

STAtOR

periodiek van de VVS jaargang 3 nummer 2 juni 2002

Modellen in de pers

De ruimte in getallen

In memoriam Richard Freling

Statistiek en genen

Vereeniging voor de Statistiek (VVS),
een gezelschap van juristen

Interview met Aharon Ben-Tal,
een robuuste optimist

Hoe ontwikkel je toetsen en examens?

Taakplanning als gat in de markt

STATOR

Jaargang 3, nummer 2, juni 2002

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

Dick den Hertog (hoofdredacteur), Wies Akkermans, Martijn Berger, Han Oud, Marc Schul, Gerrit Stemerding (eindredacteur), Fred Steutel.

Kopij en reacties richten aan

Prof. dr. ir. D. den Hertog (hoofdredacteur)
Faculteit der Economische Wetenschappen van de Katholieke Universiteit Brabant, Postbus 90153, 5000 LE Tilburg, telefoon 013 - 466 2122, <D.denHertog@kub.nl>.

Bestuur van de VVS

Prof. dr. G.T. Timmer (voorzitter) <gtimmer@ortec.nl>, prof. dr. S. J. Koopman <s.j.koopman@econ.vu.nl>, dr. A. Mooijaart (penningmeester) <mooijaart@rulfsw.leidenuniv.nl>, prof. dr. H.G. Dehling (voorzitter commissie opleidingen en examens) <dehling@math.rug.nl>, dr. J.H.L. Oud (voorzitter publicatiecommissie) <j.oud@ped.kun.nl>. Zie voor telefoonnummers en adressen de website.

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>

Advertenties

Contactpersoon: Rita Oomen, telefoon 0167 - 563401, fax 0167 - 561200, <japm.oomen@worldonline.nl>. Uiterlijk vier weken voor verschijnen te zenden aan Pharos / M. van Hootegem, Moeftonstraat 5, 6531 JS Nijmegen, telefoon 024 - 3559214, fax 024 - 559614 <hootegem@xs4all.nl>. STATOR verschijnt in maart, juni, september en december.

Ontwerp en opmaak

Pharos / M. van Hootegem, Nijmegen

Druk

Drukkerij Trioprint Nijmegen bv

Uitgever

© Vereniging voor Statistiek en Operationele Research

ISSN 1567-3383

Inhoud

- 3** Modellen in de pers.
Dick den Hertog
- 4** De ruimte in getallen.
Alfred Stein, Kirsten de Beurs, Jasja Dekker en Christien Ettema
- 9** In memoriam Richard Freling.
Albert P.M. Wagelmans
- 11** Statistiek en genen.
Sara van de Geer
- 13** Vereniging voor de Statistiek (VVS), een gezelschap van juristen.
Ida H. Stamhuis
- 18** Interview met Aharon Ben-Tal, een robuuste optimist.
Frits van Beckum en Hans Melissen
- 23** Hoe ontwikkel je toetsen en examens? Over de rol van kwantitatieve methoden.
Piet Sanders
- 28** Taakplanning als gat in de markt. Kansen voor nieuwe technologieën in it-systemen.
Martin Healey
- 30** Agenda.

MODELLEN IN DE PERS



In de pers verschijnen steeds vaker artikelen over kwantitatieve modellen en technieken en de resultaten daarvan. Het is opvallend dat je in de pers vaak óf een overwaardering óf juist een onderwaardering van wiskundige modellen aantreft. Een prachtig voorbeeld van overwaardering is de berichtgeving naar aanleiding van Khachiyans publicatie in 1979 in een Russisch tijdschrift. Vele kranten in binnen- en buitenland besteedden destijds aandacht aan zijn ellipsoïde methode voor lineaire programmering. De conclusies en toekomstverwachtingen waren vaak sterk overdreven of zaten er zelfs helemaal naast: 'Inlichtingendiensten zien hun geheime coderingssystemen verloren gaan', of, 'Theoretisch is het mogelijk een 100 procent juiste weersvoorspelling te geven'.

Een recent voorbeeld, waarin de pers van een zowel sterke overwaardering als van een onderwaardering blijk gaf, betreft het analyseren van de partijprogramma's door het CPB in de afgelopen verkiezingstijd. Aan de ene kant werd in diverse kranten gesuggereerd alsof hetgeen het CPB berekende de echte toekomst zal zijn, ondanks de nuances die het CPB zelf aanbracht. Aan de andere kant waren er krantenartikelen die ongenuanceerd het werk van het CPB, waarin modellen een belangrijke rol spelen, totaal waardeeloos vonden.

Hoe komt dit? Brengen wij de boodschap verkeerd over of geven de journalisten de boodschap verkeerd door? Over het eerste heb ik al eerder iets geschreven in STAtOR. Ik wil nu vooral benadrukken dat er ook bij de media veel te verbeteren valt.

Veel journalisten hebben weinig kaas gegeten van wiskunde en wiskundige principes. Het is bedroevend hoe 'ongecijferd' vele journalisten zijn. Een prachtig boek in dit verband is *Een wiskundige leest de krant* van John Allen Paulos.*

Paulos ergert zich aan het feit dat vele getallen in de krant staan zonder bronvermelding. Komen de cijfers uit een politierapport? Een vage steekproef? Een breed wetenschappelijk onderzoek? Of zijn het gekleurde gegevens van een actiegroep? Volgens Paulos zijn de meeste journalisten onvoldoende toegerust om getallen in hun context te zien. Getalsfouten zijn vele malen ernstiger dan een spelfout, stelt Paulos

Overigens zijn er ook positieve uitzonderingen. Ik heb diverse uitstekende artikelen over kwantitatieve methoden en technieken gelezen, geschreven door journalisten die wisten waar ze het over hadden!

Om de journalisten beter te informeren over ons vakgebied, heeft de redactie besloten om STAtOR ook te sturen naar diverse kranten en tijdschriften. Wellicht zult u in de toekomst meer en (nog) betere artikelen over ons vakgebied in de media aantreffen!

Veel leesplezier!

Dick den Hertog
hoofdredacteur

*Paulos, J.A. *Een wiskundige leest de krant. Mathematische overdenkingen bij krantenberichten*. Amsterdam: Bert Bakker, 1995.

De ruimte in getallen

Stel je voor dat je iets wilt weten over een gebied, een land, een regio of een terrein.

Vroeger zou je een atlas pakken, een encyclopedie, of gaan praten met de mensen ter plaatse om zo een beeld te krijgen van wat je kunt verwachten. Tegenwoordig hebben we nieuwe mogelijkheden. Soms wat verfijnder, soms wat specifiek. In ieder geval hebben we nu meer mogelijkheden om kwantitatieve uitspraken te doen. Een belangrijk hulpmiddel hierbij is de ruimtelijke statistiek.

ALFRED STEIN, KIRSTEN DE BEURS, JASJA DEKKER EN CHRISTIEN ETTEMA

Ruimtelijke statistiek richt zich op een goed en verantwoord gebruik van ruimtelijke gegevens. Ruimtelijke statistiek helpt om ruimtelijke variatie te kwantificeren, en laat zien of een gebied gevarieerd of eentonig is. Het kan dan gaan om continue gegevens, zoals een enkelvoudig kenmerk dat de kwaliteit van bodem, water of lucht beschrijft. Maar het kan ook de lokale variatie van een enkele diersoort betreffen of een biodiversiteitsindex; ruimtelijke statistiek kan ook verspreiding van ziektes beschrijven. Vroeg of laat heb je daarvoor informatie nodig: ruimtelijke informatie. In de ruimtelijke statistiek kijken we naar gegevens die in de ruimte beschikbaar zijn of verzameld moeten worden. We proberen daarmee een kwantitatieve uitspraak te krijgen over een ruimtelijk kenmerk. Essentieel is daarbij dat aan ieder getal een ruimtelijke coördinaat gekoppeld is.

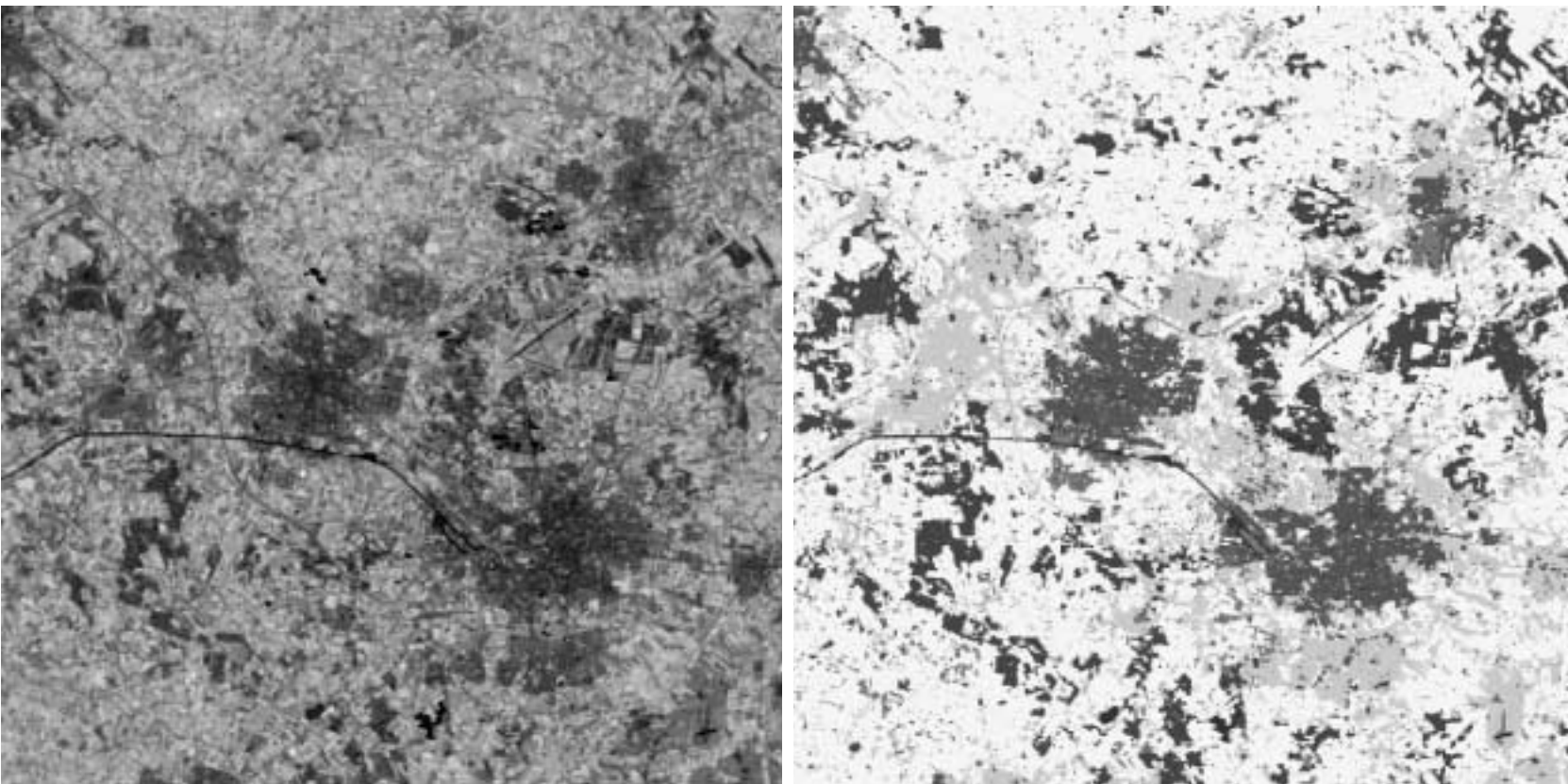
Remote sensing

Stel dat we de beschikking hebben over een satellietbeeld van een gebied. Een satellietbeeld bestaat uit een groot aantal (ruimtelijke) pixels. Iedere pixel beslaat een ruimte van bijvoorbeeld 30 bij 30 m. Op iedere pixel wordt via een nauwe opening van het spectrum gekeken naar de reflectie van het zonlicht, van radarsignalen of van warmte. Het aantal openingen (spectrale banden) is als regel beperkt, vaak is naar 6 banden gekeken. Concreet betekent dit dat er een multivariaat gegevensbestand is, met 6 variabelen en evenveel waarnemingen als er pixels zijn. Op grond van dit bestand is het belangrijk om een segmentatie uit te voeren. De segmenten sporen dan, als het goed is, met objecten (percelen, wijken, meren) in de werkelijkheid. Een gesegmenteerd beeld bestaat kortom uit samenhangende homogene eenheden. Homogeniteit is gedefinieerd in de zin van

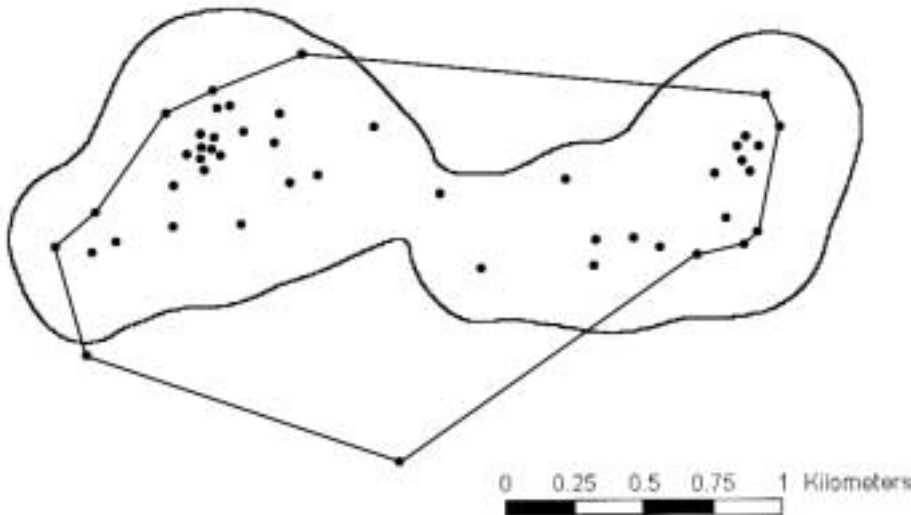
een lage variantie binnen segmenten. Er zijn verschillende manieren om een gesegmenteerd beeld te verkrijgen uit een origineel beeld. Een bekend voorbeeld is de *split and merge* techniek. Een meestal rechthoekig satellietbeeld wordt opgedeeld in steeds kleinere vierkante deelstukjes, totdat zo'n deelstukje voldoende homogeen is. Dat kan leiden tot grote deelstukjes in homogene gedeelten van een gebied en kleine stukjes in heterogene gedeelten van het gebied. Vervolgens worden afzonderlijke deelstukjes weer samengevoegd, als zo'n samenvoeging op basis van hetzelfde homogeniteitscriterium is toegelaten. In het voorbeeld zien we hoe dat uitpakt voor een beeld van Twente (Figuur 1). Het originele beeld wordt eerst gegeven, vervolgens een gesegmenteerd beeld. Er zijn duidelijk homogene segmenten te ontdekken, corresponderend met objecten

in werkelijkheid, die het zonlicht homogeen reflecteren.

