

STAtOR

periodiek van de VVS jaargang 9 nummer 3, september 2008

Ups en downs van het klimaat sinds 1000 AD

Statistiek in de medische praktijk

Polling, productie en prioriteiten

Is poker een kans- of behendigheidsspel?

In gesprek met Alexander Rinnooy Kan

European Network for Business and
Industrial Statistics

STATOR

Jaargang 9, nummer 3, september 2008

STATOR is een uitgave van de Vereniging voor Statistiek en Operationele Research (VVS). STATOR wil leden, bedrijven en overige geïnteresseerden op de hoogte houden van ontwikkelingen en nieuws over toepassingen van statistiek en operationele research. Verschijnt 4 keer per jaar.

Redactie

Goos Kant (hoofdredacteur), Ana Isabel Barros, Mirjam Moerbeek, Gerrit Stemerding (eindredacteur), Fred Steutel, Hilde Tobi, Marnix Zoutenbier.

Kopij en reacties richten aan

Prof. dr. G. Kant (hoofdredacteur), Faculteit der Economische Wetenschappen van de Universiteit van Tilburg, Postbus 90153, 5000 LE Tilburg, telefoon 013 - 4668234, mobiel 06-11045089, <G.Kant@uvt.nl>.

Bestuur van de VVS

Voorzitter: prof. dr. R. Gill <voorzitter@vvs-or.nl>
Secretaris: dr. C.G.H. Diks <c.g.h.diks@uva.nl>
Penningmeester: prof. dr. ir. C.A.G.M. van Montfort <kvmontfort@feweb.vu.nl>
Statistische dag: prof. dr. A.W. van der Vaart <aad@cs.vu.nl>
Namens de Bedrijfssectie (BDS):
prof. dr. R.J.M.M. Does <R.J.M.M.Does@uva.nl>
Namens de Biometrische Sectie (BMS):
prof. dr. A.H. Zwinderman <a.h.zwinderman@amc.uva.nl>
Namens de Economische Sectie (ECS):
dr. P.H.F.M. van Casteren <casteren@fee.uva.nl>
Namens het Ned. Genootschap voor Besliskunde (NGB):
prof. dr. J.J. van de Klundert <j.vandeklundert@math.uni-maastricht.nl>
Namens de Sectie Mathematische Statistiek (SMS):
dr. P.J.C. Spreij <spreij@science.uva.nl>
Namens de Sociaal Wetenschappelijke Sectie (SWS):
prof. dr. J.K. Vermunt <j.k.vermunt@uvt.nl>

Leden- en abonnementenadministratie van de VVS

VVS, Postbus 2095, 2990 DB Barendrecht, telefoon 0180 - 623796, fax 0180 - 623670, <admin@vvs-or.nl>. Raadpleeg onze website over hoe u lid kunt worden van de VVS of een abonnement kunt nemen op STATOR of op een van de andere periodieken.

VVS-website

<http://www.vvs-or.nl>

Advertentieacquisitie

Marieke Klein, p/a Vrije Universiteit, afdeling Econometrie & Operationele Research, De Boelelaan 1105, 1085 HV Amsterdam, <adverteren.stator@vvs-or.nl>. STATOR verschijnt in maart, juni, september en december.

Ontwerp en opmaak

Pharos / M. van Hootegem, Nijmegen

Uitgever

© Vereniging voor Statistiek en Operationele Research
ISSN 1567-3383

Inhoud

3 Een nieuw seizoen, een nieuwe aflevering van STATOR

4 Ups en downs van het klimaat sinds 1000 AD
Nanne Weber

10 Statistiek in de medische praktijk
Mark J. van der Laan

14 Polling, productie en prioriteiten
Erik Winands

18 Is poker een kans- of behendigheidsspel?
Ruud Hendrickx

22 Tweeënzestig jaar STATOR? - column
Fred Steutel

24 'OR-succesverhalen beter onder de aandacht brengen.' In gesprek met Alexander Rinnooy Kan
Ana Isabel Barros en Goos Kant

27 De paradox - column
Onno Boxma

29 In memoriam Theo Runnenburg (1932-2008)
Fred Steutel

30 European Network for Business and Industrial Statistics (ENBIS)
Ronald Does

31 Agenda



Een nieuw seizoen, een nieuwe aflevering van STATOR

Na een zomer vol sportevenementen volgt een herfst zonder grote internationale gebeurtenissen. De dagen worden korter en de temperatuur gaat omlaag. Tijd om eens lekker bij de verwarming of open haard te gaan zitten met een goed boek, of deze *STATOR*.

De temperatuur varieert niet alleen over de seizoenen van een jaar, maar ook over de jaren heen. De bijdrage van Nanne Weber gaat in op het combineren van verschillende bronnen voor het reconstrueren van de temperatuur over de laatste 1000 jaar.

Wie niet houdt van een goed boek kan tijdens de lange herfstavonden natuurlijk samen met wat vrienden een gezelschapsspel spelen. Poker bijvoorbeeld. De vraag rijst nogal eens of dit een kans- of een behendigheidsspel is. In de bijdrage van Ruud Hendrickx wordt verder op deze vraag ingegaan.

Bij een gezelschapsspel gaat het natuurlijk om het spel en niet om de knikkers. Maar het winnen van een prestigieuze prijs binnen ons onderzoeksgebied is uiteraard een flinke opsteker. Zo staan we in deze *STATOR* kort stil bij het spoorboekje van de NS, dat de Franz Edelman Award heeft gekregen. Ook is er een bijdrage van Mark van

der Laan, een van de twee winnaars van de Van Dantzigprijs 2005.

Je kunt je af vragen hoe lang je moet wachten voor je zelf zo'n prijs zult ontvangen. Erik Winands gaat in zijn bijdrage in op de toepassing van wachtrijmodellen in productieomgevingen waar prioriteiten aan bepaalde producten en klanten worden gegeven.

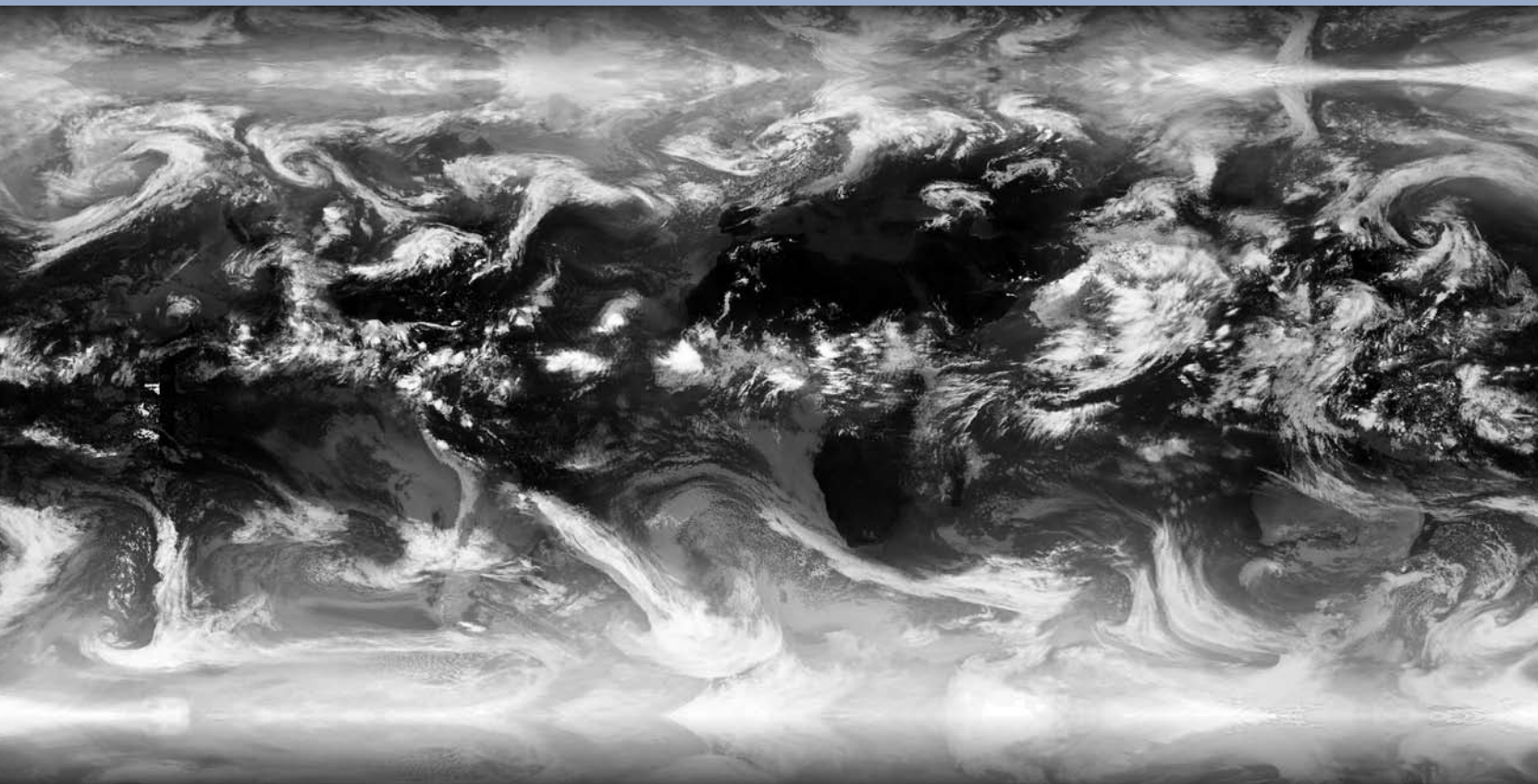
In deze *STATOR* is verder opgenomen een interview met Alexander Rinnooy Kan, één van de OR-coryfeeën in Nederland. In dit interview wordt onder meer ingegaan op de bekendheid van de OR in Nederland en onder studenten.

Fred Steutel verhaalt in zijn column over het eerste nummer van *Statistica Neerlandica*. Dit tijdschrift heeft ondertussen de respectabele leeftijd van 62 jaar bereikt.

Onno Boxma's column geeft een kort overzicht van enkele bekende paradoxen. Wie hier nog niet genoeg van heeft, kan op 23 oktober terecht bij EURANDOM. Een mooie gelegenheid voor het invullen van een herfstavond.

Veel leesplezier!

De redactie.



UPS EN DOWNS VAN HET KLIMAAT SINDS 1000 AD

Er zitten veel haken en ogen aan reconstructies van de temperatuur op basis van ‘proxy’ gegevens (indirecte klimaatinformatie). Toch willen we graag weten hoe de temperatuur in het verleden gevarieerd heeft. De onzekerheid kan ten dele gekwantificeerd worden met behulp van statistiek. Door zorgvuldig hiernaar te kijken en door informatie uit verschillende, onafhankelijke bronnen met elkaar te combineren kunnen toch conclusies getrokken worden. De recente warme jaren springen er écht uit.

NANNE WEBER

Verandert het klimaat of niet? Om deze vraag te beantwoorden kijken we meestal naar de wereldgemiddelde temperatuur, omdat dit een goede graadmeter is voor de forceringen die op

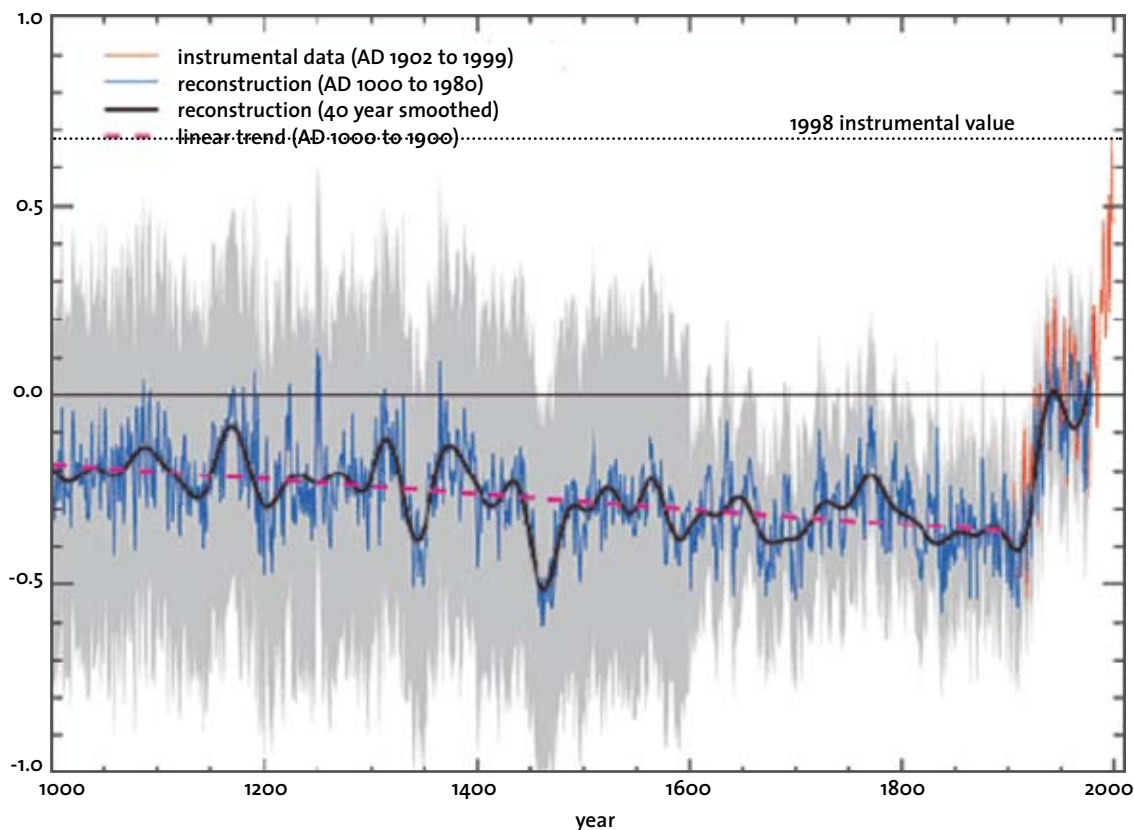
het klimaat inwerken. Pas na 1880 wordt er op zoveel plaatsen op aarde gemeten, dat we een betrouwbaar beeld hebben van het verloop van de wereldgemiddelde temperatuur. Zo heeft men

vastgesteld dat het klimaat niet constant is, maar variabel, en dat het warmer wordt op aarde.

