

Boekbesprekingen

Redactie

Alex J. Koning

postadres: Econometrisch Instituut
Erasmus Universiteit Rotterdam
Postbus 1738
3000 DR Rotterdam
telefoon: 010-4081268/2231
fax: 010-4527746
e-mail: koning@few.eur.nl

Besproken boeken in Kwantitatieve Methoden nr. 63

Chilès, J.-P., Delfiner, P.

Geostatistics: modeling spatial uncertainty

Stein, M.L.

Interpolation of spatial data. Some methods for kriging

Kirkpatrick, L., Feeney, B.

A simple guide to SPSS for Windows version 8

Lohr, S.

Sampling. Design and analysis

Tijms, H.

Spelen met kansen

Hossack, I.B., Pollard, J.H., Zehnwirth, B.

Introductory statistics with applications in general insurance

Voelkl, K.E., Gerber, S.

Using SPSS for Windows. Data analysis and graphics

Huisman, M.

Item nonresponse: occurrence, causes, and imputation of missing answers to test items

Hosmer, D.W., Lemeshow, S.

Applied survival analysis: regression modeling of time to event data

Bijleveld, C.J.H., van der Kamp, L., Mooijaart, A., van der Kloot, W.A., van der Leeden, R., van der Burg, E.

Longitudinal data analysis. Designs, models and methods

Good, P.I.

Resampling methods. A practical guide to data analysis

Bradley, W.J., Schaefer, K.C.

The uses and misuses of data and models. The mathematization of the human sciences

Chilès, J.-P., Delfiner, P.

Geostatistics: modeling spatial uncertainty

1998, Wiley series in probability and mathematical statistics,
John Wiley & Sons Ltd., Chichester,
695 blz, \$ 125.00, ISBN 0-471-08315-1.

Recent zijn er verschillende boeken verschenen op het gebied van geostatistiek en kriging. Onder andere het boek van Chilès en Delfiner, en ook dat van Stein (zie elders in deze rubriek). Het boek van Chilès en Delfiner komt voort uit de praktijk van de geostatistiek, en leunt sterk op de Franse school, met name op de werken van Matheron. Het boek bevat veel aspecten die we tegenwoordig op het gebied van geostatistiek tegen komen. Het eerste hoofdstuk behandelt variogrammen, vanaf de wolk van variogram waarden tot modellen voor variogrammen en variogrammen voor niet-stationaire situaties. Het leunt weliswaar op de covariantie functie, maar dit concept komt in dit boek nauwelijks aan de orde.

Het tweede hoofdstuk richt zich op kriging, d.w.z. lineair voorspellen. De verschillen tussen simpel kriging, d.w.z. voorspellen met een bekend gemiddelde, en ordinary kriging, voorspellen met constant maar onbekend gemiddelde, worden helder beschreven. Ook is er aandacht aan een keuze van de omgevingsgrootte.

In het volgende hoofdstuk komen intrinsieke modellen van orde k aan de orde, en daarmee gegeneraliseerde covariantie functies en universal kriging. Voor zover ik kan beoordelen is dat voor het eerst dat deze belangrijke en interessante concepten zo goed en helder beschreven zijn. Coëfficiënten worden geschat met de momenten methode.

Als multivariate methoden komen cokriging aan de orde en multivariate interpolatie, d.w.z. het voorspellen van meerdere variabelen tegelijkertijd op hetzelfde punt. Cokriging wordt zowel met een constante verwachting behandeld als met een trend. Ook verschijnen in dit hoofdstuk de ruimte-tijd modellen, zij het summier.

Niet-lineaire methodes zijn met name gebaseerd op het iso-factoriele model. Belangrijkste daarbij is dat blokgemiddelden kunnen worden voorspeld (zoals de oogst op een veld op basis van enkele metingen) en dat de kansen op overschrijding van (milieu)normen worden bepaald.

Een fors hoofdstuk is gewijd aan conditionele simulaties. Dit geeft een zeer helder en volledig overzicht van hoe random fields kunnen worden gesimuleerd en wat we er mee kunnen doen (foutenvoortplanting analyseren, simulatie studies uitvoeren). Met name dit onderdeel over simulaties sprak me erg aan, door de zorgvuldige en gedetailleerde aanpak.

Een laatste korte hoofdstuk is gewijd aan schaal effecten en invers modelleren. Beide onderdelen spelen op dit moment een belangrijke rol in landbouw en milieu onderzoek. De auteurs behandelen het vrij kort, maar het geeft zeker een goede inleiding in het onderwerp.

Dit boek is een uitstekend boek. Helder en op een manier opgeschreven dat het nauw aansluit bij landbouw en milieu problemen en bij de mijnbouw en de hydrogeologische praktijk. Inmiddels is al een hele generatie studenten in Wageningen op-

geleid met de concepten die nu voor het eerst goed in boekvorm beschreven zijn. Op sommige plaatsen is het boek wellicht niet meer helemaal up-to-date, met name als het gaat om ruimtelijk bemonsteren, waar recent duidelijk voortgang geboekt is. Ook model-based geostatistiek, zoals beschreven in het artikel van Diggle en anderen ontbreekt. Aan de andere kant zijn de voorbeelden wel overtuigend, vaak afkomstig uit de hydro-geologie, en met name de beschrijving van ruimtelijke variatie rond de kanaal tunnel is prachtig en toont het belang van geostatistiek voor de moderne praktijk nog eens levendig aan. Al met al is het boek een echte aanrader, voor allen die vanuit praktisch of statistisch perspectief belangstelling hebben in ruimtelijke fenomenen.

prof. dr A. Stein

Universiteit Wageningen, ITC Enschede

Stein, M.L.

Interpolation of spatial data. Some methods for kriging

1999, Springer texts in statistics,

Springer-Verlag, Berlin,

275 blz, DM 98.00, ISBN 0-387-98629-4.

Via de VVS

25% korting!

Recent zijn er verschillende boeken verschenen op het gebied van geostatistiek en kriging. Onder andere het boek van Stein, en ook dat van Chilès en Delfiner (zie elders in deze rubriek). Het boek van Stein (geen familie) is sterk wiskundig van aard en concentreert zich vooral op één aspect van geostatistiek, namelijk kriging (lineair voorspellen). Het is zeker geen gemakkelijk boek en leunt sterk op probabilistische en wiskundige grondslagen.

Het eerste hoofdstuk behandelt lineaire voorspelling als een soort inleiding. De bekende regressie vergelijkingen van kriging (BLUPs) worden gepresenteerd, alsmede de voorkeur van de auteur voor het Matérn co-variante model. Uitgangspunt zijn Hilbert ruimten, en de auteur gaat op basis daarvan in op wiskundige aspecten van interpolatie.

In eigenschappen van random fields, het tweede hoofdstuk, gaat de auteur in op covariante functies en de ontbindingen ervan met name met behulp van spectraal dichtheden. De turning bands methode wordt gebruikt als een methode om een random field te simuleren. Het variogram wordt summier behandeld, evenals gegeneraliseerde random fields.

Asymptotische eigenschappen van lineaire voorspellers betreffen met name de uitbreidingen van eindige naar oneindige roosters. Veel aandacht wordt besteed aan het verschil tussen interpolatie en extrapolatie en aan het voorspellen met de verkeerde spectraaldichtheid.

Hoofdstuk 4 behandelt equivalentie tussen Gaussische maten en voorspellingen. Deze equivalentie wordt eerst bewezen, en vervolgens toegepast, niet zozeer op een praktijkvoorbeeld, maar om bijvoorbeeld asymptotische optimaliteit van BLUPs te bewijzen.

In het vijfde hoofdstuk, integratie van random fields, wordt nog eens gekeken naar asymptotische eigenschappen van het gemiddelde, hoe dit verbeterd kan worden, met name op een oneindig rooster van punten.

