

STAtOR

periodiek van de VVS jaargang 4 nummer 4 december 2003

Schadeverzekeraars: toereikendheid van reserves
voor toekomstige verplichtingen

Opsporen van sneller en beter: modelling through...

Een onoverbrugbare kloof;
statistiek in de negentiende eeuw

De kracht van de wachtrijtheorie

Levensverwachting;
het CBS bepaalt hoe oud we worden

In Memoriam Prof.dr.ir. Wim Bakker

Gelijke kansen voor iedereen

STATOR

Jaargang 4, nummer 4, december 2003

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

Redactie

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

Kopij en reacties richten aan

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

Bestuur van de VVS

A.W. van der Vaart (voorzitter) <aad@cs.vu.nl>; S.J. Koopman (penningmeester) <s.j.koopman@econ.vu.nl>; namens de Bedrijfssectie (BDS) P. Banens <banens@cqm.nl>; namens de biometrische sectie (BMS) A. Stein <alfred.stein@wur.nl>; namens de economische Sectie (ECS) P.H.F.M. van Casteren <casteren@fee.uva.nl>; namens het Ned. Genootschap voor Besliskunde (NGB) H. Fleuren <fleuren@uvt.nl>; namens de Sectie Mathematische Statistiek (SMS) P. Spreij <spreij@science.uva.nl>; namens de Sectie Statistische Programmatuur (SSP) S.H. Heisterkamp <sh.heisterkamp@rivm.nl>; namens de Sociaal Wetenschappelijke Sectie (SWS) C. Glas <c.a.w.glas@edte.utwente.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

Contactpersoon: Rita Oomen, telefoon 0167 - 563401, fax 0167 - 561200, <japm.oomen@worldonline.nl>. Uiterlijk vier weken voor verschijnen te zenden aan Pharos, Moeflonstraat 5, 6531 JS Nijmegen, telefoon 024 - 3559214, e-mail <hooategem@xs4all.nl>. STATOR verschijnt in maart, juni, september en december.

Ontwerp en opmaak

Pharos / M. van Hooategem, Nijmegen

Uitgever

© Vereniging voor Statistiek en Operationele Research
ISSN 1567-3383

Inhoud

- 3** Studenten en *STATOR*.
Dick den Hertog
- 4** Schadeverzekeraars: toereikendheid van reserves voor toekomstige verplichtingen.
Arno Gabriël
- 9** Opsporen van sneller en beter: modelling through...
Leo Kroon
- 14** Een onoverbrugbare kloof; statistiek in de negentiende-eeuw.
Ida Stamhuis
- 20** De kracht van de wachtrijtheorie. Column.
Onno Boxma
- 22** Levensverwachting; het CBS bepaalt hoe oud we worden.
Kees Prins
- 28** In Memoriam Prof.dr.ir. Wim Bakker.
Henk Zijm en Hein Fleuren
- 29** Gelijke kansen voor iedereen. Column.
Fred Steutel
- 30** Agenda



Studenten en STAtOR

Toen ik zo'n vier jaar geleden bij het departement Econometrie & Operations Research van de Universiteit van Tilburg ging werken, viel het me al snel op dat studenten graag toepassingen willen zien. De meeste studenten stellen keer op keer de vraag: wat kan ik in de praktijk met de theorie die ik nu leer? Op zich een zeer terechte vraag en in ons specialisme is het ook niet moeilijk om toepassingen van de theorie te laten zien. Alhoewel, in de leerboeken op mijn eigen vakgebied Operations Research staan vaak wel praktische voorbeelden, maar het zijn opvallend vaak Mickey Mouse problemen. Niet echt overtuigend voor de studenten.

Aan de andere kant zie ik ook het gevaar dat we op de universiteit studenten gaan klaarstomen voor de praktijk. Dat we hen technieken aanleren in plaats van denkmethodes en dat zij zich te vroeg willen of moeten specialiseren. Ik denk ook niet dat het bedrijfsleven hierop zit te wachten. Het onderwijs moet er in de allereerste plaats op gericht zijn om mensen te leren denken en te leren leren, op het verwerven van niet-beroepsgerichte kennis en vaardigheden die nuttig blijven ook als telkens weer de functies in het bedrijfsleven veranderen. In dit verband kun je je zelfs

afvragen of een stage lopen in een bedrijf wel echt past bij een universitaire opleiding!

Het moet mogelijk zijn om in zeg 10 procent van de college tijd de student het vergezicht van de toepassingen te tonen. Om hen vanuit de praktijk te enthousiasmeren. Dan blijft er 90 procent college tijd over om hen de schoonheid van de theorie te laten zien.

Studenten, en andere geïnteresseerden, laten zien wat ons vakgebied in de praktijk aan mooie toepassingen oplevert, daar wil *STAtOR* graag een rol in spelen. Kopieer gerust eens een artikel uit *STAtOR* dat aansluit bij uw college en geef het aan uw studenten. De redactie van *STAtOR* wil op korte termijn de ouderejaars-studenten benaderen voor een gratis abonnement op *STAtOR*. Vertel het maar vast aan uw studenten.

Veel leesplezier!

Dick den Hertog
hoofdredacteur



SCHADEVERZEKERAARS

toereikendheid van reserves voor toekomstige verplichtingen

Met het sluiten van een verzekeringsovereenkomst neemt de schadeverzekeraar een financieel risico over van de verzekerde. Daarbij kan het gaan om de risico van brand, diefstal, aansprakelijkheid of ziektekosten. Uit de afgesloten verzekeringsovereenkomsten komen voor de verzekeraar verplichtingen voort die enkele maanden tot tientallen jaren in de toekomst kunnen doorlopen. De verzekeraar houdt reserves aan om te kunnen voldoen aan zijn verplichtingen. Het is een taak van de actuaris om de toereikendheid van deze reserves te toetsen. De kwantitatieve methoden die voor deze toetsing worden gebruikt, maakten in de afgelopen jaren een snelle ontwikkeling door.

ARNO GABRIËL

Verzekeringen delen wij in Nederland in drieën: levensverzekeringen, natura-uitvaartverzekeringen en schadeverzekeringen. Bij de eerste twee betreft het verzekerde risico uitsluitend het leven van de verzekerde. Schadeverzekeringen vormen de restcategorie: hieronder vallen alle risico's die geen betrekking hebben op het leven van verzekerde. Schadeverzekeringen bieden daarmee dekking voor een grote variatie aan risico's. Enerzijds zijn dit de zaaksschaden zoals brand, inbraak, diefstal, auto, aansprakelijkheid en rechtsbij-

stand. Anderzijds zijn dit de zorg- en invaliditeitsverzekeringen zoals ziektekosten (medische hulp) en arbeidsongeschiktheid (verlies van inkomen).

In Nederland zijn circa 260 schadeverzekeraars actief; hun gezamenlijke bruto premie-inkomen in 2001 bedroeg 18 miljard euro (bron: Pensioen- en verzekeringskamer, PVK).

Reserve voor toekomstige uitkeringen

Met het afsluiten van een schadeverzekering draagt de verzekerde het in de polisvoorwaarden

omschreven risico over op de verzekeraar. In ruil hiervoor betaalt de verzekerde de verzekeraar een premie. De verplichtingen die de verzekeraar gedurende een boekjaar jegens verzekerde heeft, houden niet op bij het sluiten van de boeken. In de nieuwe verslagperiode heeft de verzekeraar nog altijd verplichtingen op schaden die hebben plaatsgevonden in het oude boekjaar. Zo vindt er op bepaalde schaden een uitloop plaats. Een arbeidsongeschiktheidschade kan bijvoorbeeld leiden tot een verzekerde uitkering gedurende enkele tientallen jaren. Daarnaast zijn er ook geheel nieuwe schademeldingen in het nieuwe boekjaar. Een aanrijding in december wordt pas in de eerste maanden van het nieuwe jaar verwerkt. Een claim uit een aansprakelijkheidsverzekering kan zich zelfs pas vele jaren later openbaren (bijvoorbeeld asbest). Tenslotte kan ook in het nieuwe boekjaar blijken dat de eerder gemaakte inschatting van de omvang van een schadeclaim onjuist is gebleken.

Voor het nakomen van de toekomstige verplichtingen naar verzekerden houdt de verzekeraar reserves aan. Voor de rechtspositie van de verzekerden én voor de continuïteit van de verzekeraar is het van groot belang dat de schadereserves groot genoeg zijn voor het nakomen van deze toekomstige uitkeringen op schaden die plaats hebben gevonden in voorbije jaren. Ultimo 2001 bedroeg de omvang

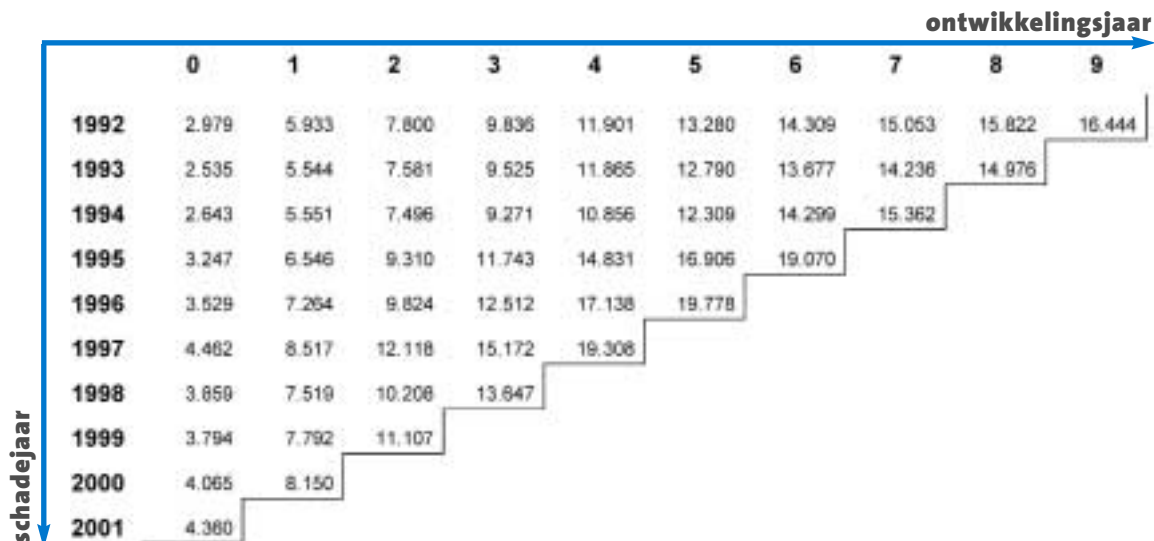
van de reserves van de Nederlandse schadeverzekeraars tezamen 18,7 miljard euro (bron: PVK).

Toereikendheid reserves

De controle op de toereikendheid van de aangelegde reserves is een taak van de actuaaris. Anders dan bij de levensverzekeraar, waar gebruik kan worden gemaakt van overlevingstafels van de Nederlandse bevolking, zijn voor schadeverzekeraars doorgaans geen beschrijvende statistieken van de verzekerde risico's voorhanden. Voor de kwantitatieve toetsing van de reserves van de schadeverzekeraar wordt daarom vaak een trendanalyse uitgevoerd op basis van de gegevens over de gerealiseerde historische schaden. De methoden die de actuaaris hierbij gebruikt maken momenteel een snelle ontwikkeling door, waarbij de traditionele mechanische schattingsmethoden steeds vaker worden vervangen door statistische methoden. Alvorens deze methoden te bespreken wordt eerst de uitlooptriehoek geïntroduceerd. Dit is de typische vorm waarin de schadegegevens doorgaans worden opgetekend.

Uitlooptriehoek

Figuur 1 laat een uitlooptriehoek met betalingen zien behorend bij een portefeuille met aansprakelijkheidsverzekeringen. In dit voorbeeld betreft de informatie de cumulatieve betalingen voor de



Figuur 1.

schadejaren 1991 tot en met 2000, zoals deze op 31 december 2000 bekend waren. Voor het oudste jaar, 1991, is de meeste informatie beschikbaar. Hoe meer recent het schadejaar, hoe minder uitloopjaren in de gegevens bekend zijn. Binnen de driehoek kunnen doorsnijdingen worden gemaakt naar drie verschillende jaren: schadejaren (verticaal), ontwikkelingsjaren (horizontaal) en kalender-/boekjaren (diagonaal). De driehoek 'groeit' ieder jaar met één nieuwe diagonaal.

De driehoek in figuur 1 laat zien dat ook in de latere ontwikkelingsjaren (de punt rechtsboven) de cumulatieve betalingen nog toenemen. Dit is typisch voor een *long tail* risico als aansprakelijkheid: de schadebetalingen kunnen tot tientallen jaren in de toekomst doorlopen. Andere long tail risico's zijn arbeidsongeschiktheid en letsel. Brand-, transport- en reisverzekeringen zijn daarentegen typische *short tail* risico's.

Mechanische methoden: chain ladder

Voor het maken van een schatting van de omvang van de toekomstige uitkeringen kan de actuaaris een keus maken uit één van de traditionele mechanische methoden. De bekendste daarvan is de *pure chain ladder*-methode, welke hierna wordt

toegelicht. Een alternatief voor de *chain ladder*-methode vormt de separatiemethode. Beide kennen vele varianten. (Kaas, Goovaerts, 1998).

De *chain ladder*-methode schat de toekomstige betalingen tot en met het laatste ontwikkelingsjaar. Elk schadejaar wordt van links naar rechts uitontwikkeld zodat de uitlopdriehoek wordt aangevuld tot een vierkant. De betalingen in ontwikkelingsjaar d worden vastgesteld door de betalingen in voorgaand ontwikkelingsjaar $d-1$ te vermenigvuldigen met de *link ratio* voor die periode. De *link ratio* is een factor en wordt voor elk ontwikkelingsjaar vastgesteld als de gemiddelde historische groei van de betalingen in die periode.

In figuur 2 is deze methode toegepast op de gegevens uit figuur 1. De onderste rij in figuur 2 geeft de link ratios per ontwikkelingsperiode. Voor ontwikkelingsjaar 7 bedraagt de link ratio $(15.053 + 14.236 + 15.362) / (14.309 + 13.677 + 14.299) = 1,06$. De betalingen voor schadejaar 1994 zullen in boekjaar 2001 naar schatting dus toenemen van 19.070 tot $19.070 * 1,06 = 20.214$. In boekjaar 2002 zullen de betalingen voor schadejaar 1994 naar schatting toenemen tot $20.214 * 1,05 = 21.224$.