Statistische beeldsegmentatie is nog steeds in ontwikkeling. Hier speelt, meer dan waar ook, het element van data-kwaliteit een rol. Er zijn nieuwe ontwikkelingen, waarbij het aantal banden dramatisch toeneemt, de zogenaamde hyperspectraalbanden. Het zichtbare (en nabij infrarode) licht wordt onderverdeeld in 100 of meer banden. Zo'n groot aantal banden vertoont onderling een grote samenhang. Deze informatie zou dan gebruikt kunnen worden bij de beeldsegmentatie. Andere ontwikkelingen laten het aantal pixels toenemen, de resolutie is al in de buurt van 1 m^2 . Tenslotte vragen radarbeelden weer om een geheel eigen benadering, met aandacht voor faseverschillen, zodat er eigenlijk sprake is van een beeld vol complexe getallen.



Figuur 1. Een satellietbeeld van Twente (links) en het beeld nadat een segmentatie en classificatie is toegepast (rechts)



Ecologische gegevens

Ruimtelijke statistiek is stilaan een breed en belangrijk deelgebied binnen de statistiek geworden, met veel dwarsverbanden met andere disciplines. Bekende voorbeelden zijn het milieu, de landbouw en de ecologie, maar ook in de econometrie zijn ruimtelijke studies gangbaar. Veelal gaat het om het kwantificeren van ruimtelijke informatie, vaak speelt daarbij op de achtergrond een ruimtelijke beheers- of beleidsvraag. Dat kan iets eenvoudigs zijn: hoeveel waarnemingen moeten we verzamelen, of iets ingewikkelders, zoals: hoe moeten we een gebied waar veel wild voorkomt optimaal beheren?

In een recente studie hebben we gekeken naar de verspreiding van vossen (*Vulpes vulpes*) in de duinen van Wassenaar. Dergelijke studies zijn lastig omdat betrouwbare gegevens moeilijk te krijgen zijn. Sommige vossen zijn gemakkelijk te volgen

Figuur 2. Links een vergelijking van een minimaal convex polygoon (dunne lijn) en 95 % kernel schatter voor de homerange van een juveniele mannelijke vos, genaamd Gé, in de periode van 1 oktober tot 31 december 1997; rechts een rode vos

omdat ze een zendertje hebben en zodoende gegevens opleveren met een hoge tijdsdichtheid. Andere vossen kom je toevallig tegen, weer andere worden ergens door een passant gesignaleerd en leveren gegevens met een veel lagere tijdsdichtheid. De centrale vraag in deze studie was of er misschien sprake is van 2 soorten vossen, namelijk huiselijke vossen en zwervers. Huiselijke vossen hebben een klein leefgebied, zwervers een veel groter. Met andere woorden: kunnen we een twee-toppige verdeling van het leefgebied vaststellen?

Analyse van deze gegevens ging als volgt. Als eerste stap heeft een harmonisatie van gegevens plaatsgevonden door een geschikte weging toe te passen. Tevens zijn de afzonderlijke vossen geïdentificeerd, zodat een vos herkend kon worden. Als tweede stap is gekeken wat de verspreiding was per vos (zie Figuur 2 voor de verspreiding van Gé). Dat is zowel gedaan via een minimaal convex polygoon als met een kernel dichtheid. Deze procedure leidde tenslotte inderdaad tot het veronderstellen van 2 populaties.

Dit voorbeeld laat zien dat er in de eerste fase vooral gewerkt wordt aan de gegevenskwaliteit. Deze wordt hier gekenmerkt door ruimtelijke resolutie (hoe nauwkeurig is de plaatsbepaling), door de temporele resolutie (konden we een vos volgen via een zendertje en bijvoorbeeld iedere minuut een signaal opvangen), door het vermogen om afzonderlijke dieren te onderscheiden en door de vorm van rapportage (spotten of incidentele waarneming). Na harmonisatie van de gegevens kon een standaard manier van analyseren worden toegepast, namelijk kernel dichtheids-schatting. Dergelijke methoden zijn gewoon beschikbaar via een pakket als S-Plus.

Ruimtelijk bemonsteren

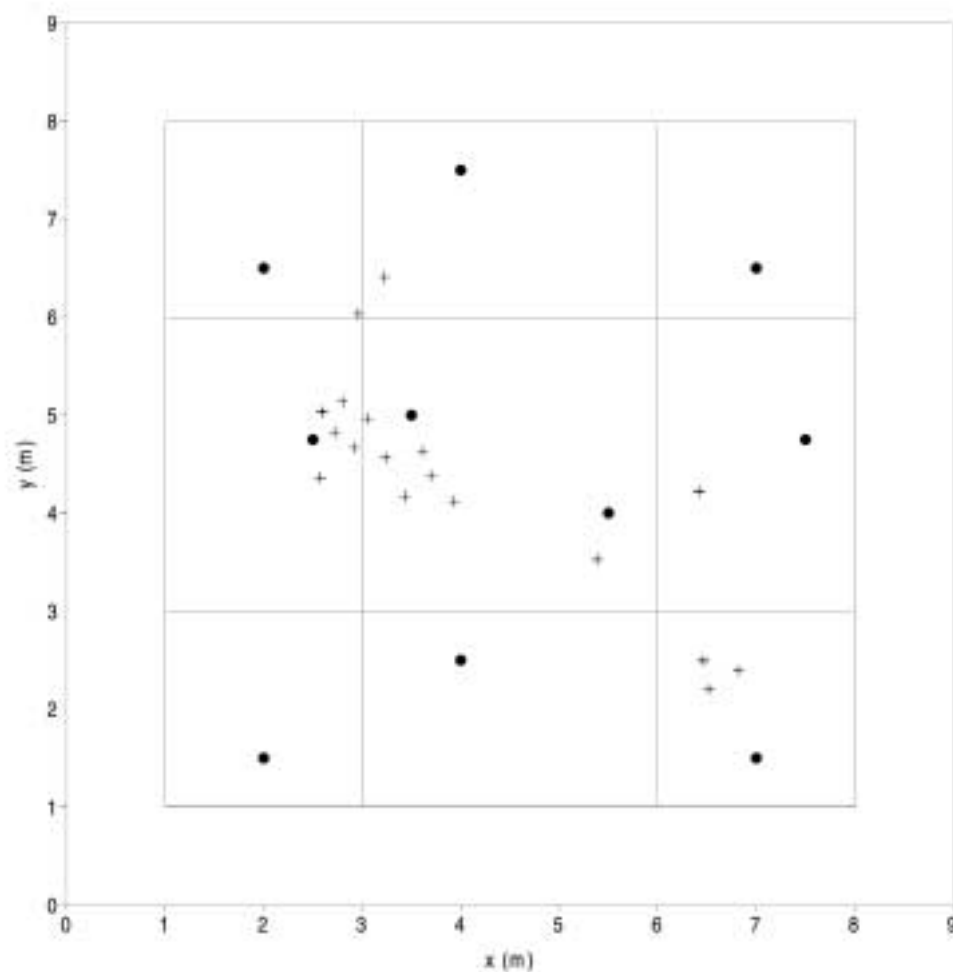
Een belangrijk aandachtspunt in de ruimtelijke statistiek betreft het doen van waarnemingen. Vaak moeten waarnemingen op de een of andere manier optimaal zijn, om zo veel mogelijk informatie krijgen met de beschikbare hulpmiddelen. Er zijn veel argumenten om hierbij een statistisch schema te kiezen, dat gedefinieerd is als een schema dat althans ergens een toevalscomponent kent. Het kan dus een totaal toevallig schema zijn, waarbij ieder punt opnieuw geloot wordt, of een rooster van waarnemingen, waarbij de oorsprong via loting wordt bepaald. Zo'n schema kan echter lastig te realiseren zijn. Geen waarnemer doet graag een waarneming in een weiland met een loslopende stier. Bovendien kan tijdens het uit-

voeren van een waarneming blijken dat een waarneming niet kan worden uitgevoerd - in bodemverontreinigingsstudies kan op sommige plaatsen puin liggen, er kan een gebouw staan, etc. Maar ook kan er al informatie bekend zijn in de vorm van betrouwbare gegevens. Als alternatief is in recente studies de Monte Carlo achtige techniek *simulated annealing* toegepast om het optimale bemonsteringsschema te bepalen.

In een recente studie hadden we belangstelling voor het opzetten van een schema voor het bemonsteren van nematoden (kleine bodemdieren) in een veld van 27×27 m in de Bovenbuurtse Weilanden, tussen Wageningen en Bennekom. Het veld werd vrijgemaakt van vegetatie en werd vervolgens verdeeld in negen 9×9 m plots. Hieraan werden op basis van loting drie behandelingen toegekend: T_1 geen vegetatie, T_2 mono-specifiek gras (*Agrostis capillaris*), en T_3 poly-specifiek gras (een mengsel van 15 soorten gras en onkruid). Het primaire doel bestond uit het schatten van de parameters die de ruimtelijke samenhang beschrijven via een zogenaamd variogram van een aantal bodemvariabelen. Op basis van logistieke en financiële beperkingen is het maximum aantal ruimtelijke monsters vastgesteld op 252, dus op 28 waarnemingen in ieder van de negen plots.

Het ruimtelijk bemonsteringsschema werd als volgt opgezet. Er zijn 10 afstandsklassen van steeds 0,5m gedefinieerd (0.0 - 0.5 m, ..., 4.5-5.0 m) en een 11^e klasse voor puntenparen die verder dan 5.0 m uit elkaar lagen. Een regelmatige verdeling van de puntenparen over de afstandsklassen correspondeert met 34.4 puntenparen per afstandsklasse.

Een zuivere toepassing van het simulated annealing algoritme zou leiden tot een clustering van de waarnemingspunten. Omdat dat voor het vervolgonderzoek minder wenselijk is, zijn 10 punten per plot vast gekozen, terwijl de overige 18 punten via simulated annealing zijn geloot. Dit heeft geleid tot het schema in Figuur 3.



Conclusie

Ruimtelijke statistiek, of beter: de statistiek van ruimtelijke gegevens, is zich verder aan het ontwikkelen. De bovenstaande voorbeelden zijn slechts een topje van de ijsberg: tegenwoordig wordt veel aandacht besteed aan de dynamiek van ruimtelijke fenomenen, aan een geschikt gebruik van vóórinformatie en vooral ook aan het gebruik van bodem-, plant- en grondwatermodellen in relatie tot ruimtelijke informatie. Op de achtergrond bestaat een gedegen fundament van kanstheorie. Zeker met de opkomst van geografische informatiesystemen gaan de ontwikkelingen erg snel. Hierbij is niet alleen sprake van eenrichtingsverkeer waarbij statistische methoden worden gebruikt om ruimtelijke problemen op te lossen, maar er is ook sprake van een ontwikke-

Figuur 3. Het geoptimaliseerde bemonsteringsschema voor een van de 9 plots in de Bovenbuurtse Weilanden met 10 vaste punten (●) en 18 variabele punten (+)

ling van statistische methoden op basis van ruimtelijke vragen.

ALFRED STEIN is hoogleraar aan Wageningen Universiteit en bij het International Institute for Geo-Information Science and Earth Observation ITC in Enschede, <alfred.stein@bodlan.beng.wau.nl>.

KIRSTEN DE BEURS en JASJA DEKKERS waren afstudeerders bij de leerstoelgroep wiskundige en statistische methoden aan Wageningen Universiteit.

CHRISTIEN ETTEMA was als postdoc verbonden bij de leerstoelgroep bodemkwaliteit van Wageningen Universiteit.



In memoriam Richard Freling

ALBERT P.M. WAGELMANS

Op 29 januari 2002 overleed Richard Freling, wetenschappelijk onderzoeker bij het Econometrisch Instituut, Erasmus Universiteit Rotterdam. Hij was reeds enige tijd ernstig ziek. Richard is slechts 34 jaar oud geworden.

