

STAtOR

periodiek van de VVSOR jaargang 20, nummer 4, december 2019

Het meten van discriminatie in algoritmische besluitvorming

Gedetailleerde en tijdige statistieken over de Nederlandse samenleving

Wiskunde op de Wallen

Statistici tegen alcoholmisbruik

Kansrekening, een echt wiskundevak voor havo en vwo

Heisenberg; over statistiek, onzekerheid en filosofie

Ranglijstjes; 'Duisternis en Computergestuurde Willekeur'



STATOR

Jaargang 20, nummer 4, december 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, Joep Burger, Kristiaan Glorie, Caroline Jagtenberg, Guus Luijben (eindredacteur), Kerry Malone, Richard Starmans, Gerrit Stermerdink (eindredacteur), Vanessa Torres van Grinsven en Sanne Willems. Vaste medewerkers: 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. Laura Zwep, db@vvsor.nl; Penningmeester: Judith ter Schure MSc, penningmeester@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, september en december.

Ontwerp en opmaak

Pharos, Nijmegen

Uitgever

© Vereniging voor Statistiek en Operations Research
ISSN 1567-3383

OUDEJAARSOVERPEINZINGEN

Dit is het laatste nummer van onze 20e jaargang, een jaar waarin we regelmatig stilstonden bij de impact die STATOR in al die 20 jaar heeft getracht te brengen. Blij verrast zijn we door de vele zeer positieve reacties die we kregen, onze dank daarvoor. Vooral de unaniem jubelende toon die we in de vele bijdragen in nr. 2019-1 ervoeren deed ons soms een beetje blozen.

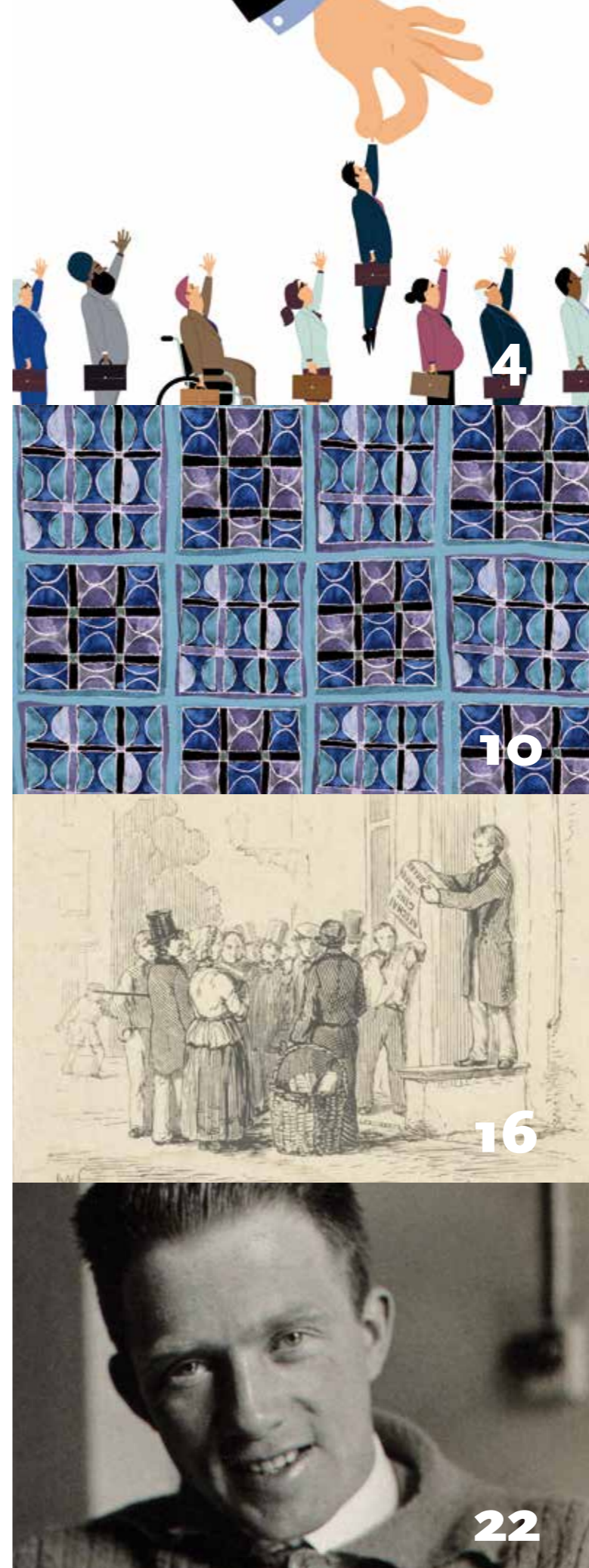
Zoals het aan het einde van het jaar gebruikelijk is, kijken we in deze redactionele inleiding niet alleen terug, maar vooral ook vooruit. En dan zien we allereerst de Annual Meeting van 2020 met de intrigerende titel *Causality and Complexity*. Meer daarover zult u kunnen lezen in nr. 2020-1, maar wij hopen dat u nu al de datum van 12 maart reserveert voor een bezoek aan Utrecht.

Zeer actueel is de vraag of het toenemend gebruik van algoritmes tot onbedoelde discriminatie kan leiden. Stan van Loon vertelt in zijn bijdrage hoe dit gemeten kan worden. Jan van den Brakel c.s. vertellen ons hoe statistieken over de Nederlandse samenleving gedetailleerd en tijdig kunnen worden samengesteld. Wist u dat statistici zich in het verleden sterk hebben gemaakt tegen alcoholmisbruik? Lees daarover in het interessante verhaal van Ronald van der Bie. Ten slotte een bijdrage van Richard Starmans, waarin de rol van Heisenberg in statistiek, onzekerheid en filosofie aan bod komt. De meesten van ons kennen Werner Heisenberg als de fysicus die op zeer jonge leeftijd belangrijke bijdragen aan de kwantummechanica leverde. Maar er is meer dan de naar hem genoemde onzekerheidsrelatie. Zou het toeval zijn dat we dit artikel kunnen zien als een voorbereiding op het thema van de Annual Meeting?

Ditmaal drie columns: Henk Tijms geeft zijn mening over de hervorming van het rekenonderwijs en de plannen om statistiek een belangrijke plaats in het curriculum te geven, Gerard Sierksma sluit min of meer aan bij de vraagstelling van Stan van Loon en vertelt over computergestuurde willekeur bij ranglijsten, en Gerrit Stermerdink haalt een herinnering op aan hoe zijn goedbedoelde aanwijzingen voor een veilige route door het centrum van Amsterdam werden gezien als een "Guide for the Red Light District".

Verder in dit nummer de aankondigingen van de NGB-LNMB-bijeenkomst en de Annual Meeting van onze vereniging, nieuws van de Young Statisticians en oproepen om kandidaten te nomineren voor de Van Dantzig prijs en de Van Zwet en Hemelrijk Awards.

De redactie wenst u veel leesplezier rond de jaarwisseling en hoopt velen van u te ontmoeten op de Annual Meeting.



INHOUD

- 2 Redactioneel
- 4 Het meten van discriminatie in algoritmische besluitvorming | STAN VAN LOON
- 9 VVSOR Annual Meeting 2019
- 10 Gedetailleerde en tijdige statistieken over de Nederlandse samenleving | JAN VAN DEN BRAKEL, HARM JAN BOONSTRA, SABINE KRIEG & MARC SMEETS
- 15 Wiskunde op de Wallen; Of hoe goede bedoelingen wel eens onverwacht kunnen uitpakken – column | GERRIT STEMERDINK
- 16 Statistici tegen alcoholmisbruik | RONALD VAN DER BIE
- 19 Oproep om kandidaten te nomineren voor de Jan Hemelrijk en Willem R. van Zwet Awards 2019
Oproep voor Nominaties Van Dantzig Prijs 2020
- 20 Kansrekening, een echt wiskundevak voor havo en vwo – column | HENK TIJMS
- 22 Heisenberg; over statistiek, onzekerheid en filosofie | RICHARD STARMANS
- 28 Ranglijstjes; 'Duisternis en Computergestuurde Willekeur' – column | GERARD SIERKSMA
- 30 Annelieke Baller gepromoveerd
Young Statisticians are looking back at a great 2019
Wisseling in de redactie
- 31 LNB & NGB Seminar op 15 januari 2010
- 32 63rd ISI World Statistics Congress 2021 in Den Haag



HET METEN VAN DISCRIMINATIE IN ALGORITMISCHE BESLUITVORMING

Kunstmatige Intelligentie wordt de laatste jaren steeds vaker gebruikt om beslissingen te nemen over keuzes die de levens van mensen beïnvloeden. Dit kan gaan om het bepalen welke online-advertentie het meest geschikt is voor een persoon, of iemand goed zal presteren tijdens een bepaalde baan of wat de kans is dat iemand (nogmaals) een strafbaar feit zal plegen. Een risico is dat deze algoritmes bepaalde groepen benadelen, bijvoorbeeld op basis van geslacht of etniciteit. Is het mogelijk om discriminatie in algoritmische besluitvorming te meten?

STAN VAN LOON

Veel mensen zien Kunstmatige Intelligentie (KI) als eerlijk en niet discriminerend. 'Het is een computer die de beslissing maakt, computers kunnen geen fouten maken.' Helaas is dit niet altijd het geval.

Om te begrijpen hoe algoritmes discriminerend kunnen zijn, kijken we eerst naar de manier waarop deze algoritmes gecreëerd worden. Algoritmes worden vaak 'getraind' met behulp van data uit het verleden. Het al-

goritme leert om de redeneringswijze van mensen over te nemen. Op deze manier kan een algoritme, wanneer deze nieuwe data ziet, zelf een keuze maken over de actie die benodigd is bij deze specifieke persoon. Als de data die het algoritme gebruikt discriminerend zijn, zal het algoritme deze discriminatie hoogstwaarschijnlijk overnemen.

Amazon

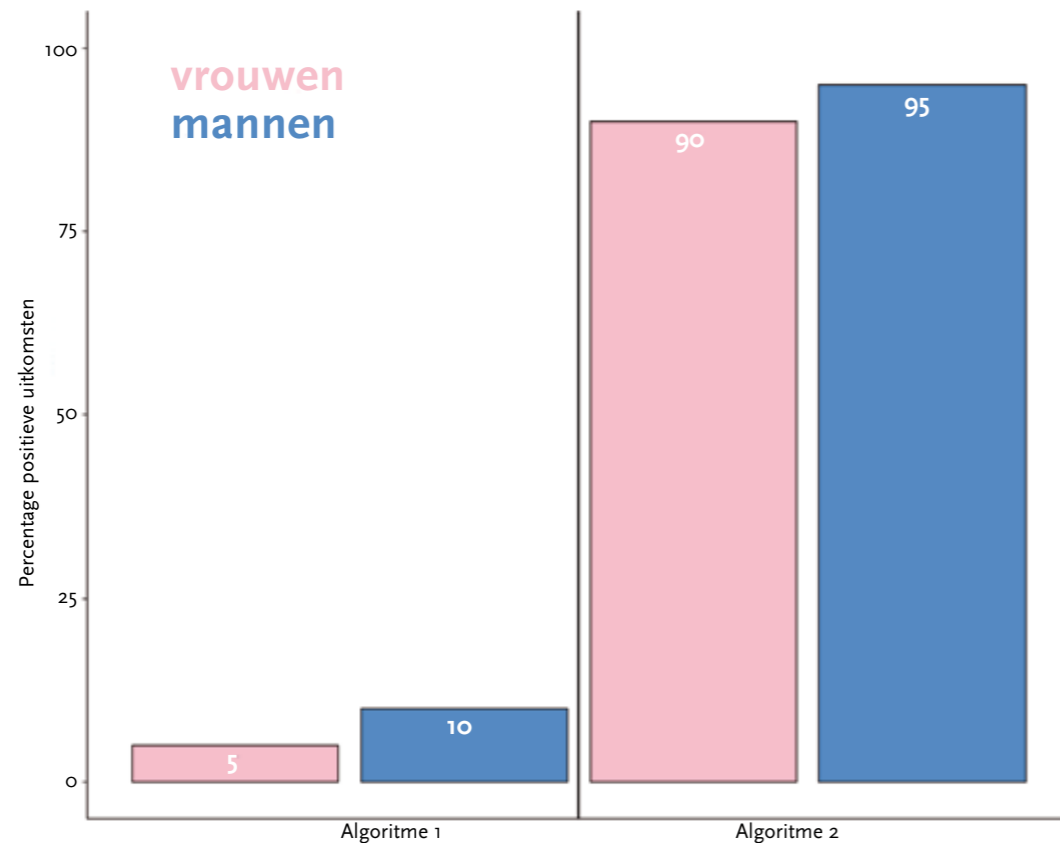
Een goed voorbeeld van een algoritme dat mensen discrimineerde, is een algoritme van Amazon. Amazon wilde haar *recruitment* proces verbeteren door een KI in te zetten om mensen te selecteren voor openstaande vacatures. De meest geschikte mensen werden vervolgens uitgenodigd voor een sollicitatiegesprek. Het algoritme in deze KI werd getraind met data over welke mensen er wel en niet werden aangenomen bij Amazon in de afgelopen tien jaar. In deze periode werden vooral (blanke) mannen aangenomen. Ondanks dat het niet impliciet verteld werd van welk geslacht of etniciteit deze mensen waren, wist het algoritme te discrimineren op geslacht en mogelijk zelfs op etniciteit. Het algoritme leerde op basis van de data die het kreeg, dat het een positieve eigenschap was om een man te zijn. Het kan bijvoorbeeld zo zijn dat het algoritme aan de hobby's van een persoon kan afleiden dat het een man of een vrouw is. Denk aan iemand die voetbal kijken en gamen als hobby's heeft. Deze persoon heeft een veel grotere kans om een man te zijn dan iemand die shoppen als hobby heeft. Het ligt dus vaak niet aan het algoritme of aan de KI zelf, maar aan de data waar het van leert.

Het is dus mogelijk dat Amazon in deze tienjarige periode onbedoeld heeft gediscrimineerd op geslacht en op etniciteit. Amazon probeerde, na het ontdekken van de onbedoelde discriminatie, het algoritme aan te passen

om minder discriminerend te zijn. Helaas was het algoritme zo discriminerend dat het niet meer gebruikt kon worden voor het selecteren van mogelijke werknemers. Los van een deukje in de reputatie, leed Amazon geen ernstige gevolgen. Amazon werd niet aangeklaagd voor het discrimineren op basis van geslacht of op basis van etniciteit. (Dastin, 2018)

Discriminatie in kaart brengen

Aan de hand van deze en andere leermomenten zijn er meerdere onderzoeken uitgevoerd om discriminatie in algoritmes tegen te gaan of te verminderen. In elk van deze onderzoeken werden er bepaalde processen bedacht die discriminatie in algoritmes konden verminderen. Hoeveel discriminatie tegengegaan of verminderd werd, werd vaak berekend met een formule bedacht door een samenwerking tussen de University of Toronto en Microsoft in 2013. Zij bedachten de volgende maatstaf voor het meten van discriminatie in algoritmes: 'De kans dat iemand van de niet-beschermde groep een positieve uitkomst heeft minus de kans dat iemand van de beschermde groep een positieve uitkomst heeft.' In het geval van de recidivisten-dataset (deze dataset wordt nader besproken) op basis van geslachtsdiscriminatie, kunnen wij dit als volgt interpreteren: 'De kans dat een man geclassificeerd zal worden als een recidivist minus de kans dat een vrouw geclassificeerd zal worden als een recidivist.' Deze maatstaf is alleen van toepassing op de uiteindelijke keuze die een algoritme maakt en kan een eenzijdig beeld geven van de keuzes die een algoritme maakt om tot de beslissing te komen. Bovendien is het mogelijk dat deze maatstaf een verkeerd beeld geeft van de hoeveelheid discriminatie die een algoritme bevat. Dit laten we zien aan de hand van een fictief voorbeeld.



Figuur 1. De kans op een positieve uitkomst voor mannen en vrouwen in twee verschillende algoritmes

In figuur 1 zijn twee verschillende algoritmes te zien. In het eerste algoritme hebben de vrouwen een 5% kans om geïdentificeerd te worden als een recidivist en de mannen een 10% kans. In het tweede algoritme zijn deze kansen respectievelijk 90% en 95%. Beide algoritmes zullen als 5% discriminerend beschouwd worden volgens de huidige maatstaf ($10 - 5 = 95 - 90 = 5$). Echter krijgen mannen een twee keer zo hoge kans op een positieve uitkomst in algoritme 1, in algoritme 2 is deze kans ongeveer 1,06 keer zo hoog. Dit geeft aan dat deze maatstaf een vervormd beeld kan geven van de hoeveelheid discriminatie die zich in een algoritme bevindt. Om dit te voorkomen, is er een methode ontwikkeld die een beter en completer beeld geeft van discriminatie in algoritmes. Aan de hand van een model dat met logistische regressie de kans op recidive schat, worden er vier alternatieve manieren gegeven om discriminatie in algoritmes zichtbaar te maken.