De totale geschatte verplichtingen worden nu bepaald door de geschatte cumulatieve be-

schadejaar	ontwikkelingsjaar									
	0	1	2	3	4	5	6	7	8	9
1992	2.979	5.933	7.800	9.836	11.901	13.280	14.309	15.053	15.822	16.444
1993	2.535	5.544	7.581	9.525	11.865	12.790	13.677	14.236	14.976	15.575
1994	2.643	5.551	7.496	9.271	10.856	12.309	14.298	15.362	16.130	16.775
1995	3.247	6.546	9.310	11.743	14.831	16.906	19.070	20.214	21.224	22.073
1996	3.529	7.264	9.824	12.512	17.138	19.778	21.953	23.270	24.434	25.411
1997	4.462	8.517	12.118	15.172	19.308	21.817	24.217	25.670	26.954	28.032
1998	3.859	7.519	10.208	13.647	17.195	19.430	21.568	22.862	24.005	24.965
1999	3.794	7.792	11.107	14.106	17.773	20.084	22.293	23.631	24.812	25.805
2000	4.055	8.150	11.247	14.284	17.998	20.338	22.575	23.930	25.126	26.131
2001	4.350	8.807	12.153	15.434	19.447	21.975	24.393	25.856	27.149	28.235
	fr	2,02	1,38	1,27	1,26	1,13	1,11	1,06	1,05	1,04

Figuur 2.

talingen in ontwikkelingsjaar 9 (rechter kolom) de werkelijke betalingen aan het eind van 2000 (diagonaal) in mindering te brengen. Indien deze schatting lager is dan de feitelijk door de verzekeraar aangehouden reserves kunnen we concluderen dat deze reserves toereikend zijn.

Chain ladder methode: voors en tegens

Eenvoud is het grote voordeel van de *chain ladder*-methode. Er komt geen gecompliceerde modellering aan te pas, een kleine spreadsheet-berekening is voldoende. De tijd voor het uitvoeren van een *chain ladder*-berekening is beperkt. Bovendien genereert de methode, onder voorwaarden, zuivere schattingen voor de reserve en kan een verdelingsvrije standaarddeviatie voor de resultaten worden berekend (Mack, 1993).

Duidelijk moge tevens zijn dat de *chain ladder*-methode veel tekortkomingen kent, waaronder:

- De reserve wordt alleen geschat op basis van de ontwikkeling in de uitloopjaren (= horizontale ontwikkeling). Andere ontwikkelingen zoals inflatie worden niet meegenomen. Het niet modelleren van deze ontwikkelingen kan resulteren in een slechte *fit* met de gegevens en een onderschatting van de toekomstige verplichtingen.
- De uitloop van de schadeverplichtingen hoeft niet te zijn beëindigd na het laatste ontwikkelingsjaar in de gegevens. De *chain ladder*-methode biedt niet de mogelijkheid deze staartuitloop te schatten.
- De *chain ladder* levert alleen een verdelingsloze puntschatter voor de toekomstige schadebetalingen. Er kunnen geen uitspraken worden gedaan over de betrouwbaarheidsintervallen waarbinnen deze schatter zich bevindt.

Het gebruik van statistische methoden en regressie biedt de mogelijkheid tot een verbeterde modellering en het verkrijgen van meer informatie uit de uitlooptriehoek met gegevens.

Statistische modellering

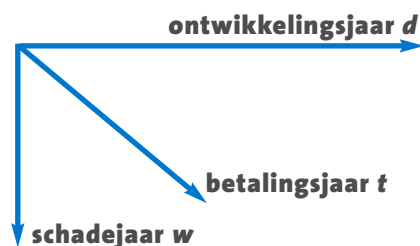
De recente jaren hebben een verbeterde toegang

tot historische gegevensbestanden bij verzekeraars laten zien. In plaats van waarnemingen per jaar zijn vaak ook driehoeken met waarnemingen per kwartaal of maand beschikbaar. De gegevens kunnen ook steeds beter worden gesplitst naar de verschillende verzekeringsproducten en/ of risico's. Met het verbeteren van de kwaliteit en kwantiteit van de gegevens is de statistische modellering van de toekomstige verplichtingen zich sterk aan het ontwikkelen.

De statistische methoden berusten veelal op regressietechnieken waarbij de storingsterm verondersteld wordt een dikstaartige verdeling te volgen. Voor het uitvoeren van de regressie-analyse kan gebruik worden gemaakt van statistische programmatuur, zoals het pakket GLIM. Echter, steeds meer toegepaste software is verkrijgbaar voor het uitvoeren van regressie-analyses op een dataset met de structuur van een uitlooptriehoek. Een bekend pakket is het Australische ICRFS.

ICRFS biedt de mogelijkheid een lognormale regressie op de gegevens in de uitlooptriehoek uit te voeren. Het model kan hierbij trends in drie verschillende richtingen schatten: ontwikkelingsjaar (d), schadejaar (w) en betalingsjaar (t) (zie figuur 3). Dit in tegenstelling tot de *chain ladder* methode, waarin slechts trend d wordt bestudeerd. De trend w volgt de risico-exposure. Deze risicomaat is afhankelijk van de ontwikkeling in onder andere het aantal verzekerde polissen of de polisvoorwaarden. De trend t kan worden gezien als ontwikkeling in inflatie.

Merk op dat de drie trendrichtingen niet onafhankelijk zijn. De vector t kan worden geïnterpreteerd als de som van d en w . Meerdere statistische



Figuur 3.

oplossingen zijn derhalve mogelijk en er bestaat geen algoritme dat na een druk op de knop met hét resultaat komt. Het bouwen van het best passende regressiemodel geschiedt derhalve door middel van een iteratief proces. Aan de hand van residuenplots worden trendbreuken in de drie richtingen d , w en t bestudeerd (zie figuur 4 voor een schermafdrruk). Met kennis van historische schade-ontwikkeling en inflatie legt de actuaaris hierbij restricties op aan de trends. De fit van het regressiemodel wordt getoetst met de gebruikelijke statistieken, waaronder R-kwadraat, informatie-criteria en verdeling van de storingstermen. De stabiliteit van een model wordt getoetst door het toe te passen op dezelfde uitlooptriehoek, echter nu zonder de onderste diagonaal. In het voorbeeld schat het model dan, op basis van de gegevens tot en met 1999, de schadebetalingen in het jaar 2000. Deze kunnen worden vergeleken met de realisaties in 2000.

Als een passend model is gevonden wordt een schatting van de toekomstige betalingen gemaakt. Door het voortzetten van de trends kunnen ook schattingen worden gemaakt voor de perioden ná de laatste ontwikkelingsperiode in de gegevens-driehoek. Naast de standaarddeviatie worden ook de kwantilen bij de verdeling van de totale toekomstige betalingen gegenereerd. Dankzij deze informatie kunnen ook uitspraken worden

gedaan over de mate van toereikendheid van de reserves van de schadeverzekeraar, in plaats van alleen het wel of niet toereikend zijn.

Het Financieel Toetsingskader

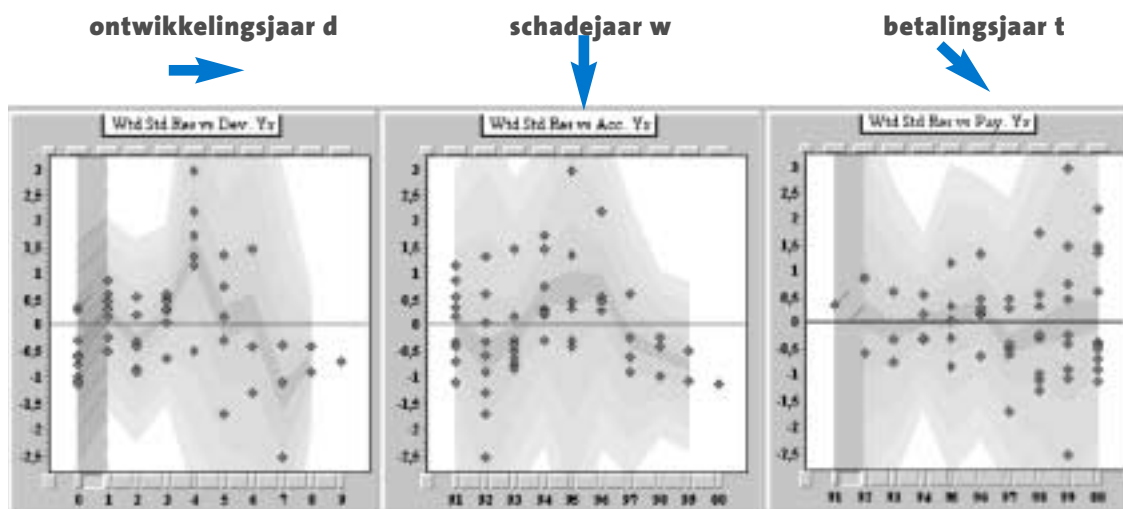
Vanaf 2006 moeten alle verzekeraars in Nederland voldoen aan het Financieel Toetsingskader (FTK). Dit is de nieuwe richtlijn die de PVK uit hoofde van haar rol als toezichthouder oplegt aan de solvabiliteitspositie van verzekeraars en de actuariële toetsing daarvan. Binnen het FTK spelen statistische toereikendheidstoetsen op korte, middellange en lange termijn een grote rol. De komende jaren zullen de statistische schademodelen voor de individuele verzekeringsproducten zich dan ook verder ontwikkelen tot het niveau van de gehele portefeuille van de verzekeraar. Voor de actuaaris ligt hierbij de grootste uitdaging in de modellering van de onderlinge correlaties tussen de verschillende verzekerde risico's.

LITERATUUR

Kaas R. & Goovaerts M.J. (1998). *Inleiding risictheorie*. Amsterdam: Universiteit van Amsterdam

Mack Th. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, Vol. 23, No. 2, 1993

ARNO GABRIËL is werkzaam voor Ernst & Young Actuarissen in Utrecht. E-mail: <Arno.Gabriel@nl.ey.com>.



Figuur 4..



Foto: P. Bosch

Opsporen van sneller en beter: modelling through...

Er gaat bijna geen dag voorbij of de Nederlandse Spoorwegen krijgen aandacht van de media: te lage punctualiteit, tekort aan materieel en personeel, rondjes om de kerk, blaadjes op de rails, vierkante wielen, zomer-, herfst- en winterdienstregelingen, agressie in en rond de trein, enzovoort. De 16 miljoen directieleden van de NS veronderstellen soms wat al te gemakkelijk dat de problemen op een relatief simpele wijze opgelost kunnen worden, wellicht zoals bij een modelspoorbaan. Hoewel Märklin-management natuurlijk erg leuk is, gaat het in de praktijk niet om het laten rijden van treinen, maar om het vervoeren van grote aantallen reizigers, en het liefst op een zo comfortabel mogelijke wijze.

LEO KROON

Hoewel Nederland een relatief klein landje is, is het spoorwegsysteem een bijzonder complex systeem, zowel de planning ervan als de uitvoering. Kortom, het is een interessant toepassingsgebied voor tal van Operations Research (OR) modellen.

Op de korte termijn moeten daar overigens geen wonderen van worden verwacht: verbeteringen van de punctualiteit zullen voorlopig vooral het resultaat moeten zijn van vergroting van de betrouwbaarheid van de infrastructuur en het

materieel en ook van meer materieel en personeel. Maar op de langere termijn kan de toepassing van OR modellen een nuttige rol spelen bij het verbeteren van de kwaliteit van het spoorwegsysteem, zowel bij de strategische capaciteitsplanning ervan als bij de operationele vervoersplanning en de real-time bijsturing.

De kwadratuur van de cirkel

De planning en de uitvoering van het spoorwegsysteem dienen er op gericht te zijn om de Reis van de Klant zo goed mogelijk te faciliteren. Daarbij spelen vooral de volgende logistieke aspecten een belangrijke rol: een korte reistijd, zo veel mogelijk rechtstreekse verbindingen, hoge frequenties, een hoge punctualiteit, en een grote zitplaatskans. Daarnaast spelen ook andere aspecten een rol, zoals goede informatie vooraf en gedurende de reis, en gunstige tarieven.

Het streven naar al deze aspecten tegelijkertijd is echter vergelijkbaar met het zoeken naar de kwadratuur van de cirkel, zeker op de korte termijn. Dit komt doordat in ieder geval de logistieke doelstellingen met elkaar conflicteren. Er is in feite sprake van een overbepaald systeem met meer beperkingen dan variabelen, en dat leidt meestal niet tot een toegelaten oplossing.

Enkele voorbeelden. Vooral een korte reistijd is voor de reizigers heel belangrijk. Een korte reistijd vergt hoge snelheden, weinig haltingen, veel rechtstreekse verbindingen, en veel mooie aansluitingen. Maar hoge snelheden, en vooral hoge snelheidsverschillen, conflicteren met hoge frequenties. Veel rechtstreekse verbindingen vergen lange lijnen en veel haltingen. Maar lange lijnen zijn weer slecht voor de punctualiteit, en veel haltingen zijn slecht voor de reistijden. Tenslotte zijn ook veel mooie aansluitingen minder goed voor de punctualiteit, en staan ze een efficiënt gebruik van de infrastructuur in de weg.

