

# STATATOR

---

periodiek van de VvS+OR jaargang 15, nummer 3, november 2014

Statistiek en schaatsen;  
met Principal Component Analysis  
naar een eerlijker allroundtoernooi

Een dynamisch voorraadprobleem;  
wat te doen als gerapporteerde data  
inexact zijn?

OR voor betere bloeddorstroming

Bayes, Price en de orde van de dingen

---

STATOR is een uitgave van de Vereniging voor Statistiek en Operationele Research (VvS+OR). STATOR wil leden, bedrijven en overige geïnteresseerden op de hoogte houden van ontwikkelingen en nieuws over toepassingen van statistiek en operationele research. Verschijnt 3 keer per jaar.

**Redactie**

Joaquim Gromicho (hoofdredacteur), Annelieke Baller, Ana Isabel Barros, Johan van Leeuwen, Guus Luijben (eindredacteur), Richard Starmans, Gerrit Stermerdink (eindredacteur), Hilde Tobi en Vanessa Torres van Grinsven. Vaste medewerkers: Fred Steutel 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, telefoon 020-5986010, mobiel 06-55886747, <j.a.dossantos.gromicho@vu.nl>.

**Bestuur van de VvS+OR**

Voorzitter: prof. dr. Jacqueline Meulman <president@vvs-or.nl>  
 Secretaris: dr. Fetsje Bijma <bestuur@vvs-or.nl>  
 Penningmeester: dr. Ad Ridder <penningmeester@vvs-or.nl>  
 Overige bestuursleden: prof. dr. Eric Cator (SMS), prof. dr. Jeanine Houwing-Duistermaat (BMS), Maarten Kampert MSc., dr. John Poppelaars (NGB), dr. Michel van de Velden (ECS), dr. Jelte Wicherts (SWS).

**Leden- en abonnementenadministratie van de VvS+OR**

VvS+OR, Postbus 244, 6700 AE Wageningen, telefoon 0317-419572, e-mail <admin@vvs-or.nl>.  
 Raadpleeg onze website over hoe u lid kunt worden van de VvS+OR of een abonnement kunt nemen op STATOR.

**VvS+OR-website**

www.vvs-or.nl

**Sociale media**

Wilt u uw vakgenoten ontmoeten en wilt u discussiëren over actuele thema's, volg dan de VvS+OR en de Young Statisticians via LinkedIn, Facebook, Twitter en Flickr.  
 Sluit u aan bij de LinkedIn-groep van VvS+OR of Young Statisticians; bekijk foto's op <www.flickr.com/photos/vvs-or/sets>; like onze Facebook-pagina; volg de President van VvS+OR op <https://twitter.com/#!/dutchstat>.

**Advertentieacquisitie**

M. van Hootegem <hootegem@xs4all.nl>  
 STATOR verschijnt in maart, juli en november.

**Ontwerp en opmaak**

Pharos, Nijmegen

**Uitgever**

© Vereniging voor Statistiek en Operationele Research  
 ISSN 1567-3383

**INHOUD**

- 3** Over schaatsen, schatten en schema's – redactioneel.
- 4** Statistiek en schaatsen; met Principal Component Analysis naar een eerlijker allroundtoernooi.  
**Miriam Loois**
- 10** Op een mooie pinksterdag – column.  
**Fred Steutel**
- 11** Een dynamisch voorraadprobleem; wat te doen als gerapporteerde data inexact zijn?  
**Frans de Ruiter**
- 15** Schijnagenda's – column.  
**Johan van Leeuwen**
- 16** OR voor betere bloeddorstomming.  
**Luuk Besselink, Samuel van Brummelen, Wim de Kort, Patricia Zonneveld & Nico van Dijk**
- 21** Young Statisticians
- 22** Lotto nonsens; de wereld wil bedrogen worden – column.  
**Henk Tijms**
- 24** Bayes, Price en de orde van de dingen.  
**Richard Starmans**
- 27** IM Guus Boender (1955–2014).  
**Gerrit Timmer**
- 28** Oproep voor voordrachten voor de VvS+OR Master's Thesis Award en de Willem R. van Zwet Award voor een uitmuntende Ph.D. thesis.  
LNMB-conferentie 'The mathematics of operations research', 13–15 januari 2015.  
Dag voor Statistiek en OR 2015, 26 maart 2015.  
Oproep voor voordrachten voor de Van Dantzigprijs 2015.

**OVER SCHAATSEN, SCHATTEN EN SCHEMA'S**

Dit nummer van STATOR opent met een verrassende toepassing van Principal Component Analysis. De vrijwel nooit gebruikte Nederlandse naam hiervoor is Hoofdassenmethode. Deze techniek werd in de jaren 30 ontwikkeld door test-theoretici als Louis Thurstone en wordt ook wel eens op een wat simplistische manier omschreven als 'het beschrijven van een best passend geraamte als alleen de buitenkant zichtbaar is'. Omdat de techniek berust op het bepalen van de eigenwaarden en eigenvectoren van een correlatiematrix kon men oorspronkelijk geen grotere problemen dan met ongeveer 5-8 variabelen aan. Maar met de komst van computers kon het theoretische concept ook praktisch uitgewerkt worden. Vooral in de sociale wetenschappen is het een veel gebruikte techniek. Miriam Loois past hem hier toe op een analyse van de prestaties van schaatsers op de verschillende afstanden in een allround toernooi. U zult na het lezen met heel andere ogen naar de komende toernooien kijken.

Zijn er eigenlijk wel exacte data? Als we die vraag heel principieel stellen moet het antwoord natuurlijk nee zijn, er zit altijd een fundamentele quantumvariatie in alles wat we meten. Gelukkig is die meestal te verwaarlozen in onze alledaagse problemen, maar ook dan hebben we regelmatig te maken met niet-exacte gegevens. Frans de Ruiter beschrijft een methode waarbij een voldoende voorraadbeheer kan plaatsvinden als de vraaggegevens niet exact bekend zijn. Dat komt vaker voor dan u denkt. Stelt u zich maar eens een aantal supermarkten voor die vanuit een centraal magazijn bevoorrad moeten worden en waarbij de bestelling van één supermarkt niet op tijd binnen is.

Nico van Dijk en zijn mede-auteurs hebben een probleem van de bloedbanken bestudeerd. Men wil de bloeddonoren, die altijd vrijwilligers zijn, natuurlijk niet laten wachten als ze komen voor een donatie. Maar dat kan alleen als er een veel te grote staf aanwezig is. In dit artikel wordt gezocht naar een methode waarbij met een minimale staf-inzet een maximale dienstbaarheid naar de donoren wordt bereikt.

Het Theorema van Bayes hoeft aan de statistici onder ons niet te worden uitgelegd. Wat echter niet algemeen bekend is, is dat het pas na de dood van Bayes is gepubliceerd door toedoen van Price. In ons vierde grote artikel beschrijft Richard Starmans de geschiedenis van het theorema en ziet daarbij verrassende parallellen met de introductie van de evolutietheorie.

Verder wordt dit nummer gevuld door onze vaste columnisten Johan van Leeuwen, Henk Tijms en Fred Steutel die vertelt over een lunch met Arnon Grunberg.

De OR-gemeenschap heeft een groot verlies geleden met het overlijden van Guus Boenders. Gerrit Timmer herdenkt hem.

Ten slotte nog dit: volgend jaar bestaat de VvS+OR 70 jaar! Wij willen daar in het eerste nummer van 2015 extra aandacht aan besteden. Aan een aantal oud-voorzitters is gevraagd om een herinnering aan hun periode met de vereniging. Maar wij verwelkomen ook soortgelijke bijdragen van ieder ander lid. De deadline voor het aanleveren van uw herinnering is de eerste week van januari.

Om af te sluiten met een variant op de gebruikelijke tekst: *wij wensen u veel schrijfplezier!*

*De STATOR-redactie*

# STATISTIEK EN SCHAATSEN met Principal Component Analysis naar een eerlijker allroundtoernooi

De Internationale Schaatsunie ISU heeft in juni 2014 besloten dat op Europese allroundschaatstoernooien de 10 kilometer voor mannen en 5 kilometer voor vrouwen worden vervangen door een 1.000 meter. Sven Kramer is woedend en noemt het een 'supersprint' in plaats van een allroundtoernooi. Heeft hij gelijk? Worden sprinters in het nieuwe systeem bevoordeeld? Of hebben stayers in het huidige systeem juist voordeel? Statistiek kan ons helpen om daar een uitspraak over te doen. En om een eerlijker allroundsysteem vorm te geven.

MIRIAM LOOIS

Voor de ene schaatser lijkt de 500 meter soms al te lang, terwijl de andere pas op gang komt halverwege de 10 kilometer. Gevoelsmatig weten we allemaal dat er verschillende soorten rijders zijn, maar we kunnen dat ook wetenschappelijk onderbouwen. Een wiskundige methode waarmee dit kan is *principal component analysis*. Met principal component analysis kan een dataset met gecorreleerde variabelen getransformeerd worden in een set met ongecorreleerde variabelen, de principale componenten. De eerste principale component verklaart het meest van de variantie in de data, gevolgd door de tweede component, die loodrecht staat op de eerste.

Ik heb deze techniek toegepast op de persoonlijke records van 2.449 mannelijke schaatsers op 6 afstanden (500, 1.000, 1.500, 3.000, 5.000 en 10.000 meter) uit de database van Jakub Majerski. Eerst worden de data genormaliseerd, zodat de gemiddelde tijd per afstand 0 is en de standaarddeviatie 1. Dan wordt deze matrix via *singular value decomposition* geschreven als het product van drie matrices  $U$ ,  $\Sigma$  en  $V$ .

$$\bar{P}_{i,d} = \sum_j U_{ij} \Sigma_{jj} V_{dj}$$

- $\bar{P}_{i,d}$  = Genormaliseerd persoonlijk record van schaatser  $i$  op afstand  $d$ ;
- $U_{ij}$  = Interactie van schaatser  $i$  met principale component  $j$ ;
- $\Sigma_{jj}$  = Principale component  $j$ ;
- $V_{dj}$  = Interactie van persoonlijk record op afstand  $d$  met principale component  $j$ .

De matrix  $\Sigma$  is diagonaal en bevat de principale componenten geordend van groot naar klein. Bij elke principale component hoort een kolom van  $U$  die wat zegt over de interactie van de schaatser met de principale component, en een kolom van  $V$  die wat zegt over de interactie met het persoonlijke record op een afstand.

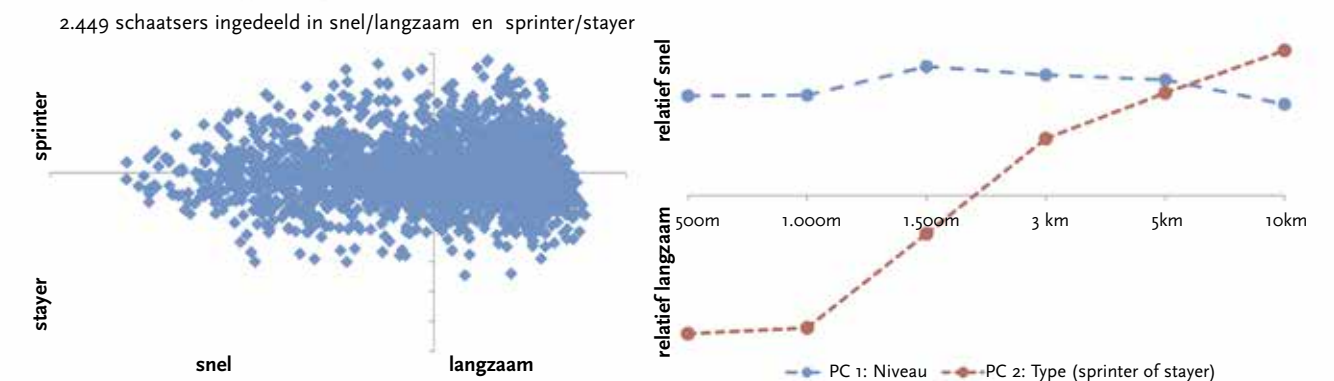
De eerste twee principale componenten verklaren samen 90% van de verschillen tussen de tijden van schaatsers. De eerste principale component zegt iets over het algemene niveau van de schaatser en verklaart 66% van de variantie. De ene schaatser is nu eenmaal sneller dan de andere, ongeacht de afstand. De tweede factor zegt

Sven Kramer op de 10.000 meter Olympische Winterspelen 2010 in Vancouver.



iets over het feit of de schaatser een sprinter of een stayer is. Sprinters zijn relatief sneller op de korte afstanden, stayers zijn juist beter op de langere afstanden. Deze principale component verklaart 24% van de variantie. De andere principale componenten verklaren allemaal minder dan 5% van de variantie en zeggen iets over de

relatieve snelheid op de middenafstanden (pc 3) en de relatieve snelheid op individuele afstanden (pc 4, 5 en 6). De eerste twee principale componenten zijn grafisch weergegeven in figuur 1. De linker figuur toont een spreidingsgrafiek waarbij de eerste kolom van matrix  $U$  is afgezet tegen de tweede kolom van  $U$ . De rechter figuur



Figuur 1. (links) Schaatsers indelen met behulp van de eerste twee principale componenten, niveau en type; (rechts) de eerste twee kolommen van matrix  $V$ .



	NIVEAU: SNEL (1) OF LANGZAAM (-1)	TYPE: STAYER (1) OF SPRINTER (-1)
Kramer	1,00	0,73
Verwey	1,00	0,03
Blokhuijsen	0,99	0,38
Uytdehaage	0,98	0,59
Verheijen	0,98	0,85
Wennemars	0,99	-0,60
Tuitert	0,99	-0,50

Tabel 1. Nederlandse (oud) topschaatsers en hun niveau en type. De schaatsters zijn geordend van snel naar langzaam en van stayer naar sprinter. Op basis van hun positie in die lijst krijgen ze een niveau en type toegekend tussen -1 en 1.