Richard Freling heeft een groot gedeelte van zijn leven op de Erasmus Universiteit doorgebracht. Hij begon daar in 1987 met de studie Econometrie en koos in de doctoraalfase de afstudeerrichting Bedrijfseconometrie. Als student viel hij aanvankelijk weinig op, maar aan het eind van zijn studie meldde hij zich als één van de weinige studenten aan voor - in het kader van het ERASMUS uitwisselingsprogramma - een afstudeerproject aan de Universiteit van Lissabon. Onder begeleiding van professor José Pinto Paixão schreef hij daar zijn afstudeerscriptie *Vehicle scheduling with time constraint* (zie voor een artikelversie Freling en Pinto Paixão 1995), waarmee hij in 1992 zijn studie in Rotterdam voltooide.

Het liefst was hij weer naar Lissabon terugge-

keerd. In Portugal was Richards belangstelling voor wetenschappelijk onderzoek gewekt. Gesprekken met Paixão leidden tot de start van zijn promotieonderzoek over integratie van voertuig- en personeelsplanning in het openbaar vervoer. Bovendien had Richard in Lissabon zijn toekomstige vrouw Teresa leren kennen, die nog enige tijd nodig had om haar studie psychologie af te ronden. Aan de Universiteit van Lissabon bleek geen promotieplaats beschikbaar, maar Richard wist dankzij Guus Boender, zijn latere promotor, een aio-aanstelling in Rotterdam te bemachtigen. Zij spraken af dat het promotieonderzoek voorlopig in Lissabon zou worden uitgevoerd. Boender typeerde Richard als 'een bijzonder aardige jongen' en geeft daarmee kernachtig weer hoe Richard bij anderen overkwam. Hij was een rustige, geduldige en betrouwbare persoon met een positieve uitstraling. Iemand die zeer geliefd was bij velen die met hem te maken hadden.

Volgens plan werden de eerste anderhalf jaar van het onderzoek in Lissabon uitgevoerd, waarna

Richard naar Rotterdam terugkeerde, samen met Teresa met wie hij inmiddels getrouwd was. Bij het Econometrisch Instituut werd het proefschrift (Freling 1997) voltooid en met succes verdedigd. Het aardige is dat de Portugees-Nederlandse inbreng in het proefschrift niet alleen tot uitdrukking kwam in het feit dat Paixão en Boender beide als promotor optraden, maar ook in het feit dat de ontwikkelde technieken werden toegepast op probleeminstaties bij zowel CARRIS als de RET, de openbaar vervoerbedrijven van Lissabon respectievelijk Rotterdam. Het promotie-onderzoek leverde diverse publicaties op, waaronder een artikel in *Transportation Science* (Freling et al 2001).

Na zijn promotie trad Richard in dienst bij ORTEC Consultants BV te Gouda, met als bijzondere taak te adviseren over het gebruik van geavanceerde operations research technieken bij de afdelingen Luchtvaart en Logistiek. Ook kwam hij in aanraking met planningsproblemen bij de spoorwegen (zie onder andere Lentink et al 2000) en ontwikkelde hij interesse voor revenue management (zie bijvoorbeeld De Boer et al 2002). De aanstelling bij ORTEC was voor vier dagen per week. De vijfde werkdag bracht hij meestal door op het Econometrisch Instituut om wetenschappelijk onderzoek te verrichten, waaraan hij veel plezier en voldoening beleefde. Hoewel hij succesvol was in zijn werk als consultant, wilde hij zich richten op meer fundamentele vragen en wilde hij grondiger te werk gaan dan binnen de consultancy-wereld vaak mogelijk is. Deze behoefte tot verdieping kwam ook tot uiting in Richard's privé-leven waarin spirituele zaken, in het bijzonder de Brahma Kumaris, een steeds grotere rol waren gaan spelen.

Richards wens om meer tijd te besteden aan wetenschappelijk onderzoek kon in 1999 worden gerealiseerd via een aanstelling als onderzoeker bij het Econometrisch Instituut. Naast zijn onderzoeksactiviteiten leverde hij ook een bijdrage aan het onderwijs. Dankzij zijn praktische kennis,

geduld en heldere manier van uitleggen oogste hij veel waardering van studenten, vooral bij de begeleiding van werkcollege's (case studies) en scripties. Maar niet alleen studenten hebben veel van Richard opgestoken, ook collega's hebben van hem geleerd. Dankzij zijn specifieke expertise konden projecten worden uitgevoerd die voorheen niet tot de mogelijkheden behoorden. Als geen ander verstond hij de kunst om een goed evenwicht te vinden tussen wat theoretisch bekend en wat praktisch mogelijk is. Een mooi voorbeeld daarvan is zijn bijdrage aan het RINTEL Pilot-project, over het optimaal rangeren van treinstellen, dat in opdracht van NS Reizigers werd uitgevoerd (Freling et al, in voorbereiding).

Richard genoot van zijn werk op de universiteit en het leverde een aantal mooie resultaten op (naast reeds genoemde publicaties onder andere Freling et al 1999, Freling et al 2000 en Huisman et al 2001). Het is hard om te moeten constateren dat hij ziek werd in een periode dat hij zich in zijn werk zo voorspoedig ontwikkelde. Verschillende projecten die nog door hem zijn opgestart zullen nu door zijn collega's worden voltooid. Maar er zijn ook plannen die zonder hem niet meer gerealiseerd zullen worden.

Richard's wetenschappelijke werk wordt gekenmerkt door het feit dat het zowel theoretisch solide als praktisch relevant is. Zijn voornaamste bijdrage aan de besliskunde bestaat uit het pionierswerk dat hij heeft verricht met betrekking tot het onderzoek naar de integratie van planningsprocessen in het openbaar vervoer. Helaas was zijn wetenschappelijke carrière van korte duur.

Op 23 mei organiseerde het Erasmus Center for Optimization in Public Transport in samenwerking met het Nederlands Genootschap voor Besliskunde een conferentie over optimalisering in het openbaar vervoer. Behalve enkele Nederlands experts hielden drie gerenommeerde

buitenlandse sprekers een voordracht. Richard was aanvankelijk de drijvende kracht achter de organisatie van de conferentie, maar heeft dat werk vanwege zijn ziekte niet kunnen voltooien. Uiteraard werd op deze dag op gepaste wijze stil gestaan bij zijn overlijden.

LITERATUUR

Een selectie uit de publicaties van Richard Freling

Boer, S.V. de, Freling, R. en Piersma, N. (2002).

Stochastic Programming for multiple-leg network revenue management. *European Journal of Operational Research*, 137, 72-92.

Freling, R. (1997). *Models and Techniques for Integrating Vehicle and Crew Scheduling*. Amsterdam: Thesis Publishers.

Freling, R., Huisman D. en Wagelmans, A.P.M. (2000). *Models and Algorithms for Integration of Vehicle and Crew Scheduling*. Econometric Institute Report 2000-10/A. Rotterdam: Erasmus Universiteit Rotterdam (geaccepteerd voor publicatie in *Journal of Scheduling*).

Freling, R., Lentink, R.M., Kroon, L.G. en Huisman, D. Passenger Train Shunting in a Railway Station (in voorbereiding).

Freling, R. en Pinto Paixão, J.M. (1995). Vehicle Scheduling with Time Constraint. In: Daduna, J.R., Branco, I. en Paixão, J. (eds.). *Computer-Aided Scheduling of Public Transport*. Berlin: Springer, 130-144.

Freling, R., Romeijn, H.E., Romero Morales, D. en Wagelmans, A.P.M. (1999). *A Branch and Price Algorithm for the Multi-Period Single-Sourcing Problem*. Econometric Institute Report 9941/A. Rotterdam: Erasmus Universiteit Rotterdam (geaccepteerd voor publicatie in *Operations Research*).

Freling, R., Pinto Paixão, J.M. en Wagelmans, A.P.M. (2001). Models and Algorithms for Single-Depot Vehicle Scheduling. *Transportation Science*, 35, 165-180.

Huisman, D., Freling, R. en Wagelmans, A.P.M. (2001). *A Dynamic Approach to Vehicle Scheduling*. Econometric Institute Report EI2001-17. Rotterdam: Erasmus Universiteit Rotterdam (onder revisie voor *Transportation Science*).

Lentink, R.M., Odijk, M.A., Freling, R. en Wit, J.S de (2000). Use of Operations Research to Facilitate and Improve Railway Planning. *Computers in Railways VII*. Southampton: WIT Press, 231-239.

ALBERT P.M. WAGELMANS is bijzonder hoogleraar wiskundige beslisplanning aan het Econometrisch Instituut, Erasmus Universiteit Rotterdam, Postbus 1738, 3000 DR, Rotterdam <wagelmans@few.eur.nl>

Illustratie: Toon Hartogs

Sara van de Geer

Statistiek en genen

Vroeger ging men biologie studeren omdat men een hekel had aan wiskunde en statistiek, maar tegenwoordig zijn geavanceerde wiskundige modellen en methoden uit de statistiek een fundamentele component van biologisch onderzoek. Woorden van deze strekking werden onlangs geuit tijdens een bijeenkomst van de *task force bio-informatics*. De tijden zijn veranderd!

Moderne statistiek verschilt in belangrijke opzichten van de klassieke statistiek. Ik noem een voorbeeld. Voor de statistische analyse van DNA micro-arrays zijn technieken nodig die zo'n 10.000 gemeten gen-expressieniveaus terugbrengen tot een overzichtelijk aantal relevante variabelen. Een hachelijke onderneming voor een klassiek statisticus, die opgevoed is met de regel dat het aantal parameters nooit groter mag zijn, en liefst veel kleiner, dan het aantal waarnemingen. De statistische analyse van micro-array gegevens is dus wezenlijk anders dan de statistische analyse aan het begin van de vorige eeuw, van Gregor Mendels bewering dat er genen bestaan. (Deze bewering werd door middel van het toen pas geïntroduceerde begrip statistische correlatie onderzocht.)

Grote datasets

Bio-informatica is een vak waarbij bekeken wordt hoe de enorme hoeveelheden gegevens, over bijvoorbeeld het menselijk genoom, kunnen worden opgeslagen, toegankelijk gemaakt en gebruikt. Het is natuurlijk niet het enige vak waar men met grote datasets te maken heeft. Overall wordt ijverig verzameld: financiële data, kosmische data, marketing data, sensor en satelliet data, etc. De toename van de geheugencapaciteit en rekenkracht van computers speelt hier een belangrijke rol. Voor de

statistiek zijn complexe datasets natuurlijk een prachtige uitdaging! Hoog-dimensionele ruimtes moeten op een slimme manier worden doorzocht, er wordt geprojecteerd, geselecteerd, geclusterd en geconcentreerd. Maar veel potentiële toepassers hebben dit statistische gegons niet gehoord en ontwikkelen zelf technieken voor hun hoog-dimensionele problemen. Daarbij dreigt het gevaar dat men eerst door de valkuil van het perfectionisme moet, dat wil zeggen dat een specifieke complexe situatie zo wordt gemodelleerd dat alle geobserveerde verschijnselen zijn verklaard. In de wereld van neurale netwerken wordt het vermijden van deze valkuil wel *optimal brain damage* genoemd.

Ouderwetse statistiek

Ondanks de nieuwe rol van de zich vernieuwende statistiek blijft er een ouderwets beeld van het vak de overhand voeren. 'Statistiek, dat is toch foutenanalyse?' hoorde ik laatst uit chemische hoek. Dit was misschien 100 jaar geleden zo voor een sterrenkundige die de baan van een planeet moest berekenen op grond van een klein aantal waarnemingen. Tegenwoordig worden de afstanden tot miljoenen sterrenstelsels bepaald, en de helderheid van honderd miljoenen hemellichamen.⁵ Natuurlijk behoudt de klassieke 'foutenanalyse' zijn waarde, maar er is veel meer onder de zon!

Een niet geheel up-to-date beeld van statistiek kwam ik ook tegen in een overigens belangwekkend proefschrift van Zielman over de analyse van scheef-symmetrische gegevens.⁶ Daar wordt beweerd dat er voor de meeste multi-dimensionele schalings-technieken die gebruikt worden in de sociale wetenschappen, nog steeds geen statistische theorie is.

Het proefschrift van de Mast gaat juist in op de grote waarde van statistische methoden bij kwaliteitsverbetering in de industrie.³ Dit proefschrift stelt onder andere de cultuur aan de kaak om problemen volgens een vaststaand protocol te lijf te gaan. Zo'n protocol komt op mij over als de 'LOSOP' methode van Winnie-the-Pooh (Lokaliseer,

Observeer, Stel vragen, Opper mogelijke oplossingen, Pas de oplossing toe).¹ De invulling van de eigenlijke statistische analyse blijkt daarbij meestal heel bescheiden en traditioneel te zijn.

Tellen

In ander empirisch onderzoek lijkt de statistische analyse terug te worden gebracht tot enkel tellen. In de oratie van Huizinga lees ik: 'Men kijkt welk gedeelte van de mensen rookt, kijkt welk gedeelte van de bevolking longkanker krijgt, en constateert dat rokers vaker longkanker krijgen. In feite komt dit vakgebied neer op leren tellen.'⁴ Huizinga noemt bovenstaande een voorbeeld van 'de getalsmatige aanpak'. Hij wijst overigens ook op de hoog-dimensionaliteit van bijvoorbeeld DNA data, en zegt te verwachten dat daardoor 'de rol van de subjectiviteit in de wetenschap meer geprononceerd zal worden'!

Beschouwen we meer in het algemeen het onderkennen en modelleren van onzekerheid, dan blijkt bijvoorbeeld uit de statistieken dat van een steekproef van 93 artikelen in 1992 op het gebied van health economics, zo'n 24 procent nalaat een gevoeligheidsanalyse uit te voeren.²

Statistiek schijnt niet in ieders genen ingebakken te zitten.