Het ideale spoorwegsysteem bestaat dus niet, en er moeten daarom voortdurend keuzes gemaakt

1×10^6 reizigers per werkdag
 14×10^9 reizigerskilometers per jaar
2.800 kilometer spoor
2.650 bakken (wagons)
350 stations
3.100 machinisten
5.000 treinritten per werkdag
3.700 conducteurs en treinsurveillanten

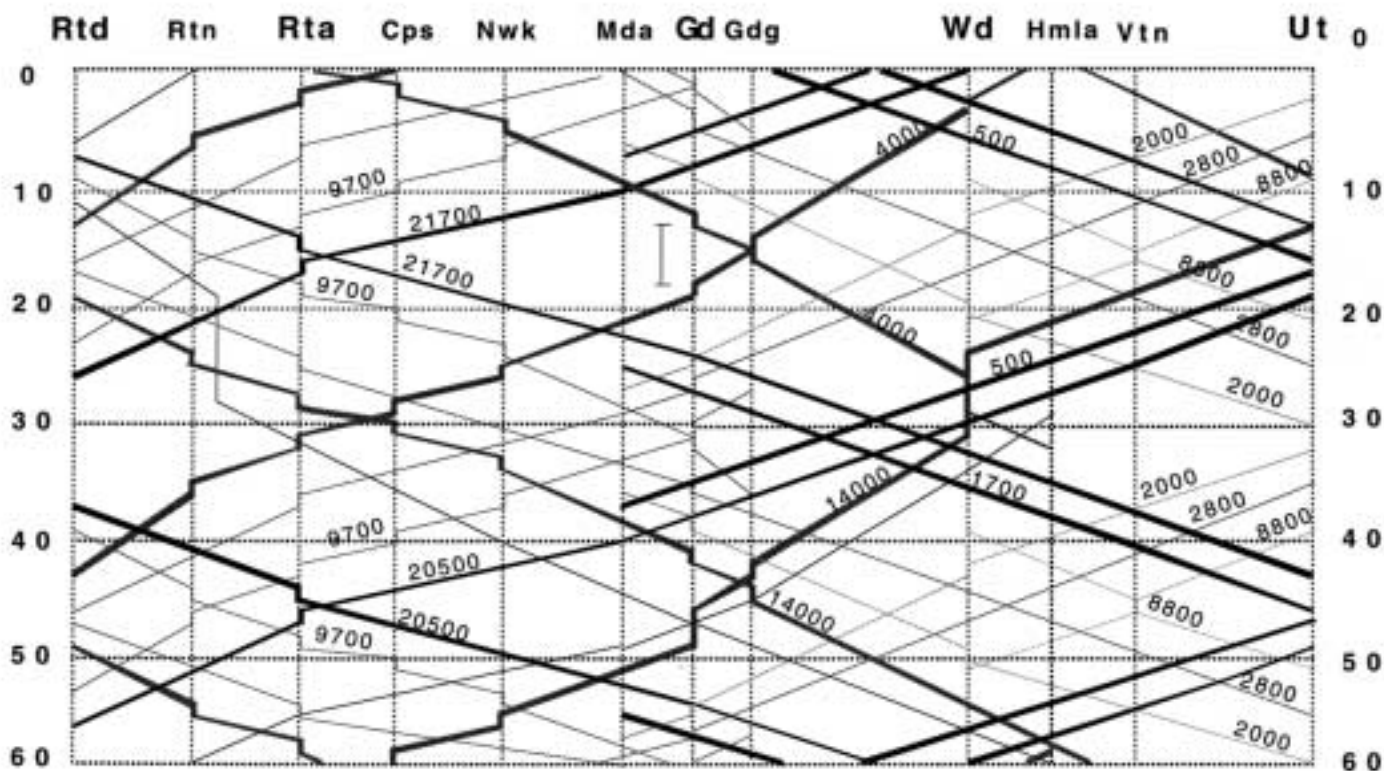
worden bij het ontwerpen van een dergelijk systeem. Daarbij is het dan natuurlijk handig om te beschikken over geschikte tools die modelmatige ondersteuning kunnen bieden bij het bepalen van de relevante *trade offs*.

Operationele planning

De globale structuur van het spoorwegsysteem wordt bepaald door de infrastructuur en de lijnvoering: iedere lijn heeft een beginpunt en een eindpunt, alsmede een frequentie en een type (IC, Snel, Stop). Relevante criteria bij het bepalen van de lijnvoering zijn de reistijden en de aangeboden rechtstreekse verbindingen voor de reizigers, alsmede de operationele kosten voor de vervoerder.

Om grote aantallen reizigers over het spoor te kunnen vervoeren, moeten verder veel operationele zaken gepland worden. Enkele daarvan volgen hieronder.:

- In de dienstregeling wordt bepaald hoe laat de treinen rijden. De dienstregeling van NS Reizigers is in grote lijnen cyclisch met een cycluslengte van een uur. Figuur 1 toont een klein stukje van de dienstregeling in de vorm van een tijd-weg-diagram voor het baanvak Rotterdam (Rtd) - Utrecht (Ut). In verticale richting staat de tijd en in horizontale richting staat de weg. De diagonale lijnen representeren de relaties tussen tijd en weg voor de verschillende treinen. De getallen in de figuur geven de treinnummers in algemene vorm weer.



Figuur 1: Tijd-weg-diagram voor het baanvak Rotterdam - Utrecht

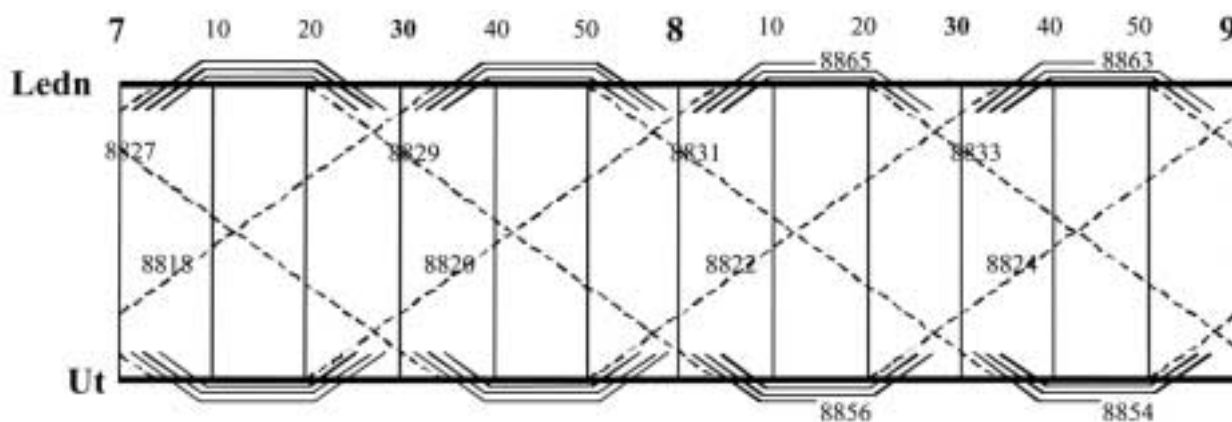
- De perrontoewijzingen leggen vast op welke perons de treinen op de stations aankomen en vertrekken. Evenals de dienstregeling is de perrontoewijzing in hoge mate cyclisch met een cyclustijdsduur van een uur.
- De materieelinzet bepaalt op welke manier de verschillende treinritten van materieel worden voorzien. Belangrijke aspecten hierbij zijn: een hoge effectieve zitplaatscapaciteit, niet te veel bakkilometers, en ook niet te vaak aftrappen en bijplaatsen in verband met de robuustheid van het systeem. Bovendien moet het onderhoud van het materieel gefaciliteerd worden.
- De personeelsinzet bepaalt hoe de treinen van rijdend personeel worden voorzien. Elke machinist en conducteur werkt in diensten vanuit een vaste standplaats. Per standplaats worden de diensten verwerkt tot roosters. Zowel de diensten als de roosters dienen aan tal van complexe regels te voldoen. Daarbij moeten de diensten en de roosters natuurlijk wel zo efficiënt mogelijk zijn.
- De rangeerplanning vormt het sluitstuk van de

logistieke planning. Iedere wijziging in de dienstregeling en/of de materieelinzet vergt aanpassingen van een of meerdere rangeerplannen. De rangeerplanning wordt steeds complexer, o.a. door de instroom van veel nieuw materieel en door de steeds strenger wordende geluidsnormen gedurende de nacht.

Lange termijn planning

Naast de voorgaand beschreven meer operationele planningsproblemen moet er natuurlijk ook naar de langere termijn gekeken worden. Hoe moet het spoorwegsysteem er op de langere termijn uit zien om aan de dan bestaande vervoersvraag te kunnen voldoen, en hoeveel infrastructuur, materieel en personeel zijn daar dan voor nodig? Een belangrijk punt hierbij is dat de vervoersvraag voor een deel bepaald wordt door de kwaliteit van het geleverde product.

Kenmerkend voor dergelijke strategische vraagstukken is dat ze een zeer lange planningshorizon hebben: alleen al het aanleggen van nieuwe



Figuur 2: Materieelinzet op de stoptrein Utrecht (Ut) - Leiden (Ledn)

infrastructuur kost vele jaren, zie de aanleg van de Betuweroute en de Hogesnelheidslijn, en als de infrastructuur eenmaal is aangelegd, dan ligt die er over honderd jaar waarschijnlijk nog. Rollend materieel heeft een levertijd van enkele jaren en een levensduur van tientallen jaren. Vanwege de zeer lange planningshorizon, die uiteraard ook een grote mate van onzekerheid met zich meebrengt, zijn uitgebreide gevoeligheidsanalyses van groot belang.

Modelmatige ondersteuning van de planning

Hoewel er bij het maken van de eerder genoemde operationele plannen gebruik gemaakt wordt van bijna oneindig veel computers (voor het vastleggen, controleren, presenteren en distribueren van de plannen), berust het bedenken van deze plannen tot dusverre nog grotendeels op het vakmanschap en de ervaring van de planners. De komende tijd zal nog veel onderzoek op het gebied van de geautomatiseerde modelmatige planningsondersteuning noodzakelijk zijn.

Maar dankzij de recente ontwikkelingen op het gebied van hard- en software kunnen bepaalde plannings- en schedulingproblemen, waar men enkele jaren geleden nog niet van durfde te dromen, momenteel 'optimaal' opgelost worden. Helaas zijn de plannings- en schedulingproble-

men in de spoorwegwereld dermate omvangrijk dat voorlopig alleen deelproblemen opgelost kunnen worden. De menselijke planners zullen daarom voorlopig zeker ook onmisbaar blijven.

Wat betreft de geautomatiseerde modelmatige planningsondersteuning is er de laatste jaren bijvoorbeeld vooruitgang geboekt op het gebied van het ontwerpen van dienstregelingen en perrontoewijzingen, en op het gebied van de inzetplanning van materieel en personeel. Voor een voorbeeld van dat laatste kan verwezen worden naar een van de voorgaande nummers van STATOR.

Figuur 2 representeert een klein deel van de materieelinzet op de stoptreinen op het traject Utrecht (Ut) - Leiden (Ledn). Er is hier weer sprake van een tijd-weg-diagram. De stippellijnen representeren de treinen, voorzien van de bijbehorende treinnummers, en de 'bogen' in de figuur zijn de treinstellen. Zo rijdt trein 8829 met drie treinstellen van Leiden naar Utrecht. In Utrecht wordt één treinstel afgekoppeld, dat pas in de middagspits weer wordt ingezet op trein 8856. De rest van de trein rijdt met twee treinstellen terug naar Leiden als trein 8824.

Het plannen van een dergelijke relatief eenvoudige materieelinzet van treinstellen van een enkel materieeltype op een enkele lijn kan gemodelleerd worden met een simpele uitbreiding van het Minimum Cost Flow model. Wanneer er echter sprake is van treinstellen van verschillende mate-

rieeltypen per trein, dan wordt ook de volgorde van de treinstellen in iedere trein belangrijk. Dit laatste heeft te maken met de rangeermogelijkheden per trein. Om met dit aspect ook modelmatig rekening te kunnen houden dient gebruik gemaakt te worden van een complexe uitbreiding van het Integer Multi-Commodity Minimum Cost Flow model. Met dergelijke modellen zijn inmiddels voor meerdere treinseries materieelomlopen berekend die zowel op de effectieve zitplaatscapaciteit als op de operationele kosten beter scoren dan de handmatig bepaalde materieelomlopen. Maar meer onderzoek op dit gebied is nodig, want de benodigde rekentijden voor dergelijke complexe modellen lopen snel op met de omvang van de probleeminstanties.

Simulatiemodellen kunnen een rol spelen bij het analyseren van de robuustheid en de stabiliteit van het spoorwegsysteem. De robuustheid geeft aan hoe gevoelig het systeem is voor externe verstoringen. Dat wil zeggen: in welke mate kunnen verstoringen zich in ruimte en tijd verspreiden. De stabiliteit van een systeem geeft aan hoeveel moeite het kost om het systeem weer te laten terugkeren naar de evenwichtstoestand, nadat het daar door verstoringen van afgeweken is.

Tenslotte lijken OR-modellen ook een belangrijke rol te kunnen gaan spelen bij de real-time bijsturing van het vervoersproces. De bijsturing dient er voor te zorgen dat de uitvoering van de dienstregeling zo veel mogelijk volgens de dienstregeling verloopt. Binnen de bijsturing moeten, op het moment dat er verstoringen plaatsvinden, onder tijdsdruk beslissingen genomen worden om de effecten van de verstoringen zo veel mogelijk te beperken. In verstoorde situaties is vooral het bijsturen van materieel en personeel een lastige klus. Modelmatige beslissingsondersteuning op dit punt is in de spoorwegwereld momenteel nog een vrijwel onontgonnen gebied. Maar in de luchtvaartwereld heeft onderzoek op dit gebied al tot echt werkende beslissingsondersteunende systemen geleid.

Slotopmerkingen

Hoewel dat voor de reizigers op dit moment nog onvoldoende waarneembaar is, waait er de binnen de spoorwegwereld een innoverende wind. Zo wordt er binnen de afdeling Logistiek van NS Reizigers gestreefd naar een verbetering van de kwaliteit van de logistieke plannen, alsmede naar een kortere doorlooptijd van het logistieke planproces. Daarbij is ook een rol weggelegd voor de toepassing van OR-modellen en technieken. Een kortere doorlooptijd is van belang in verband met de flexibiliteit van het systeem: hoe korter die doorlooptijd, hoe sneller het systeem kan reageren op veranderingen in de omgeving. Andere innovatieve verbeterprojecten in de spoorwegwereld hebben betrekking op het beschikbaar hebben van een bedrijfszekere infrastructuur en van voldoende bedrijfszekere materieel. Ook de ontwikkeling van de elektronische chipcard voor het openbaar vervoer is een goed voorbeeld. Deze kaart moet voor een grotere sociale veiligheid in het openbaar vervoer gaan zorgen. Helaas kosten wezenlijke verbeteringen in omvangrijke en complexe systemen veel tijd, zeker als tijdens de verbouwing de verkoop gewoon door gaat.