Het laatste, en langste hoofdstuk, betreft voorspellingen met geschatte parameters. Micro-ergodiciteit wordt gebruikt in relatie tot Gaussische maten, en vervolgens worden likelihood methoden gebruikt om parameters te schatten, van met name de Matérn co-variatie functie. Stein noemt dit de plug-in voorspellingen, en hun kwaliteiten worden vervolgens geëvalueerd met behulp van gangbare methoden.

Het boek begint waar de praktijk van de geostatistiek ophoudt en de theorie begint. Het boek is als monograph indrukwekkend, de diepgang die bereikt wordt is zeker in de vaak op de praktijk gerichte ruimtelijk statisticus uniek. Het is een eerste solide poging om een meer wiskundig fundament te leggen onder een vaak wat heuristisch gebruik van ruimtelijke interpolatie. Het is een boek dat ik zeker aan wiskundig gerichte AIOs aan zou raden, als een heerlijke kluit om regelmatig in bijten. Veel aspecten hebben nog wel een duidelijke vertaalslag naar de praktijk nodig. In het boek staan ook geen voorbeelden, wel opgaven die echter vooral verdere uitbreidingen van de theorie betreffen. Dit gegeven zijnde denk ik dat we het boek zeker als een waardevolle bijdrage aan de literatuur over een boeiend en fascinerend onderwerp kunnen beschouwen.

*prof. dr A. Stein
Universiteit Wageningen, ITC Enschede*

Kirkpatrick, L., Feeney, B.

A simple guide to SPSS for Windows version 8

1999, South Western, Belmont, California,
144 blz, £ 11.50, ISBN 0-534-36460-8.

In deze publicatie vindt men een weliswaar duidelijke maar uiterst summiere inleiding in het gebruik van de Windows versie van SPSS. De opzet is dat het gebruikt wordt naast een elementaire statistiektekst en daartoe wordt een twaalfstal standaardopgaven, gegoten in een psychologische context, uitgewerkt binnen SPSS versie 8.

De desbetreffende twaalf statistische onderwerpen omvatten histogrammen, allerlei ANOVA situaties, lineaire regressie, en een chi-kwadraat toets voor onafhankelijkheid.

Opvallend is het ontbreken van een uitwerking van een normaliteitstoets en van een residuenonderzoek.

De opdrachten zijn zo elementair, dat ze volledig uitvoerbaar zijn met de via de muis bereikbare menus. Daarnaast worden de equivalente syntax commando's gegeven. De verkregen output is voorzien van een summiere interpretatie. In een appendix wordt nog het zeer nuttige opslaan en binnenhalen van databestanden uitgelegd, alsmede de constructie in SPSS van nieuwe grootheden die berekend worden uit reeds

bekende variabelen. Het gidsje wordt door twee nietjes bij elkaar gehouden en bevat geen index. Het is geschikt als een allereerste kennismaking met SPSS.

*dr H.W.M. Hendriks
Subfaculteit Wiskunde
Katholieke Universiteit Nijmegen*

Lohr, S.

Sampling. Design and analysis

1999, South Western, Belmont, California,
450 blz, £ 25.00, ISBN 0-534-35361-4.

Het betreft een cursus die een aantal aspecten behandelt van het steekproefsgewijze onderzoeken van een concreet bestaande populatie. De besproken onderwerpen zijn nogal standaard, zoals men bijvoorbeeld vindt in W.G. Cochran, *Sampling Techniques*. Laten we het bijzondere van het boek aanstippen. De auteur doet een poging lezers van verschillende niveaus aan te spreken. Met name is er een groot deel dat zuiver 'design-gebaseerd' is, waarmee bedoeld wordt dat men hoofdzakelijk geïnteresseerd is in de eigenschappen van de concreet bestaande populatie, en niet zo zeer in de eigenschappen van een denkbeeldig model waarvan de concrete populatie een realisatie is. Dit gedeelte is bedoeld voor alle lezers. Verder zijn er verspreid in het boek model-gebaseerde benaderingen. Deze zijn wat mathematischer getint en zijn gericht tot de lezers die al een achtergrond hebben in de mathematische statistiek, en de verdieping van de mathematische model-gebaseerde statistiek heeft betrekking op het bijzondere van waarnemingen in een concrete populatie, zoals de kansstructuur volgens welke een individu in een steekproefneming voorkomt, nonresponsie, het bestaan van exact bekende populatiegrootheden. Zo gauw de wiskunde de overhand krijgt of er in termen van modellen geredeneerd wordt is de betreffende paragraaf, opgave of hoofdstuk gemarkeerd.

Er is een grote collectie opgaven opgenomen. Bij iedere opgave is duidelijk gemaakt voor welk lezersniveau dit bedoeld is. Er zijn geen uitwerkingen opgenomen. Innovatief is de bijvoeging van een CD-ROM die de databestanden in het boek bevat en het SURVEY softwarepakket. SURVEY simuleert gegevens voortkomend uit een door de gebruiker gedefinieerd onderzoek ten behoeve van een fictieve kabeltelevisie-maatschappij in een gegeven fictieve provincie, opgedeeld in een tweetal steden en een landelijk gebied met een totaal inwonertal van 103441, wonend in 31989 huizen enz.. Ook kan men de proportie (random) nonrespondenten ingeven. Een gedeelte van de opgaven nodigt uit om zelfstandig onderzoek te doen in deze populatie.

Het boek is prettig geschreven en ieder hoofdstuk wordt prima ingeleid en gemotiveerd. Er zijn vele waarschuwingen voor complicaties bij het nemen van steekproeven, en die kun je maar beter ter harte nemen, als je wilt uitkomen boven het 'beschrijvende' niveau, in de geest van: 'wij hebben 2860 personen aangeschreven, waarvan er 1314

gereageerd hebben en onze bevindingen waren dat zus en zoveel procent van de respondenten zus en zo'n antwoord heeft gegeven'. Het effect ervan is dat de aandacht van de lezer vastgehouden wordt.

De auteur probeerde, naar eigen zeggen, zoveel mogelijke echte data te gebruiken. De meeste voorbeelden hebben dus betrekking op gegevens uit werkelijk uitgevoerde onderzoeken. Wat ik echter in het boek mis, met name om de door boven vermelde waarschuwingen veroorzaakte psychische druk te verlichten, is een verslag van een of ander daadwerkelijk bevolkingsonderzoek, waarin een met redenen omklede toelichting op het ontwerpvoorstel, de uitvoering van het onderzoek, de methode van afwikkeling van nonresponsie en ontbrekende items en non-sampling fouten en de verdere toegepaste analysemethode.

*dr H.W.M. Hendriks
Subfaculteit Wiskunde
Katholieke Universiteit Nijmegen*

**Tijms, H.
Spelen met kansen**

1999, Epsilon Uitgaven, Utrecht,
208 blz, fl 37.50, ISBN 90-5041-051-0.

Dit boek geeft een uitstekende inleiding in de kansrekening. Het is uniek vanwege de ruime aandacht die wordt besteed aan het ontwikkelen van inzicht en gevoel voor kansen aan de hand van veel interessante voorbeelden en toepassingen. Hierbij wordt nadrukkelijk gebruik gemaakt van computersimulatie, dat een modern en perfect hulpmiddel is om kansbegrippen en resultaten te illustreren en te verhelderen. Centraal in het boek staan de twee 'natuurwetten' van de kansrekening, namelijk de wet van de grote aantallen en de centrale limietstelling.

Het boek is helder en vlot geschreven. Het is zeer geschikt om te gebruiken bij een inleidende cursus kansrekening voor niet-exacte studies. Het materiaal in het boek leent zich prima voor onderwijsvormen, waarbij studenten meer zelfstandig de stof moeten bestuderen en eigen maken. In het begin van het boek wordt een studieroute voorgesteld en het boek bevat veel leuke en stimulerende opgaven. Voor exacte studies, waar het tempo bij een inleidende cursus hoger zal liggen, is het beperkte aantal onderwerpen en resultaten dat in het boek wordt behandeld, een bezwaar. Hier kan het boek prima worden gebruikt naast een traditioneel boek om het ontwikkelen van inzicht en gevoel voor kansen te ondersteunen. Dit is namelijk ook voor exacte studenten uiterst belangrijk.

