

STAtOR

periodiek van de VVS jaargang 8 nummer 2, juni 2007

De analyse van empirische data met
behulp van informatieve hypothesen

Optimalisering in moderne
telecommunicatie netwerken

Rolgordijnen en Operations Research:
Theorie en Praktijk

Black is beautiful

De complexiteit van de
Vereenvoudigde Arbeidstijdenwet

Randomized Response Technieken
voor sociaal gevoelige vragen in
beleidsonderzoek

Christiaan Huygens

STATOR

Jaargang 8, nummer 2, juni 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 (hoofddirecteur), Ana Isabel Barros, Mirjam Moerbeek, Gerrit Stemerdink (eindredacteur), Fred Steutel, Hilde Tobi, Marnix Zoutenbier.

Kopij en reacties richten aan

Prof. dr. G. Kant (hoofddirecteur), 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: A.W. van der Vaart <aad@cs.vu.nl>
Penningmeester: S. J. Koopman <s.j.koopman@feweb.vu.nl>
Namens de Bedrijfssectie (BDS):
P. Banens <banens@cqm.nl>
Namens de Biometrische Sectie (BMS):
A.H. Zwinderman <a.h.zwinderman@amc.uva.nl>
Namens de Economische Sectie (ECS):
P.H.F.M. van Casteren <casteren@fee.uva.nl>
Namens het Ned. Genootschap voor Besliskunde (NGB):
J. van de Klundert <j.vandeklundert@math.unimaas.nl>
Namens de Sectie Mathematische Statistiek (SMS):
P. Spreij <spreij@science.uva.nl>
Namens de Sociaal Wetenschappelijke Sectie (SWS):
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** 'Wat een interessant vak'
- 4** De analyse van empirische data met behulp van informatieve hypothesen
Herbert Hoijsink
- 8** Optimalisering in moderne telecommunicatie netwerken
Stan van Hoesel
- 13** Rolgordijnen en Operations Research: Theorie en Praktijk
Edgar de Gelder
- 18** Black is beautiful - column
Fred Steutel
- 20** De complexiteit van de Vereenvoudigde Arbeidstijdenwet
Laurens Fijn van Draat, John Poppelaars en Gerhard Post
- 25** Randomized Response Technieken voor sociaal gevoelige vragen in beleids-onderzoek
Gerty Lensvelt-Mulders
- 30** Christiaan Huygens - column
Onno Boxma
- 32** Agenda



‘Wat een interessant vak’

Wie het trefwoord ‘stator’ intikt bij Google krijgt 32.700 Nederlandse pagina’s, die maar voor klein deel naar STATOR verwijzen. Een groot aantal verwijst naar een bedrijfshulpverleningsfirma met onze naam, iets anders genoteerd, maar een veel groter aantal, het merendeel, bevat informatie over wat je een stilstaande rotor zou kunnen noemen, in feite een permanente magneet, van belang bij sommige electromotoren. De aarde al geheel is ook een stator.

Is het blad STATOR dan niet populair? Toch wel, er zijn ook tientallen sites die wel over ons gaan. Bovendien er komen regelmatig reacties binnen bij auteurs of redactieleden, die bijval of andere vormen van tevredenheid uitdrukken in verband met de inhoud van ons blad. Zo betuigden twee medewerkers van een zuivelbedrijf hun blijdschap met een artikel over proefopzetten. Dat was precies wat ze nodig hadden, konden ze direct toepassen. Een jurist, advocaat, schreef: ‘Al enige tijd lees ik met veel plezier de stukjes van STATOR. Wat een interessant vak ‘kansrekening’. Ik heb me nooit gerealiseerd op hoeveel terreinen kansrekening toegepast kan worden. Heel leuk geschreven en veel nieuwe gezichtspunten voor een leek op dat terrein. Ik hoop ze nog lang aan te treffen.’

Zou *Statistica Neerlandica* ook dat soort bijval ten deel vallen? Ik denk van niet, maar wij hebben het ook een stuk makkelijker. Statistiek en Operations Research zijn beide voor een groot deel toepassingen van de kansrekening, op haar beurt onderdeel van de wiskunde. Terwijl wij voornamelijk aansprekende en makkelijk te begrijpen toepassingen laten zien, moet ook de theoretische kant van STA en OR ergens aan de orde komen. Dat gebeurt in *Statistica Neerlandica*, dat bovendien het nadeel heeft dat mensen die een ‘doorbraak’ te melden hebben dat liever in een bekend buitenlands tijdschrift doen dan in een Nederlands blad, ook al is dat geheel Engelstalig. De VVS moet *Statistica* in stand en liefst in ere houden, maar dat is niet in de eerste plaats onze opgave.

Wij moeten voor leesbare stukken zorgen, waar toepassers iets aan hebben en waar ook de geïnteresseerde leek iets van kan leren. Afgaande op de reacties zijn we daar de afgelopen zes-en-een-half jaar redelijk in geslaagd. Reden tot een bescheiden tevredenheid.

De redactie



De analyse van empirische data met behulp van informatieve hypothesen*

HERBERT HOIJTINK

Nul-, alternatieve en informatieve hypothesen

Veel onderzoekers maken gebruik van hypothesetoetsing bij de beantwoording van hun onderzoeksvragen. Veel tijdschriften verwachten een significant toetsingsresultaat voordat ze publicatie van een wetenschappelijk artikel willen overwegen. Toch is het de vraag of het klassieke vergelijk van een nul-hypothese ('er gebeurt niets') met een alternatieve hypothese ('er gebeurt iets maar ik weet niet wat') veel bruikbare informatie oplevert. Cohen (1994) heeft het probleem prachtig gekarakteriseerd: de significantie van de nul-hypothese 'de aarde is rond' zal bij een steekproef van voldoende grootte kleiner zijn dan .05. De conclusie zal dus worden 'de aarde is niet rond', waarmee we in essentie niets bruikbaars geleerd hebben. Er kan meer worden geleerd als er informatieve hypothesen ('dit gebeurt er') met betrekking tot de toestand in een populatie worden opgesteld, om vervolgens met behulp van empirische data de steun voor elke informatieve hypothese te bepalen. In het geval van de aarde zou dit, in volgorde van anciënniteit kunnen leiden tot de informatieve hypothesen 'de aarde is plat', 'de aarde is rond' en 'de aarde is ovaal'.

Binnen de sociale wetenschappen kunnen informatieve hypothesen vaak worden geconstrueerd door statistische modellen nader te specificeren met behulp van ongelijkheidsrestricties.

Lucas (2003) beschrijft een experiment waarin hij meet hoe competent de deelnemers hun leider vinden. Er zijn vijf condities:

1. Een groep waarin *at random* een man tot leider wordt benoemd.
2. Een groep waarin *at random* een vrouw tot leider wordt benoemd.
3. Een groep waarin op grond van bewezen vaardigheden een man tot leider wordt benoemd.
4. Een groep waarin op grond van bewezen vaardigheden een vrouw tot leider wordt benoemd.
5. Een groep die zo is bewerkt dat een vrouwelijke leider normaal en legitiem wordt gevonden (institutionalisering), en waarin vervolgens een vrouw tot leider wordt benoemd.

In Tabel 1 worden de competentiescores nader beschreven. Lucas (2003) formuleert onder andere de volgende drie deelhypothesen met betrekking tot de gemiddelde competentiescores (μ_g) in de vijf experimentele groepen ($g=1,\dots,5$):

Groep	M	SD	N
1	2.33	1.86	30
2	1.33	1.15	30
3	3.20	1.79	30
4	2.23	1.45	30
5	3.23	1.50	30

Tabel 1: Competentiescores: gemiddelde (M), standaarddeviatie (SD) en steekproefgrootte (N)

1. Het sexe-effect: $\mu_1 > \mu_2$ en $\mu_3 > \mu_4$
2. Het benoemingseffect: $\mu_1 < \mu_3$ en $\mu_2 < \mu_4$
3. Het institutionaliseringseffect: $\mu_5 > \{\mu_1, \mu_4\}$

Het institutionaliseringseffect zegt dat de competentiescores van geïstitutionaliseerde vrouwelijke leiders groter is dan die van vrouwen die op grond van bewezen vaardigheden tot leider zijn benoemd, en groter dan die van mannen die *at random* tot leider zijn benoemd. Dit lijkt een vreemde hypothese omdat groep 2 en 3 er niet bij zijn betrokken. Echter, dit lost zich op door naar de combinatie van deelhypothese 1. en 3. te kijken. Dan wordt duidelijk dat geïstitutionaliseerde vrouwelijke leiders hogere competentiescores hebben dan vrouwen die op grond van bewezen vaardigheden danwel *at random* zijn benoemd, ook hogere scores dan mannen die *at random* zijn benoemd, maar niet noodzakelijkerwijs hoger dan mannen die op grond van bewezen vaardigheden zijn benoemd (mannelijk leiderschap is namelijk van nature al geïstitutionaliseerd).

Lucas (2003) is geïnteresseerd in elke deelhypothese afzonderlijk en in combinaties van deze hypothesen wat leidt tot een totaal van zeven niet-geneste informatieve hypothesen (zie Tabel 2).

Model	Posterior kans
Sexe-effect	.05
Benoemingseffect	.05
Institutionaliseringseffect	.04
Sexe en Benoeming	.15
Sexe en Institutionalisering	.17
Benoeming en Institutionalisering	.17
Sexe, Benoeming en Institutionalisering	.37

Tabel 2: Modellen en posterior kansen

Anraku (1999), Klugkist, Laudy en Hoijtink (2005) en Klugkist, Kato en Hoijtink (2005) beschrijven informatiecriteria die gebruikt kunnen worden om de beste van deze sets van (gecombineerde)

hypothesen te selecteren. De eerste gebruikt hiervoor een aanpassing van Akaike's informatie-maat met een *penalty* die rekening houdt met de ongelijkheidsrestricties, de laatsten gebruiken de Bayes factor (Kass en Raftery, 1995) waarin de ongelijkheidsrestricties automatisch worden verdisconteerd. Lucas (2003) was niet op de hoogte van deze mogelijkheden en heeft zijn conclusies omtrent de (combinaties) van hypothesen noodgedwongen gedestilleerd uit paarsgewijze vergelijkingen van gemiddelden.

Ook buiten de hierboven geschetste *one-way* anova context kunnen informatieve hypothesen worden geformuleerd met behulp van ongelijkheidsrestricties: Kato en Hoijtink (2006) baseren zich op multi-level modellen; Laudy en Hoijtink (2007) op modellen voor kruistabellen; en Laudy, Zoccolillo, Baillargeon, Boom, Tremblay, en Hoijtink (2005) op latente klasse modellen. In al deze publicaties wordt de Bayes factor gebruikt om de beste van de competitieve informatieve modellen te selecteren. In de volgende sectie zal nader op de Bayes factor worden ingegaan.

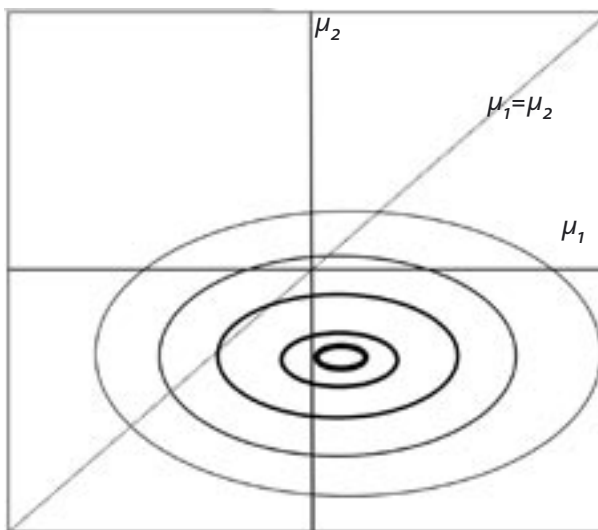
Informatieve hypothesen en à priori verdelingen

Gebruikmakend van Bayes' theorema kan de Bayes factor (Kass and Raftery, 1995) als volgt worden gedefinieerd:

$$BF_{om} = \frac{p(x|H_o)}{p(x|H_m)} = \frac{p(x|\theta)p(\theta|H_o)/p(\theta|x,H_o)}{p(x|\theta)p(\theta|H_m)/p(\theta|x,H_m)} \quad (1)$$

waarbij H_o een ongerestricteerd model is, en H_m voor $m=1, \dots, M$ een ongelijkheidsgerestricteerd model. Zoals gebruikelijk representeert x de data en θ de modelparameters. Zoals kan worden gezien is de Bayes factor niets anders dan de ratio van de kans op de data gegeven respectievelijk H_o en H_m .