Als er maar voldoende prioriteit aan de verbetering en de innovatie van het spoorwegsysteem gegeven wordt zal dat op termijn leiden tot een kwalitatief hoogstaand systeem dat kan concurreren met de beste in de wereld. Ik zie daarbij ook een duidelijke rol weggelegd voor de toepassing van OR-modellen, zowel bij de strategische capaciteitsplanning als bij de operationele vervoersplanning en bij de real-time bijsturing.

LEO KROON is adviseur bij de afdeling Logistiek van NS Reizigers. Daarnaast is als bijzonder hoogleraar verbonden aan de Faculteit Bedrijfskunde aan de Erasmus Universiteit Rotterdam. E-mail: <lg.kroon@reizigers.ns.nl>.

* Een verkorte versie van de oratie die de auteur op 21 september 2001 heeft uitgesproken bij het aanvaarden van het ambt van bijzonder hoogleraar van de Erasmus Universiteit Rotterdam.

Statistiek in de negentiende eeuw,

Een onoverbrugbare kloof

In de negentiende eeuw werden twee vormen van statistiek bedreven. De ene vorm werd gekenmerkt door juristen die zich vooral toelegden op de systematische en uitvoerige beschrijving van een staat en de andere door een meer wiskundige aanpak waarbij statistiek beperkt bleef tot kwantitatieve gegevens over de bevolking en kansrekening gebruikt werd voor de interpretatie van de gegevens. Tussen de juridische en wiskundige aanpak gaapte een onoverbrugbare kloof.



IDA H. STAMHUIS

In de negentiende eeuw was er sprake van twee vormen van statistiek die twee essentieel verschillende intellectuele tradities vertegenwoordigden: die van de exacte wetenschappen en die van de letteren. In de jaren vijftig van de twintigste eeuw trok de Engelse natuurkundige en schrijver C.P. Snow de aandacht met zijn geschrift *The Two Worlds* waarin hij de vinger legde op de zere plek van de miscommunicatie tussen de werelden van

de exacte wetenschappen en van de letteren. In vorige artikelen heb ik die twee vormen van statistiek besproken: die van de juristen-statistici, die zich in 1857 in de Vereeniging voor de Statistiek organiseerden, en die van de wiskundige Lobatto.¹ In dit artikel breng ik de onoverbrugbare kloof tussen deze twee manieren van het bedrijven van statistiek ter sprake en de consequenties voor de officiële statistiekbeoefening in Nederland.²

Twee werelden

Juristen kregen, vaak verplicht, onderwijs in de statistiek, waarbij onder statistiek de systematische beschrijving van een staat werd verstaan. Om een goede indruk van een staat te krijgen, moest zo'n beschrijving zo compleet mogelijk zijn en een staat werd daardoor hoofdzakelijk op kwalitatieve wijze behandeld. Er werden kwantitatieve elementen opgenomen, wanneer die beschikbaar waren, maar de behandeling was in essentie kwalitatief. Dit kennisgebied stond in de Duitse traditie, die zowel *Staatkunde* als *Statistik* werd genoemd. Aan de andere kant was er de wiskundige Rehuel Lobatto, die zich bij het bedrijven van statistiek hoofdzakelijk beperkte tot kwantitatieve gegevens over de bevolking en bij de interpretatie van die gegevens kansrekening gebruikte. Hij werkte in de traditie van de 'Politieke Rekenkunde', waartoe ook de levensverzekeringswiskunde behoorde.

De juristen en Lobatto hadden wel beide het idee dat ze met statistiek bezig waren. De *doelstellingen* van beide manieren van statistiek vertoonden grote overeenkomsten: een groeiende behoefte aan een meer abstract en formeel overheidsbestuur, gebaseerd op duidelijke en correcte informatie. De *middelen* om deze doelen te bereiken waren vooralsnog verschillend. Op twee manieren kan ik laten zien hoe verschillend beider benadering van de statistiek was en hoe groot de kloof tussen beide manieren van denken. Eerst behandel ik hoe ze op elkaar reageerden en daarna hoe ze reageerden op L.A.J. Quetelet (1796-1866), de Belgische statisticus die beide benaderingen heeft geprobeerd te verenigen.

Reactie op elkaar

Een manier om te weten te komen hoe Lobatto en de juristenstatistici reageerden op elkaars wijze van statistiek bedrijven, is na te gaan hoe de sectie statistiek in het zogeheten *Jaarboekje van Lobatto* werd ontvangen door de juristenstatistici en wat de reactie van Lobatto daar dan weer op was. In 1826 had

Lobatto toestemming gekregen om een jaarboekje uit te geven met allerlei wetenswaardigheden. Dit boekwerk bevatte tevens jaarlijks een sectie statistiek, waarin allerlei populatiegegevens werden opgenomen, zoals sterfte- en geboorteaantallen. Ook bevatte het een sterftetafel, gebaseerd op sterftegegevens van de bevolking van Amsterdam. Opvallend is dat verschillende tijdschriften die meer of minder gewijd waren aan de statistiek op de wijze van de juristenstatistici Lobatto's *Jaarboekje* niet bespraken. Ik heb alleen een bespreking gevonden in *De Vriend des Vaderlands*. In 1838, dertien jaar na de eerste jaargang, heeft de jurist J. Ackersdijck, hoogleraar in de statistiek te Utrecht, het statistische gedeelte van dit jaarboekje uitgebreid besproken. Wat zou zijn commentaar op Lobatto's statistiek zijn?

Eerst maakte hij een opmerking over het belang van de statistiek in het algemeen: 'Het geldt niet ijdele bespiegelingen of berekeningen; maar de kennis van werkelijke waarheden van hoog aanbelang, welke menigvuldig ten grondslag dienen moeten voor maatregelen, die op het welzijn der burgers onvermijdelijken invloed hebben; zoodat, wanneer men uit gebrek van kennis dier daadzaken mistast de gevolgen altijd noodlottig zijn.' Voor Ackersdijck was statistiek dus veel meer dan populatiestatistiek. Het ging om feiten die van belang waren voor het staatsbestuur.

De inhoud van de sectie statistiek van het *Jaarboekje* beschreef hij als volgt: 'De meeste der mededeelingen betreffen de bevolking.' Hij besprak daarna elk onderwerp en elke tabel minutieus en deed voorstellen tot kleine wijzigingen, zoals het aanbrengen van een andere categorie in een tabel, of hij wees op het risico van het dubbel tellen van gegevens. Hij onderstreepte ook het belang van een sterftetafel voor levensverzekeringen. Populatiestatistiek waren voor Ackersdijck zeker de moeite waard, maar hij schreef ook: 'Hiermee afstapende van het onderwerp der bevolking, kunnen wij den wensch niet onderdrukken, dat omtrent

beeldwijze te doen kennen, welke tot eene aanwenselijke waarde van W kan leiden, geldende voor alle waarden van a , en waardoor tevens de juistheid der door Ponsse verkregen uitkomst zal bevestigd worden.

Herneem wij te dien einde de algemeene form. (3)

$$W = \frac{1}{\pi} \int_{-\pi}^{\pi} e^{-ax} e^{-1} \frac{\sin ax}{x} \left[\int_{-\pi}^{\pi} y e^{ax} e^{-1} dx \right]^2 dx,$$

waarin geen bijzondere aanneming ten aanzien der functie y is aangenomen.

Stellen wij $y^2 - 1 = i$, dan is voorts

$$\int_{-\pi}^{\pi} y e^{ax} e^{-1} dx = \int_{-\pi}^{\pi} y \cos ax dx + i \int_{-\pi}^{\pi} y \sin ax dx.$$

Ontwikkel wij thans $\cos ax$, $\sin ax$ in de bekende reeksen, en stellen wij hierbij kortwiskunde,

$$\int_{-\pi}^{\pi} y x^k dx = k' \int_{-\pi}^{\pi} y x^k dx = k'' \int_{-\pi}^{\pi} y x^k dx = k''' \text{ enz.}$$

$$\int_{-\pi}^{\pi} y dx = k_1 \int_{-\pi}^{\pi} y x^2 dx = k_2 \int_{-\pi}^{\pi} y x^4 dx = k_3 \text{ enz.}$$

dan komt er

$$\int_{-\pi}^{\pi} y \cos ax dx = 1 - \frac{1}{2} a^2 k' + \frac{1}{2 \cdot 3 \cdot 4} a^4 k'' + \text{enz.} = 1 - k,$$

$$\int_{-\pi}^{\pi} y \sin ax dx = ka - \frac{1}{1 \cdot 3} a^3 k_1 + \frac{1}{2 \cdot 3 \cdot 4 \cdot 5} a^5 k_2 + \text{enz.} = B,$$

$$\text{Dus } \int_{-\pi}^{\pi} y e^{ax} e^{-1} dx = 1 - (k - iB) = e^{2a} (1 - (k - iB))$$

$$= e^{-2a} (1 - (k - iB)) = e^{2a} (1 - k + iB).$$

Uit de ontwikkeling van den exponent van a , zal men niet weinig moeite inzien dat de term P nihilair is de evenen, en de term Q nihilair is de onevenen magten van a zal bezien. Men zal dus mogen stellen

$$P = p_1 a^2 + p_2 a^4 + p_3 a^6 + \text{enz.}$$

$$Q = q_1 a + q_2 a^3 + q_3 a^5 + \text{enz.}$$

Bladzijde met wiskunde uit het Artikel van Lobatto uit 1860, 'Over de Waarschijnlijkheid van Gemiddelde Uitkomsten uit een Groot Aantal Waarnemingen'.

vele andere onderwerpen voor de statistiek van ons land mededeelingen in het jaarboekje mogten gevonden worden.' Hij dacht daarbij aan statistieken over onderwijsinstellingen, over in- en uitvoer, en aan mededeelingen omtrent de rechtspleging.

De juristenstatisticus J.K.W. Quarles van Ufford schreef in 1850 dat hij het een groot gemis vond dat Lobatto's jaarboekje zich beperkte tot bevolkingsgegevens en niets vermeldde over zaken als landbouw, handel, nijverheid, financiën en koloniën. Hij sprak de hoop uit dat de overheid een uitvoeriger statistisch jaarboekje zou gaan uitgeven, hetgeen vanaf 1851 ook is gebeurd. Hierop zou Lobatto hebben

gereageerd met de volgende zinsnede: 'De criticus heeft gelijk, wat de zaak betreft, maar ik mag geenerlei statistiek medelen, dan waarop men staat maken kan.' Volgens dezelfde bron heeft deze kritiek ook bijgedragen aan het beëindigen van het *Jaarboekje*, aangezien 'nieuwe inzichten over de inrigting van dit Jaarboekje [...] hem de taak uit handen namen'. Lobatto zelf schreef een aantal jaar later aan Quetelet dat zijn jaarboekje sinds 1850 niet meer verschenen was en wel hoofdzakelijk omdat er een statistisch jaarboekje door het statistisch bureau van het ministerie van Binnenlandse Zaken werd uitgegeven. Uit archiefonderzoek blijkt dat dit bureau, dat in 1848 was opgericht, al in 1849 plannen had opgesteld om een statistisch jaarboekje uit te geven. De minister onthief toen Lobatto van zijn taak om het jaarboekje samen te stellen.

Uit bovenstaande episode volgt dat de juristenstatistici bepaalden welke onderwerpen tot de statistiek behoorden. Lobatto bestreed dat niet; hij kwam alleen op voor de zorgvuldigheid waarmee het kwantificeren plaats behoorde te vinden en vond dat op dat moment voldoende reden om zich bijna geheel tot populatiestatistieken te beperken. Indirect beschuldigde hij daarmee de juristenstatistici van onvoldoende zorgvuldigheid in het samenstellen van hun statistieken.

Quetelet

Quetelet heeft geprobeerd beide benaderingen met elkaar te verenigen. Wie was Quetelet en wat deed hij aan statistiek? Quetelet was een Belgisch wiskundige die in 1823 tijdens een bezoek aan Parijs in contact was gekomen met beroemde wiskundigen als P.S. de Laplace, die zich met kansrekening en mogelijke maatschappelijke statistische toepassingen bezighielden. Nadien is ook hij zich met statistiek gaan bezighouden en heeft veel invloed uitgeoefend, niet alleen in Nederland, maar ook bijvoorbeeld in Engeland waar hij indirect betrokken was bij de oprichting van de nog altijd bestaande *Royal Statistical Society*.

In 1835 verscheen zijn belangrijkste werk *Sur l'homme et le développement de ses facultés. – Essai de physique sociale*. Hierin presenteerde hij vele statistische gegevens over zowel fysieke als morele kwaliteiten van de mens en liet zien dat veel van deze gegevens rond hun gemiddelde waren verdeeld volgens de foutenwet uit de astronomie, ofwel de normale verdeling. Hij definieerde de *gemiddelde mens* ofwel *l'homme moyen* als de mens waarbij alle te onderscheiden grootheden een gemiddelde waarde hadden. Deze beschouwde hij als het ideaaltype mens. Omdat hij deze mens als het *zwaartepunt van de maatschappij* beschouwde, zou die het onderwerp van studie moeten zijn in een nieuwe wetenschap die hij *sociale fysica* noemde. Kennis van de gemiddelde mens zou inzicht in de mensheid en de maatschappij opleveren.

Quetelet verzamelde voor zijn studies statistische gegevens uit een groot aantal landen en merkte dat deze gegevens meestal niet of nauwelijks vergelijkbaar waren. Ze waren volgens verschillende criteria en categorieën verzameld en ingedeeld. Dit was voor hem de reden om het voortouw te nemen bij het organiseren van internationale statistische congressen, waarvan de eerste in 1853 onder zijn voorzitterschap te Brussel werd gehouden. Er zijn negen van deze congressen geweest waarvan één in Londen in 1860 en ook één in Den Haag, in 1869. Deze congressen werden met name bezocht door ambtenaren en die waren meestal uit de wereld van de juristenstatistici afkomstig.

Lobatto en Quetelet

Welke kenmerken van Quetelets wijze van statistiek bedrijven zorgden ervoor dat beide Nederlandse richtingen zich in zijn benadering herkenden en waaruit blijkt dan die onmachtigheid in het navolgen?

Wat de kwantitatief-mathematische richting betreft, is het volgende van belang. Lobatto en Quetelet waren beiden wiskundigen en verder leeftijdgenoten. Het eerste contact dateert van vòòr 1825. In die tijd was Quetelet nog veel meer

met wiskunde bezig dan later toen hij veel meer door zijn statistische en andere activiteiten in beslag werd genomen. Lobatto legde het contact en dit was het begin van een levenslange vriendschap, waarbij vele brieven werden gewisseld; van de kant van Lobatto in elk geval zo'n 120. In de brieven komt veel persoonlijke informatie voor; de wiskunde nam een belangrijke plaats in en ook de statistiek.