NOTEN

1. Allen, E.A. en Allen, S.D. (1995). *Winnie-the-Pooh on Problem Solving*. E.P. Dutton, Inc.- New York.
2. Briggs, A. en Sculpher, M. (1995). Sensitivity analysis in economic evaluation: a review of published studies. *Health economics* 4, 355-371.
3. de Mast, J. (2002). *Quality Improvement from the Viewpoint of Statistical Method*. Proefschrift, Universiteit van Amsterdam.
4. Huizinga, T.W.J. (2002). *Experimentele Reumatologie*. Oratie, Universiteit Leiden.
5. Zie www.sdss.org
6. Zielman, A.J.W. (2002). *The Analysis of Skew-Symmetry in Proximity and Dominance Data*. Proefschrift, Universiteit Leiden.

SARA VAN DE GEER is hoogleraar Kansrekening en Statistiek bij het Mathematisch Instituut, Universiteit Leiden, <geer@math.leidenuniv.nl>.



Vereeniging voor de Statistiek (VVS) een gezelschap van juristen

Sedert 1945 bestaat in Nederland de Vereniging voor Statistiek (en Operationele Research), afgekort tot VVS. De oprichters waren er zich waarschijnlijk niet van bewust dat zij een naam kozen, die reeds een eeuw eerder een vereniging sierde: de Vereeniging voor de Statistiek in Nederland. Alleen de spelling van het woord Vereeniging is veranderd, maar de afkorting VVS is dezelfde gebleven. Ida Stamhuis verdiepte zich in het ontstaan en vergaan van de Vereeniging.

IDA H. STAMHUIS

De eerste aanwijsbare activiteiten met als doel een statistische organisatie op te richten dateren uit 1846. Een dergelijke oprichting vormde een gesprekstema op de vergadering van de *Sectie der Regtsgeleerdheid en Staatswetenschappen* van het *Provinciaal Utrechts Genootschap voor Kunsten en Wetenschappen (PUG)*. De hoogleraar in de statistiek aan de Utrechtse universiteit, J. Ackersdijk bracht dit onderwerp te berde in de jaarvergadering van 1846. Hij betoogde dat de overheid zich de laatste jaren weinig aan de statistiek gelegen had laten liggen en dat was voor hem aanleiding voor

te stellen een particuliere vereniging te laten voorzien in deze leemte. Een jaar later op het zogenaamde *Landhuishoudkundig Congres* werd duidelijk dat nog meer organisaties zich de zaak van de statistiek hadden aangetrokken. Er hing blijkbaar iets in de lucht en het argument dat steeds werd gebruikt was dat de overheid haar statistische plicht verzaakte. Er was in 1826 door de overheid wel een *Statistisch Bureau* opgericht en een *Statistische Commissie* ingesteld, maar die hadden niet de kans gekregen om activiteiten te ontplooiën. In 1830 na de afsplitsing van België, toen de invloed van de

beroemde Belgische statisticus Adolphe Quetelet was verminderd, waren ze in rook opgegaan.

Behoeftte aan feitelijke kennis

Hoe kan het dat er halverwege de negentiende eeuw meerdere stemmen klonken van personen die de statistiek zo'n warm hart toedroegen dat ze daaraan een vereniging wilden wijden? Dat had er mee te maken dat er vanaf 1802 colleges in de statistiek werden gegeven, eerst alleen in Leiden door Adriaan Kluit, maar later ook aan andere universiteiten. In Leiden werd in 1815 bepaald, dat colleges in de statistiek verplicht waren voor alle studenten in de juridische wetenschappen. Het gevolg was dat een behoorlijk contingent aan afgestudeerden ontstond, dat met de statistiek had kennisgemaakt en van het belang ervan was overtuigd.

Vanuit het perspectief van de hedendaagse statistiek is het ronduit verbazingwekkend dat de rechtenfaculteit als de aangewezen faculteit werd beschouwd voor dergelijke colleges. Wat was daarvan de reden? Evenals nu kwamen juristen toen vaak in overheidsfuncties terecht en moesten overheidsbeleid ontwerpen en uitvoeren. In Duitsland was het idee ontstaan dat aan de basis van overheidsbeleid feitelijke kennis van een staat ten grondslag moet liggen. Men was daar gaan nadenken over hoe deze feitelijke kennis eruit zou moeten zien en hoe deze moest worden gesystematiseerd, zodat ook verschillende staatkundige eenheden met elkaar zouden kunnen worden vergeleken. Daaraan was in het achttiende-eeuwse Duitsland, bestaande uit vele staten en staatjes, grote behoefte. Statistiek was dus aanvankelijk een systematische beschrijving van een staat en die betekenis zit ook nog in het woord *statistiek*.

De eerste statistiekcolleges

Wat behandelde de eerste hoogleraar in de statistiek in Nederland, Adriaan Kluit, in zijn colleges? Wij zijn daarvan heel nauwkeurig op de hoogte, omdat zijn eigen aantekeningen en collegediktaten

van enkele van zijn studenten in de Universiteitsbibliotheek van Leiden worden bewaard. Een collegediktaat van toen gaf vrijwel letterlijk weer wat de hoogleraar op zijn college onderwees. De titels van de collegediktaten zijn: *Statistiek van Nederland*, *Statistiek der Bataafsche Republiek*, *Voorlezingen over de Staathuishoudkunde der Nederlanden* en *Bouwstoffen voor een Nederduitsch Collegie over Staathuishoudkunde*. Statistiek en staathuishoudkunde waren beiden vakgebieden in ontwikkeling en men maakte er aanvankelijk geen onderscheid tussen.

De statistiek bevatte volgens Kluit de kennis die nodig zou zijn om de 'waare kragten', ofwel de huishouding, ofwel de macht en de welvaart van een land te leren kennen. Kluit onderscheidde twee aspecten aan de huishouding van een staat, namelijk het *natuurlijke* en het *zedelijke*. Onder de natuurlijke gesteldheid van een land verstond hij haar ligging, grenzen en naburen, grootte en klimaat. Bij het zedelijke behandelde hij allereerst de bewoners van het land, onder andere in staatkundig opzicht, maar ook hun aantal en hun karakter, evenals hun middelen van bestaan en hun welstand. Deze konden worden behandeld aan de hand van een bespreking van visserij, akkerbouw, veeteelt, industrie, koophandel en geldzaken. Als tweede onderdeel van het zedelijke beschouwde Kluit de regering.

Zowel economische, antropologische, sociologische als geografische onderwerpen kwamen bij Kluits bespreking van een staat aan bod.

De rol van kwantitatief redeneren

Kluit maakte bij zijn bespreking soms gebruik van getallen, maar meestal niet. De reden daarvoor zal zijn geweest dat er over een bepaald onderwerp geen getalsmatige gegevens beschikbaar waren, of dat het zich niet leende voor een weergave in getallen. Voor Kluit speelde het getal geen belangrijke rol. In de statistiek ging het hem er in de eerste plaats om een indruk te krijgen hoe de feitelij-

ke toestand van een staat was, op welke manier dat gebeurde kwam op de tweede plaats.

De behandeling van het *bevolkingsaantal* is het meest kwantitatieve gedeelte van Kluits statistiek. Kluit onderscheidde meerdere methoden om dat te doen, bijvoorbeeld het tellen van de strijdbare mannen en dit aantal met 4 of 5 te vermenigvuldigen. Verder was er de methode gebaseerd op 'de theorie der probabiliteit van 't leven'. Deze methode berustte erop dat er een constante verhouding bestond tussen het aantal inwoners en het jaarlijks aantal gestorvenen. In de steden was de sterfte hoger; voor de grote steden zou de verhouding tussen het jaarlijkse sterfteaantal en het bevolkingsaantal tussen 1 op 24 en 1 op 28 zijn, en op het platteland tussen 1 op 40 en 1 op 42.

Kluit bracht het bevolkingsaantal met andere gegevens in verband. Hij vond het belangrijk om te weten of een land voldoende bevolkt was. In die tijd was de mening algemeen dat het bevolkingsaantal een maat is voor de macht van een land. De aantallen vertoonden tussen landen grote verschillen, voor Zweden 230 per vierkante mijl, Portugal 1325 en Nederland leverde het hoogste aantal: 4800. Voor deze verschillen voerde Kluit een aantal verklaringen aan. Zo oefende volgens hem de godsdienst invloed uit op de omvang van de bevolking. In Nederland had de religie positief gewerkt. Hetzelfde gold, dankzij de traditionele vrijheid, voor de regeringsvorm. Het tegengestelde gold voor het hoge niveau van de belastingen.

'Zaakkennis van staatsaangelegenheden'

Na Kluit ontstonden er al spoedig twee vakken: statistiek en staathuishoudkunde. Een flink aantal juristen had onderwijs in deze vakken gehad en raakte van het belang van de statistiek overtuigd. Dat was de achtergrond van de eerste gesprekken over de wenselijkheid van een *Vereeniging voor Statistiek*. Deze discussies liepen uiteindelijk op niets uit, want toen de overheid in 1848 een *Statistisch Bureau* instelde, voelde men zich niet

langer geroepen om er zich mee bezig te houden.

Langs een andere weg is er toch een statistisch genootschap ontstaan. Het *Staatkundig en Staathuishoudkundig Jaarboekje* heeft daarbij een belangrijke rol gespeeld. Het jaarboekje was een initiatief van Jeronimo De Bosch Kemper. De grondwetswijziging van 1848, waardoor de Tweede Kamer rechtstreeks werd verkozen, was voor deze jurist aanleiding met het jaarboekje te beginnen. Het verscheen voor het eerst in 1849. Volgens hem was het noodzakelijk dat de kiezers op de hoogte waren van de onderwerpen die een rol speelden in de landelijke politiek

In het Voorberigt van de tweede jaargang werd ingegaan op de inhoud en de bedoeling van het jaarboekje. Het ging om het opdoen van 'zaakkennis van staatsaangelegenheden', en om het 'opmerkzaam letten op verschijnselen, - oorzaken na te vorschen, ten einde gevolgen te kunnen berekenen, - gedurig vergelijken, ten einde wetten op te sporen'. Daarvoor kon de wetenschap der statistiek worden gebruikt, die 'in dezelfde verhouding staat tot de staatkunde en staathuishoudkunde als de natuurkunde tot de geneeskunde'. Het succes van de natuurkunde in de ontwikkeling van de geneeskunde gold dus als voorbeeld voor de rol die de statistiek in de juridische en economische wetenschappen zou moeten spelen.

Oprichting van de VVS

In 1857 benutte de Leidse hoogleraar in de statistiek Simon Vissering de jaarlijkse vergadering van de redactie van het jaarboekje om voor te stellen dat formeel zou worden beslist een Nederlandsch statistisch genootschap op te richten, dat zich zou gaan bezighouden met het verzamelen en het verspreiden van statistische kennis. Het lukte Vissering niet om de overige medewerkers te winnen voor zijn ideeën. Wel werd een besluit genomen dat leidde tot de minder officiële oprichting van een statistisch genootschap. Men besliste dat het aantal medewerkers aan het jaarboekje zou worden



uitgebreid. Vanaf 1857 werd het jaarboekje volgens het titelblad door de *Vereeniging voor de Statistiek* uitgegeven, zodat dit jaar kan worden beschouwd als het jaar van ontstaan. De Bosch Kemper werd benoemd tot president. In 1862 kreeg de vereniging rechtspersoonlijkheid. De aanleiding om dat aan te vragen was dat een *Rijkscommissie voor de Statistiek*, opgericht in 1858, in 1861 al weer was opgeheven.

Sociale vraagstukken

In 1864 stelde de Utrechtse hoogleraar in de statistiek Otto van Rees voor naast de jaarvergadering over de inhoud van het jaarboekje een tweede vergadering te beleggen. Van Rees wilde de vereniging een meer vakinhoudelijke functie geven. Hij stelde voor dat op een dergelijke extra vergadering van gedachten zou worden gewisseld over statistische en staathuishoudkundige onderwerpen. De vergadering voelde hier voor mits het onderwerpen zou betreffen 'die in het volksleven ingrijpen'. De besprekingen zouden licht (moeten) verspreiden over den toestand der arbeidende klasse.

Op de eerste extra vergadering werden drie onderwerpen aan de orde gesteld, namelijk arbeidersverenigingen, het voor en tegen van kosteloos onderwijs, en het ongunstige sterftcijfer in Nederland. Hieruit blijkt dat er binnen de *Vereeniging voor de Statistiek* belangstelling bestond voor sociale vraagstukken. Die interesse is altijd blijven bestaan. Eveneens illustreert de keuze van de onderwerpen, dat de thema's die werden besproken niet alleen de statistiek betroffen.

In 1866 stelde Vissering de vereniging voor een *Algemeene Statistiek*

van Nederland gebaseerd op gegevens over de jaren 1850 tot 1865, te gaan bewerken en uitgeven. Er werd een commissie van redactie benoemd onder zijn voorzitterschap. Het langdurige project (in 1874 rapporteerde de commissie voor de negende keer over haar werkzaamheden) is niet klaargekomen. Van de geplande vijf delen zijn er slechts twee verschenen. Om een dergelijke onderneming te volvoeren zou een overheidsinstelling nodig zijn.