De dataset

We zullen deze nieuwe methode illustreren aan de hand

van de dataset van 'The Florida Department of Corrections' (ProPublica, 1996). In deze dataset staan gegevens over mensen die eenmaal of meerdere malen een strafbaar feit gepleegd hebben. Van deze dataset worden hun leeftijd, aantal eerdere arrestaties, of er geweld plaatsvond in de laatste arrestatie en hoe ernstig de daad is waarvan de persoon beschuldigd wordt gebruikt als variabelen die het algoritme meeneemt in de berekeningen. Aan de hand van deze 'variabelen' bepaalt het algoritme hoe groot de kans is dat elk specifiek persoon nogmaals een strafbaar feit zal plegen, dit is de kans op recidive. Deze berekende kans wordt ook wel de 'modelscore' genoemd.

De data worden gesplitst in train-data om het model op te *fitten* en een test-dataset om de nauwkeurigheid van het model te testen. De maatstaven die gebruikt worden om de graad van discriminatie te laten zien, zijn allemaal metingen op de resultaten van de logistische regressie uitgevoerd op de testdata.

We gebruiken logistische regressie om de kans op recidive te voorspellen en het meten van discriminatie te illustreren. Deze logistische regressie gebruikt de eerdergenoemde variabelen om de kans op recidive te

voorspellen. De formule van de logistische regressie ziet er als volgt uit:

$$\text{Log}\left(\frac{p}{1-p}\right) = 0,55 - \text{Graad} \cdot 0,31 - \text{Leeftijd} \cdot 0,04 + \text{Misdrijven} \cdot 0,15 + \text{Geweld} \cdot 3,24 + \text{Geslacht} \cdot 0,21$$

- p = De kans op recidive.
- Graad = De graad van het misdrijf. Dit is een getal van 0 tot 11. Hoe groter dit getal, hoe ernstiger het misdrijf.
- Misdrijven = Aantal eerder gepleegde misdrijven.
- Geweld = 0, geen geweld bij het misdrijf. 1, anders.
- Geslacht = 0, als Vrouw. 1, als Man.

Discriminatie weergegeven in de modelscores

Om discriminatie op basis van geslacht weer te geven kunnen de modelscore-verdelingen van mannen en vrouwen met elkaar vergeleken worden. Dit wordt gedaan in figuur 2. Hier is te zien dat een logistische regressie (die geslacht ook daadwerkelijk als een voorspellende variabelen meeneemt) over het algemeen mannen een hogere modelscore geeft. Dit betekent dat de logistische regressie mannen sneller als recidivist zal classificeren dan vrouwen. Dit komt doordat er later een cutoff gezet wordt voor het bepalen of iemand als mogelijke recidivist bestempeld wordt. Hoe dit precies werkt wordt later in dit artikel uitgelegd.

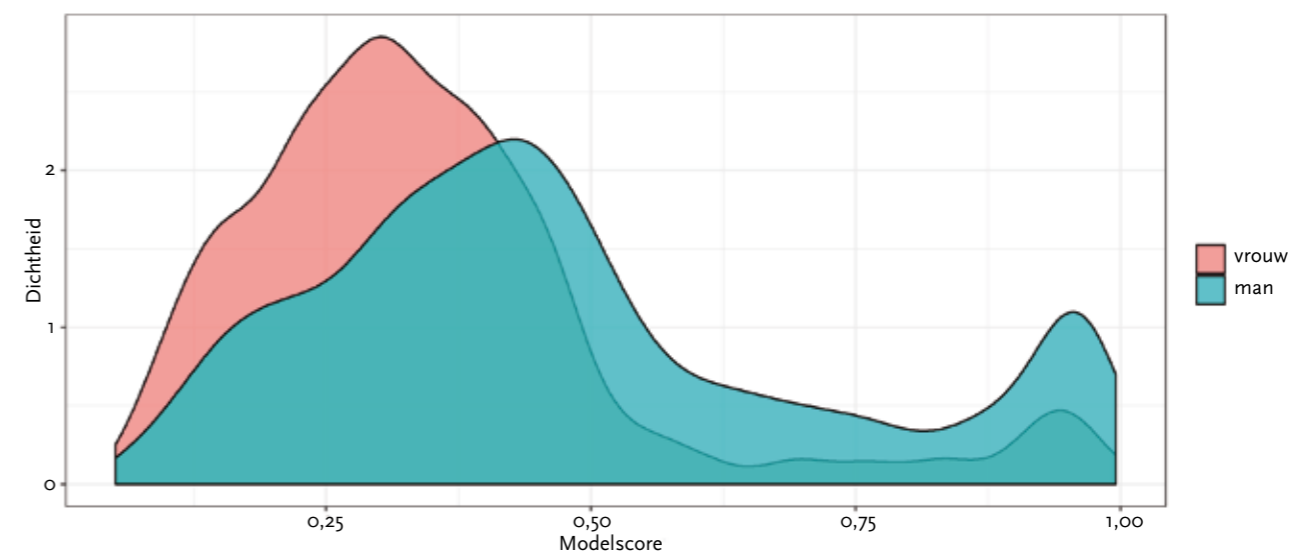
Het is natuurlijk mogelijk dat de twee verdelingen erg op elkaar lijken en elkaar grotendeels overlappen. Als dit het geval is kan de Kolmogorov-Smirnov test gebruikt worden om te testen of de twee verdelingen significant van elkaar verschillen. In dit geval heeft

de Kolmogorov-Smirnov test een p-waarde lager dan $2,2 \times 10^{-16}$, een p-waarde die nauwelijks van nul te onderscheiden is. Dit geeft aan dat de twee verdelingen hoogstwaarschijnlijk van elkaar verschillen en dat er onderscheid gemaakt wordt tussen mannen en vrouwen in deze logistische regressie.

AUC_D

Om op een makkelijker te begrijpen maatstaf voor discriminatie te komen, kunnen de modelscores van de mannen en de vrouwen met elkaar vergeleken worden en worden omgezet in een getal. Dit wordt gedaan met behulp van de Area Under the Curve of AUC in het kort. De AUC wordt oorspronkelijk gebruikt als een maatstaf om de nauwkeurigheid van een model te bepalen. De AUC geeft weer hoe groot de kans is dat een willekeurig persoon die uiteindelijk een recidivist blijkt te zijn een hogere modelscore heeft gekregen dan iemand die uiteindelijk geen recidivist blijkt te zijn.

In dit geval wordt de AUC gebruikt als een maatstaf voor discriminatie, vanaf nu wordt dit de AUC_D genoemd. De AUC_D geeft de kans weer dat een willekeurige man een hogere modelscore heeft dan een willekeurig gekozen vrouw. In het geval dat de AUC_D gelijk is aan 1, heeft elke man een hogere modelscore dan elke vrouw. In het geval dat de AUC_D gelijk is aan 0, geeft dat aan dat elke vrouw een hogere modelscore heeft dan elke man. Als de AUC_D berekend wordt over figuur 2 komt hier een waarde uit van ongeveer 0,69. Wat aangeeft dat een willekeurige man uit de testdata een 69% kans heeft om een hogere



Figuur 2. De modelscore-verdelingen van een logistische regressie van de recidivisten-dataset

modelscore te krijgen dan een willekeurige vrouw uit de testdata.

Op deze manier kan de discriminatie in de modelscores van een algoritme al weergegeven worden met een enkel getal.

Discriminatie weergeven in de keuze

Nadat de modelscores voor elk individu berekend zijn kan er een *cutoff* bepaald worden. Deze *cutoff* bepaalt of een specifieke waarde van elke modelscore gelabeld wordt als 'positief' of als 'negatief'. In het geval van de recidivisten-dataset worden alle individuen van de testdata met een modelscore gelijk of lager dan 0,5 geclassificeerd als 'niet-recidivisten' en alle individuen in de testdata met een modelscore boven de 0,5 worden gelabeld als een recidivist. Uiteindelijk kan er gekeken worden naar de spreiding van recidivisten en niet-recidivisten en van mannen en vrouwen. Om deze spreiding weer te geven kan een kruistabel opgezet worden. Een voorbeeld van een kruistabel is te zien in tabel 1.

		VOORSPELLING	
		Niet-recidivist	Recidivist
GESLACHT	Vrouw	298	46
	Man	955	505

Tabel 1. Kruistabel van de keuze van de logistische regressie op de testdata

Aan deze kruistabel is te zien dat mannen relatief vaker worden geclassificeerd als recidivisten volgens de logistische regressie. Als we dit percentage van de mannen vergelijken met het daadwerkelijke percentage van de mannen uit de dataset, valt op dat deze aanzienlijk lager is: 0,35 volgens het algoritme en 0,47 in de dataset. Voor de vrouwen zijn deze percentages respectievelijk 0,13 en 0,36. Over het algemeen classificeert de logistische regressie mensen sneller als niet-recidivisten, dit heeft te maken met de vaste *cutoff* bij een score van 0,50. In figuur 2 valt op dat er veel waarnemingen waren net onder de 0,50. Dit kan verklaren waarom er bij deze *cutoff* minder men-

sen geclassificeerd worden als recidivist. Hierdoor hebben vrouwen over het algemeen een kans van maar liefst 13% om geclassificeerd te worden als recidivist. Dit algoritme toont dus sterke positieve discriminatie tegenover vrouwen. Mogelijk wordt dit veroorzaakt door het meenemen van het geslacht in de logistische regressie. Wat zal er gebeuren als deze variabele niet wordt gebruikt?

Helaas, zelfs als de variabele van het geslacht uit de logistische regressie gehaald wordt, blijkt de logistische regressie nog steeds discriminerend te zijn ten opzichte van de feitelijke waarnemingen. Volgens de regressie hebben mannen een 32% kans en vrouwen een 19% kans om geclassificeerd te worden als een recidivist. Deze kansen liggen al dicht bij elkaar en liggen ook dicht bij daadwerkelijke waarnemingen, maar laten nog steeds een duidelijk nadeel zien voor mannen. Dit heeft te maken met de andere variabelen die de logistische regressie gebruikt om de kans op recidive te voorspellen. Aan de hand van deze variabelen is het mogelijk om te voorspellen of een persoon een man of vrouw is. Dit gebruikt het algoritme om alsnog een onderscheid te kunnen maken tussen mannen en vrouwen.

Een andere manier om discriminatie in een getal uit te drukken is volgens de *selection rate*. De *selection rate* in dit geval geeft de verhouding aan tussen de percentages van mannen en vrouwen die geclassificeerd worden als recidivisten. De berekening voor de *selection rate* gaat als volgt:

$$\frac{505 / (955 + 505)}{46 / (298 + 46)} \approx 2,5$$

Deze 2,5 geeft weer dat mannen, na berekening door de logistische regressie, een 2,5 keer zo grote kans hebben om als een recidivist geclassificeerd te worden ten opzichte van vrouwen. Om te kijken of hier sprake is van onterechte discriminatie, gebruiken we de 80%-regel om te bepalen of er een kans is dat er onethisch gediscrimineerd wordt. Deze regel wordt door de overheid van Amerika gehanteerd om te bepalen of er mogelijk sprake is van discriminatie. Volgens de 80%-regel mag dit getal variëren tussen de 0,80 en de 1,25. Alles wat buiten dit gebied valt, kan mogelijk gezien worden als discriminatie. Het is natuurlijk mogelijk dat dit niet het geval is. Deze maatstaf is alleen een indicatie die aangeeft of verder onderzoek nodig is. (Workplace Dynamics, LLC, 2009)

Als laatste maatstaf kan er een Chi-kwadraat test gedaan worden over de kruistabel. Deze test geeft aan of er een verband is tussen de twee variabelen. In dit

geval zijn dat de beslissing van de logistische regressie en het geslacht. Als uit de Chi-kwadraat test blijkt dat er een verband is, zou dit in aanmerking komen voor discriminatie. In dit geval geeft ook de Chi-kwadraat test een p-waarde die vrijwel gelijk is aan nul en dit duidt op mogelijke discriminatie binnen het algoritme.

Met deze methode kunnen niet alleen logistische regressies, maar vele soorten algoritmes onderzocht worden op mogelijke discriminatie, zoals neurale netwerken en random forests. Deze methode richt zich niet alleen op discriminatie op basis van geslacht, maar is toepasbaar op bijvoorbeeld rassendiscriminatie en op leeftijdsdiscriminatie.

Samenvattend hebben we vier alternatieven gegeven om discriminatie door algoritmes inzichtelijk te maken. Door middel van een grafiek van de modelscore verdelingen, twee statistische toetsen, de AUC_D en de 80%-regel, kan er een compleet beeld gecreëerd worden van de hoeveelheid discriminatie. De resultaten van de maatstaven van één model kunnen dan weer vergeleken worden met die van een ander model om te kijken of één van de modellen minder discriminerend is dan de andere.

LITERATUUR

- Dataset: Broward County Clerk's Office, Broward County Sheriff's Office, Florida Department of Corrections, ProPublica. (2019). COMPAS Recidivism Risk Score Data and Analysis. ProPublica Data Store. <https://bit.ly/STAtOR2qoTjU>
- Calmon, F., Wei, D., & Vinzamuri, B., Varshney, K., (2017). *Optimized Pre-processing for Discrimination Prevention*. <https://bit.ly/STAtOR35WiBTY>
- Dastin, J. (2018). *Amazon scraps secret AI recruiting tool that showed bias against women*. <https://bit.ly/STAtOR2rXmYzG>
- Pedreschi, D., Ruggieri, S., & Turini, F., (2009). *Measuring Discrimination in Socially-Sensitive Decision Records*. <http://pages.di.unipi.it/ruggieri/Papers/sdm09.pdf>
- Pedreschi, D., Ruggieri, S., & Turini, F., (2012). *A Study of Top-K Measures for Discrimination Discovery*. <http://pages.di.unipi.it/ruggieri/Papers/sac2012.pdf>
- Workplace Dynamics, LLC., (2009). *Adverse Impact Analysis; Four-Fifths Rule*. <https://bit.ly/STAtOR2PaE3hn>
- Zemel, R., Wu, Y., Swerky, K., & Pitassi, T., (2013). *Learning Fair Representations*. <https://bit.ly/STAtOR2Rk4oOk>

STAN VAN LOON is in 2019 afgestudeerd bij de studie Toegepaste Wiskunde op de Hogeschool van Amsterdam. Tijdens zijn afstudeerstage deed hij, in samenwerking met Pegasystems, onderzoek naar een methode om discriminatie in algoritmes meetbaar te maken. Dit artikel is een samenvatting van het afstudeeronderzoek. Het hele onderzoek is te vinden via <https://bit.ly/STAtOR2P19V7Y>
E-mail: stanvanloon1998@gmail.com



2020

VVSOR ANNUAL MEETING

donderdag 12 maart 2020
in de Gertrudiskapel in Utrecht



Causality & Complexity

Het bestuur van de VVSOR hoopt op 12 maart 2020 veel leden en andere belangstellenden te ontmoeten op de VVSOR Annual Meeting in de Gertrudiskapel – In De Driehoek congres- en vergadercentrum in Utrecht. Het thema van de dag is Causality and Complexity. Verdere bijzonderheden kunt u vinden in de volgende editie van STAtOR en te zijner tijd op de website van de VVSOR. Zet de datum van 12 maart alvast in uw agenda.



GEDETAILLEERDE EN TIJDIGE STATISTIEKEN OVER DE NEDERLANDSE SAMENLEVING

Officiële statistieken gepubliceerd door statistiekbureaus zijn traditioneel gebaseerd op schattingsmethoden uit de steekproeftheorie die parametervrij zijn. Het nadeel van deze technieken is dat alleen met relatief grote steekproefomvangenvoldoend nauwkeurige cijfers kunnen worden gemaakt. Dit leidt ertoe dat met deze methoden alleen cijfers op een betrekkelijk hoog aggregatieniveau voor een vrij lange referentieperiode kunnen worden gepubliceerd. Recente ontwikkelingen maken modelgebaseerde schattingsmethoden steeds aantrekkelijker voor statistiekbureaus om meer gedetailleerde schattingen te maken die sneller beschikbaar komen omdat ze over een kortere verslagperiode gaan. Hierdoor wordt statistische informatie voor veel gebruikers en beleidsmakers een stuk relevanter. In dit artikel beschrijven we het gebruik en de ontwikkelingen van multilevel- en tijdreeksmodellen om officiële statistieken in meer detail en sneller te kunnen produceren.

