

STATOR

periodiek van de VWS jaargang 8 nummer 4, december 2007

Het economisch beslissingsprobleem
van de bescherming tegen overstromingen

Multivariate meta-analyse: wat te doen
bij meerdere uitkomsten per studie?

Robuuste optimalisatie
van retail-assortimenten

Statistici, kansrekenaars
en andere wiskundigen gelauwerd

Quantum computation en optimalisering

Het refereeproces

Nederlands-Vlaamse Vereniging voor
Ordinatie en Classificatie (VOC)

Opnieuw Lucia de B.

STATOR

Jaargang 8, nummer 4, december 2007

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

Redactie

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

Kopij en reacties richten aan

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

Bestuur van de VVS

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

Leden- en abonnementenadministratie van de VVS

VVS, Postbus 2095, 2990 DB Barendrecht, telefoon 0180 - 623796, fax 0180 - 623670, e-mail <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

Monique van Hootegem, Moeflonstraat 5, 6531 JS Nijmegen, e-mail <hootegem@xs4all.nl>. STATOR verschijnt in april, juni, september en december.

Ontwerp en opmaak

Pharos / M. van Hootegem, Nijmegen

Uitgever

© Vereniging voor Statistiek en Operationele Research
ISSN 1567-3383

Inhoud

- 3** Metastatistiek
- 4** Opnieuw bekeken. Het economisch beslissingsprobleem van de bescherming tegen overstromingen
Carel J.J. Eijgenraam
- 8** Multivariate meta-analyse: wat te doen bij meerdere uitkomsten per studie?
Lidia Arends
- 13** Robuuste optimalisatie van retail-assortimenten
Robert Rooderkerk
- 18** Statistici, kansrekenaars en andere wiskundigen gelauwerd - *column*
Fred Steutel
- 20** Quantum computation en optimalisering
David Bulger
- 25** Het refereeproces - *column*
Onno Boxma
- 28** Een revolutionaire methode om van een zelfselecte steekproef een goede steekproef te maken? Open brief aan Wiebe Draijer et al
- 29** Nederlands-Vlaamse Vereniging voor Ordinatie en Classificatie (VOC)
- 30** VVS-nieuws
- 31** Opnieuw Lucia de B.
- 32** Agenda



Metastatistiek

Meta- betekent volgens Van Dale: ‘handelend over het in het grondwoord genoemde’. Zo betekent metataal ‘taal waarin over taal gesproken wordt’ en metamathematica (niet in Van Dale): ‘het bestuderen van wiskunde met wiskunde’. Deze aflevering van *STATOR* heeft iets van meta-statistiek: statistiek over statistiek en statistici over statistici. Hierbij staat ‘statistiek’ steeds voor kansrekening, statistiek en operations research: Kastor.

Eén van de artikelen heeft niet alleen het woord meta-analyse in de titel, het artikel gaat ook inderdaad over statistiek-statistiek: hoe kunnen bestaande statistische studies met statistiek worden verbeterd? Het actuele artikel over de hoogte van onze dijken bericht onder meer over de vraag hoe de statistiek vijftig jaar geleden beter had gekund – gelukkig kan het nog steeds beter. In *Quantum Computation* daarentegen wordt juist geschetst hoe het in de toekomst beter zou kunnen gaan in de Operations Research, speciaal in de optimalisering. Er zijn weliswaar nog geen bruikbare kwantumcomputers (ook niet in Van Dale), maar, als ze beschikbaar zouden zijn, zouden we veel grotere problemen kunnen oplossen en niet-zo-grote problemen veel sneller.

In de columns wordt nooit statistiek bedreven; ze gaan óver statistiek: metastatistiek. Eén column gaat over het refereeproces: wat vindt de ene beoefenaar van het vak van het werk van een andere, en hoe gaat dat in zijn werk: metameta-statistiek. De andere column gaat ook niet over de statistiek zelf, maar over haar beoefenaren, statistici (en andere wiskundigen) gelauwerd, vergeten, jong gestorven of juist heel oud geworden.

Als sluitstuk van al dit statistische zelfonderzoek komt de affaire Lucia de B. weer aan de orde. De dag- en weekbladen hebben uitvoerig aandacht geschonken aan de omstreden procesgang en het herzieningsverzoek. Zelfs Fokke en Sukke hebben commentaar op de niet zo verheffende, statistische kant van deze kwestie. Daar willen wij niet achterblijven!

Dit alles is natuurlijk ook een taak van *STATOR*: niet zelf statistiek, kansrekening of OR beoefenen, maar het vak laten zien en aan ‘de man’ brengen, laten kijken in de keuken van de statistiek.

Veel leesplezier!
De redactie



OPNIEUW BEKEKEN

Het economisch beslissingsprobleem van de bescherming tegen overstromingen

Foto: Pieter Bosch

CAREL J.J. EIJGENRAAM

Tijdens voorlichtingsbijeenkomsten in 1965 voor de studie econometrie aan de Universiteit van Amsterdam werden voorbeelden genoemd om duidelijk te maken waar besliskunde over gaat. Een van die voorbeelden was het probleem hoe hoog je het beste een dijk kan maken. Dit probleem was eind 1953 opgelost door prof. dr. D. van Dantzig (Van Dantzig, 1956 en 1960), een man die mede daarom wel de vader van de Nederlandse besliskunde wordt genoemd. Terecht wijdde *STATOR* dan ook uitgebreid aandacht aan de vijftigste verjaardag van deze historische prestatie (Oud en Stermerdink, 2003). De veiligheidsnorm die in de Deltawet van 1958 aan de grootste dijkkringen in Nederland werd toegekend, namelijk een maximale overstromingskans van 1/10.000 per jaar, was mede gebaseerd op zijn berekening, ook al liet Van Dantzig er in zijn bijdrage aan het rapport van de Deltacommissie geen twijfel over bestaan dat hij deze norm onverantwoord laag vond. De veiligheidsnorm is ongewijzigd overgenomen in de huidige Wet op de waterkering en ook de vei-

ligheidsnormen voor andere dijkkringen zijn er op een wat losse manier van afgeleid.

Timing

In 2003 ben ik gevraagd om een kosten-batenanalyse te maken van het project Ruimte voor de Rivier (Eijgenraam, 2005) en ik herinnerde mij dat er een theoretische oplossing moest bestaan voor het probleem van de optimale dijkhoogte. Bestudering van het *Econometrica*-artikel uit 1956 leerde mij echter dat Van Dantzigs oplossing onvolledig was. Hij had zich niet afgevraagd op welk moment een dijkverhoging het beste kon plaatsvinden. Dat was in 1953 natuurlijk ook niet relevant, want het was toen duidelijk dat er een groot tekort aan veiligheid was. Maar of er in 2003 uit economisch oogpunt inderdaad voldoende aanleiding was voor verkleining van de feitelijke overstromingskansen in het rivierengebied, met als gevolg een investering van ruim 2 miljard

euro, was vooraf niet zo duidelijk. In mijn poging om dit *timing*-probleem op te lossen was het eerste dat ik ontdekte, dat ook Van Dantzigs oplossing van het hoeveelheidprobleem (de optimale dijkhoogte) niet juist was. Dit laatste is gemakkelijk in te zien.

Een voorraadprobleem

Het probleem van de optimale dijkhoogte kan worden gezien als een voorraadprobleem. We hebben ons met de dijk een zekere voorraad bescherming verschaft. Deze voorraad ‘verdamp’t echter in de loop der tijd en wel om twee redenen. Ten eerste neemt de overstromingskans toe door bijvoorbeeld bodemdaling, zeespiegelstijging en klimaatverandering waardoor de piekafvoeren van de rivieren groter worden. De tweede reden is dat er groei optreedt van bevolking en kapitaalgoederen waardoor de schade bij overstromen stijgt. Om beide redenen stijgt de verwachte schade per jaar (kans maal gevolg). Aanvulling van de voorraad bescherming, dat wil zeggen: verhoging van de dijk, brengt naast variabele ook hoge vaste kosten met zich mee. Vandaar dat we de voorraad niet continu aanvullen, maar slechts met flinke tussenpozen een heleboel tegelijk. De optimale bestelgrootte is hier de optimale dijkverhoging.

De sleutel tot de oplossing is een moment te vinden waarop een exacte herhaling optreedt van een vorige situatie. Op dat moment is dezelfde beslissing optimaal als die op het overeenkomstige vorige moment. Afgezien van een aanloopperiode kunnen we het beslissingsprobleem op lange termijn nu oplossen als het probleem voor

de eerste periode gegeven het feit dat we deze beslissing daarna ongewijzigd kunnen herhalen voor volgende perioden.

Van Dantzigs oplossing

De oplossing die Van Dantzig aandroeg volgt in principe deze aanpak en is gebaseerd op het idee dat de overstromingskans na iedere dijkverhoging weer gelijk moet zijn aan die direct na de vorige verhoging. Maar dan is de situatie niet gelijk, want de potentiële schade bij overstromen en dus de verwachte schade is inmiddels toegenomen door groei van de bevolking en de kapitaalgoederenvoorraad. Gelijk houden van overstromingskansen in een systeem met economische groei kan dus nooit de optimale oplossing zijn. Met dit laatste zeg ik niets nieuws. In waterstaatskringen wordt al lang gezegd dat de veiligheidsnormen zouden moeten worden aangescherpt omdat de schade bij overstromen sinds 1953 is toegenomen. Maar ook in die kringen heeft men niet beseft dat dit pleidooi voor herziening van overstromingskansen tevens inhoudt dat een oplossing die zegt dat overstromingskansen constant gehouden moeten worden, niet juist kan zijn.

De correcte redenering

Uitgangspunten

Omdat absolute veiligheid tegen overstromen in Nederland onmogelijk is te bereiken, kan veiligheid eigenlijk alleen gedefinieerd worden met behulp van zijn tegendeel: veiligheid is de accep-

tabel geachte kans op schade. Er is dus altijd restschade. Maar er is ook altijd minder kans op schade mogelijk, zij het tegen steeds hogere kosten. Kiezen is dus nodig en mogelijk. Dit maakt de keuze van de veiligheidsnorm tot een economisch probleem. Beslissen over een veiligheidsnorm is echter een politieke zaak, omdat daarbij vele waarderingen een rol spelen van zaken die geen marktprijs hebben, zoals mensenlevens. Niettemin ga ik er verder van uit dat we de schade bij een overstroming kunnen uitdrukken in een bedrag aan geld.

De centrale vraag is dan de volgende:

Bij welke omvang van investeren, bijvoorbeeld in dijkverhoging, gaan de maatschappelijke kosten van de investering de maatschappelijke baten van de daling van de verwachte schade overtreffen? Bij die omvang gaan we niet verder met investeren. In feite kiezen we dan op basis van minimalisatie van de som van de restschade en de investeringskosten, omdat dit het hoogste restinkomen overlaat voor andere uitgaven.

Wat de oplossing van dit probleem lastig maakt, is dat er voortdurend de bovengenoemde veranderingen optreden. Stijging van de overstromingskans en groei van de schade bij overstromen versterken elkaar in de zin dat beide de verwachte schade per jaar doen stijgen (kans op overstromen per jaar \times schade bij overstromen).¹ Het is duidelijk dat deze veranderingen leiden tot meer dan één beslissing over dijkverhoging in de loop der tijd. En tevens, dat de huidige beslissing invloed heeft op het tijdstip en de omvang van de volgende beslissing.

Naast de snelheid van de groei van de verwachte schade zijn er twee andere zaken van grote invloed op de beslissing:

- disconteringsvoet ($\delta > 0$), deze werkt in de richting van het zoveel mogelijk uitstellen van investeringsuitgaven;

- vaste investeringskosten, dit werkt in de richting van zo veel mogelijk tegelijk doen. Wegens de vaste kosten doen we dijkverhogingen in sprongen, met als gevolg dat het beschermingsniveau niet constant is.

Het bovenstaande maakt duidelijk dat we concreet twee vragen moeten beantwoorden: *wanneer* en *hoeveel* investeren?

Oplossing

Als we uitgaan van een constante vorm van de exponentiële kansverdeling van waterstanden, een constante groeivoet van de schade en constante investeringskosten die alleen afhangen van de omvang van de verhoging, dan ligt het voor de hand dat de oplossing in de herhaalperiodes ook constant is (zie voor alle bewijzen Eijgenraam 2005 of 2006). De variabele waar het om draait ligt eigenlijk al besloten in de formulering van de centrale vraag, namelijk de verwachte schade. Het is optimaal dat de verwachte schade S_t zich beweegt in een vast interval (S^+, s^-). Daarbij is s^- de ondergrens van het 'beschermingsinterval', dat wil zeggen het punt waar de verwachte schade het grootste is en we dus overgaan tot dijkverhoging. Als deze grens overschreden gaat worden, moet een vaste dijkverhoging gereedkomen zodat het hoge beschermingsniveau S^+ wordt bereikt met een lage verwachte schade. Door de vaste omvang van de dijkverhoging en de vaste groeivoet van de schade ligt ook de periode tussen twee investeringsmomenten vast. Met deze oplossing van het periodieke deel van het probleem is de hoeveelheidvraag beantwoord.