De ruimte laat niet toe om uitgebreid in te gaan op het ledenbestand van de *Vereeniging*. Economisch-geïnteresseerde juristen en overheidsambtenaren hebben deze organisatie altijd gedomineerd en de omvang van de vereniging is groot geweest; in het topjaar 1889 waren er 746 personen lid. Verder moet worden opgemerkt dat de vereniging gedurende zijn gehele bestaan altijd heeft getracht de overheid ertoe te brengen een volwaardige overheidsstatistiek op poten te zetten. In 1884 richtte de Vereeniging het *Statistisch Instituut* op, omdat de overheid opnieuw niet bereid was om de statistiek tot haar taak te rekenen. Hoewel het instituut hoogwaardig bemenst werd (Anthony Beaujon werd directeur en hoogleraar statistiek aan de Amsterdamse universiteit) bleken de middelen te beperkt om effectief te kunnen opereren.

'Waartoe spannen wij ons in?'

Uiteindelijk, in 1892, na vragen in de Tweede Kamer van onder meer het socialistische kamerlid F.J. Domela Nieuwenhuis, werd de nu nog altijd bestaande *Centrale Commissie voor de Statistiek* ingesteld. De oprichting van het *Centraal Bureau voor de Statistiek (CBS)* in 1899 was daarna een logisch gevolg.

De vraag was toen, of er nog een taak voor de *Vereeniging* over bleef. Het *Statistisch Instituut* werd opgeheven, maar wat moest er verder gebeuren? In de speciaal bijeengeroepen vergadering deelde de voorzitter mee, dat het bestuur zich had afgevraagd of de *Vereeniging* zich zou moeten ontbinden. Dat vond het bestuur echter te ver gaan, hoewel



Simon Vissering (1818-1888)

één van haar belangrijke doelstellingen, een goed georganiseerde overheidsstatistiek, bereikt leek. Om twee redenen vond het bestuur opheffing geen juiste reactie. In de eerste plaats omdat het nog maar de vraag was of de overheid zou voortgaan op deze ingeslagen weg. De voorzitter: *‘Men zou kunnen zeggen: de Staat doet nu het werk, waartoe spannen wij ons in? Doch in’t algemeen kan men zeggen, dat concurrentie ook hier, wat ons betreft op bescheiden schaal, nuttig is. En dan; wij hebben het meer gezien; door de lotswisselingen der politiek kon wel eens het door de Regeering nu zoo kloek opgerichte gebouw te gronde gaan. Reeds eens is een Centrale Rijkscommissie voor de Statistiek begraven. Welnu dan is altijd onze Vereeniging, wanneer zij haar kader behoudt, daar, om de draad weder op te vatten en den statistischen arbeid voort te zetten.’*

De tweede reden was dat ze op een voor de hand liggende wijze haar interessegebied kon bijstellen. Ze zou voortaan niet langer worden aangeduid als *Vereeniging voor de Statistiek*, maar als *Vereeniging voor Staathuishoudkunde en Statistiek*. De voorstellen van het bestuur werden bij acclamatie aangenomen.

In de jaren voorafgaand aan 1892 waren op de jaarvergaderingen als regel zowel statistische als economische onderwerpen aan de orde geweest. Na de naamsverandering in 1892 stonden er alleen economische onderwerpen op het programma. Meer dan een halve eeuw later, in 1950, veranderde de vereniging haar naam. De directe aanleiding was de oprichting van de huidige *Vereniging voor Statistiek* in 1945. De *Vereeniging voor Staathuishoudkunde en Statistiek* kreeg toen de naam waaronder ze nu nog bestaat: *Vereniging voor de Staathuishoudkunde*. Daarmee werd pas in 1950 een stand van zaken bevestigd die feitelijk bestond vanaf 1892.

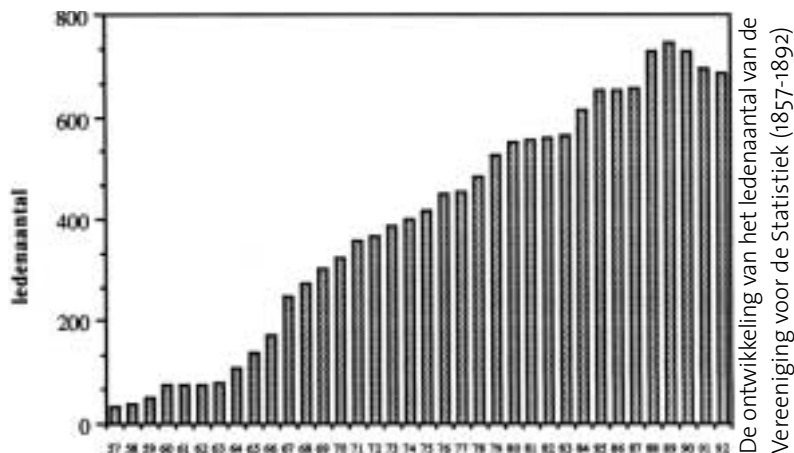
Minimale rol van de wiskunde

De *Vereeniging voor de Statistiek* was een geheel andere vereniging dan de huidige VVS. De personen die in 1945 de huidige VVS oprichtten, waren wis-

kundig georiënteerd en dat kon je van de leden van de *Vereeniging voor de Statistiek* niet zeggen. Waren het aanvankelijk juristen met een statistische belangstelling; later waren ze meer economisch georiënteerd. Was de statistiek aanvankelijk niet exclusief getalsmatig; in de loop van de negentiende eeuw ging het getal een overheersende rol in de statistiek spelen. De rol van de wiskunde bleef echter minimaal. De oprichters van de huidige VVS zullen vermoedelijk helemaal niet geweten hebben dat de *Vereeniging voor Staathuishoudkunde*, die tegenwoordig de economen in Nederland organiseert, als *Vereeniging voor de Statistiek* is begonnen.

Echter, het is niet zo, dat de statistiek van nu niets meer met de statistiek van toen te maken heeft. Het huidige CBS is een resultaat van de inspanningen van negentiende eeuwse statistici en in de activiteiten van het CBS is dat nog terug te zien. Streeft het CBS er ook tegenwoordig niet naar om een systematische beschrijving van de samenleving te geven? De vorm waarin dat gegoten wordt; daaraan is veel veranderd. Systematisch heeft zich ontwikkeld naar getalsmatig en de aandacht is verlegd van de compleetheid van de data naar de methode van het verzamelen en de analyse ervan.

Ida H. Stamhuis is in 1989 gepromoveerd op het proefschrift ‘Cijfers en Aequaties’ en ‘Kennis der Staatskrachten’. Statistiek in Nederland in de Negentiende Eeuw. Zij werkt als wetenschapshistoricus bij de Afdeling Algemene Vorming, Faculteit Exacte Wetenschappen, van de Vrije Universiteit Amsterdam, <stamhuis@nat.vu.nl>.





Aharon Ben-Tal

een robuuste optimist

Aharon (Ronny) Ben-Tal is een vooraanstaand wetenschapper op het gebied van robuuste optimalisering. Na zijn promotie in 1973 aan de *Northwestern University* werkt hij sinds 1975 als hoogleraar Operations Research aan de *Faculty of Industrial Engineering and Management* van het Technion in Haifa. Hij is medeauteur van twee boeken over niet-lineair programmeren en heeft ruim 80 publicaties op zijn naam staan. Ook is hij redacteur van een viertal vooraanstaande tijdschriften. Vorig jaar bracht hij een sabbatical aan de TUD door.

FRITS VAN BECKUM EN HANS MELISSEN

Hoe zou u uw eigen gebied, uw specialisme, willen karakteriseren?

Als ik één woord zou moeten geven, dan zou dat 'optimalisering' zijn. Dit woord omvat meer dan wat wordt verstaan onder 'mathematisch programmeren'. Voor mij betekent het ook optimale besturing en ook in zekere mate de approximatie-theorie; het zou stochastisch programmeren kunnen omvatten. Ook het optimaliseren in ruimten

die niet noodzakelijk eindig-dimensionaal zijn kan en moet er eigenlijk toe gerekend worden. Onder mathematisch programmeren verstaan we meestal optimalisering van een doelfunctie met een eindig aantal nevenvoorwaarden in een eindig aantal dimensies. Mijn werkterrein is dus optimalisering, verbonden met sommige onderwerpen uit de fundamentele wiskunde, onder andere de reële analyse, lineaire algebra en kans-

rekening. Daarnaast vormt de functionaalanalyse, en in het bijzonder convexe analyse, de werkelijke ruggengraat van optimalisering, althans vanuit theoretisch oogpunt.

Ik zou me niet uitsluitend willen rekenen tot het theoretische, het toegepaste of het numerieke kamp. Mijn eerste vijftien wetenschappelijke jaren waren voornamelijk aan theoretisch werk gewijd, maar een jaar of tien geleden besloot ik een duik te nemen in het koude bad van de toepassingen. Doordat ik voornamelijk in de niet-lineaire optimalisatie werkte, merkte ik dat de natuurlijke voedingsbodem voor de toepassing daarvan lag in de engineering. Daar is het het meeste nodig en daar ligt dus ook de grootste kracht. Totaal anders dan in de traditionele OR, die zijn meeste toepassingen vindt in de zakenwereld, de economie en de logistiek. Deze toepassingen zijn vaak lineair, terwijl je in de engineering door de fysica en de techniek meteen in niet-lineariteiten belandt.

Ik heb voor het eerst aan die niet-lineaire toepassingen gewerkt tijdens mijn *sabbatical* aan de universiteit van Michigan in Ann Arbor, waar ik twee jaar doorbracht aan het *College of Engineering*. Ik gaf daar een cursus *Advanced Optimization Methods*. Tot mijn verbazing en die van anderen verschenen er, naast de gebruikelijke wiskundestudenten, zo'n zes, zeven hoogleraren van allerlei afdelingen als elektrotechniek, scheepsbouwkunde, luchtvaarttechniek en werktuigbouwkunde, elk met een paar studenten. En daar is mijn werk op dit gebied begonnen. Door contacten die daar ontstonden heeft het zich allemaal ontwikkeld en ben ik jaren bezig geweest met toepassingen van optimalisatie, bijvoorbeeld bij het ontwerpen van vakwerken en optimale vormen. Mijn ervaring uit die periode is, dat een goed toegepast probleem altijd interessante wiskundige vragen oproept, waardoor je automatisch terechtkomt in belangwekkende theorie. Je kunt toepassing en theorie dus niet scheiden, het is één grote cirkel.

Er wordt wel beweerd dat alle noodzakelijke wiskunde inmiddels ontwikkeld is; alle gereedschappen liggen klaar en het is nu tijd voor het toepassen op problemen.

Dat is een controversiële bewering. Ik zou zeggen dat dit niet het geval is. Als je kijkt in het pas bij SIAM verschenen boek van Arkadi Nemirovski en mijzelf, *Lectures on Modern Convex Optimization*, dan presenteren we daar iets 'moderns', technieken die niet meer dan een jaar of tien geleden op het toneel verschenen zijn, op een manier die fundamenteel verschilt van wat optimalisering was en nu is. Het leidt tot terreinen, ook van wiskunde, die tot dan toe ofwel niet beschikbaar waren, ofwel niet relevant geacht werden voor de optimalisatie. Daarbij doel ik op semi-definiete programmering en de zogenaamde kegel-programmering. In lineair programmeren is de eenvoudigste kegel het eerste orthant, de verzameling van vectoren met niet-negatieve kentallen. Andere interessante kegels zijn de tweede-orde kegel, dat is een kegel die in drie dimensies een ijscohortje zou zijn, de zogenaamde Lorentzkegel, die een jaar of vijftig geleden al door zuiver wiskundigen bestudeerd blijkt te zijn. De interessantste en belangrijkste is de zogenaamde semi-definiete kegel. Dat is de verzameling van alle matrices die positief semi-definiet zijn. Deze kegel induceert een partiële ordening op de verzameling van matrices.

Ik kan een soort filosofie geven van wat wij denken dat moderne optimalisering is. De oorsprong van wiskundig programmeren was het lineair programmeren. Je optimaliseert dan een lineaire doelfunctie onder een verzameling van lineaire ongelijkheden. Hiervoor heeft George B. Dantzig eind veertiger jaren de simplexmethode ontwikkeld en dat was zo'n succes dat lineair programmeren de meest gebruikte techniek in de toegepaste wiskunde geworden is. In het begin van de vijftiger jaren kwam daar het artikel van Kuhn-Tucker bij. Hierin worden de doelfunctie en de nevenvoorwaarden niet-lineair verondersteld.

Dit gezichtspunt heeft de weg gebaad voor de volgende dertig, vijfendertig jaar. Dit noemen we de *f-g*-formulering: optimaliseer een functie $f(\mathbf{x})$ voor een vector \mathbf{x} onder voorwaarden $g(\mathbf{x}) \leq 0$. Hoe minder we hoeven aan te nemen over de functies f en g hoe beter. Alles hangt nu af van de lokale eigenschappen van deze functies, de eerste en tweede afgeleiden. Daarop zijn algoritmen gebouwd.