Het boek omvat 7 hoofdstukken. Het eerste hoofdstuk geeft een korte inleiding en bevat een lesroute voor het bestuderen van het boek. In hoofdstuk 2 worden een aantal uitdagende kansproblemen geformuleerd, met name om het gevoel voor kansen bij de lezer te testen. Een leuk en berucht probleem is bijvoorbeeld het quizdeurenprobleem, waarin achter een van de drie gesloten deuren de hoofdprijs staat en de

quizmaster je in de gelegenheid stelt om van gekozen deur te veranderen nadat hij een deur heeft geopend waarachter de hoofdprijs niet staat. De oplossingen voor de problemen in dit hoofdstuk worden gegeven in de volgende hoofdstukken. Dit leidt tot een sterke samenhang van de gepresenteerde stof. In het volgende hoofdstuk staat de wet van de grote aantallen centraal. Aan de hand van voorbeelden zoals het werpen met een munt en het roulettespel wordt inzicht verschaft in deze belangrijke wet. Computersimulaties laten perfect zien wat deze wet wel (en niet) vertelt over het verloop van kansprocessen. Er is ook aandacht voor het zelf schrijven van een simulatieprogramma. De Pascal code voor een aantal problemen is ter illustratie opgenomen in de tekst. In hoofdstuk 4 komt het gebruik van computersimulatie voor het oplossen van kansproblemen uitgebreid aan de orde en daarin maakt de lezer ook kennis met een aantal belangrijke oplossingsmethoden uit de kansrekening. Boeiende problemen uit bijvoorbeeld de casino- en loterijfeer worden behandeld. De problemen worden opgelost zowel door het opzetten van een simulatiemodel als door het analyseren van het kansmodel. In hoofdstuk 5 worden een aantal bekende kansverdelingen behandeld en wordt inzicht gegeven in de verschillende soorten situaties waarin deze kansverdelingen van toepassing zijn. Hoofdstuk 6 is gewijd aan de normale verdeling en de centrale limietstelling. Aan de hand van kansdiagrammen voor het werpen met (on)zuivere dobbelstenen wordt de werking van de centrale limietstelling prachtig geïllustreerd. In het laatste hoofdstuk komt het gebruik van kansbomen aan de orde. Deze aanpak wordt toegepast op het quizdeurenprobleem. Achterin het boek zijn tenslotte een aantal appendices te vinden. In deze appendices wordt nader ingegaan op begrippen, die in de voorbeelden op natuurlijke manier te voorschijn zijn gekomen (zoals bijvoorbeeld voorwaardelijke kansen en de verwachtingswaarde). De presentatie van de stof in de appendices is wat formeler dan in de rest van het boek.

Dit is een uniek boek over kansrekening. Het is niet alleen geschikt als leerboek, maar het is aan te bevelen aan iedereen die spelenderwijs kennis wil maken met de fascinerende wereld van de kansrekening.

*I.J.B.F. Adan
Faculteit Wiskunde en Informatica
Technische Universiteit Eindhoven*

Hossack, I.B., Pollard, J.H., Zehnwrith, B.

Introductory statistics with applications in general insurance

1999, Cambridge University Press, Cambridge,
282 blz, £ 22.95, ISBN 0-521-65534-X.

Afgezien van tabellen en opgaven bevat dit boekje bevat ca. 240 bladzijden uitleg. Dat is niet veel voor een cursus statistiek, waarbij dan ook nog voor de verzekeringswereld specifieke toepassingen aan de orde moeten komen. Van de lezer die het zelfstandig wil doorwerken wordt een behoorlijk inzicht en begrip verwacht. Veel uitleg gaat aan de hand van voorbeelden. Niet altijd volgt daarna de formele beschrijving. De inleiding

over wiskunde-voorkennis lijkt me niet erg op zijn plaats. Wie deze nodig heeft, zal in de rest van het boek (althans zonder docent of begeleider) vastlopen. Recente ervaring met, zeg, wiskunde B op VWO-niveau lijkt noodzakelijk. Verder blijken de auteurs uit te gaan van een lezer die met de verschillende begrippen uit de verzekeringswereld bekend is.

Hoofdstuk 2–6 bevatten beginselen van de kansrekening. Bij de verdelingen ligt de nadruk op die die in de verzekeringswereld een rol spelen. Dat zijn dan in de eerste plaats de normale (met een wel erg uitgebreide kansentabel van maar liefst vijf bladzijden - en dat in deze tijd van computers!), de log-normale, Pareto-, gamma-, Poisson-, binomiale en negatief-binomiale verdeling.

Het principe van toetsen wordt aan de hand van de binomiale verdeling gintroducteerd. In hetzelfde hoofdstuk 6 wordt dan de Poisson-verdeling gebruikt voor het toelichten van de meest aannemelijke schatters en de betrouwbaarheidsintervallen. De momenten-methode wordt verbaal beschreven, maar een voorbeeld ontbreekt. De hoofdstukken 7–11 zijn geheel aan verzekeringsonderwerpen gewijd.

De volledige inhoudsopgave luidt:

1. Introduction and mathematical preliminaries
2. Elementary probability
3. Random variables and their distributions
4. Location and dispersion
5. Statistical distributions in general insurance work
6. Inferences from general insurance data
7. The risk premium
8. Experience rating
9. Simulation
10. Estimation of outstanding claim provisions
11. Elementary risk theory

Het elementaire gedeelte zit aardig in elkaar. De uitleg is echter verre van wijdlogig, wat de beginner voor problemen zal stellen. Het aantal opgaven (alle met antwoord) is voor een lesboek erg klein.

In de hoofdstukken 7–11 treffen we een afwisseling aan van uitleg met case-studies en zuivere theorie. Hier en daar is er veel aandacht voor details betreffende berekeningen, die soms niet moeilijker zijn dan optellen en vermenigvuldigen. Dan weer wordt met vanzelfsprekendheid gegrepen naar integralen, de normale benadering en gamma-verdelingen.

De opmaak is verzorgd en meestal overzichtelijk. Tabellen op zijn kant echter vind ik minder gelukkig (en soms vermijdbaar), en een enkele keer is de volgorde tussen tabel, figuur en begeleidende tekst niet handig. Sommige paragrafen zijn gemerkt met één of twee sterretjes om alternatieve leerroutes door het boek mogelijk te maken. Helaas komt het een aantal malen voor dat verderop gebruik wordt gemaakt van zulke paragrafen zonder dat expliciet naar de gemerkte (en dus misschien ongelezen) delen wordt verwezen.

Het boekje is zeker leuk als inspiratiebron voor een docent die een statistiekcursus binnen een actuariële opleiding verzorgt of voor een docent schadeverzekering die een statistische aanpak wat wil illustreren. Voor zelfstudie lijkt het me niet geschikt. Als cursusmateriaal is het op zichzelf te mager, maar aangevuld met extra opgaven, extra theorie en een levende instructeur kan het zeker worden gebruikt. De hoofdstukken 1–6 zijn te kort voor het aanbrengen van kennis op het gebied van wiskunde en statistiek op het niveau waarop er in de hoofdstukken 7–11 soms gebruik van wordt gemaakt. De titel zou dan ook eerder omgekeerd moeten worden gelezen: toepassingen van statistiek in verzekeringen met een bloemlezing van wat statistiek. In ieder geval is het boekje bruikbaar om te worden getoond aan een student die vraagt waarom hij toch al die statistiek moet leren.

J.M. Buhrman

Hogeschool voor Economische Studies

Voelkl, K.E., Gerber, S.