Het tijdstip waarop begonnen wordt met investeren, hangt mede af van de begintoestand op tijdstip o . Hier ligt namelijk een absolute grens: we kunnen niet meer in het verleden investeren. Er zijn nu twee mogelijkheden, de feitelijk verwachte schade S_o ligt aan de goede kant van s^- , of aan de verkeerde kant. In het eerste geval kunnen

we wachten tot het moment waarop de bestelgrens wordt bereikt. Op het moment van investeren geldt de volgende rendementsvoorwaarde:

$$s^- - S^+ = \delta I(\hat{u})$$

waarin, naast de al eerder in de tekst gedefinieerde variabelen, I staat voor de investeringskosten en \hat{u} voor de omvang van de optimale dijkverhoging. Dit is het bekende criterium voor het eerste jaarsrendement (FYRR) en het is dit criterium dat in Van Dantzig (1956) ontbreekt. In tegenstelling tot wat Van Dantzig dacht (zie daarvoor de grafiek in Oud en Stemerink, 2003), spelen de totale investeringskosten, en daarmee de vaste kosten bij investeren, dus wel degelijk een rol in de optimale oplossing. Dat komt omdat het besluit over de vaste kosten pas genomen wordt op het moment dat besloten wordt om te investeren. Dit is een welbekend resultaat uit formules over de optimale bestelgrootte.

De andere mogelijkheid is dat de verwachte schade S_0 al wel groter is dan de grenswaarde s^- . Er moet nu onmiddellijk een veiligheidsachterstand worden ingehaald met eenmalig een extra-grote investering. Dit is het enige geval dat Van Dantzig heeft behandeld. Omdat het besluit om op tijdstip nul te investeren in dit geval al genomen is, doen de vaste kosten nu niet meer ter zake en is de hoeveelheid alleen nog afhankelijk van de marginale kosten. Maar de correcte formule voor de hoeveelheid wijkt af van die van Van Dantzig. Bij Van Dantzig wordt van de disconteringsvoet eerst de economische groei afgetrokken. Van Dantzig gebruikte 4% voor de disconteringsvoet en 2,5% voor de economische groei. Als we echter achteraf de gerealiseerde economische groei uitrekenen over de laatste 50 jaar, dan komen we praktisch uit op 4%. Als we dat cijfer zouden gebruiken, dan zien we dat Van Dantzig's vergelijking een oneindig hoge

dijk oplevert en bij een groeivoet iets hoger dan 4% geen oplossing meer heeft. Er is echter geen enkele reden waarom de groei binnen een afzonderlijke dijkkring kleiner zou zijn dan de disconteringsvoet. Toch leidt de aanleg van Almere niet tot een onoplosbaar veiligheidsprobleem. In de juiste formule vindt deze aftrek niet plaats en treedt dit probleem dus niet op.²

Beleidsimplicaties

Het op lange termijn constant houden van de verwachte schade impliceert in een omgeving met economische groei dat de optimale overstromingskansen even hard dalen als de economische groei. Wettelijk liggen de normen voor de overstromingskansen echter al sinds 1958 vast en er is geen goede reden om aan te nemen dat ze zelfs toen op een economisch gezien goede hoogte zijn vastgesteld. Er is sinds 1953 echter zoveel veranderd dat een herberekening van Van Dantzig's oorspronkelijke berekening geen praktische betekenis meer heeft. Een duidelijk beleidsadvies is om het economisch optimale veiligheidsniveau met recente cijfers te gaan uitrekenen voor alle dijkeringen in Nederland en aan de hand daarvan een maatschappelijk debat te voeren over nieuwe wettelijke veiligheidsnormen.

Dit is wat nu aan de gang is in het project Waterveiligheid 21e eeuw, dat de staatssecretaris van Verkeer en Waterstaat eind 2005 is gestart. In het kader van dit project is inmiddels een studie gestart om de in dit artikel beschreven theorie uit te breiden naar meer ingewikkelde gevallen. Enige vragen over en aanzetten tot praktische toepassing van deze theorie staan in een CPB-notitie (2005). Daarin staan ook de uitkomsten vermeld voor 20 dijkeringen in het rivierengebied, gebaseerd op het onderzoek beschreven in Eijgenraam (2005).

NOTEN

1. De volgende cijfers geven een idee over de orde van grootte van de veranderingen in de verwachte schade. De schade bij overstromen stijgt ongeveer even snel als de reële groei van het nationaal inkomen, zeg op lange termijn met een ongeveer stabiele bevolking tegen de 2% per jaar. In het gebied van de Bovenrivieren, ruwweg oostelijk Nederland, heeft de stijging van de overstromingskans een effect van 1% per jaar op de stijging van de verwachte schade. In het gebied van de Benedenrivieren, ruwweg het westelijk deel van Nederland, stijgt dit effect vanaf het midden van het land tot 2 en plaatselijk tot 5% per jaar op de lijn Krimpen-Moerdijk. Zonder ingrijpen stijgt de verwachte schade dus in totaal met ongeveer 3% per jaar langs de Bovenrivieren en met 4 tot 7% per jaar langs de Benedenrivieren.
2. Als we rekening houden met een realistische investeringsfunctie waarin de kosten van een dijkverhoging met een vaste omvang afhangen van de hoogte van de aanwezige dijk en dus bij iedere verhoging toenemen, dan verschijnt er wel weer een ondergrens voor de disconteringsvoet die afhangt van de groeivoet van de investeringskosten. Dit betekent dat er in werkelijkheid ook een in de loop der tijd stijgende ondergrens is voor de verwachte schade, als we tenminste willen voorkomen dat de kosten gaan ontploffen.

LITERATUUR

CPB (2005) *Urgentie van acties omtrent veiligheid tegen overstromen*. CPB Notitie, 30 juni 2005.

Dantzig, D. van (1956), Economic decision problems for flood prevention. *Econometrica* 24, pp 276-287.

Dantzig, D. van (1960), *Het economisch beslissingsprobleem betreffende de veiligheid van Nederland tegen stormvloeden*. Rapport van de Delta Commissie, Bijdrage II.2, pp 59-110.

Eijgenraam, C.J.J. (2005), *Veiligheid tegen overstromen, Kosten-batenanalyse voor Ruimte voor de Rivier deel 1*. CPB Document 82, CPB, Den Haag.

Eijgenraam, C.J.J. (2006), *Optimal safety standards for dike-ring areas*. CPB Discussion Paper 62, CPB, Den Haag.

Oud, H. en G. Stemerding (2003), Interview met prof. drs. J. Kriens; oud-medewerker van Van Dantzig en mede-oprichter van de beslissonderzoeksgroep in Nederland.

STATOR 4-1, pp 4-9.

CAREL J.J. EIJGENRAAM is programmaleider ruimtelijke economie bij het Centraal Planbureau in Den Haag
E-mail: <c.j.eijgenraam@cpb.nl>.



LIDIA ARENDS

Voor veel situaties waarin een meta-analyse een multivariate uitkomst heeft, ontbreken nog geschikte meta-analytische methoden of zijn deze nog onderontwikkeld. Lidia Arends heeft in haar proefschrift* statistische methoden ontwikkeld, die geschikt zijn om meta-analytische data met een multivariaat karakter op een optimale manier te analyseren. Met de door haar voorgestelde methoden kan men veelomvattender onderzoeksvragen beantwoorden dan met de standaard univariate methoden die normaal worden gebruikt. De door haar voorgestelde methoden zijn relatief gemakkelijk te gebruiken voor de meeste onderzoekers, omdat ze met standaard software zijn uit te voeren.



MULTIVARIATE META-ANALYSE

Wat te doen bij meerdere uitkomsten per studie?

Een meta-analyse is een onderzoek waarin studies over eenzelfde fenomeen worden samengevoegd. Dit wordt meestal gedaan om een secuurder uitkomst te verkrijgen. Er worden bijvoorbeeld allerlei gepubliceerde studies verzameld die het behandelingseffect van paracetamol bij hoofdpijn rapporteren. Met behulp van statistische methoden worden de gevonden effecten van paracetamol vervolgens gemiddeld. Hierbij worden de resultaten van grote studies veelal zwaarder gewogen dan de resultaten van kleine studies. Een meta-analyse kan worden gedefinieerd als een kwantitatieve samenvoeging van de resultaten van gerelateerde maar onderling onafhankelijke studies. Als een meta-analyse gericht is op slechts één uitkomstmaat in al die aparte studies, dan is men het er in de wetenschappelijke wereld redelijk over eens welke statistische methoden men moet gebruiken om een algemeen gemiddeld resultaat te berekenen. Het zogenoemde univariate random effecten model heeft in dat geval de voorkeur. In zo'n model wordt er rekening mee gehouden dat de studieresultaten meer van elkaar zullen verschillen dan op grond van uitsluitend toeval: de tussen-studie-heterogeniteit. Immers, de ene studie zal zijn uitgevoerd bij zie-

kere patiënten, bij een hogere dosis paracetamol, in een ander seizoen, enz, enz dan een andere studie.

Maar wat als je percentages zieken en percentages overledenen per studie in een behandelde en onbehandelde groep hebt? Of als percentages overledenen per studie op meerdere tijdstippen worden weergegeven? Sterker nog, de ene studie kan bovendien deze overlijdenspercentages na 1, 2 en 3 jaar rapporteren, terwijl een andere studie de overlijdenspercentages na 0.5, 1.5 en 2.5 jaar geeft. Zo zijn er nog vele praktische situaties te noemen waar in de individuele studies meerdere uitkomstmaten worden gepresenteerd. In dat geval is het univariaat uitvoeren van een meta-analyse voor elke uitkomstmaat of elk tijdstip apart vaak niet optimaal. Het tegelijkertijd analyseren van alle uitkomstmaten met multivariate methoden is dan raadzaam.

Relatie behandelingseffect en onderliggend risico

Een eerste toepassing van een meta-analyse met een multivariate uitkomst is een onderzoek naar

de relatie tussen het effect van de behandeling en specifieke patiëntkenmerken. Het probleem daarbij is dat we in een meta-analyse meestal data hebben op studieniveau, dus bijvoorbeeld het percentage zieken in de behandelde groep en in de onbehandelde groep. Hoe kunnen we nu onderzoeken of de behandeling beter werkt bij ziekere patiënten dan bij minder zieke patiënten? Als we alleen maar gegevens op studieniveau hebben, kunnen we daarvoor per studie het risico van de patiënten berekenen. In studies met veel sterke, jonge personen zullen veel minder mensen ziek worden dan in studies met veel kwetsbare, oude personen. Door in elke studie te kijken naar het percentage zieken in de onbehandelde controlegroep, zou je een indicatie kunnen krijgen van hoe ziek de groep patiënten per studie is. We noemen dit percentage zieken in de onbehandelde controlegroep het onderliggende risico van de mensen in die specifieke studie. Aannemende dat het allemaal gerandomiseerde studies zijn, zal dit gelden voor zowel de onbehandelde als voor de behandelde groep. We kijken alleen naar de onbehandelde groep, omdat deze goed vergelijkbaar zal zijn tussen de verschillende studies.

Naast de berekening van het onderliggende risico kunnen we ook de behandelingseffecten in de verschillende studies berekenen. Dit is bijvoorbeeld het (relatief) verschil tussen de uitkomst in de behandelde en de onbehandelde groep. Let op, dit is dus het verschil tussen het percentage zieken in de behandelde groep en het zg. onderliggende risico van de patiënten in deze studie.

Een eenvoudige manier om vervolgens de relatie tussen het behandelingseffect en het onderliggende risico te schatten, is door het berekenen van een gewogen kleinste kwadraten regressielijn, waarbij de behandelingseffecten in de verschillende studies worden geschat op basis van de geschatte onderliggende risico's. Deze conventionele methode heeft verschillende manco's en er is

dan ook veel statistische kritiek op deze methode geleverd. Met name het negeren van de meetfout die in de meting van de onderliggende risico's zit, kan leiden tot een vertekening. Als er bijvoorbeeld per toeval in een studie heel veel mensen in de controlegroep ziek worden, zal dit zowel tot uiting komen in een extra hoog onderliggend risico als in een extra hoog behandelingseffect. Het behandelingseffect komt immers tot stand door het vergelijken van het aantal zieken in de behandelde groep versus de onbehandelde groep en als die laatste per toeval te hoog uitvalt, zal het aantal zieken in de behandelde groep al snel lager uitvallen, ook als de behandeling zelf geen effect heeft.

Met een bivariate random-effecten meta-analyse kan de relatie tussen het behandelingseffect en het onderliggende risico van de patiënten op een statistisch betere manier worden geschat. Hierbij wordt de relatie berekend tussen beide variabelen, waarbij ze allebei als het ware zijn 'ontdaan' van hun meetfout. Deze analyses kunnen worden uitgevoerd met behulp van de BUGS implementatie van Markov Chain Monte Carlo (MCMC) numerieke integratietechnieken. De methode kan echter ook worden uitgevoerd met standaardsoftware, zoals bijvoorbeeld het lineair gemengd model in SAS Proc Mixed. Het voordeel van BUGS is dat het exacte meetfoutenmodel kan worden gespecificeerd. De procedure Proc Mixed van SAS is echter eenvoudiger in het gebruik.