Is de recente opwarming een trend of slechts onderdeel van een terugkerend patroon, waarin warme en koude fases elkaar afwisselen? Dit is te bepalen door naar de oorzaken van de opwarming te kijken. Ook is het zinvol om wat verder terug te gaan in de tijd. Dit kan door gebruik te maken van indirecte informatie over het klimaat zoals historische bronnen of natuurlijke archieven, bijvoorbeeld boomringen, ijskernen of oceaan sedimenten. Deze indirecte informatie over het klimaat noemen we 'proxy' gegevens (plaatsvervangende informatie, in plaats van de meetgegevens die er niet zijn voor de pre-instrumentele periode). In het derde klimaatrapport van het Intergovernmental

Panel on Climate Change (IPCC), het klimaatpanel van de Verenigde Naties, wordt een klein aantal reconstructies besproken van de gemiddelde temperatuur op het noordelijk halfrond tijdens de afgelopen 1000 jaar. Een van deze reconstructies (Mann et al, 1999) werd opgenomen in de veelgelezen *Samenvatting voor Beleidsmakers*. Deze grafiek werd na het verschijnen van het derde klimaatrapport in 2001 een beeldmerk in het klimaatdebat. En als zodanig ook onderwerp van een heftig publiek debat.

De grafiek uit de IPCC-samenvatting toont schommelingen van tienden van graden, waarbij de temperatuur meestal iets onder het gemiddelde over de instrumentele periode ligt. De 11de eeuw was relatief warm, daarna zien we een



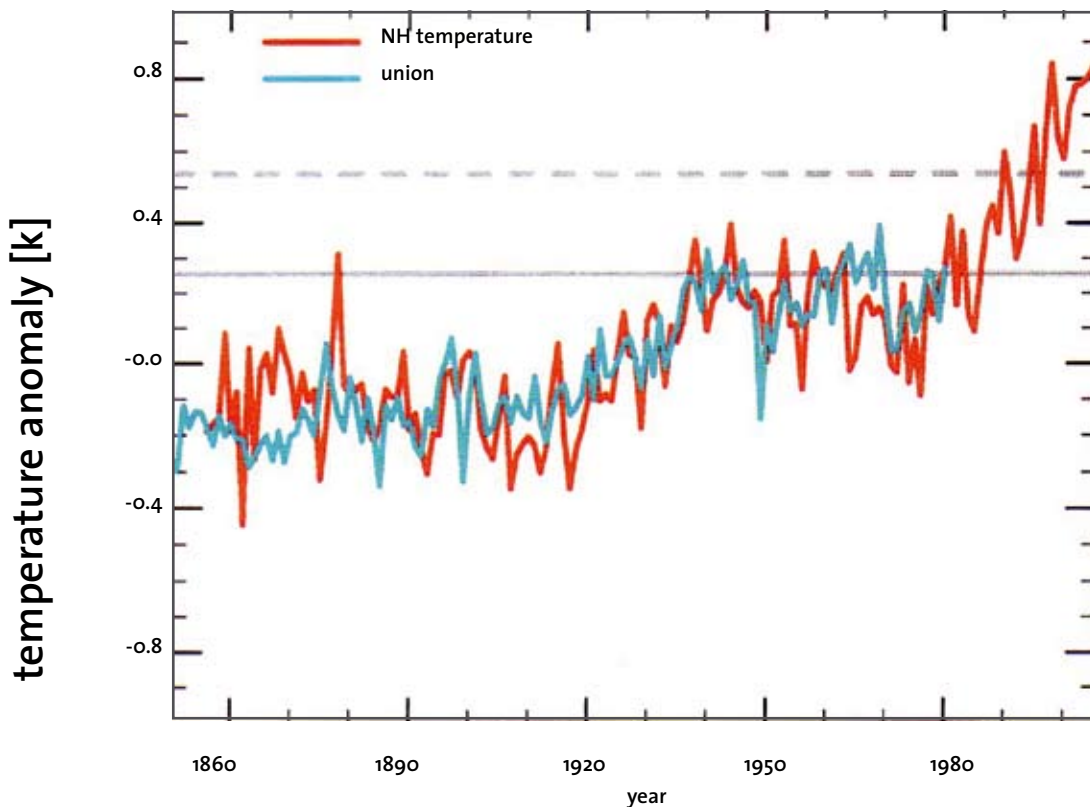
Figuur 1. Het verloop van de gemiddelde temperatuur op het Noordelijk Halfrond (afwijkingen ten opzichte van de gemiddelde temperatuur over de periode 1961-1990) volgens Mann et al. (1999). Het grijze gebied geeft de 95% betrouwbaarheidsintervallen weer, op basis van de hoeveelheid variantie in de gemeten temperatuur die verklaard wordt door de proxies. De onzekerheid in de reconstructie neemt toe in de periode voor 1600 omdat er dan minder lokale proxy reeksen zijn. Bron: IPCC, 2001.

geleidelijke afkoeling tot de 19de eeuw. De vroege 20ste eeuw is weer vergelijkbaar met de 11de eeuw, maar voor de late 20ste eeuw ligt dit anders. Vooral de laatste jaren ligt de temperatuur ver boven waarden die in het afgelopen millennium normaal waren. Een dergelijke curve wordt wel 'hockeystick' genoemd: eerst neemt de temperatuur geleidelijk af (de steel van de hockeystick) om vervolgens sterk te stijgen (het blad). Er zijn inmiddels een tiental reconstructies, die een paar

honderd tot tweeduizend jaar teruggaan in de tijd. Allen laten een vergelijkbaar temperatuurverloop zien.

Hoe maak je een reconstructie van de temperatuur?

Dit gebeurt alleen voor het Noordelijk Halfrond, omdat er voor het Zuidelijk Halfrond te weinig



Figuur 2. De gereconstrueerde temperatuur (blauw; 1850-1980) en de gemeten temperatuur (rood; 1856-2005). De correlatie tussen beide reeksen is 0.69 (met trend) en 0.44 (zonder trend), terwijl de standaardafwijking van het residu over de kalibratie periode 0.14°C is.

De grijze lijn geeft de hoogste afwijking in de temperatuur (ten opzichte van het gemiddelde over 1900-1960) die in de pre-industriële periode voorkomt – dit is in de 11de eeuw. De grijze gestippelde lijn geeft dit maximum plus twee keer de standaardafwijking. Bron: Juckes et al, 2007.

informatie is. De simpelste methode, die in de meeste studies wordt toegepast, gaat uit van het middelen van de lokale reeksen. De lokale reeksen worden eerst gestandaardiseerd, dat wil zeggen dat ze elk een gemiddelde van nul en een standaardafwijking van één krijgen. Hierdoor wegen alle reeksen even zwaar mee in het gemiddelde.

Dit gemiddelde wordt vervolgens voorzien van een schaalverdeling (gekalibreerd) door aan te nemen dat de temperatuur T lineair afhangt van het proxy gemiddelde P . Hierbij wordt de evenredigheidsconstante gegeven door de regressie van T op P of door te stellen dat de standaardafwijkingen $\sigma(T)$ en $\sigma(P)$ hetzelfde moeten zijn over de periode van overlap. Deze loopt in het algemeen van 1880 (het begin van de gemeten reeks) tot 1980 (het eindpunt van veel proxy reeksen). Idealiter wordt de overlapperiode in tweeën gesplitst. Eén deel wordt dan gebruikt voor de kalibratie en het resterende deel om de statistische significantie te toetsen. Als de toets goed uitvalt, dan gaat men ervan uit dat het gevonden verband ook geldt voor de periode waarvoor geen metingen beschikbaar zijn.

Een ingewikkelder methode ontbindt de in het verleden opgetreden klimaatvariaties in een aantal ruimtelijke patronen, zowel voor de proxy reeksen als voor de lokale instrumentele metingen. Het voordeel van deze methode is dat gaten in de ruimtelijke bedekking van de proxy data opgevuld worden en dat gebruik gemaakt kan worden van niet-lokale verbanden tussen proxies en temperatuur. Een nadeel is dat de methode minder transparant is. Deze methode wordt toegepast door Mann et al. Als eersten reconstrueerden zij daarmee de ruimtelijke verdeling van temperatuurvariaties. Dit is alleen mogelijk voor de afgelopen zeshonderd jaar, waarvoor relatief veel lokale reeksen beschikbaar zijn. Voor de afgelopen duizend jaar levert deze methode slechts een reconstructie van de gemiddelde temperatuur op. Ook nu wordt de schaalverdeling meestal ver-

kregen door inverse regressie. Lokale reeksen, die goed correleren met de gemiddelde temperatuur op het noordelijk halfrond, tellen dan zwaarder mee. Deze methode is daarom iets gevoeliger voor de selectie van lokale reeksen dan de middelingsmethode.

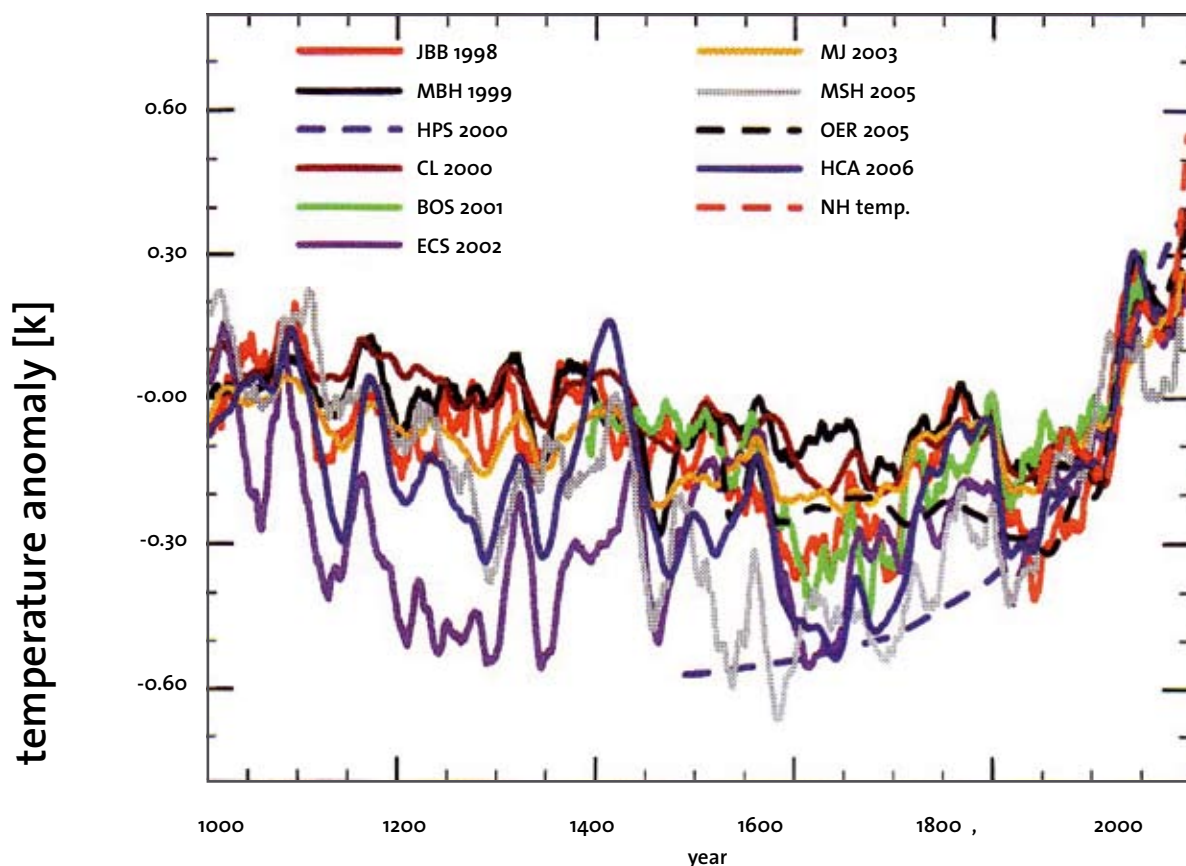
Er bestaan verschillende maten om de nauwkeurigheid van een reconstructie uit te drukken. Eén maat is de correlatie van de afwijkingen ten opzichte van het gemiddelde van elk van beide reeksen. De correlatie blijkt te variëren tussen 0,3 en 0,7, afhankelijk van de gebruikte proxy reeksen (Juckes et al, 2007). Een eenvoudige t-test leert dat dit voor tijdreeksen met zo'n honderd jaarlijkse waarden significant is met een significantieniveau van tenminste 95%. Een andere maat is de standaardafwijking van het verschil tussen beide reeksen (het residu). Deze ligt tussen 0,15 en 0,2 graden Celsius. De fout is dus kleiner is dan de temperatuurvariaties tijdens het afgelopen millennium, maar niet veel kleiner. Bovenop deze statistische onzekerheid moet men uiteraard rekening houden met niet-kwantificeerbare beperkingen van de data zelf, zoals een mogelijke slechte representatie van lange tijdschalen of veranderingen in de relatie tussen proxy grootte en klimaat.