Using SPSS for Windows. Data analysis and graphics

1999, Springer-Verlag, Berlin,

265 blz., DM 69.00, ISBN 0-387-98563-8.

Via de VVS
25% korting!

Dit boek kan op twee manieren als handleiding gebruikt worden: 'stand-alone' voor mensen die al wat van statistiek afweten en zelfstandig SPSS for Windows willen leren, alsook als handleiding naast een boek in het statistiek onderwijs. Qua opbouw wordt 'The New Statistical Analysis of Data' door Anderson en Finn gevolgd, maar de gehanteerde terminologie is standaard, dus ook andere begeleidende studieboeken zijn denkbaar.

Het boek gebruikt SPSS for Windows versie 8.

Het boek bestaat uit vijf delen. Het eerste deel van het boek is een algemene inleiding op SPSS en gaat over opstarten, afsluiten, en omgaan met ASCII en SPSS datafiles. Ook het construeren van nieuwe variabelen wordt behandeld. Het is jammer dat geen aandacht wordt besteed aan kalenderdata. Nu kan iemand lang zoeken voordat hij uitgevonden heeft hoe hij, bijvoorbeeld, leeftijd kan uitrekenen op basis van geboortedatum en datum van meting.

Het tweede deel van het boek gaat over SPSS procedures voor beschrijvende statistiek. Zoals de ondertitel van het boek al suggereert, worden zaken als histogrammen, stem-and-leaf-, box-and-whisker- en scatter-plots uitgebreid getoond. Hierbij wordt

niet alleen ingegaan op hoe je de plaatjes op je scherm krijgt, maar ook op de interpretatie. De auteurs melden op blz. 44 dat, ten tijde van het schrijven, kwartielen niet altijd accuraat werden berekend door SPSS. Dit is waarschijnlijk nog steeds het geval.

Het derde deel gaat over waarschijnlijkheidstheorie en kan overgeslagen worden als het boek alleen gebruikt wordt voor het leren van data-analyse. In dit deel leer je het genereren van random getallen, het vinden van kansen en z -waarden voor de familie van normale verdelingen, alsook het trekken van een steekproef uit een normale verdeling. Eén van de histogrammen laat goed zien dat 50 trekkingen uit een dergelijke verdeling niet een duidelijke klokcurve hoeven op te leveren.

In het vierde deel, over inferentiële statistiek, laten de auteurs zien hoe je met SPSS vragen over populatiekarakteristieken kunt beantwoorden. Daarbij gaat het zowel om het onderzoeken van één populatie (ook met gepaarde metingen) als om het vergelijken van twee populaties met betrekking tot gemiddelde, proportie en lokatie (de teken-toets).

Het vijfde deel tenslotte, bevat een drietal hoofdstukken over uiteenlopende onderwerpen: inferentie voor categoriale gegevens (chi-kwadraat toetsen, de coëfficiënten λ en γ), regressie analyse (enkelvoudige en multiple regressie, met o.a. een voorbeeld met dummy-codering) en het vergelijken van meerdere populaties (een-weg-ANOVA, ook niet-parametrisch).

De structuur van elk hoofdstuk is hetzelfde: een korte introductie gevolgd door voorbeelden die de lezer kan nadoen door de benodigde datafiles van Internet te halen en de genummerde instructies te volgen. Elk hoofdstuk wordt afgesloten met een aantal oefeningen. De datasets behorende bij deze oefeningen zijn ook op Internet beschikbaar. De instructies zijn erg duidelijk en zelfs te volgen als je leest zonder SPSS onder handbereik, mede doordat veel dialoogschermen als illustratie in het boek zijn opgenomen.

Het boek bevat een groot aantal voorbeelden, met slechts een enkele inconsequentie (bijvoorbeeld figuur 9.1 waarin het gemiddelde in de uitvoer verschilt van het gemiddelde in de tekst). Alle voorbeelden worden uitgewerkt met 'clicks'. Het is mogelijk syntaxfiles te schrijven en zo optimaal gebruik te maken van de mogelijkheden van SPSS, maar dit wordt alleen gemeld en niet besproken. De uitvoer wordt over het algemeen uitgebreid besproken. Echter, de adjusted R^2 (op uitvoer in hoofdstuk 15), wordt niet toegelicht.

Een sterk punt van dit boek is dat het laat zien dat zaken die standaard in de uitvoer zijn opgenomen niet per definitie ook zinvol zijn, bijvoorbeeld cumulatieve frequenties bij nominale variabelen.

Alhoewel dit boek SPSS versie 8 gebruikt is het ook bruikbaar voor versie 9. Met betrekking tot de behandelde stof in het boek zijn de wijzigingen beperkt maar opvallend: het pull-down menu 'statistics', heet in versie 9 'analyse', en het menu 'summarize' heet in versie 9 'descriptive statistics'.

Voor een ervaren statisticus die SPSS for Windows aan haar toolkit wil toevoegen, is dit boek wat langzaam en elementair. Het boek geeft nauwelijks een indruk van wat er aan analysemethoden in SPSS beschikbaar is, zodat het voor een statisticus minder

geschikt is als naslagwerk.

Voor studenten is dit boek zeer geschikt, zowel als handleiding SPSS naast een inleidend boek als 'stand-alone' manual om zelfstandig SPSS for Windows te leren. Het boek bevat veel goed uitgewerkte voorbeelden en een behoorlijke index waardoor het voor beginners geschikt is als naslagwerk.

H. Tobi

*Klinische Epidemiologie en Biostatistiek
Vrije Universiteit Amsterdam*

Huisman, M.

Item nonresponse: occurrence, causes, and imputation of missing answers to test items

1999, M & T series,

DSWO Press, Leiden,

200 blz, fl 41.30, ISBN 90-6695-144-3.

Non-response in enquêtes is een probleem dat volop in de belangstelling staat. Toegenomen aandacht voor dit onderwerp in tijdschriften en andere publicaties, maar ook een congres als 'International Conference on Survey Nonresponse 99', dat in oktober plaatsvond met een ruime vertegenwoordiging van Nederlandse statistici, zijn daar getuigen van. Eén van de redenen van die toegenomen aandacht is de groeiende actualiteit van het non-response probleem: het publiek is kennelijk 'enquête-moe', maar bovendien, nieuwe media als internet die in toenemende mate worden benut voor het afnemen van enquêtes blijken om voor de hand liggende reden ook tot veel hogere non-respons te leiden.

Het non-respons probleem kan op een aantal manieren benaderd worden, die alle ook in het proefschrift van Mark Huisman aan bod komen. Het eerste hoofdstuk geeft een inleiding in de materie, en behandelt veel voorkomende oorzaken van non-respons, tezamen met daarmee samenhangende patronen van ontbrekende waarnemingen. In het tweede hoofdstuk wordt het vraagstuk van non-respons vertekening (bias) aan de hand van een zestal dataverzamelingen uit empirisch gedragswetenschappelijk onderzoek besproken: zijn er systematische verschillen tussen de antwoorden van respondenten en van niet-respondenten. Hoofdstuk drie bespreekt vervolgens de meest directe methode om de ontdekken of dit type vertekening aanwezig is, en wat de oorzaak er van kan zijn: het herbenaderen van non-respondenten. En hoofdstuk vier is daar dan weer de logische vervolgstap op: nadat eerst een onderscheid wordt gemaakt tussen diverse bronnen van non-respons, worden daarna de gegevens uit het vervolgonderzoek gebruikt om de vertekening als gevolg van de non-respons te verminderen.

In het tweede deel van het proefschrift, dat de hoofdstukken 5 t/m 7 omvat, wordt een andere benadering van het non-respons probleem beschreven: die van de 'imputatie', het schatten van de ontbrekende gegevens om daarmee de dataverzameling aan te vullen. Het zal duidelijk zijn dat voor dit schattingsprobleem meer of minder

geavanceerde 'imputatiemodellen' gebruikt kunnen worden. In hoofdstuk 5 worden een aantal eenvoudige modellen besproken en in een simulatiestudie op hun kwaliteit onderzocht. Hoofdstuk 6 bespreekt meer geavanceerde imputatiemodellen, zoals IRT-modellen.