JAN VAN DEN BRAKEL, HARM JAN BOONSTRA, SABINE KRIEG & MARC SMEETS

De taak van statistiekbureaus is het verzamelen en publiceren van betrouwbare statistische informatie over sociaaleconomische ontwikkelingen in een samenleving. Hierbij gaat het om variabelen die gedefinieerd zijn als totalen of gemiddelden. Bijvoorbeeld het percentage werklozen uitgesplitst naar regio of demografische ach-

tergrondkenmerken. Veel statistiekbureaus, waaronder het Centraal Bureau voor de Statistiek (CBS), verzamelen de benodigde data (voor zover niet beschikbaar via registers) via kanssteekproeven en maken gebruik van traditionele 'design'-gebaseerde schattingsmethoden voor het maken van statistische informatie. Deze methoden zijn

hoofdzakelijk gebaseerd op het kansmechanisme van het steekproefontwerp. Statistische modellen spelen hierbij geen of slechts een ondergeschikte rol. Dergelijke schatters corrigeren voor de ongelijke insluitkansen uit het steekproefontwerp en voor selectieve non-respons.

De precisie van deze schatters kan verbeterd worden door gebruik te maken van gerelateerde hulpinformatie waarvan de populatietotalen bekend zijn uit bijvoorbeeld registers. Deze schatters worden gemotiveerd met een lineair regressiemodel dat het verband tussen de doel- en hulpvariabelen beschrijft en zijn in de literatuur bekend als gegeneraliseerde regressie (GREG) schatters. Hoewel GREG-schatters worden gemotiveerd door een lineair regressiemodel is de methodiek design-gebaseerd omdat eigenschappen van de schatters, zoals verwachting en variantie, worden geëvalueerd met betrekking tot het kansmechanisme van het steekproefontwerp in plaats van het veronderstelde lineaire model.

In de literatuur zijn ook veel schatters voor eindige populatieparameters bekend die gebaseerd zijn op een statistisch model. Over het algemeen zijn statistiekbureaus, waaronder het CBS, terughoudend in het gebruik van deze modelgebaseerde schattingsmethoden. Bij veel statistiekbureaus is het uitgangspunt dat officiële statistieken, die worden gebruikt voor het nemen van beleidsmatige beslissingen, bij voorkeur zijn gebaseerd op empirie en data en zo min mogelijk op modelaannames waarvan de validiteit moeilijk verifieerbaar is.

Design-gebaseerde schatters zijn met name bij grote steekproefomvangenvolgeschikt voor het maken van statistische informatie omdat in deze situatie design-varianties klein zijn. Bij kleine steekproefomvangenvol zijn deze schatters vanwege grote design-varianties vaak onbruikbaar. De relevantie van statistische informatie neemt echter toe met de mate van detail, frequentie en tijdigheid van de cijfers. Cijfers op een gedetailleerd regionaal niveau op maandbasis zijn doorgaans relevanter dan cijfers op nationaal niveau op jaarbasis. Gedetailleerde uitsplitsingen in combinatie met een korte referentieperiode leiden snel tot situaties waarbij slechts weinig waarnemingen beschikbaar zijn. Dan heeft het de voorkeur om informatie te gebruiken uit andere bronnen, deelpopulaties of perio-

den, die via een statistisch model aan elkaar gerelateerd worden. Deelpopulaties of perioden waarvoor onvoldoende waarnemingen beschikbaar zijn om betrouwbare design-gebaseerde schattingen te maken worden kleine domeinen genoemd. Rao en Molina (2015) geven een uitgebreid overzicht van methoden voor het schatten over kleine domeinen.

Recente ontwikkelingen maken modelgebaseerde schattingsmethoden steeds aantrekkelijker voor statistiekbureaus. Naast de toenemende vraag naar gedetailleerde en tijdige cijfers is er een constante druk op statistiekbureaus om de kosten en lastendruk voor de berichtgevers te verlagen door meer gebruik te maken van informatie uit registers en bigdatabronnen.

Kleindomeinschatters voor de beroepsbevolking

Literatuur over kleindomeinschatters is hoofdzakelijk gebaseerd op multilevelmodellen die gebruik maken van cross-sectionele correlaties tussen domeinen met als doel de precisie van een domeinschatter te verbeteren met informatie uit andere domeinen. De meeste surveys op statistiekbureaus worden bovendien herhaaldelijk uitgevoerd. Tijdreeksmodellen zijn dan relevant omdat hiermee informatie uit voorgaande perioden gebruikt kan worden om de precisie van de schatting voor de laatste periode te verbeteren. Door tijdreeksen voor meerdere domeinen te modelleren in één multivariaat tijdreeksmodel ontstaat de mogelijkheid om zowel cross-sectionele als temporele correlaties te gebruiken om de precisie van domeinschatters te verbeteren.

Onderzoek naar het gebruik van dit soort modelgebaseerde schatters op het CBS volgde twee aanpakken. De eerste aanpak was het ontwikkelen van een cross-sectioneel multilevel model voor het maken van jaarcijfers over de beroepsbevolking op gemeenteniveau, gebaseerd op de Enquête Beroepsbevolking (EBB). Hiervoor wordt een model op persoonsniveau toegepast dat bestaat uit drie componenten. De eerste component is een regressiecomponent zoals bij de GREG-schatters waarin veel achtergrondkenmerken worden meegenomen. De tweede

component bestaat uit de gemeente-effecten. Vanwege de geringe hoeveelheid data in veel gemeenten worden de gemeente-effecten gemodelleerd in een tweede 'level' van het model, waardoor effectief informatie uit andere gemeenten wordt gebruikt om de schattingen voor een specifieke gemeente te verbeteren. De derde component is een ruisterm om de resterende variatie te verklaren. Het multilevel model wordt jaarlijks geschat en sinds 2015 worden hiermee de officiële gemeentelijke cijfers over de beroepsbevolking gemaakt.

In een tweede aanpak is een tijdreeksmodel ontwikkeld voor het schatten van maandcijfers over de beroepsbevolking op basis van de EBB waarbij gebruik gemaakt wordt van temporele correlaties. De EBB is een doorlopend onderzoek gebaseerd op een roterend panelontwerp. De respondenten worden in vijf peilingen waargenomen, steeds met een interval van één kwartaal. Dit survey ontwerp heeft drie problemen:

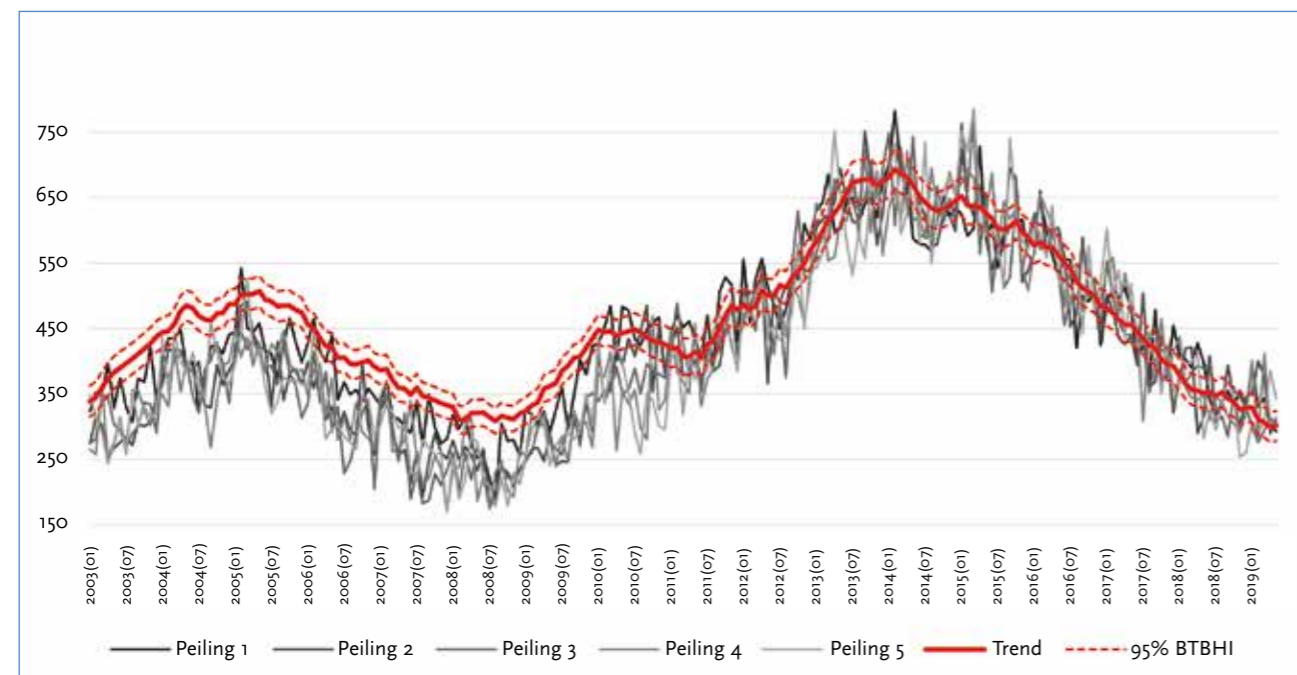
- De steekproefomvang is te klein om met de GREG-schatter voldoende precieze schattingen te maken voor maandcijfers over de beroepsbevolking.
- Er zijn systematische verschillen tussen de opeenvolgende peilingen. De werkloze beroepsbevolking wordt

bijvoorbeeld systematisch hoger geschat met de data van de eerste peiling ten opzichte van de daaropvolgende vier peilingen. Dit fenomeen wordt in de literatuur 'rotation group bias' (RGB) genoemd (Bailar, 1975).

- Door veranderingen in de manier waarop data worden verzameld en het veldwerk wordt uitgevoerd ontstaan er breuken in de waargenomen tijdreeksen.

Als oplossing voor deze drie problemen is een vijfdimensionaal structureel tijdreeksmodel ontwikkeld. Het rotatieschema van het panelontwerp impliceert dat iedere maand data worden verzameld in vijf onafhankelijke steekproeven: de steekproef die voor het eerst wordt waargenomen, de steekproef die drie maanden geleden is getrokken en voor de tweede keer wordt waargenomen, et cetera. Op basis hiervan kunnen iedere maand vijf onafhankelijke schattingen worden gemaakt voor de onbekende populatieparameter. De vijf tijdreeksen die op deze manier worden waargenomen vormen de input voor het tijdreeksmodel.

Het vijfdimensionale tijdreeksmodel bestaat uit vier componenten. De eerste component modelleert de onbekende populatieparameter als de som van een stochastische trend voor de laagfrequente variatie, een seizoen-



Figuur 1. Maandcijfers totale werkloze beroepsbevolking (in duizend) op nationaal niveau. De reeksen van de peilingen zijn beïnvloed door de breuken van 2010 en 2012. Voor de eerste peiling treden deze breuken op in januari 2010 en in april 2012. Voor de vervolppeilingen is dat 3 tot 12 maanden later

component voor de cyclische variatie en een witte ruis voor de onverklaarde hoogfrequente variatie. Via deze component wordt steekproefinformatie uit het verleden gebruikt om de precisie van de schattingen te verbeteren. De tweede component modelleert de RGB. Onder de aanname dat de waarnemingen verkregen bij de eerste peiling onvertekend zijn, wordt het niveauverschil tussen de tijdreeksen van de tweede, derde, vierde en vijfde peiling ten opzichte van de eerste peiling gemodelleerd met vier verschillende random walk processen. Deze component zorgt ervoor dat de schattingen voor de populatieparameter op het niveau van de tijdreeks van de eerste peiling liggen. De derde component modelleert de breuken ten gevolge van veranderingen in het surveyproces. De vierde component modelleert de steekproefruis waarbij rekening wordt gehouden met de autocorrelatie die ontstaat door de paneloverlap. Zie Van den Brakel en Krieg (2015) voor een technische verantwoording van deze schattingsmethodiek.

Sinds 2010 wordt het hierboven beschreven tijdreeksmodel gebruikt voor het schatten van maandcijfers over de beroepsbevolking. Hierbij gaat het om de totale werkloze en werkzame beroepsbevolking op nationaal niveau en een uitsplitsing naar de kruising van geslacht en leeftijd in drie categorieën. Figuur 1 illustreert de vijf inputreeksen van het tijdreeksmodel en de outputreeks voor de totale werkloze beroepsbevolking op nationaal niveau. De outputreeks is gecorrigeerd voor de breuken in 2010 en 2012 ten gevolge van veranderingen in de dataverzameling. Vóór 2012 schat het model de aantallen zoals deze geweest zouden zijn met de nieuwe waarnemingsmethode. Dit model reduceert de standaardfouten met circa 20% op nationaal niveau en circa 50% op het niveau van de zes domeinen.

Om naast temporele informatie ook gebruik te maken van cross-sectionele informatie kunnen de vijfdimensionale tijdreeksen voor de zes domeinen worden gecombineerd in één 30-dimensionaal structureel tijdreeksmodel. Op een vergelijkbare manier kan het model worden uitgebreid met hulpreeksen, bijvoorbeeld het aantal geregistreerde uitkeringsgerechtigden. Hiermee kan de precisie van de modelschattingen verder worden verbeterd. Zie Van den Brakel en Krieg (2016).

In Boonstra en Van den Brakel (2019) wordt een alternatieve aanpak gevolgd. In plaats van een state-space model dat wordt gefit met een Kalman filter, wordt een

structureel tijdreeksmodel voor de vijf peilingen in de twaalf provincies geschat met een Bayesiaans multilevel tijdreeksmodel via een MCMC-simulatie.

Kleindomeinschatters voor andere onderzoeken

Behalve bij de EBB worden kleindomeinschatters ook bij een aantal andere statistieken toegepast, of wordt daar op dit moment onderzoek naar gedaan.

Het Consumenten Conjunctuur Onderzoek (CCO) meet maandelijks het consumentenvertrouwen van de Nederlandse bevolking. Het consumentenvertrouwen wordt berekend op basis van vijf vragen over de financiële situatie van het eigen huishouden en de economische situatie van ons land. Naast het consumentenvertrouwen zelf worden ook tijdreeksen over deze vijf onderliggende vragen gepubliceerd. Sinds 2017 worden het consumentenvertrouwen en de onderliggende variabelen geschat met een structureel tijdreeksmodel. Dit model lijkt in grote lijnen op het model dat voor de maandcijfers over de beroepsbevolking toegepast wordt. Natuurlijk is het model aangepast om met de design-aspecten van het CCO rekening te houden.

Het Verplaatsingsonderzoek is een jaarlijks steekproefonderzoek waarin personen wordt gevraagd naar hun verplaatsingen op een bepaalde dag. Ook hier is behoefte aan gedetailleerde schattingen naar andere motief en vervoerwijze van de verplaatsing en naar persoonskenmerken zoals geslacht en leeftijd. Daarnaast worden trends over de tijd geschat. Hiervoor is recentelijk een multilevel tijdreeksmodel ontwikkeld waarmee zowel cross-sectionele als temporele verbanden worden gemodelleerd. Het model corrigeert bovendien voor breuken die zijn ontstaan door veranderingen in het surveyproces. Innovatief aan deze toepassing is dat ruim 500 tijdreeksen in één Bayesiaans multilevel tijdreeksmodel worden gecombineerd. Deze methodiek is in 2019 in gebruik genomen.

Met het Schoolverlatersonderzoek (SVO) worden verschillende aspecten van de aansluiting van werk en opleiding gemeten. Hierover worden gedetailleerde cijfers naar onderwijsinstelling of arbeidsmarktregio en opleidingsrichting gemaakt. Het SVO wordt de afgelopen jaren onder alle mbo-schoolverlaters uitgevoerd. Omdat slechts een deel van hen uiteindelijk respondeert, moeten

de ontbrekende gegevens worden bijgeschat. Hiervoor wordt, net zoals voor de gemeentelijke jaarcijfers over de beroepsbevolking, een multilevel model op persoonsniveau toegepast. In dit geval bevat het model naast een uitgebreide set van registervariabelen meerdere gemiddelde effecten voor indelingen naar school, regio en opleiding.

Het toepassen van kleindomeinschatters bij bedrijfsstatistiek is vaak complexer. Hierbij moet rekening gehouden worden met grote schaalverschillen tussen kleine en grote bedrijven. In dit geval is de doelvariabele scheef verdeeld. Daarnaast kan de doelvariabele in sommige surveys vaak 0 zijn. Dit is het geval bij de investerings- en R&D-statistiek. In de afgelopen jaren is onderzocht hoe modelgebaseerde domeinschatters gebruikt kunnen worden voor het regionaliseren van deze bedrijfsstatistiek. Hierbij worden schattingen gemaakt per COROP-gebied en gemeente. De resultaten zijn veelbelovend en op dit moment onderzoeken we samen met de betrokken afdelingen de mogelijkheden om deze schatters in productie te nemen.