Trivariate meta-analyse

Een ander voorbeeld van een meta-analyse met meerdere uitkomsten is een situatie waarin het effect van een operatie wordt vergeleken met een behandeling met medicijnen. Het probleem hierbij is dat op korte termijn de operatie een

lagere overleving zal geven dan medicamenteuze behandeling (door de risico's van de ingreep), terwijl dit op langere termijn omgekeerd zal zijn. We keken daarom met een meta-analyse naar het korte- en lange-termijneffect van beide behandelingen op het krijgen van een beroerte bij patiënten die daar een verhoogd risico op hebben. Er zijn drie uitkomstmaten per studie beschikbaar, die we alle drie samen analyseren met een algemeen lineair gemengd model. Voor zover ons bekend, is dit het eerste gepubliceerde voorbeeld van een multivariate random-effecten meta-analyse waarin meer dan twee uitkomsten samen worden geanalyseerd¹. Er zitten veel voordelen aan een multivariate analyse ten opzichte van univariate analyses waarin elke uitkomst apart wordt geanalyseerd. De multivariate benadering laat de relaties tussen de verschillende uitkomsten zien en geeft eenvoudige functies voor het schatten van afgeleide behandelingseffecten zoals de cumulatieve overlevingskans ratio als een functie van de tijd dat de patiënten na de behandeling in het onderzoek zijn gevolgd. Bovendien maken de resultaten van een multivariate benadering het mogelijk om ook de relatie te schatten tussen de verschillende behandelingseffectparameters en het onderliggende risico, zoals we ook in het eerste voorbeeld lieten zien.

Bij het 'fitten' van het trivariate model in het standaard lineaire-gemengd-model programma van SAS kan men gebruik maken van de benaderende ('approximate') likelihood. In een paar speciale gevallen is ook het gebruik van de 'exacte' likelihood mogelijk in standaard software. Voor de multivariate situatie hebben we het model tevens gefit met Bayesiaanse methoden om de binomiale verdeling te specificeren van het aantal beroertes na de operatie en de Poisson verdelingen te specificeren voor het aantal beroertes op de lange termijn in beide behandelingsgroepen. De resultaten van het gebruik van de benaderen-

de ('approximate') en de 'exacte' likelihood blijken in dit voorbeeld zeer vergelijkbaar te zijn.

Meta-analyse van ROC curve data

Een heel andere toepassing van multivariate meta-analyse is de meta-analyse van ROC curve data. In een ROC curve wordt gekeken hoe goed een test werkt. Dit wordt weergegeven in twee maten, de sensitiviteit en de specificiteit. De sensitiviteit geeft aan hoeveel procent van de echte zieken door de test ook als 'ziek' wordt gediagnosticeerd. De specificiteit geeft aan hoeveel procent van de echte gezonden door de test ook als 'gezond' wordt gediagnosticeerd. Een goede test heeft uiteraard zowel een hoge sensitiviteit als een hoge specificiteit.

Stel dat er per studie één paar geschatte sensitiviteit en specificiteit beschikbaar is en dat je op grond van die enkele punten over meerdere studies een gezamenlijke ROC curve wilt schatten. Meta-analyse van ROC curve data wordt gewoonlijk gedaan met vaste effecten modellen. Ondanks dat er enkele random-effect modellen zijn gepubliceerd om een meta-analyse van ROC curve data uit te voeren, worden deze modellen vanwege hun complexiteit in de praktijk niet vaak toegepast. Multivariate random-effecten meta-analyse van ROC-curve data kan echter worden uitgevoerd met standaard software. De sensitiviteiten en de specificiteiten van de diagnostische studies worden samen geanalyseerd met een tweedimensionaal random-effecten model. Wel moet worden bedacht dat er verschillende alternatieven zijn om de geschatte bivariate normale verdeling te karakteriseren door een regressielijn of een zogenaamde samengevatte ROC curve. Wil je bijvoorbeeld de sensitiviteit schatten op basis van een bepaalde specificiteit, of wil je de specificiteit schatten

op basis van een bepaalde sensitiviteit? Dit kan heel verschillende samengevatte ROC curven opleveren!

Meta-analyse van overlevingsduur data

Het algemene lineaire-gemengde model is ook geschikt voor meta-analyse van overlevingsduur data, waarbij elke klinische trial meerdere survivalpercentages met bijbehorende standaardfouten presenteert gedurende een bepaalde periode nadat de behandeling is gestart. In de praktijk zijn de tijden waarop de gegevens zijn gemeten en het aantal metingen per studie vaak zeer verschillend. Daarom reduceren onderzoekers vaak de overlevingscurve tot één of meerdere vaste tijdstippen, bijvoorbeeld de overleving na één jaar in alle studies die daar een schatting voor geven. En los daarvan kijkt men dan naar de overleving na drie jaar, voor zover de studies daar iets over vermelden. In dat geval kunnen de data voor elk van de gekozen tijdstippen apart worden geanalyseerd met het standaard univariate random- (of vaste) effecten model. Het uitvoeren van aparte analyses voor elk tijdstip en dus het uitvoeren van veel meta-analyses, is echter inefficiënt en kan leiden tot foute conclusies. Er zijn in de literatuur betere methoden voorgesteld, maar dat zijn allemaal vaste-effecten modellen. Eén van de meest aanbevolen (vaste-effecten) methoden is het model van Dear². Dear stelt een algemeen lineair model voor, waarbij de overlevingspercentages worden geschat op basis van de follow-up tijd, de soort behandeling en de studie. Al deze predictoren worden als categorische covariaten in het model gestopt. De parameters worden geschat met GLS (generalised weighted least squares). Wij hebben de GLS-methode van Dear uitgebreid tot een multivariaat random-effecten model, dat kan worden toegepast op data met een willekeurig aantal overlevingsschattingen en waarbij wil-

lekeurige tijdsintervallen tussen de schattingen bestaan, mogelijk verschillend tussen de studies. Dit maakt het voor degene die de meta-analyse doet mogelijk om alle beschikbare data te analyseren, zoals ze worden weergegeven in de publicaties, zonder de noodzaak van inter- of intrapolatie naar vaste tijdstippen. De methode past in het raamwerk van het algemene lineaire gemengde model met een aanpassing voor deze specifieke situatie, omdat de correlaties tussen de verschillende overlevingsschattingen per studie ook moeten worden geschat.

Samenvattend kunnen we zeggen dat veel bestaande meta-analyse methoden (al dan niet met meerdere uitkomstmaten) kunnen worden uitgevoerd met het algemeen lineair-random-effects model. In de *Tutorial in Biostatistics over geavanceerde methoden op het gebied van meta-analyse van Van Houwelingen et al*³ staat uitgebreid omschreven hoe men dit allemaal in de praktijk kan uitvoeren.

Noot

* L.R. Arends (2006) *Multivariate meta-analysis: modelling the heterogeneity. Mixing apples and oranges: dangerous or delicious?* Proefschrift, Erasmus Universiteit Rotterdam.

LITERATUUR

1. Arends, L.R., Z. Voko & T. Stijnen T. (2003), Combining multiple outcome measures in a meta-analysis: an application. *Statistics in Medicine* 22-8, pp 1335-1353.
2. Dear, K.B.G. (1994), Iterative generalized least squares for meta-analysis of survival data at multiple times. *Biometrics* 50, pp 989-1002.
3. Houwelingen, H.C. van, L.R. Arends & T. Stijnen (2002), Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* 21-4, pp 589-624.

LIDIA ARENDS promoveerde in 2006 aan de Erasmus Universiteit Rotterdam, onder begeleiding van prof. dr. Theo Stijnen. Ze is werkzaam als universitair docent M&T bij het Instituut voor Psychologie aan de Erasmus Universiteit Rotterdam en als Biostatisticus bij Epidemiologie & Biostatistiek, Erasmus MC, Rotterdam. E-mail: <Arends@fsw.eur.nl> of <L.Arends@erasmusmc.nl>



ROBUUSTE OPTIMALISATIE VAN RETAIL ASSORTIMENTEN

Hoe kunnen retailers assortimenten selecteren die robuust zijn tegen de onzekerheid in productverkoop? Retailers zien zich geplaagd voor de vraag welke producten ze moeten opnemen in het assortiment van een categorie om de winst te maximaliseren. Een complicerende factor in de optimalisatie is dat de winstbijdrage van elk product onzeker is, aangezien de vraag naar het product niet met 100% zekerheid voorspeld kan worden. In dit artikel bespreken we een methode waarmee retailers assortimenten kunnen selecteren die robuust zijn tegen de onzekerheid in winstbijdrages. Daartoe formuleren we een robuuste variant van het Retail Assortiment Optimalisatieprobleem welke de balans zoekt tussen de verwachte winst van een assortiment en de daarmee gepaard gaande onzekerheid.

ROBERT ROODERKERK

Het assortiment is een belangrijk element van de retailer's marketing mix. Een belangrijke vraag voor de retailer is welk assortiment hij moet kiezen binnen een productcategorie (bijvoorbeeld tandpasta's) om de categorie winst te maximali-

seren en de hoeveelheid beschikbare schapruimte niet te overschrijden. We noemen dit het Retail Assortiment Optimalisatieprobleem (RAO). We nemen aan dat de structurele winstbijdrages van individuele producten onafhankelijk zijn van de

aanwezigheid van andere producten. Als gevolg hiervan kunnen we RAO als een lineair knapzakprobleem formuleren (Kellerer, Pferschy en Pisinger 2004). Vaak worden de winstbijdrages van de producten gebaseerd op een model van productverkoop. Zoals in elk model is ook hier typisch onzekerheid in de modelschattingen te verwachten. Dit resulteert in onzekerheid in de input van het RAO-probleem. Door aan te nemen dat de winstbijdrages in een onzekerheidsset U liggen, kunnen we een robuuste variant van het Retail Assortiment Optimalisatieprobleem (RAO_{Robuust}) formuleren:

$$\begin{aligned} \max_x \quad & \min_{p \in U} \sum_{k=1}^n p_k x_k \\ \text{s.t.} \quad & \sum_{k=1}^n w_k x_k \leq c \\ & x_k \in \{0,1\} \quad k=1, \dots, n. \end{aligned} \quad (1)$$

Hierin is p_k de winstbijdrage van product k ($=1, \dots, n$), w_k de hoeveelheid schapruimte die product k inneemt indien het in het assortiment opgenomen wordt en c de hoeveelheid beschikbare schapruimte voor de categorie. De beslissingvariabelen x_k zijn 1 indien product k in het assortiment opgenomen wordt, 0 anders. Merk op dat het doel van de robuuste variant is om het assortiment te vinden waarvoor de slechts mogelijke winstrealisatie bij de gegeven onzekerheidsset maximaal is. We maximaliseren zeg maar het *worst case* scenario.

Op basis van Ben-Tal en Nemirovski (1998) kunnen we (RAO_{Robuust}) in combinatie met een ellipsoïde onzekerheidsset U voor de winstbijdrages herformuleren tot het volgende Robuuste Retail Assortiment Optimalisatieprobleem (RRAO):

$$\begin{aligned} \max_x \quad & \sum_{k=1}^n \bar{p}_k x_k - r \sqrt{\sum_{k=1}^n \sum_{k'=1}^n \theta_{kk'} x_k x_{k'}} \\ \text{s.t.} \quad & (1) \text{ en } (2) \end{aligned}$$

Hierin is \bar{p}_k de verwachte winstbijdrage van product k en representeert $\theta_{kk'}$ de covariantie in de winstbijdrages van product k en k' . De doelstellingsfunctie vertegenwoordigt nu de robuuste winst van een assortiment. Deze robuuste winst balanceert de verwachte (i.e. nominale) winst en de standaarddeviatie van de winst. De parameter r bepaalt nu de penalty op onzekerheid: als de beslissingsnemer meer risicoaversief is dan beboet hij/zij variantie (i.e. onzekerheid) in de winst meer. Maar hoe kiezen we een waarde voor r ?

We veronderstellen dat de winstrealisaties een (multivariate) normale verdeling volgen (i.e. $p \sim N(\bar{p}, \Theta)$). Stel dat we de kans, dat de daadwerkelijke winstrealisatie van een assortiment lager is dan de bijbehorende robuuste winst, willen beperken tot α . Hieruit volgt een waarde voor r . Bijvoorbeeld een α van 5% impliceert een waarde voor r van 1.64. Merk op dat de waarde voor r , de penalty op onzekerheid, hoger zal zijn indien de beslissingsnemer meer zekerheid wil en dus een lagere waarde voor α eist.

Synthetisch voorbeeld

We contrasteren de nominale en robuuste variant nu met een eenvoudig voorbeeld van een Retail Assortiment Optimalisatieprobleem, beschreven in Tabel 1. In dit geval leggen de drie producten (i.e. items) evenveel beslag op de beschikbare schapruimte maar kunnen er maar maximaal twee producten tegelijk in het assortiment opgenomen worden.