De zwakkere en sterkere punten van het boek? Ze liggen steeds heel dicht bij elkaar. Het kenmerkende van het proefschrift is dat alle hoofdstukken (op de inleiding en epiloog na) geschreven zijn als bijdragen voor tijdschriften (hoofdstukken 2 en 3 zijn al in die vorm verschenen, in respectievelijk 'Quality and Quantity: European Journal of Methodology' en 'Kwantitatieve Methoden'). Het aantrekkelijke van die opzet is dat ieder hoofdstuk goed op zich zelf te lezen is; de interpedentie tussen verschillende hoofdstukken is minimaal. Dit voordeel is tegelijkertijd een nadeel voor andere lezers: wil je het proefschrift als 'boek' in één keer uitlezen, dan vind je wel erg veel redundante informatie op je weg. Verder geldt dat het boek niet in de eerste plaats als een overzicht van hedendaagse inzichten op het gebied van item non-respons te beschouwen is. Het is in z'n benadering selectiever, en laat zich dan ook vooral lezen als een zeer volledige gevalstudie: hoe in empirisch onderzoek het verschijnsel non-respons zo goed mogelijk voorkomen kan worden en, indien het toch zou optreden, hoe de gevolgen tot een minimum beperkt kunnen worden. Door die invalshoek zal het boek dan ook eerder toegepaste dan theoretische onderzoekers aanspreken. Het proefschrift is er toegankelijk genoeg voor: het is helder geschreven en zorgvuldig samengesteld (op één nogal curieuze, welhaast Freudiaanse, vergissing na: volgens de inhoudsopgave gaat hoofdstuk 3 als enige over het 'respons-probleem').

*D. Tempelaar
Capaciteitsgroep Kwantitatieve Economie
Universiteit Maastricht*

Hosmer, D.W., Lemeshow, S.

Applied survival analysis: regression modeling of time to event data

1999, Wiley series in probability and statistics,

John Wiley & Sons Ltd., Chichester,

386 blz, \$ 79.95, ISBN 0-471-15410-5.

Dit boek gaat over de analyse van 'overlevingsduren': wachttijden op een bepaalde gebeurtenis. Zulke wachttijden kunnen gecensureerd zijn en dat vraagt om een speciale aanpak. De context van dit boek is medisch, de methoden zijn uiteraard bruikbaar ook in andere toepassingsgebieden.

De stof is gepresenteerd in 9 hoofdstukken. De eerste hoofdstuk (27 blz) introduceert de problematiek en basisbegrippen, zoals de censurering. Aan de hand van een voorbeeld worden de overeenkomsten en verschillen met de lineaire en de logistische regressie besproken. Ook worden hier drie data sets gintroduceerd die het boek als voorbeelden gebruikt. Twee gaan over overleving en AIDS, het derde over overleving na een hartaanval. De data sets kan men op de website van de uitgever

ophalen. Hoofdstuk 2 (61 blz) bespreekt het schatten van overlevingsfunctie en het vergelijken van onafhankelijke steekproeven. Hoofdstukken 3 tot en met 7 (251 blz), de kern van het boek, zijn gewijd aan het semi-parametrische regressie-model van Cox. Aan bod komen het model, het schatten van parameters en van de overlevingsfunctie (Hoofdstuk 3), het interpreteren van het geschatte model (Hoofdstuk 4), het opbouwen van het model en het selecteren van de variabelen (Hoofdstuk 5), de evaluatie van de modelfit (Hoofdstuk 6), het gestratificeerde model en het model met tijdsvariërende verklarende variabelen (Hoofdstuk 7). Hoofdstuk 8 (57 blz) gaat over parametrische regressie-modellen. In Hoofdstuk 9 (48 blz) komen diverse topics zoals frailty- en additieve modellen aan bod. Elk hoofdstuk eindigt met opgaven. Meeste hebben betrekking op de analyse van een van de gebruikte data sets en zijn sterk praktisch gericht. Antwoorden op de vraagstukken geeft het boek niet. Het boek bevat verder summieri aanhangsels over de delta-methode en over telprocessen.

Het boek doet sterk denken aan het succesvolle boek over de logistische regressie [1] van dezelfde auteurs. Dezelfde lay-out, maar vooral dezelfde degelijke en rustige opbouw, met zorgvuldig uitgewerkte voorbeelden en goed bruikbare opgaven. Het is duidelijk gericht op de praktijk. Een voorbeeld van het laatste is een uitgebreide - en mijns inziens nuttige - bespreking van de keuzes bij de grafische presentatie van het geschatte regressiemodel. Het boek veronderstelt een goede basiskennis van statistiek. Het is iets minder wiskundig en compleet dan het boek van Klein en Moeschberger [2], maar mijns inziens toegankelijker.

Samenvattend: Ik verwacht dat het boek heel goed te gebruiken is bij het op de praktijk gerichte onderwijs. Ook voor de zelfstudie vind ik het een aanrader.

Verwijzingen:

- [1] Hosmer, D.W. and Lemeshow, S. (1989). *Applied Logistic Regression*. John Wiley and Sons.
- [2] Klein, J.P. and Moeschberger, M.L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer.

*dr V. Fidler
Vakgroep Medische Statistiek
Rijksuniversiteit Groningen*

Bijleveld, C.J.H., van der Kamp, L., Mooijaart, A., van der Kloot, W.A., van der Leeden, R., van der Burg, E.

Longitudinal data analysis. Designs, models and methods

1998, Sage, Thousand Oaks, California,
448 blz, £ 18.99, ISBN 0-7619-5538-0.

Dit boek, waaraan in totaal zes auteurs hebben meegewerkt, geeft een overzicht van een aantal belangrijke technieken die gebruikt kunnen worden om longitudinale data in kaart te brengen (datareductie), te modelleren en te analyseren.

Longitudinale data wordt in het boek gerepresenteerd als een databox, een driedimensionale datamatrix. Gegeven deze matrix zijn verschillende modes waarin de data

bestudeerd kan worden met elk zijn voor- en nadelen. Deze worden duidelijk uitgelegd en vergeleken. Er wordt uitgegaan van een discrete tijd. De reden hiervoor is dat in gedrags- en sociaal onderzoek dit meestal het geval is en dat het boek juist op dit type onderzoek gericht is. Daarnaast zijn continue tijd analysetechnieken in de regel gecompliceerder. Voor elk type data, dat wil zeggen numeriek of categorisch, in de longitudinale dataset worden verschillende technieken behandeld en vergeleken.

Een korte, niet geheel volledige, opsomming van de onderwerpen die aan bod komen in het boek is: Steekproefontwerpen en andere methodologische aspecten voor longitudinale studies, Optimale schaaltechnieken zoals Multiële correspondentieanalyse, Niet-lineaire principale componenten analyse, Niet-lineaire gegeneraliseerde canonische analyse en Multivariate dynamische analyse voor longitudinale data, Univariate en Multivariate variantie-analyse (waaronder herhaalde metingen ANOVA, MANOVA), Structural equation modellen zoals Linear structural equation modellen, Autoregressieve modellen, Groeicurve modellen, Multilevel modellen, Log-lineaire modellen, Markov modelleringen. Zonder al te veel op de achterliggende statistiek en wiskunde in te gaan, worden de ideeën achter de technieken gegeven, rijk gelardeerd met voorbeelden uit de praktijk. Aan het eind van elk hoofdstuk wordt er aandacht besteedt aan het aspect van ontbrekende gegevens en op welke manier je hiermee om kan gaan.