Toekomstige ontwikkelingen

Zoals aangegeven in de inleiding, neemt de relevantie van statistische informatie toe met de mate van detail, frequentie en snelheid waarmee cijfers beschikbaar komen. Modelgebaseerde domeinschatters bieden de mogelijkheid om meer relevante statistische output te produceren zonder de dataverzamelingskosten extravagant te laten stijgen. Het ligt dan ook voor de hand dat het CBS door zal gaan met het verder ontwikkelen van dergelijke schattingsmethodieken.

Deze schattingsmethodieken bieden ook de mogelijkheid om efficiënt gebruik te maken van gerelateerde informatie uit zogenaamde bigdatabronnen. Een belangrijk aspect van dit soort nieuwe databronnen is dat ze vaak sneller en met een veel hogere frequentie beschikbaar komen dan herhaald waargenomen steekproefonderzoeken. Google trends kunnen bijvoorbeeld op dag- of weekbasis worden afgeleid. Een potentiële toekomstige toepassing is om reeksen gebaseerd op surveys te combineren met snel beschikbare hulpreeksen uit alternatieve databronnen. Hierdoor ontstaat de mogelijkheid om in real time nauwkeurige voorlopige schattingen te maken voor de doelvariabelen van het steekproefonderzoek. Eerste resultaten daartoe zijn beschreven in Van den Brakel et al. (2017) voor het consumentenvertrouwen en Schiavoni et al. (2019) voor de maandcijfers van de beroepsbevolking.

LITERATUUR

- Bailar, B. A. (1975). The effects of rotation group bias on estimates from panel surveys. *Journal of the American Statistical Association*, 70, 23–30.
- Boonstra, H. J., & Brakel, J. A. van den. (in press). Estimation of level and change for unemployment using structural time series models. *Survey Methodology*.
- Brakel, J. A. van den, & Krieg, S. (2015). Dealing with small sample sizes, rotation group bias and discontinuities in a rotating panel design. *Survey Methodology*, 41, 267–296.
- Brakel, J. A. van den & S. Krieg (2016). Small area estimation with state-space common factor models for rotating panels. *Journal of the Royal Statistical Society Series A*, 179, 763–791.
- Brakel, J. A. van den, Söhler, E., Daas P., & Buelens B. (2017). Social media as a data source for official statistics; the Dutch Consumer Confidence Index. *Survey Methodology*, 43, 183–210.
- Rao, J. N. K., & Molina, I. (2015). *Small Area Estimation*. Wiley, New York.
- Schiavoni, C., Palm, F., Smeekes, S., & Brakel, J. A. van den. (2019). *A dynamic factor model approach to incorporate Big Data in state space models for official statistics*. Discussion paper January, 2019. Heerlen: Statistics Netherlands.

JAN VAN DEN BRAKEL heeft biometrie gestudeerd aan de Universiteit van Wageningen en werkt momenteel als methodoloog bij het CBS en als bijzonder hoogleraar Survey Methodology bij de vakgroep Kwantitatieve Economie van de Universiteit van Maastricht. Zijn belangrijkste onderzoeksgebieden zijn steekproeftheorie, ontwerp en analyse van proeven, tijdreeksanalyse, kleindomeinschatters en het gebruik van big data in de productie van officiële statistiek. Email: jbrl@cbs.nl

HARM JAN BOONSTRA heeft natuurkunde gestudeerd aan de Rijksuniversiteit Groningen en heeft daar ook zijn promotieonderzoek gedaan. Hij werkt sinds 1999 als onderzoeker bij het Centraal Bureau voor de Statistiek en heeft zicht gespecialiseerd in Bayesiaanse schattingsmethoden toegepast op de officiële statistiek. E-mail: hbta@cbs.nl

SABINE KRIEG heeft wiskunde gestudeerd in Jena (Duitsland). Nadat ze een aantal jaren onderzoek gedaan heeft op wiskundig gebied op de universiteiten van Essen (Duitsland) en Groningen, werkt ze sinds 1997 als methodoloog op het CBS. Ze is gespecialiseerd in steekproeftheorie, modelgebaseerd schatten en seizoenscorrectie. E-mail: skrg@cbs.nl

MARC SMEETS heeft wiskunde gestudeerd aan de Technische Universiteit Eindhoven en heeft daar ook promotieonderzoek gedaan. Hij werkt sinds 2001 als methodoloog bij het Centraal Bureau voor de Statistiek en is gespecialiseerd in modelgebaseerde schattingsmethoden, steekproeftheorie en steekproefcoördinatie. E-mail: mset@cbs.nl



Foto: Ben Bender CC

WISKUNDE OP DE WALLEN

Of hoe goede bedoelingen wel eens onverwacht kunnen uitpakken

Eind jaren tachtig van de vorige eeuw was ik oprichter en eerste voorzitter van de Europese gebruikersclub van een Amerikaans softwarepakket voor wiskunde en statistiek. Na een eerste meeting in Frankfurt waren de Amerikanen erg tevreden over de input van hun gebruikers en werd besloten tot een tweede bijeenkomst in Amsterdam. Aan mij de eer een en ander te regelen.

Door de vele goede connecties binnen de VVS lukte het mij om het Trippenhuys aan de Kloveniersburgwal te huren voor deze conferentie. Deze eerbiedwaardige zetel van de Koninklijke Nederlandse Academie van Wetenschappen was indertijd behoorlijk exclusief en de ongeveer 100 deelnemers keken hun ogen uit.

Zoals bij de meeste conferenties was er ook een hoofdkwartier geregeld, een hotel waar de organisatoren verbleven en waar op de avond voor de start een welkomstborrel werd geschonken. Dat was het Victoriahotel, pal tegenover het station. Hotels geven grote kortingen op hun kamerprijzen als je als organisatie een flink contingent afneemt, dus het merendeel van de deelnemers verbleef daar ook.

De Amsterdamse Wallen vormden in die tijd nog een behoorlijke *no go area*, in het bijzonder de Zeedijk. Een groot deel van de handel in harddrugs speelde zich daar af. Zelfs politieagenten liepen er altijd met minimaal twee man tegelijk en liefst meer. Nu zijn wiskundigen in het dagelijks leven weleens wereldvreemd en rechtlijnig in

hun denken. Ik was dan ook bang dat men op de kaart zou kijken en een rechte lijn zou trekken van Victoriahotel naar Trippenhuys: die kortste weg viel nagenoeg samen met de Zeedijk.

Om te voorkomen dat mijn deelnemers, al keuvelend over zaken als optimalisatie van software om matrices te inverteren op vectorcomputers, in de problemen zouden komen had ik mijn voorzorgen genomen. Iedereen kreeg in het informatiepakket dat van tevoren was toegezonden een kaartje van de Amsterdamse binnenstad. Daarop was, met uitleg waarom, de minder gevaarlijke route via Damrak, Dam en Hoogstraat duidelijk aangeven.

De conferentie verliep voortreffelijk, er was een goede sfeer en voor zover bij de organisatie bekend werd niemand beroofd of lastig gevallen. De Amerikanen bedankten mij opgelucht voor de goede voorzorgen, ze kenden de reputatie van het Wallengebied en waren toch wel benauwd geweest dat ze aansprakelijk gehouden zouden worden bij incidenten.

Maar de mooiste dank kreeg ik bij het afscheid van twee Duitse deelnemers: zij hadden nog nooit een conferentie meegemaakt waar de organisatie zo duidelijk en expliciet had aangegeven waar zij *the nice ladies* konden vinden.

GERRIT STEMERDINK is eindredacteur van STAtOR. E-mail: gjsterdink@hotmail.com

STATISTICI TEGEN ALCOHOLMISBRUIK

Negentiende-eeuwse statistici waren vaak arts, advocaat, onderwijzer of predikant, maar ook wel ambtenaar of hoogleraar. Zij bekommerden zich beroepshalve, maar ook vanuit een overtuiging, om de medemens. Zij namen misstanden waar, publiceerden daarover en hoopten op een luisterend oor bij beleidsmakers. Zij ontpopten zich als woordvoerders van de verenigingen die zich bemoeiden met het onderwijs en het drankmisbruik.

RONALD VAN DER BIE

Al in het allereerste nummer van het *Tijdschrift voor staathuishoudkunde en statistiek* (1841) stond een groot artikel over 'Het misbruik van sterken drank. Geen oorzaak, maar gevolg.' De onbekende auteur – waarschijnlijk de oprichter van het blad, mr. B. W. A. E. baron Sloet tot Oldhuis – nam daar de stelling voor zijn rekening dat drankmisbruik het gevolg was van pauperisme. De arme dronk om de troosteloosheid van zijn leven te vergeten. Sloet had de statistiek aan zijn kant: zo hoog als het verbruik rond 1825 was, bijna 10 liter pure alcohol per hoofd, zo hoog is het daarna nooit meer geweest.

Drankgebruik beteugelen

De verbanden waren hem al langer bekend. Om dezelfde reden dat onder de minst welvarende volken het meest werd gedronken, zag hij ook dat in de meest verarmde plaatsen het aantal tapperijen het grootst was. Om het

drankgebruik te beteugelen stelde hij de oprichting van matigheidsgenootschappen naar Amerikaans voorbeeld voor. Wetgeving om het drankgebruik te reguleren, was op dat moment geen reële optie, beducht als de regering was dat de schatkist accijnsinkomsten zou mislopen. Omwille van de schatkist wilde de overheid volgens de schrijver 'liever op de zedeloosheid der natie blijven voort speculeren.'

Niet onbetwist, wel relevant

Het CBS heeft altijd een grote expertise over het drankvraagstuk in huis gehad. Henri Methorst (1868-1955), oprichter en directeur-generaal van het CBS, sprak in 1900 op een congres in Parijs over het alcoholverbruik in ons land in de negentiende eeuw. De drankconsumptie in de periode 1831-1899 heeft hij op basis van de accijnsgegevens geschat. Het is een van de oudste tijdreeksen van



Man verkondigt de afschaffing van sterke dranken aan mensen op straat. Houtgravure naar Isaac Weissenbruch, circa 1910. Collectie: Rijksmuseum Amsterdam

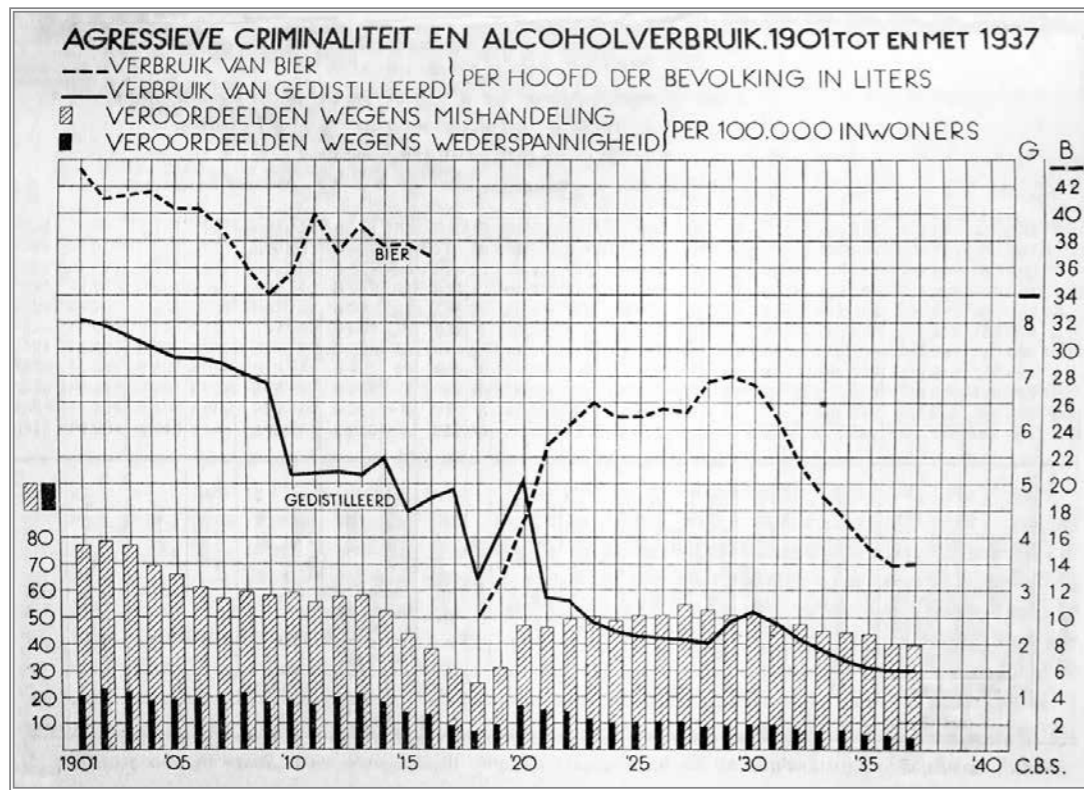
de sociale statistiek. De daling van het verbruik die hij constateerde, schreef hij toe aan drie factoren. Allereerst de invoering van de Drankwet (1881) die het aantal alcoholverkooppunten beperkte. Daarnaast de toename van de sportbeoefening, die met matigheid gepaard ging. En ten derde het succes van de georganiseerde drankbestrijding, die hij onderdeel noemde van een breder *réveil social*. Daarmee waren volgens hem maatschappelijke ontwikkelingen – bestaanszekerheid, vast werk, een vast leefpatroon, het opkomende verenigingsleven – meer dan wetgeving er verantwoordelijk voor dat het alcoholgebruik daalde.

Na 1920 werd het CBS verantwoordelijk voor de publicatie van de nieuwe statistiek van het verbruik van gedistilleerd. Deze statistiek stond als niet erg betrouwbaar bekend. Het ministerie van Binnenlandse Zaken was lang verantwoordelijk voor de publicatie maar is in 1900, vanwege de communicatieproblemen, gestopt met de uitgave. Drankbestrijders, beleidsmakers en onderzoek-

kers zijn echter altijd om de verbruiksgegevens blijven vragen. Dankzij de tussenkomst van het CBS is de verbruiksstatistiek nog twee keer verschenen, in aangepaste vorm. Onbetwist waren de cijfers nog steeds niet, maar Methorst was doordrongen van het feit dat het maatschappelijke debat gediend was met relevante, niet per se onbetwiste statistieken.

Blaue front

De nieuwe verbruiksstatistiek verscheen in de glorie dagen van de georganiseerde drankbestrijding. Dit blauwe front had in de eerste decennia van de twintigste eeuw jaren grote successen geboekt. De eerste was de aanpassing van de Drankwet in 1904. De nieuwe wet gaf gemeenten de mogelijkheid de uitgifte van vergunningen voor de verkoop van sterke drank te beperken en het aantal verkooppunten vast te stellen tot onder het wettelijk toe-



Grafiek alcoholverbruik 1901-1937. Bron: CBS

gestane maximum. In 1919 is een blauwe staatscommissie gevormd die moest onderzoeken of het mogelijk was de productie van consumptiealcohol te beperken zonder de industrie te knakken. De commissie wilde van elke branderij en stokerij weten hoeveel ze produceerden, om vervolgens te kunnen bepalen hoeveel de productie kon worden beperkt. Bedrijven weigerden evenwel gegevens te verstrekken en de commissie stond met lege handen. Ten einde raad benaderde die het CBS. Het bureau had gedetailleerde gegevens over de productie van gedistilleerd verzameld voor de productiestatistiek die spoedig gepubliceerd zou worden, maar het stond die gegevens niet af. De Wet op de Statistiek van voortbrenging en verbruik stond dat domweg niet toe. Methorst had de ondernemers die zijn vragenlijst moesten invullen herhaaldelijk de garantie op geheimhouding gegeven. Hij kon de commissie dus niet helpen.

Tinbergen toont rol afschaffersbeweging

In de jaren dertig en veertig hebben diverse CBS-onderzoekers modelanalyses uitgevoerd van het bier- en gedistilleerdverbruik. In 1942 was het Jan Tinbergen die rekende aan het alcoholverbruik. Wat Methorst nog gaf als verklaring voor de daling van het verbruik, kon Tinbergen aantonen als een factor in een mathematisch model. Het verbruik was in dat model een functie van de prijs van het product, het inkomen van de consument, de gemiddelde temperatuur en een trendfactor, waarin de invloed van de drankbestrijding naar voren kwam. Deze trendfactor

berekende hij door de invloeden van prijs, inkomen en het weer van de verbruikscurve af te trekken. Tinbergen becijferde die daling op 1,6 procent per jaar (1925-1935).