Product 2 en 3 hebben een grote, positieve covariantie, wat het gevolg kan zijn van een vergelijkbare gevoeligheid voor schokken in de vraag. Bijvoorbeeld, als dit kledingartikelen zouden zijn, dan kunnen product 2 en 3 dezelfde kleur hebben (e.g. beide rood), welke in de mode kan zijn (positieve vraagschok) of niet (negatieve vraagschok).

Item (k)	Schapruimte (w_k)	Verwachte winst (\bar{p}_k)	Winst covariantiematrix (Θ)			Beschikbare schapruimte (c)
			Item 1	Item 2	Item 3	
1	1	8	1	-1.5	-1	2
2	1	9		4	5	
3	1	10			9	

Tabel 1: Voorbeeld van een Retail Assortiment Optimalisatieprobleem

Product 1 heeft een negatieve covariantie met product 2 en 3, wat een indicatie kan zijn van een tegengestelde gevoeligheid voor vraagschokken. Als we het eerdere voorbeeld volgen, dan zou product 1 bijvoorbeeld een andere kleur kunnen hebben (e.g. blauw) dan product 2 en 3 (rood), en rood en blauw zijn om en om in de mode. Voor elk van de toegelaten assortimenten beschrijft Tabel 2 de nominale winst, variantie van de winst en de robuuste winst.

Het blijkt dat de optimale nominale oplossing, {2,3}, niet de optimale oplossing is van het Robuuste Retail Assortiment Optimalisatieprobleem. Dit komt door de hoge variantie die met dit assortiment gepaard gaat. De optimale robuuste oplossing is {1, 2}. Omdat product 1 en 2 een sterke negatieve covariantie hebben (zie Tabel 1), is

Oplossing	Nominale winst	Variantie ($x^T \Theta x$)	Robuuste winst
{1}	8.00	1.00	6.36
{2}	9.00	4.00	5.72
{3}	10.00	9.00	5.08
{1, 2}	17.00	2.00	<u>14.68</u>
{1, 3}	18.00	8.00	13.36
{2, 3}	<u>19.00</u>	23.00	11.13

Tabel 2: Toegestane oplossingen voor het Retail Assortiment Optimalisatieprobleem: nominale vs. robuuste winst bij $r = 1.64$ ($\alpha = 5\%$).

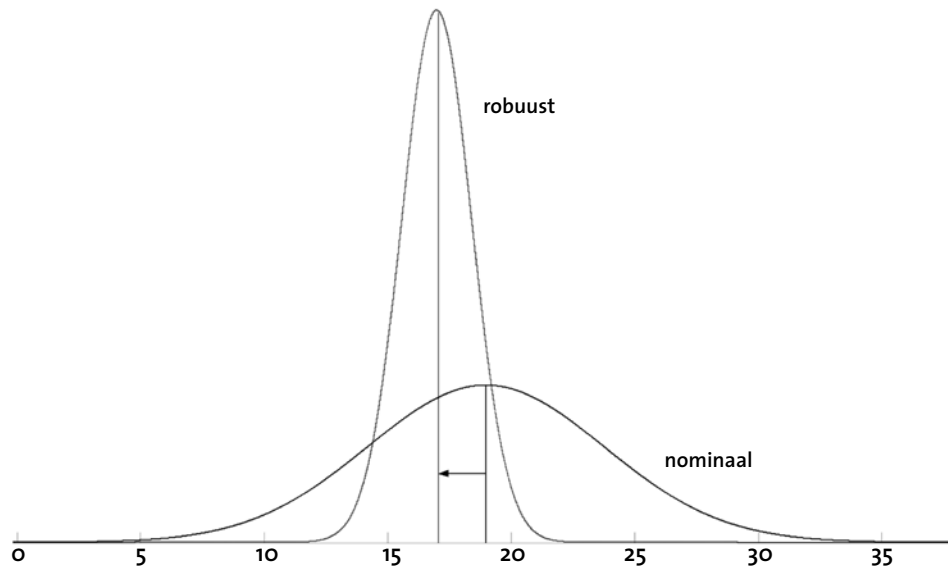
het opnemen van beide producten een *hedging mechanisme* voor de retailer. Als we de optimale robuuste oplossing, {1, 2}, vergelijken met de optimale nominale oplossing, {1, 3}, dan zien we dat de het robuuste assortiment een veel lagere variantie (2.00) heeft dan de winstrealisatie van het nominale assortiment (23.00). Dit wordt geïllustreerd in Figuur 1, waarin de winstverdeling van het optimale nominale en robuuste assortiment vergeleken worden.

Figuur 1 laat zien dat de hoeveelheid onzekerheid met betrekking tot de winstrealisatie aanzienlijk verminderd kan worden door een assortiment te kiezen dat tot slechts een kleine daling in de verwachte winst leidt. In feite, de verwachte winst van het optimale robuuste assortiment (17.00) is slechts 10.5% minder dan die van het optimale nominale assortiment (19.00), maar de robuuste winst (14.68) is 31.9% meer dan die van het optimale nominale assortiment (11.13).

Dit voorbeeld laat het potentiële voordeel zien van de robuuste optimalisatie om oplossingen voor het Retail Assortiment Optimalisatieprobleem te vinden die een relatief klein verlies in verwachte winst opofferen voor een substantiële reductie in onzekerheid.

Empirische Toepassing

We hebben het potentieel van de robuuste aan-



Figuur 1: Winst Verdeling van het Optimale Nominale en Robuuste Assortiment

pak ook bepaald in een empirische toepassing met betrekking tot de categorie keukenrollen. IRI Frankrijk heeft wekelijkse scanner-data beschikbaar gesteld voor 54 winkels van een grote landelijke Franse retailer voor de periode van september 2002 tot september 2005. In de hele keten worden 21 verschillende producten verkocht, gemiddeld is er per winkel voor 15 producten plaats. De doelstelling is om per winkel het assortiment te optimaliseren zodat de categorie winst gemaximaliseerd wordt. In de afwezigheid van marge-data, veronderstellen we gelijke marges (i.e. 25%). Op basis van de scanner-data hebben we een hiërarchisch Bayesiaans model van wekelijkse productverkoppen geschat (Rossi en Allenby, 1993). De verdeling van de winstbijdrages is vervolgens gebaseerd op de posterior verdeling van de wekelijks verkopen. Hierbij is aangenomen dat het een gewone week betreft met reguliere prijzen en zonder promotionele activiteiten.

Door volledige enumeratie hebben we per winkel het optimale nominale en robuuste assortiment bepaald. Om meer inzicht te krijgen in de verschillen tussen deze twee soorten assortimenten vergelijken we ze op basis van de nominale

en robuuste winst. Vanzelfsprekend leidt het optimale robuuste assortiment in vergelijking met het optimale nominale assortiment tot minder nominale winst maar tot meer robuuste winst. De vraag is echter hoeveel? Wat zijn de kosten en baten van de robuuste aanpak?

Figuur 2 contrasteert per winkel de relatieve toename in robuuste winst (witte staven) met de relatieve afname in nominale winst (zwarte staven) wanneer men het optimale nominale assortiment vervangt door het optimale robuuste assortiment. De winkels zijn gesorteerd naar relatieve toename in robuuste winst. Allereerst merken we op dat de optimale robuuste en nominale assortimenten voor 24 (van de 54) winkels hetzelfde zijn. Voor een aanzienlijk deel van de overgebleven 30 winkels leidt het optimale robuuste assortiment tot een grote toename in de robuuste winst (grote witte staaf) maar een kleine afname in de nominale winst (kleine zwarte staaf). Dit geldt in het bijzonder voor de winkels aan de uiterste rechterkant. Voor deze winkels kan een substantiële reductie in de onzekerheid gerealiseerd worden tegen relatief lagen kosten (i.e. afname in verwachte winst). Echter het is interessant om te zien

dat het tegenovergestelde ook opgaat voor sommige winkels. Hier zijn aanzienlijke offers vereist om een afname in onzekerheid te realiseren. Dit mogelijke nadeel van de robuuste aanpak is tot nu toe nog niet onderstreept in de literatuur.

Tot Slot

In dit artikel zijn de potentiële voor- en nadelen van robuuste optimalisatie in kaart gebracht in het domein van retail assortiment optimalisatie. In het onderzoek waarop dit artikel gebaseerd is, hebben we ook een grid heuristiek ontwikkeld om (bijna) optimale oplossingen te vinden voor het Robuuste Retail Assortiment Optimalisatieprobleem. Deze heuristiek doet er een fractie van de tijd over die volledige enumeratie hiervoor nodig heeft. Bovendien blijken de

gevonden oplossingen bij een fijnmazig grid altijd optimaal te zijn.

De auteur dankt IRI Frankrijk voor het beschikbaar stellen van de data.

LITERATUUR

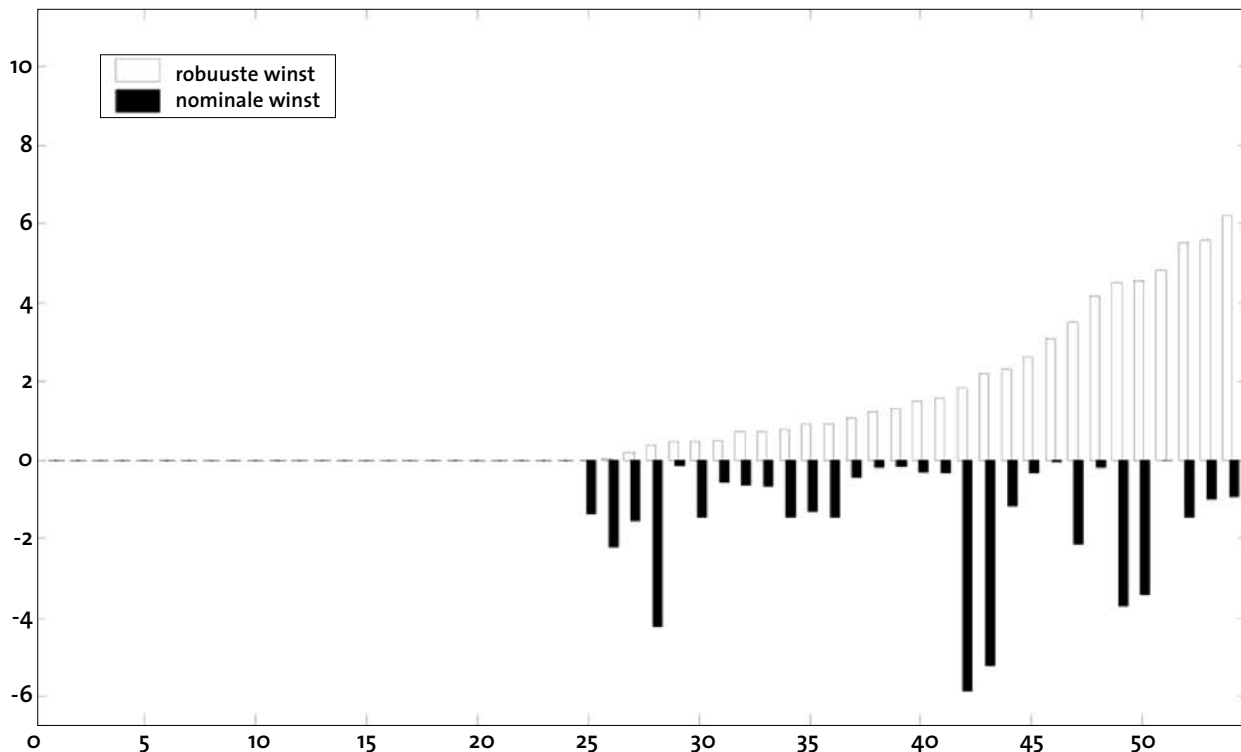
Het artikel is gebaseerd op: Rooderkerk, R.P., H.J. van Heerde & Dick den Hertog (2007), *Robust Optimization of Retail Assortments*. Working Paper. Tilburg University.

Ben-Tal, A. en Nemirovski, A. (1998), Robust Convex Optimization, *Mathematics of Operations Research*, Volume 23, Issue 4, pp 769-805.

Kellerer, H., Pferschy, U. en Pisinger, D. (2004), *Knapsack Problems*, Berlin: Springer-Verlag.

Rossi, P.E. en Allenby, G.M. (1993), A Bayesian Approach to Estimating Household Parameters, *Journal of Marketing Research*, Volume 30, Issue 2, pp 171-182.

ROBERT ROODERKERK is werkzaam als Assistant Professor of Marketing & CentER fellow aan de Universiteit van Tilburg. E-mail: <R.P.Rooderkerk@uvt.nl>.



Figuur 2: Relatieve winstvergelijking van optimale robuuste vs. nominale assortimenten



column

Foto: Pieter Bosch

STATISTICI, KANSREKENAARS EN ANDERE WISKUNDIGEN GELAUWERD

FRED STEUTEL

Sommige wiskundigen zijn heel oud geworden: De Nederlander Dirk Struik werd 106 en de Oostenrijker Vietoris 110, maar sommige anderen leefden heel kort en werden toch beroemd. Anderzijds hoef je om beroemd te worden in de kansrekening of statistiek niet per se heel oud te worden of heel jong te sterven. Een paar voorbeelden.