Het boek is bedoelt voor onderzoekers die longitudinaal onderzoek verrichten, en dan met name voor onderzoekers op het gebied van psychologie, pedagogiek, geriatrie, medische psychologie en sociologie. De harde statistiek en formules zijn tot een minimum beperkt en daarom is een grote statistische achtergrond niet vereist om het boek te lezen. De schrijvers zijn er goed in geslaagd om de geavanceerde technieken eenvoudig uit te leggen zodanig dat ze met behulp van statistische software zoals SPSS toegepast kunnen worden. Door het lezen van dit boek kan een onderzoeker snel een overzicht krijgen wat er allemaal aan recente technieken voorradig zijn voor longitudinale data die hem/haar kunnen helpen om zijn onderzoeksvragen te beantwoorden.

Zoals gewoonlijk het geval is bij sociaal onderzoek gaat men uit van multivariate data. Er wordt dan ook nauwelijks aandacht besteedt aan univariate technieken uit de tijdreeksanalyse zoals autocorrelaties en trendanalyses.

Ondanks het feit dat het boek door zes auteurs is geschreven is er duidelijk geprobeerd (met succes) om de het hele boek in dezelfde lijn te schrijven. Het boek leest prettig, zeker door de interessante en aansprekende voorbeelden.

Een sterk punt van het boek is dat het een scala aan (recente) technieken geeft die voor handen zijn om longitudinale data aan te pakken. Naast het beschrijven van worden de technieken ook naast elkaar gelegd en vergeleken. Er wordt veel aandacht besteed aan de interpretatie van de uitkomsten. Daarnaast wordt er, zoals eerder gezegd, ruimschoots aandacht besteedt aan ontbrekende gegevens. Dit is in andere boeken vaak een ondergeschoven kindje omdat het een lastig onderwerp blijft. Het boek is zeker de moeite waard om te lezen, en dan met name voor onderzoekers die beschikken over

een standaard statistiek kennis en hun horizon willen verbreden.

dr C. Oudshoorn

TNO - Preventie en Gezondheid

Good, P.I.

Resampling methods. A practical guide to data analysis

1998, Scientific computing,

Birkhäuser, Basel,

332 blz, DM 138.00, ISBN 3-7643-4091-6.

Zoals de subtitel aangeeft, is dit boek bedoeld als een praktische gids voor data-analyse met bootstrap- en permutatiemethoden. Het beoogt een inleiding te zijn, die leesbaar is zonder dat daarvoor kennis van geavanceerde wiskunde en statistiek nodig is. Drie resampling methoden komen uitgebreid aan de orde. Allereerst de bootstrap-methode, waarbij uit de voorliggende steekproef voortdurend nieuwe steekproeven getrokken worden, en de variabiliteit van de schattingen beschouwd wordt als een indicatie van de steekproevenvariantie. Daarbij zijn twee varianten te onderscheiden: de nonparametrische bootstrap, waarbij de steekproeven getrokken worden uit de voorliggende data, en de parametrische bootstrap, waarbij de steekproeven gesimuleerd worden op basis van een statistisch model, gebaseerd op de voorliggende data. De tweede resampling methode die behandeld wordt is de permutatie-methode. De permutatiemethode kan onder andere gebruikt worden om hypothesen te toetsen over groepsverschillen. Daartoe wordt eerst een relevante toetsingsgrootheid berekend. Vervolgens worden de gegevens volgens een toevalsproces herhaald geordend, en daardoor at random opnieuw over de groepen verdeeld. Telkens wordt de betreffende toetsingsgrootheid opnieuw berekend, en de verdeling van de op toeval gebaseerde toetsingsgrootheden wordt beschouwd als de steekproevenverdeling van de betreffende toetsingsgrootheid. Tenslotte behandelt Good density estimation. Density estimation kan onder andere gebruikt worden bij klassificatieproblemen, zoals het schatten van mengsels van verdelingen, met als doel het aantal en de vorm van de onderliggende verdelingen te schatten.

Het boek is zeer uitgebreid, met veel voorbeelden. De expositie van de verschillende resampling-methoden wordt steeds opgehangen aan kleine voorbeelden, waarbij de dataset gegeven wordt (voor wie een hekel heeft aan het intikken van data: ze zijn ook beschikbaar via Internet). Daarnaast worden regelmatig stroomschema's en stappenplannen gegeven (bijvoorbeeld "Five steps to a Permutation Test", in drie bladzijden behandeld en toegelicht). De voorbeelden worden uitgerekend met een eigen programma (ook op te halen via Internet) en met resamplingprocedures beschikbaar in Stata, SAS of StatExact. Voor programmeurs wordt in een bijlage ook een serie C++ procedures gegeven, plus SAS en S-Plus code.

Het boek gaat duidelijk verder dan een eenvoudige inleiding, zoals bijvoorbeeld het boek van Mooney en Duval (1993) in de bekende Sage 'groene boekjes' reeks.

Good geeft een goede en leesbare inleiding in resampling methoden, maar behandelt daarna ook meer gecompliceerde methoden zoals punt- en intervalschattingen, bias-correctie, smoothing, en de geïtereerde bootstrap. Daarnaast gaat Good in op het onderscheidingsvermogen van resampling methoden, multivariate analyses, en gecensureerde data.

Good is geen nieuwteling op het gebied van resampling methoden. Eerder schreef hij het boek *Permutation Tests* (1994), een meer technische inleiding. Het boek is goed geschreven, en prettig leesbaar. Het niveau is echter nogal wisselend. In de Preface stelt Good dat het boek bedoeld is voor "Physicians and physicians in training, nurses and nursing students, business persons, business majors, research workers and students in the biological and social sciences", maar ook voor "...advanced students in biology, dentistry, medicine, psychology, sociology, and public health", en ook nog voor "industrial statisticians, statistical consultants, and research workers". De poging om al deze groepen te bedienen is mijns inziens niet helemaal geslaagd. Voor de beginners vangt het boek aan met een bespreking van beschrijvende statistische grootheden, legt uit wat het gemiddelde en de standaard deviatie is, en bespreekt het verschil tussen steekproef en populatie. Ongeveer halverwege het boek vinden we een gedetailleerde uitleg van multivariate analyse en Hotellings T^2 . Het boek eindigt met classificatieproblemen, mengverdelingen en gecensureerde gegevens. Het uitlegen van een gemiddelde en een standaard afwijking zonder ingewikkelde formules kan heel goed. Ongeveer halverwege het boek blijken complexe formules toch niet gemist te kunnen worden, en gelukkig geeft Good die dan ook. Maar naar mijn inschatting markeert dit ook een punt waar de inleiding overgaat in een tekst voor meer gevorderden, waar de "nurses and business persons" ongetwijfeld afhaken. De echte statistici daarentegen zullen uiteindelijk niet kunnen zonder Efron en Tibshirani (1993) of een vergelijkbaar boek. Daarmee is de ecologische nis voor dit boek gegeven. De echte beginners zou ik voor een toegankelijke inleiding in resampling methoden het boek van Mooney en Duval aanraden (1993). Daarna is het boek van Good een bijzonder goede aanvulling, omdat hij voor een scala aan statistische procedures de resampling variant bespreekt, met veel aandacht voor praktische toepassingen, inclusief de aannamen en keuzes die in de analysefase gemaakt moeten worden. Mathematisch georiënteerde statistici zullen zich meer thuis voelen bij het eerder genoemde boek van Efron en Tibshirani (1993).

Verwijzingen:

- [1] Efron, B. and Tibshirani, R. (1993). *An introduction to the bootstrap*. London: Chapman & Hall.
- [2] Good, P.I. (1994). *Permutation tests*. New York/Berlin: Springer.
- [3] Mooney, C.Z., Duval, R.D. (1993). *Bootstrapping: a nonparametric approach to statistical inference*. Newbury Park, CA: Sage.

prof. dr J.J. Hox
Dept. of Methodology and Statistics
Utrecht University

Bradley, W.J., Schaefer, K.C.