De proxy data

Temperatuur reconstructies verschillen ook in de gebruikte gegevens. Sommige reconstructies gebruiken alleen boomringen, terwijl anderen een combinatie van verschillende soorten proxy gegevens gebruiken. De meeste studies gebruiken alleen reeksen met jaarlijkse resolutie (deze zijn makkelijker te kalibreren dan reeksen met een lagere tijdsresolutie, door het grotere aantal punten in de periode van overlap met instrumentele gegevens). Er zijn echter hooguit enige tientallen van dergelijke lange reeksen en daarom zijn verschillende reconstructies op basis van jaarlijkse

proxies voor een deel gebaseerd op dezelfde lokale gegevens. Deze reconstructies zijn niet onafhankelijk van elkaar en dus is het niet zo verrassend dat ze een vergelijkbaar temperatuurverloop laten zien. Studies, die ook of vooral reeksen met lagere resolutie gebruiken, voegen wezenlijk informatie

toe. Daarnaast zijn er kortere reconstructies, die gebaseerd zijn op volledig onafhankelijke data zoals boorgaten en gletsjerlengtes. Deze geven een aanvullend beeld van de temperatuur sinds de Kleine IJstijd.



Figuur 3. Het temperatuurverloop (21-jaar lopend gemiddelde, afwijkingen ten opzichte van de gemiddelde temperatuur in de periode 1900-1960) volgens verschillende studies.

Enkele daarvan gebruiken alleen jaarlijkse proxyreeksen (JBB1998, MBH1999, BOS2001, ECS2002, MJ2003), anderen gebruiken ook of vooral reeksen met een lagere tijdsresolutie (CL2000, MSH2005, HCA2006). Twee reconstructies zijn gebaseerd op volledig onafhankelijke gegevens, namelijk gletsjerlengtes (OER2005) en temperatuurmetingen in boorgaten (HPS2000). De gemeten temperatuur is ook aangegeven (rode lijn). Bron: Juckes et al, 2007.

Kritiek op de hockeystick

De grafiek van Mann et al heeft aanleiding gegeven tot heftige debatten op internet, in de media en in de wetenschappelijke literatuur. Volgens de critici was het onmogelijk deze te reproduceren, doordat de gebruikte data en programmatuur niet goed waren gearchiveerd. Mann et al zouden bovendien fouten gemaakt hebben in zowel de reconstructiemethode als in de dataselectie, die het eindresultaat sterk beïnvloeden (McIntyre and McKittrick, 2005). Essentieel in de kritiek op de statistische methode is dat er een ongebruikelijke standaardisatie is toegepast (namelijk voor de deelperiode 1902-1980 in plaats van de hele periode). Hierdoor heeft het eerste, dominante ruimtelijk patroon een tijdreeks met een uitgesproken hockeystickvorm. Latere studies hebben echter laten zien dat de reconstructie zelf nauwelijks gevoelig is voor deze keuze. De reden hiervoor is dat een dergelijke reconstructie opgebouwd is uit de tijdreeksen van meerdere ruimtelijke patronen, waarbij de hockeystick blijkt voort te komen uit het eerste patroon (in het geval van standaardisatie voor deelperiode) of uit bijdragen van alle patronen (hele periode). Dit resultaat is achteraf gezien niet verrassend, als men bedenkt dat ook de op de middelingsmethode gebaseerde reconstructies een hockeystickvorm vertonen.

Waarom ups en downs?

Ook in het vierde IPCC Klimaatrapport (2007) wordt gesteld dat de hierboven besproken temperatuur reconstructies het ongewone karakter van de recente opwarming bevestigen.

Het gegeven dat de recente opwarming ongewoon is, zegt niets over de oorzaken van deze opwarming; is het de menselijke invloed op het

klimaat of moeten we denken aan natuurlijke oorzaken? Grootschalige temperatuur variaties hebben soms te maken met interne processen, zoals El Niño, maar meestal hangen ze samen met variaties in de stralingsforcering. Deze variaties zijn weer gerelateerd aan externe factoren zoals vulkaanuitbarstingen, zonneactiviteit en veranderingen in de concentratie van broeikasgassen in de atmosfeer. Van deze factoren is goed (broeikaseffect) tot redelijk (vulkanen) tot matig (zonne-activiteit) bekend hoe ze inwerken op het klimaat. Ook zijn er metingen en reconstructies beschikbaar. Uit modelexperimenten en statistische analyses blijkt dat natuurlijke forceringen veel van de variaties tijdens de afgelopen duizend jaar kunnen verklaren, waarbij de menselijk invloed in de loop van de 20ste eeuw een steeds grotere rol gaat spelen. De recente opwarming is dus een trend die voorlopig wel zal doorzetten.

Een nadere uitwerking van de hier besproken punten is te vinden in Juckes et al (2007). Dit artikel is verschenen in een 'free-access' tijdschrift *Climate of the Past* (www.clim-past.net). Ook de klimaatrapporten van IPCC zijn via internet vrij toegankelijk (www.ipcc.ch).

LITERATUUR:

Juckes, M.N., Allen, M.R., Briffa, K.R., Esper, J., Hegerl, G.C., Moberg, A., Osborn, T.J., and Weber, S.L. (2007). Millennial temperature reconstruction intercomparison and evaluation, *Climate of the Past*, 3, pp. 591-609.
Mann, M.E., Bradley, R.S., and Hughes, M.K. (1999). Northern Hemisphere temperatures during the past millenium: inferences, uncertainties and limitations, *Geophys. Res. Lett.*, 26, pp. 759-762.
McIntyre, S. and McKittrick, R. (2005). Hockey sticks, principal components, and spurious significance, *Geophys. Res. Lett.*, 32, L03710, doi:10.1029/2004GL02.

NANNE WEBER is verbonden aan de Universiteit Utrecht, Faculteit Geowetenschappen, als bijzonder hoogleraar Klimaatmodellering en Klimaatanalyse. Zij is tevens werkzaam bij het KNMI als senior onderzoeker.
E-mail: weber@geo.uu.nl



STATISTIEK IN DE MEDISCHE PRAKTIJK

De Van Dantzigprijs 2005 werd uitgereikt aan biostatisticus Mark van der Laan. De prestigieuze Van Dantzigprijs wordt eens in de vijf jaar uitgereikt aan een jonge onderzoeker in de statistiek of operations research. Mark van der Laan analyseert de effecten van medische behandelingen. Hem is gevraagd enkele onderzoekprogramma's te beschrijven en de uitdagingen die hem inspireren tot verder onderzoek.

MARK J. VAN DER LAAN

Mijn statistiecarrière begon als promovendus bij Richard Gill in Utrecht, waar ik in december 1993 ben gepromoveerd. Daar kreeg ik een degelijke opleiding in statistiekonderzoek. Mijn proefschrift ging over theoretische analyse van schatters in semiparametrische modellen, in het bijzonder van schatters voor de bivariate overlevingsfunctie gebaseerd op rechtsgecensureerde

bivariate data.

Direct na mijn promotie ging ik bij James Robins (Harvard School of Public Health) werken. Hij liet mij kennis maken met een belangrijke relatie tussen efficiëntietheorie en de constructie van schattingsfuncties toegesneden op het schatten van een speciale parameter van de onderliggende verdeling.

Biostatistiek

Na aankomst in Berkeley ben ik snel naar de afdeling Biostatistiek verhuisd. Onderzoekfondsen moest ik nu bij het National Institute of Health aanvragen, en dat betekende dat onderzoekvoorstellen geïnspireerd moesten zijn door medische- of gezondheidstoepassingen en dat er samengewerkt diende te worden met wetenschappers uit dat vakgebied.

Een samenwerking met medici van UCSF (geleid door de arts Steve Deeks en mijn oud-promovenda Maya Petersen) betreft de vraag ‘Wanneer overschakelen naar een nieuw geneesmiddel?’ voor HIV-patiënten met een virus dat immuun is voor het gebruikte geneesmiddel. Tot dan toe was het idee dat het geneesmiddel onmiddellijk vervangen moest worden, maar Steve Deeks had gemerkt dat bij veel van zijn patiënten het immuunsysteem goed in stand bleef bij voortgezet gebruik van het ‘falende’ geneesmiddel. Dat wil zeggen hoewel het virus toenam, bleef de CD4-telling stabiel.

Onze statistische analyse betrof het bepalen van een oorzakelijk verband tussen de tijd tot de vervanging van het geneesmiddel en de gezondheidsontwikkeling bij de patiënt en vast te stellen wanneer van geneesmiddel te veranderen aan de hand van de CD4-telling. De analyse bevestigde het vermoeden van Steve Deeks dat dit een unieke eigenschap is van *protease* remmers, die andere veel gebruikte geneesmiddelen niet hebben. De analyse liet in feite duidelijk zien dat bij niet-*protease* remmers uitstel van medicijnvervanging schadelijk is. Deze conclusies hebben voor de hand liggende implicaties voor de AIDS-epidemie in Afrika.

Schatten van de *scientific* parameter

Ik zal een kort overzicht geven van de statistische onderwerpen waaraan ik heb gewerkt. Direct na mijn promotie raakte ik geïnteresseerd in de algemene theorie voor het construeren van lokaal efficiënte schatters gebaseerd op gecensureerde waarnemingen, gericht op causale effecten van behandelingen. Dit onderzoek is begonnen door Robins en zijn medewerkers, met de bedoeling om dit toe te passen op allerlei schattingsproblemen met hoogdimensionale en tijdafhankelijke gecensureerde en onvolledige gegevensstructuren. Dit werk resulteerde in het boek *Unified methods for censored longitudinal data and causality* (2002) dat ik samen met James Robins schreef.

Deze theorie was vooral gericht op ‘gladde’ aspecten van de onderliggende verdeling, waarvoor een efficiëntietheorie bestaat. Daarom betrof mijn tweede grote onderzoeksproject aan de data aangepaste schattingsmethoden voor niet-gladde aspecten van de verdeling, zoals een niet-gespecificeerde regressiefunctie gebaseerd op gecensureerde waarnemingen, of het schatten van een causaal effect van een medicijndosis zonder de parametrische vorm te specificeren van de dosis-respons curve. Dit betekent dat de literatuur wel een grote klasse van semi-parametrische modellen opleverde die de parameter in kwestie modelleerde en de *nuisance parameter* ongespecificeerd liet, maar geen methoden voor aan de data aangepaste selectie van het model. Zo kwam met name het probleem van een aan de waarnemingen aangepaste selectie van een model met een marginalenstructuur voor causale effecten onder onze aandacht. Mijn onderzoek betreft methoden die een aan de waarnemingen aangepaste selectie

mogelijk maken van het model voor een *scientific* (*user supplied*) parameter. Omdat de bestaande literatuur over aan data aangepast schatten vooral gericht was op niet-parametrische regressie en dichtheidsschatters gebaseerd op ongecensureerde data, was het essentieel om een algemene samenhangende methodologie te ontwikkelen gebaseerd op 'verlies', die zelfs verliesfuncties toelaat die geïndiceerd worden door ruisparameters. Deze methodologie maakt de constructie mogelijk van minimax adaptieve schatters.

Een belangrijke component van deze theorie betreft de studie van kruisvalidatie als een krachtig gereedschap voor het selecteren, uit de kandidaat-schatters, van de *scientific* parameter in kwestie, waarbij deze schatters geïndiceerd kunnen zijn door verschillende algoritmen of modellen. Dit werk resulteerde in twee artikelen met Aad van der Vaart waarin *finite-sample* resultaten worden behaald voor kruisvalidatie bij algemene (onbekende) verliesfuncties en *minimax* adaptieve schatters waarbij kruisvalidatie wordt gebruikt om modellen te selecteren.

De methodologie voor het schatten van functies door Robins en coauteurs, zoals uitgewerkt in het bovengenoemde boek, ziet af van het maximum-likelihoodbeginsel. Dit betekent in het bijzonder dat er, in het geval van meervoudige oplossingen, geen criteria zijn om uit deze oplossingen te kiezen.

Pas kort geleden hebben we een algemene *targeted* maximum likelihood methodologie ontwikkeld voor het schatten van de *scientific* parameter, die de actuele likelihood maximaliseert, maar maximaliseert in de richting die de grootste reductie oplevert in de onzuiverheid van de *scientific* parameter in kwestie. Daardoor geven deze *targeted* maximum likelihood dichtheidsschatters aanleiding tot *substitution* schatters voor de gezochte *scientific* parameter die dezelfde theoretische eigenschappen hebben als de lokaalef-

ficiënte (robuuste) schatters in de methodologie voor het schatten van functies. Maar de *targeted* maximum likelihood schatter heeft belangrijke voordelen; hij heeft, doordat de log-likelihood als criterium wordt gebruikt, niet het probleem van meervoudige oplossingen, er is geen herparametrisering nodig van de *efficient influence curve* in termen van een schattingsfunctie (wat vaak onmogelijk is). Tenslotte kunnen, naast hun aantrekkelijke eenvoud, deze schatters met standaard software berekend worden.