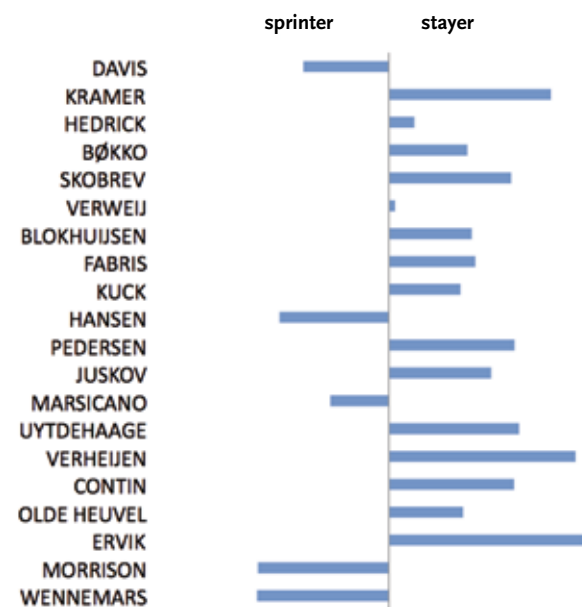
geeft de eerste twee kolommen van matrix V weer.

Deze twee principale componenten kun je op basis van percentielen in de verdeling uitdrukken in een getal tussen -1 en 1. Zie tabel 1 met Nederlandse toppers. Ze hebben allemaal een niveau dat dicht bij 1 ligt. Kramer en Verheijen blijken in de analyse, niet geheel onverwacht, echte stayers, terwijl Tuitert en Wennemars aan de sprintkant van het spectrum zitten. Verweij is een relatief neutrale schaatser.

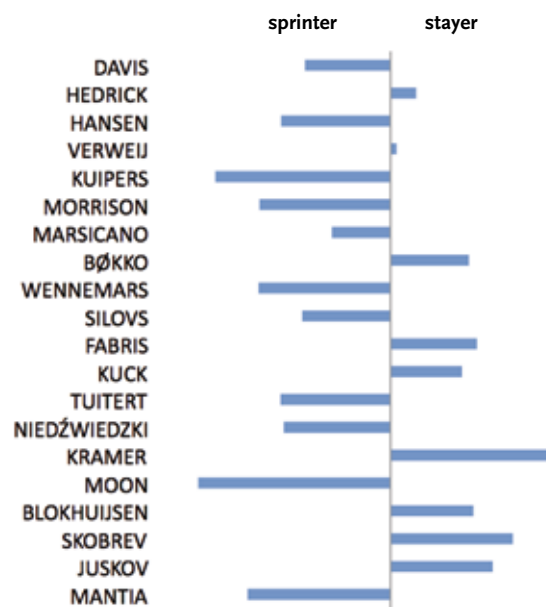
### Is het allroundtoernooi wel eerlijk?

In een allroundtoernooi worden vier afstanden geschaatst. Bij de mannen zijn dat op dit moment de 500, 1.500, 5.000 en 10.000 meter. De vrouwen rijden de 500, 1.500, 3.000 en 5.000 meter. Elke schaatser krijgt punten voor elke afstand. Per afstand wordt de tijd omgezet naar een tijd op de 500 meter, en deze tijden worden bij elkaar opgeteld. Degene met het minste aantal punten wint. Een achterstand van één seconde op de 500 meter moet dus worden gecompenseerd door 10 seconden op de 5 kilometer.

top 20 huidig allroundsysteem  
(500 | 1.500 | 5.000 | 10.000)



top 20 nieuw allroundsysteem  
(500 | 1.000 | 1.500 | 5.000)



Figuur 2. Top 20 gebaseerd op persoonlijke records in het huidige systeem en in de nieuwe opzet.

Nu de hamvraag: kunnen we zeggen of een bepaald type schaatser in het voordeel is op een allroundtoernooi? Als we in figuur 2 kijken naar de top 20 van de wereld in het huidige systeem, gebaseerd op persoonlijke records (de zogenoemde Adelskalender), dan zien we relatief veel stayers. Davis is de aanvoerder van deze lijst, gevolgd door Kramer. In het nieuwe systeem, waarin de 10 kilometer is vervangen door de 1.000 meter, komen juist veel sprinters bovendrijven. Kramer moet genoeg nemen met een 15e plek. Geen wonder dat hij boos is!

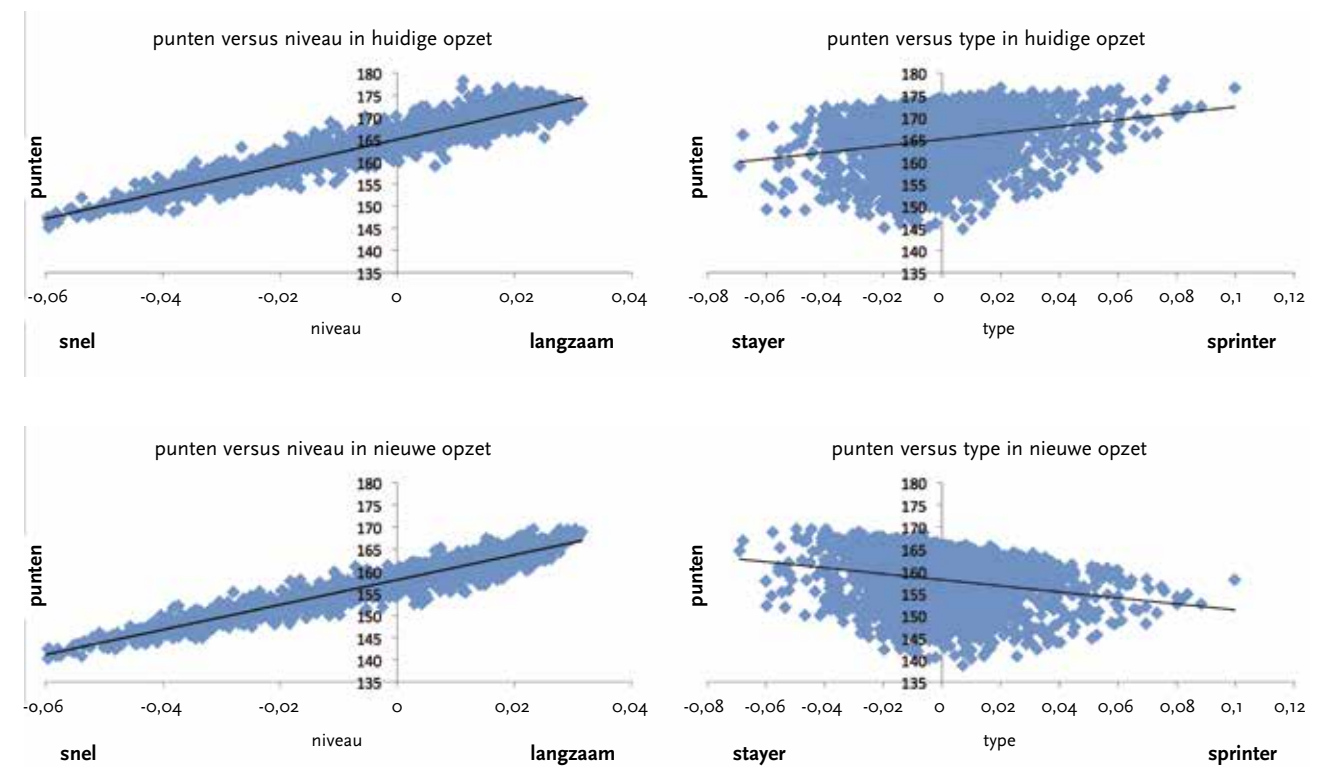
In wiskundige termen: in het huidige allroundtoernooi is de correlatie tussen puntenaantal en het niveau van de schaatser 96% en de correlatie met de principale component 'type' is 24%. Dat betekent dat stayers in het voordeel zijn. Winnen in het nieuwe allroundtoernooi hangt voor 97% samen met het algemene niveau van de schaatser en voor 23% met de vraag of hij een sprinter is. Zowel het nieuwe als het huidige systeem bevoordeelt dus een groep schaatsters. De samenhang tussen puntenaantal en niveau en type is te zien in figuur 3.

Naast het besluit om de langste afstand te vervangen door de 1.000 meter heeft de ISU besloten dat tijdens de Olympische Spelen van 2018 de 500 meter maar één keer zal worden gereden. Nu rijden de schaatsters deze afstand twee keer, een keer startend vanuit de binnenbaan en een keer startend vanuit de buitenbaan. Zowel Kamst e.a. (2010) als Hjort (1994) hebben aangetoond dat er een significant binnenbaan-buitenbaan-verschil is. Ook dit besluit leidt dus niet tot een eerlijker toernooi.

### Naar een eerlijker allroundtoernooi?

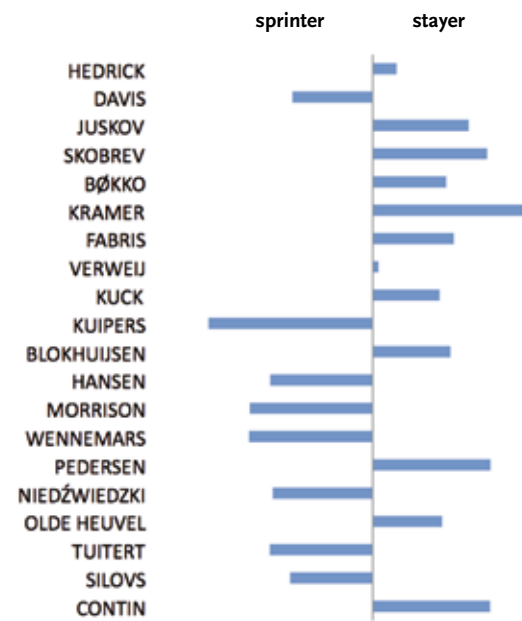
Ik heb twee alternatieven bekeken. In het eerste alternatief wordt de 10 kilometer niet door de 1.000 meter, maar door de 3 kilometer vervangen. In het tweede alternatief wordt de 10 kilometer helemaal niet vervangen door een andere afstand; hij valt simpelweg.

In figuur 4 zien we weer de top 20 op basis van persoonlijke records. Bij het eerste alternatief zien

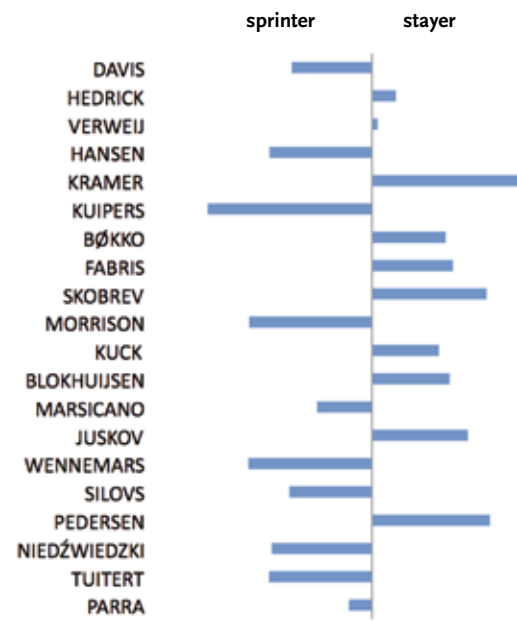


Figuur 3. Relatie tussen punten in de huidige en nieuwe opzet en het niveau en type schaatser. Een negatief 'niveau' hoort bij een relatief snelle schaatser. Een positief 'type' representeert een sprinter, een negatief 'type' een stayer. De grafieken bevatten trendlijnen.

top 20 alternatief 1  
(500 | 1.500 | 3.000 | 5.000)



top 20 alternatief 2  
(500 | 1.500 | 5.000)

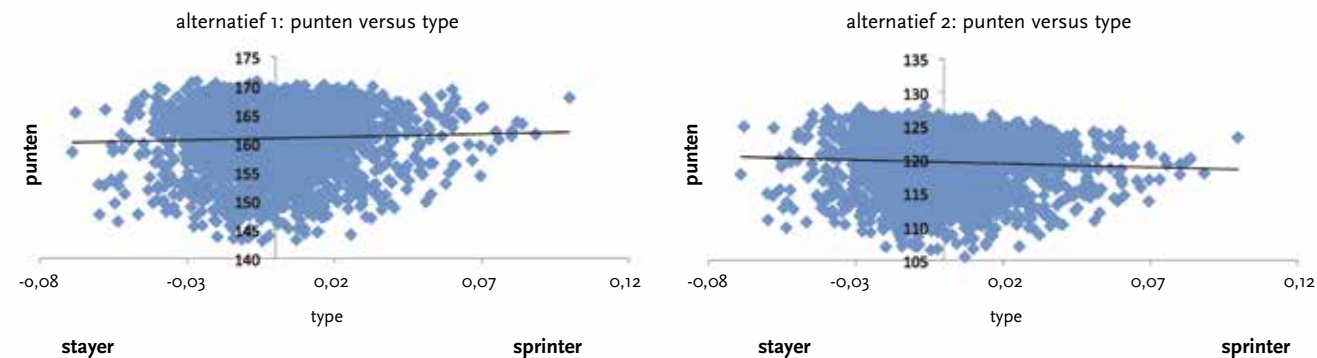


Figuur 4. De top 20 op basis van persoonlijke records voor twee alternatieve systemen.

we al meer sprinters, hoewel de top 10 nog steeds gedomineerd wordt door stayers. Bij het tweede alternatief is de verdeling tussen sprinters en stayers aan de top nog evenwichtiger.

Dat de alternatieven evenwichtiger zijn zien we

ook als we naar figuur 5 en tabel 2 met correlaties kijken. Winnen hangt nu voor maar liefst 99% samen met het algemene niveau van de schaatser en de vraag of hij een sprinter of een stayer is heeft veel minder invloed.



Figuur 5. Relatie tussen punten in de twee alternatieven en het niveau en type schaatser.

	HUIDIG SYSTEEM 500-1.500-5.000-10.000	NIEUW SYSTEEM 500-1.000-1.500-5.000	ALTERNATIEF 1 500-1.500-3.000-5.000	ALTERNATIEF 2 500-1.500-5.000
NIVEAU	0,96	0,97	0,99	0,99
TYPE SCHAATSER	0,24	- 0,23	0,04	- 0,05

Tabel 2. Correlatie tussen puntenaantal en niveau en type in de verschillende systemen. Een positieve correlatie betekent dat stayers in het voordeel zijn, bij een negatieve correlatie zijn sprinters in het voordeel.