De laatste tijd was een aantal bekende statistici (kansrekenaars) prominent in het nieuws: Harry Kesten, Srinivasa S. R. Varadhan en Vincent Doblin. Twee van hen ontvingen belangrijke prijzen, de derde is al 67 jaar dood. Voor de toegepaste statisticus zal van deze coryfeeën mogelijk geen enkele naam bekend voorkomen; daarom zowel over de prijswinnaars als over de prijzen wat achtergrondinformatie.

Harry Kesten werd eredoctor aan de universiteit Paris Sud voor zijn baanbrekende werk op verschillende terreinen van de kansrekening: percolatietheorie, *random walks* en vertakkingsprocessen. Kesten (1931) studeerde wiskunde in Amsterdam, onder meer bij Van Dantzig en De Bruijn. Hij emigreerde in 1955 naar de VS, waar hij in 1958 promoveerde bij Mark Kac aan de Cornell universiteit; hij werkt daar nog steeds. In Amsterdam werkte hij intensief samen met Theo Runnenburg. Zij schreven een prachtig artikel over wachtrijproblemen met prioriteiten, '*Priority in waiting line problems*'. Zij waren begin jaren vijftig - wiskundig - onafscheidelijk. In 1981 ontving Kesten, uit handen van Runnenburg, de prestigieuze Brouwermedaille, genoemd naar de beroemde Nederlandse wiskundige Luitzen Egbertus Jan

Brouwer, ter wiens nagedachtenis onlangs een postzegel is uitgegeven met een onderschrift in logische symbolen, dat gelezen kan worden als 'de regel: òf A is waar òf niet-A is waar, is niet bewijsbaar'. Eerder ontving Kesten de Pólya-prijs en de Steele-prijs. Hij was nog vrij onlangs op bezoek bij Eurandom in Eindhoven.

Varadhan ontving dit jaar de Abelprijs voor zijn werk op het gebied van *large deviations*, een klasse van limietstellingen, die veel subtieler zijn dan bijvoorbeeld de centrale limietstelling. Het gaat dan over de kans op extreme situaties, omstandigheden die zich zeer zelden voordoen. Er zijn tal van praktische toepassingen, in de verzekeringsbranche en bij de constructie van boorplatforms. Varadhan werd in 1940 geboren in Madras, India. Hij studeerde ook in India, maar is sinds 1963 verbonden aan het Courant Instituut in New York. De Abelprijs is te beschouwen als de Nobelprijs voor wiskunde. Immers er bestaat geen Nobelprijs voor wiskunde. Er zijn allerlei geruchten over het waarom niet, en over de veronderstelde afkeer van Nobel van wiskundigen. De prijs wordt sinds 2003 elk jaar in Oslo uitgereikt door de koning van Noorwegen; de bijbehorende geldprijs bedraagt 6.000.000 Noorse kronen, ongeveer een miljoen dollar. De prijs is genoemd naar Niels Henrik Abel, een Noorse wiskundige met een nogal tragische geschiedenis: hij stierf in 1829, straatarm en 26 jaar oud, aan tuberculose, op de dag dat hem een hoogleraarschap in Berlijn werd aangeboden. Hij is bekend door zijn oplossing van vijfdegraads vergelijkingen en leeft voort in het woord 'abels' (commutatief) als in 'abelse groepen'. Op de campus van de universiteit van Oslo zijn twee gebouwen naar wiskundigen genoemd: het Lie Hus, naar Sophus Lie bekend van de Lie Groepen (we kennen de naam Lie van Trygve Lie, de eerste secretarisgeneraal van de VN), en het Abels Hus, genoemd naar Niels Henrik Abel.

De derde kansrekenaar is alweer een tragische figuur: Wolfgang Doblin, zoon van Alfred



Vincent Doblin

Döbblin, de auteur van *Berlin Alexanderplatz*. Mijn aandacht werd op hem gevestigd door een lang artikel van Bru en Yor (1). Onlangs hebben Agnes Handwerk en Harrie Willems een documentaire over Doblin gemaakt met de titel *A mathematician rediscovered* (2). De spelling van de achternaam van Wolfgang is anders dan die van Alfred, zonderumlaut. Wolfgang, als jood naar Frankrijk uitgeweken, was in 1936 Fransman geworden en had zijn voornaam in Vincent veranderd. Wolfgang (Vincent) Doblin kwam in het nieuws toen de Académie Française in 2000 een verzegelde brief (*pli cacheté*) opende die Doblin als militair in februari 1940 naar Parijs had gestuurd. Het onderwerp van het handschrift betrof stochastische differentiaalvergelijkingen (3). Toen zijn eenheid door Duitsers was omsingeld, pleegde hij bij het Franse dorp Housseras op 21 juni 1940 zelfmoord. Hij was vijfentwintig jaar. Zijn vader heeft de oorlog overleefd. In Housseras staat een monument voor Alfred en zijn vrouw, en ook één voor Vincent: *mort pour la FRANCE*.

De jongst gestorven beroemde wiskundige is Evariste Galois (van de Galois-theorie). Hij stierf, twintig jaar oud, aan de gevolgen van een duel.

LITERATUUR

1. Bru, B. & M. Yor (2002), Comments on the life and mathematical legacy of Wolfgang Doeblin. *Finance and Stochastics* 6, pp 3-47
2. Handwerk A. & H. Willems (2007), *Wolfgang Doeblin, A Mathematician Rediscovered* (DVD), Springer, VideoMATH
3. Willems, H. (2002), Verzegelde formules. *Nieuw Archief voor Wiskunde*, 5-3, pp 18-20

FRED STEUTEL is emeritus hoogleraar kansrekening aan DE TU EINDHOVEN. Hij is redacteur van STATOR.
E-mail: <f.w.steutel@tue.nl>

QUANTUM COMPUTATION EN OPTIMALISERING

De geschiedenis van de Operations Research kenmerkt zich door voortdurende vernieuwingen in de algoritmes, waardoor steeds krachtiger oplosmethodes beschikbaar komen. We kunnen nu een zeer fundamentele uitbreiding van onze 'toolbox' tegemoet zien: *quantum computation*. Deze techniek beperkt zich nu nog hoofdzakelijk tot fysici maar zal zeer binnenkort door ingenieurs worden overgenomen en snel daarna in productie komen. Tegen die tijd zullen we de tot nu toe meest succesvolle technieken in de OR dienen te heroverwegen.

DAVID BULGER

Wat is quantum computation?

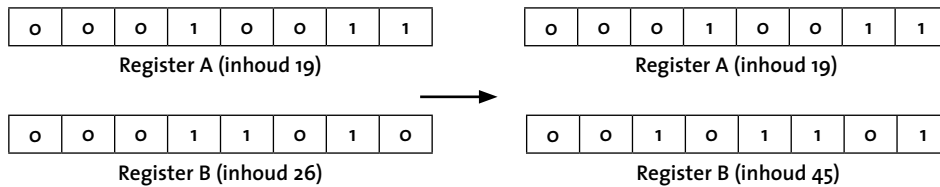
Het idee achter quantum computation is door gebruik te maken van de raadselachtige eigenschappen van de kwantumtheorie daarmee de hedendaagse informatie-technologie grotere mogelijkheden te geven. Op deze manier zijn we ook van gloeilamp naar laser en van koperdraad naar halfgeleiders gekomen.

Alle computers, vanaf de eerste buizenmachines tot de modernste hedendaagse PC, werken op basis van hetzelfde principe: gegevens zijn als een reeks nullen en enen (bits) gecodeerd. Geheugen

en processoren zijn er op gebouwd deze reeksen te manipuleren. Een van de ingebouwde basisfuncties van iedere microprocessor telt bijvoorbeeld het binaire getal in één register op bij dat in een ander (Figuur 1).

Quantum computation gaat verder door de gebruikelijke bits te vervangen door kwantum bits, *qubits* genaamd, die de kwantumeigenschappen *superpositie* en *verstrengeling* vertonen.

Superpositie betekent dat een qubit zich bevindt in een ongedefinieerde toestand, een mengeling van 0 en 1. Dit lijkt op een random toestand maar is dat niet helemaal. Als een qubit wordt 'gelezen'



Figuur 1: Een opteloperatie van register A naar register B op een klassieke computer

zal er een 0 of een 1 tevoorschijn komen. Maar zolang we dit niet doen kan een qubit gelijktijdig zowel 0 als een 1 opslaan. Superpositie is een interessant verschijnsel, ietwat verrassend, maar niet moeilijk te begrijpen. Het kan als zodanig niet veel bijdragen aan het vergroten van de rekenkracht van een computer.

Verstrengeling daarentegen is het echte werkpaard van *quantum computation*, en een geweldige uitdaging voor ieder intuïtief begrip. Stel we hebben twee qubits: door superpositie bevat elk een kwantum mix van de 0 en 1 toestanden. Verstrengeling zorgt ervoor dat deze superposities niet onafhankelijk hoeven te zijn. Als we de beide qubits als een paar beschouwen hebben we de superpositie van de vier toestanden 00, 01, 10 en 11; maar we kunnen ook een superpositie van de twee toestanden 00 en 11 hebben, of van de drie toestanden 00, 01 en 11.

Een kwantumcomputer zal numerieke en andere gegevens op dezelfde manier opslaan en bewerken als een 'gewone' computer. Daarnaast echter kunnen kwantumcomputers gebruik maken van superpositie en verstrengeling van deze gegevens. Door hier op een slimme manier gebruik van te kunnen algoritmes ontworpen worden die de interferentie van kwantumtoestanden gebruiken om *short cuts* te vinden voor problemen.

Shors algoritme

Het algoritme dat Peter Shor in 1994 publiceerde

wordt als een lichtend voorbeeld gezien door iedereen die zich bezighoudt met het ontwikkelen van kwantum rekenmethoden. Dit algoritme kan een groot getal in zijn factoren ontbinden. De werking is – nog slechts op zeer kleine schaal – gedemonstreerd door een IBM groep in 2001: met een prototype van een 7-qubit computer werd 15 ontbonden in zijn factoren 3 en 5.

Het algoritme gaat uit van de veronderstelling dat het te ontbinden getal het product is van twee grote, onbekende, priemgetallen. Zonder quantum computation is de factorisering exponentieel afhankelijk van het aantal cijfers van de kleinste factor. Momenteel hebben we het dan al snel over zo'n 20 cijfers. Hoewel die exponentiële complexiteit niet is bewezen, gaan we er wel van uit: 2200 jaar onderzoek sinds Eratosthenes heeft nog geen snellere methode opgeleverd.

Tenzij we *kwantum* algoritmes gebruiken. Het doorslaggevende idee van Shor was het gebruik van getaltheorie om het factoriseringsprobleem om te zetten naar het probleem van het vinden van de periode van een periodieke functie. Dat probleem kan veel efficiënter op een kwantumcomputer worden opgelost dan op een 'gewone', door het gebruik van de *kwantum Fourier transformatie*. De complexiteit van Shors algoritme is nu slechts polynomiaal afhankelijk van het aantal cijfers.

Hiermee heeft Shor in feite gedemonstreerd dat voor bepaalde vraagstukken een kwantumcomputer exponentieel sneller kan zijn dan een gewone. Dat is in twee opzichten revolutionair. Ten eerste berust moderne cryptografie op de

praktische onuitvoerbaarheid van het factoriseren van zeer grote getallen, maar met een kwantumcomputer zou dit vervallen. Ten tweede toont Shors algoritme aan dat we op een compleet andere manier moeten gaan denken over dit soort problemen.

Grovers algoritme

Een ander welbekend kwantum algoritme is dat van Grover, gepubliceerd in 1996. Hij beoogde hiermee het *simple search* probleem op te lossen. Dit is een erg algemeen probleem: we willen uit een domein van mogelijke antwoorden op een vraag het correcte antwoord vinden. We hebben daarbij wel de mogelijkheid het correcte antwoord te herkennen, maar niet om het rechtstreeks te berekenen.

Stel dat we een database hebben met namen en telefoonnummers, alfabetisch gesorteerd op naam. We willen nu bij een telefoonnummer de bijbehorende naam zoeken. Tenzij we een index van de database hebben gesorteerd op nummer helpt geen enkele slimme methode ons. We zullen alle namen één voor één moeten bekijken tot we het door ons gezochte telefoonnummer hebben gevonden. Als de database N regels bevat en het gezochte nummer is willekeurig, dan zal deze *brute-force* methode gemiddeld $N/2$ zoekpogingen vergen.

In een kwantumcomputer kan zo'n database worden doorzocht met Grovers algoritme dat gebruik maakt van superpositie en verstrengeling. Hierdoor wordt het aantal zoekpogingen teruggebracht tot ongeveer $\pi\sqrt{N}/4$. (Een 'supergepositioneerde' zoekvraag kijkt gelijktijdig naar iedere regel van de database. Dat lijkt op bedrog, maar het punt is wel dat de totale zoekvraag net zo snel kan worden beantwoord dan wanneer er slechts naar één regel zou zijn gekeken.)