Na de oorlog nam het alcoholverbruik aanvankelijk slechts langzaam toe. Na 1970 steeg het razendsnel, van 5,5 liter per hoofd (1970) tot 9,1 liter in 1979. Na 1980 daalde de alcoholconsumptie vrijwel jaarlijks tot 7,0 liter per hoofd in de jaren 2014-2016. Met een dergelijk verbruik behoort ons land tot de matigste lidstaten van de Europese Unie. De strategieën en middelen waarmee dit is bereikt, waren geen andere dan in de negentiende eeuw, al zijn het niet meer de dominees en de onderwijzers die de weg wijzen. Het was een combinatie van wetgeving (hoge accijnzen, beperking van de drankverkoop aan jongeren), overheidscampagnes die wezen op de gezondheidsrisico's en sociale schade van alcoholmisbruik, het permanente morele appel op matigheid ('drank maakt meer kapot dan je lief is') en voorlichting, bijvoorbeeld op scholen.

Dit artikel is eerder verschenen in de CBS-publicatie *Het licht van de statistiek* ter gelegenheid van 120 jaar CBS.

RONALD VAN DER BIE studeerde geschiedenis en economie en promoveerde in 1995 aan de Vrije Universiteit Amsterdam. Sinds 1996 werkzaam als wetenschappelijk redacteur bij het Centraal Bureau voor de Statistiek te Den Haag. Hij publiceerde onder andere over economische geschiedenis van Nederland (historische nationale rekeningen), tabaksteelt en de (sociale) geschiedenis van het CBS.
E-mail: rj.vanderbie@cbs.nl



UITMUNTENDE MASTER'S OF PH.D. THESIS BEGELEID?

OPROEP OM KANDIDATEN TE NOMINEREN VOOR DE JAN HEMELRIJK EN WILLEM R. VAN ZWET AWARDS 2019

Ter bekroning van een uitzonderlijke afstudeerprestatie aan een Nederlandse instelling voor wetenschappelijk onderwijs/hoger beroepsonderwijs looft de VVSOR al vanaf 1989 een scriptieprijs uit. In 2014 kreeg deze de naam *Jan Hemelrijk Award*. Sinds 2012 is er ook een prijs voor dissertaties: de *Willem R. van Zwet Award*.

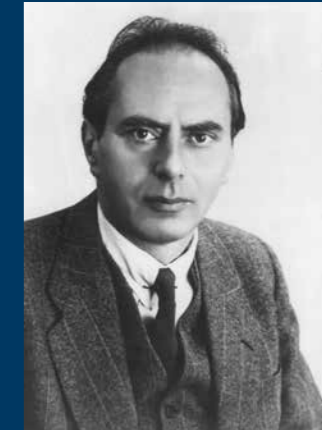
De VVSOR roept op tot nominaties voor deze prijzen. Beide prijzen bestaan uit een oorkonde en een geldbedrag van 1000 Euro. De prijzen zullen worden uitgereikt tijdens de annual meeting van de VVSOR, op 12 maart 2020. Genomineerd kunnen worden personen die van september 2017 tot en met augustus 2019 zijn afgestudeerd respectievelijk gepromoveerd en die nog niet eerder zijn genomineerd.

Hierbij worden supervisors (begeleiders) opgeroepen om een uitmuntende afstudeerscriptie (Master) of dissertatie (Ph.D.) te nomineren voor de Jan Hemelrijk dan wel Willem R. van Zwet Award 2019.

De indiening van een nominatie dient vergezeld te gaan van een aanbevelingsbrief van de supervisor van de genomineerde. De precieze procedure voor beide prijzen, alsmede de reglementen en het nominatieformulier zijn te downloaden op de website van de VVSOR, www.vvsor.nl. De nominatie dient uiterlijk 19 januari 2020 binnen te zijn.

Namens de VVSOR,
Prof. dr. Eric Cator, juryvoorzitter Jan Hemelrijk Award
Prof. dr. Mark de Rooij, juryvoorzitter Willem R. van Zwet Award
Dr. Sander Scholtus, Secretaris der beide jury's

Oproep voor nominaties VAN DANTZIGPRIJS 2020



De VVSOR reikt eenmaal in de vijf jaar de Van Dantzig Prijs uit, de hoogste prijs in de Nederlandse Statistiek en Operations Research. Deze prijs is ingesteld ter nagedachtenis aan prof. dr. David van Dantzig (1900 – 1959), de grondlegger van de mathematische statistiek in Nederland. De eerstvolgende uitreiking zal plaatsvinden tijdens de Annual Meeting van de VVSOR in 2020. De prijs bestaat uit een in brons gegoten medaille met de beeltenis van Van Dantzig in reliëf.

Genomineerden voor de Van Dantzigprijs 2020 zijn op 1 januari 2020 niet ouder is dan 40 jaar en hebben de afgelopen vijf jaar een zeer noemenswaardige bijdrage – hetzij theoretisch, hetzij praktisch van aard – geleverd aan de statistiek of operations research. De nominatie dient uiterlijk 19 januari 2020 binnen te zijn.

Eerdere winnaars van de Van Dantzigprijs zijn Willem van Zwet (1970), Ton van Meurs (1975), Arie Hordijk (1980), Alexander Rinnooy Kan (1985), Richard Gill (1990), Geert Ridder (1995), Aad van der Vaart (2000), Sem Borst & Mark van der Laan (2005), Peter Grünwald & Harry van Zanten (2010) en Bert Zwart (2015).

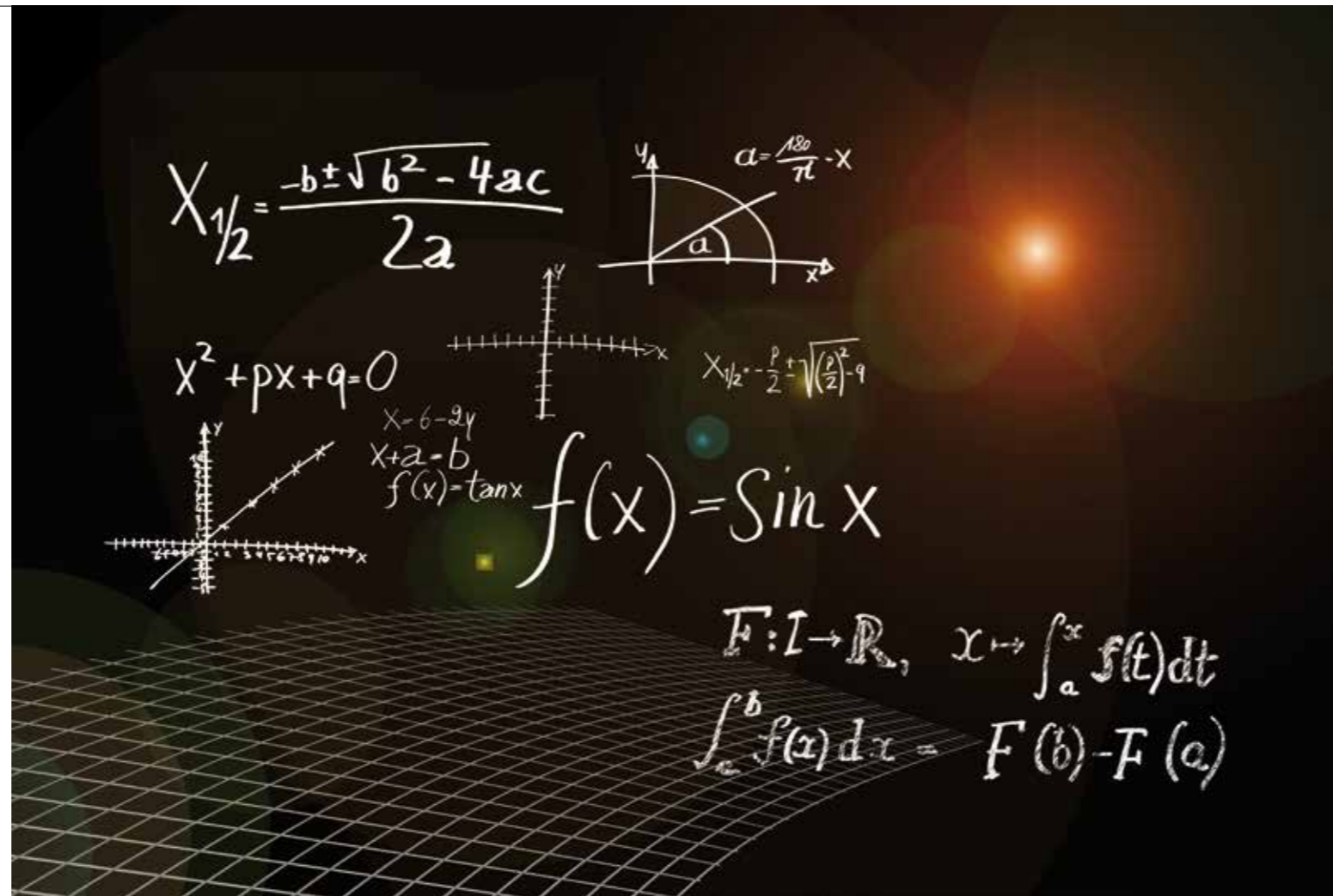
Voel u vrij om deze oproep te verspreiden naar mogelijk geïnteresseerden. Nominaties en vragen over het proces kunnen worden gericht aan het dagelijks bestuur van de VVSOR, c.q. de jury voor de prijs, via [<vandantzig@vvsor.nl>](mailto:vandantzig@vvsor.nl).

KANSREKENING, een echt wiskundevak voor havo en vwo

Op 10 oktober 2019 heeft de coördinatiegroep Curriculum.nu de voorstellen voor herziening van de landelijke kerndoelen en eindtermen voor het primair en voortgezet onderwijs overhandigd aan Arie Slob, de minister voor Basis- en Voortgezet Onderwijs. Ontwikkelteams bestaande uit leraren en schoolleiders werden ingesteld voor negen leergebieden waaronder het leergebied rekenen en wiskunde.

Curriculum.nu heeft vanaf het begin veel weerstand ontmoet vanuit het veld, mede omdat vakinhoudelijke verenigingen van leraren niet of nauwelijks betrokken werden bij de werkzaamheden van de ontwikkelteams. Zo kwam het ontwikkelteam voor het leergebied Rekenen & Wiskunde in de media zwaar onder vuur te liggen door haar voorstel om drastisch te snijden in het breukenonderwijs op de basisschool en het rekenen met breuken te verschuiven naar het voortgezet onderwijs. Behalve de vermeende moeilijkheidsgraad van breukenonderwijs op de basisschool, was het argument van het ontwikkelteam rekenen en wiskunde dat op deze wijze ruimte in de bovenbouw van de basisschool vrijgemaakt kon worden voor statistische data-analyse. Een weinig zinnig voorstel. Op de basisschool hoort statistiek niet thuis. Kinderen op de basisschool missen daarvoor de rijpheid en ook basisschoolleerkrachten zijn niet toegelegd op het onderwijzen van statistiek. De basisschool is er om de leerling de basisbewerkingen van het rekenen op gedegen wijze bij te brengen inclusief het rekenen met breuken. Wel is het zo dat een klein beetje kansrekening op de basisschool het breukenonderwijs zou verrijken.

Wat havo en vwo betreft was het oorspronkelijke voorstel van het ontwikkelteam om veel meer aandacht te besteden aan statistische technieken dan momenteel het geval is en alle leerlingen – dus niet alleen in wiskunde A – te onderwijzen in hypothese-toetsing, p -waardes en significantieniveaus. De VVSOR heeft op dit voorstel een heldere en degelijk onderbouwde reactie gegeven <<http://bit.ly/STATOR-2ONuODB>>; een reactie die haar uitwerking op het eindvoorstel van het ontwikkelteam re-



kenen en wiskunde niet gemist heeft. De bottom line van de VVSOR reactie is: statistiek van een meer geavanceerd karakter kun je beter bewaren tot op hbo en universiteit. Wel op havo en vwo aandacht voor statistische data-analyse en voor wat je statistische geletterdheid zou kunnen noemen: begrijpen wanneer getallen in een context wel of niet kloppen of betrouwbaar zijn, denk aan misleidende grafieken om statistische boodschappen over te brengen en aan de betrouwbaarheid van statistische gegevensverzameling door peilingen. Onderwerpen die ook heel goed in samenhang met lessen economie of maatschappijleer behandeld kunnen worden. Ook als de voorstellen van Curriculum.nu naar de prullenbak worden verwezen, is het document van VVSOR over de rol van kansrekening en statistiek in het voortgezet onderwijs heel nuttig om hierover discussie te voeren samen met de Nederlandse Vereniging voor Wiskundeleraren.

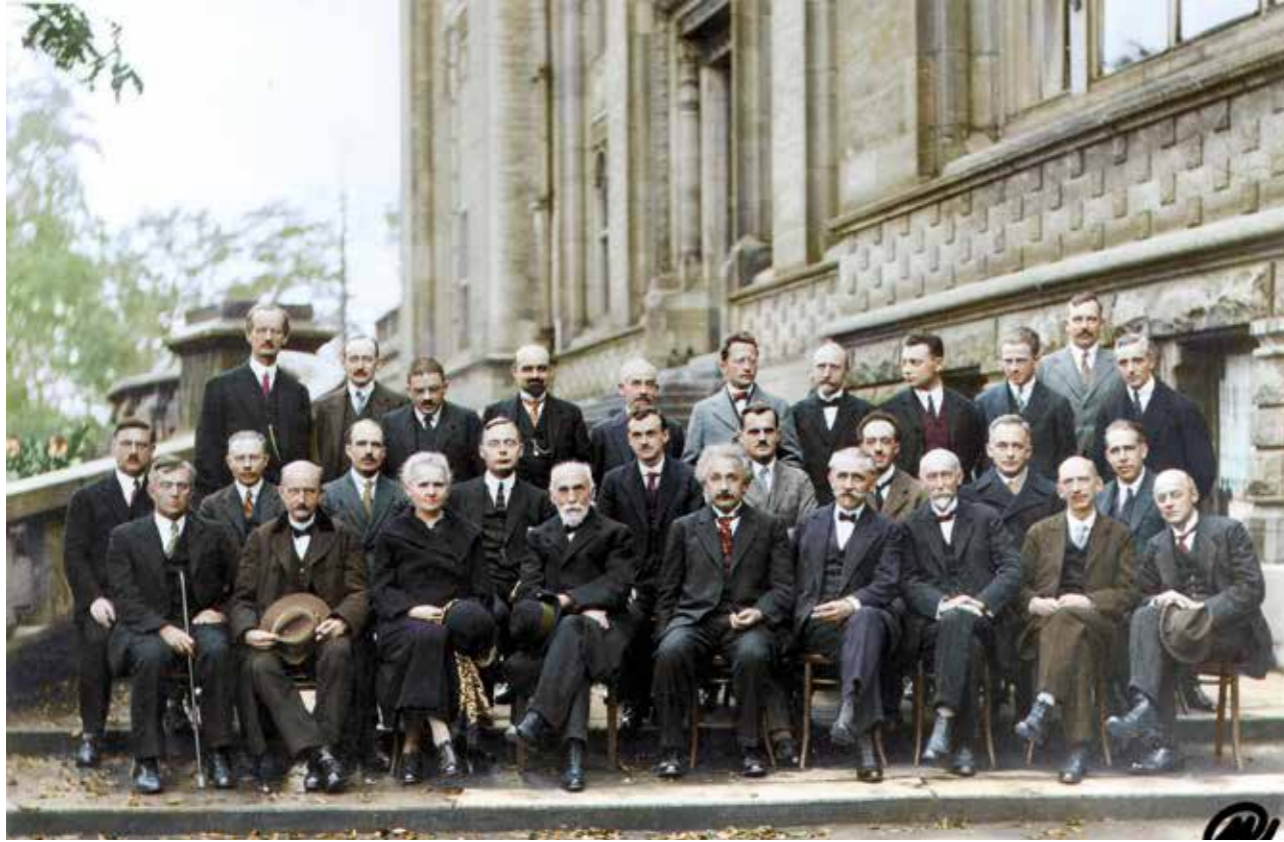
In deze column wil ik ten slotte mijn persoonlijk standpunt geven ten aanzien van kansrekening en statistiek op havo en vwo: laat de statistiek over aan hbo

en universiteit, maar voer wel kansrekening in als een verplicht wiskundevak voor alle leerlingen in de bovenbouw van havo en vwo. De invulling van statistiek op hbo en universiteit is sterk afhankelijk van de studierichting, maar heeft wel kansrekening als gemeenschappelijke basis. Een dergelijk standpunt werd in feite al verkondigd in de jaren zeventig door de legendarische wiskunde-didacticus Hans Freudenthal die op de bladzijden 610 en 613 van zijn standaardwerk *Mathematics as an Educational Task* schreef '... pervade all mathematics by probability at an early stage – as soon as the children get to know about fractions ...', en '... I simply do not understand the philosophy of those who propose to teach highschool children a lot of statistical techniques ...'