Naar onze mening – de boodschap die wij in ons boek brengen – was dit een wat minder gelukkige historische ontwikkeling. Het was een megalomane sprong van lineair programmeren, dat heel specifiek is, naar de zeer abstracte *f-g*-formulering. Bij lineair programmeren kun je voor een probleem meteen het duale probleem opschrijven en de optimaliteitsvoorwaarden. Alles is uiterst expliciet en berekenbaar. Er is een calculus aanwezig. Voor elk niet-lineair probleem moet je echter iedere keer vanaf scratch beginnen: je moet het duale probleem berekenen, de optimaliteitsvoorwaarde vinden, of de noodzakelijkheidsvoorwaarde verifiëren, de Lagrangiaan zoeken, de Lagrangiaan optimaliseren. In de meeste gevallen lukt dat niet. Het is alsof we niet kunnen differentiëren en we voor iedere berekening terug moeten naar de definitie. We kennen echter ondertussen de afgeleide van een aantal functies, sinus, e-macht, en we hebben differentiatierregels voor producten, samenstellingen, etc. Dat noemen we calculus. Welnu, niet-lineaire programmering heeft geen calculus. In moderne optimalisering zeggen we: we gaan terug naar lineair programmeren. Je hebt een lineaire doelfunctie (dat kan, zonder verlies van algemeenheid) en je hebt $A\mathbf{x} = \mathbf{b}$ in een niet-negatieve kegel. We generaliseren nu niet de lineariteit van $A\mathbf{x} = \mathbf{b}$, maar de partiële ordening die door de kegel is gegeven. We behouden het affiene deel, en vervangen de ordening; $A\mathbf{x} = \mathbf{b}$ komt nu in een meer algemene kegel te liggen, zoals een semi-definiëte kegel of een tweede-orde kegel.

Op deze manier kun je een krachtige theorie creëren met calculusregels zoals de lineaire pro-

grammering die heeft. Nu kan men denken dat dit een veel zwakker model is dan de meer algemene niet-lineaire beschrijving van het *f-g*-model, maar dat blijkt niet het geval te zijn. Veel problemen blijken naar deze vorm omgeschreven te kunnen worden, en wat interessant is: het omvat de belangrijkste toepassingen.

Aan de andere kant is er een hele middengeneratie in de optimalisering die moet veranderen. Zij moeten nieuwe dingen leren. Het zijn niet langer de traditionele gereedschappen die opgeld doen. Sommige gereedschappen zijn moeilijker, dus de acceptatie in de optimaliseringsgemeenschap gaat langzaam. De wiskunde is anders dan alles wat je geleerd hebt. Bijgevolg hebben we met onze optimalisering een veel bredere aantrekkelijkheid voor andere delen van de wiskunde en voor toepassingen waarvoor we vroeger vrijwel niets te betekenen hadden. Ondertussen zijn deze technieken al behoorlijk geaccepteerd in twee andere contreien: in de combinatorische optimalisering, en bij de besturingsmensen.

Als je bijvoorbeeld in een besturingsprobleem de stabiliteit wilt verifiëren, dan heb je ofwel de Lyapunov- ofwel de Riccati-vergelijking als karakterisering van de stabiliteit. Als nu sommige gegevens in het dynamisch systeem met onzekerheid behept zijn en je wilt stabiliteit hebben voor iedere trajectorie, ongeacht de realisatie van de data binnen deze onzekerheid, dan krijg je in plaats van één Lyapunov-ongelijkheid nu een continuüm van zulke ongelijkheden, voor elke mogelijke realisatie van de data één. Dit is een NP-lastig probleem, hetgeen betekent dat je op geen enkele manier kunt garanderen dat het probleem in eindige tijd op te lossen is. Maar semi-definiëte relaxatie biedt tegenwoordig de mogelijkheid om dit soort problemen aan te pakken en een zeer goede benadering van de oplossing te vinden. Onzekerheid in de data speelt een belangrijke rol in wat wij de robuuste optimalisatiemethodologie noemen.



Ook de *state of the art* in algoritmen om kwadratische problemen op te lossen, schrijdt met rasse schreden voort; er is nu zeer betrouwbare software voor kwadratisch programmeren met kwadratische nevenvoorwaarden. Dit was vijf jaar geleden beslist nog niet beschikbaar. We hebben dus zowel theorie als nieuwe toepassingsrichtingen, nieuwe software. De combinatie van dit alles zal volgens mij een grote impact hebben.

Waarschijnlijk gaat deze methodologie, met al zijn toepassingen, een aparte tak worden, en dat zou een aderlating voor de OR kunnen betekenen. Aan de andere kant is OR niet een discipline die op zich van groot belang is. Het is een gebied van toepassingen. Een discipline die niet puur wetenschappelijk is meet zijn succes af aan de mate waarin het in andere disciplines wordt opgepikt en daar succes heeft. Het oorspronkelijke onderwerp sijpelt dus door van het ene gebied naar het andere en kan eventueel zelfs uitsterven. Dit is een teken van succes, niet van falen. Want gebieden die niet relevant zijn voor andere gebieden koken in hun eigen sop gaar, ze slinken, maar ver-

dwijnen doen ze niet. Toen Operations Research zich als discipline ontwikkelde was er nauwelijks Informatica, noch als afdeling, noch als discipline. OR verbond zich met discrete wiskunde en combinatorische optimalisatie en werd een hoofdonderwerp binnen de Informatica. Zaken als netwerken en convexiteit, die van meet af aan in de OR zaten, werden belangrijke onderwerpen in de Informatica. Mensen die op deze onderwerpen bijvoorbeeld een promotieonderzoek doen, vind je tegenwoordig niet meer in een OR- maar in een Informatica-afdeling. Nou, is dat een teken van falen van de OR, of juist van succes?

U bent ook verbonden aan het MINERVA Optimization Center in het Technion, en zelfs een van de directeuren.

Ja, in feite een van de oprichters. Laat ik uitleggen wat dit Center is. Het begon als een Optimalisatie Laboratorium. Het allereerste project was van een bedrijf dat een apparaat bouwde waarmee een cardioloog een beeld van een hart kan maken: tijdens een operatie raakt hij met een sonde het hart aan op verschillende plaatsen en het apparaat moet uit die gegevens de geometrie van het hart uittekenen. In drie maanden hebben we daar een algoritme voor gemaakt, waardoor dat bedrijf een enorme voorsprong in de markt kreeg; het was echt een *success story*. Wij ving er vijftienduizend dollar voor, een lachertje natuurlijk. We waren toen nog behoorlijk naïef. Later hebben we bijvoorbeeld projecten gedaan voor het belangrijkste telefoonbedrijf van Israël, onder andere door te adviseren waar je centrales moet plaatsen. Voor het elektriciteitsbedrijf hebben we een groot project gedaan om ze te helpen bij de beslissing in welke vorm van energie ze voor de komende vijftien jaar zouden moeten investeren.

We merkten spoedig dat we eigenlijk werkten als een commercieel bureau. Helaas is het Technion, en universitaire instellingen in het algemeen, niet de beste plek voor zulke werk-

zaamheden, vanwege de trage bestuursstructuur en een belangentegenstelling. Ik besloot me van de industriële projecten te distantiëren. Drie studenten van me zijn als *spin-off* bedrijfje geleidelijk losgekomen van het Technion. Nu hebben ze zo'n vijftig personen aan het werk met een vestiging in de Verenigde Staten, en zijn bij erg grote projecten betrokken. Wij besloten dus terug te keren naar werk op basis van wetenschappelijke fondsen, dus fundamenteel onderzoek, waar geen gevaar bestaat voor conflicterende belangen. Eind vijftiger jaren was er een Duits-Israëlische vergoedingsregeling voor de slachtoffers van de Holocaust, die deels bestond uit gelden voor maatschappelijk nut in bredere zin. Een deel van die gelden is besteed aan het oprichten van MINERVA centra in Israël: wetenschappelijke *centers of excellence* ter ondersteuning van Israëlische universiteiten. Wij hebben ons ingespannen om zo'n *MINERVA Center* te worden. Dat is gelukt en aldus werden wij het *MINERVA Optimization Center*. Dit houdt in dat we een donatie hebben gekregen van twee miljoen DM, waarvan we de rente ontvangen. Natuurlijk niet een bedrag om het centrum van te runnen, maar het is een leuke aanvulling als je andere toelagen hebt. Zo hebben we een flinke toelage gehad van de Europese Commissie in een Espritproject, in feite de eerste Esprit-toelage voor Israël. Dat was een uiterst interessant project over medische beeldverwerking. Het project liep ongeveer drieëneenhalf jaar en onze groep ontwikkelde een nieuw iteratief reconstructie-algoritme voor positron-emissietomografie.

De andere inkomsten die we in het lab hebben, komen van de *Science Foundation*, te vergelijken met STW hier, en een Duits-Israëlische *Science Foundation* waar groepen uit deze twee landen in participeren. Soortgelijke *Foundations* heeft Israël ook met de Verenigde Staten, en met Canada. Ik hoop dat er in de toekomst ook een Nederland-Israëlische *Science Foundation* zal ontstaan. Als er op ons gebied een initiatief vanuit Nederland zou

komen, zou ik alles doen wat in mijn vermogen ligt om dit van Israëlische kant te bevorderen, want het is werkelijk erg vruchtbaar voor beide kanten en in het bijzonder voor jonge mensen.

Nu bent u hier voor een jaar op sabbatical in Delft. Was er een speciale reden om voor Delft te kiezen? Iedere zes jaar mag je een jaar op *sabbatical*. Dit is mijn vierde keer. Er zijn verschillende speciale redenen. Ten eerste wilde ik in Europa zijn. Ik heb alles bij elkaar al zeven jaar in Noord-Amerika doorgebracht. De Europese manier van leven bevalt me beter dan de Amerikaanse. Ten tweede heb ik contacten met de mensen hier in Delft. De derde reden is dat mijn jongste zoon al vier jaar in Nederland woont. Hij is balletdanser, lid van het Nederlands Dans Theater. Ik wilde eens wat meer bij hem in de buurt zijn.

Verder kan ik hier goed werken aan robuuste optimalisering, aan bepaalde besturingsproblemen. Ik werk ook samen met Kees Roos aan tweede orde kegel optimalisering, en recentelijk doe ik toepassingen uit de meer klassieke Operationele Research, onderwerpen die toegepast worden in *Supply Chain Management*. Een keten van ondernemingen, waarin ieder zowel klant als leverancier is, en waarin beslissingen niet worden genomen door één onderneming, want die hebben gevolgen voor ieder ander. Als je dit uitbuit, dan kan ieder lid van de keten ook de totale situatie verbeteren. Dit is werkelijk een van de meest actuele onderwerpen in *Operational Management*.

FRITS VAN BECKUM is universitair docent bij de leerstoel Applied Analysis and Mathematical Physics van de Universiteit Twente en hoofdredacteur van ITW Nieuws, <f.p.h.vanbeckum@math.utwente.nl>.

Hans Melissen is universitair hoofddocent bij de leerstoel Optimization Technology van de Technische Universiteit Delft en redacteur van ITW Nieuws en het Nieuw Archief voor Wiskunde, <j.b.m.melissen@its.tudelft.nl>.

Dit interview is eerder in gewijzigde vorm verschenen in *ITW Nieuws*.

Hoe ontwikkel je toetsen en examens?

Over de rol van kwantitatieve methoden

PIET SANDERS

De Citogroep is een instituut dat meetinstrumenten, dat wil zeggen toetsen en examens, ontwikkelt voor het onderwijs. Toetsen en examens worden vooral gebruikt om de kennis en vaardigheden van leerlingen te beoordelen. Meestal gebeurt dat op momenten dat er een beslissing genomen moet worden over een volgende stap in het onderwijs of om een onderwijsperiode af te sluiten.

In deze bijdrage wordt eerst een korte beschrijving gegeven van de meetinstrumenten. Die beschrijving beoogt duidelijk te maken dat bij de ontwikkeling van de meetinstrumenten het gebruik van kwantitatieve methoden een essentieel onderdeel is.

Het Psychometrisch Onderzoek- en Kenniscentrum van de Citogroep - waaraan het tweede deel van deze bijdrage is gewijd - is verantwoordelijk voor de toepassing van bestaande en de ontwikkeling van nieuwe kwantitatieve methoden .



De Citogroep, voorheen Cito (Centraal Instituut voor Toetsontwikkeling), is in 1968 opgericht om toetsen en examens voor het onderwijs te ontwikkelen en is nu het grootste instituut op dit gebied in Nederland. Bij de Citogroep, gevestigd in Arnhem, werken zeshonderd medewerkers. Daarnaast zijn bij de ontwikkeling van toetsen en examens ongeveer tweeduizend leerkrachten betrokken. Wereldwijd gezien behoort de Citogroep tot de drie grootste bedrijven die toetsen en examens ontwikkelen. In 2001 bedroeg de jaaromzet 44 miljoen euro.

Eindtoets basisonderwijs

Toetsen en examens worden vooral gebruikt om kennis en vaardigheden van leerlingen te beoordelen. Meestal gebeurt dat op momenten dat er een beslissing genomen moet worden over een volgende stap in het onderwijs of om een onderwijsperiode af te sluiten. Een bekende toets uit het primair onderwijs is de Eindtoets basisonderwijs, de 'Cito-toets', aan het eind van groep 8. In 2000 nam 85 procent van de scholen, met ongeveer 160.000 leerlingen, deel aan de toets. De Eindtoets basisonderwijs is geen examen waar leerlingen voor kunnen slagen of zakken. De toets is bedoeld om het schoolsucces in het voortgezet onderwijs te voorspellen en uit onderzoek blijkt dat ze daar goed in slaagt. De toets bestaat uit de onderdelen rekenen, taal, informatieverwerking en wereldoriëntatie die elk 60 meerkeuzevragen omvatten. De vragen worden op optisch leesbare antwoordbladen beantwoord. Na verwerking ontvangen de scholen een schoolrapport en per leerling een leerlingrapport. Hierin staat onder andere de door de leerling behaalde standardscore. Dat is een lineaire transformatie van de totaalscore die afgebeeld wordt op een schaal met een gemiddelde van 535 en een standaarddeviatie van 10. De resultaten op de Eindtoets worden behalve door de Citogroep ook door anderen gebruikt om te rapporteren over hoe de school het doet. Aan dat laatste wordt de laatste jaren in de media steeds meer aandacht besteed.