### Conclusies

Terug nu naar de woede van Kramer. Hij had gelijk: de maatregel om de 10 kilometer te vervangen door de 1.000 meter bevoordeelt de sprinters fors. Maar bij de huidige vormgeving van het allroundtoernooi zijn stayers juist erg in het voordeel.

Met behulp van principal component analysis is het mogelijk om een eerlijker allroundtoernooi vorm te geven, waarbij het algemene niveau van de schaatser daadwerkelijk de doorslag geeft. Dit kan door de 10 kilometer door de 3 kilometer te vervangen in plaats van door de 1.000 meter. Een andere mogelijkheid is om de 10 kilometer helemaal niet te vervangen en een toernooi over drie afstanden te rijden.

Moge de beste winnen!

### LITERATUUR

- Jakub Majerski's speedskating database: <<http://www.sskating.com/list.php?season=2013/2014&lm=m&dist=Big6>>.  
Belangrijkste ISU-besluiten op een rij (ISU-congres Dublin 2014): <<http://knsb.nl/nieuws/belangrijkste-isu-besluiten-op-een-rij/>>.  
Kamst, R., Kuper, G. H., & Sierksma, G. (2010). The Olympic 500 meter speed skating; the inner-outer lane difference. *Statistica Neerlandica*, 64(4), 448-459.  
Hjort, N. J. (1994). *Should the Olympic Sprint skaters run the 500 meter twice?* Statistical Research Report. Oslo: Institute of Mathematics, University of Oslo.

MIRIAM LOOIS heeft de Master Theoretische Natuurkunde gevolgd aan de Universiteit Utrecht en de Master Actuarial Science and Mathematical Finance aan de Universiteit van Amsterdam. Ze werkt nu als Asset Liability Manager bij pensioenuitvoerder PGGM.

E-mail: <[miriamloois@gmail.com](mailto:miriamloois@gmail.com)>

Ook binnen de financiële sector wordt principal component analysis gebruikt, bijvoorbeeld voor het beschrijven van renterisico. De rentecurve met rentes bij verschillende looptijden wordt gebruikt voor het waarderen van verplichtingen van verzekeraars en pensioenfondsen. Bewegingen in deze curve vormen daarom (potentieel) een risico. Als principal component analysis wordt toegepast op de veranderingen in de rente komen er ook twee belangrijke factoren naar boven. De eerste is een verandering in het niveau van de rente; alle rentes stijgen of dalen. Dit is te vergelijken met het niveau van de schaatser. De tweede is een draaiing van de rente; de korte rente stijgt en de lange rente daalt of vice versa. Hier is een parallel te trekken met het type schaatser. Deze factor 'draait' als het ware de tijden op de verschillende afstanden; een snellere tijd op de korte afstanden en een langzamere tijd op de lange afstanden of vice versa.





## Op een mooie pinksterdag

Op tweede pinksterdag jongstleden waren mijn vrouw Vita en ik voor de lunch te gast bij Arnon Grunberg op een terras van het Hilton Amsterdam. Dat is zo gekomen. In een van Grunbergs 'Voetnoten' op de voorpagina van *de Volkskrant* stond de volgende oproep.

Wetenschappers die niet met hun studenten vrijen leveren een belangrijke bijdrage aan de vooruitgang van de mensheid. Daarom roep ik vandaag de David Lurie-prijs in het leven voor wetenschappers die over weinig didactische vaardigheden beschikken, maar desalniettemin weigeren seks te hebben met studenten. U kunt tot 1 mei 2014 wetenschappers nomineren voor de David Lurie-prijs. Graag met uitvoerige motivatie. De prijs is een diner met mij in een sterrenrestaurant.

De prijs is vernoemd naar de hoofdpersoon David Lurie in het boek *Disgrace* van J. M. Coetzee; die hoofdpersoon vrijt wel met studentes. Ik heb mijzelf als volgt aangemeld.

In 1969 was ik *assistant professor* in de statistiek aan de universiteit van Texas in Austin. Op mijn spreekuur kwam eens een meisjesstudent met een mandje. Ze zette dat mandje gedecideerd op mijn bureau; er bleek een baby in te liggen. Hoewel ik te kort in Austin was om van vaderschap beschuldigd te kunnen worden, ging er toch een zekere dreiging van uit. Maar, het meisje was uiterst vriendelijk: ze kwam alleen vertellen dat ze niet meer op mijn college zou komen: de baby had haar aandacht nodig. Omdat er dus geen enkele aanleiding was voor 'disgrace', denk ik in aanmerking te komen voor de David Lurie-prijs: hele generaties Eindhovense studenten zullen willen

getuigen van mijn didactische tekortkomingen, en ik ben niet met dat meisje, of met welke meisjesstudent dan ook, naar bed geweest.

Rest mij nog een vraag: mag ik naar het beloofde etentje mijn vrouw meenemen? Zij eet niet zo veel en is vegetariër – ik niet.

### Nominatievoetnoot

Eind vorige maand maakte ik bekend dat ik de David Lurie-prijs zou uitreiken aan een universitair docent of professor met matige didactische kwaliteiten die niettemin nooit seks met zijn of haar studenten heeft gehad. Een onafhankelijke jury heeft inmiddels al besloten dat de David Lurie-prijs 2014 gaat naar F. W. Steutel uit Eindhoven. Steutel heeft zichzelf genomineerd en is naar eigen zeggen nooit met een 'meisjesstudent' naar bed geweest. Of hij wel met een jongensstudent naar bed is geweest, is vooralsnog onduidelijk. Steutel zal zijn vrouw meenemen naar de prijsuitreiking. Hij schreef me: 'Zij eet niet zoveel en zij is vegetariër – ik niet'.

We hebben zeer smakelijk geluncht, onder her genot van een glaasje Gewürztraminer en heel gezellig gepraat. We zijn als vrienden uit elkaar gegaan en onderhouden sindsdien een vrij regelmatig e-mailcontact. Arnon Grunberg wil statistiek leren; dat heeft hij mij bij herhaling verzekerd. Ik heb hem ter inleiding het allereerste boek van William Feller aangeraden: *Introduction to Probability Theory and its Applications*, Volume I (1950). Er bestaat geen (mathematische) statistiek zonder kansrekening.

FRED STEUTEL is emeritus hoogleraar kansrekening aan de TU Eindhoven.

E-mail: <fsteutel@xs4all.nl>



# EEN DYNAMISCH VOORRAADPROBLEEM

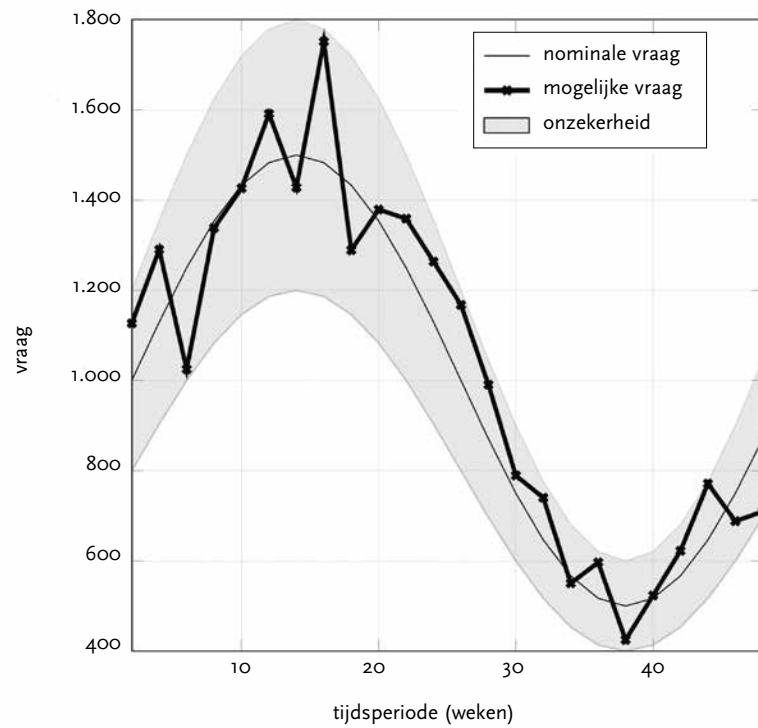
## Wat te doen als gerapporteerde data inexact zijn?

FRANS DE RUITER

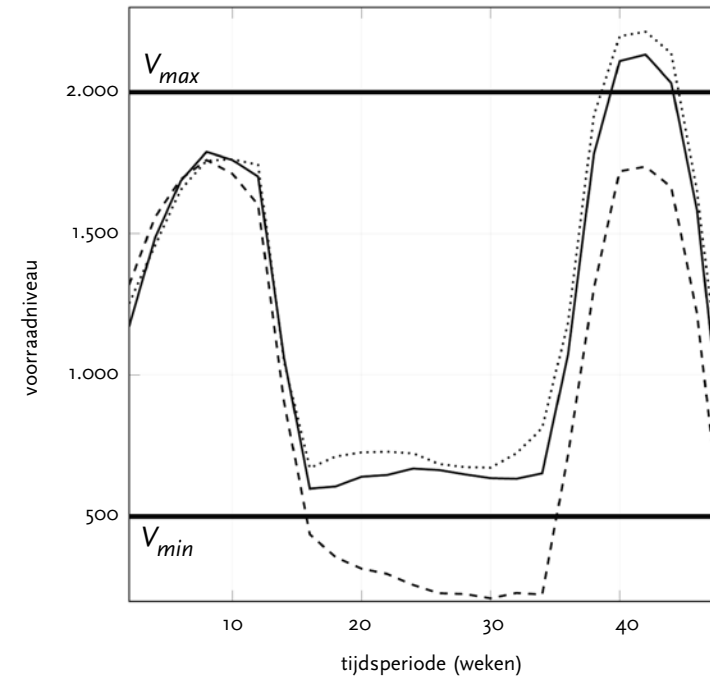
Bij productie-voorraadproblemen is de vraag naar de producten vaak onzeker. Dit maakt het opstellen van een productieplan dat aan alle vraag voldoet lastig, zeker als het plan moet voldoen aan restricties op productie- en voorraadniveaus. Het is door die restricties niet mogelijk om de productie voor de hele planningshorizon al van tevoren vast te stellen. Dit hoeft ook niet, omdat we elke periode kunnen bepalen hoeveel we gaan produceren. Dit probleem wordt daarom een *multi-stage probleem* genoemd. In dit multi-stage probleem is op tijdstip  $t$  bekend wat de vraag in alle voorgaande perioden  $1, 2$  tot en met  $t-1$  is geweest. Multi-stage technieken gebruiken deze data om de productie voor periode  $t$  te bepalen. Al deze technieken nemen echter aan dat de verkregen data exact zijn. In de praktijk blijkt dat deze data vaak inexact zijn: de werkelijke vraag verschilt van de gerapporteerde waarde. De inexacte aard van de gerapporteerde data kan grote gevolgen hebben voor het productieplan. In De Ruiter et al. (2014) is een methode ontwikkeld om dergelijke multi-stage voorraadproblemen met inexact verkregen data op te lossen.

### Een toepassing: het productie-voorraad probleem

Ter illustratie van onze methode gebruiken we een bekend voorraadmodel uit het artikel van Ben-Tal, El Ghaoui en Nemirovski (2004). Het model bestaat uit één opslagfaciliteit en drie fabrieken die het product produceren. Om de week neemt het management een beslissing over de productie. In totaal moet de productie gepland worden voor 48 weken. De vraag naar de producten is seizoensgebonden, met een maximum vraag in de winter. Voor ons voorbeeld hebben we verschillende onzekerheidsniveaus getest van 2,5%, 5%, 10% en 20%. Een onzekerheidsniveau van 20% wil zeggen dat de werkelijke vraag tot 20% meer of minder kan zijn dan de door de manager verwachte vraag, de zogenaamde nominale vraag. Dit onzekerheidsniveau wordt opgesteld in samenspraak met het management en indien mogelijk aan de hand van historische data. Het productieplan moet zo opgesteld worden dat men aan alle vraag in het onzekerheidsgebied kan voldoen. Zie figuur 1 voor een grafische weergave van dit onzekerheidsgebied.



Figuur 1. Onzekerheid in de vraag. De grijs gekleurde strip geeft de onzekerheid van 20% weer.



Figuur 2. Het negeren van de inexacte aard van data heeft gevolgen voor het voorraadniveau.

Hierbij moeten we ook rekening houden met productiekosten, productiecapaciteit per fabriek, een minimum voorraadniveau en de opslagcapaciteit.

### Het robuuste optimalisatiemodel

Alle restricties in het optimalisatieprobleem voor het productie-voorraadprobleem kunnen als lineaire ongelijkheden worden geschreven. Helaas weten we de vraag in ons model niet precies en wijkt deze af van de nominale waarde. Meestal is de informatie over de data erg miniem en kunnen we daarom niet vertrouwen op stochastische technieken. Daarom baseren we ons hier op robuuste optimalisatietechnieken waarbij het productieplan moet voldoen aan de restricties voor alle mogelijke vraagpatronen in het onzekerheidsgebied. Het eerste artikel dat dit idee ter sprake bracht is geschreven door Soyster (1978), maar intensief onderzoek naar robuuste optimalisatietechnieken kwam pas eind jaren 90 op gang door een reeks artikelen van de grondleggers van het onderzoeksveld: Aharon Ben-Tal, Laurent El Ghaoui en Arkadi Nemirovski. Diezelfde onderzoekers hebben recentelijk een boek geschreven (Ben-Tal et al., 2009) waarin een grote verscheidenheid aan robuuste optimalisatietechnieken wordt uitgelegd.