The uses and misuses of data and models. The mathematization of the human sciences

1998, Sage, Thousand Oaks, California,
xii+212 blz, £ 12.99, ISBN 0-7619-0922-2.

Dit boek gaat in op het gebruik en misbruik van data en modellen in de sociale wetenschappen. Het doel van het boek is om ethische normen en principes te formuleren voor het gebruik van data en modellen in de sociale wetenschappen. Het is geschreven voor onderzoekers en studenten in de sociale en beslissingswetenschappen en beleidsmedewerkers.

Het boek bestaat uit twee delen. Het eerste deel "Foundations" kent drie hoofdstukken.

1. Oracles, Norms and Science
2. Modeling
3. Dreams and Disappointments

Het tweede deel is: "the information cycle" en bestaat uit de volgende hoofdstukken.

4. A priori influences on the information cycle
5. Measurements on human information
6. Limitations of measurement in the Social Sciences
7. Information for Inferences: What are social science data?
8. Causality
9. Models and policy making.

In het eerste deel gaan de auteurs in op de problemen van sociaal wetenschappelijk onderzoek in een praktische beleidsgeoriënteerde setting. Sociaal wetenschappelijk onderzoek en sociaal wetenschappelijke informatie spelen een grote rol in het hedendaagse leven en daarom is het nodig dat er op een ethisch verantwoorde manier mee wordt omgegaan. Voor onderzoek wordt vaak gebruikt gemaakt van een modellering van de werkelijkheid. De auteurs geven aan dat modellering beperkingen kent, omdat het onmogelijk is om alle informatie over de werkelijkheid in het model op te nemen. Dit kan ertoe leiden dat er resultaten worden geproduceerd die niet erg bruikbaar of zelfs misleidend zijn. In het derde hoofdstuk wordt aangegeven dat de verwachtingen voor sociaal wetenschappelijk onderzoek vaak hoog zijn, maar dat deze verwachtingen niet waargemaakt worden: modellen alleen zijn vaak niet voldoende om de praktijk te veranderen.

In het tweede deel van het boek willen de auteurs richtlijnen geven voor het uitvoeren van verantwoord sociaal wetenschappelijk onderzoek. De richtlijnen moeten zijn gebaseerd op twee punten, namelijk het respecteren van de beperkingen van sociaal wetenschappelijk onderzoek en de onderzoekscontext. Als uitgangspunt voor het opstellen van de richtlijnen, maken de auteurs gebruik van de "information cycle" (vergelijkbaar met de empirische cyclus van A.D. de Groot).

In hoofdstuk 4 wordt ingegaan op de waarden en assumpties die een rol spelen bij het opstellen van een model. Het gaat hierbij niet alleen om statistische assumpties, maar vooral over de waarden en normen van een onderzoeker. De volgende drie hoofdstukken (5, 6 en 7) gaan over het meten en analyseren in de sociale wetenschappen. Het bevat een uitleg van de verschillende meetniveaus die kunnen worden onderscheiden en de beperkingen van metingen in de sociale wetenschappen. De auteurs zijn in deze hoofdstukken zeer strikt in de leer: men mag alleen analyses doen op het niveau van de gevonden data. Helaas ontbreekt in deze hoofdstukken informatie over methoden om met behulp van een model, zoals item response modellen, de verkregen informatie naar een hoger meetniveau te brengen of over de consequenties van het gebruik van een hoger meetniveau als het gaat om de conclusies van een onderzoek. In hoofdstuk 7 wordt ingegaan op de verschillende aspecten van het analyseren van de gegevens met als belangrijkste doel het verkrijgen van een conclusie. Daarbij wordt gekeken naar onder meer de gevolgen van het gebruik van analysemethoden. Deze hebben veelal a-priori assumpties, zo veronderstelt regressieanalyse een lineaire relatie tussen de variabelen, die beperkingen opleggen aan de mogelijkheden tot het doen van uitspraken. In de meeste gevallen is niet bekend hoe de relatie in de werkelijkheid is en het is dus niet altijd eenvoudig om aan te geven of deze aanname wel zo goed is. De auteurs pleiten er daarom ook voor om onderzoek te doen naar de assumpties van de analyses en vooral om goed te kijken naar de residuen.

In hoofdstuk 8 wordt ingegaan op de vraag naar causaliteit. De auteurs geven aan dat kennis omtrent causaliteit zeer belangrijk is: het stelt ons in staat om zaken daadwerkelijk te kunnen veranderen, maar het kunnen concluderen of er sprake is van causaliteit is niet makkelijk, soms zelfs onmogelijk. De stellingen van Lazarsfeld worden aangehaald, maar daarbij wordt aangegeven dat het nodig is om een mechanisme te kunnen geven hoe de causaliteit plaatsvindt. Een hoe-vraag moet worden beantwoord naast het voldoen aan de eisen van Lazarsfeld. De auteurs geven ook aan dat lang niet altijd kan worden voldaan aan de eisen van Lazarsfeld. Want hoe stel je vast dat er geen andere variabele is die de relatie tussen twee andere beïnvloedt?

Het laatste hoofdstuk gaat over het omzetten van onderzoek naar beleid. In dit hoofdstuk wordt ingegaan op de teleurstellingen die sociaal wetenschappelijk onderzoek te weeg hebben gebracht. In de jaren 60 was toch de verwachting dat onderzoek veranderingen te weeg konden brengen, maar dat bleek maar al te vaak niet zo te zijn. De verwachtingen van sociaal wetenschappelijk onderzoek waren veelal gebaseerd op de successen van natuurwetenschappelijk onderzoek: de mogelijkheid van een maakbare sociale wereld werd als doel gezien. Om teleurstellingen te voorkomen moeten de opdrachtgevers van een onderzoek al in een vroeg stadium worden betrokken bij het

onderzoek zodat deze een beeld hebben van wat ze met het onderzoek zouden kunnen doen.

De normen die door de auteurs geformuleerd zijn de normen die al gelden voor het doen van onderzoek, alleen worden ze nog strenger dan gebruikelijk geformuleerd. De auteurs benadrukken nogal dat opdrachtgevers zo veel mogelijk moeten worden betrokken bij de beslissingen die voor het modelleren moeten worden genomen. Veel onderzoekers zullen echter de ervaring hebben dat een opdrachtgever dit nu juist aan de onderzoekers willen overlaten om er zelf niet over na te hoeven denken. Dit boek wil waarschijnlijk te veel van beleidmedewerkers.

Het boek zou ook gelezen moeten worden door beleidmedewerkers, zodat deze weten wat de mogelijkheden en onmogelijkheden van sociaal wetenschappelijk onderzoek zijn.

Het is overigens spijtig dat onderwerpen uit het gebied van toegepast beleidsonderzoek niet aan de orde komen. Juist in die literatuur wordt ingegaan op de belangen en doelstellingen die beleidsmedewerkers kunnen hebben die vaak diametraal staan op de doelstellingen die de auteurs van dit boek hebben.

In de inleiding geven de auteurs echter een belangrijke reden waardoor problemen kunnen ontstaan, namelijk de verregaande specialisering in de sociale wetenschappen. Hierdoor ontstaan vertaalproblemen in de verschillende stappen bij het vergaren van gegevens. De auteurs zouden daarom willen dat sociaal wetenschappers breed opgeleid zouden worden. Een visie die ik met hen deel, maar die gezien de ontwikkelingen in de financiering van het wetenschappelijk onderwijs een illusie zal moeten blijven.

*P. Moorer
ARGO b.v.
Groningen*

Binnengekomen boeken

Hieronder volgt een overzicht van de recente uitgaven die sinds de vorige aflevering van Kwantitatieve Methoden bij de redactie zijn binnengekomen. Mocht er belangstelling bestaan voor het bespreken van een boek, dan kan contact opgenomen worden met de boekbesprekingsredacteur. Indien het boek dan nog niet vergeven is, krijgt de kandidaat het boek gratis thuisgezonden. Hiervoor dient dan binnen een half jaar een recensie te worden teruggezonden. Een aantal titels is niet fysiek toegezonden, maar kan door ons besteld worden bij de uitgever. Dit heeft consequenties voor de levertijd van deze titels.