Centrale Examens

In opdracht van het Ministerie van Onderwijs ontwikkelt de Citogroep jaarlijks enkele honderden examens. In 2001 namen er in het gehele voortgezet onderwijs (vbo, mavo, havo en vwo) ongeveer 220.000 leerlingen deel aan het Centraal Examen, eerste tijdvak, dat in mei afgenomen werd. Een examen kan bestaan uit alleen maar gesloten of meerkeuze vragen, alleen maar open vragen of een combinatie van gesloten en open vragen. De examens met alleen gesloten vragen scoort de Citogroep via de optisch leesbare antwoordbladen van de leerlingen. De examens die open vragen bevatten (naast eventuele gesloten vragen), worden door docenten aan de hand van een correctievoorschrift gescoord. Deze scoring wordt tweemaal uitgevoerd: eerst door de eigen docent en vervolgens door een docent van een andere school. Van deze examens sturen de docenten de scores van vijf kandidaten naar de Citogroep.

De scores, zowel die welke rechtstreeks door de Citogroep zijn gescoord als die van de docenten, worden voor twee doeleinden gebruikt. Ten eerste om na te gaan hoe op de individuele vragen en op het totale examen gescoord is. Ten tweede om de psychometrische kwaliteit van de examens te onderzoeken. Uit de kwantitatieve analyse van de afzonderlijke vragen kan blijken dat een bepaalde vraag veel te moeilijk was, met als gevolg dat deze vraag anders wordt gescoord of zelfs vervalft.

De verdeling van de totaalscores die de leerlingen op het examen behaald hebben wordt gebruikt bij het definitief vaststellen van de cesuur, dat wil zeggen de laagste totaalscore waaraan nog een voldoende cijfer toegekend wordt.

Peilingsonderzoek

In het Peilingsonderzoek worden meetinstrumenten ontwikkeld die bij leerlingen afgenomen worden om het onderwijssysteem als geheel te beoordelen, in plaats van de individuele leerling. Het onderzoek betreft in principe alle leerdomeinen





binnen het primair onderwijs en vindt plaats in groep 8. Na ongeveer 5 jaar wordt voor ieder leerdoel het onderzoek herhaald om na te gaan of de leerprestaties vooruit of achteruit gegaan zijn. Voor de uitvoering van het onderzoek is de Citogroep afhankelijk van de medewerking van de scholen. Vandaar dat wordt getracht de scholen zo weinig mogelijk te belasten. Dit wordt gedaan door voor een peilingsonderzoek slechts een steekproef van scholen te gebruiken en de leerlingen op die scholen niet alle items en/of taken te laten uitvoeren. De proefopzetten die bij de proefafnames gebruikt worden zijn dan ook structureel incomplete designs die met speciaal voor dit doel ontwikkelde software geanalyseerd kunnen worden. De peilingsresultaten worden verspreid onder landelijke en regionale onderwijsverzorgingsinstellingen, lerarenopleidingen, de onderwijsinspectie, schrijvers en uitgevers van lesmethoden en natuurlijk de scholen voor primair onderwijs.

Psychometrisch Onderzoek- en Kenniscentrum

Bij het Psychometrisch Onderzoek- en Kenniscentrum (POK) werken momenteel 24 medewerkers: 4 in het Kenniscentrum en 20 in het Psychometrisch Onderzoekcentrum.

De specialisten van het Kenniscentrum verzamelen informatie over toetsen, examineren en evalueren. Zij ontsluiten deze informatie voor zowel intern als extern gebruik (zie ook www.citogroep.nl/pok). Het Kenniscentrum voert ook de redactie van www.toetswijzer.nl, een site die voor alle betrokkenen bij het onderwijs informatie over toetsen, examineren en evalueren ontsluit.

Van de medewerkers van het Psychometrisch Onderzoekcentrum zijn er 16 als methodoloog werkzaam. Het Onderzoekcentrum werkt samen met verschillende universiteiten en wordt geadviseerd door een Wetenschappelijke Adviesraad. De twee belangrijkste taken van het Centrum zijn het geven van psychometrische ondersteuning aan

meer dan honderd projecten en het doen van onderzoek op het gebied van de psychometrie.

Psychometrische theorieën

Het construeren van meetinstrumenten zoals toetsen en examens heeft een kwalitatieve component en een kwantitatieve component. De kwalitatieve component betreft het ontwikkelen van de vragen of opdrachten waaruit de meetinstrumenten bestaan. De kwantitatieve component betreft het analyseren van de antwoorden die leerlingen op de vragen van de meetinstrumenten gegeven hebben. Het Psychometrisch Onderzoekcentrum gebruikt bestaande en ontwikkelt nieuwe psychometrische theorieën en statistische technieken om de toetsen en examens te onderzoeken. De belangrijkste psychometrische theorieën die gebruikt worden zijn de klassieke testtheorie, de generaliseerbaarheidstheorie en de itemresponsstheorie. Het boek *Psychometrie in de Praktijk* (red. Eggen & Sanders, 1993) geeft hiervan een beschrijving. Het boek kan men via de website downloaden.

In de *klassieke testtheorie* wordt aangenomen dat de toetsscore van een persoon de som is van een ware score, de variabele waar de interesse naar uitgaat maar die niet rechtstreeks meetbaar is, en een toevallige, niet systematische, meetfout. Het centrale begrip in de klassieke testtheorie is de betrouwbaarheid, die gedefinieerd is in een populatie personen. Betrouwbaarheid is een maat voor de precisie waarmee verschillen in ware scores tussen personen geschat kunnen worden met behulp van de verschillen in toetsscores tussen die personen. Een ander belangrijk begrip in de klassieke testtheorie is de standaardmeetfout van een toets. Deze is gelijk aan de standaardafwijking van de meetfouten in een populatie personen.

Het begrippenkader van de *generaliseerbaarheidstheorie* is voor een belangrijk deel ontleend aan de variantie-analyse. In deze theorie worden waarnemingen of metingen beschreven in termen van de condities waaronder zij geobserveerd worden.

De generaliseerbaarheidstheorie, die als een uitbreiding van de klassieke testtheorie gezien kan worden, wordt voornamelijk gebruikt bij het onderzoek naar meetinstrumenten waarbij de door leerlingen uitgevoerde opdrachten of beantwoorde vragen door twee of meer beoordelaars beoordeeld zijn.

De klassieke testtheorie heeft twee belangrijke tekortkomingen. De eerste tekortkoming is dat de maten die gebruikt worden zoals itemmoeilheidsgraad, itemdiscriminatie-index en betrouwbaarheid altijd naar een verdeling of een populatie verwijzen. De tweede tekortkoming is dat de geschatte ware score toetsspecifiek is.

In de *itemresponsstheorie* wordt aan die tekortkomingen van de klassieke testtheorie, populatieafhankelijkheid en toetsspecificiteit, tegemoet gekomen. Van de vele itemresponsmodellen die de laatste jaren ontwikkeld zijn wordt een aantal in *Psychometrie in de Praktijk* besproken. Ook een model dat door medewerkers van het Psychometrisch Onderzoekcentrum is ontwikkeld, het logistisch model voor dichotoom en polytoom gescoorde items (OPLM), wordt hierin behandeld. Door de aantrekkelijke eigenschappen van itemresponsmodellen in het algemeen en het éénparameter logistisch model in het bijzonder, worden deze modellen in nagenoeg alle projecten van de Citogroep toegepast. Ook is het onderzoek in belangrijke mate gericht op de verdere ontwikkeling en toepassing van itemresponsmodellen.

Recent onderzoek

Tot slot een kort overzicht van recent onderzoek.

- Van de leerlingen wordt bij meerkeuzevragen gewoonlijk gevraagd alleen het correcte alternatief aan te geven. Er is een itemresponsmodel ontwikkeld waarbij leerlingen hun voorkeur voor elk alternatief aangeven. Hierdoor wordt gebruik gemaakt van de (statistische) informatie die de keus voor deze alternatieven geeft. Door ook de informatie uit de antwoorden op de alternatieven te gebruiken, blijkt men veel minder items nodig te hebben

om dezelfde meetprecisie te realiseren dan bij de gebruikelijke meerkeuzevragen. Implementatie van het model vereist dat de leerlingen hun antwoorden via het beeldscherm invoeren.

- Het gebruik van de computer (en itemresponsstheorie) wordt ook bij adaptief toetsen verondersteld. De laatste jaren heeft de Citogroep veel geïnvesteerd in de ontwikkeling en productie van adaptieve toetsen. Een computergestuurde adaptieve toets (CAT) is een geautomatiseerde toetsprocedure waarmee de vaardigheid van een persoon gemeten wordt met een selectie uit een verzameling items. Deze selectie wordt tijdens de afname afgestemd op het vaardigheidsniveau van de leerling. Een belangrijk voordeel van adaptief toetsen, vergeleken met de gebruikelijke (lineaire) manier van toetsen, is dat met minder items, dus ook in minder tijd, betrouwbaarder metingen van een persoon verkregen kunnen worden. Een tweede voordeel is dat personen op hun eigen niveau gemeten worden en dus niet geconfronteerd worden met items die te moeilijk of te makkelijk zijn.

- Aan toetsen worden behalve psychometrische eisen veelal ook vakinhoudelijke en praktische eisen gesteld. Het samenstellen van een toets kan dan ook als een mathematisch programmeringsprobleem worden gezien. De doelfunctie is het maximaliseren van de betrouwbaarheid met de vakinhoudelijke en praktische eisen als restricties. Om items te kunnen selecteren dient men te kunnen beschikken over itembanken. Dat zijn grote verzamelingen items waarvan de psychometrische en andere kenmerken bekend zijn en in een database opgeslagen. De Citogroep beschikt voor veel vakken over itembanken. Er zijn algoritmen en software ontwikkeld om uit zulke itembanken de gewenste toetsen te genereren.

- Leerlingen van het voortgezet onderwijs die examen doen voor bijvoorbeeld het vak wiskunde mogen verwachten dat het examen wat betreft inhoud, maar zeker ook wat betreft moeilijkheid, vergelijkbaar is met het examen van het vorig jaar

en eventueel voorgaande jaren. De psychometrische techniek waarmee onderzocht kan worden of er verschillen bestaan tussen de moeilijkheidsgraad van examens wordt aangeduid met equivalenten of normhandhaving. Wat normhandhaving van examens zo complex maakt is dat verschillen in moeilijkheidsgraad van twee examens zowel toegeschreven kunnen worden aan of (de opgaven van) het examen, of de (vaardigheid van) de leerlingen of aan beide. In samenwerking met medewerkers van de vakgroep OMD van de Universiteit Twente zijn nieuwe normhandhavingstechnieken ontwikkeld die bij de examens van het voortgezet onderwijs geïmplementeerd zijn. Normhandhaving betekent dat leerlingen niet benadeeld worden wanneer zij geconfronteerd worden met een examen dat moeilijker is dan een examen van een voorafgaand jaar.

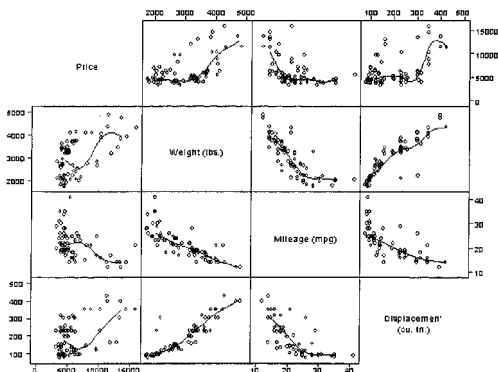
- Bij de Eindtoets basisonderwijs en het Leerlingvolgsysteem worden de resultaten die leerlingen op toetsen behalen ook gebruikt bij het beoordelen, en dus vergelijken, van resultaten van scholen. Een

probleem bij het vergelijken van scholen is echter dat zij verschillende achtergronden en omstandigheden kunnen hebben. Voor een eerlijke vergelijking dient voor die verschillen gecorrigeerd te worden. In het project Schoolzelfevaluatie is met behulp van itemresponstheorie en multilevelmodellen een methodologie ontwikkeld die het mogelijk maakt scholen te vergelijken op basis van hun toegevoegde waarde. Dit vereist wel dat de deelnemende scholen niet alleen hun toetsresultaten maar óók de achtergrondgegevens van de leerlingen en de school beschikbaar stellen. Voor het doelmatig archiveren van deze gegevens wordt momenteel een datawarehouse ingericht.

Zie ook: Eggen, T.J.H.M. en P.F. Sanders (1993), *Psychometrie in de praktijk*. Arnhem, Cito. (downloaden is mogelijk via de website www.citogroep.nl/pok).

P.F. Sanders is hoofd van het Psychometrisch Onderzoek- en Kenniscentrum van de Citogroep in Arnhem.
<Piet.Sanders@citogroep.nl>.

STATA®
Release 7
voor Windows 95/98/NT en 3.1, Macintosh en UNIX



SMIT CONSULT
Adviesbureau voor Gegevensanalyse

Postbus 220, 5150 AE Drunen
telefoon 0416 - 378 125, fax 0416 - 378 385
e-mail: info@smitconsult.nl
URL: www.smitconsult.nl

Stata is een hulpmiddel voor verwerking en analyse van gegevens, gebruikmakend van statistische methoden. Het programma is compleet en wordt gebruikt door onderzoekers op alle gebieden. Rodney Hayward van de University of Michigan's Schools of Medicine & Public Health verklaarde onlangs: "I've used a lot of statistical packages over the years, but I find that I'm using **Stata** 95% of the time now. It is wonderful! Its speed and power are much impressed, but its simplicity for beginners is perhaps one of its best features."