Grovers algoritme geeft dus een kwadratische

winst in snelheid bij het oplossen van het general search probleem. Ook kan men zeggen dat het gedeelte van het domein dat in een bepaalde tijd kan worden doorzocht kwadratisch toeneemt. Numeriek is dat wat minder dan de exponentiële winst van Shors algoritme, maar het komt meer verwonderlijk over. Het lijkt alsof we naar een computer kijken die iets kan vinden temidden van een miljoen plekken, terwijl hij er slechts duizend doorzoekt.

Quantum algoritmes voor optimalisering

De algoritmes van Shor en Grover zijn totaal verschillend van aard. Als men weet hoe beide werken krijgt men bij Shor de neiging te zeggen: 'Ik wou dat ik dat bedacht had.' Bij Grover daarentegen denkt men: 'Hoe is hij daar in 's hemelsnaam op gekomen?' Daarnaast is Shors algoritme toegepast op een specifiek probleem, al is dat dan ook een erg belangrijk probleem. Grovers algoritme is meer een basis bouwsteen die in vele tientallen verschillende toepassingen gebruikt kan worden. Aan de andere kant geeft Shor een exponentiële winst terwijl Grover niet verder komt dan kwadratisch.

Deze verschillen werken wat verwarrend als we deze twee beroemdste voorbeelden willen gebruiken als uitgangspunt voor de ontwikkeling van nieuwe kwantumalgoritmes. Het enige dat ze gemeen schijnen te hebben is hun totaal nieuwe aanpak die tegen iedere intuïtie ingaat. Algoritmes voor 'gewone' computers zijn niet meer dan afspiegelingen van bewerkingen die men ook op papier zou kunnen doen. Kwantumcomputers kunnen echter gegevens bewerken op manieren die niet als stapjes op papier kunnen worden gezien. We beschikken hiermee over zeer krachtig nieuw gereedschap, maar we hebben geen enkele intuïtief idee hoe we er mee moeten werken.

Het algoritme van Dürr en Høyer

Het eerste kwantum algoritme voor optimalisering, dat Dürr en Høyer uit 1996, is een tamelijk directe toepassing van Grovers algoritme. Het is een iteratief algoritme.

Stel we zoeken het minimum van een functie over een gegeven domein. Dürr en Høyers algoritme start met het berekenen van die functie op een willekeurig punt. Vervolgens wordt met behulp van Grovers algoritme gezocht naar een punt met een lagere waarde op de functie. In iedere volgende stap wordt vervolgens gezocht naar een nog lagere waarde. Voor discrete problemen zal op deze manier altijd het globale minimum worden gevonden.

Dit algoritme heeft dezelfde complexiteit als dat van Grover: bij een domein met 10¹² punten kan met ongeveer een miljoen operaties een uniek optimum worden gevonden. Op zich is dat niet erg indrukwekkend, hedendaagse computers presteren over het algemeen al beter. Als we echter bedenken dat dit resultaat bereikt wordt zonder dat gebruik wordt gemaakt van enige informatie over het systeem, dan mogen we ons afvragen wat er gebeurt als we dat wel zouden doen.

Quantum Basin Hopper

Recent werk van Bulger laat zien hoe de methode van Dürr en Høyer gecombineerd kan worden met een local search methode om multimodale functies te optimaliseren.

Stel dat we een efficiënte *local optimisation* techniek hebben, maar dat het probleem zeer vele lokale optima kent. Dit is weergegeven in Figuur 2, we kunnen gemakkelijk in een lokaal dal terechtkomen, maar het vinden van het allerlaagste dal kan erg moeilijk zijn. Voor dit soort problemen wordt vaak een *multi-start* heuristische methode

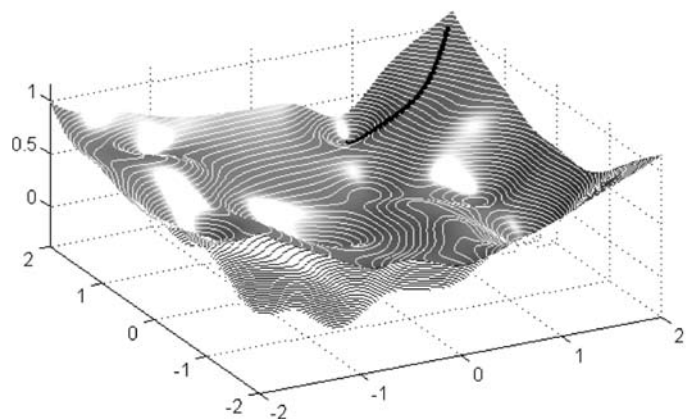
gebruikt. Het aantal 'afdalingen' dat nodig is om dat globaal laagste dal te vinden is ruwweg gelijk aan de helft van het totaal aantal dalen. Maar dit soort problemen kan in de praktijk miljarden dalen kennen.

De *quantum basin hopper* combineert the heuristische winst van de local search met de kwadratische winst van Grovers algoritme. Door de local search in een *kwantum orakel* op te nemen (dat is een kwantum subroutine die kan worden gezien als één enkele functie), wordt tegelijkertijd een aantal gesuperponeerde *local searches* uitgevoerd dat in de orde grootte ligt van de wortel uit het aantal dalen.

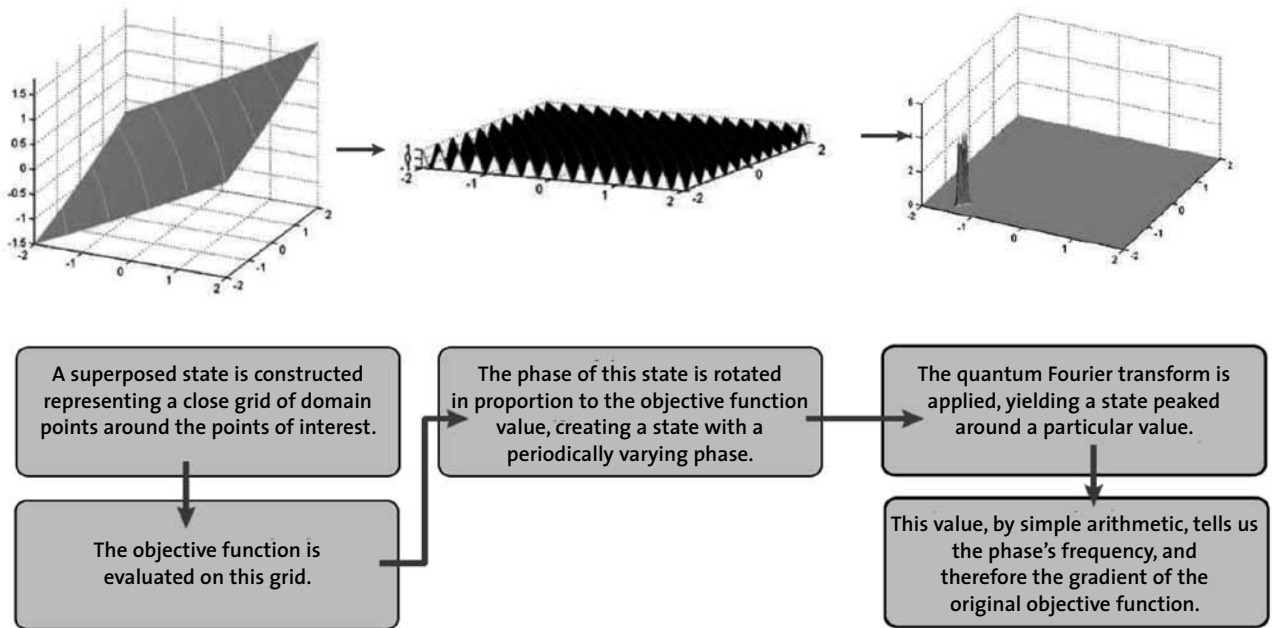
Jordans algoritme

Het recentelijk gepubliceerde *gradient estimation algorithm* van Jordan zal waarschijnlijk erg nuttig kunnen zijn in continue optimaliseringsproblemen. Het maakt gebruik van de gradiënt van een functie van meerdere variabelen.

Bedenk dat Shors algoritme het factoriseringsprobleem aanpakt door het eerst om te zetten in het probleem van het vinden van de periodiciteit. Hierop wordt een 'kwantum Fouriertransformatie' losgelaten. Jordans algoritme werkt op een soortgelijke manier. Het zet de helling van



Figuur 2: Een *local search* beginnend op (1.9, 1.9) vind een lokaal minimum op (0.83, 1.78), maar andere lokale minima zijn lager.



Figuur 3: Jordan's gradient estimation algorithm.

de objectieve functie om in een oscillerende functie en berekent dan de periode daarvan om zo die helling te vinden. Deze procedure is weergegeven in Figuur 3.

Voor het schatten van de gradiënt van een *black-box*-functie van d variabelen heeft een klassieke computer $d+1$ functie evaluaties nodig. Met een kwantumcomputer kan Jordans algoritme het in slechts een gesuperpositioneerde evaluatie. Voor hoog-dimensionele problemen kan dit een aanzienlijke winst opleveren.

De local search techniek zoals beschreven bij de *quantum basin hopper* is algemeen. Jordans algoritme kan worden ingepast in de implementatie van meer specifieke local search methodes, zoals conjugate gradiënt methodes.

Adiabatische Quantum Computation

Simulated annealing is een stochastisch niet-kwantum optimalisatie algoritme dat werkt op de manier waarop in afkoelende metalen een natuurlijke energie-optimalisatie tot stand komt. Om een metaal met een sterke structuur te maken

wordt begonnen met het verhitten van het metaal, tot soms bijna aan het smeltpunt. De afzonderlijke atomen kunnen dan relatief vrij bewegen ten opzichte van elkaar. De atomen zijn dan ook tamelijk willekeurig verdeeld. Als het metaal afkoelt concentreren de atomen zich meer rond de plekken met een lagere energie. Als dit afkoelen zeer geleidelijk geschiedt komen de atomen uiteindelijk erg dicht bij de toestand van de laagste energie. Dit zorgt voor een zeer sterke structuur van het metaal.

Adiabatische quantum computation werkt op een vergelijkbare manier, het is voor het eerst beschreven door Farhi et al in 2000. De computer wordt geïnitieerd in de *grondtoestand* van een bekende *Hamiltonian*. Die Hamiltonian beschrijft de evolutie van de toestanden van de kwantumcomputer, en komt overeen met de temperatuur in simulated annealing. De grondtoestand komt overeen met een energetische evenwichtstoestand. De Hamiltonian wordt dan langzaam veranderd van een bekende begintoestand naar een eindtoestand die bij het probleem hoort. Als die verandering geleidelijk genoeg is zal de computer in een toestand blijven die dicht bij de langzaam veranderende grondtoestand ligt. De grondtoe-

stand van de uiteindelijke Hamiltonian is dan de oplossing van het probleem.

Toepassingen van deze methode zijn beschreven door Hogg en andere onderzoekers. Ze is zelfs al geïmplementeerd (zij het op een erg klein probleem) door Steffen en zijn medewerkers in 2003.

Verwacht het onverwachte

Hoewel al meer dan een halve eeuw oud lijkt operations research nog dikwijls een mengelmoes van slimme heuristische technieken. De toepassing komt op de eerste plaats, niet de diepgravende theorie. Als het goed werkt wordt het gebruikt.

Naar alle waarschijnlijkheid zullen de op *quantum computation* berustende methoden een vergelijkbare ontwikkeling doormaken. Het is alsof het weer 1950 is: we hebben al watertandend enkele voorbeelden gezien van wat mogelijk zou kunnen zijn, maar we kunnen slechts gissen naar de dingen die nog zullen komen.