Als θ een waarde heeft die onder H_m is toegestaan, dan verdwijnt de waarschijnlijkheid $p(x|\theta)$ uit de teller en noemer van (1). Als encompassing priors (Klugkist, Kato en Hoijsink, 2005) worden gebruikt kan (1) nog verder worden vereenvoudigd. Voor een simpel model waarin slechts twee gemiddelden een rol spelen staat de encompassing prior $p(\theta|H_o)$ weergegeven in Figuur 1: het is een niet informatieve uniforme verdeling, dat wil zeggen een verdeling die vergeleken met de informatie in de data weinig tot geen informatie over de modelparameters θ bevat. De prior verdeling voor de parameters van het gerespecteerde model H_m valt eenvoudig uit de encompassing prior af te leiden: het is simpelweg dat deel van de encompassing prior in overeenstemming met de restricties en vermenigvuldigd met een constante C_m zodanig dat $\int_{\nu} p(\theta|H_m)d\theta = 1$.



Figuur 1: Priors en posteriors

In Figuur 1 is voor $H_m: \mu_1 > \mu_2$ de prior verdeling gelijk aan de onderdriehoek vermenigvuldigd met een constante $C_m=2$. Meer formeel kan worden gesteld dat voor een waarde van θ die is toegestaan onder H_m geldt dat $p(\theta|H_m) = C_m p(\theta|H_o)$.

In Figuur 1 representeren de ellipsen de isodensity contouren van de posterior verdeling van μ_1 en μ_2 onder het niet gerespecteerde model, en dezelfde contouren in de onderdriehoek de posterior verdeling voor $H_m: \mu_1 > \mu_2$. Ook hier geldt in het algemeen dat, voor waarden van θ die zijn toegestaan onder H_m , de posterior kansdichtheid van het gerespecteerde model proportioneel is aan die van het ongerespecteerde model:

$$p(\theta|x, H_m) = d_m p(\theta|x, H_o).$$

Met behulp van de afleidingen in de vorige alinea kan (1) worden herschreven tot

$$BF_{om} = d_m / c_m, \quad (2)$$

waarbij $1/d_m$ gelijk is aan de proportie van de ongerespecteerde posterior in overeenstemming met de restricties van model H_m , en $1/c_m$ de proportie van de ongerespecteerde prior in overeenstemming met de restricties. Dit leidt tot een eenvoudige schatter van BF_{om} : trek een steekproef uit zowel de ongerespecteerde posterior als prior, en gebruik deze om d_m en c_m te schatten. Deze steekproeven zijn eenvoudig te verkrijgen met behulp van Markov Chain Monte Carlo methoden (zie bijvoorbeeld Gelman, Carlin, Stern en Rubin, 2004).

De Achilleshiel van de Bayesiaanse statistiek is dat inferenties afhankelijk zijn van de keuze van de a priori verdelingen. Voor een brede klasse van ongelijkheidsgerespecteerde modellen kan worden aangetoond dat dit niet het geval is (Klugkist, Kato en Hoijsink, 2005). Dit kan aan de hand van Figuur 1 worden geïllustreerd: ongeacht de grenzen van de prior verdeling is $1/c_m$ altijd gelijk aan .50; en als de data de prior domineren (dat is vaak zo) dan is de posterior verdeling nagenoeg onafhankelijk van de gekozen prior en hangt ook $1/d_m$ niet af van de grenzen van de prior. Kortom, het

selecteren van de beste van een brede klasse van ongelijkheidsgerestricteerde modellen is nagenoeg objectief, dat wil zeggen onafhankelijk van de gekozen prior (mits deze niet informatief is voor het ongerestricteerde model).

Als wordt aangenomen dat a priori elk model even waarschijnlijk is, kan de posterior kans van elk model als volgt berekend worden:

$$p(H_m|x) = \frac{1/BF_{om}}{1/BF_{o1} + \dots + 1/BF_{oM}} \quad (3)$$

Voor de leiderschap data staan de resultaten weergegeven in Tabel 2. Zoals kan worden gezien zijn elk van de drie deelhypotheses even waarschijnlijk, combinaties van twee ook, en heeft het model waarin alle door Lucas (2003) voorspelde effecten zijn opgenomen de hoogste a posteriori kans.

Hoe verder?

In het voorgaande is aan de hand van een relatief eenvoudig voorbeeld geïllustreerd dat een grote klasse van informatieve hypothesen met behulp van de Bayes factor nagenoeg objectief kan worden geëvalueerd. Dankzij een NWO-VICI subsidie kan deze benadering de komende jaren verder worden onderzocht en uitgebouwd. Daarbij zal aandacht zijn voor: een wetenschapsfilosofische onderbouwing van Bayesiaanse modelselectie; het gebruik van ongelijkheidsrestricties in het multivariaat lineaire model en lineair structurele modellen; prior specificatie met behulp van training data (Berger en Perricchi, 2004) voor situaties waarin de eerder geschetste objectiviteit niet geldt – dit is bijvoorbeeld het geval voor restricties van het type $\mu_1 = \mu_2 = \dots$ en $\mu_g > k$ waarbij k een getal is; en het vergelijken van de Bayesiaanse benadering met hypothesetoetsing, uitbreidingen van de in Anraku (1999) geschetste benadering en klas-

sieke en Bayesiaanse predictieve benaderingen van modelselectie.

De geïnteresseerde lezer wordt verwezen naar <www.fss.uu.nl/ms/hoijtink> voor verdere informatie en software voor de analyse van anova designs en kruistabellen met behulp van ongelijkheidsrestricties.

LITERATUUR

- Anraku, K. (1999). An information criterion for parameters under a simple order restriction. *Biometrika*, 86, 141-152.
- Berger, J.O. and Perricchi, L. (2004). Training samples in objective Bayesian model selection. *Annals of Statistics*, 32, 841-869.
- Cohen, J. (1994). The Earth is Round ($p < .05$). *American Psychologist*, 49, 997-1003.
- Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2004). *Bayesian Data Analysis*. London: Chapman and Hall.
- Kass, R.E. and Raftery, A.E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773-795.
- Kato, B.S. and Hoijtink, H. (2006). A Bayesian approach to inequality constrained hierarchical models: Estimation and Model selection. *Statistical Modelling*, 6, 1-19.
- Klugkist, I., Laudy, O. and Hoijtink, H. (2005). Inequality Constrained Analysis of Variance: A Bayesian approach. *Psychological Methods*, 10, 477-493.
- Klugkist, I., Kato, B. and Hoijtink, H. (2005). Bayesian Model Selection Using Encompassing Priors. *Statistica Neerlandica*, 59, 57-69.
- Laudy, O., Zoccolillo, M., Baillargeon, R., Boom, J., Tremblay, R. and Hoijtink, H. (2005). Applications of confirmatory latent class analysis in developmental psychology. *European Journal of Developmental Psychology*, 2, 1-15.
- Laudy, O. and Hoijtink, H. (2007). Bayesian methods for the analysis of inequality constrained contingency tables. *Statistical Methods in Medical Research*.
- Lucas, J.W. (2003). Status processes and the institutionalization of women as leaders. *American Sociological Review*, 68, 464-480.

NOOT

* Onderzoek gesubsidieerd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO-VICI-453-05-002).

HERBERT HOIJTINK is hoogleraar Toegepaste Bayesiaanse Statistiek bij het departement Methoden en Technieken van de Universiteit Utrecht. E-mail: <h.hojtink@uu.nl>



Foto: Pieter Bosch

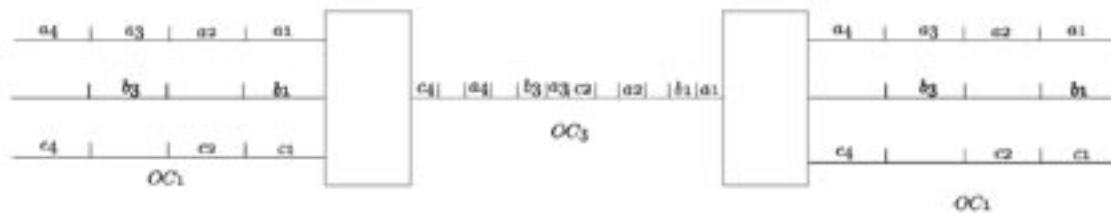
OPTIMALISERING in moderne telecommunicatie-netwerken

Het klassieke probleem in telecommunicatie is het ontwerp van netwerken: waar leggen we verbindingen en hoeveel capaciteit moeten deze hebben om aan de vraag te voldoen en voldoende bescherming tegen uitval te bieden. Hierin opgesloten ligt nog het probleem van het routeren van vraag over de verbindingen.

STAN VAN HOESEL

De overgang van analoge naar digitale transmissie heeft reeds diverse nieuwe problemen opgeleverd, waaronder het optimaal plaatsen van apparaten die signalen kunnen samenvoegen (en scheiden), zogenoemde multiplexers, in een netwerk. De introductie van licht als transmissie medium naast electriciteit heeft niet alleen nieuwe varianten van het multiplexer probleem opgeleverd (licht

kan in meerdere golflengtes over één glasfiber verzonden worden). Met de introductie van apparatuur die een signaal van golflengte kan doen veranderen, zogenoemde *wavelength converters* is weer een nieuw probleem ontstaan. Deze nieuwe problemen vertalen zich direct naar kleuringsproblemen op grafen. Deze problemen en hun oplossingsmethoden willen we hier bespreken.



Figuur 1: Time division multiplexing

De klassieke problemen.

Het ontwerp van een netwerk is in essentie niets meer dan in een graaf $G=(V,E)$ een aantal kanten te kiezen zodanig dat aan bepaalde connectiviteitseisen voldaan wordt. De connectiviteitseisen hangen af van de plaats in het netwerk: lokaal is 1-samenhang vaak genoeg. In de (inter)nationale netwerken is minstens 2-samenhang vereist, in verband met de problemen die ontstaan als ergens breuk of uitval optreedt. Een recent voorbeeld van grootschalige uitval is een aardbeving in de Chinese zee, waardoor diverse overzeese verbindingen in Azië uitvielen. De volgende stap is het bepalen van de hoeveelheid capaciteit, doorgaans gebaseerd op verwachte vraag tussen paren punten. Waar vroeger de capaciteit gemeten werd in aantallen circuits wordt dat nu door aantallen Megabits per seconde gedaan. Hierbij worden routes van het telecomverkeer vaak voor vast aangenomen bijvoorbeeld bepaald door middel van kortste pad algoritmen. Een leuk nieuw voorbeeld is de *Open Shortest Path First* (OSPF) routing in IP netwerken. Dit mechanisme bepaalt de route van een IP pakketje door de kortste weg in een graaf te bepalen, waarbij de lengtes van de verbindingen bepaald worden door de netwerkbeheerder. Deze zal dat zodanig doen dat het verkeer zo weinig mogelijk *bottlenecks* oplevert. In [5] worden algoritmen gegeven die de lengtes van de verbindingen zo goed mogelijk bepalen aan de hand van de vraag. Nagenoeg alle 'oude' problemen zijn varianten van *multi-commodity*

flow. Er wordt nog steeds veel onderzoek gedaan naar efficiënte oplossingsmethoden voor deze problemen. Surveys zijn [10] en recent [6].

Moderne problemen in op licht gebaseerde netwerken

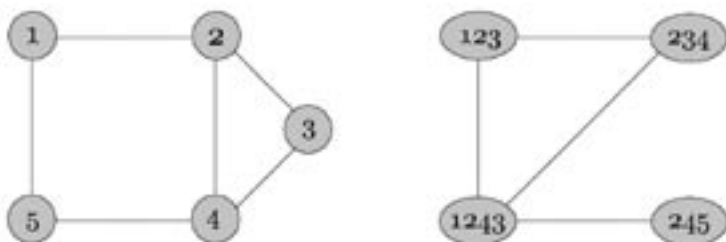
In digitale netwerken is voor het eerst het probleem van het samenvoegen van signalen bekeken, het zogenaamde multiplexen. Dit werd gedaan door meerdere (typisch drie of vier) signalen van een lage snelheid in een signaal van hoge snelheid te combineren. Dit gebeurt door middel van *Time Division Multiplexing* (TDM), waarbij de langzame signalen tijdsloten krijgen in het snelle signaal.

De introductie van licht in glasfiber heeft enorme hoeveelheden capaciteit tegen relatief lage prijzen beschikbaar gemaakt. In de glasfiber netwerken die tegenwoordig overal in de grond liggen worden probleemloos snelheden van tientallen Gbits per seconde bereikt. Daarnaast is het mogelijk om verschillende lichtfrequenties tegelijkertijd door een fiber te sturen. Alhoewel tegenwoordig honderden frequenties simultaan verstuurd kunnen worden, wordt er voornamelijk gewerkt met 16 frequenties per fiber, ieder met een snelheid van 2.5 Gbit/s. Met zogenaamde ADMs (*Add-Drop Multiplexers*) kunnen op een verbinding frequenties toegevoegd of uitgefilterd worden. Dit idee heet *Wavelength Division Multiplexing* (WDM). WDM is weliswaar technisch anders dan

TDM, het optimaliseringsvraagstuk van waar multiplexers te plaatsen en welke signalen samen te voegen is nauwelijks verschillend.