Mijn huidige onderzoek betreft in het bijzonder de toepassing van de *targeted* maximum likelihood methodologie op het ontdekken van biomarkers, klinisch onderzoek, en het ontwikkelen van een uitbreiding naar *targeted* Bayesiaans leren, want Bayesiaanse technieken lijden aan dezelfde *one size fits all* aanpak als maximum likelihood voor frequentisten.

Een ander interessegebied, geïnspireerd door de hoogdimensionale genomdata, betreft meervoudige toetsen, waarbij een groot aantal nulhypotesen wordt getoetst, terwijl de verdeling van het aantal of de fractie valse claims onder controle wordt gehouden. Onze resultaten op dit terrein hebben tot een nieuwe aanpak geleid voor meervoudige toetsen, waarbij gebruik gemaakt wordt van de afhankelijkheid van de toetsingsgrootheden om het onderscheidend vermogen van deze procedure te maximaliseren.

Ons werk mondde uit in het boek door S. Dutoit en M.J. van der Laan, *Resampling Based Multiple Testing with Applications to Genomics*, Springer Series of Statistics (2007)*.

Behandelingseffect

Ik zal een aantal praktijkuitdagingen beschrijven die de aanzet gaven tot mijn huidige onderzoekprogramma. In het algemeen is mijn onderzoek

op dit moment vooral gericht op open statistische problemen die voortkomen uit het door de FDA (Food and Drug Administration) geïnitieerde *critical path initiative*, dat betrekking heeft op het goedkeuren van medicijnen in gerandomiseerd klinisch onderzoek.

Tegenwoordig is het goedkeuren van medicijnen gebaseerd op het vergelijken van gemiddelden van klinische resultaten (of van overlevingsfuncties, als de resultaten overlevingstijden zijn) tussen de behandelingsgroep en de controlegroep, waarbij de basisfactoren (en de tijdsafhankelijke) bij de patiënt gemeten prognosefactoren worden verwaarloosd. Een efficiënte schatter zou het behandelingseffect moeten meten binnen elk van de strata zoals gedefinieerd door deze prognosefactoren en die dan middelen over de verschillende strata. Wij stellen – van te voren gespecificeerde – *targeted* maximum likelihood schatters voor, waarin a priori gespecificeerde regressiemodellen kunnen worden verwerkt of aan de data aangepaste algoritmen (hierbij wordt bijvoorbeeld eerst gekeken naar de belangrijkste prognosefactor). Deze *targeted* maximum likelihood schatters zijn lokaal efficiënt en volledig robuust tegen misspecificatie van het regressiemodel; ze zijn bovendien efficiënter dan de niet-aangepaste methoden, die doorgaans worden gebruikt of voorgeschreven door de FDA.

Een ander belangrijk onderdeel van het 'kritieke padinitiatief' van de FDA bestaat uit adaptieve proefopzetten. De gebruikelijke designs (proefopzetten) voor klinische experimenten vereisen a priori specificatie van de proefopzet, en het is doorgaans verboden om naar de data te kijken tot het experiment werkelijk wordt uitgevoerd. Dit heeft enorme frustratie tot gevolg voor de farmaceutische industrie; want geneesmiddelen worden vaak verboden, omdat het niet mogelijk was de proefopzet aan te passen bijvoorbeeld door het vergroten van de steekproefomvang, terwijl

de data suggereren dat het medicijn werkzaam is. Bovendien, als de heilige 0,05 grens niet wordt gehaald, moet een nieuw experiment worden bekostigd (wat gemakkelijk honderden miljoenen dollars extra kost) en het vorige experiment wordt genegeerd in het hele goedkeuringsproces.

Dit leidt tot interessante uitdagingen voor de statisticus. Wij kunnen bijvoorbeeld sequentiële designs construeren, waarbij de proefopzet voor iedere volgende groep patiënten mede bepaald wordt door wat is waargenomen bij eerdere groepen. In het bijzonder kunnen de kansen op toelating tot behandeling worden aangepast, of secundaire ingrepen die de naleving van het behandelingsprotocol verbeteren. Er kan zelfs de doelpopulatie, waaruit patiënten worden geselecteerd, worden veranderd, zodat de proefopzet gericht kan worden op deelgroepen waarvoor de behandeling effectief is. Natuurlijk moeten we er voor zorgen, als we zulke aanpassingen toestaan, dat de toets de kans op een valse claim vasthoudt op een van te voren gespecificeerd niveau van 0,05 (of wat de FDA maar voorschrijft).

Nogmaals, er lijkt geen gebrek te zijn aan interessante en belangrijke problemen om aan te werken. Door de jaren heen is mijn overtuiging dat statistiek, de wetenschap om van gegevens te leren, een ongelooflijk mooi, opwindend en belangrijk wetenschapsgebied is alleen maar toegenomen.

LITERATUUR

De artikelen van onze groep, met name die waar boven naar verwezen wordt, zijn te vinden op www.bepress.com/ucbbiostaten op mijn website www.stat.berkeley/laan

MARK J. VAN DER LAAN is hoogleraar aan de afdeling Biostatistics van de University of California te Berkeley.
E-mail: laan@stat.berkeley.edu



POLLING, PRODUCTIE EN PRIORITEITEN

Op 4 september 2007 promoveerde Erik Winands cum laude aan de Technische Universiteit Eindhoven na het verdedigen van zijn proefschrift *Polling, production & priorities* [3]. Onderstaand artikel bevat een greep uit de resultaten en observaties uit dit proefschrift waarin de toepassing van pollingsystemen in een productieomgeving centraal staat.

ERIK WINANDS

Regulering van datacommunicatie in computernetwerken. Afstelling van verkeerslichten bij een kruispunt. Instelling van liften in een flatgebouw. Cyclische productie op een chemische machine. Bezorgen van post in een kantoorgebouw.

Tussen bovenstaande toepassingen lijkt op het eerste gezicht geen enkel verband te bestaan. Toch is dit verband er wel degelijk. Ze kunnen allemaal gemodelleerd worden door een klassiek wachtrijmodel, het zogenaamde pollingsysteem. Pollingsystemen zijn intensief bestudeerd sinds de jaren vijftig en sindsdien zijn er meer dan duizend wetenschappelijke artikelen over gepubliceerd [2]. Een pollingsysteem kan als volgt worden beschreven. Er is één bediende die klanten

bedient in verschillende rijen, waarbij de rijen in een vaste volgorde worden bezocht. De klanten arriveren bij deze rijen volgens stochastische processen. Wanneer de bediende bij een rij vertrekt naar de volgende rij duurt dit enige tijd, die we in het vervolg als omsteltijd zullen aanduiden. Het aantal klanten dat per bezoek van een rij wordt bediend hangt af van de geïmplementeerde bedieningsstrategie. Een voorbeeld van een dergelijke strategie is de *exhaustive* strategie waarbij de bediende pas vertrekt bij een rij als er geen wachtende klanten meer zijn in deze rij. Een ander voorbeeld is de *k-limited* strategie. Volgens deze strategie vertrekt de bediende als de rij leeg is of als er k klanten zijn bediend.

Productieomgeving

Het onderzoek bestond uit een analyse van pollingsystemen in een productieomgeving waarbij prioriteiten aan bepaalde producten en klanten worden gegeven.

De volgende productiesituatie kan worden gemodelleerd met behulp van een pollingsysteem: één machine of productielijn waarop verschillende producten in een vaste volgorde worden geproduceerd en waar slechts één product tegelijkertijd geproduceerd kan worden. Deze situatie komt vaak voor in de praktijk, meestal in de chemische industrie. De machine modelleert men vervolgens met de bediende in het pollingsysteem, terwijl iedere rij in een dergelijk systeem een producttype representeert. De rij die bediend wordt komt ten slotte overeen met het producttype dat geproduceerd wordt.

Hier zullen we ons ook voornamelijk toespitsen op de toepassing van pollingsystemen in productieomgevingen, waar we veelvuldig teruggrijpen op resultaten uit mijn proefschrift *Polling, production & priorities*. Deze onevenredige grote aandacht voor eigen werk komt voort uit het feit dat ik al moeite genoeg heb om mijn eigen resultaten te begrijpen; laat staan die van anderen.

Kenmerken

Voordat ik me volledig richt op de productie-toepassing wil ik nog even terugkomen op de voorbeelden waarmee ik dit artikel ben begonnen. Mijn bewering dat deze allemaal door één en hetzelfde model kunnen worden beschreven zou de lezer de indruk kunnen geven dat er dus slecht eenmalig een wiskundige analyse van een pollingsysteem uitgevoerd zou hoeven te worden en dat vervolgens de verkregen resultaten rechtstreeks voor iedere toepassing gebruikt kunnen worden. Dit is echter zeker niet waar! De speci-

fieke kenmerken van het pollingsysteem hangen wel degelijk af van de onderliggende toepassing.

Het beste kan ik dit illustreren met een voorbeeld. Bekijk een kruispunt waarbij het verkeer gereguleerd wordt door middel van verkeerslichten. Als men dit kruispunt zou modelleren als een pollingsysteem, komt iedere rijbaan overeen met een rij in het pollingsysteem en de rijbanen met groen licht zijn de rijen die bediend worden. Zoals uzelf ongetwijfeld wel eens hebt opgemerkt kunnen er best twee of meer (niet-conflicterende) rijbanen tegelijkertijd groen krijgen. Dit impliceert dat in het bijbehorende pollingsysteem de bediende klanten bij twee of meer rijen tegelijkertijd kan bedienen. Wanneer men exact hetzelfde pollingsysteem als model voor een flatgebouw met liften zou gebruiken zou dit echter geen hout snijden. Of hebt u wel eens een lift gezien die op twee etages tegelijkertijd stopt? Daarentegen stelt deze lifttoepassing weer heel specifieke andere eisen aan de volgorde waarin de rijen worden bezocht, want een lift die zijn eigen leven gaat leiden en rechtstreeks van etage 1 naar 7 kan gaan komt men alleen tegen in Nederlandse films.

We hebben gezien dat iedere toepassing zijn eigen specifieke eisen aan het pollingsysteem stelt. Voor onze productietoepassing zijn dit de volgende twee:

- Regelmatig zijn de omsteltijden in het systeem extreem groot, vooral in de chemische industrie waar machines zorgvuldig moeten worden schoongemaakt vanwege veiligheidsredenen. De bijbehorende vraag is wat de impact van deze grote omsteltijden op de prestaties van het systeem is.
- Vaak is de productiemanager minder geïnteresseerd in gemiddeldes van prestatiematen, maar des te meer in de standaarddeviaties en de staartkansen. Wij hebben ons afgevraagd wat de invloed hiervan is op de optimale bedieningsstrategie.

We gaan nu verder in op beide kenmerken.

Grote omsteltijden

Waarschijnlijk is uw eerste ingeving – althans dat was voor mij het geval – dat een toename van omsteltijden altijd leidt tot langere wachttijden van klanten en grotere rijlengten in het systeem. Dit blijkt echter niet altijd het geval te zijn zoals het volgende voorbeeld uit mijn proefschrift laat zien.

Beschouw een symmetrisch pollingsysteem met twee rijen (zodat beide rijen statistisch identiek zijn), waar iedere rij volgens de exhaustive strategie bediend wordt en waar de vraagprocessen met Poissonprocessen kunnen worden gemodelleerd. Verder nemen we aan dat de productietijden (in iedere rij) deterministisch zijn ter grootte van 1 tijdseenheid. In dit systeem wordt de gemiddelde wachttijd $E[W]$ per klant gegeven door

$$E[W] = \frac{\rho/2}{1-\rho} + \frac{\text{Var}[S]}{2E[S]} + E[S] \frac{1/2 - \rho/4}{1-\rho},$$

waar $E[S]$ en $\text{Var}[S]$ respectievelijk het gemiddelde en de variantie van de totale omsteltijd per cyclus zijn en waar $0 < \rho < 1$ de totale bezettingsgraad van het systeem representeert.

Als we nu de totale omsteltijd in een cyclus verhogen met een constante δ , dan verandert deze formule als volgt

$$E[W] = \frac{\rho/2}{1-\rho} + \frac{\text{Var}[S]}{2E[S] + 2\delta} + (E[S] + \delta) \frac{1/2 - \rho/4}{1-\rho}.$$

De tweede en derde term in bovenstaande vergelijking zijn respectievelijk dalend en stijgend in δ . De gemiddelde wachttijd als functie van δ kan worden geminimaliseerd door differentiatie. De bijbehorende optimale waarde δ^* wordt vervolgens gegeven door

$$\delta^* = \max\left(\sqrt{\text{Var}[S] \frac{1-\rho}{1-\rho/2}} - E[S], 0\right).$$

Nu zien we duidelijk dat als

$$\frac{\text{Var}[S]}{E[S]^2} > \frac{1-\rho/2}{1-\rho},$$

dat dan de optimale δ^* (waarbij $E[W]$ minimaal is) groter dan nul is. Dit impliceert dat de individuele wachttijden van de klanten dalen door een toename van de totale omsteltijd met de (positieve) constante δ^* .

Dit voorbeeld laat zien dat de veel gehoorde bewering dat een reductie van de omsteltijden in een productiesysteem altijd leidt tot lagere wachttijden en kortere rijlengten niet waar is. Ondanks het feit dat deze tegen-intuïtieve observatie al eerder is gemaakt in de literatuur, besteden de meeste boeken op het gebied van productie- en voorraadtheorie hier geen enkele aandacht aan. Hoewel het zeker waar is dat in de meeste systemen een reductie van de omsteltijden leidt tot een hogere prestatie, zou een korte voetnoot die bovenstaand effect beschrijft geen kwaad kunnen.