Dürr, C., & P. Høyer (1996), A Quantum Algorithm for Finding the Minimum, *CoRR quant-ph*, 9607014: (

Farhi, E., J. Goldstone, S. Gutmann & M. Sipser (2000), Quantum Computation by Adiabatic Evolution, *arXiv:quant-ph/0001106v1*

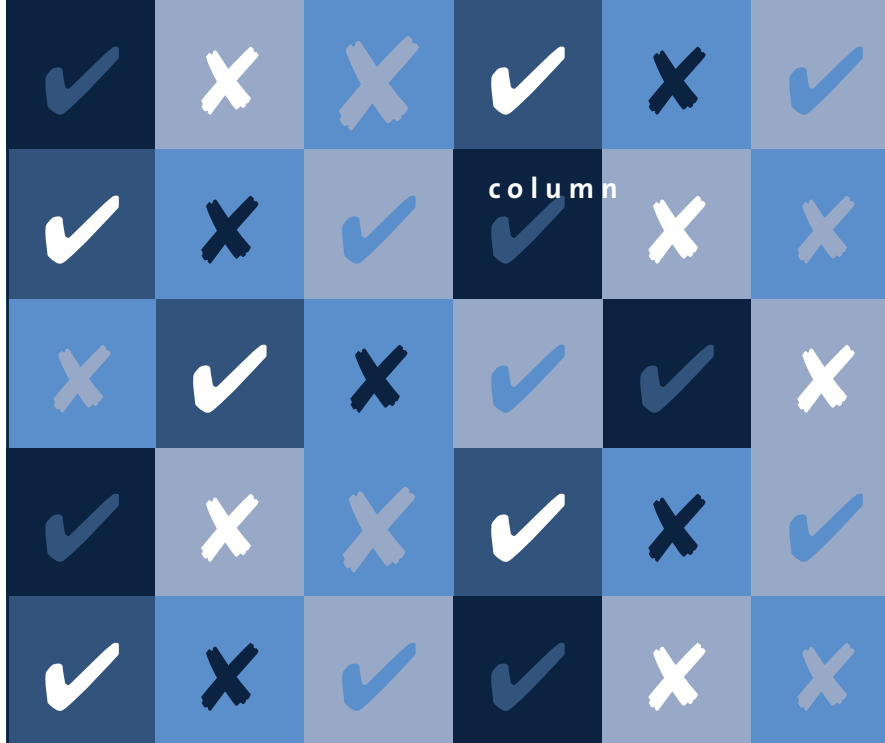
Grover, Lov K. (1996), A fast quantum mechanical algorithm for database search, *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp 212-219, May 22-24, 1996, Philadelphia, Pennsylvania, US [doi>10.1145/237814.237866

Shor, P.W. (1994), Algorithms for quantum computation: Discrete logarithms and factoring, *Proceedings 35nd Annual Symposium on Foundations of Computer Science*, IEEE Computer Society Press, pp 124-134.

Steffen, M., W. van Dam, T. Hogg, G. Breyta & I.

Chuang (2003), Experimental implementation of an adiabatic quantum optimization algorithm, *Physical Review Letters*, 90, 067903.

DAVID BULGER is docent en onderzoeker aan de Macquarie University in Sydney, Australië. zijn belangrijkste onderzoeksterreinen zijn de 'quantum computational optimisation algorithms' en bio-informatica. E-mail: <dbulger@efs.mq.edu.au>.



HET REFEREEPROCES

ONNO BOXMA

Een paar jaar geleden bood ik het artikel 'A brilliant analysis of the D/D/1 queue' ter publicatie aan bij het tijdschrift *Stochastics of LOW quality (SLOW)*. Ik was enthousiast over de verkregen resultaten, zoals wellicht doorschemert in de abstract, ook al was die met de mij zo kenmerkende bescheidenheid geschreven: 'This groundbreaking paper presents the first in-depth analysis of the waiting time distribution in the D/D/1 queue, a fundamental queueing model with deterministic arrivals and services. It rigorously proves that, in steady state, all waiting times are zero. This remarkable fact is confirmed via lengthy simulations.'

Natuurlijk werd dit juweeltje geaccepteerd, maar het refereeproces liet heel veel te wensen over, zoals u wel kunt opmaken uit deze passage van de acceptatiebrief: 'Helaas kostte het ons grote moeite iemand te vinden die bereid was een van uw onleesbare artikelen door te ploegen. Gisteren ontvingen wij bericht van de referee die

deze ondankbare taak wel op zich had genomen. Hij laat weten dat het verhaal wat hem betreft mag worden geaccepteerd. Hij heeft verder geen commentaar, omdat hij het eigenlijk niet gelezen heeft. Ik neem zijn aanbeveling graag over.'

Jammer genoeg is dit soort ervaringen niet uitzonderlijk, en is de e-mail die ik vorige week ontving dat wel: 'Geachte Professor Boxma, Deze email betreft uw artikel 'A few new views on queues', dat u vorige week heeft aangeboden aan het tijdschrift *Fundamental and Applied Stochastics (FAST)*. Bijgevoegd is een refereerapport. De referee merkt op dat u exact hetzelfde model analyseert als Professor Smarter in 1763, met het kleine verschil dat uw analyse fout is. De referee geeft ook aan hoe het allemaal veel algemener gekund had, en voegt een lijst van 5 bladzijden met typfouten toe. Op basis van zijn rapport heb ik helaas moeten besluiten, uw artikel af te wijzen.'

Kijk, zo'n e-mail willen we elke dag wel ontvangen: een scherp en duidelijk refereerapport, geproduceerd binnen een week!

Als onderzoekers hebben we allemaal voortdurend te maken met het refereeproces. De ervaringen zijn wisselend, maar het overheersende gevoel lijkt te zijn: ons peer review proces, met beoordelingen door collega's, is verre van perfect maar er is geen betere manier. Na bijna 35 jaar ervaringen met het refereeproces in alle mogelijke hoedanigheden weet ik nog steeds niet zeker hoe het peer review-proces kan worden verbeterd, maar durf ik het aan eens een wilde suggestie voor een verbetering te doen (sommige lezers zullen vermoeden dat ik in mijn 'naïeve periode' zit, en anderen weten dat ik altijd al naïef ben geweest; hoe dan ook, het voorstel is serieus bedoeld en ik ben oprecht geïnteresseerd in uw mening erover).

Plussen en minnen

Allereerst enige positieve constatering. Ondanks een voor velen toenemende werkdruk en pressie

om de kostbare tijd te besteden aan eigen onderzoek, is het nog steeds mogelijk bijna elk aangeboden verhaal door een stuk of twee mensen te laten beoordelen. Daarbij moet worden bedacht dat dit in ons vakgebied nagenoeg altijd zonder wezenlijke financiële tegemoetkomingen gaat; de beloning is dat men de wetenschappelijke vruchten plukt van het feit dat men – gedwongen – een artikel heel grondig doorploegt. De additionele morele reden om een artikel te beoordelen is dat we allemaal zelf ook artikelen aanbieden, en hopen dat anderen bereid zijn er een gedegen rapport over te schrijven.

Maar dan nu de negatieve constatering. De gemiddelde tijd tussen indiening van een artikel bij een tijdschrift en een eerste beslissing ligt in ons vakgebied ergens tussen de 6 en 12 maanden. Vooral voor promovendi is het buitengewoon vervelend dat zij soms een jaar moeten wachten op een rapport.

Soms is zo'n rapport dan ook nog eens heel oppervlakkig. Bij congressen met proceedings is de lange review termijn het probleem niet; de datum waarop het programmacomité een beslissing neemt ligt vast. Dat brengt extra tijdsdruk voor de referee met zich mee, en het percentage oppervlakkige rapporten is dan ook heel hoog. Ik beperk me in deze column graag tot het peer review-proces bij tijdschriften, ook al omdat congresproceedings in het vakgebied van STATOR geen overheersende rol spelen.

Een oplossing

1. Alle uitgevers van goede (ISI-lijst) tijdschriften in de stochastiek en besliskunde stellen een gezamenlijke database samen van auteurs en referees. Deze database bevat slechts één getal per persoon: zijn of haar 'krediet'.
2. Een persoon P die in de database wordt ingevoerd krijgt als startwaarde C kredietpunten.
3. Elke keer dat P een artikel indient bij een tijd-

schrift van het consortium, worden S punten afgetrokken van P 's totaal.

4. Elke keer dat P een artikel beoordeelt dat is ingediend bij een tijdschrift van het consortium, kent de betreffende hoofdredacteur een aantal (maximaal B) kredietpunten toe aan P . Dit geldt voor elk der reviewrondes.

Het aantal punten hangt af van het aantal maanden dat P nodig had om een rapport te produceren, en van de mate waarin het rapport nuttig was voor het nemen van een beslissing. Dit laatste is natuurlijk enigszins arbitrair maar een verantwoordelijkheid van de hoofdredacteur, die tevens de lengte en moeilijkheidsgraad van het artikel kan meewegen; een relatie tussen 'rapport is ingeleverd voor de gestelde deadline' en aantal punten is natuurlijk makkelijk te leggen.

En nu dan waar het om gaat.

5. Als P 's krediet onder nul komt, dan neemt geen enkel tijdschrift van het consortium een artikel van hem meer in behandeling – tot P weer krediet heeft opgebouwd door op bevredigende wijze refereewerk te doen.

Als P beduidend meer refereewerk doet dan hij artikelen aanbiedt, zal zijn krediet hoge waarden kunnen bereiken. Dat lijkt me ongewenst, vandaar:

6. Zodra P een krediet C_{\max} heeft bereikt, springt P 's krediet (eventueel: als P akkoord gaat) terug naar C . Ter compensatie bieden de verzamelde uitgevers P op dat moment gratis toegang tot bepaalde tijdschriften aan, of de aanschaf van boeken tegen een sterk gereduceerde prijs.

Overwegingen

Schrandere lezers zoals u zullen begrijpen dat we hier een model bespreken dat nauw verwant is aan dat van een verzekeringsbedrijf dat premies

int en uitbetalingen doet bij schadeclaims, en waarbij de kans op ruïnering bij start in C een belangrijk onderzoeksitem is.

Dat soort verzekeringsmodellen kan ons leren hoe groot de diverse parameters (startkrediet C , maximaal krediet C_{\max} , straf S voor het indienen van een artikel, beloning B voor refereewerk) moeten worden gekozen. Die keuzes zijn essentieel voor een goed functioneren van het systeem.

De samenwerking tussen diverse uitgevers kan een achilleshiel van het plan zijn. Maar de voordelen voor hen zijn evident: ze krijgen snellere en betere refererapporten. De gemiddelde auteur profiteert daar ook sterk van.

Verder ontmoedigt het systeem het verscheidene malen indienen van hetzelfde artikel. Te vaak zie je dat auteurs eerst hoog inzetten, en geleidelijk afzakken van de *Annals of Applied Probability* naar het niveau van de *Liechtenstein Journal of Trivial Probabilities* – een spoor van geïrriteerde referees achterlatend.

Tot slot ter illustratie de volgende, voor de verandering waargebeurde, anecdote. Ooit heb ik als redacteur van een tijdschrift een verhaal afgewezen, waarin reeds op pagina 2 de kansen groter dan 1 werden. Twee weken later diende de auteur, die zich door zo'n kleinigheid niet uit het veld liet slaan, het verhaal opnieuw in. De kansen bleven inderdaad inmiddels onder de 1, en wel zover dat ze op pagina 7 historisch lage waarden zoals -100 bereikten ...

Mocht de lezer kritiek hebben op het voorstel, of verbeteringen willen suggereren, dan houd ik me ten zeerste aanbevolen. Als u uw e-mail, gericht aan boxma@win.tue.nl, voorziet van de header 'uw prachtige voorstel', dan zal het niet verloren gaan tussen de ladingen andere fanmail.

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

Een revolutionaire methode om van een zelfselecte steekproef een goede steekproef te maken?

Open brief aan Wiebe Draijer (McKinsey & Company), Alexander Rinnooy Kan (SER) en Paul Schnabel (SCP) over de onderzoeksmethode van 21minuten.nl

Geachte heer Draijer, geachte heer Rinnooy Kan, geachte heer Schnabel,

Onlangs heeft u in de media veel aandacht gehad met het internetonderzoek van 21minuten.nl. Op deze website zijn honderdduizend enquêtes ingevuld door mensen die zichzelf geselecteerd hebben. Nu is bekend dat zo'n 'zelfselecte' steekproef een groot risico op vertekeningen heeft. In het algemeen is zo'n steekproef immers geen doorsnee van de bevolking. Door rekenwerk te doen, de technische term hiervoor is het 'herwegen van de data', meent u dat uw herwogen zelfselecte steekproef toch een goed beeld geeft. U gebruikt voor de herweging een kleinere steekproef, een internetpanel van 5.600 mensen die ook de vragenlijst van 21minuten.nl hebben ingevuld. Op internet gaat het gerucht dat uw methode is gebaseerd op een idee van Nobelprijswinnaar Daniel McFadden.

U schrijft dat u het grootste opinieonderzoek van Nederland heeft uitgevoerd. Maar groot is in de statistiek niet per se goed. In Amerika is in 1936 een onderzoek gehouden waarbij meer dan 2 miljoen mensen ondervraagd werden. Veel meer dus dan in uw onderzoek dat 0,1 miljoen ingevulde vragenlijsten had. Dit Amerikaanse onderzoek, de *Literary Digest presidential poll*, is beroemd geworden als voorbeeld van een groot maar slecht onderzoek.

Mijn eerste vraag aan u is: kunt u uitleggen waarom u denkt dat uw methode statistisch gezien goed is? Uitleg is nodig omdat u onbedoeld studenten in de war kunt brengen. Die leren in colleges statistiek dat het bij steekproeven niet primair gaat om een grote steekproef, maar om een goede, bijvoorbeeld een aselechte, steekproef. Een zelfselecte steekproef is over het algemeen een slechte steekproef. Waarom denkt u dat uw methode, die bestaat uit het herwegen van een zelfselecte steekproef, statistisch gezien goed is?