We bekijken het volgende probleem: als in een netwerk de vraag naar bandbreedte tussen diverse paren punten bekend is (in eenheden van 2.5 Gb/s) en de route van vraag tussen twee eindpunten vastligt, dan is het probleem om te kijken welke paden met elkaar in één fiber gecombineerd kunnen worden. Merk op dat twee paden in dezelfde fiber niet van dezelfde frequentie gebruik kunnen maken. In grafen terminologie: Gegeven een graaf $G=(V,E)$ en paden $P_1, \dots, P_K$ in G . Gevraagd een kleuring (met zo weinig mogelijk kleuren) van de paden zodanig dat paden die een verbinding gemeen hebben verschillende kleuren gebruiken. Dit probleem is in feite een standaard kleuring probleem. De transformatie gaat als volgt. De punten van de nieuwe graaf worden gevormd door de paden. Twee paden worden verbonden door een kant als ze in G een verbinding gemeen hebben. De zo geconstrueerde graaf wordt ook wel conflictgraaf genoemd: conflicten tussen twee objecten worden gemodelleerd door middel van kanten. Merk op dat we in feite ongericht verkeer beschouwen. In de modellen is dat niet echt relevant, omdat gericht verkeer gemakkelijk als ongericht verkeer gemodelleerd kan worden bij vaste routes.

In onderstaand voorbeeld worden, in een netwerk van vijf punten, vier paden geïdentificeerd



Figuur 2: Constructie van de conflictgraaf

aan welke een kleur toegekend dient te worden. Dit zijn de paden 123, 234, 245 en 1243. De bijbehorende conflictgraaf is ook gegeven.

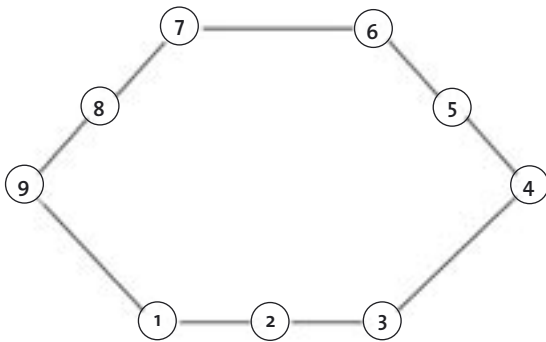
Merk op dat in dat geval routes tussen puntenparen per definitie vastliggen. Speciale netwerken (topologieën) vinden veel toepassing in de huidige netwerken. De boom is op laag niveau populair, en daarnaast wordt de ring veel gebruikt in hogere lagen van netwerken. Het probleem is NP-moeilijk, in beide topologieën, zie [2] en [3]. De ring is een populaire structuur in SONET/SDH netwerken, doordat deze 2-samenhangend is. De zogenoemde *Self-Healing Ring* wordt dan ook in de meeste SONET/SDH netwerken gebruikt. Er zijn niet of nauwelijks oplossingsmethoden voor dit probleem bekend. In [2] worden *greedy* heuristieken besproken, die als voornaamste doel hebben het gat tussen het gebruikte aantal kleuren en de maximum *load* (aantal signalen over een



Figuur 3. Toplaag SDH netwerk in Nederland.

verbinding, een ondergrens) zo klein mogelijk te houden.

Het probleem wordt een stukje moeilijker als ook nog de route van het verkeer (het pad) bepaald moet worden. Tegelijkertijd dienen nog steeds kleuren aan de paden toegekend te worden zodanig dat geen conflicten optreden. Er zijn nu verschillende optimaliseringscriteria mogelijk: het minimaliseren van het aantal gebruikte kleuren en het minimaliseren van het aantal ADMs (elke keer dat een kleur toegevoegd of verwijderd wordt dient een ADM geplaatst te worden). Onderstaand voorbeeld geeft aan dat de criteria 'minimum aantal kleuren' en 'minimum aantal ADMs' niet simultaan kunnen worden geoptimaliseerd.

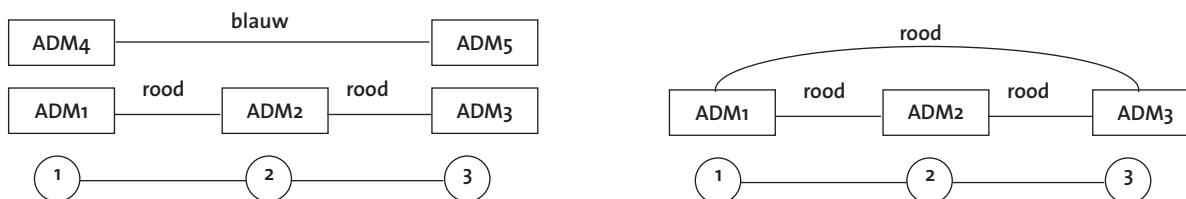


Figuur 4. Kleuren versus ADMs.

Er dient een verbinding te zijn tussen ieder paar in de puntverzameling {1,2,3}, in de puntverzameling {4,5,6}, en in de puntverzameling {7,8,9}. Dit kan op twee manieren: eerst direct 12, 23, 13 etc. Dit kost

per groep van drie punten twee kleuren en 5 ADMs. In totaal dus twee kleuren en 15 ADMs. Een alternatief is om de verbinding tussen 1 en 3 via de route 34567891 te laten lopen. Nu kan deze groep toe met één kleur. Dit heeft als voordeel dat er voor deze groep slechts 3 ADMs nodig zijn. Als nadeel geldt dat iedere groep nu zijn eigen kleur heeft en dat er dus 3 kleuren nodig zijn en in totaal 9 ADMs.

Een relaxatie van het kleuringsprobleem is de minimalisatie van de grootste load per verbinding. De load is een voor de hand liggende ondergrens van het aantal benodigde kleuren. Ook dit probleem is NP-moeilijk, met als interessante uitzondering de ring indien de vraag in kleine eenheden is. Dit probleem is polynomiaal oplosbaar, zie [12]. Hier wordt ook een approximatie algoritme besproken voor algemene vraag in ringen, met absolute garantie voor de afwijking van het optimum. Het laatste probleem dat we willen bespreken is een uitbreiding op bovenstaande problemen: we kunnen zogenaamde *light-wave converters* toevoegen, zie [4]. Deze apparaten kunnen een de golflengte van een gegevensstroom wijzigen. In grafenterminologie: we kunnen een pad opsplitsen en ieder deel zijn eigen kleur geven. Dit kan zowel winst opleveren in aantal benodigde kleuren als aantal benodigde ADMs. De kostenstructuur van netwerken is zodanig dat we zo weinig mogelijk gebruik willen maken van apparatuur en dat dus het terugbrengen van het aantal ADMs de voorkeur heeft.



Figuur 5. De twee oplossingen

Voor alle drie genoemde problemen zijn niet of nauwelijks goede algoritmen bekend op grote structuren. Er zijn wat NP-moeilijkheids resultaten en wat approximatie-algoritmen in bovengenoemde literatuur. Daarnaast behandelt [8] enkele modellen en [9] ondergrenzen gebaseerd op LP. Uitdagingen zijn er nog volop in het vinden van goede exacte en approximatieve algoritmen. Goede oplossingen leveren de telecomoperators handenvol geld op qua besparingen en dus is het commercieel interessant.

Er zijn nog twee variaties die onderwerp van onderzoek zijn op dit moment:

- *multicasting*, het versturen van data naar verschillende ontvangers tegelijkertijd;
- het dynamisch routeren van verkeer, dit hoort sinds kort tot de mogelijkheden omdat ADMs en *light-wave converters* tegenwoordig softwarematig op afstand te bedienen zijn. Dit geeft mogelijkheden voor on-line algoritmen.

Andere problemen

Ook in de mobiele telefonie kwamen nieuwe problemen voor: het bepalen van de frequentie waarover antennes moeten communiceren, zie [1]. Grappig is dat in beide gevallen electromagnetische golven gebruikt worden en dat in beide gevallen kleuringsproblemen (weliswaar van totaal verschillend type) optreden als te optimaliseren probleem.

Een laatste probleemtype is het prijzen van verbindingen in netwerken. Hierbij moet een operator proberen de prijzen van zijn verbindingen zodanig vast te stellen dat aan de ene kant klanten van zijn verbindingen gebruik maken, en aan de andere kant de revenuen zo groot mogelijk zijn. Dit is een speltheoretisch probleem, dat met zogenaamde bilevel programs gemodelleerd kan worden. Hier is de afgelopen tijd veel aandacht aan besteed in de literatuur, zie voor een overzicht [7].

Een zeer recent boek dat over alle (oude en nieuwe) onderwerpen in de telecommunicatie informatie bevat is [11].

LITERATUUR

1. K.I. Aardal, C.P.M. van Hoesel, A.M.C.A. Koster, C.Mannino, A.Sassano. Models and solution techniques for the frequency assignment problem. *4OR*, 1(4):261--317, 2003.
2. I. Caragiannis, C. Kaklamanis, P. Persiano. Wavelength routing in all-optical tree networks: A survey. *Computers and Informatics*, 20:95--120, 2001.
3. T. Erlebach, K.Jansen. The complexity of path coloring and call scheduling. *Theoretical Computer Science*, 255(1-2):119--137, 2001.
4. T. Erlebach, S. Stefanakos. *Wavelength conversion in networks with bounded treewidth*. Tik report 132, Institut für Technische Informatik und Kommunikationsnetze, ETH Zürich, 2002.
5. B. Fortz, M. Thorup. Internet traffic engineering by optimizing OSPF weights. *INFOCOM* (2), pages 519--528, 2000.
6. C.P.M. van Hoesel. Optimization in telecommunication networks. *Statistica Neerlandica*, 59(2):180--205, 2004.
7. C.P.M. van Hoesel. An overview of stackelberg pricing in networks. *European journal of Operational Research*, to appear, 2007.
8. A.M.C.A. Koster, M. Scheffel. *A routing and network dimensioning strategy to reduce wavelength continuity conflicts in all-optical networks*. In Proceedings of INOC 2007, Spa, Belgium, 2007. International Network Optimization Conference.
9. A.M.C.A. Koster, A. Zymolka. Tight LP-based lower bounds for wavelength conversion in optical networks. *Statistica Neerlandica*, 61(1), 2007.
10. M. Minoux. Network design and optimum network design problems: Models, solution methods and applications. *Networks*, 19:313--360, 1989.
11. M.G.C. Resende, P.M.Pardalos (eds.). *Handbook of Optimization in Telecommunications*. Springer, 2006.
12. A. Schrijver, P. Seymour, P. Winkler. The ring loading problem. *SIAM J. Discrete Mathematics*, 11(1):1--14, 1998.

STAN VAN HOESSEL is hoogleraar Operation Research aan de Faculteit der Economische Wetenschappen en Bedrijfskunde van de Universiteit Maastricht. Email: <s.vanhoesel@ke.unimaas.nl>

Rolgordijnen en Operations Research: **THEORIE EN PRAKTIJK**

Rolgordijnen en operations research (OR); op het eerste gezicht lijken deze twee weinig met elkaar te maken te hebben. Maar als u zich realiseert dat een rolgordijn niets meer is dan een op maat gezaagde buis met daar omheen gerold een op maat gesneden stuk stof en één van de standaard OR-problemen het snijprobleem is, is de link wellicht al een stuk duidelijker.*

EDGAR DE GELDER

Dit artikel is geschreven op basis van een onderzoek naar afvalminimalisatie bij één van de grootste rolgordijnfabrikanten van Nederland [2] en het daaruit volgende implementatietraject. De fabrikant, die omwille van bescherming van bedrijfsgevoelige informatie niet zal worden genoemd, levert

klantspecifieke rolgordijnen. Dat wil zeggen dat de klant naast het kiezen van een bepaald dessin ook de afmetingen kan bepalen (in mm nauwkeurig). Als gevolg hiervan is iedere order uniek en is het onmogelijk voorgezaagde buizen en/of voorgesneden stukken stof op voorraad te houden.

In plaats daarvan worden standaardbuizen met een lengte van 5 m en standaardrollen stof van 25 m lang en 2 m breed op voorraad gehouden en worden daaruit de benodigde stukken gezaagd respectievelijk gesneden. Dit proces is wel gemechaniseerd maar niet geautomatiseerd. Dat wil zeggen dat elke buis en elk stuk stof afzonderlijk wordt gezaagd en gesneden waarbij de operator de volgorde van de orders bepaalt. Het feit dat de operator wordt afgerekend op snelheid en kwaliteit heeft tot gevolg dat het onvermijdelijke afval (overgebleven stukjes buis/stof die niet meer gebruikt kunnen worden) enorm is. Kortom, een 'ideale' situatie waar een 'OR-hart' sneller van gaat kloppen.