Voor we verder gaan even een intermezzo.

U weet ongetwijfeld dat de Engelse vertaling voor het Nederlandse woord (wacht)rij *queue* is. Echter ook in het Nederlands mag het woord *queue* als geldig synoniem voor het woord wachtrij (binnen computertoepassingen) gebruikt worden. Het is leuk om te vermelden dat het woord *queue* vanuit een taalkundige optiek een bijzonder woord is. Het is namelijk het enige eenlettergrepige woord waarin vier klinkers achter elkaar staan met alle letters op oneven alfabetplaatsen [1]. Maar nu weer terug naar onze pollingsystemen.

Bedieningsdiscipline

Vaak eisen klanten dat hun order binnen een bepaalde termijn geleverd wordt door de producent, bijvoorbeeld binnen twee werkdagen. In deze situatie wordt de prestatie van het systeem niet zozeer bepaald door de gemiddelde levertijd: het is niet zo belangrijk of deze één of anderhalve werkdag bedraagt. De fractie van de producten dat echter daadwerkelijk binnen de twee dagen geleverd wordt – en dus niet naar een concurrent vertrekt – is daarentegen van het grootste belang. We zullen derhalve ook een bedieningsdiscipline moeten implementeren die niet alleen de gemiddeldes, maar zeker ook de standaarddeviaties en staartkansen van de prestatiematen kan beheersen. Een geschikte bedieningsdiscipline hiervoor is de *k-limited* strategie, die door het begrenzen van de afzonderlijke productie-intervallen er voor zorgt dat de standaarddeviatie van de cycluslengten binnen de perken blijft. Dit wordt geïllustreerd door het volgende voorbeeld.

Bekijk een pollingsysteem met twee rijen waar producten gevraagd worden volgens Poissonprocessen met gelijke intensiteiten voor beide producten van 0.375. Verder zijn zowel de productie- als omsteltijden voor beide producten exponentieel verdeeld met gemiddeldes van 1 en 0.25. Om de totale kosten te berekenen gaan we uit van een standaard lineaire kostenfunctie die zowel voorraad- als boetekosten meeneemt. Neem vervolgens aan dat product 1 een product is met hoge kosten waarvoor we de exhaustive discipline gebruiken, terwijl product 2 een goedkoper product is waarvoor de *k-limited* discipline is geïmplementeerd. In mijn proefschrift heb ik laten zien dat in dit geval de kosten maar liefst 35 procent lager zijn dan de kosten in de situatie waar beide producten volgens de exhaustive strategie worden geproduceerd.

Conclusie

Mijn hoop is dat dit artikel de lezer een idee heeft gegeven van de toepassing van pollingsystemen binnen productieomgevingen. Hoewel pollingsystemen al meer dan vijftig jaar worden bestudeerd, liggen er nog steeds veel open problemen. Vanuit de productietoepassing is het probleem waar producten sequentieel twee machines met cyclische productie moeten doorlopen heel relevant en helaas nog niet bekeken in de literatuur. Dit multi-stage probleem zou overeenkomen met een wachtrijsysteem waar twee pollingsystemen in serie staan en waar de output van het eerste systeem de input van de tweede vormt. Intuïtief lijkt de *k-limited* strategie opnieuw een goede bedieningsdiscipline voor het eerste pollingsysteem, omdat op deze manier het aankomstproces van het tweede systeem wordt gestabiliseerd. Maar kunnen we dit ook kwantificeren? En bestaan er nog betere disciplines? En welke bedieningsdiscipline moeten we voor het tweede pollingsysteem kiezen? Wordt ongetwijfeld vervolgd!

Dit artikel is gebaseerd op het promotieonderzoek dat ik de afgelopen 4 jaar heb uitgevoerd aan de Technische Universiteit Eindhoven bij de faculteit Wiskunde en Informatica en de faculteit Technologie Management onder de begeleiding van Ivo Adan, Geert-Jan van Houtum, Onno Boxma en Ton de Kok.

LITERATUUR

1. Battus, (2007). *Opperlans woordenboek* (Querido Uitgeverij, Emmen).
2. Takagi, H., (2000). Analysis and application of polling models (In *Performance Evaluation: Origins and Directions*, G. Haring, C. Lindemann and M. Reiser (eds.), Lecture Notes in Computer Science, vol. 1769, Springer, Berlin, pp. 423-442).
3. Winands, E.M.M., (2007). *Polling, production & priorities* (PhD Thesis, Technische Universiteit Eindhoven).

ERIK WINANDS is werkzaam als risk validator bij Rabobank Nederland te Utrecht en part-time als universitair docent aan de Vrije Universiteit te Amsterdam. E-mail: erik.winands@rabobank.com.



IS POKER EEN KANS- OF BEHENDIGHEIDSSPEL?

Bij roulette wordt het spelresultaat volledig bepaald door een toevalselement en dus is roulette een puur kansspel. Bij schaken daarentegen speelt het toeval in het geheel geen rol en zijn het juist de strategiekeuzes van de spelers die de uitslag bepalen, waardoor schaken een puur behendigheidsspel is. Poker bevindt zich tussen deze twee uitersten in, omdat zowel het toeval als de spelers zelf invloed hebben op de resultaten. De Wet op de Kansspelen legt restricties op aan het commercieel exploiteren van spelen waarbij de invloed van de spelers niet overwegend is. In dit artikel zetten we uiteen hoe deze invloed te meten is en welke conclusies we hieruit kunnen trekken voor poker.

RUUD HENDRICKX

De Wet op de Kansspelen stelt dat het verboden is om '...gelegenheid te geven om mede te dingen naar prijzen of premies, indien de aanwijzing

der winnaars geschiedt door enige kansbepaling waarop de deelnemers geen overwegende invloed kunnen uitoefenen, tenzij daar ingevolge deze

wet vergunning is verleend.'

Spelen die onder deze passage vallen, noemen we kansspelen. Alle andere spelen zijn per definitie behendigheidsspelen. De kwalificatie 'overwegende' laat uiteraard veel ruimte voor discussie. Dreef (2005a) heeft uitvoerig uiteengezet hoe het behendighedsniveau van spelen wiskundig gemodelleerd kan worden. In dit artikel geven we allereerst een korte samenvatting van dit algemene theoretische kader. Daarna presenteren we enkele numerieke resultaten voor diverse pokervarianten. Tenslotte bespreken we het pokerexperiment dat in het kader van dit onderzoek in samenwerking met VARA Nieuwslicht is uitgevoerd.

Theoretisch model

Impliciet opgesloten in het concept van behendigheid is dat spelers zich de strategische kneepjes van een spel eigen kunnen maken. Door te leren kunnen zij gemiddeld betere resultaten behalen dan spelers die het spel net pas leren kennen. Met dit effect van leren geven we een concrete invulling aan de invloed van de spelers op het spelresultaat.

Om te bepalen of de spelers een overwegende invloed op hun resultaat kunnen uitoefenen, voeren we het volgende gedachtenexperiment uit. Neem een startsituatie met allemaal beginners, d.w.z. spelers die de spelregels wel kennen maar nog geen benul hebben van wat verstandige strategieën zijn in het spel. Nu vervangen we één van die beginners door een optimale speler. Dit is een speler die geleerd heeft het spel zodanig goed te spelen dat zijn winst – gegeven dat hij nog steeds tegen diezelfde beginners speelt – maximaal is. Het verschil tussen deze maximale winst en de winst die de oorspronkelijke beginner had in zijn positie noemen we het leereffect.

Als volgende stap vervangen we de optimale

speler door een fictieve speler. Dit is een speler die van tevoren te horen krijgt wat de uitkomst is van alle kansmomenten in het spel en op basis van deze informatie zijn winst maximaliseert tegen weer diezelfde beginners. Het verschil in winst tussen de fictieve speler en de optimale speler is volledig toe te schrijven aan (kennis van de uitkomst van) het toevalselement en we noemen dit verschil dan ook het toevalseffect.

De relatieve behendigheid van een spel definiëren we vervolgens als het quotiënt van het leereffect en de som van het leereffect en het toevalseffect. Is dit quotiënt hoog, dan is het leereffect relatief groot, hebben de spelers een overwegende invloed op hun resultaat en is het spel een behendigheidspel. Is het quotiënt laag, dan overheerst het toevalseffect, hebben de spelers blijkbaar geen overwegende invloed en is het spel een kansspel. De schaal van relatieve behendigheid loopt van 0 (roulette) tot 1 (schaken). Voor een uitgebreide discussie over onze methodiek voor het meten van behendigheid verwijzen we naar Dreef (2005b).

Merk op dat een specificatie van de beginnersstrategie waartegen de resultaten van de drie spelertypen worden afgemeten een vrij te kiezen parameter van het model is. Onze methode staat of valt niet met een specifieke keuze van deze strategie. De behendighedsanalyse kan toegepast worden voor elke mogelijke invulling voor de beginnersstrategie. Hiermee hebben we de algemene discussie of een spel nou een kans- of behendigheidspel is weten te reduceren tot de specifieke vraag tegenover welke populatie tegenspelers je bereid bent de resultaten van de drie spelertypen af te zetten.

Waar precies de grens getrokken moet worden tussen kans- en behendigheidspel, is geen onderdeel van het model. Het model beoogt louter een ordening aan te brengen in spelen op basis van hun relatieve behendigheid. Het vaststellen van een concrete grens is de taak van de rech-

ter. Een grens van 0.2 zou consistent zijn met jurisprudentie op het gebied van de Wet op de Kansspelen. Met name de recente zaak rondom het managementspel Grand Prix Manager 2003, dat uiteindelijk door de rechter werd bestempeld als behendigheidsspel, speelt hierbij een rol.

Resultaten poker

Om de relatieve behendigheid van een spel te bepalen, is het noodzakelijk om de verwachte winst van de beginner, de optimale speler en de fictieve speler te berekenen. Bij poker heeft deze berekening heel wat voeten in de aarde, zoals we beschrijven in Hilbers et al (2008). Omdat het aantal mogelijke strategieën zo gigantisch groot is, zal de computer bij zelfs de simpelste pokervariant al snel in de problemen komen wat betreft de benodigde rekentijd en geheugenruimte. Daarom is het noodzakelijk om enerzijds de benodigde rekenpartij efficiënt aan te pakken en anderzijds om de grootte van de te doorzoeken strategieënruimte in te perken door het groeperen van min of meer gelijkwaardige strategieën.

Om het effect van de diverse spelparameters op de relatieve behendigheid van poker te meten, hebben we als uitgangspunt een groot aantal varianten van een eenvoudig pokerspel doorge-rekend. In dit zogenaamde Straight Poker krijgen alle spelers in één keer hun volledige hand, bestaande uit vijf kaarten. Er wordt gewerkt met een vaste startinzet (ante) en er is slechts één biedronde, waarin spelers wel meerdere malen aan de beurt kunnen komen wanneer er verhoogd wordt. De variabele spelparameters zijn de grootte van de inzet (bet), het aantal keren dat de inzet verhoogd mag worden (raise) en uiteraard het aantal spelers. Ook hebben we de keuze voor de beginnersstrategie gevarieerd.

Uit de analyse komt naar voren dat Straight Poker behendiger wordt naarmate de inzet groter

wordt. Het aantal toegestane verhogingen van de inzet blijkt nauwelijks van invloed. Straight Poker valt door te rekenen tot en met vier spelers en het blijkt dat de relatieve behendigheid van het spel sterk stijgt naarmate er meer spelers zijn. Het effect van de specifieke keuze van de beginnersstrategie op de relatieve behendigheid is niet uitgesproken groot. In alle varianten ligt de relatieve behendigheid boven de grens van 0.2. Voor vier spelers kan deze zelfs boven de 0.65 uitkomen. Hiermee is zelfs het op het eerste gezicht eenvoudige spel Straight Poker te classificeren als behendigheidsspel.

Pokerspellen die in de praktijk gespeeld worden zijn veelal gecompliceerder dan Straight Poker. Een in Nederland zeer populaire variant is Texas Hold'em. De extra complexiteit ten opzichte van Straight Poker zit in de gehanteerde biedstructuur en de manier waarop de uiteindelijke hand van iedere speler wordt samengesteld. Als we ons beperken tot een vereenvoudigde variant van twee-persoons Texas Hold'em waarin de vijf gemeenschappelijke kaarten (flop, turn en river) tegelijk getoond worden en er maar twee biedronden zijn, vinden we een relatieve behendigheid van 0.35-0.70, afhankelijk van de gekozen beginnersstrategie. Hieruit kunnen we met een hoge mate van zekerheid de conclusie trekken dat het uitgebreide Texas Hold'em zoals dat in de praktijk wordt gespeeld, met een groter aantal biedronden en meer spelers, een behendigheidsspel is.

Experiment VARA Nieuwslicht

In het kader van het onderzoek naar de behendigheid van (*fixed limit*) Texas Hold'em hebben collega's Borm en Van der Genugten in samenwerking met het televisieprogramma VARA Nieuwslicht in december 2007 een experiment uitgevoerd. Hiermee hebben we de problematiek van het meten van behendigheid van poker en

onze methodiek hiervoor voor een breed publiek kunnen etaleren. Voor een uitgebreid verslag van dit experiment verwijzen we naar het rapport Maaten et al (2008).