### Robuust dynamische oplossingen

In ons productie-voorraadprobleem hoeven sommige beslissingen pas genomen te worden als een deel van de

onzekere data bekend is. Voor deze *wait-and-see* beslissingen in periode  $t$  kunnen we onze beslissingen baseren op vraagcijfers die beschikbaar zijn in de perioden 1 tot en met  $t-1$ . De *wait-and-see* variabelen in periode  $t$  zijn dan een wiskundige functie van de beschikbare data op dat moment. Dit model heet het *Adjustable Robust Optimization (ARO)* model (Ben-Tal et al., 2014). In dit model bedoelen we met *adjustable* dat sommige variabelen zich aanpassen aan de geobserveerde data. Vaak maken deze *wait-and-see* variabelen het probleem vele malen complexer, maar voor sommige gevallen kunnen we deze problemen nog efficiënt oplossen. Een van die gevallen is wanneer er geen onzekerheid is in de coëfficiënten van de *wait-and-see* variabelen en de mogelijke wiskundige functies voor deze variabelen lineair zijn. De coëfficiënten van deze lineaire functies zijn dan nieuwe optimalisatievariabelen die bepaald worden door ons model. De waarde voor de *wait-and-see* variabele kan dan bepaald worden zodra de data beschikbaar zijn. Met ARO-technieken kan dit model herschreven worden tot een efficiënt oplosbaar probleem zoals beschreven in Ben-Tal et al. (2014). In dit geval is het resulterende model een lineair programmeringsprobleem dat efficiënt op te lossen is met simplex of inwendige-punt methoden.

### Inexacte data in de praktijk.

Inexacte data zijn nog steeds een groot probleem in de praktijk. Zo heeft volgens recent onderzoek door Haug et al. (2011) nog niet de helft van de bedrijven vertrouwen in

hun data. Nog belangrijker, 75% procent van de bedrijven heeft kosten kunnen identificeren die terug te leiden zijn op afwijkingen in de data. Voor voorraadmodellen zijn er zelfs studies, zoals DeHoratius en Raman (2008), die zeggen dat 65% van de gebruikte data inaccuraat zijn. Samenvattend kunnen we vaststellen dat verkregen data vaak inexact zijn.

### Gevolgen van inexact verkregen data

Stel we zijn naïef en nemen aan dat de verkregen data voor de beslisregels exact zijn. We bepalen vervolgens met bestaande ARO-technieken een dynamisch productieplan. We hebben dan een werkend productieplan, maar dat werkt alleen als alle verkregen data exact zijn. In het onderstaande voorbeeld laten we zien dat de oplossingen niet bruikbaar zijn voor het management als er inexacte data gebruikt worden in de beslisregels. Ter illustratie hebben we in ons voorbeeld verlate rapporten over de vraagcijfers als oorzaak van inexacte data. Volledige informatie is dan pas laat beschikbaar; op het moment dat de beslissing genomen moet worden is er alleen een schatting van de werkelijke vraag. Als in periode  $t$  slechts 9 van de 10 winkels de vraagcijfers hebben gerapporteerd voor een bepaalde periode, dan is er nog steeds een bepaalde onzekerheid en dus zijn de verkregen data inexact. Resultaten in De Ruiter et al. (2014) laten zien dat de beslisregels niet aan alle restricties voldoen als naïef wordt aangenomen dat de verkregen data exact zijn. Zo kan de opslagcapaciteit overschreden worden of kunnen

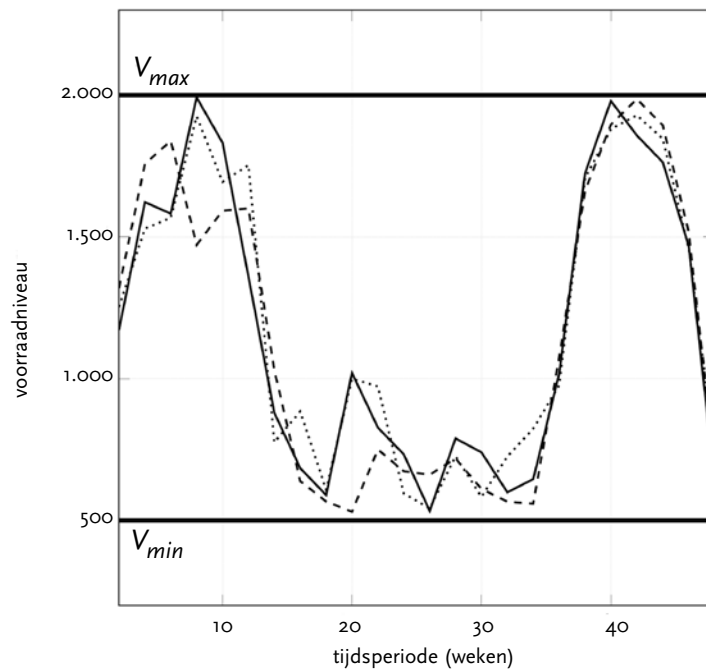
we onder het gewenste minimum voorraadniveau uitkomen. In figuur 2 zijn de gevolgen op het voorraadniveau zichtbaar voor drie potentiële scenario's uit ons onzekerheidsgebied. In totaal hebben we 100 verschillende vraagpatronen gesimuleerd en trad het voorraadniveau tot in 80% van de gevallen buiten de gewenste grenzen.

### Beslisregels gebaseerd op inexacte data

Het negeren van inexacte data kan dus ernstige gevolgen hebben. Een mogelijkheid is om alle data die inexact zijn niet mee te nemen in ons model. In dat geval gooien we echter waardevolle informatie weg. Het is zelfs zo, dat zonder de informatie in deze data we niet eens een plan kunnen opstellen dat aan alle restricties voldoet. We willen daarom in ons voorraadmodel wel inexacte data gebruiken, maar tegelijkertijd ook rekening houden met het feit dat de verkregen data inexact zijn.

In De Ruiter et al. (2014) is een methode ontwikkeld die als uitgangspunt heeft dat de beslisregels gebaseerd zijn op schatters in plaats van de werkelijke waarden. Voor de afwijking tussen de schatter en de werkelijke waarde van de vraag stellen we wederom een onzekerheidsgebied op. Zo kunnen we bijvoorbeeld het effect bestuderen als er in een bepaalde periode nog 10% onzekerheid is over onze data. Dit betekent dat de werkelijke vraag tot 10% meer of minder kan zijn dan weergegeven wordt door de beschikbare data. We krijgen nu een geheel nieuw onzekerheidsgebied dat zowel parameters, de schatters en de afwijking van de schatters bevat.





Figuur 3. Voorraadniveaus met beslissingsregels gebaseerd op inexacte data: geen overschrijdingen van grenzen.

In principe zijn we weer terug bij een bekend ARO-model, zij het met een nieuw type onzekerheidsgebied. Het fundamentele verschil met het ARO-model is dat beslissingsregels op schatters gebaseerd zijn in plaats van de werkelijke parameterwaarden, die in de praktijk niet geobserveerd kunnen worden. Er kunnen nu verschillende technieken gebruikt worden om dit probleem weer om te schrijven tot een efficiënt oplosbaar probleem. In De Ruiter et al. (2014) wordt gebruik gemaakt van de zogeheten Fenchel dualiteit theorie, die recentelijk toegepast is op robuuste optimalisatie door Ben-Tal et al (2014). Het resulterende probleem wordt weer een lineair programmeringsprobleem dat we efficiënt op kunnen lossen met simplex- of inwendige punt methoden.

Het resulterende productieplan blijft nu binnen de grenzen voor het voorraadniveau blijft, zoals te zien is in figuur 3.

## Conclusie

Bij het maken van beslissingen moeten we er rekening mee houden dat verkregen data inexact kunnen zijn. Doen we dit niet, dan kan er aan verschillende restricties in het model niet worden voldaan. Zo kan de opslagcapaciteit overschreden worden of kan het voorraadniveau beneden gewenste minimumwaarden zakken. Voor multi-stage problemen hebben we beslissingsregels gecreëerd die rekening houden met de inexacte aard van data. In De Ruiter (2014) wordt een model beschreven dat dit probleem, met beslissingsregels gebaseerd op inexacte data, efficiënt op kan lossen. Met deze methode kunnen we de inexacte, desalniettemin

waardevolle, data gebruiken om het productieplan op te stellen. Deze nieuwe methode is niet alleen toepasbaar op voorraadproblemen, maar is ook te gebruiken voor andere multi-stage problemen waar verkregen data inexact zijn. Voorbeelden hiervan zijn medische toepassingen en voor de financiële wereld in portfolio management.

## LITERATUUR

- Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2014). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2), 351–376.
- Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust optimization*. Oxfordshire: Princeton University Press.
- DeHoratius, N., & Raman, A. (2008). Inventory record inaccuracy: an empirical analysis. *Management Science*, 54(4), 627–641.
- Haug, A., Zachariassen, F., & Liempd, D. van (2011). The costs of poor data quality. *Journal of Industrial Engineering and Management*, 4(2), 168–193.
- Ruiter, F. J. C. T. de, Ben-Tal, A., Brekelmans, R. C. M., & Hertog, D. den (2014). *Adjustable robust optimizations with decision rules based on inexact revealed data*. (CentER Discussion Paper, 2014-003). Tilburg: Tilburg University, Operations Research.
- Soyster, A. L. (1973). Technical note – convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5), 1154–1157.

FRANS DE RUITER rondde in 2013 zijn master Operations Research aan Tilburg University af en in 2014 zijn master Applicable Mathematics aan de London School of Economics. Hij ontving in 2013 de VVS+OR master thesis prijs voor 'Adjustable robust optimization based on inexact information'. Momenteel is hij PhD student aan Tilburg University waar hij onderzoek doet naar 'Anticipative robust decision making with optimal learning'.  
E-mail: <F.J.C.T.deRuiter@uvt.nl>



## Schijnagenda's

Stel je bent tandarts met een bestand van achthonderd patiënten. De meeste afspraken betreffen halfjaarlijkse controles, een bron van gemiddelde maar zekere inkomsten. Verder houd je wat ruimte voor enkele acute gevallen en plan je het zetten van een kroon vier weken vooruit. Mevrouw de Berg heeft kiespijn, maar kan het verdragen tot volgende week vrijdag. Dan heb je een gaatje.

Toch rooster je mensen en geen afspraken. Zo worden de halfjaarlijkse controles regelmatig vergeten (zogeheten *no shows*) en vallen kostbare gaten die je niet meer kunt opvullen. De OR-termen voor de grootte van het bestand en de agenda met afspraken zijn *panel size* en *schedule*. Er is een actieve beweging die oproept om *panel sizes* te vergroten en niet langer te scheduleren. De OR-term hiervoor is *open access* (niet te verwarren met het openbaar maken van publicaties). De tandarts zou dan 's ochtends de mouwen opstropen om vervolgens maar te kijken wie zich meldt.

Toegegeven, voor een tandarts lijkt dit een brug te ver. De openaccess-beweging binnen de OR richt zich dan ook met name op huisartspraktijken, waar de wachttijden in termen van dagen inmiddels onacceptabel lang zijn. De huisartsen moeten de panel size vergroten, minder plannen, en met grote kans de zieke patiënten die zich melden dezelfde dag nog ontvangen. Ook het *noshows*-probleem is dan opgelost.

Kijken we eenmaal met een openaccess-bril dan zien we misstanden en mogelijkheden, niet alleen in

de medische wereld. Kappers, restaurants, hotels, vliegtuigmaatschappijen, collega's die zeggen altijd druk te zijn. Volgeboekte agenda's domineren ons leven en is dat niet te ver doorgeschoten? Als vuistregel zou ik voorstellen om te waken voor schijnagenda's. De schijnagenda bevat afspraken die geen inhoud kennen of terugkerende afspraken tussen twee of meerdere partijen, terwijl er maar een partij is voor wie het werkelijk van belang is. Laat ik putten uit eigen praktijk. Traditionele werkcolleges (instructies) behoren tot die categorie: studenten willen iets leren en spreken daarvoor op een vast tijdstip af met de docent. Het werkcollege krijgt daarmee een plaats in de weekagenda van de student en de docent. Maar studenten komen nauwelijks nog naar werkcolleges. Afschaffen. Laat ze vragen stellen op online-fora en stel *office hours* in. Zij die zich melden zullen echt iets te vragen hebben.

De huisarts, collega en docent die open access gaan betalen wel een prijs. Je moet bereid zijn om de zaken te nemen zoals ze komen, zoals de jazzmuzikant en eerste hulpverpleger gewoon zijn. Je kunt je niet langer verschuilen achter een overspannen façade. En hoewel we in Nederland op uitnodigingen krampachtig vragen of je wel of niet komt en we de biefstukjes per vier keer tachtig gram inkopen, moet er in agenda's toch ruimte te vinden zijn.