De interesses van Quetelet en Lobatto voor waarschijnlijkheidsrekening en statistiek ontwikkelden zich echter in verschillende richtingen. Lobatto merkte dit verschil in belangstelling op en hij bewonderde Quetelet vanwege diens brede interesse. Hij heeft Quetelet eens vol respect een universeel mens genoemd, die evenzeer bekwaam was in de letteren als in de exacte wetenschappen. In Lobatto's formulering van zijn gedachten naar aanleiding van het verschijnen van Quetelets hoofdwerk *Sur l'homme* komt heel karakteristiek zowel dit verschil tussen beide persoonlijkheden tot uitdrukking, als de bewondering die Lobatto juist daarom voor Quetelet koesterde. Lobatto schreef '[...] ben ik erg benieuwd naar de inhoud, speciaal de mathematische theorie over de bevolking, waarmee ik ook zelf nogal druk mee bezig ben. Ik bewonder altijd het gemak waarmee je met evenveel gemak zulke verschillende onderwerpen als toepassingen van de exacte wetenschappen bespreekt' en toen Lobatto van het boek kennis had kunnen nemen '[...] bewonder ik de moed waarmee je zulke veelomvattende en lastige onderzoeken hebt ondernomen.'

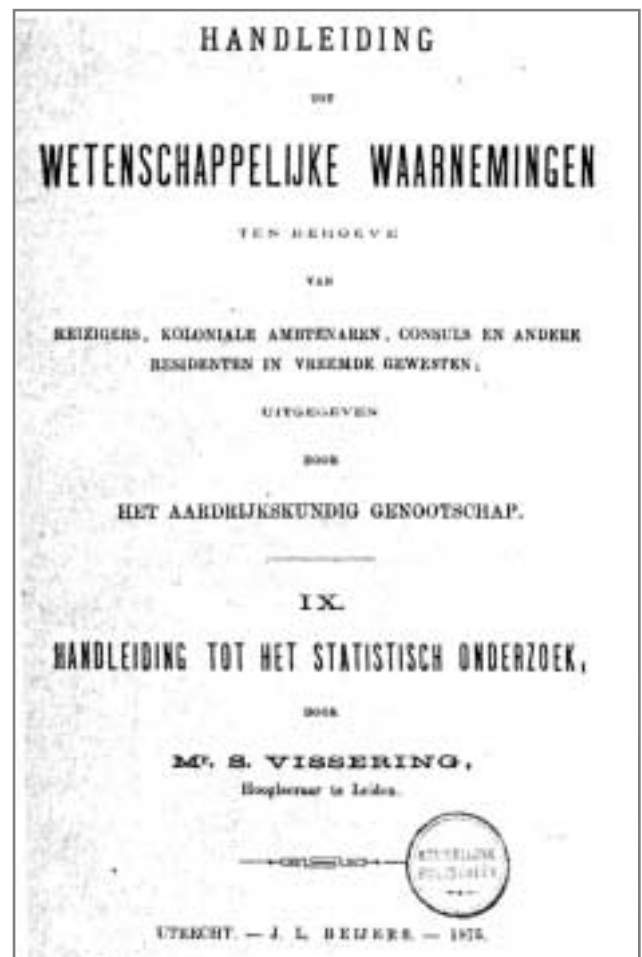
Hoewel er een levendige briefwisseling tussen beide geleerden heeft bestaan en Lobatto onder de indruk was van Quetelets capaciteiten, is de invloed van Quetelet op Lobatto beperkt is gebleven. Lobatto voelde aan dat hij Quetelet niet kon navolgen in diens enthousiasme voor de statistiek als basis voor een wetenschap van mens en maatschappij. Hij was en bleef een wiskundige, ook in zijn belangstelling voor de statistiek.

De VVS en Quetelet

Welke kenmerken van Quetelets wijze van statistiek bedrijven zorgden ervoor dat de andere Nederlandse richting in de statistiek zich in zijn benadering herkende en waaruit blijkt daar de onmachtigheid in het navolgen?

'Ik beschouw U en eer U als mijn meester' schreef Vissering aan Quetelet in 1869, het jaar dat de internationale statistische conferentie in Nederland werd gehouden. Tijdens dit congres, waar Quetelet eregast zal zijn geweest, werd hij geridderd in de Orde van de Nederlandse Leeuw, een onderscheiding waarin de juristenstatistici ongetwijfeld de hand hebben gehad. Dit congres was door deze statistici georganiseerd. Vissering was vice-voorzitter en sprak de deelnemers toe met een rede getiteld *Limites de la statistique*. Aan deze internationale statistische beweging namen zowel Quetelet als de juristenstatistici deel; het internationaal afstemmen van statistiek was beider doel. Ze verschilden echter wanneer het om toepassen van kansrekening in de statistiek ging. Vissering was daar dan wel niet op tegen, maar de manier van denken was hem zo vreemd dat hij zich die benadering ook later niet eigen heeft weten te maken. Hoewel hij het getal in de statistiek in de loop van de tijd steeds belangrijker is gaan vinden, moest hij niet veel van een gemiddelde hebben, omdat het, zo beargumenteerde hij dat, geen onderscheid maakte tussen de hoge en lage standen in de samenleving.

Verder ondersteunde hij zijn pleidooi om veel waarnemingen te doen met de wet die hij volgens zichzelf bij Quetelet gevonden had. Hij formuleerde die wet als volgt: 'dat de waarschijnlijkheid van het voorkomen van een bepaald feit toeneemt in de verhouding van den vierkantswortel van het getal waargenomen feiten van dien aard.' Vissering citeerde in een voetnoot, hoe Quetelet deze regel geformuleerd had en dat was nogal anders: 'de nauwkeurigheid van de resultaten is recht evenredig met de vierkantswortel uit het aantal waarnemingen.' De nauwkeurigheid was voor Quetelet



Titelblad 'Handleiding tot het Statistisch Onderzoek' (Vissering)

een getal dat bij de normale verdeling hoort en dat recht evenredig is met de waarschijnlijke fout. Daaronder verstond Quetelet die afwijking van de verwachtingswaarde, waarvoor geldt dat een willekeurige meting met evenveel kans een grotere of een kleinere afwijking van de verwachtingswaarde heeft. De nauwkeurigheid is dus een graadmeter voor de breedte van de bijbehorende normale curve en staat in directe relatie tot de standaarddeviatie. Vissering liet bij het aanhalen van Quetelet de term nauwkeurigheid achterwege. De wijze waarop Vissering die wet formuleerde, is niet eenduidig te interpreteren. Hij begreep wel dat het belangrijk is om veel waarnemingen te doen om zo een nauwkeuriger resultaat te verkrijgen. Hij wist ook dat de wortel uit het aantal waarnemingen hierbij een speciale rol speelt. Echter, de

exacte wiskundige inhoud van de aangehaalde opmerking van Quetelet ging aan hem voorbij. Vissering bewonderde Quetelet, maar hij begreep hem niet.

Conclusie

De manier waarop Lobatto met statistiek bezig was en de wijze van Vissering en de andere juristenstatistici was te verschillend om in een nieuwe geïntegreerde vorm van statistiek te resulteren. Op zijn minst zouden ze zich daartoe vertrouwd moeten maken met elkaars denken. Het is duidelijk geworden dat dat niet van hen kon worden verwacht. Dat ze in twee verschillende werelden leefden en dat de kloof hiertussen niet te overbruggen viel, is ook gebleken uit hun reactie op elkaar en op Quetelet die ze beide bewonderden. Hun twee werelden doen denken aan de *Two Worlds* van C.P. Snow.

Zoals besproken, moest in 1849 het jaarboekje van Lobatto het afleggen tegen een jaarboekje uit-

gegeven door een statistisch overheidsbureau waar Lobatto geen deel van uitmaakte. Lobatto's rol was uitgespeeld en de juristenstatistici, georganiseerd in de *Vereeniging voor de Statistiek* maakten de dienst uit. Dit zou tot gevolg hebben dat een meer wiskundige benadering voorlopig geen enkele kans zou krijgen in de overheidsstatistiek. Dat gebeurde pas veel later, in de twintigste eeuw. Het ontstaan van een *Vereniging voor de Statistiek*, waarin de kansrekening centraal zou staan, moest zelfs wachten tot 1945.

NOTEN

1. *STATOR* 3 (2002) nr. 2, 13-17 en *STATOR* 4(2003)nr. 1, 15-20.
2. Voor dit artikel is relevant de bundel: Paul M.M. Klep and Ida H. Stamhuis, *The Statistical Mind in a Pre-Statistical Era. The Netherlands 1750-1850*. Amsterdam: Aksant, 2002, met name hoofdstuk 3: Ida H. Stamhuis, 'An Unbridgeable Gap between Two Approaches to Statistics', pp. 71-98.

IDA H. STAMHUIS doceert wetenschapsgeschiedenis aan de Vrije Universiteit Amsterdam. E-mail: <stamhuis@nat.vu.nl>.

Statistiek voor de psychologie

Jules L. Ellis

NIEUW

Een zeer toegankelijke inleiding voor het hoger onderwijs in vier delen:



Deel 1: Verdeling van een variabele, samenhang tussen twee variabelen
ISBN 90 5352 861 x, 240 pagina's, € 25,- (verschijnt maart 2004)

Deel 2: Toetsen voor twee gemiddelden en toetsingstheorie
ISBN 90 5352 923 3, 296 pagina's, € 29,75 (verschijnt januari 2004)

Deel 3: Variantieanalyse
ISBN 90 5352 863 6, 288 pagina's, € 29,50

Deel 4: GLM en non-parametrische toetsen
ISBN 90 5352 864 4, 320 pagina's, € 34,50

Kijk voor uitgebreide informatie op:
www.statistiekvoordepsychologie.nl

Ook verkrijgbaar via de boekhandel. Docenten kunnen beoordelingsexemplaren aanvragen bij Uitgeverij Boom, 020 – 622 61 07, info@uitgeverijboom.nl.



De kracht van de wachtrijtheorie



ONNO BOXMA

In het vorige nummer van *STATOR* schreven Emile Aarts en Benjamin Jansen over '50 jaar georganiseerde OR: oude dame of jonge heer?'. Als 50 jaar oud wordt gevonden - waarmee ik persoonlijk om begrijpelijke redenen grote moeite heb - dan is de wachtrijtheorie helemaal op sterven na dood. Zij is namelijk geboren in 1917, toen de Deense ingenieur Agner Krarup Erlang een dimensioneringsprobleem van telefooncentrales bestudeerde: hoeveel lijnen dienen in een centrale te worden aangelegd opdat de kans hooguit ϵ is dat een oproep 'verloren gaat'. Hij veronderstelde dat telefoonoproepen optreden volgens een Poissonproces met intensiteit λ , dus met onafhankelijke, exponentieel verdeelde tussenpozen met gemiddelde $1/\lambda$. Tevens veronderstelde hij dat de gespreksduren onafhankelijke, exponentieel verdeelde stochastische variabelen zijn, met gemiddelde $1/\mu$. De gemiddelde hoeveelheid per tijdseenheid aangeboden 'verkeer' is dan $\rho := \lambda/\mu$.

Erlang [1] (zie ook [2]) leidde af dat, in evenwicht, de kans op k bezette lijnen in een centrale met N lijnen gelijk is aan

$$p_k = \frac{\frac{\rho^k}{k!}}{\sum_{j=0}^N \frac{\rho^j}{j!}}, \quad k = 0, 1, \dots, N.$$

Voor de benodigde kansrekening raadpleegde hij de *Encyclopaedia Britannica*. Erlang leunde zwaar op wat hij 'de theorie van statistisch evenwicht' noemde.

In moderne termen zouden we zeggen: het aantal bezette lijnen is een Markovproces, zelfs een geboorte- en sterfteproces, en de balansvergelijkingen voor dat geboorte- en sterfteproces leiden onmiddellijk tot bovenstaande formule. Erlang noemt Markov overigens niet, hoewel de eerste publicatie over Markovketens al uit 1909 stamt.

Wat is nu de kans dat een oproep alle lijnen bezet aantreft, dus 'verloren gaat'? Erlangs antwoord was p_N , dus de kans dat alle n lijnen op een *willekeurige* moment bezet zijn. Hij redeneerde waarschijnlijk dat deze twee kansen gelijk zijn, wegens de geheugenloosheid van de exponentiële aankomstintervallen - 65 jaar voordat Wolff formeel PASTA (Poisson Arrivals See Time Averages) bewees. Veertig jaar na Erlang bewees Sevast'janov dat Erlangs formule zelfs blijft gelden als de gespreksduren een *willekeurige* kansverdeling hebben met hetzelfde gemiddelde $1/\mu$. Deze *ongevoeligheid* voor de gespreksduurverdeling is een diep en buitengewoon nuttig resultaat.

Springlevend

Nu, 86 jaar na Erlang, is de wachtrijtheorie nog springlevend. Haar levenselixer is misschien wel een drietal heel natuurlijke basismodellen, die in allerlei vermommingen en generalisaties telkens weer opduiken, en waarvan we het gedrag heel goed begrijpen door grondige bestudering met een breed scala aan technieken. Eén van die basis-



modellen van de wachtrijtheorie is het bovengeschetste Erlang verliessysteem - een wachtrijstelsysteem waarin ironisch genoeg nooit gewacht wordt (de andere basismodellen zijn in mijn optiek de M/G/1 rij en het Jackson netwerk).

Eén van haar vermommingen is het machine-reparatiemodel: N machines functioneren, onafhankelijk van elkaar, gedurende een exponentieel verdeelde tijd met verwachting $1/\mu$, en gaan dan stuk. Er is één monteur beschikbaar. Deze repariert machines in volgorde van stuk gaan. De reparatietijden zijn exponentieel verdeeld met verwachting $1/\lambda$. De balansvergelijkingen voor de kansverdeling van het aantal werkende machines zijn *identiek* aan die voor het aantal bezette telefoonlijnen, en inderdaad leert enig nadenken dat het aantal werkende machines zich stochastisch hetzelfde gedraagt als het aantal bezette lijnen. In de computerwereld spreekt men trouwens over het computer-terminal model (waarbij de werkende machines worden vervangen door terminalgebruikers die na een exponentiële denktijd een job genereren, en de monteur door een processor). Het afgelopen jaar doken varianten van deze modellen op bij Eindhovense studentenprojecten bij zulke verschillende bedrijven als ECT, Philips en Siemens-VDO Automotive. Kortom, Erlangs pionierswerk is buitengewoon vruchtbaar gebleken.