Het doel van dit experiment was tweeledig. In de eerste plaats wilden we een schatting van het theoretische leereffect en het toevalseffect zoals zich die in de praktijk voordoen, om daarmee een empirische relatieve behendigheid te bepalen. Wel zij opgemerkt dat het noodgedwongen geringe aantal waarnemingen van slechts 36 gespeelde spellen statistisch onvoldoende is om tot een aanvaardbaar redelijke schatting te komen. In de tweede plaats wilden we zicht krijgen op welke strategie een beginner hanteert, aangezien juist de beginnersstrategie een cruciale rol speelt in ons gedachtenexperiment.

Aan het experiment namen achttien spelers deel, verdeeld over drie tafels van elk zes spelers. Zestien spelers waren (relatieve) beginners. Zij kregen van tevoren een korte uitleg over de regels van Texas Hold'em. Daarnaast was er een ervaren speler, die van tevoren wist dat hij alleen tegen beginners zou spelen. Tenslotte was er een dummyspeler, die telkens via een oortje te horen kreeg welke actie hij moest ondernemen. Deze werd ingefluisterd door een expert die de handen die zijn tegenstanders kon zien en al van tevoren de spellen had geanalyseerd. Met deze laatste twee spelers, de ervaren speler en de dummyspeler plus volledig geïnformeerde expert op de achtergrond, hebben we een praktische invulling gegeven aan de optimale en fictieve speler die in de theorie een rol spelen.

Aan alle tafels werden gelijktijdig dezelfde spellen met dezelfde kaartverdeling gespeeld. De optimale en fictieve speler zaten telkens op verschillende tafels in dezelfde positie. Op deze wijze konden hun resultaten op de spellen rechtstreeks met elkaar vergeleken worden, alsmede met de beginner die op de derde tafel in dezelfde positie zat. Vergelijking van hun behaalde winsten na in

totaal 36 spellen resulteerde uiteindelijk in een primaire (naïeve) schatting van 0.05 voor het behendighedeniveau. Hierbij zij nogmaals aangetekend dat met deze schatting een hoge variatie gepaard gaat door het geringe aantal waarnemingen. De secundaire schatting, na reflectie met betrekking tot statistische uitschieters, bedroeg 0.26.

Wat betreft de beginnersstrategie zijn een aantal interessante kenmerken naar voren gekomen in het experiment. Allereerst houden beginners geen rekening met de positie die zij innemen aan tafel. Verder passen ze zelden vóór de flop. Ook zullen ze, zolang ze nog in het spel zijn, veelal pogen om zo goedkoop mogelijk de river te bereiken; zelfs met zeer goede kaarten zal een beginner niet snel verhogen. Tenslotte kunnen we stellen dat beginners weliswaar soms bluffen, maar dat de timing hiervan willekeurig is en niet, zoals bij ervaren spelers het geval is, gebaseerd op een analyse van de situatie.

Concluderend kunnen we stellen dat de resultaten van het experiment geen aanleiding geven om onze theoretische conclusie – dat Texas Hold'em een behendigheidsspel is – te herzien. Verder is er nuttige informatie boven tafel gekomen over hoe beginners spelen.

LITERATUUR

- Dreef, M. (2005a). Kans- of behendigheidsspel? Wiskunde helpt de rechter bij uitspraken op het terrein van kansspelen. *STAtOR*, december 2005/4.
- Dreef, M. (2005b). *Skill and strategy in games*. Proefschrift, Universiteit van Tilburg.
- Hilbers, P., R. Hendrickx, P. Borm en B. van der Genugten (2008). *Skill in poker*. Working paper.
- Maaten, R., P. Borm, B. van der Genugten en R. Hendrickx (2008). *Behendighedeniveau van Texas Hold'em* – Een experiment uitgevoerd in samenwerking met VARA-Nieuwslicht. URL: http://lyrawwww.uvt.nl/~pemborm/Rapport_Nieuwslicht.pdf

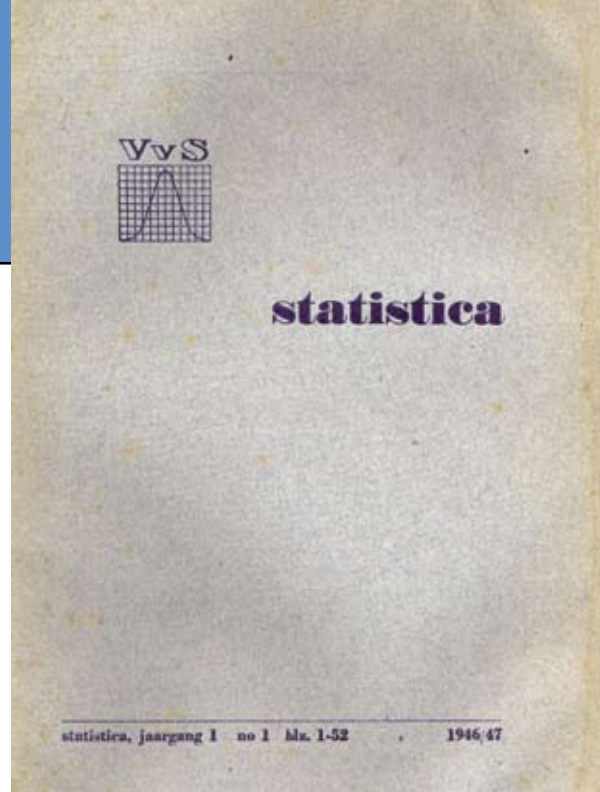
RUUD HENDRICKX is als universitair docent verbonden aan de faculteit Econometrie en OR van de Universiteit van Tilburg. E-mail: ruud@uvt.nl

Tweeënzestig jaar STATOR ?

FRED STEUTEL

Onlangs kreeg ik het allereerste (losse) nummer van *Statistica* in handen: 'statistica, jaargang 1 no 1 blz. 1-52, 1946/47'. Het ziet er heel simpel uit en een beetje smoezelig. Niet van ouderdom, maar door de uiterst sobere lay-out en de kwaliteit van het papier: simpel grijs kftje en gedrukt op 'oorlogspapier'. Maar hiermee is de statistische bescheidenheid uitgeput. *Statistica* is overtuigd van haar belang en vol vertrouwen in de toekomst – zie hieronder. De toevoeging 'Neerlandica' kwam pas in 1951. Die was nodig, omdat er ook een Italiaans tijdschrift *Statistica* (Bologna) bestond – al sinds 1941 – en nog steeds bestaat.

Het eerste *Statistica*-nummer roept allerlei nostalgische gevoelens op. De oorlog was net voorbij en iedereen wilde aan de slag. *Statistica* 1 nr. 1 oogt een beetje als *STATOR* 1, nr. 1. Het heeft ook een voorwoord van het VVS-bestuur (mede ondertekend door de net aangetreden Redactie), dat zeer voortvarend opent met 'Terecht wint de overtuiging, dat de statistische wetenschap van groote beteekenis, ja zelfs onmisbaar is voor een juiste oplossing van vele en velerlei vraagstukken, meer en meer veld.' Er is sprake van 'den afgeloopen oorlog', en 'De Vereeniging voor Statistiek'. Er wordt aangekondigd dat het tijdschrift 'elke twee maanden zal verschijnen' en er wordt een 'lange historische reeks' in het vooruitzicht gesteld. De laatste claim is niet helemaal waargemaakt; de eerste jaargang telde vijf nummers en daarna waren het er jarenlang vier; de magere jaren rond de eeuwwisseling leverden maar drie nummers per jaar op; nu zijn het er weer vier.



Het eerste nummer bevat vier paginagrote advertenties, in één waarvan het boek *De gedachting van de Statistica* door (bestuurs- en redactielid) prof. dr. S.T. Bok wordt aangeboden voor fl. 5,- (ing.). Dit boek wordt vervolgens – lovend – besproken in de sectie Boekbesprekingen, door bestuurs- en redactielid A.R. v. d. B. Met het woord 'statistica' werd toen het vakgebied van de theoretische statistiek bedoeld.

In de redactie en in het bestuur van de 'Vereeniging' zaten toen mensen die alleen de ouderen onder ons zich nog zullen herinneren: Van der Burg, Enters, Van Ettinger, Lisman en Sittig; twee van hen zouden later erelid van de VVS worden. De laatstgenoemde, J. (Hans) Sittig, is op 9 juni van dit jaar overleden, 96 jaar oud. Hij leidde het Adviesbureau voor Toegepaste Statistiek in het Bouwcentrum in Rotterdam. Een markante man, die onder meer bekend geworden is door het schrijven van een boek, samen met de wiskundige Hans Freudenthal, over het optimaliseren van confectiematen: *De Juiste Maat* (Stafleu, Leiden, 1951). Het onderzoek werd verricht in opdracht van de Bijenkorf. Drs. A.R. van der Burg werkte destijds bij het adviesbureau Ir. B.W. Berenschot,

dat in de tijd dat ik daar op school ging voornamelijk in Hengelo gevestigd was. Twee broers Berenschot gingen daar ook op school.

Ook de inhoud van *Statistica* 1 nr. 1 lijkt meer op de inhoud van *STATOR* dan op die van de huidige *Statistica Neerlandica*; weinig wiskundige formules en heel praktische onderwerpen: Toleranties in de Machinebouw, Betrouwbaarheid en Stukscontrole, Opinieonderzoek en de Verkiezingen. Er is al sprake van kwaliteitscontrole en van 'Weersverwachting op langen termijn'. Computers waren er nog niet; er wordt veel gewerkt met tabellen en nomogrammen.

Het eerste nummer van *STATOR* verscheen in april 2000. Het opent met een 'Ten geleide' van het VVS-bestuur. Deze inleiding is aanzienlijk bescheidener dan het voorwoord in het eerste nummer van *Statistica*. In *STATOR* 1 nr. 1 staat voorzichtig, bijna bedremmeld, 'Het bestuur heeft een aantal geïnteresseerden bereid gevonden een poging te doen dit blad gestalte te geven. Het voorliggende nummer is uitdrukkelijk bedoeld als proef'. Hierbij wordt de hoop uitgesproken dat het blad 'bij bedrijven en andere geïnteresseerden, tot nuttige informatie en leesplezier zal leiden'. Het lijkt er op dat *STATOR* de bescheiden verwachtingen heeft waar gemaakt.

Maar, hoe zal het er in 2062 jaar uitzien? Zal dan ook met nostalgie en zelfs vertedering worden gekeken naar het eerste nummer *STATOR*? Zal de VVS nog bestaan, zal *STATOR* nog bestaan, zullen de eerste redacteurs in veler herinnering voortleven en zal minstens één van onze in 2000 zeer jonge redactieleden nog in leven zijn? Zal de boekdrukkunst nog bestaan, of zullen alle letters elektronisch zijn en is *STATOR* op een blader-scherm te lezen?

De tijd van de grote verwachtingen lijkt voorbij.

FRED STEUTEL is emeritus hoogleraar kansrekening aan de TU Eindhoven. Hij is redacteur van Stator

Spoorboekje NS wint Franz Edelman Award

Een team van onderzoekers naar de dienstregeling van de Nederlandse Spoorwegen (NS), Erasmus Universiteit in Rotterdam, University of Padova (Italië) en het Centrum Wiskunde & Informatica (CWI) heeft de prestigieuze 2008 Frans Edelman Award in de wacht gesleept voor het spoorboekje dat eind 2006 werd ingevoerd. Hun operationeel onderzoek en analyse op het reorganiseren van een intensiefgebruikt regionaal spoorwegnetwerk voor forensen werd bekroond als de beste toepassing van 'analytische methoden bij het optimaliseren van bedrijfsprocessen'. Deze innovatieprijs werd afgelopen april in Baltimore (VS) uitgereikt door INFORMS, Institute for Operations Research and the Management Sciences (INFORMS), de Amerikaanse vereniging voor Operations Research en Management Sciences.

Bij de ontwikkeling van de dienstregeling voor de treinen van NS over één van de drukste spoor-netwerken van Europa moet rekening gehouden worden met veel operationele factoren, zoals beschikbare sporen, knooppunten, perrons, treinen en personeel. Met gebruik van geavanceerde besliskundige modellen is het gelukt om een robuuste dienstregeling te ontwikkelen waarin meer treinen rijden en vertragingen gereduceerd zijn. Bovendien heeft het team ook het dienstrooster van het NS-personeel en de benutting van het materieel verbeterd, is de punctualiteit gestegen van 84,8% in 2006 tot 87% in 2007 en is er een grotere tevredenheid onder de reizigers. Tot slot is het aantal passagiers toegenomen met 2% tijdens de eerste helft van 2007.

De *STATOR*-redactie feliciteert het team van harte met deze geweldige prestatie. In *STATOR* 2007-3 is teamlid Dennis Huisman nader ingegaan op de technieken voor het maken van een nieuwe dienstregeling.