Kansrekening is een echt wiskundevak met axioma's en stellingen die volgen uit de axioma's en definities, waar statistiek meer een apart vak is met een eigen specifieke inhoud en aanpak. Evenals de Euclidische meetkunde is kansrekening een wiskundevak dat de leerling in aanraking brengt met logisch redeneren en dit doet met

boeiende en motiverende problemen en toepassingen. Voor de meer exacte en technische studierichtingen aan hbo en universiteit is kansrekening een essentiële schakel in het curriculum, maar ook in de opleiding van juristen en medici is enige basale kennis van (Bayesiaanse) kansrekening geen overbodige luxe. Al met al voldoende redeneren om op havo en vwo kansrekening als een verplicht wiskundevak in te voeren. Wel als een vak dat op een moderne leest geschoeid is. Geen overmatig gebruik van combinatoriek, maar wel veel aandacht voor conditionele kansen en Bayesiaans redeneren. Uiteraard mag in het vak de behandeling van belangrijke discrete kansverdelingen zoals de binomiale verdeling en de Poisson-verdeling niet ontbreken evenals de behandeling van de normale verdeling met daaraan gekoppeld de statistische z-score methode en de centrale limietstelling. Verder zou ik ervoor willen pleiten om in een moderne opzet van het vak kansrekening ook computersimulatie van kansproblemen te betrekken, niet alleen uit didactisch oogpunt nuttig maar ook nuttig als oefening in modelmatig en algoritmisch denken. Het is waar dat kansrekening moeilijker te onderwijzen is dan andere wiskundevakken op de middelbare school. Kansrekening kent geen kookboekrecepten maar vereist creatief en flexibel denken om voor het beschouwde probleem de juiste aanpak te vinden. Zelfs ervaren kansrekenaars zien niet altijd meteen de juiste aanpak. Bij invoering van kansrekening als verplicht vak op havo en vwo zou ik ervoor willen pleiten om het vak niet in het centraal eindexamen op te nemen maar alleen voor het schoolexamen te laten tellen. Kansrekening is typisch een vak dat het best tot zijn recht komt en het meest leerzaam is als het niet in het nauwe keurslijf van eindexametoetsen wordt gedwongen. Het is een wiskundevak waar leerlingen in hun verdere leven veel aan zullen hebben.

HENK TIJMS is emeritus hoogleraar operations research aan de Vrije Universiteit en auteur van diverse leerboeken over operations research en kansrekening.
E-mail: h.c.tijms@xs4all.nl



De meest beroemde Solvay-conferentie was de vijfde, zij werd gehouden in oktober 1927; 17 van de 29 aanwezigen kregen ooit een Nobelprijs. Er werd vooral gediscussieerd over de nieuwe kwantumtheorie. Einstein (middenvoor) en Bohr (tweede rij rechts) discussieerden over de implicaties van het onzekerheidsprincipe van Heisenberg (derde van rechts bovenste rij). De historische foto is ingekleurd

HEISENBERG

over statistiek, onzekerheid en filosofie

RICHARD STARMANS

De vaststelling dat de Duitse natuurkundige, filosoof en Nobelprijslareaat Werner Heisenberg (1901-1976) in de ideeëngeschiedenis een onwrikbare positie inneemt, klinkt welhaast als een gemeenplaats. In de jaren twintig van de vorige eeuw groeide hij uit tot één van de pioniers en grondleggers van de kwantummechanica. Het ontstaan van deze fysische theorie in de jaren 1900-1930 vormt – op de wetenschappelijke revolutie van de 17e eeuw na – waarschijnlijk de meeste bestudeerde en gedocumenteerde periode uit de wetenschapsgeschiedenis. Desondanks, of wellicht juist daardoor, is een rigide, onomstreden afbakening of demarcatie van die beginperiode geen sinecure. Gemakshalve identificeren we deze ‘vroege jaren’ hier met de periode 1913–1927, die ruwweg begon met Bohrs atoommodel (1913) en ‘eindigde’ met de Kopenhaagse interpretatie van de kwantummechanica door Bohr en Heisenberg en de beroemde Solvay-confe-

rentie in 1927. Het moge duidelijk zijn dat de Deense fysicus Niels Bohr (1885–1962) bij dit alles in menig opzicht als scharnier- en ijkpunt fungeerde, maar de opkomst van Heisenberg en de betekenis van zijn vroege werk kunnen het beste tegen deze achtergrond worden bezien. Al op zeer jonge leeftijd maakte Heisenberg deel uit van de intellectuele voorhoede van theoretisch fysici, wat vaak leidde tot unieke vormen van samenwerking met uiteenlopende fameuze tijdgenoten. Zo gaf hij samen met de eveneens Duitse fysici Max Born (1882–1970) en Pascual Jordan (1902–1980) gestalte aan de befaamde *Dreimännerarbeit*.

In zijn vriend Wolfgang Pauli (1900-1958) vond Heisenberg een intellectuele antagonist met wie hij een bijkans levenslange tweespraak voerde. Legendarisch werd evenwel zijn succesvolle, maar ook moeizame samenwerking met Niels Bohr, die jarenlang de rol van mentor

en in zekere zin van mecenas vervulde. In enkele jaren verschenen in rap tempo vele theoretische bijdragen, die direct of indirect aan Heisenberg cum suis kunnen worden toegeschreven en die gebaseerd waren op een door nagenoeg alle betrokkenen gedeelde kritiek op Bohrs oorspronkelijke atoommodel; de matrixmechanica, de probabilistische interpretatie van de Schrödingervergelijking, de onzekerheid- of onbepaaldheidsrelatie en de genoemde Kopenhaagse interpretatie.

Door de uitvoerige documentatie van dit alles werd al spoedig zichtbaar dat de reikwijdte van Heisenbergs werk die van de natuurkunde in engere zin ruim oversteepte. Zo heeft het belangrijke implicaties voor de verhouding tussen statistiek en natuurkunde, voor de wetenschapsfilosofie in het algemeen, en bovenal blijkt veel van zijn wijsgerige oeuvre tot op de dag van vandaag uiterst relevant. Het wordt dan ook nog steeds bestudeerd, heruitgegeven en vertaald. Nog onlangs publiceerde uitgeverij Lemniscaat het boek *Het deel en het geheel* (2019), een vertaling van Heisenbergs precies vijftig jaar geleden verschenen bundel *Der Teil und das Ganze* (1969). Het boek, vertaald en ingeleid door Maarten van Buuren, bevat een selectie van essays, voordachten, korte aantekeningen en transcripties van gesprekken die Heisenberg voerde met beroemde tijdgenoten over natuurkunde, politiek, maar ook over de filosofische problemen die het tijdsgewricht waarin hij leefde beheersten. De bloemlezing biedt een unieke blik op de intellectuele duizendpoot Heisenberg en zijn wetenschappelijke, filosofische en levensbeschouwelijke drijfveren. Enige aspecten van Heisenbergs fysische en niet-fysische werk worden in dit essay kort voor het voetlicht gebracht.

Genealogie van de onzekerheid

Hedendaagse historici en filosofen van de statistiek als Ian Hacking, Stephen Stigler, Lorenz Kruger en Nancy Cartwright zijn het erover eens dat de Probabilistische Revolutie – die zich in de tweede helft van de 19e eeuw aandeede – mede door Heisenberg, Born en Bohr in een stroomversnelling geraakte, om niet te zeggen voltooid werd. (Kruger, 1987, 1991). Vrijwel alle wetenschappen ondergingen de afgelopen honderd jaar een probabilistische wending, dat wil zeggen hun methoden, technieken en concepten werden gaandeweg steeds meer gebaseerd op kansrekening en statistiek. Sommige van hun kern-

begrippen, theorieën en modellen hebben enkel in een probabilistische context nog betekenis of relevantie. Een volgende stap is dan de gedachte dat uiteindelijk ook het achterliggende wereldbeeld probabilistisch wordt, waarbij onzekerheid niet langer in laplacianse zin louter epistemisch is, maar ook ontologisch dient te worden geduid. Het door Hacking gepostuleerde proces van een *erosion of determinism* zou zijn voltooid en onzekerheid is irreducibel geworden. Dat wil zeggen, dat de onzekerheid van de kennis voortkomt uit wezenlijke kenmerken van de structuur van de werkelijkheid en niet is te wijten aan onze vermogens om kennis te verwerven, zoals beperkte waarneming, opslagcapaciteit, rekenkracht en uitdrukingskracht van de taal. Dit alles zou voor een belangrijk deel door het hierboven genoemde werk van Heisenberg cum suis tot stand zijn gebracht. Een dimensie van onzekerheid die hier nadrukkelijk op de voorgrond treedt is uiteraard die van onbepaaldheid, het gegeven dat een eigenschap of toestand (nog) niet vastligt of is gerealiseerd c.q. ook niet onafhankelijk van de waarnemer is te bepalen. Plaats, impuls, tijd en energie zijn louter probabilistisch gedefinieerd.

Tegen de visie van Hacking c.s. op het voltooien van de probabilistische revolutie is natuurlijk wel het een en ander in te brengen, al was het maar het opwerpen van de klassieke vraag of dergelijke veranderingen niet veeleer evolutionair dan revolutionair moeten worden begrepen. Bovendien ondervond de Kopenhaagse interpretatie, die beoogde enkele problemen rond de golf-deeltjes dualiteit en het meetprobleem van de kwantummechanica het hoofd te bieden, behoorlijke weerstand onder fysici, soms op filosofische gronden. Concurrerende benaderingen als de *many-worlds*-interpretatie en de *hidden-variable*-theorie dienden zich aan. Daarnaast kan worden opgemerkt dat de relatie tussen statistiek en natuurkunde natuurlijk niet nieuw was. Reeds een halve eeuw eerder hadden Gibbs, Maxwell en Boltzmann gebruik gemaakt van de ‘oude’ statistiek van Laplace en Quetelet bij de kinetische gas-theorie en statistische mechanica. De katalysator kwam juist met Plancks kwantificering van de energie, waarbij hij teruggreep op Boltzmann en de thermodynamica. Einstein paste het principe in 1905 toe in zijn paper over het foto-elektrisch effect, Bohr in 1913 bij zijn atoommodel, dat kan worden beschouwd als het vertrekpunt voor de hier geschetste ontwikkelingen. Tot slot moet worden aangetekend dat vanuit methodologisch perspectief de door Claude Shannon (1916–2001) eigenhandig in gang



(vanaf links) Werner Heisenberg in 1925; Werner Heisenberg in Leipzig, omstreeks 1930; op skivakantie met Felix Bloch in 1933. Collectie: B. Blum-Heisenberg

In 2001 bracht Duitsland een postzegel uit ter gelegenheid van de honderdste geboortedag van Werner Heisenberg

gezette fluwelen revolutie van data in context en flux als het feitelijke sluitstuk van voornoemde revolutie kan worden beschouwd (Starmans, 2017). Hoe dan ook, duidelijk is in ieder geval dat het denken over onzekerheid een lange genealogie kent in de westerse ideeëngeschiedenis en een moeizaam emancipatieproces heeft doorlopen. De conceptie van onzekerheid werd lange tijd nogal pejoratief geduid; uitgerekend dankzij (de wisselwerking van) statistiek en natuurkunde werd voornoemd proces in een aantal achtereenvolgende stappen voltooid (Starmans, 2014). De door Hacking c.s. genoemde fysische invloed is dan ook evident.

Eenheid in verscheidenheid

Belangrijker dan het voeren van een vermoeiend debat over een prioriteitstelling is het kiezen van een invalshoek die minder door annexatiedrang lijkt te worden gevoed en die de vanuit historisch-wijsgerig perspectief zeer dwingende verbondenheid tussen fysica en statistiek kan benadrukken. Vertrekpunt hierbij is de dikwijls gekoesterde en geprotegeerde opvatting dat uiteindelijk slechts een ultieme of hogere waarheid gedacht kan worden, welke in de loop van de geschiedenis binnen de filosofie, de godsdienst, de politiek of anderszins binnen de maatschappelijke orde, op verschillende wijzen gestalte heeft gekregen. Binnen de wetenschap manifesteert dit denkbeeld zich onder meer in het streven naar de eenheid van kennis: eenheidswetenschap, unificatie, integratie en synthese. Of te wel, zoveel mogelijk verschijnselen verklaren

vanuit een en dezelfde theorie, zoveel mogelijk theorieën onderbrengen in een en dezelfde supertheorie of 'theorie van (bijna) alles', en – niet in de laatste plaats – zoveel mogelijk disciplines herleiden tot een en dezelfde discipline, die dan dikwijls als de meest abstracte wetenschap en van de weeromstuit ook als de meest fundamentele wordt beschouwd. Dit alles vindt zijn oorsprong nadrukkelijk in aloude wijsgerige preoccupaties met het zoeken naar een Eerste Filosofie, die de grondslagen legt of ten minste onderzoekt voor alle andere disciplines, een streven dat teruggaat tot Aristoteles, Descartes en Kant en dat zelfs nog in de 20e eeuw bij Edmund Husserl is terug te vinden. Ook binnen de wetenschappen zelf zijn vele pogingen ondernomen een vorm van reductionisme te bewerkstelligen. De wijze waarop zulks gestalte zou moeten krijgen heeft in de ideeëngeschiedenis een rijkgeschaakt pallet aan visies opgeleverd, waarvan er ten minste twee hier relevant zijn. De eerste betreft de natuurkunde, die vanuit een opvatting van fysicalistisch reductionisme de meest fundamentele wetenschap is en achtereenvolgens scheikunde, biologie, neurowetenschap, (cognitieve) psychologie en sociologie insluit. Via brugwetten en correspondentieregels kan de gewenste integratie worden tot stand gebracht. De unificatie is primair ontologisch, en anders gezegd, verticaal gericht. Tegelijkertijd kan ook de statistiek sedert de probabilistische revolutie aanspraak maken op de status van unificator, aangezien alle wetenschappen, zoals reeds betoogd, een probabilistische wending hebben ondergaan. Hier is de integratie niet primair of noodzakelijk ontologisch, maar allereerst methodologisch en horizontaal gericht. Zeker in tijden

van *data science* vormt deze status van de statistiek inclusief het paradigma van *statistical learning* volgens velen ten minste een krachtige narratief. Zo beschouwd zijn fysica en statistiek dus concurrerend, of milder geformuleerd complementair waar het gaat om de oude zoektocht naar eenheid in verscheidenheid. Met enige goede wil kan worden gesteld dat in het werk van Heisenberg en Bohr beide op zeer specifieke wijze samenkomen. Bohrs cruciale betekenis in de onderhavige problematiek betreft juist zijn streven op verschillende momenten de rol van unificator op te eisen, al bleek hij voor veel filosofen veeleer een dwaallicht, dat een wijsgerige verwarring veroorzaakte, die bijna honderd jaar later nog steeds aanleiding geeft tot debat (Starmans, 2014). Wie *Het deel en het geheel* ter hand neemt zal constateren dat Heisenberg nog een stap verder gaat en als een Platonist pur sang ook in metafysische zin een queeste naar de diepere orde van de dingen onderneemt, die sterk aan James Joule doet denken en nog meer aan Max Planck, wiens werk als Archimedisches beginpunt van het ontstaan van de kwantummechanica kan worden opgevat (Starmans, 2017)

De matrixmechanica en Stiglers Wet

De wetenschapsgeschiedenis kent een schier onuitputtelijke lijst met uiteenlopende resultaten, die alle getooid gaan met de naam van de al dan niet vermeende ontdekker, uitvinder of bedenker ervan; de constante van Planck, het getal van Avogadro, de wetten van Kepler, het vermoeden van Fermat, de these van Church, de Turing-machine,