Acquisitie is een vak op zich ...

Dat niet iedereen de kracht en schoonheid van Erlangs formule kan waarderen, toont de volgende anekdote, die u tevens mijn afkeer van parkeergarages verklaart - een afkeer die lang niet alleen voortkomt uit het feit dat m'n armen doorgaans net te kort zijn om het kaartje uit de automaat te pakken. Toen ik in 1985 in dienst trad van het CWI, ontmoette ik daar Ko Anthonisse. Ko wist goed verbanden te leggen tussen problemen uit het dagelijks leven en bekende technieken op het gebied van mathematische programmering. Eén van zijn vruchtbare contacten (althans, tot ik me ermee bemoeide) was met een bureau van de gemeente Amsterdam

op het gebied van verkeer en vervoer, en hij stelde voor dat ik hem eens naar dat bureau zou vergezellen. We werden te woord gestaan door een wat morsige man. Nadat hij een uur met Ko over deterministische planningsproblemen had gesproken, vroeg ik of hij ook wel eens behoefte had aan stochastiek.

Hij keek of ik hem een oneerbaar voorstel had gedaan, maar antwoordde tenslotte dat men dacht over de aanleg van een parkeergarage, waarbij de vraag rees hoeveel parkeerplaatsen moesten worden gemaakt opdat zelfs op zaterdagmiddag de capaciteit vrijwel zeker toereikend zou zijn. 'Dat is precies een Erlang verliessysteem', kraaide ik enthousiast, en ik begon de Erlang-formule op een blaadje te schrijven. 'Rustig maar', zei hij - pas later besepte ik dat hij de faculteiten als uitroeptekens had opgevat - 'zo moeilijk hoeft het niet, die garage kunnen we wel simuleren'. 'Ja maar', wierp ik nog zwakjes tegen, 'als u me vertelt wat λ en μ zijn op zaterdagmiddag, kan ik u precies voorrekenen hoe groot het aantal parkeerplaatsen moet zijn opdat een auto met hooguit kans ε verloren gaat'. Dat laatste, een standaarduitdrukking uit de telefonie, bleek een ongelukkig woordgebruik. 'Auto's gaan niet verloren in Amsterdam', sprak hij ijzig, en ik kon gaan.

Buiten gekomen bleek dat fietsen wel verloren kunnen gaan in Amsterdam...

LITERATUUR

1. A.K. Erlang. Løsning af nogle Problemer fra Sandsynlighedsregningen af Betydning for de automatiske Telefoncentraler. *Elektroteknikeren* 13(1917), p. 5. Ook verschenen als 'Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges'. *Post Office Electrical Engineer's Journal* 10 (1917-1918) 189-197.
2. E. Brockmeyer, H.L. Halstrøm and Arne Jensen. The life and works of A.K. Erlang. *Transactions of the Danish Academy of Technical Sciences* 1948, No. 2. Reprinted in *Acta Polytechnica Scandinavica, Applied Mathematics and Computing Machinery Series* No. 6 (AP287, 1960).

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



Foto: P. Bosch

LEVENSV ERWACHTING

het CBS bepaalt hoe oud we worden

Volgens het CBS worden mannen gemiddeld 76,0 jaar oud, en vrouwen 80,7 jaar. In 2030 zullen we gemiddeld nog ouder worden: mannen 78,6 jaar, vrouwen 81,7. Je kunt je afvragen hoe het CBS dit weet. De getallen zullen wel berekend zijn, maar hoe doet het CBS dat dan? In dit artikel wordt nader op deze berekeningswijze ingegaan. Het CBS baseert zich voor de berekening van de levensverwachting op waarnemingen die in een model worden verwerkt. Bij cijfers die betrekking hebben op de toekomst, zijn voorts prognosehypothese en extrapolatietechnieken in het geding.

KEES PRINS¹

Op buitenlandse diplomaten en buitenlandse NAVO-militairen na, behoort iedereen die in Nederland woont te zijn ingeschreven in de basisadministratie van zijn woonplaats. Bevolkingsregisters in Nederland dateren uit midden 19de eeuw. Bij de laatste grote wijziging, met de invoer

ring van de Wet GBA in 1994, zijn de bevolkingsregisters geautomatiseerd en hebben ze de naam 'basisadministratie' gekregen.

De gemeenten sturen het CBS gegevens van alle demografische veranderingen die in de gemeentelijke basisadministraties zijn verwerkt.

Deze gegevens omvatten onder meer het geslacht, de geboortedatum en de overlijdensdatum. Met behulp van deze gegevens maakt het CBS statistieken over geboorte, sterfte², immigratie enz. (CBS, 1999). Daarnaast verstrekken de gemeenten het CBS jaarlijks een overzicht van een aantal gegevens van alle inwoners van de gemeente, zoals het geslacht en de geboortedatum.

Voordat de levensverwachting kan worden bepaald, berekent het CBS eerst voor elke leeftijd het sterftequotiënt. Dit is het quotiënt van het aantal mensen dat in een kalenderjaar op zekere leeftijd overlijdt en het aantal inwoners van die leeftijd aan het begin van dat kalenderjaar. Zowel de teller als de noemer van deze quotiënten worden afgeleid af uit de waarnemingen zoals beschreven in de vorige alinea. Het sterftequotiënt heet ook wel sterftekans, de kans dat iemand van een bepaalde leeftijd binnen twaalf maanden sterft.

Het model: de overlevingstafel

Om de levensverwachting te kunnen berekenen, stelt het CBS met behulp van de sterftequotiënten een overlevingstafel op (vroeger sterftetafel genoemd). Een overlevingstafel start met een fictieve groep 100000 levendgeboren jongens of meisjes; zie tabel 1, waarin een gedeelte van de overlevingstafel vrouwen in 2002 is weergegeven. Er is berekend dat 0,408% van de meisjes die in 2002 zijn geboren, in dat kalenderjaar zijn overleden. Dit betekent voor de overlevingstafel dat van de 100000 nuljarigen er 408 in het jaar van geboorte overlijden, en dat dus 100 000 - 408 = 99592 kinderen de leeftijd 0,5 bereiken³. Van de meisjes die in 2001 zijn geboren, is in 2002 0,062% gestorven. Van de 99592 0,5-jarige kinderen in de overlevingstafel zullen er op die leeftijd dus 0,00062 x 99592 = 62 overlijden. Dat betekent dat er aan het eind van het eerste volledige jaar dus 99592 - 62 = 99530 kinderen van gemiddeld 1,5 jaar oud, over zijn. Alle aantallen zijn afgerond op gehelen, waardoor berekeningen afrondverschillen kunnen vertonen.

Op deze wijze is voor alle leeftijden het aantal overlevenden berekend. De levensverwachting e_0 wordt nu bepaald als het rekenkundige gemiddelde van de leeftijden waarop de mensen in de overlevingstafelbevolking dood gaan. In formule:

$$e_0 = (408 \times 0,25 + 62 \times 1 + 31 \times 2 + 14 \times 3 + \dots) / 100000 = 80,69.$$

Hierbij is verondersteld dat kinderen die in het kalenderjaar van geboorte overlijden, gemiddeld drie maanden oud worden, kinderen die overlijden in het kalenderjaar dat volgt op hun geboortjaar, gemiddeld 1 jaar, enz. De formule komt overeen met de verwachtingswaarde uit de wiskundige statistiek. Om die reden wordt de levensverwachting met het symbool e aangeduid⁴. De levensverwachting kan ook worden opgevat als het gemiddelde aantal jaren dat de overlevingstafelbevolking leeft. De levensverwachting die op deze manier is berekend, wordt de levensverwachting bij geboorte

Leeftijd	Geboortearjaar	Sterftequotiënt	Levenden	Overledenen	Levensverwachting (jaren)
0	2002	0,00408	100000	408	80,69
0,5	2001	0,00062	99592	62	80,52
1,5	2000	0,00031	99530	31	79,57
2,5	1999	0,00014	99499	14	78,59
3,5	1998	0,00020	99485	20	77,60
4,5	1997	0,00015	99465	14	76,62
5,5	1996	0,00012	99450	12	75,63
6,5	1995	0,00011	99439	10	74,64
7,5	1994	0,00008	99428	8	73,65
8,5	1993	0,00008	99420	8	72,65
9,5	1992	0,00005	99412	5	71,66
10,5	1991	0,00008	99407	8	70,66
70,5	1931	0,01696	83510	1416	14,92
71,5	1930	0,01817	82094	1492	14,17
72,5	1929	0,02168	80602	1748	13,42
73,5	1928	0,02371	78854	1870	12,71
74,5	1927	0,02528	76985	1946	12,00
75,5	1926	0,02929	75039	2198	11,30
76,5	1925	0,03347	72841	2438	10,63
77,5	1924	0,03726	70403	2623	9,98
78,5	1923	0,04281	67780	2902	9,35
79,5	1922	0,04669	64878	3029	8,74

Tabel 1. Overlevingstafel vrouwen, 2002. Bron: CBS.

genoemd, omdat de berekening bij leeftijd 0 start. Het is ook mogelijk om op een andere leeftijd te beginnen. Zo is het gemiddeld aantal jaren dat een 75,5-jarigen vrouw nog zal leven, gelijk aan

$$e_{75,5} = ((75039 + 72841)/2 + (72841 + 70403)/2 + \dots) / 75039 = 11,30.$$

Een vrouw die nu 75,5 jaar oud is, zal naar verwachting dus bijna 87 jaar oud worden. In de overlevingstafel is zo voor iedere leeftijd het gemiddelde aantal nog te leven jaren weergegeven.

Interpretatie: minder eenvoudig dan het lijkt

Hoe bruikbaar is de levensverwachting die uit de overlevingstafel is berekend? Op grond van het voorafgaande kan men zeggen dat dit getal het aantal jaren aangeeft dat een groep mensen van een zekere leeftijd gemiddeld nog heeft te leven *mits de sterftequotiënten van hun leeftijd en hogere leeftijden niet meer zullen veranderen*. Hier stuiten we op een probleem. In de praktijk blijkt namelijk dat de leeftijdsspecifieke sterftequotiënten in de tijd veranderen. Dit betekent dat de levensverwachting in eerste plaats een indicator is voor de sterfte in een zekere periode of een zeker kalenderjaar. In tijden van oorlog,terreur of hongersnood kan de levensverwachting zeer sterk dalen. Een voorbeeld is de Oekraïne, waar in 1931 de levensverwachting bij geboorte voor mannen 43,5 jaar was, in 1933 (toen als gevolg van de landbouwcollectivisatie zich een grote hongersnood voordeed) niet meer dan 7,3 jaar, en in 1935 weer 46,4 jaar. Bedacht moet worden dat deze getallen zijn afgeleid uit de sterftequotiënten per jaar. In 1933 waren de sterftequotiënten voor alle leeftijden hoog, waardoor de levensverwachting extreem laag was. De kinderen die in 1931 zijn geboren, zijn, in verband met hongersnood en oorlogshandelingen gemiddeld mogelijk minder oud geworden dan 43,5 jaar. Het sterftequotiënt voor 1,5-jarigen was in 1933 immers veel hoger dan in 1931. Op een analoge manier zal de gemiddelde leeftijd bij sterfte van

de kinderen die in 1933 zijn geboren, hoger zijn dan de voorspelde 7,3 jaar. Aan de veronderstelling dat de sterftequotiënten in de toekomst niet veranderen, is hier dus duidelijk niet voldaan.

Vergeleken met de cijfers van de Oekraïne veranderen die van Nederland van jaar op jaar nauwelijks. Toch kan worden gesteld dat de levensverwachtingen die het CBS jaarlijks presenteert, een onderschatting geven van de gemiddelde levensduur van een generatie. Op basis van de overlevingstafels die zijn gebaseerd op de sterftequotiënten over eind 19e eeuw, zou een jongetje dat toen is geboren, gemiddeld 47,5 jaar oud worden. In werkelijkheid zijn de mannen die in die tijd zijn geboren, vijf jaar ouder geworden dan volgens sterftetafels ten tijde van hun geboorte: de sterftequotiënten zijn voor alle leeftijden in de tijd kleiner zijn geworden. Over de hele twintigste eeuw bezien heeft zich een trendmatige daling van de sterftequotiënten voorgedaan.

De invloed van immigratie en emigratie

Het sterftetafelmodel veronderstelt dat de sterfte in een vaste groep personen wordt nagegaan, waarbij er tussentijds geen personen bijkomen of de groep verlaten. In de praktijk zijn immigratie en emigratie echter geen verwaarloosbare verschijnselen. De oplossing voor dit probleem is dat immigranten vanaf de dag van hun vestiging in Nederland gaan meedoen met het bepalen van de sterftequotiënten voor hun leeftijd, terwijl emigranten vanaf de dag dat zij Nederland verlaten niet langer meedoen. De wijze waarop de levensverwachting wordt berekend, verandert natuurlijk niet.

Men vermoedt dat de wijze waarop immigranten en emigranten bij de overlevingstafel worden betrokken, van geringe invloed is op de hoogte van de levensverwachting. Het is lastig om deze invloed precies te becijferen. Dat komt doordat het CBS geen inzicht heeft in de sterfte van emigranten, zodat het ondoenlijk is om die bij het opstellen

van de Nederlandse overlevingstafels te betrekken. Van immigranten zou de invloed theoretisch wel kunnen worden becijferd, maar voorzichtigheid is hier op zijn plaats. Zo blijkt dat de sterftequotiënten van oudere Turken en Marokkanen aanmerkelijk lager zijn dan die van autochtone Nederlanders; men vermoedt dat veel Turken en Marokkanen kort voor hun dood besluiten naar hun vaderland terug te keren. De lage sterftcijfers voor oude Turken en Marokkanen leiden tot een lichte overschatting van de levensverwachting in de overlevingstafels van Nederland.