JOHAN VAN LEEUWAARDEN is hoogleraar wiskunde aan de TU Eindhoven.  
E-mail: <j.s.h.v.leeuwaarden@TUE.nl>





## OR VOOR BETERE BLOEDDOORSTROMING

LUUK BESSELINK, SAMUEL VAN BRUMMELEN, WIM DE KORT, PATRICIA ZONNEVELD & NICO VAN DIJK

Sanquin Bloedvoorziening heeft anno 2014 meer dan 50 vaste afnamecentra en meer dan 100 mobiele locaties voor de afname van bloed. Deze centra verwerken jaarlijks bijna 500.000 volbloeddonaties en 300.000 plasmadonaties. Bloeddonaties vinden plaats op vrijwillige basis. Donoren dienen dus zo klantvriendelijk mogelijk te worden behandeld. Onnodige vertragingen zijn uit den boze en wachttijden zouden tot een minimum beperkt moeten worden. Echter, met het oog op kosten is het eveneens van belang inzet van personeel bij een donorcentrum te minimaliseren. Tot dusverre wordt personeelscapaciteit bij donorcentra bepaald op basis van een simpele werklustberekening met een beperkt aantal shifts. Sanquin Bloedvoorziening en donorcentra worden dus geconfronteerd met een tweeledige en deels tegenstrijdige uitdaging:

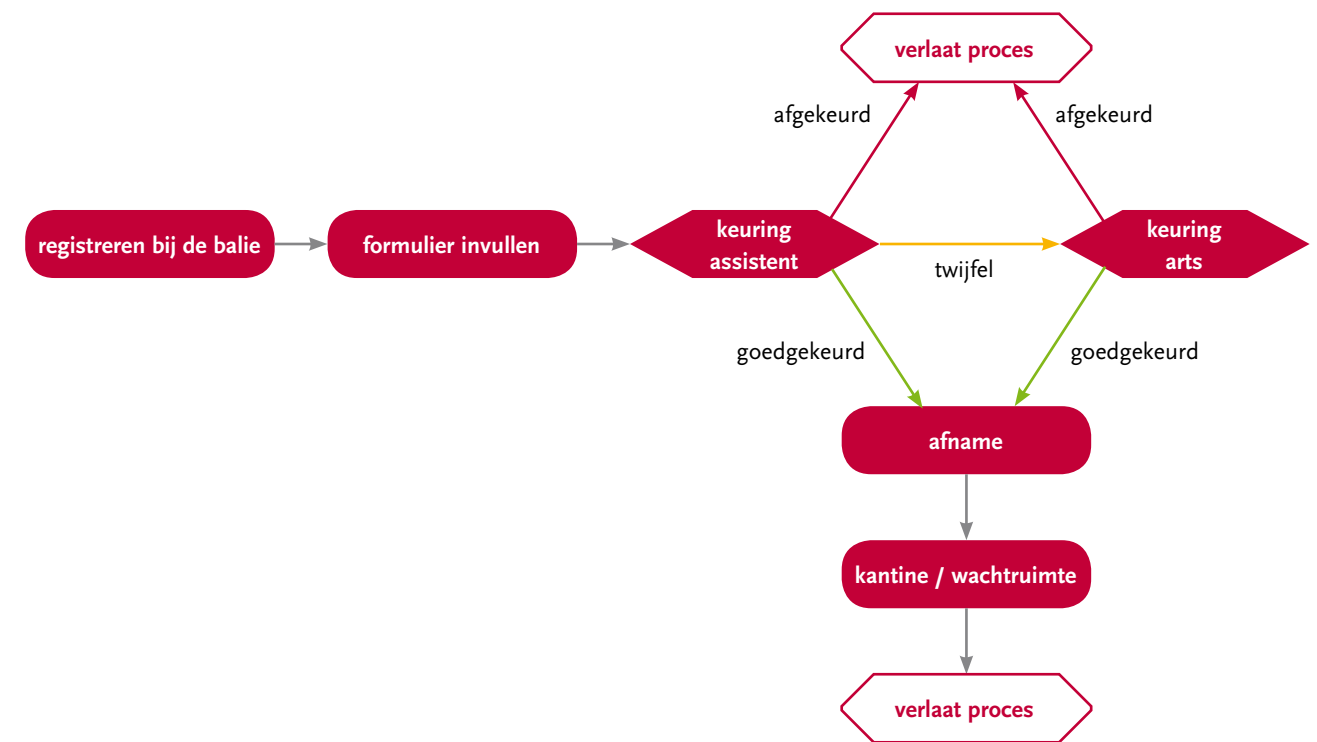
1. Het bepalen van benodigde personeelscapaciteit en shifts voor het behalen van productienormen; en
2. Het integreren van wachttijden voor donoren en wachttijdnormen.

### Het donatieproces

Bij binnenkomst bij Sanquin registreert een donor zich bij de balie. Van de baliemedewerker krijgt de donor een formulier die hij/zij invult in een wachtruimte. Door een keuringsmedewerker wordt hij vervolgens vanuit de wachtruimte meegenomen naar een keuringskamer om het formulier te bespreken en om de bloeddruk en HB-waarde te bepalen. Bij de keuring wordt de donor goedgekeurd, afgekeurd of doorverwezen naar de arts. Figuur 1 geeft het donatieproces bij Sanquin schematisch weer.

Donoren die zijn goedgekeurd gaan naar het afnamegedeelte. Onderscheid is hier te maken in volbloeden plasmadonoren. De procestijd van afname verschilt aanmerkelijk, ruwweg 15 en 45 minuten.

Donoren komen niet gelijkmatig verspreid over een dag aan. Plasmadonoren komen op afspraak en kunnen zodoende uniform over een dag ingepland worden. Voor volbloeddonoren wil Sanquin uit oogpunt van



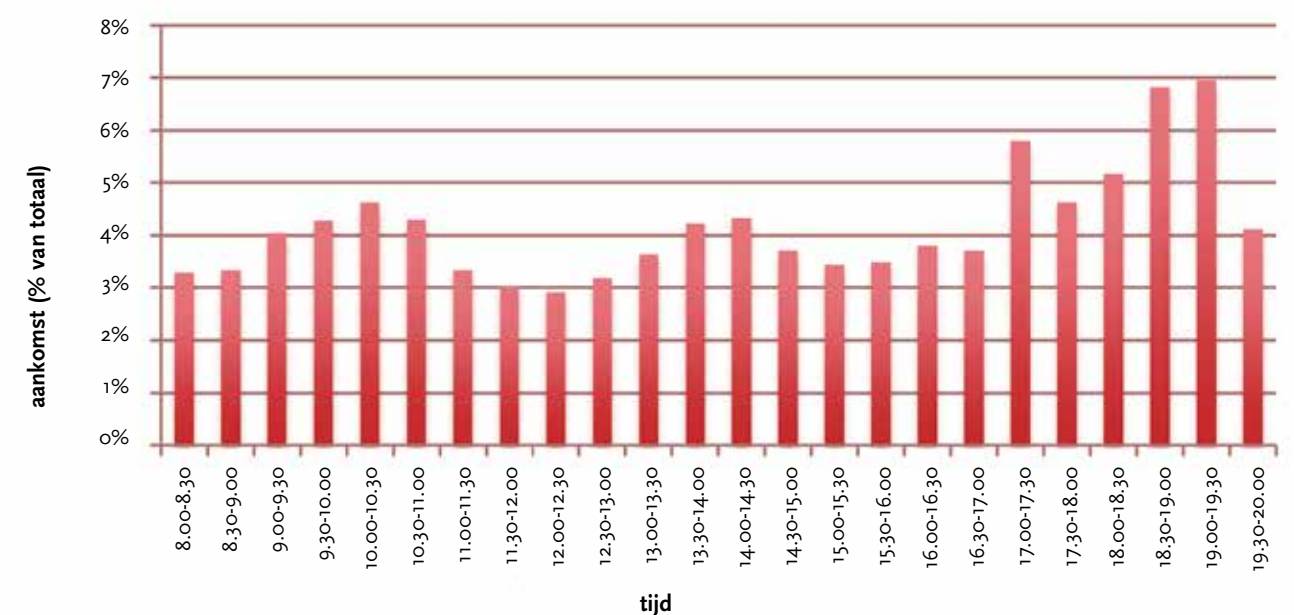
Figuur 1. Schematische weergave van het donatieproces.

donorvriendelijkheid de mogelijkheid van vrije inloop behouden. (Wel wordt binnenkort de mogelijkheid geboden om afspraken te maken via internet.)

Deze vrije inloop van volbloeddonoren vertoont een duidelijk aankomstpatroon, zoals te zien in figuur 2. Kortere shifts met meer startpunten in combinatie

met planning door middel van een Integer Linear Programming-model (ILP) zouden hierbij kunnen leiden tot een besparing in personeelsinzet.

In Blake en Shimla (2014) wordt enkel op basis van een productienorm bepaald hoeveel medewerkers er nodig zijn in elke fase van een bloeddonorcentrum.



Figuur 2. Voorbeeld aankomstpatroon.

## Wachtproblematiek

Bij het stellen van productienormen worden wacht- en doorlooptijden niet expliciet meegenomen. Wachtijd; theoretische inzichten en simulatie zijn hiertoe gewenst. Bretthauer en Coté (1998) geven een algemene formulering om personeel te plannen voor zorginstellingen, met een donorcentrum als voorbeeld. Deze referentie richt zich op het strategische en tactisch niveau en niet, zoals in dit artikel, op het operationele niveau. Sanquin stelde de volgende doelstelling:

**De ontwikkeling van een geïntegreerde tool voor bepaling en planning van personele capaciteit en productienormen, rekening houdend met wacht- en doorlooptijdsnormen.**

Bij deze doelstelling wordt geappelleerd aan twee klassieke OR-gebieden:

- ILP
- Wachtijdtheorie

## ILP-Model

Voor de planning van shifts is een ILP-model opgesteld (zie kader 1).

### Doelfunctie

Minimaliseer loonkosten op basis van arbeidsuren met:

- Diverse mogelijkheden van shifts.
- Verschillende normen: productienormen en wachtijdsnormen.

In kader 1 is het corresponderende (I)LP-model in standaardvorm weergegeven.

De coëfficiënten  $k_{s,p}$  representeren kosten verbonden aan het inzetten van de verschillende shifts, weergegeven in tabel 1. Hierbij wordt rekening gehouden met een pauze van een half uur bij shiftduren langer dan vijf uur en dat 1 langere shift de voorkeur heeft boven 2 kortere.

SHIFTDUUR	KOSTEN
2	2,00
3	3,00
4	3,99
5	4,99
6	6,50
7	7,50
8	8,49
9	9,49

Tabel 1. Kosten shifts.

Minimaliseer:

$$C = \sum_s \sum_h \sum_p x_{s,h,p} \cdot k_{s,p}$$

Onder voorwaarden:

- (1)  $y_{h,p} \geq b_{h,p}$ ,  $\forall (h,p)$
- (2)  $y_{h,p} \geq mn_{h,p}$ ,  $\forall (h,p)$
- (3)  $y_{h,p} \cdot a_h = 0$ ,  $\forall (h,p)$
- (4)  $\sum_s \sum_h x_{s,h,p} \cdot q_{s,h,h'} = y_{h',p}$ ,  $\forall (h',p)$
- (5)  $x_{s,h,p} \in \mathbb{N}$ ,  $\forall (s,h,p)$

Met

- s: shiftlengte in halve uren  
 p: de verschillende werknemertypes: artsen en assistenten  
 h: het halve uur in de te plannen periode  
 $x_{s,h,p}$ : het aantal startende shifts s voor werknemertype p in half uur h  
 $k_{s,p}$ : de kosten behorend bij shift s en type p  
 $y_{h,p}$ : het aantal werknemers in dienst van type p in half uur h  
 $b_{h,p}$ : het aantal benodigde werknemers van type p in half uur h  
 $mn_{h,p}$ : het minimum aantal werknemers van type p in half uur h  
 $a_h$ : 1 als het centrum gesloten is in half uur h, anders 0  
 $q_{s,h,h'}$ : 1 als  $h \leq h' < s + h$ , anders 0

Kader 1. (I)LP-model.

## Integratie wachtijdsnormen

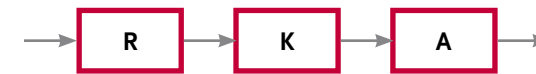
Voor de integratie van wachtijdsnormen zijn successievelijk diverse wachtijdsmodellen toegepast en onderzocht, zoals weergegeven in tabel 2, waarbij het laatstgenoemde QNA model hieronder kort zal worden toegelicht. Elk van deze modellen zijn op verschillende toepassingsniveaus gevalideerd maar dienen ook nadrukkelijk als benaderend te worden gezien.

MODEL	
M   M   1	Voor eerste wachtijdsinzichten
M   M   s	Rekening houdend met meerdere medewerkers
M   G   s	Rekening houdend met niet exponentiële procestijden
M   M   $s_i$ , $i=1,2,3$	Rekening houdend met de opdeling in verschillende fasen
QNA - netwerk	Rekening houdend met niet-exponentiële aspecten en interacties tussen fasen

Tabel 2. Wachtijdsmodellen.

Het QNA model (zie Whitt, 1983) laat zowel meerdere achtereenvolgende fasen (zie figuur 3, R: registratie, K: keuring, A: afname) als interacties daartussen toe op basis van variatiecoëfficiënten. De variatiecoëfficiënten van de servicetijden zijn gegeven door  $C_{si}$  voor  $i = 1, 2, 3$ . De variatiecoëfficiënten van de aankomstprocessen bij fasen 1, 2 en 3 zijn:

$$C_{a1}^2 = 1, \quad C_{a(i+1)}^2 = C_{ai}^2 = 1 + (1 - \rho_i^2)(C_{ai}^2 - 1) + \frac{\rho_i^2}{\sqrt{s_i}}(C_{si}^2 - 1).$$



Figuur 3. Schematische weergave van een donorcentrum.

## Opbouw tool

Vanuit Sanquin bestond de wens om voor de berekening van bovenstaande methodes een gebruiksvriendelijke tool te ontwikkelen, om invoering van de nieuwe planingswijze te vergemakkelijken. Deze tool moest universeel zijn, waardoor managers van verschillende donatiecentra deze zouden kunnen gebruiken.

Om een goede capaciteitsbepaling te berekenen zijn er per afname locatie verschillende gegevens nodig: informatie over de openingstijden van het centrum, welke shiftduren er mogelijk zijn en hoeveel donoren er worden verwacht. Hiermee wordt het minimaal aantal benodigde werknemers bepaald. Dit kan op drie manieren worden in-

gegeven. Op basis van een productienorm, op basis van een wachttijdsnorm of door middel van eigen invoer. In figuur 4 is een schematische weergave van de tool getoond.

Bij de berekening op basis van een productienorm wordt het verwacht aantal donoren per tijdseenheid gedeeld door de productienorm om tot een benodigd aantal werknemers per tijdseenheid te komen. In de huidige situatie is de productienorm ( $\eta$ ) 2.0.

