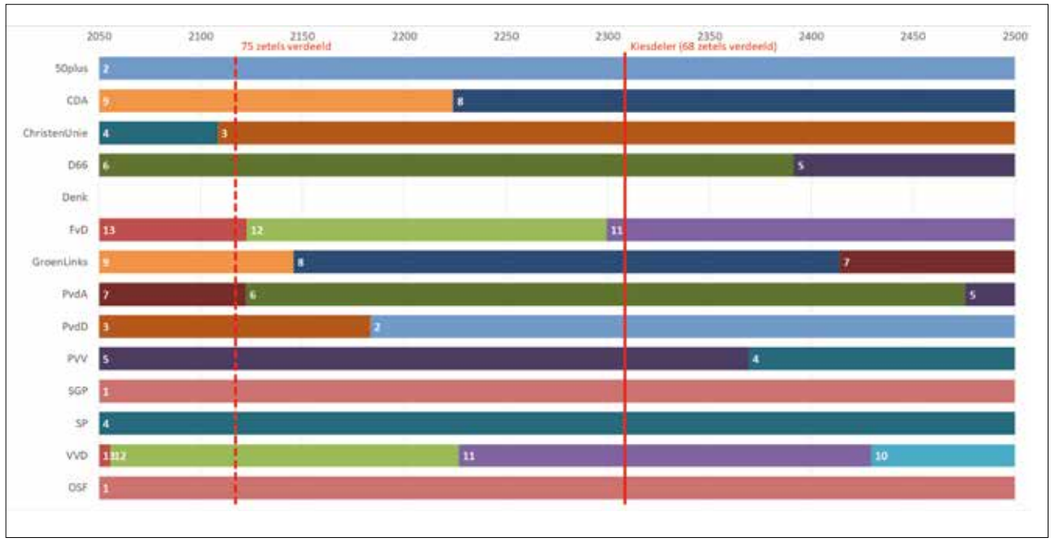
De verschuivingen vallen, niet helemaal verrassend, in 4 blokken uiteen. Al voor de verkiezing in 2011 was bekend dat OSF en 50Plus een afspraak hadden gemaakt. In Flevoland, Noord-Holland en in Limburg is, zoals linksonder in de figuur is te zien, een stem van 50Plus naar de OSF gegaan. Die drie stemmen vertegenwoordigden samen een stemwaarde van 829 en dat was voldoende om OSF aan een zetel te helpen, terwijl 50Plus nog genoeg overhield voor de eigen zetel. Zonder verdere verschuivingen zou die extra zetel voor de OSF ten koste zijn gegaan van de SP.

De verschuivingen tussen D66, PvdA, GroenLinks, PvdD en de SP, in het blok rechtsboven in de figuur, lijken ingegeven te zijn om dat laatste te voorkomen. Met D66 als de grootste sponsor en SP de grootste ontvanger. En die verschuivingen zouden op zich voldoende zijn geweest om de laatste restzetel weer naar de SP te halen, nu ten koste van de PVV.

Waarschijnlijk hebben de coalitiepartijen van destijds ook zitten rekenen en kans gezien de laatste restzetel toch weer naar de PVV te halen. Politiek gezien is de stem van SGP naar de VVD in Zuid-Holland opmerkelijk te noemen, maar het is denkbaar dat er een tussenstation is gekozen: de SGP brengt een stem uit op de CDA, en het CDA een stem op de VVD. Hoe het ook zij, het nettoresultaat in deze provincie is dat er een stemwaarde van 641,



Figuur 2. Strategisch stemmen van de Provinciale Staten voor de Eerste Kamer in 2011



Figuur 3. Basisscenario voor het strategisch stemmen van de Provinciale Staten voor de Eerste Kamer in 2019