Levensverwachting in de toekomst

Behalve bevolkingsstatistieken maakt het CBS bevolkingsprognoses. Deze prognoses geven niet alleen informatie over de toekomstige aantallen inwoners van Nederland tot het midden van deze eeuw, maar ook over de jaarlijkse aantallen geborenen, overledenen, immigranten en emigranten. Sterker, het CBS leidt eerst de jaarlijkse aantallen geborenen enz. af voordat de stand van de bevolking kan worden doorgerekend. Uiteraard zijn de toekomstige aantallen geborenen, overledenen, immigranten en emigranten gebaseerd op veronderstellingen, die op hun beurt zijn gebaseerd op waarnemingen uit het verleden. Om de toekomstige levensverwachting bij de geboorte te kunnen bepalen, maakt het CBS sinds kort gebruik van een combinatie van een trendmodel en een verklarend model (Van Hoorn en De Beer, 2001). Naast een trendmatige ontwikkeling zijn de verklarende componenten: roken, verkeersongevallen en de introductie van antibiotica in het model opgenomen. Natuurlijk zijn veel meer verschijnselen van belang voor het sterfteverloop, maar bij gebrek aan betrouwbare kwantitatieve informatie is gekozen voor een beperkt model.

Met behulp van dit model kan voor ieder moment in de toekomst de levensverwachting bij de geboorte worden *ingesteld*. Als dit is gebeurd, worden

de leeftijdsspecifieke sterftequotiënten geëxtrapolerd op zodanige wijze dat deze op logische wijze aansluiten bij de waargenomen quotiënten. Voor geboorte⁵ en emigratie worden eveneens toekomstige quotiënten afgeleid. Alleen voor immigratie worden direct absolute aantallen voorspeld.

Het opstellen van een bevolkingsprognose gaat nu als volgt. Het CBS gaat uit van het meest recente moment waarover de samenstelling van de bevolking bekend is, zeg 1 januari 2003. Voor het kalenderjaar 2003 worden de aantallen levendgeborenen, overledenen en emigranten, verdeeld naar leeftijd en geslacht, bepaald door de voorspelde quotiënten toe te passen op de bevolking van 1 januari 2003.

Het aantal inwoners op 1 januari 2004 wordt voorspeld door het aantal inwoners op 1 januari 2003 te vermeerderen met de verwachte aantallen levendgeborenen en immigranten, en te verminderen met de verwachte aantallen overledenen en emigranten. Dit gebeurt voor mannen en voor vrouwen voor alle leeftijden afzonderlijk.

Als het verwachte aantal inwoners per 1 januari 2004 naar leeftijd en geslacht bekend is, kunnen op dezelfde wijze als voor het kalenderjaar 2003 de aantallen levendgeborenen, sterfgevallen, immigranten en emigranten over het kalenderjaar 2004 worden berekend, en daarmee de bevolking per 1 januari 2005. Dit proces kan voor alle toekomstige kalenderjaren worden herhaald. Het laatste jaar waarvoor het CBS de toekomstige bevolking bepaalt, is 2050.

Levensverwachting voor generaties

We hebben het gehad over de gemiddelde leeftijd waarop de generatie die eind 19de eeuw is geboren, is overleden. Dit geeft aan dat het CBS ook overlevingstafels voor generaties samenstelt (zie Tas 2002). In een dergelijke tafel worden niet de sterftequotiënten gebruikt die over een bevolking gedurende een bepaalde periode zijn waargenomen, maar de sterftequotiënten die per jaar zijn waargenomen. Voor de overlevingstafel van de generatie 1900 is

voor het sterftequotiënt voor 0-jarigen dat van 0-jarigen in 1900 genomen, voor het sterftequotiënt van 0,5-jarigen dat van 0,5-jarigen in 1901, voor het sterftequotiënt van 73,5-jarigen dat van 73,5-jarigen in 1974, enz. De opstelling van de overlevingstafel gaat verder op precies dezelfde wijze als eerder in dit artikel beschreven. De sterftequotiënten in zo'n generatie-overlevingstafel hebben dus betrekking op de werkelijke generatie van 1900.

Een generatie-overlevingstafel (zoals in tabel 2, waarin een gedeelte van de overlevingstafel van de vrouwen generatie 1917 is afgebeeld) laat de sterfte zien van mensen die in een zelfde kalenderjaar zijn geboren en in verschillende kalenderjaren zijn overleden. Een periode-overlevingstafel (zoals in tabel 1) laat de sterfte zien van mensen die in een zelfde kalenderjaar (2002 in tabel 1) zijn overleden en in verschillende jaren zijn geboren.

De interpretatie van een generatie-overlevingstafel is eigenlijk eenvoudiger dan die van een periode-

tafel waarvan de sterftequotiënten betrekking hebben op een kalenderjaar, dat wil zeggen: de mensen die in dat kalenderjaar tot de bevolking van Nederland behoren resp. in dat jaar sterven. De levensverwachting in een generatie-overlevingstafel is gelijk aan de gemiddelde leeftijd waarop die generatie overlijdt, onder voorbehoud van de kanttekening bij immigranten en emigranten. Maar het nadeel van een generatie-overlevingstafel zal ook direct duidelijk zijn: er moet ruim honderd jaar worden gewacht voordat zo'n tafel 'af' is, en zeker negentig jaar, zo niet langer, voordat een enigszins redelijk beeld van het sterftepatroon van een generatie is verkregen. Voor actuele sterfteontwikkelingen is een generatie-overlevingstafel niet geschikt.

Bij het opstellen van bevolkingsprognoses maakt het CBS om deze reden geen gebruik van levensverwachtingen die zijn gebaseerd op generatie-overlevingstafels. Ook worden er geen uitspraken gedaan over hoe oud een toekomstige generatie gemiddeld gaat worden. Om de gemiddelde leeftijd bij overlijden te bepalen van kinderen die bijvoorbeeld in 2003 zijn of worden geboren, zouden immers de volgende leeftijdsspecifieke sterftequotiënten 'bekend' moeten zijn: van 0-jarigen over het kalenderjaar 2003, van 0,5-jarigen over 2004, van 20,5-jarigen over 2024, van 81,5-jarigen over 2085, enz. Zover in de toekomst kijken zou slechts tot speculatieve uitspraken leiden.

Zijn er geen andere cijfers?

Na lezing van het voorafgaande zou de lezer zich kunnen afvragen waarom het CBS zo'n moeilijke berekeningswijze heeft gekozen. Zijn er geen andere cijfers te bedenken die minder lastig te interpreteren zijn? Is, bijvoorbeeld de gemiddelde leeftijd van de mensen die in een kalenderjaar sterven, niet een eenvoudiger te interpreteren maat? Eenvoudig is deze maat wel, maar hij is niet zo eenvoudig te interpreteren. Dat blijkt bijvoorbeeld als we dit cijfer voor de gemeenten van Nederland zouden berekenen. In een jonge gemeente als

Leeftijd	Jaar van-overlijden	Sterfte-quotiënt	Levend	Overledenen
0	1917	0,06040	100000	6040
0,5	1918	0,05642	93960	5301
1,5	1919	0,02276	88659	2018
2,5	1920	0,00665	86641	576
3,5	1921	0,00368	86065	317
4,5	1922	0,00232	85748	199
5,5	1923	0,00176	85549	151
6,5	1924	0,00147	85399	126
7,5	1925	0,00132	85273	113
8,5	1926	0,00128	85160	109
9,5	1927	0,00120	85051	102
70,5	1988	0,01830	67406	1234
71,5	1989	0,02075	66173	1373
72,5	1990	0,02219	64800	1438
73,5	1991	0,02467	63362	1563
74,5	1992	0,02701	61799	1669
75,5	1993	0,03170	60129	1906
76,5	1994	0,03414	58223	1988
77,5	1995	0,03881	56235	2182
78,5	1996	0,04166	54053	2252
79,5	1997	0,04864	51801	2520

Tabel 2. Generatie-overlevingstafel vrouwen, generatie 1917.

Almere, waar naar verhouding veel jonge gezinnen wonen, zijn de overledenen gemiddeld jonger dan in een gemeente met een verhoudingsgewijs oude bevolking, zoals Wassenaar. Dat betekent niet dat inwoners van Wassenaar gemiddeld ouder zouden worden dan die in Almere. Voor vergelijking met andere bevolkingen (in ruimte of tijd) voldoet deze maat dus niet.

Uit de overlevingstafel afgeleide kwartielen of andere percentielen geven soms additioneel inzicht, bijvoorbeeld in de generatie-overlevingstafels. Zo kan uit de overlevingstafel van vrouwen die in 1917 zijn geboren worden afgeleid dat van de oorspronkelijke 100.000 personen (fictief aantal) er op 75+-jarige leeftijd nog 60.129 in leven waren. Een in 1917 geboren vrouw die dus vóór 1 januari 1993 is overleden, behoort dus tot de eerste 40% overledenen van haar generatie. Anders dan voor de levensverwachting hoeft voor de berekening van percentielen de generatie-overlevingstafel niet helemaal 'af' te zijn. Merk overigens op dat in de overlevingstafel van 2002 het percentage overlevende 75+-jarige vrouwen veel groter is dan 40%, namelijk 75%. In dit cijfer wordt de sterke afname van de sterftequotiënten in de afgelopen eeuw weerspiegeld.

Zijn de overledenen gemiddeld echt zo oud geworden?

In het voorafgaande is aangegeven dat de levensverwachting de gemiddelde leeftijd bij overlijden is van de overlevingstafelbevolking. Mannen in de overlevingstafel zijn dus gemiddeld 76,0 jaar oud als ze dood gaan, vrouwen gemiddeld 80,7 jaar. Maar hoe zit het met de gemiddelde leeftijd van mensen die in 2002 zijn gestorven? Is die ook zo hoog? Het blijkt dat deze een paar jaar lager ligt dan de levensverwachting. Mannen die in 2002 overleden, waren gemiddeld ruim 72 jaar oud, bijna vier jaar jonger dan de levensverwachting. Vrouwen overleden in 2002 gemiddeld bijna 2,5 jaar eerder dan de levensverwachting aangeeft. De vraag dringt zich op waarom de overlevingstafelbevolking zo oud

is in vergelijking met de werkelijke (huidige) bevolking. Om dat te kunnen begrijpen moeten we eerst nog iets over de overlevingstafelbevolking zeggen. Het is de bevolking die op den duur zou ontstaan als er ieder jaar precies 100.000 jongens en 100.000 meisjes zouden worden geboren, de sterftequotiënten van jaar op jaar hetzelfde zouden zijn en er jaarlijks voor iedere leeftijd evenveel immigranten als emigranten zouden zijn. In dat geval zou de overlevingstafel ieder jaar precies dezelfde aantallen bevatten, en daardoor zou het aantal 0,5-jarigen uit de huidige bevolking gelijk zijn aan het aantal 0,5-jarigen uit de overlevingstafel van vorig jaar, het aantal 1,5-jarigen nu zou gelijk zijn aan het aantal 1,5-jarigen uit de tafel van twee jaar geleden, enz. Het zijn eisen waaraan de bevolking van Nederland in de twintigste eeuw beslist niet heeft voldaan. Het belangrijkste effect is dat door fluctuaties in aantallen geboorten, immigranten en emigranten de bevolking in 2002 jonger was dan volgens de overlevingstafel was verwacht. Dat betekent (analoog aan de situatie in Almere) dat er meer jonge mensen sterven dan verwacht en dus dat de leeftijd bij overlijden lager uitvalt.

Conclusie

Voor het beschrijven van de huidige en toekomstige sterfte maakt het CBS gebruik van het overlevingstafelmodel, waarvan de uitkomsten op zichzelf soms lastig te interpreteren zijn, maar die vooral bij het opstellen van bevolkingsprognoses en het vergelijken in de tijd en met andere landen goede diensten bewijst.

KEES PRINS is als wiskundige werkzaam bij de sector Ontwikkeling en Ondersteuning van de divisie Sociale en Ruimtelijke Statistiek van het Centraal Bureau voor de Statistiek. E-mail: <kpns@cbs.nl>

LITERATUUR

Centraal Bureau voor de Statistiek (1999), *Index special Bevolking*, augustus 1999.

Hoorn, W. van, en J. de Beer (2001), *Bevolkingsprognose 2000-2050: prognosemodel voor de sterfte. CBS Maandstatistiek van de bevolking*, juli 2001, blz. 10-15.

IN MEMORIAM

Prof.dr.ir. Wim Bakker

Op vrijdag 3 oktober j.l. is professor Wim Bakker overleden. Hij is ruim 25 jaar aan de universiteit Twente verbonden geweest, als hoogleraar Productieorganisatie binnen de toenmalige faculteit Werktuigbouwkunde (nu deel van CTW). In 1966 was hij de opvolger van professor van Hasselt. Gedurende de laatste 8 jaren van zijn loopbaan was hij tevens decaan van de faculteit Werktuigbouwkunde. In 1990 werd hij opgevolgd door Henk Zijm als hoogleraar Productie- en Operations Management.

Wim Bakker heeft het vakgebied Productieorganisatie op een bewonderenswaardige wijze uitgebouwd. Hij was werktuigbouwkundige van huis uit en promoveerde bij Van Hasselt. Zijn benadering was multi-disciplinair, een mengsel van werktuigbouwkunde, bedrijfskunde en wiskunde, in het bijzonder de Operations Research. Hij had een uitstekend oog voor het aanwenden van goede theorie in praktische problemen. In Nederland was hij een pionier in het gebruik van kwantitatieve methoden binnen de productie-organisatie en daarmee eigenlijk van de introductie van wat we nu noemen Operations Management. In zijn intrede uit 1966 'Over het nut van detailonderzoek' toonde hij zich daarnaast scherp bewust van het belang van een goede structurering van de infor-

matiestromen in het bedrijf, in een tijd waarin de informatietechnologie haar grote opmars in de bedrijfssystemen nog moest beginnen. Ook het wetenschappelijk onderzoek kreeg onder zijn leiding vorm; enkele van zijn oud-promovendi, waaronder Erik-Joost de Bruijn, Gerard Gaalman en Hein Fleuren, zijn nu hoogleraar. Een zeer groot aantal studenten studeerde in zijn leerstoel af.