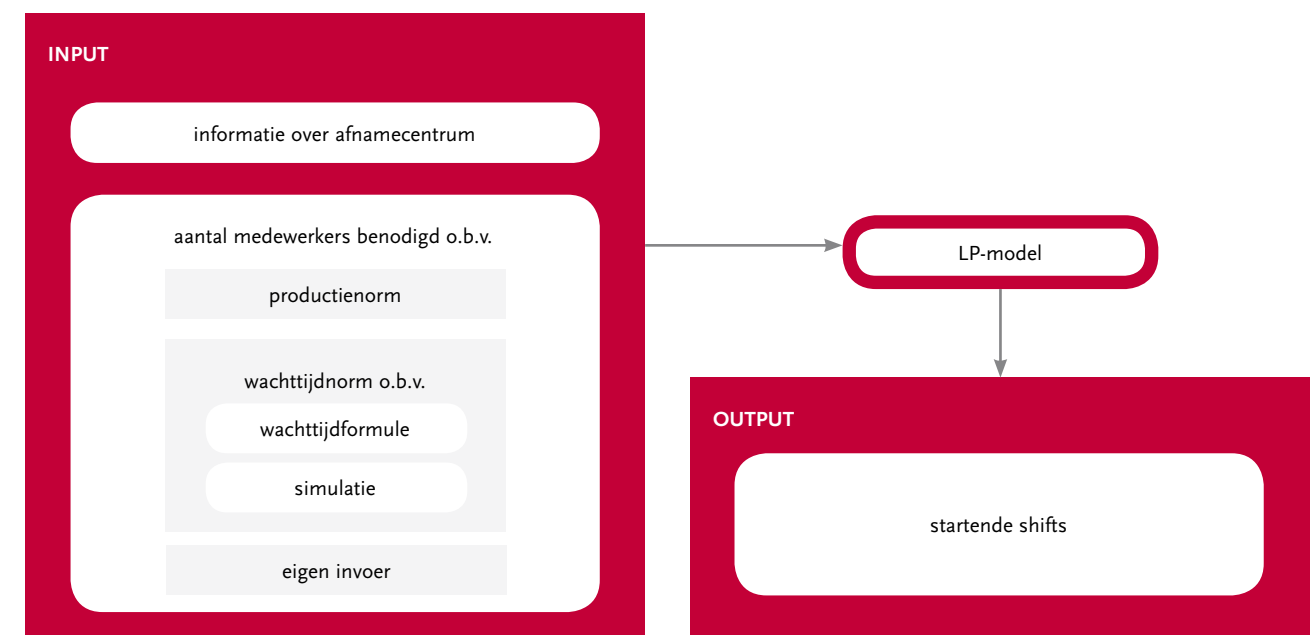
De berekening op basis van wachttijdsnormen is op twee manieren opgenomen: door middel van wachtijdsnormen en door middel van simulatie. Onder wachtijdsnormen zijn verschillende modellen mogelijk, zoals weergegeven in tabel 2. Met een simulatiemodel kunnen ook verplaatsingen van medewerkers worden gemodelleerd en geëvalueerd.

Bij eigen invoer wordt gebruikers de mogelijkheid gegeven om zelf het aantal benodigde medewerkers per half uur in te voeren.

## Testcase

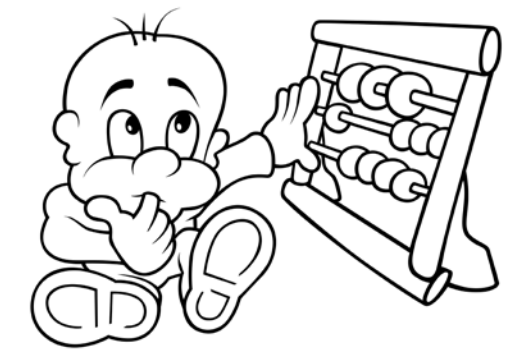
Om een indicatie te kunnen geven over de te behalen resultaten met de tool, zal voor de mobiele afname locatie (zonder plasma-afname) in Tiel het effect van de volgende scenario's op het totaal aantal arbeidsuren worden geïllustreerd op basis van:

1. het invoeren van een doorlooptijdsnorm, en
2. het inzetten van kortere, flexibele werkshifts
3. het verhogen van de productienorm van medewerkers.



Figuur 4. Schematische weergave van de voor Sanquin ontwikkelde tool.





Voordat het effect van deze scenario's op de huidige situatie kan worden bekeken, is het van belang de huidige situatie even toe te lichten. In Tiel is alleen een shift van een gehele werkdag mogelijk (12.30u – 20.00u). Er is sprake van een uur 'afsluittijd' en een half uur voorbereidingstijd. Zo komt de effectieve werktijd op 12.00u – 21.00u. Er wordt gewerkt met een productienorm van 2,0; iedere werknemer kan twee volbloeddones per uur verwerken. Het effect van de beleidswijzigingen wordt bepaald aan de hand van 3 modellen en 9 scenario's.

#### Modellen:

- A. Huidige planning (12.00u – 21.00u);
- B. Flexibele planning met 2-, 3-, 4-, 5-, 6-, 7-, 8- en 9-uur durende shifts; en
- C. Flexibele planning met 3-, 4-, 5-, 6-, 7-, 8- en 9-uur durende shifts.

#### Scenario's:

- 1.  $\eta = 2,0$ ;
- 2.  $\eta = 2,5$ ;
- 3.  $\eta = 3,0$ ;
- 4.  $\mu = 2,0$ , gemiddelde wachttijd elk half uur kleiner dan 5 minuten, QNA;
- 5.  $\mu = 2,0$ , gemiddelde wachttijd elk half uur kleiner dan 15 minuten, QNA;
- 6.  $\mu = 2,5$ , gemiddelde wachttijd elk half uur kleiner dan 5 minuten, QNA;
- 7.  $\mu = 2,5$ , gemiddelde wachttijd elk half uur kleiner dan 15 minuten, QNA;

- 8.  $\mu = 3,0$ , gemiddelde wachttijd elk half uur kleiner dan 5 minuten, QNA;
- 9.  $\mu = 3,0$ , gemiddelde wachttijd elk half uur kleiner dan 15 minuten, QNA.

Hierbij staat  $\eta$  voor de productienorm en  $\mu$  voor de servicecapaciteit. Er is een cruciaal verschil tussen beide begrippen: de productienorm is de norm, zoals vastgesteld door Sanquin, die aangeeft hoeveel donoren gemiddeld één medewerker per uur zou moeten verwerken. De servicecapaciteit is het aantal donoren dat een medewerker per uur op basis van werklust zou kunnen verwerken.

Bij scenario's 2 en 3 wordt gekeken naar het verhogen van de productienorm, bij scenario's 4 en 5 naar het invoeren van een doorlooptijdnorm en bij scenario 6 tot en met 9 naar het verhogen van de servicecapaciteit na invoering van een doorlooptijdnorm.

De huidige situatie staat gelijk aan model A, scenario 1. In tabel 3 staan de effecten van bovenstaande beleidswijzigingen op het totaal aantal arbeidsuren in de huidige situatie. Het verhogen van de productienorm van werknemers leidt logischerwijs tot een afname en de invoer van een wachttijdnorm in eerste instantie tot een toename in aantal arbeidsuren. Maar in combinatie met verhoging van de servicecapaciteit kan alsnog winst worden behaald. Verder komt het positieve effect van flexibele planning naar voren en blijkt dat model C niet veel onderdoet voor model B.

SCENARIO	MODEL A	MODEL B	MODEL C
1: $\eta = 2,0$	-	- 13%	- 11%
2: $\eta = 2,5$	- 10%	- 26%	- 23%
3: $\eta = 3,0$	- 30%	- 30%	- 30%
4: $\mu = 2,0$ , E[W] < 5 min, QNA	+ 50%	+ 31%	+ 34%
5: $\mu = 2,0$ , E[W] < 15 min, QNA	+ 40%	+ 16%	+ 20%
6: $\mu = 2,5$ , E[W] < 5 min, QNA	+ 30%	+ 12%	+ 16%
7: $\mu = 2,5$ , E[W] < 15 min, QNA	+ 10%	- 6%	- 3%
8: $\mu = 3,0$ , E[W] < 5 min, QNA	+ 10%	- 3%	- 1%
9: $\mu = 3,0$ , E[W] < 15 min, QNA	0 %	- 17%	- 14%

Tabel 3. Effect van scenario's op totaal aantal arbeidsuren t.o.v. huidige situatie.

#### Conclusies op basis van de test case.

1. Flexibel plannen leidt tot een aanmerkelijke capaciteit reductie.
  2. Wachttijdnormen lijken ook te behalen met minder arbeidsuren, mits flexibel plannen wordt ingevoerd.
  3. Het verscherpen van doorlooptijdnormen brengt weinig arbeidsuren met zich mee, mits de servicecapaciteit daadwerkelijk 2,5 of hoger is.
- Vergelijkbare bevindingen vonden plaats voor vaste, middelgrote (M) en grote (L, XL) afnameplaatsen.

De toepassing en combinatie van twee klassieke OR gebieden blijkt aldus van state-of-the-art-waarde voor praktische toepassing.

#### LITERATUUR

- Blake, J. T., & Shimla, S. (2014). Determining staffing requirements for blood donor clinics: the Canadian blood services experience, *Transfusion*, 54(3), 814–820.
- Brethauer, K. T., & Côté, M. J. (1998). A Model for Planning Resource Requirements in Health Care Organizations, *Decision Sciences*, 29(1), 243-270.
- Whitt, W. L. (1983). The queuing network analyzer, *The Bell System Technical Journal*, 62(9), 2779–2815.

LUUK BESSELINK is student Econometrics & Operations Research aan de Vrije Universiteit Amsterdam; onderwerp van zijn bachelorscriptie (Universiteit van Amsterdam) was 'Capaciteitsbepaling Sanquin'.  
E-mail: <p.l.besselink@student.vu.nl>

SAMUEL VAN BRUMMELEN is promovendus bij Stochastic Operations Research (SOR) en het Centre for Healthcare Operations Improvement & Research (CHOIR), Universiteit Twente.  
E-mail: <s.p.j.vanbrummelen@utwente.nl>

WIM DE KORT is hoofdonderzoeker van donorstudies bij Sanquin en hoogleraar donorgeneeskunde aan de Universiteit van Amsterdam  
E-mail: <w.dekort@sanquin.nl>

PATRICIA ZONNEVELD is Junior Consultant bij KWPS en is in 2013 afgestudeerd aan de masteropleiding Econometrics & Operations Research (Universiteit van Amsterdam) op het onderwerp 'Capaciteitsplanning en wachttijdbepaling in de donorcentra van Sanquin'.  
E-mail: <zonneveld@kwps.nl>

NICO VAN DIJK is hoogleraar bij SOR en CHOIR aan de Universiteit Twente en bij Kwantitatieve Economie aan de Universiteit van Amsterdam  
E-mail: <n.m.vandijk@utwente.nl>

#### LUCIA DE B MOVIE

The Lucia de B movie is a psychological thriller, based on one of the most controversial court cases in the Netherlands. Lucia, a nurse, is sentenced to life imprisonment for allegedly having killed at least seven babies and elderly people. The controversies around the use of statistics in this case is of particular interest to us of course. We have organized a movie night where we watched the movie first and afterwards discussed the case with Richard Gill. This was a very successful event with about 25 Young Statisticians present. Afterwards we pursued the discussion and evening in a local bar in Leiden.

#### CALL FOR NEW BOARD MEMBERS

The current board of the Young Statisticians has been in place for over a year. Currently we are looking for new enthusiastic young statisticians who want to replace the board. Are you interested in organizing and coordinating nice activities for your fellow young statisticians? Don't hesitate to contact us. For more information e-mail Nynke (chair@youngstatisticians.nl)

#### YOUNG STATISTICIANS BOARD

- Nynke Krol – chair  
<chair@youngstatisticians.nl>
- Iris Eekhout – internal affairs  
<info@youngstatisticians.nl>
- Sanne Willems – external affairs  
<extern@youngstatisticians.nl>
- Nadia Vendrig – money affairs  
<treasurer@youngstatisticians.nl>

Stay updated on future events and **SIGN UP TO THE MAILINGLIST** info@youngstatisticians.nl. Please let us know if you are interested in becoming an active member!

## LOTTO NONSENS DE WERELD WIL BEDROGEN WORDEN

In mijn column heb ik wel eens slukreclame gemaakt voor mijn boek *Understanding Probability* als zijnde een innovatieve aanpak om kansrekening te onderwijzen aan beginnende studenten. Soms denk ik dat ik beter een ander boek had kunnen schrijven. Die gedachte kwam voor het eerst bij mij op toen ik de lottoboeken van Gail Howard zag. Stuk voor stuk bestsellers. De oudere lezers onder ons die bekend zijn met de vroegere Amerikaanse bioscoopwesterns herinneren zich vast wel de *snake oil salesman* die bijna altijd in de Western ten tonele werd gevoerd. Dit was meestal een wat oudere vadsig personage met de gave van het woord die de argeloze boeren en buitenlui allerlei waardeloze gezondheidsdrankjes en zalfjes wist aan te smeren. Gail Howard handelt met haar lottosystemen in de beste Amerikaanse tradities van de *snake oil salesman*. Het zogenaamde systeem van gebalanceerde getallen is één van haar paradepaardjes. Haar trouwe volgerschare imponeert ze met de mededeling dat dit lottosysteem gebaseerd is op het principe van de *bell curve* dat ze toeschrijft aan de grootste wiskundige aller tijden, Carl Friedrich Gauss.

Hoe werkt dit lottosysteem? Het is geniaal in zijn eenvoud. Laten we voor de uitleg de lotto 6/49 nemen waarin zes verschillende getallen random worden getrokken uit de getallen 1 tot en met 49. Voor deze lotto word je geadviseerd de zes getallen zo te kiezen dat de som van de getallen tussen 117 en 183 ligt. De advies berust op het feit dat de som van de zes getrokken getallen bij goede benadering normaal verdeeld is, zoals uit simulatie blijkt. De verwachtingswaarde en de spreiding van deze benaderende normale verdeling zijn 150 en

32,8. Bij de normale verdeling met verwachtingswaarde  $m$  en spreiding  $s$  ligt ongeveer 68% van de kansmassa tussen  $m-s$  en  $m+s$ . Vandaar het advies de zes getallen zo te kiezen dat de som tussen 117 en 183 ligt. Dit zou de kans verhogen om zo een prijs in de lotto winnen. Klinkklare onzin natuurlijk. Vergeten wordt te zeggen dat er veel meer combinaties van zes getallen zijn met een som tussen 117 en 183 in het midden van de verdeling van de som dan er bijvoorbeeld combinaties van zes getallen zijn met een som tussen 21 en 87 in het uiteinde van de verdeling. Als je de lotto speelt, dan wed je op de zes individuele getallen die getrokken gaan worden en niet op wat hun som zal zijn.