Kripke-modellen, de Bernoulli-verdeling, de Gauss-kromme, de theorie van Wegener, etcetera. En, in het licht van de onderhavige thematiek: Pauli's principe, de Schrödinger-vergelijking, Bohrs model en Heisenbergs onzekerheidsrelatie. Opvallend is niet alleen de veelsoortigheid van de resultaten, die alle als wetenschappelijke *output* worden erkend en te boek gesteld, maar zeker ook het gemak en de vanzelfsprekendheid waarmee deze exclusief met één persoon worden verbonden. Zulke eenduidige typering of eponiemen onttrekken onmiskenbaar de namen van vele erflaters van onze beschaving aan de vergetelheid, maar doen zelden volledig recht aan de complexe en weerbarstige historische werkelijkheid. De reeds genoemde Amerikaanse historicus van de statistiek Stephen M. Stigler (1941) gaat nog een stap verder en stelt onomwonden dat wetenschappelijke vondsten altijd ten onrechte aan een persoon worden toegeschreven: 'No scientific discovery is named after its original discoverer.' Zonder schroom benoemt hij deze boude uitspraak als 'Stigler's Law of Eponymy' in een gelijknamig artikel uit 1980. Door de evidente en destructieve zelfreferentie van de wet, moeten we Stiglers actie veeleer als een kwinkslag en knipoog beschouwen, dan als een bedenkelijk gebrek aan bescheidenheid. Om elk misverstand hierover uit te sluiten benadrukt Stigler bovendien dat alle argumenten die hij aandraagt voor zijn 'Wet van de verkeerde vernoeming' zijn terug te vinden in het werk van de in die tijd toonaangevende wetenschapshistoricus Robert K. Merton (1910–2003). In klassieke publicaties als *Science, Technology and Society in Seventeenth Century England* (1938) en *The Sociology of Science* (1973) kiest Merton een externalistische benadering. Hij beschouwt de wetenschap als sociaal systeem en onderzoekt de wisselwerking tussen het wetenschappelijk bedrijf en cultuur en samenleving; de *stakeholders*, de instituties, de waarden. Daarbij besteedt hij vooral aandacht aan receptie, waardering en beloningssystemen. Aan de hand van uiteenlopende historische studies ontrafelt hij vele misverstanden, anomalieën en paradoxen uit de wetenschapsgeschiedenis op een voor de onbevungen *scientist* vaak ontnuchterende wijze. Zo toont hij aan dat sommige fundamentele concepten en ideeën vrijwel unaniem worden geaccepteerd en overgenomen terwijl de uitvinder ervan wordt vergeten of bewust genegeerd. Tegelijkertijd vindt dikwijls een soort epische verdichting plaats. Met name in zijn artikel *The Matthew Effect in Science: The reward and communication systems of science* uit 1968 betoogt Merton dat beroemde wetenschappers vaak ten onrechte krediet krijgen voor hun bijdragen, terwijl de lagere Olympiërs het omgekeerde ervaren. Met de term

Matthew-effect verwijst hij openlijk naar de passage uit het Nieuwe Testament (Matteüs 25:14-30), door economen nogal eens geciteerd ter illustratie van het verschijnsel dat de rijken steeds rijker en de armen steeds armer worden. Merton veraanschouwt niet alleen dat prioriteitsgeschillen tussen wetenschappers op twijfelachtige wijze worden beslecht, maar gaat ook uitvoerig in op het verschijnsel eponymie, waarvan hij de vele manifestaties onderzoekt en waarin hij een gelaagde structuur onderkent. Hoe dan ook, de juistheid van Stiglers visie wordt in het werk van Heisenberg op verschillende manieren op treffende wijze geïllustreerd en herbevestigd. Zijn matrixmechanica uit 1922 is daarvan een saillant voorbeeld. Bohrs atoommodel uit 1913 was in zekere zin buitengewoon 'aanschouwelijk'. Elektronen bewogen in banen om de kern als in een klein Copernicaans zonnestelsel met concentrische cirkels. Uitgaande van stationaire toestanden kan een verandering van baan als een kwantsprong worden getypeerd, waarbij de frequentie van de straling evenredig is met het verschil in energie van de voornoemde toestanden. Omdat het steeds lastiger werd om nieuwe kwantumverschijnselen in te passen en met name kwantumovergangen te beschrijven met dit model, besloot Heisenberg zich in de traditie van Ernst Mach te beperken tot direct waarneembare grootheden en fysische eigenschappen als matrix te representeren. Dat kon echter pas nadat Born in Heisenbergs berekeningen een niet-commutatieve vermenigvuldiging had herkend zoals bij de matrixvermenigvuldiging die hij zich nog herinnerde van toen hij wiskunde en natuurkunde studeerde. Ook de rol van Pascual Jordan bleek cruciaal bij de uiteindelijke totstandkoming. Juist vanwege deze succesvolle *Dreimännerarbeit* beschouwde Heisenberg het als een historische onjuistheid dat Born en Jordan niet meedeelden in de Nobelprijs, die hem in 1932 werd toegekend. Ook in *Het deel en het geheel* wordt fraai geschetst hoe nieuwe ideeën ontkiemen en tot stand komen, soms informeel, via toevallige ontmoetingen, hoe ze worden ontwikkeld en verspreid en hoe uiteindelijk de door Merton geschetste perverse beloningssystemen in het gunstigste geval tot epische verdichting kunnen leiden, maar in andere scenario's tot geschiedvervalsing of zelfs bedrog.

Heisenberg en de filosofie

In het begin van de 20e eeuw was de verhouding tussen filosofie en natuurkunde nog betrekkelijk harmonieus, ondanks – of wellicht doordat – relativiteitstheorie en kwantummechanica grote consequenties hadden voor

de aanschouwelijkheid van het wereldbeeld, de vertrouwde categorieën van ruimte, tijd en causaliteit. Vele fysici waaronder Einstein, Bohr en Planck waren evenals voorgangers Ernst Mach en Max Planck filosofisch zeer onderlegd en legden zich erop toe de wijsgerige implicaties van hun werk te doorgronden. De betreurde controverse tussen de filosoof Henri Bergson en Einstein in de vroege jaren twintig was in dit opzicht veeleer een incident (Starmans, 2015). Die filosofische dimensie bleek ook op de genoemde Solvayconferentie, waarbij 29 genodigden aanwezig waren, waarvan er zeventien de Nobelprijs hadden gewonnen of later zouden winnen. Hier werd de status van de kwantummechanica definitief gevestigd, maar het was ook een culminatiepunt van de intellectuele spanningen tussen enerzijds Albert Einstein, Perrin en andere denkers die vaak als wetenschappelijk realist worden aangeduid en anderzijds het kamp van Niels Bohr, die niet geheel terecht doorgaans wordt gekarakteriseerd als instrumentalist in de traditie van P.W. Bridgeman, de latere Nobelprijswinnaar wiens *The Logic of Modern Physics* eveneens in 1927 verscheen. Meer nog dan zijn illustere voorgangers en tijdgenoten begreep Heisenberg de implicaties van de soms paradoxale en onbegrijpelijke aspecten van de kwantummechanica en zocht nadrukkelijk de dialoog met een breder intellectueel publiek. Zijn *Physics and Philosophy; the Revolution in Modern Science* uit 1962 werd een moderne klassieker. Ook de reeks gesprekken die hij in 1962 en 1963 voerde met Thomas Kuhn, samen met Karl Popper de beroemdste wetenschapsfilosoof van de twintigste eeuw, leidde tot een serie belangwekkende tijdsdocumenten. (Kuhn, 1962, 1963). Heisenbergs had kennis genomen van Kuhns *The Structure of Scientific Revolutions* uit 1962 en stelde tegenover Kuhns paradigmatische opvattingen over 'gesloten systemen'. Hoe dan ook anomalieën en paradoxen die om een verklaring riepen, waren er volop in de kwantummechanica. De opvatting dat de snelheid of impuls en de plaats van een deeltje nooit beide exact op hetzelfde moment kunnen worden bepaald; dat de nauwkeurigheid waarmee de ene grootte wordt gemeten omgekeerd evenredig is met de nauwkeurigheid waarmee de andere, 'complementaire' grootte wordt gemeten. Het inzicht dat een deeltje 'verandert' louter doordat het wordt waargenomen, dat noch het verleden noch de toekomst daarmee 'vastlagen' en het determinisme voorgoed voorbij lijkt. Dat alles is vanuit een klassieke epistemologie problematisch. Heisenberg zag ook hoe de wisselwerking tussen de meetproblematiek en golfdeeltjes dualiteit ertoe leidde dat Bohr stevast verkeerd werd begrepen. Al eerder had deze in zijn principe van comple-

mentariteit uitgedrukt dat golven en deeltjes twee zijden van c.q. perspectieven op hetzelfde verschijnsel vormen en met zijn correspondentieprincipe aangegeven dat bij grote kwantumgetallen de klassieke mechanica van de macroscopische wereld toereikend is. Hij ging echter in zekere zin verder dan Heisenberg, die alleen stelde dat onzekerheid vooral door de genoemde energieoverdracht werd veroorzaakt. Bohr beschouwde waarnemer en experiment als een enkel, samenhangend systeem: met het meten wordt het systeem vastgelegd. Een niet-gemeten deeltje is als het ware onbepaald, heeft geen geschiedenis, etc. Bohr zou volgens criticasters het aloude principe van intelligibiliteit ondermijnen, de natuurlijke categorie van oorzaak-gevolg-relaties opheffen en de poort openen naar subjectivisme, holisme en Oosterse filosofie. De intuïtie van het kenbare stond onder druk. De kloof tussen filosofie en natuurwetenschap leek daarmee te worden verdiept, terwijl Bohr zich juist bekommerde om de consequenties voor het wetenschappelijk wereldbeeld en een zinvolle interpretatie nastreefde door in een conceptuele analyse onder meer de beperkingen van de op de macroscopische wereld gebaseerde natuurlijke taal te analyseren. Anders dan vaak wordt beweerd was hij veeleer een wetenschappelijk realist 'zonder aanschouwelijkheid' dan een instrumentalist, en al helemaal geen relativist. Heisenberg begreep beter dan zijn tijdgenoten wat er hier op het spel stond inzake de teloorgang van de aanschouwelijkheid van het wereldbeeld, de grenzen van de kennis en vooral het zoeken naar eenheid en een diepere orde. Al deze aspecten lopen als een rode draad door *Het Deel en het Geheel* en komen onder meer naar voren in essays als 'Kwantummechanica en Kant' (1930), 'Begrijpen in de natuurkunde' (1920), 'Kwantummechanica en een gesprek met Einstein' (1925), 'Natuurwetenschap en religie' (1927) of 'Elementaire deeltjes en platoonse filosofie' (1961).

Epiloog

Heisenberg leefde en werkte in een zeer bewogen tijdsgewricht. Tijdens de Tweede Wereldoorlog heeft hij in tegenstelling tot vele andere wetenschappers en intellectuelen Duitsland niet verlaten. Zijn rol bij het Duitse atoomprogramma, zijn al dan niet vermeende pogingen dit af te remmen of zelfs te obstrueren, alsmede zijn eigen getuigenissen over deze periode worden met enige regelmaat weer opgerakeld en leiden dan tot enig debat. Ook zijn relatie met Bohr zou hierdoor zijn beïnvloed en hun enigszins mysterieuze ontmoeting in Kopenhagen

in 1941 leidde tot een persoonlijke breuk, maar gaf bovendien aanleiding tot vele speculaties. Het inspireerde de Britse dramaturg Michael Frayn tot het schrijven van het toneelstuk *Copenhagen* (1998), waarin hij Heisenberg, Bohr en diens vrouw Margrethe jaren later laat terugkijken op die ontmoeting en een viertal mogelijke versies of 'hypothetische scenario's' van die gebeurtenis uitwerkt. Het stuk werd ook in Nederland door achtereenvolgens het Noord Nederlands Toneel (1999) en het Nationale Toneel (2009) op de planken gebracht. Laatstgenoemde productie zou bovendien de Toneel Publieksprijs 2009 in de wacht slepen.

Hoe dan ook, dit alles heeft uiteindelijk slechts een geringe smet op Heisenbergs reputatie geworpen. Maarten van Buuren heeft dan ook beslist gelijk als hij de filosofie van Heisenberg eerst en vooral typeert als een poging de orde der dingen te vatten in de beste wijsgerige traditie, waarbij het hele spectrum van ontologie, metafysica, epistemologie en ethiek aan de orde komt. Dat de Duitse natuurkundige het begrijpen c.q. herstellen van de fysische orde zag als een eerste, noodzakelijke stap om de waanzin van het era waarin hij leefde te begrijpen en de politieke, maatschappelijke orde te herstellen, kan wellicht worden geduid als een vorm van hooggestemd utopisme, maar vloeit hoe dan ook rechtstreeks voort uit zijn rijkgeschakeerde filosofie.

LITERATUUR

- AIP (1962, 1963). Interviews of Werner Heisenberg by Thomas S. Kuhn and John L. Heilbron (Transcript: Niels Bohr Library & Archives). College Park, MD: American Institute of Physics (AIP). <https://www.aip.org/history-programs/niels-bohr-library/oral-histories/4661-1>
- Heisenberg, W. (1958). *Physics and Philosophy: The Revolution in Modern Science*. New York: Harper.
- Starmans, R. J. C. M. (2011). Models, Inference and Truth; Probabilistic Reasoning in the Information Era. In Mark J. Van der Laan & Sherri Rose (Eds.) *Targeted Learning; Causal Inference for Observational and Experimental Data* (pp. 1-20). Springer Series in Statistics. New York: Springer.
- Starmans, R. J. C. M. (2014). De weg naar Kopenhagen; statistiek, natuurkunde en onzekerheid. *STATOR*, 15(2), 18-22.
- Starmans, R. J. C. M. (2015). Over tijd en relativiteit; Einstein versus Bergson. *Filosofie*, 25(4).
- Starmans, R. J. C. M. (2017). Van Heraclitus tot Shannon: de fluwelen revolutie van data in context en flux. *STATOR*, 18(1), 22-31.

RICHARD STARMANS is verbonden aan de Faculteit Bèta-wetenschappen (Department of Information and Computing Sciences) van de Universiteit Utrecht en aan Tilburg University. Hij doet onderzoek op het snijvlak van filosofie, statistiek en informatica.
E-mail: starmans@cs.uu.nl

RANGLIJSTJES

‘Duisternis en Computergestuurde Willekeur’

Het is dringen geblazen aan de top van menig ranglijstje; de onderlinge verschillen zijn miniem en liggen steeds vaker binnen de foutmarges van de gehanteerde meet-systemen. Nog een dik half jaar en minimaal de halve wereld is in de ban van de Olympische Zomerspelen. De verhitte strijd om de olympische startbewijzen lijkt een prima voorbereiding op de medaillejacht in het dan tropisch hete Tokio.

Nou ja, de halve wereld. In 2016 in Brazilië deden 205 ‘normale’ landen mee; Koeweit was toen ‘abnormaal’ en mocht deelnemen onder de vlag Onafhankelijk (sic!) Olympisch Deelnemer om te onderstrepen dat Koeweitse sporters werden gemanipuleerd door de politiek. Verder was er nog een team van landloze vluchtelingen. Ook nobel. Een kleine rekensom leert dat het International Olympic Committee van mening was dat er 224 normale landen zijn. Immers, het Committee beweert volgens wiki dat 76 landen nog nooit een medaille hebben gescoord terwijl 148 landen dat wel deden. Inderdaad, $76 + 148 = 224$. Duidelijk. Dat van ‘minimaal de halve wereld’ slaat maximaal op de omvang van de wereldbevolking. Straks in Tokio gaan een dikke 1.000 atleten de medailles verdelen op meer dan 300 sportonderdelen. *All over the world* valt er veel te juichen.

Nog even iets meer over de Spelen in 2016. Van de 11.551 atleten in Rio kwamen er 242 uit Nederland. Op de medaillelandenranglijst, de landen alfabetisch (eigenlijk $\alpha\beta\gamma$ met α =goud, β =zilver, γ =brons) gerangschikt, staat Nederland op plek 11. Niet gek toch? Hoewel, even rekenen! Australië, een sportland waar we ons graag aan spiegelen, staat op 10 met 29 medailles en 421 deelnemers. Even over de duim betekent dit dat 6,89% van de deelnemende Aussies met metaal huiswaarts keerde. En voor Nederland, met z’n 19 medailles en 242 deelnemers, is deze effectiviteitsscore 7,85%, een tikkie beter dan de *down unders*. Ik ga nog even door. België, het land waarvoor we graag de lachspiegel hanteren: 108 atleten en 6 medailles en dus een effectiviteit van 5,56%. Duidelijk. Maar dan die Duitsers: 422 atleten en 42 plakken, plek 5 op die landenranglijst en een effectiviteit van 9,95%, bijna 1 op de 10 Duitse atleten won olympisch Rio-metaal. De Belgen hebben nog een lange weg te gaan. Er zit overigens wel een vertekening in deze getallen, want meerdere atleten gingen met meerdere medailles naar huis, plus dat we

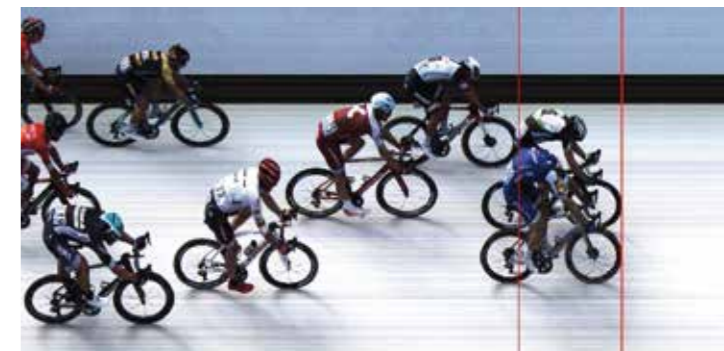
goud, zilver en brons als gelijkwaardig hebben gerekend.