UITGEBRACHTE STEMMEN	Drenthe	Flevoland	Fryslân	Gelderland	Groningen	Limburg	Noord-Brabant	Noord-Holland	Overijssel	Utrecht	Zeeland	Zuid-Holland	Bonaire	Sint Eustatius	Saba	Stem-waarde	Kale zetels	Rest-zetels
50plus	1	2	1	2	1	1	2	1	1	1	2	2				5099	2	
CDA	5	3	8	7	3	9	8	4	9	5	7	4				20016	8	1
ChristenUnie	3	3	3	4	4		1	1	4	4	2	3				8433	3	
D66	2	2	2	4	3	3	5	6	3	5	1	5				14348	6	
Denk		1						1		1		1				1563	0	
FvD	6	8	6	8	5	7	9	9	6	6	5	11				27593	11	2
GroenLinks	4	4	3	6	6	4	5	9	5	8	2	5				19311	8	1
PvdA	6	3	6	5	5	3	3	6	4	4	4	4				14855	6	1
PvdD	1	2	1	2	1	2	2	3	1	2	1	2				6550	2	1
PVV	3	4	3	3	2	7	4	3	3	2	2	4				11846	5	
SGP		1		3					2	1	5	2				3825	1	
SP	3	2	2	3	4	4	5	3	3	2	2	2				10179	4	
VVD	6	6	4	8	4	5	10	9	6	8	4	10				26722	11	1
OSF	1		4		5	2	1				2		9	5	5	2785	1	
Stemwaarde	120	102	151	377	136	237	463	519	246	274	98	668	22	6	4	173125	68	7
														Kiesdeler		2308,33		

Tabel 1. Basisscenario voor het strategisch stemmen van de Provinciale Staten voor de Eerste Kamer in 2019. *We nemen aan dat de lijstverbintenis tussen ChristenUnie-SGP in Noord-Brabant zal stemmen op de ChristenUnie, de lijstverbintenis tussen 50Plus-Partij van de Ouderen in Noord-Holland zal stemmen op 50Plus en alle lokale partijen en eilanden op de Onafhankelijke Senaat Fractie (OSF)*

goed voor bijna 1/3 zetel, van de SGP naar de VVD gaat. Kijken we nog iets verder naar het totaal van de verschuivingen in het blok linksboven in de figuur, dan is te zien dat uiteindelijk van die door de SGP afgestane stemwaarde, 22 naar het CDA, 72 naar de VVD en 547 naar de PVV is gegaan. Wat daarmee de doorslaggevende bijdrage aan de tiende zetel voor de PVV is.

Door de ongeldige stem van D66 verliest deze partij een restzetel en komt deze alsnog terecht bij de SP.

Wat te verwachten voor de komende verkiezing?

De samenstelling van de Provinciale Staten en de stemwaardes zijn bekend, en beschikbaar via de Kiesraad. We kunnen dus doorrekenen wat er gaat gebeuren als we aannames doen over het stemgedrag van de Statenleden. Als je verwacht dat iedereen op zijn/haar eigen partij stemt, dan geeft tabel 1 weer wat de uitslag zal zijn op 27 mei 2019.

Met de kiesdeler gelijk aan 2308,33 (de totale stemwaarde 17.3125 gedeeld door 75 Eerste Kamerzetels) worden er 68 kale zetels verdeeld over de partijen. In het overzicht (figuur 3) geeft elk balkje aan hoeveel zetels een partij krijgt als de kiesdeler valt tussen het begin en eind van het balkje. De rode lijn is de kiesdeler en de doorkruiste balkjes tellen op tot de 68 kale zetels.

Bij het verdelen van de restzetels wordt de kiesdeler ; stap voor stap verlaagd. Zo krijgen de volgende partijen achtereenvolgende de restzetels: FvD, VVD, CDA, PvdD, Groenlinks, FvD en PvdA. Bij de onderbroken rode lijn zijn de 75 Eerste Kamerzetels verdeeld, en stopt het toewijzen van restzetels.

Het is opvallend hoe dicht de laatste drie restzetels bij elkaar liggen en hoe nipt de ChristenUnie geen restzetel behaalt (zie tabel 2). Het is daarom te verwachten dat hier *strategisch* gestemd zal worden om de verdeling van deze laatste restzetels te beïnvloeden.

Hoe gaat er strategisch gestemd worden?

Zoals gezegd zullen er vast en zeker strategische stemmen uitgebracht worden. Om te beginnen valt te verwachten dat de SGP de ChristenUnie aan een extra zetel gaat helpen. Er zijn bij de eerdere Eerste Kamerverkiezingen ook stemmen van de SGP naar de ChristenUnie gegaan. Als bijvoorbeeld de drie SGP Provinciale Staten-

	Bij # zetels	Waarde per zetel
ChristenUnie	4	2108,3
FvD	13	2122,5
GroenLinks	9	2145,7
PvdA	7	2122,1

Tabel 2. Restzetels 2019

leden in Gelderland op de ChristenUnie stemmen, krijgt de ChristenUnie er 1 kale zetel bij zonder dat de SGP haar zetel in gevaar brengt. Dit gaat dan ten koste van de restzetel van de PvdA.