Bron: www.logistiek.nl/nieuws/

‘OR-SUCCESSVERHALEN BETER ONDER DE AANDACHT BRENGEN’

In gesprek met Alexander Rinnooy Kan

Alexander Rinnooy Kan, één van de OR-coryfeeën in Nederland, verruilde in 1991 de universitaire wereld voor woordvoerder van de werkgevers bij VNO-NCW en versterkte in 1996 de financiële wereld van ING. In augustus 2006 volgde hij Herman Wijffels op als voorzitter van de SER. Sinds 1 september 2007 is hij onbezoldigd hoogleraar Policy Analysis in the Private and Public Sectors aan de Faculteit Economie en Bedrijfskunde van de Universiteit van Amsterdam. Het gaat om een nieuwe leerstoel. Rinnooy Kan is daarmee weer terug aan de universiteit waar hij in 1976 promoveerde. Verder is hij prominent lid van het Innovatieplatform. Tijdens het Lunteren Seminar 2007, van het Nederlands Genootschap Besliskunde (NGB) en het Landelijk Netwerk Mathematische Besliskunde (LNMB), trad Alexander Rinnooy Kan op als voorzitter. Alle redenen dus voor STATOR om Alexander te interviewen, één van de trouwe en enthousiaste STATOR-lezers.

ANA ISABEL BARROS & GOOS KANT

Jouw aanwezigheid bij het Lunteren-seminar laat wel doorschijnen dat je OR nog steeds interessant vindt, of niet?

Natuurlijk. Na 25 jaar bezig te zijn met OR is het niet mogelijk om niks met OR te hebben. Ik heb nog steeds veel belangstelling voor dit veld en





ben benieuwd naar de nieuwe ontwikkelingen en richtingen op OR gebied. Wat zijn eigenlijk de nieuwe methoden, de *breakthroughs*.

Wat is de relatie tussen jouw werk als voorzitter van de Sociaal-Economische Raad en OR?

Consensus bereiken is een groot deel van mijn werk. Consensus bereiken gebruikt impliciet speltheorie!

Is OR goed bekend in Nederland?

Alle aandacht voor de nieuwe dienstregeling en andere problemen bij de Nederlandse Spoorwegen heeft OR wel in de *spotlights* gezet. Artikelen in kranten en tv-aandacht lieten zien dat beslistkundig advies een belangrijke rol heeft gespeeld bij het oplossen van het beruchte conflict 'rondje om de kerk'. Ook heeft beslistkunde een bijdrage gehad bij andere NS-onderwerpen zoals punctualiteit van treinen en het nieuwe dienstrooster in 2007. Maar er zijn veel meer voorbeelden over de kracht en impact van OR in de maatschappij die niet bekend zijn, en er is zonder meer nog veel te

doen voor OR. Ik denk ook dat meer te bereiken is op het gebied van Revenue Management, wat een grote maatschappelijke impact heeft. Dit gebied is net aan het groeien en heeft veel potentie. Ook binnen de financiële wereld kan OR en econometrie veel impact bereiken.

Hoe kunnen wij OR beter bekend maken?

OR-professionals moeten hun succesverhalen beter onder de aandacht brengen. Om zoiets te bereiken moeten wij een goed journalistiek netwerk opbouwen om er voor te zorgen dat de OR-succesverhalen een standaard plek krijgen in kranten en op tv. Als de meerwaarde en bijdrage van beslistkunde duidelijk en helder weergegeven wordt, zal een reporter altijd belangstelling hebben. Maar meestal zijn OR-professionals minder geïnteresseerd in pr-activiteiten. Hier is er ook een duidelijke rol voor het NGB en *STaTOR*: actiever over deze succesverhalen vertellen.

Je was voorzitter van commissies die adviseerden over het functioneren van NWO en de invoering van het bachelor-mastersysteem in het hoger onderwijs. Dus onderwijs is altijd een onderwerp geweest dat jou aanspreekt. Hoe kijk je tegen het feit aan dat het aantal studenten die voor technische studies en OR/toegepaste wiskunde kiest, terug loopt?

Het teruglopende aantal studenten is inderdaad een zorgwekkende ontwikkeling. Nederlanders zijn vaak "risicomijdend" en kiezen daarom de eenvoudige weg. Ze denken dat wiskunde of econometrie te moeilijk is. Vreemd, want de arbeidsmarkt vraagt juist om mensen met deze achter-



Alexander Rinnooy Kan. Foto: Sjaak Ramakers

grond. NGB en universiteiten zouden hier een bijdrage kunnen leveren om de algemene *awareness* over wat OR is bij de kandidaat-studenten te profileren.

Als voorzitter van Strategische Adviescommissie over innovatie, ben je van heel dichtbij betrokken met innovatie. Heeft Nederland de potentie om innovatieland te zijn op OR gebied?

Sowieso associeert men innovatie aan het ontwikkelen van nieuwe producten. Maar innovatie kan ook procesinnovatie zijn. Nederland heeft wel op het vlak van procesinnovatie veel mogelijkheden, ook gegeven zijn grote traditie in besliskunde. Er moet hard gewerkt wordt om dit te bereiken.

Eigenlijk zouden de technieken die procesinnovatie ondersteunen deel uit moeten maken van het OR/econometrie onderwijs. Daarnaast heeft Nederland een grote traditie in de logistiek en door zijn "ideale" locatie zou Nederland hier veel moeten kunnen bereiken.

Alexander, heel hartelijk dank voor dit interview.

ALEXANDER RINNOOY KAN studeerde wiskunde in Leiden en econometrie in Amsterdam, werkte een jaar voor de *Spectrum Encyclopedie* (redacteur wiskunde) en drie jaar aan de universiteit Delft. In 1977 ging hij naar de Erasmus Universiteit Rotterdam waar hij in 1980 hoogleraar operationeel onderzoek werd, in 1983 directeur van het Econometrisch Instituut en in 1986 rector magnificus. Tussendoor bekleedde hij gasthooglerschappen in onder andere Californië (Berkeley) en Boston (MIT).

Na zijn voorzitterschap van VNO (vanaf 1991) en VNO-NCW (vanaf 1995) trad Rinnooy Kan in 1996 toe tot de raad van bestuur van bank/verzekeraar ING. In 2006 volgde hij Herman Wijffels op als voorzitter van de Sociaal-Economische Raad. Hij werd unaniem voorgedragen door de werkgevers- én de werknemersvertegenwoordigers in de SER. Rinnooy Kan is per 1 september 2007 benoemd tot onbezoldigd hoogleraar Policy Analysis in the Private and Public Sectors aan de Faculteit Economie en Bedrijfskunde (FEB) van de Universiteit van Amsterdam.

Hij heeft zich altijd voor de publieke zaak ingezet. Hij was voorzitter van Dreamstart, dat startende innovatieve ondernemers op weg moest helpen, en van de visitatiecommissie voor de publieke omroep. Hij adviseerde minister Borst over het geneesmiddelenbeleid en minister Netelenbos over de luchtvaart. Ook was hij voorzitter van commissies die adviseerden over het functioneren van NWO en de invoering van het bachelor-mastersysteem in het hoger onderwijs. Rinnooy Kan is voorzitter van de Raad van Commissarissen van Het Concertgebouw N.V. en voorzitter van de Stichting Praemium Erasmianum. Hij is sinds 2006 ambassadeur van de Belangenvereniging voor Vrijwilligers, waarbij hij streeft naar meer bekendheid en aandacht voor het vrijwilligerswerk. Ook heeft hij minder bekende functies, zo zit hij bijvoorbeeld in het Comité van Aanbeveling van de Nederlandse Wereldwijde Studenten.

DE PARADOX

ONNO BOXMA

Op de voorhoofden van twee personen worden random twee opeenvolgende natuurlijke getallen geschreven. De persoon met het laagste getal is de verliezer; hij moet de ander dit (laagste) bedrag betalen. Beiden redeneren: ik zie dat mijn opponent het getal k heeft, dus ik heb $k-1$ of $k+1$, met gelijke kansen. Als ik verlies, dan betaal ik $k-1$, maar als ik win, dan ontvang ik k . Mijn verwachte winst is dus een half. Beide partijen hebben dus een positieve winstverwachting, wat onmogelijk is. De bedenker van deze *paradox*: de fysicus Erwin Schrödinger [2, p. 177-178], bij velen ook bekend wegens de paradox van Schrödingers kat, die volgens de quantummechanica zowel levend als dood zou kunnen zijn.

In het Latijn betekent *paradoxa* 'wonderspreuken' (in strijd met de gewone mening), en het Griekse woord $\delta\omicron\xi\alpha$ betekent 'verwachting, vermoeden', zodat $\pi\alpha\rho\alpha\ \delta\omicron\xi\alpha\nu$ staat voor 'in strijd met de verwachting'. Ik heb paradoxen in de wiskunde altijd zowel fascinerend als nuttig gevonden. Het besef dat je iets nog niet volledig begrijpt is essentieel voor vooruitgang in de wetenschap, en paradoxen vormen een confronterende manier om dat besef te versterken. Het is niet voor niets dat doorbraken als de begrippen nul, oneindig, complex getal en irrationaal getal gepaard zijn gegaan met discussies over allerlei paradoxen. Bolzano heeft zelfs een heel boek (*Paradoxen des Unendlichen*) gewijd aan het begrip oneindig, en toen moest Cantor nog komen met zijn verzamelingentheorie.

De kansrekening is een tak van de wiskunde die wel heel erg rijk is aan paradoxen. We zijn dan ook gewend te denken in termen van het gemiddelde, en *afwijkingen* van het gemiddelde stellen onze intuïtie op de proef. Ook hebben we moeite met denken in termen van conditionele kansen, wat iedereen kan bevestigen die zich wel eens het hoofd heeft gebroken over de paradox van de quizmaster met de drie deuren.

Een paar mooie paradoxen uit de kansrekening

Een van de eerste kansrekeningparadoxen waar ik mee in aanraking kwam is de St. Petersburg paradox. Een speler gooit een zuivere munt net zo lang tot kop boven komt. Als dit bij de r -de worp voor het eerst gebeurt, ontvangt de speler 2^r euro. Wat is een 'eerlijke' inleg om dit spel te mogen spelen? Antwoord: oneindig.

Als student hoorde ik van Bertrand's paradox: kies random een koorde van een cirkel (een rechte lijn die twee punten van de omtrek met elkaar verbindt), en bereken de kans dat deze koorde langer is dan de zijde van de ingeschreven gelijkzijdige driehoek. De paradox is dat deze kans niet eenduidig bepaald is. Verschillende methoden om de koorde random te kiezen leiden tot verschillende antwoorden. Bijvoorbeeld: kies een punt uniform verdeeld binnen de cirkel, en neem dit punt als middelpunt van een koorde, met als ant-

woord 1/4; of kies voor een der eindpunten van de koorde een punt van de ingeschreven gelijkzijdige driehoek, en kies het andere eindpunt uniform verdeeld op de cirkel – met als antwoord 1/3.

De volgende paradox is verbonden met het min-max principe uit de speltheorie. Twee kinderen, A en B, spelen het volgende spel. Ze houden beiden tegelijkertijd een of twee vingers op. Als het totale aantal vingers van beiden samen even is, dan betaalt A aan B dit totale bedrag, en anders betaalt B aan A het bijbehorende totale (oneven) bedrag. De naïeve lezer (en wie is dat niet) denkt dat A en B quitte spelen. Dit spel is echter gunstig voor A. De beste strategie voor elk der spelers is om met kans 7/12 een vinger op te steken, en met kans 5/12 twee vingers. A wint dan gemiddeld 1/12 euro.

Zo kan ik nog wel een tijdje doorgaan, onder meer met dank aan Székely [2]. Denk maar aan de verjaardagparadox, voor zover je dat een paradox mag noemen: in een groep van 55 mensen is de kans al 0.99 dat twee mensen op dezelfde dag jarig zijn. Of aan de paradox van het eerste cijfer: het eerste cijfer van de getallen in allerlei tabellen is niet uniform verdeeld op $\{1,2,\dots,9\}$; volgens de wet van Benford is de frequentie van k niet $1/9$ maar $^{10}\log(k+1)/(k)$, voor $k=1,\dots,9$. Ook fraai zijn de paradoxen over stochastische wandelingen in 1, 2 en meer dimensies; zie bijvoorbeeld het schitterende hoofdstuk 3 van Feller [1]. In mijn eigen vakgebied, de wachtrijtheorie, is de wachtrij- of busparadox heel bekend, en zij duikt ook in allerlei vermommingen telkens weer op: als bussen volgens de dienstregeling elke 20 minuten langs een bepaalde halte moeten komen, en u naar de bushalte gaat zonder op de dienstregeling acht te slaan, dan wacht u gemiddeld langer dan 10 minuten, tenzij de dienstregeling perfect wordt aangehouden. Zodra de aankomstintervallen van bussen enige schommeling om het gemiddelde van 20 minuten vertonen, bent u slechter af.

Een avond met paradoxen en problemen

EURANDOM, het onderzoeksinstituut in de statistiek, kansrekening en stochastische besliskunde, viert dit najaar haar 10-jarig jubileum. Het doet dat niet alleen via een congres in haar eigen gebouw, op 27 en 28 augustus, maar ook met een bijeenkomst voor een breed publiek in het auditorium van de TU Eindhoven, op de avond van 23 oktober – en paradoxen zullen die avond een prominente rol spelen. De avond, getiteld *An unlikely evening with EURANDOM*, begint met een lezing over *Credit Risk* van de Leuvense hoogleraar (en voormalig EURANDOM postdoc) Wim Schoutens. Daarna zullen Remco van der Hofstad (TU Eindhoven) zijn favoriete paradoxen bespreken. Het laatste deel van de avond vult u zelf in: als u een intrigerend/leuk/uitdagend probleem heeft op het gebied van de statistiek of kansrekening (mag ook een paradox zijn!), stuur het dan naar Marlies Brangers (brangers@eurandom.tue.nl) o.v.v. *An unlikely evening with EURANDOM*. De meest interessante vragen zullen 23 oktober 's avonds worden behandeld door aan EURANDOM verbonden stochastici.