In theorie kan alles

De beschrijving in de vorige paragraaf maakt duidelijk dat we hier te maken hebben met een een- en twee-dimensionaal *Cutting Stock Problem* (CSP). Zoals al eerder gemeld is dit één van de standaard OR-problemen, maar helaas geeft de literatuur geen standaardoplossing. Het CSP is NP-moeilijk [1] met als gevolg dat er geen praktisch bruikbare optimalisatiealgoritmes bestaan. Algemeen bekend is de kolomgeneratie heuristiek van Gilmore en Gomory [3,4] en hun uitbreiding voor twee of meer dimensies [5]. Echter, deze methode levert de optimale oplossing van de lineaire relaxatie van het probleem waarna een heuristiek wordt gebruikt om de uiteindelijke oplossing geheeltallig te krijgen. Deze laatste stap levert goede oplossingen op voorwaarde dat de multipliciteit van de geselecteerde patronen hoog is. Aangezien de rolgordijnen voornamelijk worden verkocht aan particulieren is het aantal identieke rolgordijnen laag (gemiddeld 2,1). De multipliciteit van de patronen is dus nog kleiner. De oplossing van de lineaire relaxatie van het CSP zal voornamelijk waardes tussen 0 en 1

opleveren met als gevolg dat een geheeltalligheidsheuristiek waarschijnlijk oplossingen genereert die ver van het optimum afliggen.

Naast het feit dat boven beschreven redenering een kolomgeneratie methode afwijst, levert een nadere bestudering van het daadwerkelijke zaag- en snijproces nog meer voorwaarden waarmee de oplossingsmethodiek rekening moet houden.

Zagen van buizen

Het zaagproces kan worden beschreven met behulp van vier eigenschappen: verschil in standaardlengtes, relatieve orderlengte, multipliciteit en probleemgrootte. Zoals in de inleiding beschreven is er slechts één standaardlengte (5 m) die kan worden gebruikt, maar daarnaast worden reststukken boven een bepaalde lengte bewaard om eventueel bij een volgende batch alsnog te kunnen gebruiken. De gemiddelde orderlengte is ongeveer 1,2 m, wat betekent dat de relatieve orderlengte (ten opzichte van de standaardlengte) groot is. De multipliciteit van de orders is laag en de batchgrootte is gemiddeld 50 tot 100 items.

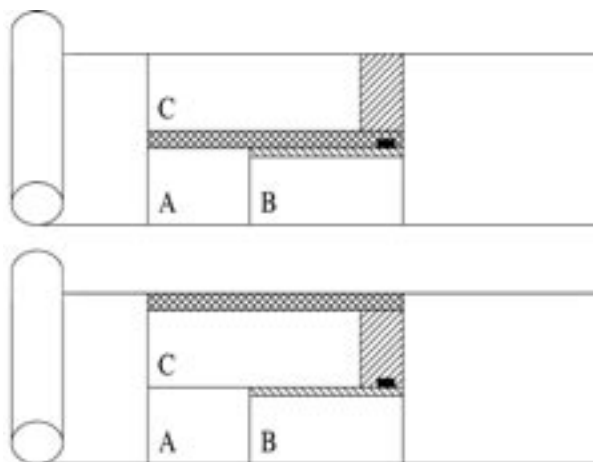
Het feit dat er naast een 'oneindige' hoeveelheid van standaardlengtes ook een aantal reststukken (met eindige hoeveelheid) kan worden gebruikt is een extra complicatie die, net als de lage multipliciteit, nogmaals onderstreept dat kolomgeneratie niet de juiste methode is. Ondanks dat de relatieve orderlengte groot is en de batchgrootte klein lijkt is een volledige enumeratie (of *branch-and-bound*) ook geen oplossing.

Er vanuit gaande dat er nooit meer dan vijf items uit één standaardlengte worden gehaald (wat hoogstwaarschijnlijk niet waar is) is het aantal mogelijke patronen namelijk gigantisch, in de orde van grootte van

$$\binom{100}{5} + \binom{100}{4} + \binom{100}{3} + \binom{100}{2} + \binom{100}{1} = \pm 80 \text{ M}$$

Deze redenering heeft geleid tot een oplossingsmethodiek welke is opgebouwd uit twee fases. In de eerste fase wordt een groot aantal veelbelovende patronen gegenereerd (circa 750) voor de standaardlengte van 5 m waaruit met behulp van *branch-and-bound* een selectie wordt gemaakt. Aangezien slechts een deel van alle mogelijke oplossingen wordt bekeken kan een geldige oplossing van het IP-probleem niet worden gegarandeerd. Derhalve wordt de hoeveelheidsrestrictie op de verschillende items gerelaxeerd met als consequentie dat de resulterende oplossing overproductie bevat.

In het begin van de tweede fase wordt overproductie uit de oplossing ‘gesneden’ door een aantal slechte patronen te laten vervallen. De overgebleven oplossing zal wederom niet voldoen aan de hoeveelheidsrestrictie. In plaats van overproductie zal er nu sprake zijn van onderproductie. De niet-geproduceerde items, een aantal standaardlengtes en de reststukken zijn input voor Gradisar’s *Sequential Heuristic Procedure*. [6] Deze heuristiek is een item-georiënteerde oplossingsmethodiek, wat wil zeggen dat het met name geschikt is voor CSP’s waarbij het aantal standaardlengtes groot is en de probleemgrootte binnen de perken blijft.



Verskillende configuraties van een patroon.

Snijden van stof

In tegenstelling tot het zaagproces bevat het snijproces nogal wat specifieke kenmerken. Naast de vier eigenschappen van het zaagproces (die hetzelfde zijn voor het snijproces) wordt het snijproces gekenmerkt door nog vier eigenschappen: orthogonaliteit, guillotinesnijden, oriëntatierestricties en fouten in de stof. Orthogonaliteit en guillotinesnijden zijn bekende eigenschappen voor 2D CSP's. Dit betekent dat alleen horizontaal en verticaal kan worden gesneden en dat elke snede moet worden afgemaakt (dus altijd de stof in zijn geheel doorsnijden, halverwege stoppen is onmogelijk). Oriëntatierestricties zijn ook niet ongebruikelijk, echter, voor sommige rolgordijnstoffen geldt dat een product gedraaid uit de stof mag worden gesneden maar dan en slechts dan als alle producten van dezelfde klant ook gedraaid gesneden worden (in verband met ‘nerven’ in de stof). Tot slot, de meest complicerende factor is het feit dat de stof weeffouten bevat welke niet in het product terecht mogen komen. Doordat stofrollen van tevoren worden gescand is wel de exacte locatie van deze fouten a priori bekend.

Ook dit 2D CSP wordt opgelost in een tweetal fases. In de eerste fase wordt, vergelijkbaar met het zaagproces, een groot aantal veelbelovende patronen gegenereerd (circa 750) waarna met behulp van *branch-and-bound* een selectie wordt gemaakt. In deze fase wordt geen rekening gehouden met eventuele fouten in de stof, maar wel met de eventuele oriëntatierestrictie. Analoog aan het zaagproces wordt de hoeveelheidsrestrictie gerelaxeerd, resulterend in overgeproduceerde items. Na het verwijderen van de slechte patronen (met overgeproduceerde items) wordt de oplossing bewaard en wordt de hele fase opnieuw uitgevoerd. Ditmaal met de overgebleven niet-geproduceerde items. Deze ‘loop’ eindigt zodra alle items precies voldoende worden geproduceerd.

De volgende fase in de oplossingsheuristiek is de door ons geïntroduceerde *Defects Avoiding Sequence* (DAS). Deze heuristiek maakt gebruik van het feit dat de volgorde waarin de verschillende patronen zullen worden gesneden nog niet bepaald is evenals de opbouw van de 2D patronen. Door deze opbouw (configuratie) van de patronen te veranderen en de onderlinge volgorde van de verschillende patronen te wisselen kan er een oplossing ontstaan waarbij de fouten in rest/afvalmateriaal terecht komen. In Figuur 1 is een voorbeeld gegeven van een patroon met 3 items (A, B en C) welke in twee verschillende configuraties kan worden gebruikt om 'om de fout heen te snijden' (het zwarte rechthoekje). Uiteraard dient het deel tussen het begin van de rol tot aan dit patroon te worden gevuld met andere patronen. Merk op dat beide configuraties kunnen schuiven zonder dat de fout in één van de items terecht komt. De DAS-heuristiek gaat op zoek naar zo'n oplossing. Mocht dit niet mogelijk zijn dan zal de heuristiek de additionele hoeveelheid afval proberen te minimaliseren.

Resultaat

Met behulp van simulaties is aangetoond dat de oplossingsmethodieken die hier in het kort zijn beschreven, aanzienlijk beter presteren dan de oude werkmethode. De hoeveelheid afval voor het 1D en 2D materiaal kan met respectievelijk ruim 80% en 50% worden gereduceerd. Deze resultaten waren aanleiding om de oplossingsmethode daadwerkelijk in praktijk toe te passen.

In praktijk valt het altijd tegen

Het toepassen van theoretische oplossingen in de praktijk is altijd lastig. Tijdens de ontwikkeling van software-tools voor het eenvoudig dagelijks oplossen van 1D en 2D CSP's bleek al snel dat de

theorie een heleboel aannames doet die in de praktijk lang niet altijd geldig zijn. Zo moet er in praktijk rekening worden gehouden met het feit dat iedere zaagsnede een verlies oplevert van 3mm. Uiteraard is dat niet veel maar het betekent wel dat een patroon dat precies passend is, in praktijk net iets te groot is.

Daarnaast blijkt dat in praktijk een inklem-breedte aanwezig is. In het 1D geval betekent dit dat er altijd een reststuk moet overblijven van minimaal 50 mm (1% extra afval). Het wiskundig 'verkleinen' van de standaard lengtes met 50 mm is voldoende. Echter, in het 2D geval (snijden van stof) geldt eenzelfde restrictie welke niet zo eenvoudig op te lossen is. Voor een rolgordijn geldt dat de hoogte iets overschreden mag worden maar dat de breedte exact moet worden afgesneden. Dit betekent dus dat een randje kleiner dan 50 mm afsnijden van de breedte van een rolgordijn onmogelijk is maar van de hoogte niet nodig is. Als gevolg daarvan dient tijdens de generatie van verschillende patronen na te worden gegaan of het betreffende patroon wel mogelijk is en voor welke configuraties dit dan geldt.

Er zijn tientallen voorbeelden van dit soort praktische randvoorwaarden welke over het algemeen goed op te lossen zijn zonder daarbij teveel aan kwaliteit van de oplossing in te leveren. Een groter probleem van de implementatie is de keuze van de nieuwe werkvolgorde. Zo zou men kunnen pleiten voor een werkvoorbereider welke de software bedient en de output (zaag- en snijplannen) gebruikt om de operators aan te sturen. Voordeel van zo'n opstelling is het feit dat slechts enkele *expert-users* de ontwikkelde software zullen gaan gebruiken. Nadeel van deze werkvolgorde is het verlies aan flexibiliteit. Wat te doen als een reststuk om onverklaarbare redenen niet aanwezig is? Wat te doen als door een productiefout een bepaald item opnieuw dient te worden geproduceerd?

De andere optie is het *real-time* optimaliseren van de verschillende CSP's. Dit betekent dat iedere operator een computer tot zijn beschikking krijgt welke hem helpt het afval te minimaliseren. Naast de enorme flexibiliteit heeft deze methodiek nog een aantal positieve neveneffecten, zoals verhoging van de productiviteit, registratie van productie en traceerbaarheid van een bepaald item. Een groot nadeel is het feit dat het softwarepakket dusdanig flexibel en makkelijk te gebruiken dient te zijn dat ook een operator met weinig verstand van computers het programma kan bedienen. Bovendien moet de optimalisatie snel zijn afgerond om de productie niet te veel te vertragen.

Aangezien flexibiliteit een enorm belangrijke factor op de productievloer is, is er gekozen voor de tweede optie. Gebruiksvriendelijke software is ontwikkeld welke het mogelijk maakt eenvoudig en snel tussentijds te heroptimaliseren als bijvoorbeeld een nieuwe weeffout is ontdekt die de scanner per abuis had overgeslagen. Diegene die denkt dat daarmee de implementatie succesvol is afgerond moeten wij teleurstellen. Veruit de grootste uitdaging van het gehele project moet dan nog komen: de acceptatie.

Ondanks dat het softwarepakket aan alle flexibiliteits- en eenvoudigheidseisen voldeed bleek het lastig te zijn om werknemers te overtuigen om de software te gaan gebruiken. De angst dat de computer hun werk zou overnemen (wat totaal niet aan de orde is) bleek een belangrijke reden om niet met de software te willen werken. Veruit de belangrijkste reden voor de negatieve houding is simpelweg de verandering. Angst voor verandering is menselijk en tijdens de acceptatie een groot struikelblok. Maar in diezelfde angst schuilt ook een groot voordeel: nu, inmiddels een jaar later, is iedereen gewend en wil niemand meer terug naar de oude situatie!