De linkse partijen zullen hun hoofd gaan breken over het terugwinnen van de restzetel van de PvdA, en het verdedigen van de restzetel van GroenLinks die ook niet zo zeker meer lijkt. Gaan ze die hulp halen bij de SP, of zelfs de Caribische eilanden? Daarnaast zal FvD de verdediging inzetten om de laatste restzetel te behouden. Wat doet Denk met haar stemmen die 2/3 zetel waard zijn? En kunnen tot slot de huidige regeringspartijen samen nog een scenario in hun voordeel uitwerken?

Tot slot

Wij kunnen met kwantitatieve analyses inzicht geven in deze problematiek en de consequenties doorrekenen van elk mogelijk scenario. Het is aan de politici om in te schatten hoe politiek haalbaar deze scenario's gaan zijn en hoe ze strategisch gaan stemmen. Eind mei zullen we zien of en hoe er strategisch gestemd is, en wat het effect is geweest op de zetelverdeling.

We kunnen slechts gissen naar de beloftes die partijen doen om een extra stem binnen te halen.[5] Ik laat het aan de lezer om zelf een mening te vormen over hoe wenselijk dit strategisch stemgedrag is, en of het democratische proces zo bedoeld is.

1. NOS: https://nos.nl/artikel/2277868-koehandel-om-zetels-eerste-kamer-is-begonnen.html
2. Kieswet: https://wetten.overheid.nl/BWBR0004627/2019-02-22/
3. Kiesraad: https://www.kiesraad.nl
4. Databank Verkiezingsuitslagen: https://www.verkiezingsuitslagen.nl
5. *de Volkskrant* van 17 juni 2011: 'Statenlid Johan Robesin uit Zeeland heeft premier Mark Rutte per brief nog eens uitdrukkelijk gewezen op de belofte om niet te ontpolderen in Zeeland. Een belofte waarvan Rutte altijd heeft gezegd, dat deze niet door hem werd gedaan in ruil voor de steun van het Statenlid.' https://www.volkskrant.nl/nieuws-achtergrond/deed-rutte-toch-ontpolderbelofte-robesin-schrijft-bezorgde-brief~bb0c58ff/

Jacob Jan Paulus is Senior Consultant en Optimalisatie Expert bij Consultants in Quantitative Methods (CQM), waar hij met kwantitatieve (optimalisatie) modellen fact-based beslissingen in het bedrijfsleven ondersteund. Vanuit persoonlijke interesse volgt hij het stemgedrag bij de Eerste Kamerverkiezing.
E-mail: JacobJan.Paulus@cqm.nl

Young Statisticians are looking back at Spring 2019

In 2019, we organized several events. We kicked-off the year with New Year's drinks in February, the fact that the first month of 2019 already had passed did not reduce the fun on this evening at Het Gegeven Paard in Utrecht. During the VVSOR annual meeting, we hosted the annual Young Statisticians Lunch before the first talks on Wednesday. After the dinner of the annual meeting, we hosted a Pub Quiz in which several teams answered questions about machine learning, privacy, GDPR, R-code and of course the publications of the VVSOR board members. In April, we visited ABN AMRO to learn about their ways of applying statistics and data science at a renowned bank. Finally, we organized a Science Café about train scheduling in June. In this symposium, three speakers explained the way trains are scheduled and the role of statistics in this process. A good start of the year, we hope to see you all at our events after the summer break!

63rd ISI World Statistics Congress 2021 in The Hague!

This major biannual conference of the International Statistical Institute will take place from 11-15 July 2021 in the World Forum in The Hague. More than 2500 participants are expected from over 130 countries. They can choose from over 1300 presentations.

This is a unique chance for statisticians from the Netherlands to acquaint with new colleagues and meet some the world's leading experts in our science.

Information: ISI Permanent Office (P.O. Box 24070, 2490 AB The Hague, phone +31–70–3375737), e-mail: isi@cbs.nl, website: isi-web.org