Na Gail Howard die zover mij bekend niet gehinderd wordt door een academische achtergrond, kwam dit jaar *The Professor from Brazil*. Zo te zien een afgestudeerd wiskundige en niet een professor zoals we die kennen uit advertenties van helderzienden in de krant, maar helemaal duidelijk is het niet. De betreffende persoon is Renato Gianella en die publiceerde in 2013 het artikel 'The Geometry of Chance: Lotto Numbers follow a Predicted Pattern' in het tijdschrift *Biometric Brazilian Journal*, voorheen *Brazilian Journal of Mathematics and Statistics*. In het voorjaar van dit jaar kreeg het artikel volop media-aandacht, waarna de auteur besloot een website op te richten waarin niet alleen voor de Braziliaanse 6/48 lotto maar voor elke andere lotto in de wereld advies gekregen kan worden hoe te spelen in de betreffende lotto. Voor mij valt geen touw vast te knopen aan het artikel van Gianella waarin de auteur zich onder meer beroept op de wet van de grote aantallen.

De methode wordt uitgelegd voor de Braziliaanse 6/48 lotto Super Sena waarin zes verschillende getallen random worden getrokken uit de getallen 1 tot en met 48. Als ik het gegoochel van Gianella goed begrepen heb, gaat zijn methode als volgt. De getallen 1 tot met 48 worden in vijf groepjes onderverdeeld en elk groepje krijgt een aparte kleur. De getallen 1 tot en met 9 krijgen de kleur geel, de getallen 10 tot en met 19 de kleur blauw, de getallen 20 tot en met 29 de kleur rood, de getallen 30 tot en met 39 de kleur violet, en de getallen 40 tot en met 48 de kleur paars. Elk rijtje van zes getallen uit de getallen 1 tot met 48 krijgt een kleurencode. Bijvoorbeeld, het rijtje 2-5-8-17-25-38 krijgt de kleurencode GGGBRV. Ter vergelijking, het rijtje 7-15-23-32-43-44 krijgt de kleurencode GBRVPP. De kern van Gianella's methode is dat sommige kleurencodes meer waarschijnlijk zijn dan andere om in de zes getallen van de loterijtrekking te verschijnen. Bijvoorbeeld, er is een kans van ongeveer 1 op 150 dat een rijtje van zes getallen uit de GGGBRV kleurengroep zal worden getrokken, terwijl de kans op een getrokken rijtje uit de kleurengroep GBRVPP ongeveer 1 op de 38 is en dus ongeveer vier keer zo groot is. Gianella probeert de lottospeler wijs te maken dat deze de kans om de jackpot te winnen kan vergroten door rijtjes in te vullen met geschikt gekozen kleurencodes die Gianella uiteraard graag levert tegen betaling. Hoewel het er hier nog dikker dan bij Gail Howard bovenop ligt dat het voorgestelde schema *bullshit* is, zijn er blijkbaar velen die Gianella's aanbeveling voor zoete koek slikken en zich niet realiseren dat er ongeveer 4 keer zoveel rijtjes met de GBRVPP kleurencode zijn dan rijtjes met de GGGBRV kleurencode.

Het is in de wereld van lottosystemen niet anders dan in de kwakzalverij waar elke keer wel weer ergens een miraculeuze behandeling opduikt. In de lotto is er maar één advies dat zinvol is en dit advies is om populaire rijtjes te vermijden. Mensen gebruiken vaak hun geboortedata, huisnummers, geluksgetallen, etc. om hun zes cijfers te kiezen. Voor de kans om alle zes getallen goed te hebben, maakt dat natuurlijk niets uit. Maar als je het rijtje 1-2-3-4-5-6 hebt ingevuld en de jackpot zou daar op vallen, dan moet je wel de jackpot met talloze anderen delen. Hetzelfde geldt voor rijtjes als 7-14-21-28-35-42. Bij het 25-jarig bestaan van de Nederlandse lotto – die zich inmiddels uit de markt heeft geprijsd door 2 euro per ingevuld rijtje te vragen waar de Duitse lotto met 1 euro volstaat – werd door de lotto bekend gemaakt dat 1-2-3-4-5-6 het meest ingevulde rijtje was in het bestaan van de lotto tot dan. Het is niet aan te raden om een dergelijk rijtje in te vullen. Nogmaals, voor de winkans maakt het niets uit, maar wel voor de uitbetaling in het uiterst onwaarschijnlijke geval dat de jackpot op dit rijtje zou vallen.

Ten slotte, de schrijver van deze column heeft iets in de aanbieding en wel een systeem voor roulette waarmee de speler niet kan verliezen. Uiteraard geef ik dit systeem niet gratis weg in deze column. Degenen die er interesse in hebben, kunnen het systeem tegen betaling bij mij bestellen, waarbij de lezers van *STATOR* 50% korting krijgen.

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



# Bayes, Price en de orde van de dingen

RICHARD STARMANS

De 19e eeuwse bioloog Thomas Huxley (1825–1895) is vandaag de dag vooral bekend als de flamboyante, maar uiterst vileine provocateur en agitator, die in de tweede helft van de 19e eeuw de evolutieleer van Charles Darwin te vuur en te zwaard verdedigde. Anders dan de terughoudende Darwin zelf, schuwde Huxley het publieke debat niet. Hij benoemde zichzelf tot Darwins Bulldog, schokte het grote publiek met soms groteske verhalen over aapmens en vocht met sardonisch genoegen polemieken uit met kerkelijke en wereldlijke autoriteiten. Zijn notoire debat met de anglicaanse bisschop Samuel Wilberforce is bijkans klassiek. Toch was Huxley allerminst een darwiniaan van het eerste uur. Hij twijfelde nadrukkelijk aan Darwins idee over gradualisme en plaatste vraagtekens bij het mechanisme van de natuurlijke selectie. Zijn verdediging van Darwin was dan ook niet vrij van eigenbelang. De zelfverklaarde agnosticus Huxley was bovenal een apologet van de conflicthypothese, de opvatting dat wetenschap en geloof altijd en onvermijdelijk met elkaar in staat van oorlog verkeren of ten minste contradictoir zijn.

Deze opvatting vond bekende en radicale protagonisten in Huxley's tijdgenoten John Draper en Andrew White en wordt daarom ook de Draper-White these genoemd. Deze manifesteert zich dikwijls in een gekleurde, onvolledige of onjuiste weergave van de ideeëngeschiedenis, waarin de verworvenheden van bij voorbeeld de Middeleeuwse filosofie worden ontkend of gebagateliseerd, het conflict tussen Galileï en de kerk uitsluitend als een godsdienstconflict wordt opgevat, en de wetenschappelijke revolutie van de 17e eeuw wordt geduid als een abrupte omwenteling, die aan eeuwen van achteruitgang of stilstand een einde maakte. De tegenaanval in de wetenschapsgeschiedenis werd dan ook al spoedig ingezet, reeds rond 1900 door de bekende natuurkundige en historicus Pierre Duhem. Ofschoon de meeste hedendaagse historici de Draper-White these als ongeënuanceerd terzijde hebben geschoven, wordt zij in de populaire literatuur van bijvoorbeeld Dennett en vooral Dawkins nog steeds hartstochtelijk omarmd. Al met al zocht en vond Huxley in Darwins werk de munitie die hij nodig had voor zijn privé-oorlog en dankt hij zijn huidige reputatie grotendeels aan zijn 'liaison' met Darwin.

Toch was zo'n bondgenootschap niet ongebruikelijk. Veel vermaarde wetenschappers uit het verleden konden terugvallen op invloedrijke vrienden of tijdgenoten, die zich al dan niet gedreven door waarheidsliefde of eigenbelang, opwierpen als steun en toeverlaat, beschermheer, geldschieter of uitgever, volgeling, exegeet of apologet. Niet zelden leidde de wijze waarop zij zich van deze uiteenlopende taken kweten ertoe dat ook hun namen in de annalen van de wetenschapsgeschiedenis konden worden bijgeschreven.

## Inverse kansen

Ook de geschiedenis van de statistiek kent zulke liaisons. De relatie tussen Thomas Bayes (1702–1761) en de politiek denker, moraaltheoloog en wiskundige Richard Price (1723–1791) is zonder twijfel een van de bekendste. Precies 250 jaar geleden, verscheen in de *Philosophical Transactions of the Royal Society* Bayes' vermaarde essay onder de ogenschijnlijk weinig sprankelende titel 'An Essay towards solving a Problem in the Doctrine of Chances. By the late Rev. Mr. Bayes, F.R.S. communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S.'. Het artikel zou uitgroeien tot één van de meest invloedrijke bijdragen in de geschiedenis van wetenschapsfilosofie en statistiek en de bakermat vormen van het huidige dominante Bayesiaanse paradigma. Het verhaal is genoegzaam bekend. Price trof in Bayes' nalatenschap een amper 20 pagina's tellende ongepubliceerde tekst aan, waarvan hij onmiddellijk relevantie en reikwijdte begreep. Hij redigeerde het artikel, gaf in 1763 een voordracht erover bij de Royal Society, en voorzag het van een belangrijke en in menig opzicht zelfs visionaire introductie. Ook in de jaren daarna zou Price zich blijven opwerpen als onvermoeibaar pleitbezorger van Bayes' ideeën over waarschijnlijkheid, inductie en redeneren, deze interpreteren en verder ontwikkelen. Hierdoor werd Richard Price veel meer dan Bayes zelf, de heraut van het huidige Bayesianisme.

In zijn retorisch zeer vernuftige introductie positioneert Bayes het essay in het vigerende 18e eeuwse debat over de nog prille waarschijnlijkheidsleer als een belangrijke reactie op en verbetering van het werk van Abraham



de Moivre (1667–1754) 'Mr. De Moivre, indeed, the great improver of this part of mathematics, has in his *Laws of chance, after Bernoulli, and to a greater degree of exactness, given rules to find the probability there is, that if a very great number of trials be made concerning any event, the proportion of the number of times it will happen, to the number of times it will fail in those trials, should differ less than by small assigned limits from the proportion of its failing in one single trial. But I know of no person who has shown how to deduce the solution of the converse problem to this; namely, "the number of times an unknown event has happened and failed being given, to find the chance that the probability of its happening should lie somewhere between any two named degrees of probability".*'

Essentieel was inderdaad Bayes' conceptie van omgekeerde of inverse waarschijnlijkheid, gebaseerd op de notie van voorwaardelijke kans, die in Bernoulli's tijd nog nauwelijks was ontwikkeld. Jakob Bernoulli (1654–1705) had een concept van directe waarschijnlijkheid ontwikkeld, waarbij op basis van een bekende verdeling de waarschijnlijkheid van empirische data kon worden berekend. Hij deed dit in zijn postuum verschenen *Ars Conjectandi* (1813), die de schriftelijke neerslag en synthese bevatte van de wijsgerige en wiskundige inspanningen die gedurende de tweede helft van de 17e eeuw door onder meer Pascal, Fermat en Huygens waren verricht. Bernoulli zette daarmee een noodzakelijke tussenstap op weg naar een notie van inverse waarschijnlijkheid, die een voor wetenschap en filosofie belangwekkender vraag mogelijk maakt. Kunnen we waarschijnlijkheidsuitspraken doen over de waarheid van een hypothese op basis van beschikbare data, bijvoorbeeld frequenties of proporties. En daarmee samenhangend: hoe verandert ons geloof in een theorie over de werkelijkheid wanneer nieuwe data beschikbaar zijn? Zo'n omkering of inverse is thans onlosmakelijk met inferentiële statistiek verbonden, maar was indertijd problematisch. Dat lag deels aan de notie van voorwaardelijke kansen. De opvatting dat kansen op gebeurtenissen feitelijk altijd voorwaardelijk zijn, dat wil zeggen, conditioneel ten opzichte van andere gebeurtenissen, achtergrondkennis of referentieklassen, was in grote lijnen bekend. Maar bij het ontwikkelen van dit concept lag het accent in eerste instantie op een

temporele relatie, die niet verder ging dan 'na gegeven voor'. Een relatie van 'voor gegeven na' was problematisch, zo niet absurd. Bayes onderkende de symmetrie en exploreerde in zijn essay de mogelijkheden van deze omkering. Op die manier konden kansen van oorzaken worden geschat, welke leidden tot de geobserveerde gebeurtenissen. Zo kon worden geredeneerd van gevolg naar oorzaak, van gegevens naar theorie, van bekend naar onbekend, van eindig naar oneindig. Hiermee gaf Bayes een belangrijke aanzet tot inferentiële statistiek.

Vervolgens benadrukt Price de wezenlijke filosofische betekenis van het essay. Hij stelt dat 'every judicious person' zal beamen dat de onderhavige problematiek niet louter 'a curious speculation in the doctrine of chances' betreft, maar opgelost dient te worden 'to a sure foundation for all our reasonings concerning past facts, and what is likely to be hereafter'. Dat geldt met name voor een ieder die 'would give a clear account of the strength of analogical or inductive reasoning; concerning, which at present, we seem to know little more than that it does sometimes in fact convince us, and at other times not'.

## Rationele theologie

Het is ontegenzeggelijk de verdienste van Price dat hij inzag welke fundamentele consequenties de omkering van de kansen door Bayes had voor de kennisleer en vooral voor het streven naar een universele rationele calculus, die Pascal, Leibnitz en Hobbes al in de 17e eeuw hadden bepleit. Price wees de Royal Society op de verbondenheid tussen het wijsgerige begrip inductie en de nog prille wiskundige notie van waarschijnlijkheid, waarmee Bayes het pad effende voor latere pogingen het inductieprobleem op te lossen, dat in die tijd uiteraard met het werk van Hume verbonden was.

De naam van David Hume (1711–1776) wordt in het essay niet genoemd. Toch vormde de Schotse scepticus, die onder meer door zijn befaamde essay *On Miracles* (1748) bekend stond als notoir tegenstander van de mogelijkheid van wonderen, een belangrijke drijfveer voor Price om een voor hem cruciale passage toe te voegen aan zijn introductie: 'The purpose I mean is, to shew what

reason we have for believing that there are in the constitution of things fixed laws according to which things happen, and that, therefore, the frame of the world must be the effect of the wisdom and power of an intelligent cause; and thus to confirm the argument taken from final causes for the existence of the Deity. It will be easy to see that the converse problem solved in this essay is more directly applicable to this purpose; for it shows us, with distinctness and precision, in every case of any particular order or recurrency of events, what reason there is to think that such recurrency or order is derived from stable causes or regulations in nature, and not from any irregularities of chance.' Zoals vele wetenschappers en filosofen in de 18e eeuw bedreef de christelijk apologet Price een vorm van rationele theologie, en wilde hij de grote en goddelijke orde der dingen door middel van wetenschap aantonen, aannemelijk maken, verhelderen of illustreren.