Tot slot

Zoals hierboven reeds aangegeven draait de software, met daarin de heuristieken zoals kort in dit artikel beschreven, inmiddels een jaar naar alle tevredenheid. Naast tevredenheid bij de werknemers is er ook tevredenheid bij de directie welke haar afvalpercentages heeft zien dalen met respectievelijk 70% en 40% voor het zaag- en snijproces. Kortom, in theorie kan alles, in praktijk valt het altijd tegen, maar gelukkig is niets onmogelijk!

* Edgar de Gelder ontving voor zijn onderzoek een eervolle vermelding van de jury van de VVS-scriptieprijs voor Statistiek en Operationele Research 2006.

LITERATUUR

1. E.E. Bischoff (1995). Cutting and packing. *European Journal of Operations Research*, 84(3), pp. 503-505.
2. E.R. de Gelder (2005). *The one and two-dimensional cutting stock problem within the roller blind production process – theory and application*. Master's thesis, Erasmus Universiteit Rotterdam.
3. P.C. Gilmore and R.E. Gomory (1961). A Linear Programming Approach To The Cutting Stock Problem. *Operations Research*, 9(60), pp. 849-859.
4. P.C. Gilmore and R.E. Gomory (1963). A Linear Programming Approach To The Cutting Stock Problem - Part II. *Operations Research*, 11(6), pp. 863-888.
5. P.C. Gilmore and R.E. Gomory (1965). Multistage Cutting Stock Problems of Two and More Dimensions. *Operations Research*, 13(1), pp. 94-120.
6. M. Gradišar, M. Kljajić, G. Resinovič, and J. Jesenko (1999). A sequential heuristic procedure for one-dimensional cutting. *European Journal of Operational Research*, 114(2), pp. 557-568.

EDGAR DE GELDER is werkzaam bij Pointlogic waar hij zich voornamelijk bezig houdt met loonkostensimulaties en TV-(reclame-)planning.
E-mail: <gelder@pointlogic.com>



Edinburgh Mathematical Society Colloquium in St Andrews in 1964. Op de vierde rij in het midden staat Aida Beatrijs Paalman-de Miranda

Black is beautiful

FRED STEUTEL

Toen ik me afvroeg wat er in een mogelijk artikel over 'Vrouwen in de Statistiek' zou kunnen staan, kwam ik op het idee om na te denken over andere herkenbare 'minderheden' in de kansrekening en statistiek – vrouwen vormen in de samenleving natuurlijk geen minderheid. Je kunt daarbij van alles bedenken: roodharigen, linkshandigen of stotteraars in de statistiek. Uiteindelijk leek 'zwarten in de statistiek' me wel wat, al is een bezwaar dat zelfs mensen die maar voor een achtste of nog minder niet-blank ('niet-Kaukasisch') zijn, zich 'zwart' noemen. We hebben eigenlijk geen geschikt woord voor 'zwarten'. Vroeger zeiden we 'negers', maar dat is nu een taboewoord. Ik besloot op internet te gaan zoeken bij 'zwarte statistici', maar dan in het Engels. Het groepje eminente zwarte statistici dat u hieronder aantreft, is subjectief gekozen; sommigen van hen heb ik persoonlijk ontmoet. Mijn losse opmerkingen over

hen zijn niet bedoeld als beknopte biografieën.

De zoekterm '*black statisticians*' leverde weinig op: pagina's over casino's en een artikel over '*Death Sentences Linked to History of Lynching in States*'. De enige bruikbare pagina betrof een interview van Albert Marshal met William Birnbaum, geen van beiden zwart. De link met '*black statisticians*' betreft een passage over de achterstelling van zwarte statistici en andere zwarte wiskundigen in, speciaal, de zuidelijke staten van de VS, waarover straks meer. Omdat het woord '*black*' in de VS al bijna even erg is als 'neger' bij ons, besloot ik '*black*' te vervangen door '*Afro-American*'. Daarnaast heb ik '*statisticians*' uitgebreid tot '*mathematicians*'.

Met de zoekterm '*Afro-American mathematicians*' vond ik veel excellente wiskundigen, en het enige wat mij te doen stond was daar de statistici (inclusief kansrekenaars) uit te selecteren.

Verreweg de beroemdste van allemaal is David Blackwell (nomen est omen). Zijn carrière begon in een periode van nog ernstige discriminatie, ook in het Noorden van de VS. In 1941 werd Blackwell benoemd aan het Institute for Advanced Studies in Princeton. Door die benoeming werd hij automatisch *visiting fellow* van Princeton University, onder heftig protest van de president van de universiteit, die eerder de beroemde zwarte zanger Paul Robeson buiten de deur had weten te houden. Blackwell mocht toch even blijven. Na korte benoemingen aan een aantal zwakkere ‘zwarte universiteiten’, belandde hij op de beste zwarte universiteit van de VS, Howard University in Washington, D.C. Later werkte hij onder meer aan UCLA en Berkeley. Kansrekenaars en mensen uit de stochastische OR (bijvoorbeeld wachttijdtheorie) kennen Blackwell van de naar hem genoemde limietstelling in de vernieuwingstheorie, de statistici kennen de ongelijkheid voor de variantie van schatters, genoemd naar Rao en Blackwell.

Na Blackwell is Charles Bell mogelijk de bekendste zwarte statisticus. Hij was actief op allerlei terreinen van de statistiek, met name *order statistics*, signaaldetectie en parameter vrije methoden. Hij was in 1964 in Nederland, waar hij ook het Mathematisch Centrum (CWI) bezocht. Ik herinner me dat hij al in de VS vrij goed Nederlands had geleerd, maar dat tot zijn verdriet bijna iedereen in het Engels met hem wilde converseren.

Bharucha-Reid heette in zijn jeugd Albert Turner Reid. Na zijn huwelijk met de Indiase Bharucha voegde hij haar naam aan de zijne toe. Ook hij bezocht het MC/CWI. Hij heeft verschillende boeken geschreven, waaronder een over Markov-processen en hun toepassingen, onder meer in de wachttijdtheorie. In dat boek heeft hij een artikel van Runnenburg overgenomen waar een fout in zat; een heel subtiele kwestie

over (on)afhankelijkheid, waar Runnenburg zelf al eens op gewezen had. Deze omstandigheid was er mogelijk de oorzaak van dat Runnenburg niet veel waardering voor Bharucha-Reid had. In de Verenigde Staten had hij een status niet zo heel veel onder die van Blackwell; misschien wel wat veel eer. Hij had nog twee bijzonderheden: hij heeft nooit een proefschrift geschreven – vond hij zonde van zijn tijd – en, volgens geruchten op het MC, droeg hij altijd een pyjama onder zijn kleren; hij had het vaak koud.

Ernest Wilkins was een expert op het gebied van *random* polynomen. Hij is in zijn jeugd een keer onheus behandeld door de American Mathematical Society, en heeft daarna nooit meer iets met die club te maken willen hebben. Ik heb hem nooit ontmoet.

Er lijken geen befaamde zwarte vrouwelijke statistici te zijn. Maar er is natuurlijk een website met *Black Women in the Mathematical Sciences*. De succesvolste die we daar vinden, is de in Engeland geboren Katherine Adebola (Kate) Okikiolu, dochter van de zeer productieve Nigeriaanse wiskundige George Okikiolu (95 vermeldingen in de *Mathematical Reviews*). Zij is opgeleid in Cambridge en aan de UCLA, en nu werkzaam aan de UC in San Diego. We vinden op deze site natuurlijk ook de Surinaams-Nederlandse Aida Beatrijs (Ietje) Paalman-de Miranda, in 1951 een jaargenoot van mij aan de Universiteit van Amsterdam. Zij is geen statisticus, maar heeft toch een artikel geschreven met Gilbert Helmberg – deels kansrekenaar – over wel-verdeelde rijen. Ik ontmoette haar nog onlangs op het oud-medewerkersfeestje van het 60-jarige CWI. *Black is beautiful!*

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



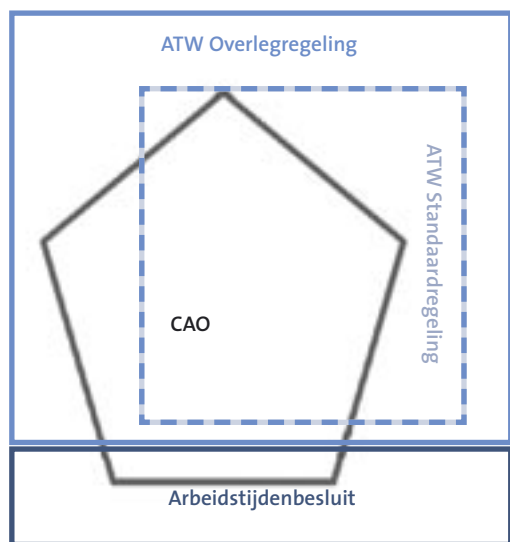
De **COMPLEXITEIT** van de Vereenvoudigde Arbeidstijdenwet

Dit artikel beschrijft een aantal recente veranderingen van de Arbeidstijdenwet en het Arbeidstijdenbesluit en de gevolgen daarvan voor het maken en controleren van dienstroosters. Met name wordt ingegaan op de gevolgen van het Jaeger-arrest, dat bepaalt dat aanwezigheidsdiensten als arbeid gerekend moeten worden. Hoewel sprake is van vereenvoudigde wetgeving kunnen daar vanuit een wiskundig perspectief vraagtekens bij gezet worden.

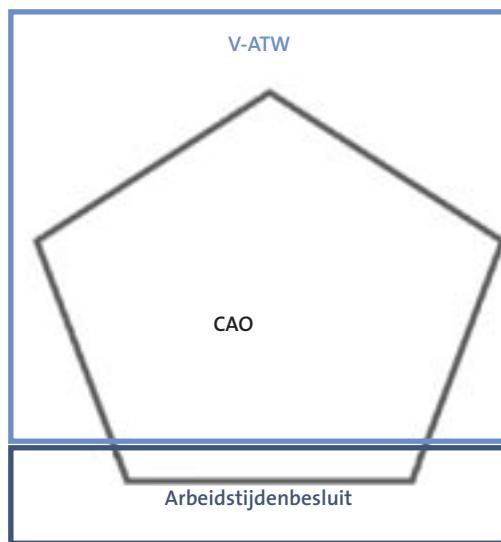
LAURENS FIJN VAN DRAAT, JOHN POPPELAARS EN GERHARD POST

Een recente uitspraak van de rechtbank in Den Haag bepaalde dat brandweerlieden minder mochten gaan werken, met behoud van het salaris. Een aantrekkelijk vooruitzicht voor de brandweerlie-

den natuurlijk, maar minder aantrekkelijk voor de gemeenten die hierdoor met een aanzienlijke kostenstijging te maken krijgen. De motivering van de uitspraak moet gezocht worden in veranderingen



Situatie voor 1 april 2007



Situatie na 1 april 2007

Figuur 1. Regelgeving rondom arbeidstijden oud en nieuw.

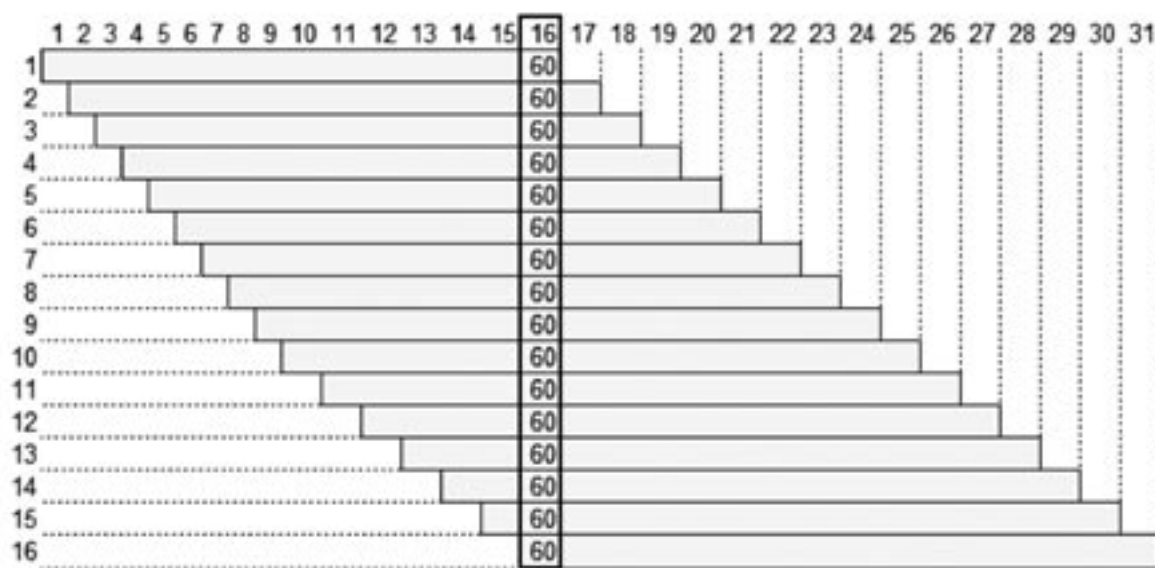
van de wetgeving rondom arbeidstijden: de invoering van de Vereenvoudigde Arbeidstijdenwet (v-ATW) op 1 april 2007 en de aanpassing van het Arbeidstijdenbesluit (ATB) in juli 2006, naar aanleiding van het Jaeger-arrest. In dit artikel gaan we in op de complexiteit van deze veranderingen. Heeft het vorige kabinet haar motto 'meer werk, minder regels' waargemaakt? We laten zien dat dit nog tegenvalt. Zo moet voor het controleren van één van de nieuwe regels uit het ATB een knapzakprobleem opgelost worden, een probleem dat in de complexiteitstheorie bekend staat als moeilijk probleem.