Eigenlijk zegt deze rekenarij niks. Er moet fors worden gecorrigeerd op van alles en nog wat. In de startblokken dus voor een eerlijke effectiviteitsranglijst. Eerlijk, effectiviteit? Maar dan wel eerst die begrippen helder definiëren. Op zich *doable*, maar voor een emeritus nauwelijks uitdagend. Leuk voor een promotie-onderzoek en een bijpassende beursaanvraag? Een promotievoorstel over ranglijstjes dus, met als doel de beantwoording van vragen zoals: Wie of wat staan of staat er straks bovenaan; Is de ranglijst ‘eerlijk’; Hoe gebruik je de ranglijst voor het selecteren van olympische atleten. En de onontkoombare validatievraag: Wat is de rol van de ranglijst in de samenleving? En zo kun je doorgaan: Is de club of de atleet, die aan het einde van de competitie bovenaan staat, ook de beste? Goeie vraag. Waarom zou dat niet zo zijn? De beste ben je toch als je, volgens de met elkaar afgesproken regels, bovenaan staat. Pas op, dat is iets te kort door de bocht. Immers, als meerdere clubs of atleten gelijk eindigen, wat dan? Oké, dan formuleren we aanvullende eisen, tot er voldoende onderscheid is en de kampioen gehuldigd kan worden. Maar zelfs als de volksliederen hebben geklonken, kan het maar zo zijn dat doping na acht jaar of zo roet in de bidon gooit en de nummer twee van de eindranglijst winnaar wordt. Nog een verhaaltje nu het toch over bidons gaat. Eric Breukink in 2012 over zijn Giro in 1989: ‘De Giro van 1989 had ik moeten winnen. Op een klimmetje van drie kilometer kreeg ik een hongerklop. Daar verloor ik de Giro. Na die rit viel ik terug naar de achtste plek. Uiteindelijk werd ik vierde. Daaruit blijkt dat ik toen de beste van het peloton was, op dat ene momentje na.’ Niet de beste, zou ik zeggen; dat ‘ene momentje’ heeft elke verliezer. *Ex aequo*: soms zelfs vier atleten op het hoogste trede van het huldigingspodium. Dat gebeurde bij het WK Turnen 2015 in Glasgow. En als de fictieve finishlijn op de finishfoto van de zevende Tour de France-etappe in 2017 ietsje anders was geprojecteerd, dan was niet Marcel Kittel maar Edvald Boasson Hagen winnaar geweest.

Crisis, zou ik zeggen. Steeds vaker liggen onderlinge topprestatieverschillen binnen de foutmarges van de meet-systemen. Zo worden de eindtijden in de schaats-sport gemeten in duizendsten van seconden, terwijl iedereen weet dat de tijdmetingen een foutmarge hebben van drieduizendste van een seconde. Zelfs olympi-



De toestelfinale op brug bij WK turnen in Glasgow in 2015 had vier winnaressen opgeleverd; de Chinese Fan Yilin, de Russinnen Viktoria Komova en Daria Spiridonova en de Amerikaanse Madison Kocian scoorden alle vier 15,366 punten



Marcel Kittel heeft in 2017 de zevende etappe in de Ronde van Frankrijk op zijn naam geschreven; als de fictieve finishlijn anders was geprojecteerd zou Edvald Boasson Hagen de winnaar zijn geweest

sche-selectiebeslissingen kunnen te maken hebben met verschillen binnen de foutmarge.

Het is dringen geblazen aan de top, niet alleen bij turnen, wielrennen en schaatsen. Grafieken van wereldrecords tonen nog nauwelijks ontwikkeling; er is volop sprake van verzadiging van topprestaties. Steve Haake, Davis James en Leon Foster lieten in 2015 in hun artikel in *Journal of Sports Sciences* fraai zien hoe in de atletiek op vrijwel alle technische nummers de prestatielimiet is bereikt. Daar komt bij dat prestatie-uitschieters zo zeldzaam geworden zijn dat de media veel te gemakkelijk de dopingkaart trekken en het publiek zelfs urineert over voorbijrijdende uitblinkers. Dit laatste overkwam viervoudig tourwinnaar Chris Froome tijdens de Tour van 2015. Bah.

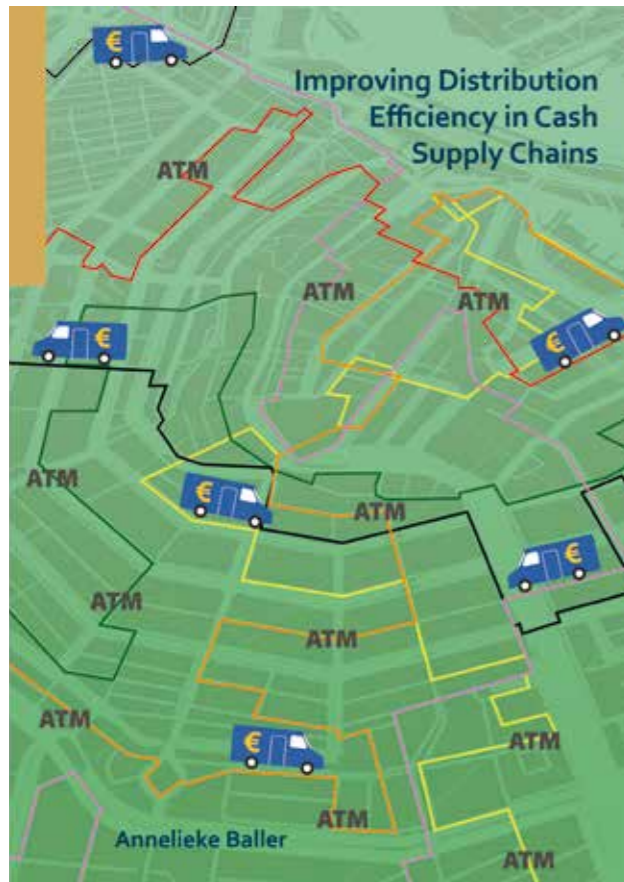
Op de dag van het Gronings Ontzet, op woensdag 28 augustus 2019, ‘bezieet Marleen Kamperman de wereld’ in een interessante column in *de Volkskrant* met als titel ‘Beurzen verdelen met de bingomolen’. Marleen is hoogleraar scheikunde in Groningen en begint haar column als volgt: ‘Ik zit in een commissie. We mogen beurzen uitdelen, geld voor onderzoek. Acht wetenschappers maken we blij. Acht van meer dan dertig gegadigden.’ Dat is dus een effectiviteitsscore van ongeveer 25%. Kom daar maar eens om als Olympiër. Hoeveel personen in de commissie zaten vertelt Marleen niet. Wel maken ze allemaal een voorkeurenranglijstje. Marleen: ‘Onze lijstjes gaan in de computer, er wordt een rekenmodel [in de volksmond heet dat tegenwoordig een algoritme] op losgelaten, er komt een generale rangschikking uit. De verschillen zijn klein, zo klein dat een beetje statisticus de vloer zou aanvegen met wat voor conclusie dan ook ...’ Volgens Marleen Kamperman ‘heerst er duisternis en computergestuurde willekeur’. Haar één na laatste alinea begint met: ‘Goed, dat was de intro. Dan komt nu mijn punt [...] Laat een commissie het kaf van het koren scheiden, al te gamele voorstellen ertussenuit filteren, en gooi de rest in een bingomolentje.’ Het idee van Marleen *gammelt*; ‘kaf’,

‘koren’, ‘al te gammel’ zijn nogal vage begrippen. Duidelijk is dat de Groningse hoogleraar tamelijk boos was op 28 augustus. En dat snap ik. Zelfs Bommen Berend – de Groningse scheldnaam van de bisschop van Munster die de stad Groningen wilde inpikken maar op 28 augustus 1672 afgeslagen werd – zou zijn laatste bom hiervoor hebben gebruikt.

Maar wat is een serieuze oplossing van deze *ex aequo*-problematiek in de topsport en -wetenschap? Voor beide domeinen geldt dat dit probleem onvoldoende wordt onderkend. Naast de willekeur van promotiebeurzenoewijzing zijn er onterechte olympische goudverliezers (schaatser Koen Verwey zilver In Sotsji met drieduizendste seconde verschil met de nummer één). Om de topsport aantrekkelijk te houden, ook voor de sponsors, zijn vergaande maatregelen nodig, zoals spelregel- en materiaal-aanpassingen. In de sport zijn de eerste voorzichtige aanpassingen zichtbaar. Zo wordt in het voetbal voorzichtig geëxperimenteerd met nieuwe spelregels om zo het lage aantal doelpunten in toernooien op te krikken (meer dan 30% van de wedstrijden in WK’s en EK’s eindigt met maximaal één doelpunt). De wetenschap loopt hierop achter; voorlopig hanteren die de computergestuurde willekeur.

P.S. Een opmerkelijke lezer van een eerdere versie van deze column wees mij erop dat je het hele zwikkie beursaanvragen ook in de versnipperaar kunt gooien, om daarna een reconstructie te starten waarbij de eerste acht teruggeconstrueerde voorstellen in de prijzen vallen. Ik ken de grapjas, hij wist heel goed dat we zo te maken hebben met een soort *NP-hard* probleem. Immers, de snippers terugpuzzelen in de oorspronkelijke staat (kopieën zijn beschikbaar!) lukt voorlopig niet, maar ... controleren of een uitkomst correct is, is een fluitje van een cent. Op mijn vraag of hij er een promotie-onderzoek in ziet met bijpassende beursaanvraag, was zijn antwoord: ‘Hoogstens een hbo-bachelorscriptie.’ Ik denk dat ik Marleen bel.

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

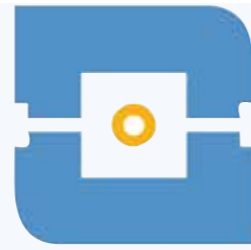


Annelieke Baller gepromoveerd

De redactie van STAtOR felicitteert haar mederedacteur Annelieke Baller van harte met haar promotie.

Zij verdedigde haar proefschrift met de titel *Improving Distribution Efficiency in Cash Supply Chains* aan de VU op 28 november 2019. Begeleiders waren Wout Dullaert, Leen Stougje en Said Dabia, onze hoofdredacteur Joaquim Gromicho was voorzitter van het Reading Committee.

Haar proefschrift beoogt inzicht te bieden in complexe distributieproblemen, efficiëntere distributieschema's te ontwikkelen en daarmee het potentiële voordeel van vernieuwende distributiestrategieën te analyseren. Zij keek daarbij naar toepassingen in het geldtransport.



Young Statisticians are looking back at a great 2019

In 2019, we organised many events for our members and statistics enthusiasts. After the summer break we kicked-off the new academic year with a Company Visit to Philips, at the High-Tech Campus in Eindhoven. A group of twenty young statisticians joined this event, where we learned how Philips uses statistical models to improve their maintenance services for medical systems and other healthcare applications. Another presentation focussed on the deep learning techniques they implemented in a new product that could improve sleep quality. We ended the day with some drinks and bites together with several data scientists from Philips, who could elaborate on the projects they worked on.

More events will be announced soon, so keep an eye on our website and Facebook page. You can also sign-up for our mailing list and get notified when new events are announced. We hope to see you all at our events in 2020!

WISSELING IN DE REDACTIE

Na lang deel te hebben uitgemaakt van onze redactie is Ana Isabel Baros teruggetreden. Wij zullen haar zorgzame inbreng missen, veel artikelen kwamen via haar binnen. Ook zorgde zij er steeds weer voor dat de winnaars van de Van Zwet en Hemelrijk Awards een artikel over hun onderwerp schreven. Dank Ana, wij hopen je nog vaak terug te zien op Annual Meetings en andere gelegenheden.

Nieuw in de redactie is Kerry Malone, wij hebben haar al ontmoet op een redeactiebijeenkomst en ze heeft al enkele contacten gelegd voor toekomstige artikelen. Welkom Kerry.

Back to school, learn about the latest developments in Operations Research

LNMB & NGB SEMINAR

OPTIMIZATION FOR AND WITH MACHINE LEARNING

Conference Center De Werelt in Lunteren,
January 15, 2020, 10.00 – 17.00

On Wednesday 15 January 2020, the Dutch OR Society (NGB) and the Dutch Network on the Mathematics of Operations Research (LNMB) will hold their joint annual seminar in Lunteren. This seminar is aimed at both practitioners and academics who are interested in the latest developments in OR. Besides an interesting program, the seminar offers an excellent opportunity to meet a large part of the Dutch OR community.

After the very successful seminars *Is Data Science a game changer for Operations Research?* (2018) and *Deep Learning and its impact on Operations Research* (2019), the 2020 seminar will focus on the interaction between optimization and machine learning. Several experts will address questions such as

- What is the role of optimization in machine learning?
- How can optimization benefit from machine learning?
- What is the potential impact of automated algorithm design methods?
- What is the role of optimization in improving the robustness of machine learning systems?
- Can machine learning be used to estimate the optimal value of complex real-life optimization problems?

The expected speakers at the seminar are:

- Marleen Balvert (Tilburg University)
- Taylan Cemgil (Deep Mind UK)
- Martina Fischetti (Vattenfall BA Wind)
- Tristan van Leeuwen (Utrecht University)
- Holger Hoos (Leiden University)
- Andrea Lodi (Polytechnique Montréal)

More information on the invited speakers and the programme will follow in the next weeks. www.lnmb.nl/conferences/2019/announcementngblnmbseminar/

Join the seminar and learn from experts all about Optimization for and with Machine Learning!



Na het 100-jarig jubileum congres in 1985 in Amsterdam, dat werd geopend in aanwezigheid van koningin Beatrix, keert het World Statistics Congress weer terug naar Nederland. Het 63e WSC zal van 11 – 15 juli 2021 in het World Forum in Den Haag plaatsvinden. Het National Organising Committee, onder co-voorzitterschap van Eric Schulte Nordholt en John Bailer, is al geruime tijd bezig dit voor te bereiden. Ook de VVSOR is hierbij actief betrokken.

World Statistics Congresses hebben hun oorsprong in het werk van de Belgische geleerde Adolphe Quetelet die in 1853 een eerste Europees congres over statistiek organiseerde. Na nog enkele van deze congressen werd in 1885 het International Statistical Institute opgericht om de continuïteit van deze congressen en de vele internationale contacten te bewerkstelligen. In de beginjaren lag de focus vooral op harmonisatie van methoden en technieken tussen de verschillende Europese landen. Op die manier wilde men waarborgen dat economische en demografische cijfers vergelijkbaar zijn. Het ISI beschikt al sinds 1913 over een Permanent Office dat in Den Haag is gevestigd. De reeks tweejaarlijkse WSC's, die oorspronkelijk ISI Sessions heetten, is slechts door de beide Wereldoorlogen onderbroken. Vanaf de jaren '50 is het ISI uitgebreid met Associations, momenteel zeven in getal, die zich bezighouden met specifieke onderwerpen zoals Mathematical Statistics, Statistical Education of Statistical Computing.

De WSC's kenmerken zich door een open en gastvrij karakter. Beginnende wetenschappers en *hot shots* men- gen zich zonder drempels in discussies en vele langdurende vriendschappen zijn hier ontstaan. Dit congres is daarom een unieke kans voor Nederlandse statistici om zonder hoge reiskosten nieuwe contacten op te doen en veel van de grootste experts ter wereld te ontmoeten. U bent welkom, niet alleen als deelnemer, ook uw ideeën over onderwerpen of sponsoring zijn waardevol.

INFORMATIE

ISI Permanent Office (P.O. Box 24070, 2490 AB The Hague, phone +31 70 33 75 737)

e-mail: isi@cbs.nl

ISI website: isi-web.org

Congres website: isi2021.org



Het 62e World Statistics Congress vond van 18 – 23 augustus 2019 plaats in Kuala Lumpur, Maleisië. Meer dan 2500 deelnemers konden kiezen uit honderden presentaties. Ook presenteerde Maleisië zich met een uitgebreid cultureel programma.

Een groot aantal deelnemers uit Nederland was in Kuala Lumpur aanwezig. Samen met ISI president John Bailer nodigen zij op deze foto alle statistici wereldwijd uit om deel te nemen aan het 63e WSC in 2021 in Den Haag.