Mijn tweede vraag aan u is: is het een idee om de uitkomsten van beide steekproeven naast elkaar te zetten? U heeft twee steekproeven ondervraagd, een grote zelfselecte steekproef die zonder herweging mogelijk niet representatief is, en een kleine steekproef die naar u schrijft wel representatief is. In plaats van de kleine steekproef te gebruiken voor het herwegen van de grote is het misschien aardig om de uitkomsten te vergelijken. Dan kunnen we de verschillen zien. Is het naast elkaar laten zien van de uitkomsten uit uw twee steekproeven niet een beter idee dan herwegen?

Als u in gewone mensentaal een uitleg kunt geven, dan zou dat zeer prettig zijn. U heeft dan wel geen Nobelprijs, of moet ik zeggen *nog* geen Nobelprijs, maar u heeft wel de gave des woords (Wiebe), wiskundig inzicht (Alexander) en verstand van opinieonderzoek (Paul). Ik heb er alle vertrouwen in dat u met zijn drieën de bovenstaande twee vragen kunt beantwoorden.

Met vriendelijke groet,
Wim van Leeuwen, sociaal-wetenschappelijk onderzoeker, Amsterdam



Nederlands-Vlaamse Vereniging voor Ordinatie en Classificatie (VOC)

EVA CEULEMANS EN MARIEKE TIMMERMAN

De Nederlands-Vlaamse Vereniging voor Ordinatie en Classificatie (VOC) is opgericht in 1989, met als doel het bevorderen van communicatie en samenwerking tussen onderzoekers die geïnteresseerd zijn in de wetenschappelijke principes, methoden en toepassingen van ordinatie- en classificatietechnieken. Bij ordinatie kan men denken aan dimensie-reductie-technieken als principale componenten analyse en multidimensional scaling, maar zeker ook aan statistische modellen als item response modellen en random coëfficiënt modellen. Bij classificatie is het doel observatie-eenheden te clusteren. Ook hiervoor zijn vele technieken beschikbaar, variërend van k-means clustering tot support vector machines.

Dat ordinatie- en classificatietechnieken wijdverspreid van nut zijn, blijkt uit de wetenschappelijke achtergrond van de ongeveer 110 VOC-leden: Ze komen uit verschillende disciplines, zoals psychometrie, biologie, economie, sociometrie, chemometrie, sensometrie, geneeskunde, gezondheidswetenschappen, letteren en archeologie. De brede toepasbaarheid van ordinatie en classificatie blijkt ook uit de professionele achtergrond van de leden: Een behoorlijk deel van hen is niet verbonden aan een universiteit, maar werkt in de industrie of in overheidsdienst.

Om de kennis en verspreiding van ordinatie- en classificatietechnieken in Nederland en Vlaanderen te bevorderen organiseert de VOC twee maal per jaar een bijeenkomst in de vorm van een lezingendag. Vaak hebben deze bijeenkomsten een specifiek thema en worden er één of twee prominente binnen- dan wel buitenlandse sprekers uitgenodigd. Deze bijeenkomsten staan ook open voor niet VOC-leden en worden in de regel goed bezocht. Alle VOC-leden ontvangen twee maal per jaar een Nieuwsbrief met daarin onder meer het programma van de komende bijeenkomst, recensies van recent verschenen boeken uit het domein, informatie over interessante congressen en workshops, en referenties naar relevante publicaties van VOC-leden.

Belangrijke kenmerken van de VOC zijn openheid en laagdrempeligheid. Zoals gezegd is de vereniging multidisciplinair in de echte zin van het woord. Ze staat open voor zowel ontwikkelaars van ordinatie- en classificatietechnieken, als voor toepassers ervan. De lezingen hebben dan ook een theoretische en een toegepaste component. Bovendien wordt er bij de organisatie van de halfjaarlijkse bijeenkomsten expliciet gezocht naar onaangeboorde bronnen in de vorm van sprekers die de VOC-leden laten kennismaken

met nieuwe technieken en toepassingen daarvan. Een goed voorbeeld hiervan is de laatste VOC-bijeenkomst die plaatsvond in co-organisatie met de Belgisch-Nederlandse Vereniging voor Kunstmatige Intelligentie (BNVKI); deze bijeenkomst verruimde het blikveld van de VOC naar het domein van de artificiële intelligentie.

De VOC speelt ook een rol op het brede internationale vlak in die zin dat ze deel uitmaakt van de International Federation of Classification Societies (IFCS). De IFCS is een internationale wetenschappelijke organisatie die classificatieonderzoek in de ruime zin van het woord zoveel mogelijk wil stimuleren. Hiertoe ondersteunt de IFCS de publicatie van twee tijdschriften die artikelen publiceren over classificatie- en ordinatietechnieken: *Journal of Classification* (JoC) en *Advances in Data Analysis and Classification* (ADAC). Voorts organiseert de IFCS om het jaar een classificatiecongres. VOC-leden krijgen hierbij niet onaardige voordelen: Zij kunnen een abonnement nemen op JoC en ADAC tegen een gunsttarief en krijgen een behoorlijke reductie op het inschrijvingsgeld voor de IFCS-congressen. Voorts kunnen jonge VOC-leden een aanvraag indienen voor een IFCS travel award die een serieuze tegemoetkoming bij reis- en verblijfskosten impliceert.

Meer informatie over de VOC vindt u op <www.voc.ac> en kunt u ook verkrijgen bij de secretaris Marieke Timmerman, <m.e.timmerman@rug.nl>.

EVA CEULEMANS is docent bij het Centrum Methodologie Pedagogisch Onderzoek van de Katholieke Universiteit Leuven.

MARIEKE TIMMERMAN is universiteir hoofddocent bij de afdeling de Methoden & Technieken van Gegevensverwerking van de Rijksuniversiteit Groningen.

VVS-NIEUWS

Statistica Neerlandica

Sinds 1 oktober 2007 maakt *Statistica Neerlandica* deel uit van de *Blackwell Synergy OnlineEarly service*. Een artikel dat geaccepteerd is, wordt online gepubliceerd op <www.blackwell-synergy.com/OnlineEarly>. De publicatiedatum is die van de online-publicatie. Zodra de papieren publicatie verschijnt, wordt het artikel in SN-online opgenomen.

Beste biometrische publicatie

De ANed Biometry Award 2008 voor de beste biometrische publicatie van een Nederlandse auteur in 2006 en 2007 staat open voor inzendingen. Deadline is 1 februari 2008. Inzendingen als pdf naar <maarten.schipper@rivm.nl>. Voor meer informatie zie <www.bms-aned.nl>.

VVS-OR ereleden

Het bestuur van de vereniging maakt de leden attent op de mogelijkheid leden van de VVS-OR voor te dragen als erelid.

Boekenrecensies

Schrijf een boekrecensie voor *Kwantitatieve Methoden*. Op de website <www.vvs-or.nl> staat een lijst met boeken die bij de redactie van KM zijn binnengekomen.

Interessant

De website www.kennisbasisstatistiek.net is een Nederlandstalige website met veel informatie over statistiek voor onderwijs en onderzoek.

Gevraagd: goede ideeën

De VVS-OR heeft een bedrag van ruim 10.000 euro beschikbaar voor het populariseren van het vakgebied, bijvoorbeeld op middelbare scholen. Het bestuur staat open voor ideeën. Dit geld is afkomstig uit de boedel van de opgeheven sectie SSP.



OPNIEUW LUCIA DE B.

De zaak Lucia de B. blijft de gemoederen bezig houden binnen verschillende beroepsgroepen. De gerechtelijke veroordeling tot levenslang met gedwongen tbs-behandeling is in twijfel getrokken door juristen, medici, toxicologen, auteurs, en statistici en kansrekenaars. Een Advocaat-generaal bij de Hoge Raad heeft zich over het herzieningsverzoek van het College van procureurs-generaal gebogen en zijn conclusie is dat er nader onderzoek moeten worden verricht. Afhankelijk van de resultaten van het onderzoek zal worden beslist of een vordering tot herziening zal worden ingediend.

Richard Gill, voorzitter van de Nederlandse Vereniging van Statistiek en lid van de Koninklijke Nederlandse Academie der Wetenschappen, en Peter Grunwald, statisticus bij het Centrum voor Wiskunde en Informatica in Amsterdam boden op 2 november 2007 de minister van Justitie, Ernst Hirsch Ballin, een petitie aan. De petitie stelt dat het bewijs tegen De B. niet overtuigend is; er wordt geen uitspraak gedaan over schuld of onschuld van De B. De tekst van deze petitie is op de volgende bladzijde afgedrukt en ook te lezen op website <www.ipetitions.com/petition/lucia/index.html>. Voor de toelichting is geciteerd uit

het persbericht van de Universiteit Leiden van 1 november 2007. De petitie was bij indiening onder meer getekend door zo'n tachtig hoogleraren statistiek of kansrekening. Inmiddels (eind december 2007) is het totaal aantal ondertekenaars gestegen tot circa duizend.

Enkele hoogleraren zijn het blijkbaar niet eens met de tekst van de petitie. Volgens *Fokke en Sukke* is de kans dat vrijwel alle hoogleraren statistiek het eens zijn natuurlijk wel heel erg klein. Hoe groot deze kans is vermelden zij echter niet, hoewel uit een andere cartoon blijkt dat zij wel verstand van kansrekening denken te hebben. Wij als redactie hebben geprobeerd deze kans in te schatten. We gaan uit van een groep van 80 hoogleraren en een kans van 99 procent dat twee hoogleraren het eens zijn. De overige 78 moeten het weer met deze twee eens zijn. Dat levert dan op .99 tot de macht 79. Een kans van 45% dus, die wij niet zo heel erg klein vinden. Maar uiteraard zal niet iedere lezer het eens zijn met deze berekening en deze conclusie. Als zelfs bij zo'n eenvoudige berekening de meningen al kunnen verschillen, dan zal het laatste woord in de zaak Lucia de B zeker nog niet gesproken zijn.

PETITIE HEROPENING ZAAK LUCIA DE B.

'De zaak Lucia de B. moet zo spoedig mogelijk worden heropend.'

'The case of Lucia de B. (life imprisonment in the Netherlands for 7 alleged murders and 3 alleged attempts) must be reopened as soon as possible.'

Zie ook: <www.ipetitions.com/petition/lucia/index.html>.

Gezaghebbende internationale wetenschappers en medici pleiten voor heropening van de zaak tegen Lucia de B.

Zij stellen dat essentiële medische en statistische inzichten niet bij de rechter terecht zijn gekomen. Dat blijkt ook uit het rapport van de Commissie Evaluatie Afgesloten Strafzaken, die haar bevindingen presenteerde na een jaar van onderzoek. De commissie constateert dat er onenigheid is tussen deskundigen maar deed geen expliciete uitspraak over wie er gelijk heeft. De indieners van de petitie Richard Gill en Peter Grunwald nemen die stap wel, zij stellen dat:

- alle competente binnen- en buitenlandse statistici die zich in de zaak verdiept hebben, tot dezelfde conclusie zijn gekomen: er zijn geen sterke statistische aanwijzingen dat De B. iets met de sterfgevallen te maken heeft gehad;
- dat twee van 's werelds meest competente medische experts op het gebied van digoxine stellen dat Baby A. niet aan een digoxinevergiftiging kan zijn overleden;
- doordat bij de veroordeling van De B. een 'schakelbewijs' is gebruikt (op zichzelf een verkapt vorm van statistiek), hiermee ook het bewijs wegvalt voor de andere 'moorden' waarvoor De B. is veroordeeld.

Uit: Persbericht Universiteit Leiden, 1 november 2007

AGENDA

4 - 7 maart 2008

GOCPS 2008 provides an international forum for presentation of new results in probability and statistics. Participants from universities, business, administration, and industry are welcome. Website <<http://gocps2008.rwth-aachen.de>>.

27 maart 2008

De Dag voor Statistiek en Besliskunde 2008 zal wederom op het CBS in Voorburg plaatsvinden, en wel op donderdag 27 maart. Sprekers en overige informatie staan binnenkort op de site <www.vvs-or.nl>. Noteer deze dag alvast in uw agenda.

26 - 28 juni 2008

We would like to invite you to participate in the **Workshop on Nonparametric Inference – WNI2008**, which will be held in Coimbra, Portugal, on June 26-28, 2008. For more information please see the website <www.mat.uc.pt/~wni2008>.

7 - 11 juli 2008

De 23e International Workshop on Statistical Modelling vindt van 7 t/m 11 juli 2008 plaats in Utrecht. Deadline voor het insturen van abstracts: 10 februari 2008. Zie: <www.fss.uu.nl/iwsm2008>.

17 - 21 augustus 2008

On behalf of ISCB and the LOC it is a great pleasure to invite you to attend the 29th annual conference in Copenhagen, 17-21 August 2008. Please visit <www.iscb2008.info> for further details.

27 - 28 augustus 2008

Oprichters van EURANDOM, voormalige postdocs en vooraanstaande sprekers vanuit industrie en wetenschap blikken terug op 10 jaar EURANDOM en kijken vooruit naar nieuwe ontwikkelingen. Voor inlichtingen, zie <www.eurandom.nl>.