Regels voor arbeidstijden

De wijzigingen van de Arbeidstijdenwet (ATW) en het Arbeidstijdenbesluit vloeien voort uit afspraken op Europees niveau (European working time directive en het Jaeger-arrest) en de ambitie van

het kabinet om meer over te laten aan werkgevers en werknemers. Naast de ATW en het ATB is ook de CAO van invloed op de inzetmogelijkheden van personeel. In Figuur 1 staan de onderlinge verhoudingen tussen ATW, ATB en CAO weergegeven.

In de oude situatie, tot 1 april 2007, was er sprake van een standaard- en een overlegregeling binnen de ATW. De standaardregeling was op iedereen van toepassing, tenzij er in de CAO of via een andere collectieve afspraak (met de Ondernemingsraad) van afgeweken was. Wel waren de (ruimere) grenzen uit de overlegregeling absoluut; die mochten in geen geval overschreden worden. In het Arbeidstijdenbesluit zijn uitzonderingen op de ATW opgenomen die noodzakelijk zijn vanwege de arbeidsomstandigheden in specifieke branches. Hierbij kan gedacht worden aan de gezondheidszorg, politie, brandweer en de horeca. Vaak gaat het om verruiming van regels uit de overlegregeling die anders tot onwerkbaar situaties zouden leiden.



Figuur 2. De controle van de arbeidstijd per 16 weken.

De v-ATW geeft voor het overgrote deel van de regels nog maar één norm. Het verschil tussen standaardregeling en overlegregeling is daarmee bijna geheel verdwenen.

Wat is er zoal veranderd?

De v-ATW beperkt zich zoveel mogelijk tot regels die nodig zijn voor de bescherming van de veiligheid, gezondheid en welzijn van de werknemer. Er staan minder regels in dan in de oude Arbeidstijdenwet. Er zijn bijvoorbeeld nog maar vier regels over de maximum arbeidstijd tegenover twaalf in de oude wet. De maximum arbeidstijd is 12 uur per dienst geworden en maximaal 60 uur per week. Een aanzienlijke verruiming omdat in de standaardregeling van de ATW maximaal 9 uur per dienst en 45 uur per week mocht worden gewerkt. In een periode van 4 weken mag een werknemer gemiddeld 55 uur per week wer-

ken (was 50 uur) en per 16 weken gemiddeld 48 uur. De periode van 16 weken is nieuw, in de oude wet was de maximale referentieperiode 13 weken voor de maximale arbeidsregels. Om deze regel te controleren zal een week waarin 60 uur wordt gewerkt 16 keer in een periode van 16 weken (over een periode van 15 weken in het verleden tot 15 weken in de toekomst) worden betrokken, zie figuur 2. Het zal duidelijk zijn dat het controleren van regels die een dergelijke lange periode beslaan een tijdrovende zaak is, zeker in een dynamische omgeving waar veel veranderingen in het rooster plaatsvinden. Overigens bestaan er CAO's met regels over maar liefst 52 weken.

Door deze verruiming van de regels neemt de inzetbaarheid van het arbeidspotentieel per jaar met 20% toe. Dat zal de werkgevers zeker aanspreken. Of die extra ruimte ook daadwerkelijk door werkgevers kan worden benut, is afhankelijk van of het werk er ook is en of het vanuit kostenoverwegingen wenselijk is. De uren die medewerkers

boven het aantal afgesproken contracturen worden ingezet zijn vaak met een extra toeslag belast. Daarnaast speelt dat als er gedurende de nacht wordt gewerkt, de maximale arbeidstijd per periode van 16 weken daalt naar 40 uur gemiddeld per week.

Naast de verruiming van de grenzen zijn ook enkele regels versoepeld en aangepast. Zo is de pauzeregeling vrijer, is de regeling voor oproepen nachtdiensten sterk verbeterd voor werknemers en is de consignatieregeling tot een goed niveau verbeterd, aldus FNV Bondgenoten.

Knapzakprobleem in het Arbeidstijdenbesluit

Ook in het ATB zijn de nodige aanpassingen doorgevoerd, met name in de regels met betrekking tot de rust na aanwezigheidsdiensten. Voornaamste aanleiding hiervoor was het Jaeger-arrest uit 2003 van het Europese Hof van Justitie dat bepaalt dat diensten, waarbij het verplicht is aanwezig te zijn, als arbeid gerekend moeten worden. Dergelijke diensten, vaak aanwezigheidsdienst of slaapdienst genoemd, worden gebruikt als snelle aanwezigheid vereist is, zoals in de gezondheidszorg en bij de brandweer. In het geval dat een werknemer één of meer aanwezigheidsdiensten uitvoert, is een regel voor wekelijkse rust in het ATB van toepassing:

‘De werkgever organiseert de arbeid zodanig dat de werknemer in elke aaneengesloten tijdruimte van 7 maal 24 uren een rusttijd heeft van tenminste 90 uren, welke rusttijd bestaat uit een onafgebroken rustperiode van tenminste 24 uren alsmede 6 onafgebroken rustperioden van tenminste 11 uren, welke 6 rustperioden in elke aangesloten tijdruimte van 7 maal 24 uren ten hoogste

eenmaal mogen worden ingekort tot 10 uren alsmede eenmaal tot 8 uren. Onafgebroken rustperioden kunnen aaneengesloten zijn.’

Om de regel correct te controleren zal iedere periode van 168 uur gecontroleerd moeten worden. De regel vertelt niet waar deze periode mag beginnen. Dit maakt natuurlijk wel verschil: de grens kan net door één van de rustperiodes lopen, die daardoor te kort wordt. Het controleren van de totale rusttijd is eenvoudig: we kunnen simpelweg alle losse blokjes rusttijd binnen iedere controle-periode bij elkaar optellen en vergelijken met de norm van 90 uur. Het vervolg van de regel betreft de verbodsbepaling van de rustperiodes. Er dient één rustperiode te zijn van tenminste 24 uur, vier van tenminste 11 uur, één van tenminste 10 en één van tenminste 8 uur. Ook dit valt met enig inzicht nog redelijk eenvoudig te controleren. De rustperiodes kunnen aflopend op lengte gesorteerd worden, en één voor één vergeleken worden met de gevraagde periodes van 24, 11, 11, 11, 11, 10 en 8 uur. Het venijn van deze regel zit ‘m echter in de staart. Door de toevoeging: *‘Onafgebroken rustperioden kunnen aaneengesloten zijn’* is de laatste controle niet meer zo eenvoudig. Wiskundig gezien is het probleem nu een (meervoudig) knapzakprobleem (*knapack problem*) geworden, welke in zijn algemeenheid moeilijk is om op te lossen.

Om het controleren van de regel te vertalen in het knapzak probleem laten we elke rustperiode van de medewerker corresponderen met een knapzak. De grootte van een knapzak is het aantal uren van de rustperiode. In deze knapzakken proberen we zeven artikelen ter grootte van 24, 11, 11, 11, 11, 10 en 8 (de vereiste rustperiodes) mee te nemen. Als dit lukt, dan hebben we aan de regel over wekelijkse rust voldaan, als het niet lukt, dan is de regel overtreden. Figuur 3



Figuur 3. De knapzakken voor rustperiodes van 32, 31 en 28 uur.

geeft de situatie weer, waarbij een medewerker slecht drie perioden van rust heeft, van respectievelijk 32, 31 en 28 uur.

Kunnen de zeven artikelen verdeeld worden over de drie knapzakken? Dat is niet direct duidelijk. Een voor de hand liggende (gretige) keus is om de 24 en 8 in de knapzak van 32 te stoppen. Enig puzzelen leert dat de overgebleven artikelen van grootte 11, 11, 11, 11 en 10 niet in de overgebleven knapzakken van 31 en 28 passen. Betekent dit dat we de regel hebben geschonden? Nee. Een andere mogelijkheid is om 11, 11 en 10 in de knapzak van 32 op te bergen. Als we dan 24 in de knapzak van 28 doen, passen 11, 11 en 8 in die van 31. Aan de regel wordt (voor deze periode van 168 uur) dus voldaan.

Conclusie

Als gevolg van het Jaeger-arrest en de v-ATW is het maximale aantal aanwezigheidsdiensten dat mag worden gelopen, voor de brandweer, gedaald van 121 naar 103 per jaar. Hierdoor loopt de beschikbare arbeidstijd per jaar met 12% terug. Dit betekent een aanzienlijke stijging in de kosten. Vanuit deze optiek lijkt de v-ATW in combinatie met het Jaeger-arrest een verslechtering. Dit is echter ook een gevolg van het perspectief van waaruit er naar de wijziging in de regels

wordt gekeken. In het voorbeeld van de brandweer wordt getracht de bestaande roosters aan te passen om ze aan de nieuwe regels te laten voldoen, met een productiviteitsdaling tot gevolg. Beter zou zijn afstand te nemen van de huidige roosters en te bekijken wat overall een betere oplossing zou zijn, bijvoorbeeld door de aanwezigheidsdienst te vervangen door een reguliere ploegendienst om de 24-uurs bezetting van de posten te garanderen. Een wiskundige benadering van deze vraagstelling kan hierin duidelijkheid verschaffen. Roosterplanningssystemen, zoals ORTEC HARMONY™, zijn in dit proces door de ingewikkelde regels in veel situaties onmisbaar; enerzijds bij het controleren van (handmatig) geconstrueerde roosters, anderszijds bij het automatisch genereren van dienstroosters.

LAURENS FIJN VAN DRAAT is in 2000 afgestudeerd als econometrist en als econoom aan de Vrije Universiteit van Amsterdam. Sindsdien werkt hij bij ORTEC bv als consultant. E-mail: <LFijnvanDraat@ortec.nl>.

JOHN POPPELAARS is senior consultant bij de afdeling Logistieke Consultancy van ORTEC bv. Hij is in 1990 afgestudeerd als econometrist aan de Erasmusuniversiteit en heeft zich gespecialiseerd in personeelsplanning. E-mail: <JPoppelaars@ortec.nl>.

GERHARD POST is part-time werkzaam bij ORTEC bv, binnen de afdeling voor algoritmiek ontwikkeling. Tevens is hij part-time universitair docent bij de afdeling Toegepaste Wiskunde van de Universiteit Twente. E-mail: <GPost@ortec.nl>.



Randomized Response Technieken voor sociaal gevoelige vragen in beleidsonderzoek*

GERTY LENSVELT-MULDERS

Het onderzoek van sociaal gevoelige onderwerpen

Sociaal gevoelig zijn alle onderwerpen die psychisch of sociaal bedreigend zijn voor respondenten. Een onderzoek is psychisch bedreigend wanneer de vragen gaan over onderwerpen die zeer persoonlijk en /of stresserend zijn voor de respondent en waarvan de antwoorden mogelijk afbreuk doen aan zijn zelfbeeld. Sociaal bedreigend is een

onderwerp wanneer er een maatschappelijk taboe op rust. Binnen onze onderzoeksgroep ontwikkelen we designs en statistische analysetechnieken voor onderzoek naar gevoelige onderwerpen, met onderzoek naar sociale fraude en integriteitsschending bij de politie (in samenwerking met de VU) als voorbeelden van sociaal gevoelige onderwerpen en onderzoek naar pesten en eenzaamheid bij leerlingen van de basisschool als voorbeeld van een psychisch gevoelig onderwerp.

Het is noodzakelijk om onderzoek te doen naar het verbeteren van instrumenten voor gevoelige vragen, omdat de kwaliteit van dit onderzoek op twee manieren wordt bedreigd: respondenten willen liever niet meewerken aan dergelijk onderzoek, wat leidt tot hoge en selecte non-response, en respondenten hebben een neiging tot het geven van sociaal wenselijke of veilige antwoorden. Beiden processen leiden tot vertekende resultaten (Rasinski, Willis, Baldwin, Yeh en Lee, 1999).