Naast zijn vele bijdragen aan de ontwikkeling van het vakgebied heeft Wim een groot aantal nevenfuncties vervuld, binnen en buiten de faculteit. Zo was hij bijvoorbeeld zeer actief binnen het KIVI en heeft hij in Europees verband tesamen met prof. Muller-Mehrbach veel gewerkt op het gebied van onderwijs in de Operations Research. Wim was een uitermate hartelijk en warm iemand, begaan met het lot van medewerkers, studenten en oud-studenten, zoals ook bleek uit de grote belangstelling toen hij in 1991 afscheid nam van de universiteit.

Wij wensen zijn vrouw Annie en kinderen veel sterkte toe bij het verwerken van dit grote verlies.

PROF.DR. HENK ZIJM, decaan faculteit Elektrotechniek, Wiskunde en Informatica, Universiteit Twente.

PROF.DR.IR. HEIN FLEUREN, director CentER Applied Research, Tilburg, Voorzitter Nederlands Genootschap voor Besliskunde.

Tas, R.F.J. (2002), Overlevenden per generatie. *CBS Maandstatistiek van de bevolking*, maart 2002, blz. 14-16.

NOTEN

- 1 De in dit artikel weergegeven opvattingen zijn die van de auteur en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek.
- 2 Deze statistiek moet niet worden verward met de statistiek van de doodsoorzaken die is gebaseerd op de z.g. doodsoorzaakverklaringen. Deze worden ingevuld door de arts die het overlijden vaststelt.
- 3 Mensen worden gemiddeld halverwege het jaar geboren, en dus zullen de kinderen die op 31 december van hun geboortjaar nog in leven zijn, dan gemiddeld een half

jaar oud zijn. Het CBS geeft in de overlevingstafels de leeftijd weer die mensen op 31 december hebben. Het voordeel van deze aanpak is dat leeftijden vertaalbaar zijn in geboortejaren en omgekeerd. Alle mensen die in de overlevingstafel dezelfde leeftijd hebben, zijn in hetzelfde jaar geboren, en mensen die in hetzelfde jaar zijn geboren, hebben in de tabel dezelfde leeftijd. Leeftijdsspecifieke sterftequotiënten hebben daarmee betrekking op mensen die in een zelfde jaar zijn geboren.

4 In de overlevingstafel is het gebruikelijk de levensverwachting met een kleine letter e aan te duiden in plaats van een hoofdletter E.

5 *Geboortequotiënten hebben betrekking op vrouwen van 15, 16, ..., 49 jaar.*

Gelijke kansen voor alles en iedereen

FRED STEUTEL

Gelijke kansen is een beginsel waar veel mensen voor te vinden zijn; zo zeer dat men niet alleen aan mensen gelijke kansen wil geven, maar ook aan gebeurtenissen, terwijl voor dat laatste vaak geen enkele reden is. Men doet dit met name in gevallen waar sprake is van twee gebeurtenissen; dan treedt een soort kruis-of-munt-mechanisme in werking. Hieronder een paar (tegen)voorbeelden.

D' Alembert

Dat bij het werpen van een geldstuk de kansen op 'kruis' (K) en 'munt' (M) gelijk zijn en dus gelijk aan een half vindt bijna iedereen vanzelfsprekend. Bij het gooien met twee geldstukken wordt al wat moeilijker. Volgens de bekende geleerde d'Alembert waren er dan drie mogelijkheden: twee keer K, twee keer M of van elk één; drie mogelijkheden, die elk kans een derde zouden moeten hebben. Dat bleek op grond van experimenten niet houdbaar. We leren nu op de middelbare school dat we KM en MK moeten onderscheiden, dat er dus vier mogelijkheden zijn, elk met kans een kwart. Bij iets ingewikkelder situaties gaat het toch weer makkelijk mis.

Kastenprobleem

Het volgende probleem was lange tijd in het nieuws en heeft de statistische gemoederen soms ernstig verhit. Het gaat om een spelprogramma op de TV, waarbij drie identiek uitziende kasten zijn uitgesteld; in één daarvan zit een aanzienlijke prijs, de andere kasten zijn leeg. De speler mag een kast aan wijzen, waarvan hij dan de inhoud

mag hebben. Als de speler een kast heeft aangegeven, maar nog niet geopend, doet de spelleider een lege kast open en vraagt aan de speler of die niet toch liever zijn kast wil ruilen tegen de derde, nog ongeopende, kast. De kwestie die de gemoederen bezig hield was: moet de speler de ruil aangaan, moet hij dat juist niet doen, of maakt het voor de kans dat hij de prijs binnenhaalt niets uit? Ook een enkele statisticus van naam ging de mist in – was tenslotte ook een mens.

Het is duidelijk dat de kans $1/3$ is dat de door de speler 'op de gok' aangewezen kast de prijs bevat. Dat zelfde geldt voor elk van de niet-aangewezen kasten. Alle kasten hebben dus kans $1/3$ en ruilen heeft dan volgens het beginsel der gelijke kansen geen zin, maar De kans dat de prijs in één van de beide andere twee kasten zit is natuurlijk $2/3$ (voor elke kast $1/3$), en de spelleider heeft een kast open gemaakt, waarvan hij wist dat hij leeg was. De kans dat de prijs in de derde (niet aangewezen, gesloten) kast zit is daarom $2/3$. Ruilen dus!

Het is eenvoudiger om als volgt te redeneren. Wie niet ruilt krijgt alle prijzen (nul of één) die in de gekozen kast zitten; wie wel ruilt krijgt alle prijzen (nul of één) die in de andere twee kasten samen zitten: ruilen! Alle kansen zijn gelijk, maar twee kansen zijn meer dan één.

Hoe is het mogelijk?

Het volgende voorbeeld lijkt verwant aan het vorige, maar is toch anders en tamelijk bizar. Een spelleider heeft in twee enveloppen briefjes gedaan met daarop geschreven een geldbedrag. In de ene envelop

zit een briefje met positief bedrag a , in de andere één met $2a$. Het getal a is onbekend; het kan tien zijn, maar ook tienduizend of tien miljoen; de enveloppen zijn van buiten niet van elkaar te onderscheiden. Een speler mag één van de enveloppen kiezen, open maken, kijken welk bedrag er op het briefje staat en dan besluiten of hij wil ruilen of niet. De vraag weer: heeft ruilen ooit enige zin?

Het is niet zo eenvoudig: als je vijftig euro op je briefje vindt, is het dan het dubbele van vijftentwintig of de helft van honderd? Als je zou weten dat de bedragen tussen de tien en de tien miljoen liggen, dan zou je soms weten wanneer je het kleinste (grootste) bedrag in handen had: als je 16 leest op het briefje, dan kan dat niet $2a$ zijn, want dan zou a kleiner dan 10 zijn; ruilen dus. Vind je 6 miljoen op het getrokken briefje, dan moet dat wel $2a$ zijn; dus niet ruilen. Maar wat te doen als je over het getal a helemaal niets weet?

Ook dan is er een truc waarmee je de kans om het grootste van de twee bedragen te pakken te krijgen groter dan $\frac{1}{2}$ kunt maken. Dat betekent dus dat je het beter kunt doen dan 'op goed geluk een envelop te kiezen en niet ruilen'. Dat gaat als

volgt. Kies een envelop, maak hem open en lees het getal, zeg A (dat is dus a of $2a$). Doe vervolgens een trekking X uit een kansverdeling met (een continue, stijgende) verdelingsfunctie F , dus met de eigenschap dat $P(X < y) = F(y)$ en $P(X > y) = 1 - F(y)$. Is nu $A > X$, ruil dan niet; immers A is 'aan de grote kant'. Is $A < X$, ruil dan wel; A is aan de kleine kant.

Wat is nu de kans dat de speler uiteindelijk het grootste bedrag ($2a$) te pakken krijgt? Dat is de kans dat hij a getrokken heeft (kans $\frac{1}{2}$) en wel ruilt plus de kans dat hij $2a$ getrokken heeft (kans $\frac{1}{2}$) en niet ruilt, dat wil zeggen (het rekenwerk voor de liefhebbers):

$$P(\text{speler krijgt } 2a) = \frac{1}{2} P(X > a) + \frac{1}{2} P(X < 2a) = \frac{1}{2} \{1 - F(a) + F(2a)\}$$

Dat is groter dan $\frac{1}{2}$, omdat $F(2a) > F(a)$. We weten niet hoe groot a is, dus we weten niet hoeveel de kans groter is dan $\frac{1}{2}$, maar groter is-ie. Is hier sprake van een wonder? Wie het weet mag het zeggen!

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

A G E N D A

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

15 januari 2003

6de Lunteren Seminar, georganiseerd door NGB en LNMB. Thema: **On-line methods: Challenges for OR in a real-time world**. Dit Seminar is onderdeel van de 3-daagse conferentie **Mathematische Besliskunde**, die voor de 29ste keer plaatsvindt. Voor informatie en inschrijving zie <www.math.leidenuniv.nl/~lnmb/conferences/lunteren2004>.

13-17 februari 2004

Het thema van de 14e Study door de International Commission on Mathematical Instruction (ICMI) is **Applications and Modelling in Mathematics Education**. Plaats: Dortmund. Nadere informatie: <www.brocku.ca/mathematics/ICMI/study14>.

19 februari 2004

Workshop on Optimization in Public Transport, Erasmus Universiteit Rotterdam. Georganiseerd door ECOPT (Erasmus Center for Optimization in Public Transport). Sprekers zijn Guy Desaulniers (GERAD, Canada), Raymond Kwan (University of Leeds, UK), Michiel Vromans en Ramon Lentink (ECOPT). Info <www.few.eur.nl/few/research/ecopt/workshop>.

16-17 april 2004

Nederlands-Belgisch Mathematisch Congres o.a.v. het Koninklijk Wiskundig Genootschap en het Belgisch Wiskundig Genootschap. Het congres vindt in Tilburg plaats. *Hoofdvoordrachten*: Neil Sloane (AT&T), Bernard De Baets (Universiteit van Gent), Stef Tijs (Universiteit van Tilburg). *Speciale voordrachten*: Beegerlezing door Manjul Bhargava (Princeton University), een veilingles door Casper de Vries (EUR), een casinoleles door Ben van der Genugten (UvT). Verder minisymposia, presentaties promotieonderzoek en korte voordrachten. Voor informatie: <www.uvt.nl/nmc2004>.

22 april 2004

De **VSO-presentatiedag** is een dag waar bureaus op het gebied van onderzoek en statistiek zich kunnen presenteren. Informatie: <www.vsonet.nl>.

26-27 mei 2004

De Katholieke Universiteit Leuven organiseert een colloquium **Risk Analysis: Statistical and Probabilistic Methods**. Voor informatie: <www.kuleuven.ac.be/ucs/colloq.htm>.

21-25 juni 2004

De 13e ECMI (European Conference on Mathematics for Industry) conferentie zal plaatsvinden in Eindhoven. Informatie: <www.ecmi2004.tue.nl>.



1-3 september 2004

Internationale conferentie Operations Research 2004 o.a.v. Gesellschaft für Operations Research (GOR) en Nederlands Genootschap voor Besliskunde (NGB). De conferentie vindt in Tilburg plaats. *Sprekers*: Prof.dr. Martin Grötschel (Berlijn), Prof.dr. Lex Schrijver (Amsterdam), Prof.dr. Gerard Sierksma (Groningen), Prof.dr. Luk Van Wassenhove (Parijs). Deadline bijdragen 30 april; aanmelden tot 31 mei 2004. Voor informatie en het programma: <www.uvt.nl/OR2004> en <OR2004@uvt.nl>.



Stata Release 8

NIEUW IN STATA 8

- » Graphics: een nieuwe look, publication quality
- » Grafische gebruikersinterface: alle commando's via een dialoog uitvoeren
- » Database management: ODBC support, meer missing values
- » Meer statistische functies: MANOVA, MANCOVA en VAR, SVAR meer...
- » Programmeerfuncties: object georiënteerd programmeren, GUI
- » Stata 8 is sneller

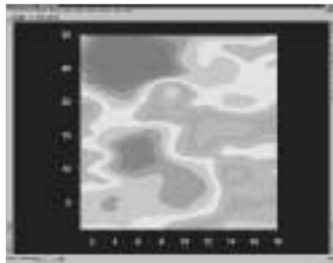
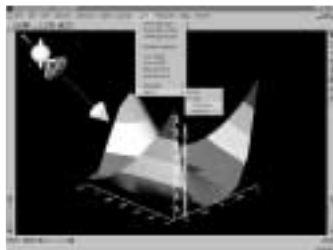


post Postbus 220, 5150 AE Drunen
bezoek Grotestaat 401a, 5142 CB Waalwijk
telefoon 0416 - 378 125 (Stata) of 0416 - 543 444,
fax 0416 378 385, www.smitconsult.nl

Smit Consult is ondergebracht bij Cosinus Computing BV

PROBABLY THE BEST STATISTICS PACKAGE CURRENTLY ON THE MARKET!

6.25	45	44
6.25	40	42
6.25	44	45
7.5	Identifier	
7.5	42	block
	FACTOR	load



The new **GenStat® for Windows® 6th Edition**, available now, boasts a long list of new features and improvements.

GenStat's reputation for the quality and accuracy of its statistical functionality is unparalleled. GenStat is developed at Rothamsted Experimental Station where many of the standard statistical techniques still in use were first discovered, and it has served the needs of the statistical community worldwide for over 30 years!

The package is developed by statisticians who understand the real needs of users and have first hand experience of dealing with challenges in the 'real world'.

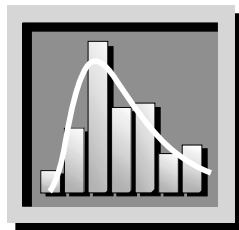
GenStat has the broadest range of statistical tools, all available in the standard package with no additional modules to install and pay for.

With its own, powerful and flexible high-level language, GenStat allows you to extend and customise its functionality virtually without limits. (Should you prefer to work with the user-friendly and familiar Windows® menu structure without writing a single line of 'code' that is, of course, also possible.)

If you have not tried GenStat recently this would be a good time to take another look! GenStat has been completely revised based on major market research and user feedback initiatives. In fact, in a recent series of focus group studies the respondents (consisting of professional users of statistical software from a variety of sectors) ranked GenStat as the best package currently on the market! Contact us to find out why...

For more information please contact:

Cosinus Computing, Postbus 220, 5150 AE Drunen, The Netherlands
tel. +31 (0) 416 543 444, fax +31 (0) 416 378 385, info@cosinus.nl
<http://www.cosinus.nl>



GenStat®

With GenStat you know you can!