Dat alles gold nadrukkelijk ook voor vele pioniers van de nog weinig gezaghebbende leer van de kansrekening en statistiek. Blaise Pascal, die zelf wel degelijk geloofde in de fundamentele betekenis van wonderen voor het Christendom, ontwikkelde zijn beroemde Weddenschap en zijn toepassing van probabilistische argumenten op een zo gewichtig terrein, was een belangrijke stimulans voor en rechtvaardiging van de nieuwe manier van redeneren en een indicatie van de reikwijdte ervan. Er zouden spoedig vele probabilistische argumenten volgen, die ook beter pasten in de bestaande rationele theologie. Illustratief is in dit opzicht het werk van de Britse arts en wiskundige John Arbuthnot (1667–1735), die onder meer Huygens' kanstheoretische werk in het Engels vertaalde. Hij ontwikkelde een primitieve notie van statistische significantie-testen, die hij onder meer toepaste in zijn *An Argument for Divine Providence, taken from the Constant Regularity observed in the Births of both Sexes*, eveneens gepubliceerd bij de Royal Academy. Hij stelt daarin dat 'Among innumerable footsteps of divine providence to be found in the works of nature, there is a very remarkable one in the exact balance that is maintained between the numbers of men and women; for by this means it is provided, that the species may never fail, nor perish, since every male may have its female, and of a proportional age. This equality of males and females is not the effect of chance but divine providence, working for a good End, which I thus demonstrate', waarna hij de man-vrouw verhouding door middel van een uitvoerige statistische analyse beargumenteert. De Duitse statisticus en demograaf Johann Süßmilch (1707–1767) zou medio 18e eeuw vergelijkbare ideeën verkondigen in zijn *The Divine order in the changes in the human sex from birth, death and reproduction of the same*. Theoretische

statistici als De Moivre en Bernouille, maar ook de Nederlandse Newtoniaan Willem Jacob 's Gravensande gaven middels een probabilistische invalshoek vorm aan de rationele theologie, die mede daardoor tot ver in de 19e eeuw zou floreren.

Evenals Huxley voerde Price dus ook een eigen agenda, al kon het contrast tussen beide inhoudelijk nauwelijks groter zijn. Bovendien koesterden Price en Hume een groot wederzijds respect en voerden een hoffelijk debat over wonderen en betrouwbaarheid van getuigenissen. Hume, die zelf nauwelijks wiskundig was onderlegd en een soort common sense kansbegrip hanteerde, had ontzag voor Price's kennis van de nieuwe leer der kansen. Hun gedachtewisseling zou een opmaat vormen voor een langdurig debat dat ook nu nog actueel is. Vele latere denkers zouden Hume kritiseren op probabilistische gronden, waaronder de 19e eeuwse pionier van de informatica Charles Babbage. Nog recentelijk hekelde de 'Bayesiaan' John Earman in zijn *Hume's abject failure* (2000) de onbekendheid van Hume met de nieuwe calculus en baseert zich onder meer op Price. Een contemporaine filosoof als Richard Swinburne verdedigt zijn theïstische visie nadrukkelijk met behulp van Bayesiaanse middelen. En onlangs beargumenteerde de historicus Stephen Stigler in zijn artikel 'The true title of Bayes' essay' (2013) dat Price eigenlijk een andere, meer informatieve titel had gewenst als retorisch wapen in de strijd met Hume.

De hoge vlucht die het Bayesiaanse paradigma nu heeft genomen in statistiek, filosofie en informatica kon Bayes kon uiteraard niet voorzien. Zoals bekend is zelfs de basale, maar beroemde naar hem vernoemde stelling in het essay niet terug te vinden en het is twijfelachtig of Bayes volledig zou hebben ingestemd met Price's apologie en exegese van zijn werk. De vooruitziende blik van zijn vriend was echter onmiskenbaar en droeg in belangrijke mate bij aan de klassieke status die het 250 jaar geleden gepubliceerde essay inmiddels heeft verworven.

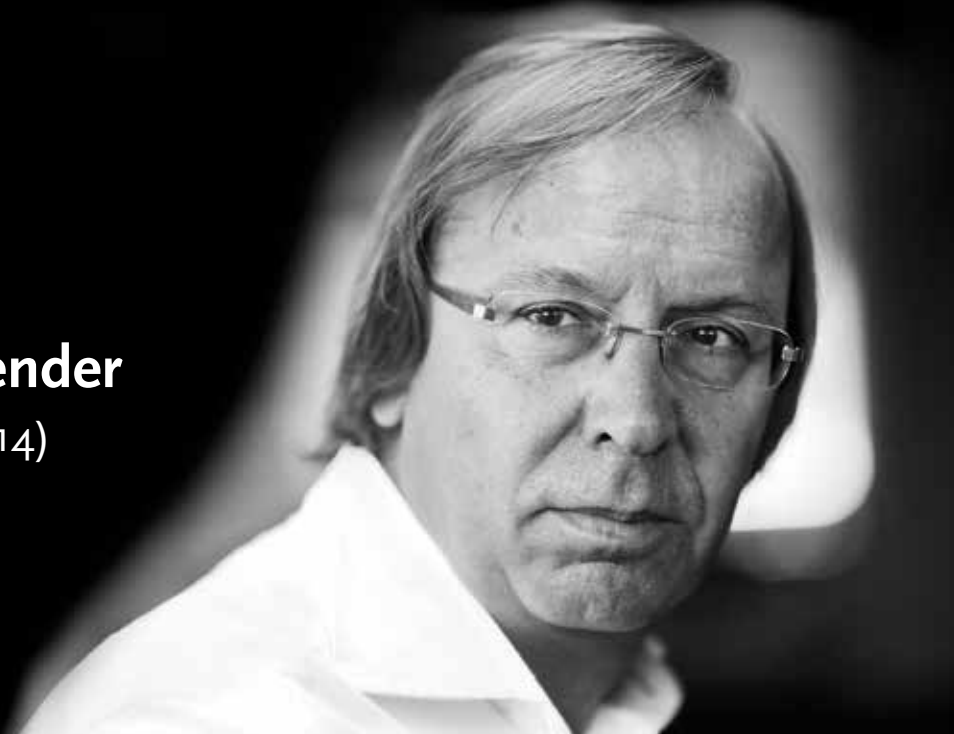
#### LITERATUUR

- Stigler, S. M. (2013). The True Title of Bayes's Essay. *Statistical Science*, 28(3), 283–288.
- Earman, J. (2000). *Hume's abject failure: The argument against miracles*. Oxford: Oxford University Press
- Swinburne, R. (2004). *The Existence of God* (Rev. ed.). Oxford: Oxford University Press.

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

#### IN MEMORIAM

## Guus Boender (1955–2014)



In de nacht van 14 op 15 september is prof. dr. Guus Boender overleden. Rondom de afgelopen jaarwisseling werd onverwacht duidelijk dat hij ongeneeslijk ziek was. Aanvankelijk had een intensieve behandeling beter dan verwachte resultaten, maar vanaf augustus verslechterde de situatie snel.

Guus Boender is geboren en getogen op Tholen om vervolgens in Rotterdam econometrie te studeren. Mede door de inspirerende begeleiding van Alexander Rinnooy Kan bleek Guus Boender een groot talent. Zijn afstudeerscriptie leidde al tot een publicatie in *Mathematical Programming*. Tijdens de promotieperiode werkte Guus Boender op het raakvlak van statistiek en optimalisatie. In 1984 resulteerde dit in een proefschrift over Bayesiaanse stopregels. Toepassingen van zijn proefschrift variëren van een realistische schatting van het aantal woorden dat Shakespeare kende tot regels over het beste moment om de zoektocht naar het globale extreem van een functie te staken.

Ook na zijn promotie hield Guus Boender altijd een rol in de wetenschappelijke wereld. In 1988 werd hij hoogleraar aan de Erasmus Universiteit en later hoogleraar Asset & Liability Management aan de Vrije Universiteit van Amsterdam. Naast een respectabel aantal publicaties was hij vooral een inspirerend begeleider en een voorbeeld voor studenten en promovendi.

Groter nog was de impact van Guus Boender op de praktische toepassing van de Operations Research en de pensioenwereld. In 1981 richtte hij samen met anderen, waaronder ondergetekende, ORTEC op. Als geen ander begreep Guus Boender de kracht en beperkingen van ons vakgebied. Hij was een meester in haar essentie.

Dat is het abstraheren van de te complexe werkelijkheid, zodanig dat het resulterende model hanteerbaar wordt, maar toch een zinvolle representatie blijft van de werkelijkheid. Alleen door deze "kunst" van het modelleren te combineren met de techniek van het "oplossen van het model" komt ons vakgebied tot glorie. Het model zag hij nooit als een black box, maar als een hulpmiddel dat ons begrip van de situatie doet toenemen, waardoor verbeterde besluitvorming mogelijk wordt. Guus Boender had een groot geloof in de kracht van deze aanpak en zag geen grens aan reikwijdte en impact van toepassingen. Met een grote gedrevenheid wilde hij dit vanuit ORTEC laten zien.

Met zijn aanpak was Guus de initiator van het Asset & Liability denken, dat in de pensioensector gemeengoed is geworden. Niet alleen als directeur en boegbeeld van ORTEC Finance, maar ook als vertrouwenspersoon was de impact van Guus Boender op de pensioenwereld groot. Hij gold nationaal en internationaal als een autoriteit, maar werd nooit een betweter. In plaats daarvan was hij innemend, ontwapenend en charmant en bracht hij met zijn originele en onconventionele ideeën velen op het juiste spoor.

Zelf heeft Guus Boender de pensioenleeftijd niet mogen bereiken. Zijn gedrevenheid en werklust zouden hem echter ook zeker hebben weerhouden om ooit met pensioen te gaan. Wij verliezen in Guus een prachtig mens, die veel heeft betekend voor ons vakgebied, de pensioenwereld en zijn omgeving. Velen zullen hem missen, maar vooral zijn vrouw Marjan en zijn familie wensen we veel sterkte bij het verwerken van dit verlies.

GERRIT TIMMER





**OPROEP voor voordrachten voor de VvS+OR Master's Thesis Award en de Willem R. van Zwet Award voor een uitstekende Ph.D. thesis**

*Master's Thesis Award*

De Master's Thesis Award heeft tot doel het bevorderen van de belangstelling voor statistiek en operations research. Wij nodigen u graag uit om een excellente afgestudeerde, die naar uw mening voor deze prijs in aanmerking komt, voor te dragen. Het gaat hierbij om personen die afgestudeerd zijn in de academische jaren 2012/2013 of 2013/2014 aan een Nederlandse universitaire instelling of HBO. Het nominatieformulier dient ondertekend te worden door de afstudeerdocent.

*Willem R. van Zwet Award*

Wij nodigen u graag uit om een excellente pas-gepromoveerde, die naar uw mening voor de Willem R. van Zwet Award in aanmerking komt, voor te dragen. Het gaat hierbij om personen die gepromoveerd zijn in de academische jaren 2012/2013 of 2013/2014 aan een Nederlandse universiteit. Het nominatieformulier dient ondertekend te worden door de promotor(es).

De sluitingsdatum voor inzendingen voor beide awards is **16 december 2014**. De prijswinnaars worden uitgenodigd de prijs (een oorkonde en € 1000) persoonlijk in ontvangst te nemen tijdens de Annual Meeting van de VvS+OR, op donderdag 26 maart 2015. Ook hun afstudeerdocenten en promotores zijn dan welkom. Reglementen en nominatieformulier zijn te downloaden op de website van de VvS+OR < [www.vvs-or.nl](http://www.vvs-or.nl) >.

In het belang van de mogelijke laureaat en van het wetenschapsgebied van statistiek en operations research en de mogelijke laureaat zien wij uw voordracht graag tegemoet.

*prof. dr. Eric Cator (juryvoorzitter VvS+OR Master Thesis Award 2014) & prof. dr. Peter D. Grünwald (juryvoorzitter Willem R. Van Zwet Ph.D. Thesis Award 2014)*

**CONFERENTIE**

**The mathematics of operations research  
13-15 januari 2015 in Lunteren**

Op de jaarlijkse conferentie van LNMB (Landelijk Netwerk Mathematische Besliskunde) spreken wederom tal van interessante, internationale wetenschappers, zoals Ehud Kalai, Alexandre Proutiere, Benjamin Van Roy en Robert Weismantel. Voor informatie zie < [www.lnmb.nl/conferences/2015/](http://www.lnmb.nl/conferences/2015/) >.

**VvS+OR  
70 JAAR**

**DAG VOOR STATISTIEK EN OR 2015  
26 maart 2015**

De jaarlijkse Dag voor Statistiek en OR 2015 zal plaatsvinden op 26 maart 2015. Verdere bijzonderheden zult u kunnen vinden in het eerstvolgende nummer van *STATOR* en op de website van de VvS+OR. Reserveer alvast deze datum, de vereniging zal in 2015 zeventig jaar bestaan! Ook zal op de Dag de 5-jaarlijkse Van Dantzigprijs worden uitgereikt.

**Oproep voor voordrachten**

**VAN DANTZIGPRIJS 2015**

Eens in de vijf jaar reikt de Vereniging voor Statistiek en Operationele Research (VvS+OR) de Van Dantzigprijs uit. De volgende uitreiking geschiedt tijdens het lustrum van de VvS+OR in 2015. De prijs is ter nagedachtenis aan prof. dr. David van Dantzig, de grondlegger van de mathematische statistiek in Nederland.

De Van Dantzigprijs 2015 wordt uitgereikt aan een vakgenoot die op 1 januari 2015 niet ouder is dan 40 jaar en die de afgelopen vijf jaren daarvoor een bij uitstek waardevolle bijdrage – hetzij theoretisch, hetzij praktisch van aard – heeft geleverd aan ons vakgebied. De 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), en Peter Grünwald & Harry van Zanten (2010).

Nominaties voor de Van Dantzigprijs 2015 kunnen voor 1 januari 2015 worden gericht aan de voorzitter van de jury, Jan Karel Lenstra, e-mail <[ijkl@cwi.nl](mailto:ijkl@cwi.nl)>.