STATA®

Nieuw en uitgebreid in **Stata 7.0**: graphics, gebruikersinterface (via de *Stata Markup* en *Control Language*), ondersteuning van namen van variabelen tot 32 karakters, survival-analyse (frailty / heterogene residuën), paneldata-analyse (Arellano-Bond schatters), clusteranalyse, en berekening van marginale effecten van vrijwel alle schatters.

Stata is een kwaliteitsprogramma. Het is goed gedocumenteerd, eenvoudig in gebruik, zeer snel en verkrijgbaar tegen een redelijke prijs. **Stata** is één programma; het kent geen modules.



IBM 701 (1953)

TAAKPLANNING ALS GAT IN DE MARKT

Kansen voor nieuwe technologieën in it-systemen

MARTIN HEALEY

De opmars van internet en e-handel gaat niet zo snel als velen hadden verwacht, maar de meeste organisaties maken er inmiddels wel gebruik van. Dit heeft het gebruik van externe communicatienetwerken, en de interactie met gebruikers via het Web zeer doen toenemen. In de meeste gevallen is daarmee een nieuw terrein betreden. De nieuwe eisen kunnen onafhankelijke systemen zijn of nieuwe front-ends, geïntegreerd met oudere systemen. In beide gevallen is het onvermijdelijk nieuwe technologie toe te passen. Soms kan dat nieuwe software op bestaande platformen zijn, maar waarschijnlijk gaat het om nieuwe hardware en besturingssystemen. [...]

De laatste jaren is veel geïnvesteerd in beheersoftware. Het belang van *overall* beheer wordt geïllustreerd door het feit dat IBM Tivoli kocht om een leidende positie te verwerven. Maar naast systeembeheer benadrukt de integratie van verschillende subsystemen in een gemeenschappelijk applicatiesysteem de noodzaak om vooruit te kijken en om de automatisering van functies zoals directory-diensten, helpdesk, dataoverdracht, enzovoort, in beschouwing te nemen. Werkstroombeheer is vele jaren van belang geweest, hoewel er mijns inziens nog weinig aandacht is besteed aan automatiseringtools. Er is zelfs meer behoefte aan het beheer van batchprocessen. Aangezien mainframesystemen veel ouder zijn dan Unix- of pc-systemen, kennen ze een lange geschiedenis van batchverwerking. Dientengevolge

zijn planningtools, met name Jess op IBM-mainframes erg krachtig en veel gebruikt. Unix en VMS zijn daarentegen in essentie 'multi-user' interactieve systemen met tamelijk zwakke batchmogelijkheden en dus relatief simpele (en eerlijk gezegd inadequate) taakplanningtools (*job-scheduling tools*). Dat moet veranderen met al die nieuwe geïntegreerde applicaties, omdat de planning- en beheertools over meerdere platformen heen moeten kunnen werken. Het is niet voldoende om alleen het mainframe-deel goed aan te pakken en de andere systemen niet. Een oplossing ligt in de handen van IBM door Jess- en Tivoli-producten te versterken en ze geschikt te maken voor meer dan z/OS. Een andere benadering is het gebruik van onafhankelijke leveranciers zoals Cybernation, die gewend zijn verder te kijken dan het mainframe en die ondersteuning op meerdere platformen bieden. In ieder geval is er iets apart nodig voor die gebruikers zonder mainframes! Het blijft verbazingwekkend hoe weinig aandacht is geschonken aan batchverwerking en taakplanning in het bijzonder, buiten de mainframe arena. De e-handel initiatieven zouden een stimulans moeten zijn om dat recht te zetten.

MARTIN HEALEY, pionier als het gaat om het ontwikkelen van op Intel gebaseerde computers en c/s-architectuur. Hij is directeur van een aantal it-bedrijven en hoogleraar aan de Universiteit van Wales.

Dit artikel is een gedeelte van een column van Martin Healey in *Computable*, 7 december 2001, nr 49, pag 27.

STATISTISCH CONSULTANT M/V

(32 TOT 40 UUR PER WEEK)

TPG Post is de grootste leverancier van postdiensten in Nederland én marktleider in internationale zakelijke post. Daarnaast is TPG Post zeer actief op het gebied van data- en documentdiensten, direct mail en e-commerce. De afdeling Kwantitatieve Ondersteuning speelt een essentiële adviserende rol bij het handhaven en uitbouwen van de strategische positie van TPG Post.

Als je op zoek bent naar

- verantwoordelijkheid
- een intellectueel uitdagende functie met veel knowhow en diversiteit
- een informele werksfeer op een relatief kleine afdeling (15 medewerkers)

Dan bieden wij je een interessante functie waarin je

- in teamverband werkt voor verschillende Business Units van TPG Post
- verantwoordelijk bent voor resultaatgerichte, statistische onderzoeks- en adviesprojecten
- adviseert over uiteenlopende statistische vraagstukken
- ruimte krijgt voor innovatie en zelfstandigheid
- uitstekende doorgroei- en loopbaanmogelijkheden hebt

Wij vragen het volgende van je

- minimaal drie jaar werkervaring op het gebied van steekproeftheorie, statistische kwaliteitsbeheersing en voorspelmethodeken
- managementpotentie

Je reactie? Voor meer informatie kun je bellen met de heer H.J. van der Brug, manager Kwantitatieve Ondersteuning, telefoon (070) 334 73 71. Je brief met c.v. kun je sturen aan TPG Post, t.a.v. mevrouw J.M.A. Braat-Toor, kamer AA302, Postbus 30250, 2500 GG Den Haag. Of per e-mail: J.M.A.Braat-Toor@tpgpost.nl



A G E N D A

Zie voor meer nieuws, conferenties, studiedagen, mededelingen van de VVS en cursussen de site van de VVS <<http://www.vvs-or.nl>>.

7-12 juli 2002

De International Association for Statistical Education (IASE) houdt haar vierjaarlijks congres in Durban, Zuid-Afrika. Het centrale thema is 'Developing a statistically literate society'. Informatie over dit congres is te vinden op: <www.beeri.org.il/icots6/>.

8-12 juli 2002

The **17th International Workshop on Statistical Modelling (IWSM)**, which will take place in Chania (Crete, Greece) concentrates on the various aspects of statistical modelling, including theoretical developments, applications and computational methods. More info can be found on <www.unl.ac.uk/iwsm/>.

24-28 augustus 2002

The **Conference for Computational Statistics, COMPSTAT 2002**, will take place from August 24th to August 28th 2002 at Humboldt-Universität zu Berlin, Germany. The conference is organized by the Institute for Statistics and Econometrics of the School for Business and Economics. Topics: Computational Finance, Statistics of E-commerce, Mining very large statistical databases, Complex Datastructures in the Biosciences, Netbased Statistics. These topics will all include methodological applications, innovative software and mathematical developments. Information is available at <www.compstat2002.de>.

25-28 augustus 2002

You are cordially invited to submit abstracts for the **International Conference on Improving**

Surveys (ICIS 2002), which will take place in Copenhagen, Denmark. Main themes are: Impact of New Technology; Quality of Surveys; Comparability of International Assessments; Comparability of Survey and Register Statistics. An abstract of no more than 500 words should be sent by e-mail to <ICIS@sfi.dk>. Information is available at <www.sfi.dk>.

9-20 september 2002

Workshop On High-Dimensional Data, Lorentz Center, Leiden University. Today's trend is to gather an enormous amount of information at a large number of instances, resulting in datasets with dimensions in the thousands or billions (think of hyperspectral imagery, financial data, or DNA microarrays). This workshop focuses on the new statistical theory required for analyzing high-dimensional data, and on its biomedical applications. Visit: <www.medstat.medfac.leidenuniv.nl/MS/HD>.

18-19 oktober 2002

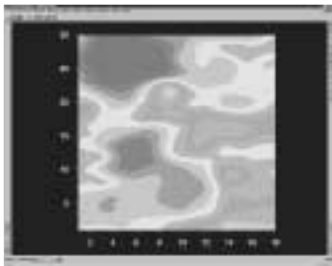
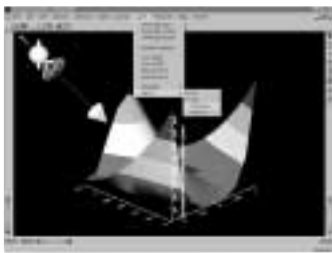
De jaarlijkse bijeenkomst van de BSS (*Belgian Statistical Society*) vindt plaats in congrescentrum Rolduc, Kerkrade. Voor meer informatie en registratie kijk op <www.luc.ac.be/censtat/BSS2002/>.

Statistische Dag 2002 De Statistische Dag heeft plaatsgevonden op 25 maart 2002. Het blauwe boekje (in pdf formaat) is online beschikbaar <www.vvs-or.nl>. De meeste secties hebben tijdens de Statistische Dag hun jaarlijkse ledenvergadering gehouden en organiseerden verder een eigen programma met sterke sprekers uit binnen- en buitenland.

De **Biometry Award 2002** is tijdens de Statistische Dag op 25 maart 2002 uitgereikt aan Hans Jansen voor het artikel 'Constructing dense genetic linkage maps'. *Theoretical and Applied Genetics*, 102, 1113-1122. Er was dit jaar een record aantal van 27 artikelen genomineerd. Meer informatie op <www.plant.wageningen-ur.nl/other/bms-aned/>.

PROBABLY THE BEST STATISTICS PACKAGE FOR TEACHING STATISTICS TODAY!

6.25	44	45
7.5	Identifier	
7.5	42	block
	FACTOR [mod	
	9 READ N; frep	



The new **GenStat® for Windows® 5th Edition**, available now, boasts a long list of new features and improvements.

Included are entirely new graphics, improved interface and spreadsheet, and dozens of new data management tools and other facilities to make it easier to manipulate your data and share it with others.

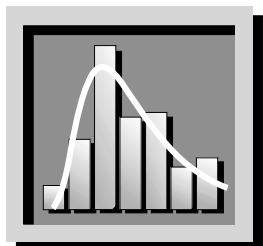
With these new features GenStat is more user-friendly than ever. To make it even easier to use we have included a multi-media tutorial which gets your students up and running in no time, enabling you to concentrate on teaching statistics, not the package!

GenStat's reputation for the quality and accuracy of its statistical functionality is unparalleled and it contains the broadest range of statistical tools, all available in the standard package with no additional modules to install and pay for. This means that GenStat is genuinely suitable for the beginner, while providing the power and flexibility required by the expert.

A student version of the package is available to support your teaching. Attractive site license and other arrangements are also available to qualifying academic organisations. Contact us for more details - and give your students and yourself the benefit of the statistical power of GenStat!

For more information please contact:

Cosinus Computing, Postbus 52, 5600 AB Eindhoven, The Netherlands
<http://www.cosinus.nl>
info@cosinus.nl



GenStat®

With GenStat you know you can!

BAKKEN RESEARCH CENTER

Medtronic is the world leader in medical technology.

Growing from our deep roots in the treatment of heart disease, we provide a wide range of products and therapies that restore life to patients all over the world.

Physicians turn to us to help them solve the most challenging, life-limiting medical problems. We will continue to partner with them as we invest in new technologies and passionately pursue our mission to restore health, extend life and alleviate pain.

Medtronic set up Bakken Research Center (BRC) in Maastricht in order to strengthen the company's clinical and technical research and development activities in Europe, the Middle East and Africa.

Working at Medtronic means working with precision, passion and a pioneering spirit. We therefore provide you with ample space in which to develop your talents. Enjoy excellent working conditions, career prospects and a high-tech environment which is known for its flexible and informal atmosphere.

The outcome of your contribution to our medical research will be far from static



BIOSTATISTICIAN

You will operate within the Arrhythmia Management (AM) department of the Bakken Research Center in Maastricht. The core activity of the AM group is the appropriate detection and treatment of arrhythmia - either a too slow or too fast heart rate - referred to as bradycardia or tachycardia, respectively. The AM team focuses on R&D, the clinical evaluation of new products and the general outcome evaluation of device-related therapies.

As a Biostatistician you will provide statistical expertise for the design and conduct of clinical studies and other research projects. You give advise about the methodological correctness of the study design, you assess the required sample sizes, you retrieve data from the clinical database and prepare needed summary reports. Using your excellent analytical skills and adequate understanding of medical and technical aspects, you perform statistical analyses of clinical study data. Through active cooperation with physicians and medical researchers you will take a leading role in the publication of study results. In general, you will contribute to the overall scientific quality of activities in the AM department by the application of state-of-the-art statistical techniques.

You are a self-starter and a generalist. A team player with excellent interpersonal skills. You easily adapt to a dynamic working environment and your colleagues speak highly of your statistical and analytical skills. A perfect command of the English language is an absolute must. Our ideal candidate combines a university degree in Mathematics with specialization in statistics or epidemiology. An in-depth knowledge of database management and a vivid interest in medical science are highly beneficial.

For further information about this long-term career-opportunity, please contact:

Mr. Jaak Minten, Manager Arrhythmia Management department at tel.nr. 043-3566824.

Are you interested and do you have the required skills?

Then we invite you to send your application and CV **within 14 days** to the attention of: Mr. M.A.J. Ebben, Manager Human Resources and Services.

Applications can also be submitted on-line. Simply surf to: www.jobs.medtronic.nl
You can find the latest information, fill out your application form and send your CV with a single mouseclick.

Bakken Research Center B.V.
Endepolsdomein 5
NL-6229 GW Maastricht



Careers with a passion for life



Medtronic
When Life Depends on Medical Technology