De randomised response techniek

In 1965 beschreef Warner een onderzoekstechniek die, door het totaal waarborgen van de privacy van respondenten, zou moeten leiden tot een vermindering van non-response en het geven van sociaal wenselijke antwoorden op gevoelige vragen. Warner's Randomised Response (RR) design kan nog steeds worden gebruikt om de algemene basisprincipes van RR technieken te demonstreren.

Een respondent krijgt twee uitspraken voorgelegd in de volgende vorm:

- Ik heb gevoelig gedrag A vertoond (wel eens XTC genomen op een feestje)
- Ik heb gevoelig gedrag A niet vertoond (nooit XTC genomen op een feestje)

Met behulp van een kansspelletje (met bijvoorbeeld dobbelstenen) wordt bepaald welke uitspraak moet worden beantwoord. Omdat alleen de respondent weet welke uitspraak hij of zij heeft beantwoord, heeft het individuele antwoord geen enkele betekenis. Door kennis van het randomisatieproces is het wel mogelijk om een populatieschatting van het gevoelige gedrag ($\hat{\pi}$) te maken.

$$\hat{\pi} = \frac{\hat{\lambda} + (p-1)}{2p-1}$$

waar de zuivere schatter van λ , $\hat{\lambda}$ de geobserveerde proportie 'waar'-antwoorden in onze steekproef is, p is de kans dat de respondent het eerste statement moet beantwoorden, en $p-1$ de kans dat het tweede statement moet worden beantwoord. De steekproefvariantie is gelijk aan

$$\text{Var}(\hat{\pi}) = \frac{\pi(1-\pi)}{n} + \frac{p(1-p)}{n(2p-1)^2}$$

Uit de tweede vergelijking blijkt ook het nadeel van de RR, namelijk dat de variantie fors toeneemt.

RR is dus veel minder efficiënt dan conventionele 'directe vragen'-designs. Wanneer we de methode van Warner gebruiken kan dit betekenen dat de steekproef tot acht maal zo groot moet worden om vergelijkbare betrouwbaarheidsintervallen te verkrijgen. Daarom zijn er nieuwe RR designs ontwikkeld, die meer efficiënt zijn (zie bijvoorbeeld Lensvelt-Mulders, Hox en Van der Heijden, 2005). Op dit moment is de best onderzochte methode de Forced Response techniek (FR).

Bij FR wordt aan de respondent slechts 1 vraag voorgelegd. Met behulp van een randomizer wordt de respondent 'geforceerd' om een antwoord te geven. Om een goede verhouding tussen het waarborgen van de privacy van respondenten enerzijds en de efficiëntie van FR anderzijds te verkrijgen kan het design het best zo worden ingericht dat p_1 , de kans op een waar antwoord, rond de 75% ligt, p_2 , de kans op een geforceerd ja-antwoord op 15% en p_3 , de kans op een geforceerd nee-antwoord op 10%.

Dit FR design is toegepast in onderzoek naar fraude met sociale uitkeringen, waarbij de *randomizer* uit 2 dobbelstenen bestond (van der Heijden, van Gils, Bouts en Hox, 2000), en in onderzoek naar integriteitsschending bij de poli-

tie, met als randomizer een digitale spinner (de spinners kunnen worden geprobeerd op <www.rlmedia.nl/portfolio/FlashExperiments/spinner.html> (Peeters, 2005, Peeters, Lensvelt-Mulders en Lasthuizen, submitted). Ook is de FR toegepast in een computergestuurd onderzoek bij kinderen in een game-based environment (Lensvelt, 2006). Het voordeel van het gebruik van de spinner is dat het heel eenvoudig wordt om het standaard dichotome RR design uit te breiden met een ander type antwoordcategorieën zoals bijvoorbeeld Likertschalen.

Validiteit van RR resultaten

Een meta-analyse van de resultaten van 35 jaar vergelijkend RR onderzoek heeft uitgewezen dat met behulp van RR meer valide uitkomsten worden verkregen dan met andere, meer conventionele dataverzamelingsmethoden, zoals schriftelijke en telefonische enquêtes en face-to-face interviews (Lensvelt-Mulders, Hox en van der Heijden, en Maas, 2005). Dat wil zeggen dat er minder sociaal wenselijke antwoorden worden gegeven op de gevoelige vragen. Dit positieve effect is groter naarmate de onderwerpen meer sociaal gevoelig zijn. Het gebruik van RR leidt daardoor tot een betere kennis van het voorkomen van sociaal gevoelig gedrag in de samenleving en een beter inzicht in sociaal gevoelige fenomenen kan bijdragen tot het verbeteren van de besluitvorming voor het beleid.

Analyse van FR data; het maken van punt-schattingen

De basisanalyse van FR resultaten is eenvoudig. Neem gevoelig gedrag X (heeft u zich wel eens

zeker voorgedaan dan u eigenlijk bent om uw uitkering niet in gevaar te brengen?) waarvoor de populatie dichotoom is. Als π de onbekende kans is op een waar ja-antwoord, dan is de kans op een ja-antwoord in de steekproef gelijk aan $\lambda = p_2 + p_1\pi$. Met behulp van deze kennis kunnen we met behulp van de geobserveerde proportie ja-antwoorden $\lambda, \hat{\pi}$ schatten

Omdat het in de praktijk van onderzoek naar gevoelige onderwerpen natuurlijk ook interessant is om, naast het voorkomen van gevoelig gedrag, ook de hoeveelheid te schatten, geven we hier ook de berekening voor het schatten van kwantitatieve variabelen (Peeters, 2005).

Neem weer het gevoelige gedrag X (word je wel eens gepest op school? nee, dat gebeurt nooit; ik word wel eens een beetje gepest, maar dan kan ik wel om lachen; ja, dat gebeurt soms wel eens en dat is niet leuk; ja, dat gebeurt best vaak en dat is niet leuk; ja, dat gebeurt elke dag en daar baal ik van).

Weer wordt er een aselechte steekproef getrokken waarbij $i = 1, \dots, i, \dots, n$. Wat we willen schatten is de onbekende populatie parameter μ_x . Om dat te doen maken we X discreet, met waarden $x_1, \dots, x_n$ en onbekende ware proporties $\pi_1, \dots, \pi_n$. In plaats van ons te focussen op de totale gemiddelde score, schatten we de populatie proporties in elke categorie apart (Peeters 2005). Oftewel we schatten π_i , waar $i = 1, \dots, i, \dots, n$.

Laat p de selectiekans voor een waar antwoord zijn en p_j de kans op een geforceerd antwoord, waarbij geldt $j = 1, \dots, i, \dots, n$ en $\sum_j p_j = 1 - p$. Als γ_i is de kans op een bepaalde kwantitatieve respons, dan blijkt dat :

$$\gamma_i = p\pi_i + p_i \quad (i = j)$$

Met behulp van deze gegevens kan weer $\hat{\pi}$ worden geschat.

RR analyses gericht op interventies

Naast het schatten van de prevalentie en frequentie van het gevoelige gedrag is het voor beleidsonderzoekers ook belangrijk om gedrag te verklaren. Met dergelijke informatie kan meer gericht beleid worden gevoerd. Met behulp van gewogen logistische regressietechnieken is het mogelijk om de FR resultaten te verklaren. Op deze manier is bijvoorbeeld regelovertreding bij sociale uitkeringen verklaard aan de hand van onder meer kennis van de regels voor het recht op een uitkering, de waargenomen redelijkheid van deze regels en de geschatte kans om gepakt te worden bij overtreding. Het standaard logistische model wordt als volgt gedefinieerd:

Als $\pi_{1|i}$ is de kans dat het antwoord 'nee' is en $\pi_{2|i}$ de kans dat het antwoord 'ja' is, beiden als functie van covariaat vector i . Laat dan de k^{de} covariaat waarde in de vector geïndexeerd zijn als i gedefinieerd als x_{ik} , waar x_{ik} zowel een continue als een dummy variabele kan zijn. Dan is β_k de regressieparameter voor de relatie tussen verklaarende variabele k en het gegeven antwoord. De standaard logistische regressie wordt dan

$$\pi_{1|i} = \frac{1}{1 + \exp(\sum_k x_{ik} \beta_k)}$$

Voor een nee-antwoord geldt dat w_i het gewicht is voor respondent i waarbij (n_{i1}, n_{i2}) is (1,0) en voor een ja-antwoord geldt (n_{i1}, n_{i2}) met (0,1). Dan wordt de loglikelihood gegeven door:

$$\log(L) = \sum_i w_i n_{i1} \log(\pi_{1|i}) + \sum_i w_i n_{i2} \log(\pi_{2|i})$$

Dit algemene model kan worden aangepast voor de RR vraag, zodanig dat kan worden gecorrigeerd voor ware en geforceerde antwoorden.

Voor meer informatie over statistische methoden om RR modellen te schatten verwijzen we

naar van den Hout and van der Heijden (2004), Van den Hout en Lensvelt-Mulders (2005) en Lensvelt-Mulders, van der Heijden, Laudy en van Gils (2006).

Het controleren voor bias in RR

Tenslotte wil ik de lezer nog een blik in de keuken van de *randomized response* techniek gunnen, door het geven van een korte samenvatting van de allerlaatste ontwikkelingen op het terrein van RR: het corrigeren voor vertekening door non-compliance met RR. De basisaanname van de RR is dat, doordat de privacy volledig beschermd is, respondenten eerlijk antwoord zullen geven op gevoelige vragen. Meta-analyse van RR-onderzoek wees uit dat RR inderdaad de meest valide schatters geeft bij onderzoek naar gevoelige onderwerpen. Datzelfde onderzoek wijst ook uit dat de RR schatters toch niet helemaal onvertekend zijn. Böckenholt en van der Heijden hebben met behulp van multivariate modellen deze vertekening gemodelleerd. Twee soorten vertekening konden daarbij worden onderscheiden: vertekening doordat de respondent toch weigert een eerlijk antwoord te geven, en vertekening omdat de respondent geen geforceerd 'ja'-antwoord wil geven. Het corrigeren voor deze vertekeningen leidt tot hogere schatters van het gevoelige gedrag. Voor een meer volledige beschrijving van deze technieken wordt verwezen naar Böckenholt en Van der Heijden (2007), Böckenholt, Barlas, en Van der Heijden (2007) en Cruijff, Van den Hout, Van der Heijden, en Böckenholt (2007).

Samenvatting en conclusie

De RR techniek is een methode voor het onderzoeken van gevoelige vragen die zowel methodologen als statistici al ruim 40 jaar bezighoudt.

Onderzoek heeft aangetoond dat RR leidt tot meer valide schatters van gevoelig gedrag, maar ook dat er nog veel onverklaarde variantie tussen studieresultaten is. Ik hoop met dit artikel duidelijk te hebben gemaakt dat met behulp van RR technieken intensievere analyses gedaan kunnen worden dan alleen rechte tellingen uitdraaien. De laatste jaren is er op het terrein van de RR statistisch enorme vooruitgang geboekt. Maar het aardige van RR is dat het onderzoek een mooi voorbeeld is van het samengaan van methoden en statistiek om meer valide inzichten in gevoelige onderwerpen te verkrijgen. Praktijkonderzoek waarbij verschillende RR designs worden vergeleken is daarvoor noodzakelijk. Wilt u meer weten over de toepassingsmogelijkheden van RR dan kunt u terecht op de website <www.randomizedresponse.nl> voor een handleiding.

LITERATUUR

Böckenholt, U. en van der Heijden, P.G.M. (2007, forthcoming). Item randomized-response models for measuring non-compliance: Risk-return perceptions, social influences, and self-protective responses. *Journal of Applied Econometrics*.

Böckenholt, U., Barlas, S., en van der Heijden, P.G.M. (2007). Do randomised-response designs eliminate response biases? An empirical study of non-compliance behavior. *Psychometrika*.

Cruijff, M.J.L.F., van den Hout, A., van der Heijden, P.G.M., en Böckenholt, U. (2007). Log-linear randomized-response models taking cheating into account. *Sociological Methods and Research*.

Lensvelt, L.G.A. (2006). *X-quizit of Game-based vragenlijstonderzoek met kinderen*. Ongepubliceerde bachelorscriptie via www.rlmedia.nl/portfolio.

Lensvelt-Mulders, G., Van der Heijden, P. G. M., Laudy, O., van Gils, Ger (2006). A validation of a computer assisted randomized response survey for measuring fraud in social security, *Journal of the Royal Statistical Society: Series A, Statistics in Society*: 69, 2, 305-318.

Lensvelt-Mulders, G.J.L.M., Hox, J.J., and Van der Heijden, P.G.M. (2005). How to improve the efficiency of randomized response designs. *Quality and Quantity*, 39, 3: 253-265

Lensvelt-Mulders, G.J.L.M., Hox, J.J., van der Heijden, P.G.M., en Maas, C.J.M. (2005). Meta-analysis of randomized response research: Thirty-five years of validation. *Sociological Methods and Research*, 33, 319-348.