Informatie over alle jubileumactiviteiten kan worden gevonden op <http://www.eurandom.tue.nl/10years.htm>

Hopelijk zie ik u op 23 oktober in Eindhoven!

LITERATUUR

1. W. Feller. An Introduction to Probability Theory and its Applications, Vol.I. Wiley, New York, 1957.[2]
2. G.J. Székely. Paradoxes in Probability Theory and Mathematical Statistics. D. Reidel Publishing Company, Dordrecht, 1986.

ONNO BOXMA is hoogleraar Stochastische Besliskunde bij de Faculteit Wiskunde en Informatica van de Technische Universiteit Eindhoven en wetenschappelijk directeur van EURANDOM. E-mail: boxma@win.tue.nl.

IN MEMORIAM THEO RUNNENBURG (1932-2008)



Johannes Theodorus (Theo) Runnenburg werd op 19 februari 1932 geboren in Amsterdam, waar zijn ouders op de Overtoom een wasserij hadden. Hij haalde zijn HBS-diploma aan de 'Tweede Vijfjarige' aan het Roelofhartplein. Daarna studeerde hij wis- en natuurkunde aan de Universiteit van Amsterdam. Zijn leermeesters in de wiskunde waren, naast Heijting en De Groot, vooral Van Dantzig en De Bruijn. Bij deze laatste haalde hij, als enige student ooit, een tien voor het vak Maattheorie (informatie van De Bruijn zelf). Na zijn afstuderen in 1954 ging hij onder Van Dantzig werken aan het Mathematisch Centrum, nu Centrum voor Wiskunde en Informatica. De eerste jaren werkte hij nauw samen met Harry Kesten. Zij schreven samen twee baanbrekende artikelen onder de titel *Priority in waiting line problems* I en II. Harry emigreerde kort daarna naar de VS.

Ik leerde Theo in september 1956 op het Mathematisch Centrum kennen; ik had net mijn kandidaatsexamen gedaan. Mijn eerste artikeltje schreef ik met Theo's hulp, over prioriteiten in een wachtrijmodel. Wij vertaalden samen één van de

eerste boeken over wachttijdtheorie: *Problèmes stochastiques posés par le phénomène de formation d'une queue d'attente à un guichet*, van F. Pollaczek, in het Nederlands.

Hij was tussen allerlei andere werkzaamheden door – met name veel werk voor het rapport van de Deltacommissie in verband met de watersnoodramp in 1953 – aan een proefschrift bezig, met Van Dantzig als promotor. Toen het werk: *On the Use of Markov Processes in One-server Waiting-time Problems and Renewal Theory* praktisch klaar was, overleed Van Dantzig (in 1959). De promotie is toen in 1960 afgehandeld door N.G. de Bruijn. Kort daarna werd Runnenburg tot lector en weinig later tot hoogleraar kansrekening benoemd. Hij gaf uiterst consciëntieus onderwijs, hij werkte zich letterlijk in het zweet en moest zich na een college verkleden.

Hij was een heel scherpzinnige wiskundige. Hij heeft aan heel diverse onderwerpen gewerkt, maar toch betrekkelijk weinig gepubliceerd. Zijn belangrijkste verdienste voor de wetenschap ligt in het begeleiden van een aantal getalenteerde promovendi. Elf mensen, onder wie ikzelf, zijn bij hem gepromoveerd; vijf daarvan zijn hoogleraar geworden.

Zijn laatste jaren werden gekenmerkt door eenzaamheid. Zijn vrouw was opgenomen in een verzorgingshuis en hij had weinig contact met zijn kinderen. Hij had veel belangstelling voor de mensen om hem heen, maar door zijn ongepolijste optreden had hij weinig blijvende vrienden gemaakt, ook onder zijn promovendi. Hij bleef wel tot het laatst in veel dingen geïnteresseerd, binnen en buiten de wiskundewereld. Op 16 april is hij overleden, drie dagen na zijn vrouw. Ik heb hem de laatste weken van zijn leven verschillende malen bezocht. Ik zal met veel waardering aan hem denken.

Fred Steutel



European Network for Business and Industrial Statistics (ENBIS)

SUCCESVOL FORUM VOOR DYNAMISCHE UITWISSELING VAN IDEEËN

In 1999 heeft een groep van zo'n 20 statistici het initiatief genomen om te komen tot een Europees netwerk van gebruikers van statistiek in het bedrijfsleven. Op 11 december 2000 werd dit netwerk in Amsterdam opgericht. ENBIS is een autonome vereniging die tot doel heeft statistische methoden te ontwikkelen en te verbeteren en deze overal in Europa toe te passen; dit alles in de breedste zin van deze woorden. ENBIS tracht dit doel te bereiken door vooral:

- het bevorderen en faciliteren van de toepassing en het begrip van statistische methoden ten behoeve van Europese bedrijven en industrieën;
- te voorzien in een forum voor de dynamische uitwisseling van ideeën en het faciliteren van netwerkactiviteiten onder beoefenaars van de statistiek (een beoefenaar van de statistiek is ieder persoon die statistische methoden gebruikt, ongeacht of deze persoon een officiële scholing heeft genoten);
- het bevorderen van interacties tussen en de professionele ontwikkeling van beoefenaars van de statistiek, zowel regionaal als internationaal;
- en met behulp van alle andere wettige middelen die het doel van ENBIS kunnen helpen bereiken.

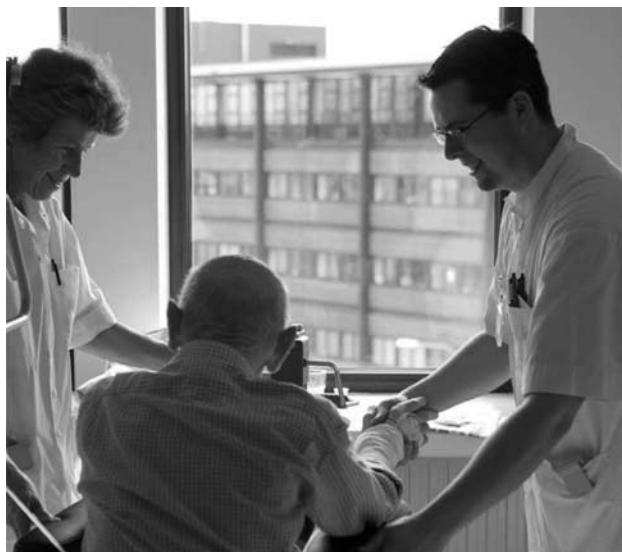
Het netwerk heeft inmiddels ruim 1300 leden uit zo'n 30 verschillende landen. Het lidmaatschap is gratis. Het jaarlijkse congres is het hoogtepunt. De afgelopen jaren vond dit congres plaats in respectievelijk Oslo, Rimini, Barcelona, Newcastle, Kopenhagen, Wroclaw, Duisburg en Athene.

Nederlanders hebben een belangrijke rol

gespeeld bij de totstandkoming van ENBIS. Jarenlang waren zij verantwoordelijk voor de website www.enbis.org, het secretariaat, het penningmeesterschap en het directoraat. Dit laatste trouwens nog steeds.

Vanaf 2002 is de Bedrijfssectie van de VVS collectief toegetreden tot ENBIS. De BDS heeft de status van lokaal netwerk en staat bekend als n-ENBIS. De deelname van BDS-leden aan de activiteiten van ENBIS is tot nu toe beperkt gebleven. Ook bij sessies verzorgd tijdens de Statistische Dag is de opkomst gering. De leden gaven, in een enquête, wel aan geïnteresseerd te zijn in bedrijfsbezoeken. De eerste gelegenheid werd hiertoe in mei 2008 geboden in het Universitair Medisch Centrum Groningen (UMCG) tijdens het Symposium over Lean Six Sigma aanpak in ziekenhuizen. Een tiental leden maakte hiervan gebruik.

Ronald Does, voorzitter BDS/n-ENBIS



AGENDA

21-25 september 2008

De 8ste jaarlijkse ontmoeting van **European Network of Business and Industrial Statistics** vindt dit jaar van 22 t/m 24 september 2008 plaats in Athene. ENBIS is te gast bij de Athens University of Economics and Business. Meer informatie: www.enbis.org/events/current/66_ENBIS8_Athens.

12 november 2008

Asset | Econometrics (voorheen Tilburgse Econometristen Vereniging) organiseert het symposium **Revenue Management**, met prof.dr. G. Koole als dagvoorzitter. Meer informatie en inschrijven: www.Revenue-Management.nl

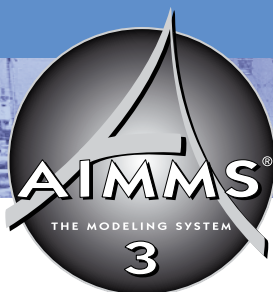
12-14 november 2008

Internationale workshop **Flexible Modelling: Smoothing and Robustness** (FMSR 2008). Algemeen thema: semi-en niet-parametrische analyses en robuuste statistische methodes. Er zijn *invited talks*, *contributed talks* en *poster sessions*. De workshop wordt gevolgd door een intensieve cursus voor doctoraatsstudenten. Zie ook: <http://wis.kuleuven.be/stat/fmsr2008.php>.

17-19 november 2008

De jaarlijkse **Bijeenkomst van Stochastici** zal gehouden worden van 17 t/m 19 november 2008 in het Congrescentrum 'De Werelt' te Lunteren. Voor registratie en overige informatie zie www.cwi.nl/events/2008/lunteren/lunteren2008.html.

Modelleren met AIMMS



AIMMS is een compleet modelleersysteem dat mensen helpt Operations Research (OR) succesvol in te zetten. Velerlei modellen (LP, MIP, NLP, MINLP, etc) zijn eenvoudig en snel in AIMMS te bouwen en op te lossen met standaard solvers of geavanceerde technieken zoals kolomgeneratie, stochastische programmering, Benders decompositie en Outer Approximation.

OR onderwijs met AIMMS

De grafische modelleeromgeving en de geïntegreerde visualisatie-mogelijkheden maken AIMMS tot een ideaal softwarepakket om te gebruiken in het onderwijs. Een compleet academisch licentiepakket kost 450 Euro. Ondersteunend materiaal is gratis beschikbaar op onze website, zoals:

- Leerboek "Optimization Modeling" met OR toepassingen van oplopende moeilijkheidsgraad
- Uitgewerkte applicatie-voorbeelden in AIMMS
- Introductie-cursussen voor zelfstudie (Tutorials)

Commerciële toepassingen

Bedrijven in uiteenlopende sectoren gebruiken AIMMS-applicaties om hun bedrijfsvoering te optimaliseren, bijvoorbeeld in productieplanning, supply chain management, netwerkontwerp, procesoptimalisatie, risicobeheersing en portfoliobeheer. Referenties zijn te vinden op onze website.

AIMMS Users Meeting

Noteer alvast de datum van de AIMMS Users Meeting in Haarlem: Vrijdag 14 November 2008.

AIMMS is een handelsmerk van
Paragon Decision Technology B.V.
Schipholweg 1, 2034 LS Haarlem
Tel. 023 5511512, info@aimms.com
www.aimms.com





Om een Triple A-rating te kunnen waarmaken, heb je Triple A-talent nodig.

Senior Model Validator

Het is tijd voor de Rabobank.

De Rabobank is al jarenlang een van de weinige commerciële banken ter wereld met een Triple A-rating. Voor zo'n rating moet je keihard werken en altijd meer doen dan er van je wordt verwacht. Dat is ook precies de instelling die de Rabobank zoekt bij nieuwe medewerkers.

De afdeling

Group Risk Management (GRM) beheert de risicoportefeuille binnen Rabobank Groep. Het is een ambitieuze afdeling die zich snel ontwikkelt. De afdeling Risk Model Validation and Methodologies (RMVM) heeft een ondersteunende rol bij het realiseren van een organisatiebrede, objectieve risicometing. RMVM is verantwoordelijk voor controle op de naleving van Basel II-regulering en de door Rabobank opgestelde interne eisen.

De functie

Je bent direct betrokken bij het valideren, controleren en testen van alle modellen met betrekking tot kredietrisico, operationeel risico, renterisico en economisch kapitaal. Ook zie je toe op de integriteit en de validiteit van de methodologieën waarmee de modellen worden ontwikkeld en onderhouden. Je werkt nauw samen met die onderdelen van Rabobank Groep, die de risicomodellen ontwikkelen en gebruiken. Teamwerk en het delen van kennis maken dus integraal deel uit van je functie.

Je expertise

Je hebt een afgeronde academische opleiding Econometrie, Wiskunde of BWI en goede kennis van econometrische en statistische technieken voor modelontwikkeling. Ook beschik je over drie tot vijf jaar relevante ervaring. Je bent aantoonbaar geïnteresseerd in kwantitatief- en statistisch-financieel onderzoek en je bezit een groot probleemoplossend vermogen.

Ons aanbod

Zoals je van een bank als de Rabobank kunt verwachten, zijn de primaire en secundaire arbeidsvoorwaarden uitstekend.

Interesse?

Voor meer informatie kun je contact opnemen met Sebastiaan van Westerop, manager Risk Model Validation and Methodologies, (030) 216 43 89. Met vragen over de procedure bel je met Miranda Niels, recruiter, (030) 215 90 93. Direct solliciteren kan via www.rabobankgroep.nl/werken. Vacaturenummer: GRM08MN013

www.rabobank.nl/werken



Rabobank