Peeters, C.F.W. (2005). *Measuring Politically Sensitive Behavior. Using Probability Theory in the Form of Randomized Response to Estimate Prevalence and Incidence of Misbehavior in the Public Sphere: A Test on Integrity Violations*. Amsterdam: Dynamics of Governance, Faculteit der Sociale Wetenschappen, Vrije Universiteit.

Peeters, C.F.W., Lensvelt-Mulders, G.J.L.M. en Lasthuizen, K. (submitted). A unified randomized-response framework for eliciting sensitive dichotomous and quantitative information.

Rasinski, K. A., Willis, G. B., Baldwin, A. K., Yeh, W., and Lee, L. (1999). Methods of datacollection, perception of risks and losses, and motivation to give truthful answers to sensitive survey questions. *Applied Cognitive Psychology*, 465-484.

Van den Hout, A. en van der Heijden, P.G.M. (2004). The Analysis of Multivariate Misclassified Data With Special Attention to Randomized Response Data. *Sociological Methods and Research* 32: 310-336.

Van den Hout, A. D. L., and Lensvelt-Mulders, G. J. L. M. (2005). On a 2 by 2 factorial design where the use of randomized response is one of the factors. *Statistica Neerlandica*, 59: 4, 434-447.

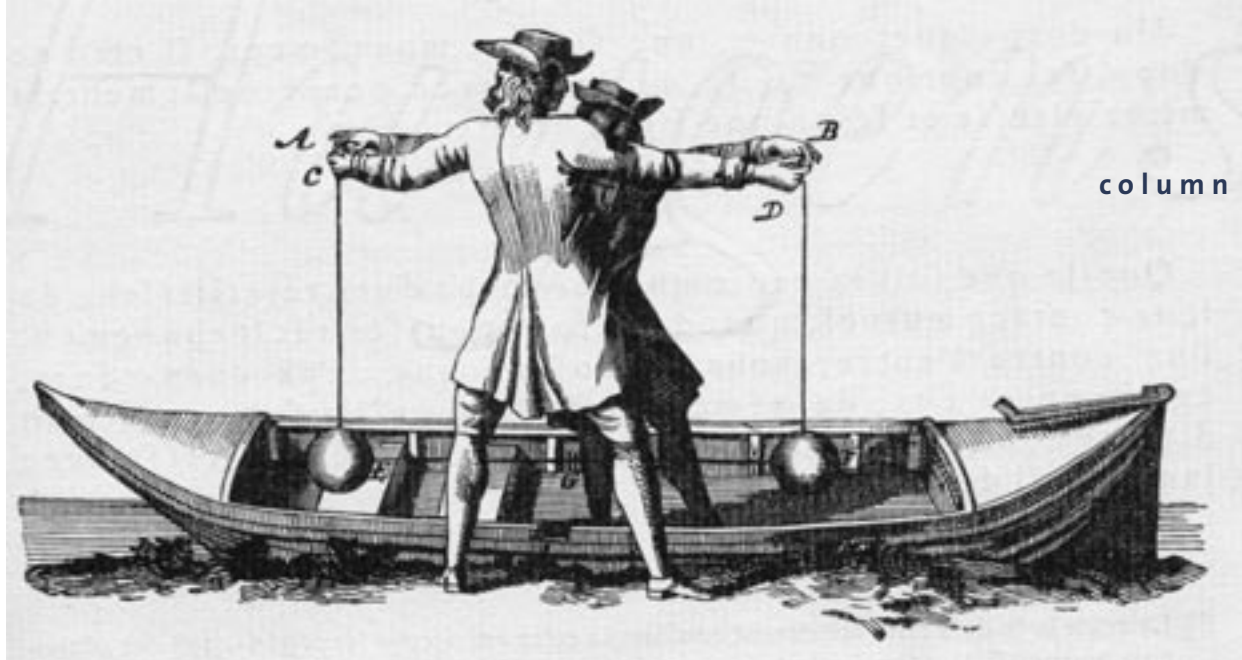
Van der Heijden, P.G.M., van Gils, G., Bouts J., en Hox, J.J. (2000). A comparison of randomized response CASI and face-to-face direct questioning: eliciting sensitive information on the context of welfare and unemployment benefit. *Sociological Methods and Research*, 28: 505-537.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60: 63-69.

Noot

* Bij het departement Methoden en Technieken van de Universiteit Utrecht loopt sinds 1998 een onderzoeksproject met als doel het valideren en verder ontwikkelen van RR procedures voor toepassing in de praktijk. Dit artikel berust op het werk van deze onderzoeksgroep.

GERTY LENSVELT-MULDERS is hoofddocent-onderzoeker bij de capaciteitsgroep Methodenleer en Statistiek van de Universiteit Utrecht.
E-mail: <G.J.L.M.Lensvelt@uu.nl>



column

Christiaan Huygens

ONNO BOXMA

'There iz a right place and a right time feur everythiing', placht Inspector Clouseau alias Peter Sellers te zeggen. Welnu, er is geen betere dag om deze column te schrijven dan 27 april 2007, en geen betere plaats dan Zuilichem. In Zuilichem, een rustig dorpje aan de Waal, staat niet alleen het huis van mijn schoonfamilie, maar het is ook de plek waar Christiaan Huygens verscheidene jaren heeft gewoond, als 'Heer van Zuylichem'.

Canon

Bij de recente discussies over de geschiedenis-canon werd door velen geklaagd over het ontbreken van Christiaan Huygens, wellicht de grootste geleerde die ons land heeft voortgebracht. Men denkt dan vooral aan zijn verhandelingen over botsingswetten, zijn uitvinding van het slingeruurwerk, zijn grote bijdragen aan de theorie van de voortplanting van het licht (het principe van Huygens), en zijn ontdekking van de ringen van Saturnus.

Wij kansrekenaars weten dat Huygens ook belangrijke bijdragen heeft geleverd aan ons vakgebied. Zo was hij degene die als eerste werkte aan het begrip mathematische verwachting. Hij deed dat in een boek, *De ratiociniis in ludo aleae* waaraan hij de laatste hand legde op 27 april 1657 – vandaag precies 350 jaar geleden! Drie jaar later verscheen het in een Nederlandse vertaling, onder de titel *Van Rekeningh in spelen van geluck*.

De laatste regels van het op 27 april 1657 gedateerde voorwoord zijn aardig, en tekenend voor de onderzoeksmores van die tijd:

'Voorts is te weeten, dat al over eenighen tijdt, sommige van de vermaertste wiskonstenaers van geheel Vranckrijck (hij bedoelt Pascal en Fermat) met dese soorte van rekeningh zijn besigh geweest, op dat niemandt hier in, de eer van de eerste inventie die de myne niet en is, my toe en schrijve. Doch sy luyden, offe wel sich onder malkanderen met veele swaere questien ter proeve stelden, soo hebben se nochtans elck sijn

maniere van uytvinding bedeckt gehouden. Soo dat ick van noode gehad heb, alles van vooren aen selfs te ondersoecken en te doorgronden: ende daerom oock noch niet verseeckert en ben, of wy hier in een selfde eerste beginsel getroffen hebben. Maer de uytkomste belangende, heb ick in veele questien ondervonden dat de myne van de haere geensins en verscheelt. U E. sal vinden dat ick in't eynde van dit tractaet, noch eenige van die questien bygevoegt hebbe, achterlaetende nochtans de werckingh; eensdeels om dat ick te veel moeyte te gemoet sagh, indien ick alles nae behooren wilde afdoen; ten anderen om dat my raetsaem dacht, iets overigh te laeten, 'twelck onse lesers (soo der eenige sijn sullen) mochte dienen tot oeffening en tijdt-verdrijf.'

U E. dienstwilligen dienaer Chr. Huygens
van Zuylichem

De laatste zin toont Huygens van een bescheiden en geestige kant. Dat Huygens er 'niet verseeckert en ben' of hij de methoden uit het boek als eerste had bedacht hangt samen met het feit dat veel geleerden uit die tijd hun methoden voor zich hielden, en slechts het probleem en de oplossing publiceerden. Ik ontleen dit voorwoord aan de zeer aardige bewerking van Huygens' boek door Wim Kleijne [2]. Men treft hier de oorspronkelijke tekst rechts in het boek aan, en de moderne tekst links. Kleijne geeft ook aanwijzingen voor het lezen van de tekst van Huygens.

De propositiones van Huygens

Het eerste 'Voorstel' (Propositio) van Huygens luidt: Als ik gelijke kans heb op a of b, dan is mij dit $(a+b)/2$ waard. Ook voorstellen 2 en 3 doen

dit soort uitspraken over verwachtingswaarden. De redenering die tot deze, nu triviaal ogende, resultaten leidt is geraffineerd. Het schijnt dat de overeenkomst met zijn berekeningen aan botsingswetten (De motu, uit 1656) Huygens het idee gaf tot de afleiding van deze resultaten in de kansrekening.

De voorstellen 4 tot en met 9 behandelen variaties op een *problem of stakes* dat Chevalier de Méré had gesteld aan Pascal en Fermat: Stel dat ik tegen een ander om drie gewonnen spelen speel, dat ik al twee spelen heb gewonnen en de ander maar één. Als wij niet zouden willen doorspelen, maar als wij wat ingezet is op een rechtmatige wijze zouden willen verdelen, dan zou ik willen weten hoeveel van de inzet aan mij zou toekomen. Huygens correspondeerde met Pascal, en pikte dit soort problemen op tijdens zijn verblijf in Parijs, midden jaren vijftig.

Ook Newton kan fouten maken

De laatste vijf voorstellen zijn gewijd aan problemen met dobbelstenen. In een prachtige lezing in het Lunterencongres van november 2006 hield Stephen Stigler een geestige verhandeling over het kansrekening werk van Huygens en Newton. Hij onthulde daarin dat ook Newton, in reactie op vragen van Samuel Pepys, werkte aan een probleem met dobbelstenen, te weten: Welk van de volgende drie gebeurtenissen heeft de grootste kans van optreden?

- A. Gooi minstens één 6 in 6 worpen met een zuivere dobbelsteen.
- B. Gooi minstens twee keer 6 in 12 worpen met een zuivere dobbelsteen.
- C. Gooi minstens drie keer 6 in 18 worpen met een zuivere dobbelsteen.

Newton gaf het goede antwoord, wat u niet zal verbazen. Pepys bleef echter doorvragen naar de

aanpak die Newton had gevolgd, en Stigler merkt op dat Newton een aantoonbaar incorrecte aanpak volgt! De geïnteresseerde lezer treft in [3] een uitgebreide bespreking aan van Newton's aanpak. Er zijn geen andere kansrekeningstudies van Newton bekend – wellicht beseftte hij dat dit vak voor hem wat te moeilijk was...

Tot slot

Huygens was een uitzonderlijk groot geleerde, een der weinige tijdgenoten van Newton die niet geheel verbleekte bij deze reus. Canon of niet, het is goed als, onder meer, nieuwe generaties wis- en natuurkundestudenten horen van zijn werk en, via boekjes als dat van Kleijne, inzicht krijgen in zijn redeneertrant. Een uitgebreide biografie over Huygens (met ook veel aandacht voor zijn familie, waaronder vader Constantijn, de dichter en diplomaat!) is geschreven door Andriesse [1]. Op pagina 376 van dat boek vindt u een moderne bespiegeling van Huygens, in diens

posthuum gepubliceerde boek *Kosmotheoros*, over de vraag of er leven is op andere planeten. Huygens spreekt in termen van kansen, en acht buitenaards leven erg waarschijnlijk. In al zijn natuurwetenschappelijk werk ging het trouwens al om de vraag of iets meer of minder waarschijnlijk is. Andriesse citeert een uitspraak van Huygens uit 1673: 'Er is niets dat we met volledige zekerheid weten, maar alles met waarschijnlijkheid, en er zijn sterke verschillen in de mate van waarschijnlijkheid.'

LITERATUUR

1. C.D. Andriesse. *Titan kan niet slapen*. Uitgeverij Contact, Amsterdam, 1993.
2. Christiaan Huygens. *Van Rekeningh in spelen van geluck*. Vertaald en toegelicht door Wim Kleijne. Epsilon Uitgaven, Utrecht, 1998.
3. S. Stigler. Isaac Newton as a probabilist. *Statistical Science* 21(3), 2006, pp. 400-403.

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>

AGENDA

25 juni - 1 juli 2007

Workshop Innovations in Mathematical Finance will be held at Hotel Alexandra in Loen, Norway, in the period 25 June – 1 July, 2007. More information in the Events section of the web page of the workshop http://www.cma.uio.no/conferences/2007/amamef_workshop.html.

29 juli - 2 augustus 2007

The LOC of ISCB-28 warmly invites you to participate in **ISCB's annual conference**. Deadline for Abstracts and Conference Awards applications: 15 Febr. 2007. Final date for early registration: 15 May 2007. Please visit www.iscb2007.gr for details on the program.

4 - 7 maart 2008

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

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 attached file and our website <http://www.mat.uc.pt/~wni2008>.