Headayat, A.S., Sloane, N.J.A., Stufken, J.
Orthogonal arrays. Theory and applications

1999, Springer series in statistics,
 Springer-Verlag, Berlin,
 440 blz, DM 126.00, ISBN 0-387-98766-5.

Via de VVS
 25% korting!

Hernandez-Lerma, O., Lassere, J.-B.
Further topics on discrete-time Markov control processes

1999, Applications of mathematics,
 Springer-Verlag, Berlin,
 310 blz, DM 119.00, ISBN 0-387-98694-4.

Via de VVS
 25% korting!

Cowell, R.G., Dawid, A.P., Lauritzen, S.L., Spiegelhalter, D.J.
Probabilistic networks and expert systems

1999, Statistics for engineering and information science,
 Springer-Verlag, Berlin,
 340 blz, DM 139.00, ISBN 0-387-98767-3.

Via de VVS
 25% korting!

Sapsford, R.
Survey research

1999, Sage, Thousand Oaks, California,
 272 blz, £ 15.99, ISBN 0-7619-5528-3.

Sirkin, R.M.
Statistics for the social sciences

1999, Sage, Thousand Oaks, California,
 576 blz, £ 24.99, ISBN 0-7619-1419-6.

Le Gall, J.-F.

Spatial branching processes, random snakes and partial differential equations

1999, Lectures in mathematics,

Birkhäuser, Basel,

176 blz, DM 44.00, ISBN 3-7643-6126-3.

Banuelos, R., Moore, Ch.N.

Probabilistic behavior of harmonic functions

1999, Progress in mathematics,

Birkhäuser, Basel,

224 blz, DM 118.00, ISBN 3-7643-6062-3.

Bramson, M., Durrett, R.T. (ed.)

Perplexing problems in probability. Festschrift in honor of Harry Kesten

1999, Progress in probability,

Birkhäuser, Basel,

408 blz, DM 158, ISBN 3-7643-4093-2.

Foata, D, Fuchs, A.

Wahrscheinlichkeitsrechnung

1999, Grundstudium Mathematik,

Birkhäuser, Basel,

400 blz, DM 98.00, ISBN 3-7643-6169-7.

Kallianpur, G., Karandikar, R.L.

Introduction to option pricing theory

1999, Birkhäuser, Basel,

280 blz, DM 128.00, ISBN 3-7643-4108-4.

Balakrishnan, N., Glaz, J.

Scan statistics and applications

1999, Statistics for industry and technology,

Birkhäuser, Basel,

346 blz, DM 168.00, ISBN 3-7643-4041-X.

Grätzer, G.

First steps in LaTeX

1999, Birkhäuser, Basel,

136 blz, DM 44.00, ISBN 3-7643-4132-7.

Kaufmann, S.

A crash course in Mathematica version 3.0

1999, Birkhäuser, Basel,

200 blz, DM 48.00, ISBN 3-7643-6127-1.

Hoffmann, K.-H., Leugering, G., Tröltzsch, F.

Optimal control of partial differential equations. International conference in Chmenitz, Germany, April 20–25, 1998

1999, International series of numerical mathematics,

Birkhäuser, Basel,

336 blz, DM 188.00, ISBN 3-7643-6151-4.

Dragan, V., Halanay, A.

Stabilization of linear systems

1999, Systems and control: foundations and applications,

Birkhäuser, Basel,

256 blz, DM 148.00, ISBN 3-7643-3970-5.

Petersen, I., Savkin, A.

Robust Kalman filtering for signals and systems with large uncertainties

1999, Control engineering,

Birkhäuser, Basel,

280 blz, DM 178, ISBN 3-7643-6137-9.

Coxon, A.P.M.

Sorting data. Collection and analysis

1999, Sage, Thousand Oaks, California,

96 blz, £ 8.99, ISBN 0-8039-7237-7.

Chilès, J.-P., Delfiner, P.

Geostatistics: modeling spatial uncertainty

1998, Wiley series in probability and mathematical statistics,
John Wiley & Sons Ltd., Chichester,
695 blz, \$ 125.00, ISBN 0-471-08315-1.

Grasman, J.

Wiskundige methoden toegepast

1999, Epsilon Uitgaven, Utrecht,
230 blz, fl 42.50, ISBN 90-5041-0053-7.

Leonard, T., Hsu, J.S.J.

Bayesian methods: an analysis for statisticians and interdisciplinary researchers

1999, Cambridge series in statistical and probabilistic mathematics,
Cambridge University Press, Cambridge,
xiv+333 blz, £ 40.00, ISBN 0-521-59417-0.

Serfozo, R.

Introduction to stochastic networks

1999, Applications of mathematics,
Springer-Verlag, Berlin,
330 blz, DM 139.00, ISBN 0-387-98773-8.

Via de VVS
25% korting!

Robert, C.P., Casella, G.

Monte Carlo statistical methods

1999, Springer texts in statistics,
Springer-Verlag, Berlin,
530 blz, DM 159.00, ISBN 0-387-98707-X.

Via de VVS
25% korting!

Efromovich, S.

Nonparametric curve estimation

1999, Springer series in statistics,
Springer-Verlag, Berlin,
425 blz, DM 148.00, ISBN 0-387-98740-1.

Via de VVS
25% korting!

Singpurwalla, N., Wilson, S.P.

Statistical methods in software engineering. Reliability and risk

1999, Springer series in statistics,

Springer-Verlag, Berlin,

310 blz, DM 139.00, ISBN 0-387-98823-8.

Via de VVS

25% korting!

Loader, C.

Local regression and likelihood

1999, Statistics and computing,

Springer-Verlag, Berlin,

305 blz, DM 129.00, ISBN 0-387-98775-4.

Via de VVS

25% korting!

Bilodeau, M., Brenner, D.

Theory of multivariate statistics

1999, Springer texts in statistics,

Springer-Verlag, Berlin,

315 blz, DM 146.00, ISBN 0-387-98739-8.

Via de VVS

25% korting!

Durrett, R.

Essentials of stochastic processes

1999, Springer texts in statistics,

Springer-Verlag, Berlin,

305 blz, DM 134.00, ISBN 0-387-98836-X.

Via de VVS

25% korting!

Billingsley, P.

Convergence of probability measures, 2nd edition

1999, Wiley series in probability and statistics,

John Wiley & Sons Ltd., Chichester,

ix+277 blz, £ 65.00, ISBN 0-471-19745-9.

Nijse, M.

Schoolkeuze en schoolloopbaan na de basisschool

1998, Vakgroep Psychologie, Rijksuniversiteit Groningen,

91 blz, fl 20.00, ISBN 90-367-1014-6.

Everitt, B.S.

Chance rules. An informal guide to probability, risk and statistics

1999, Springer-Verlag, Berlin,
202 blz, DM 49.00, ISBN 0-387-98768-1.

Via de VVS
25% korting!

Rao, C.R., Toutenburg, H.

Linear models. Least squares and alternatives

1999, Springer series in statistics,
Springer-Verlag, Berlin,
xvi+427 blz, DM 136.00, ISBN 0-387-98848-3.

Via de VVS
25% korting!

Venables, W.N., Ripley, B.D.

Modern applied statistics with S-PLUS

1999, Statistics and computing,
Springer-Verlag, Berlin,
xii+501 blz, DM 129.00, ISBN 0-387-98825-4.

Via de VVS
25% korting!

Beltrami, E.

What is random? Chance and order in mathematics and life

1999, Springer-Verlag, Berlin,
xx+201 blz, DM 39.90, ISBN 0-387-98737-1.

Via de VVS
25% korting!

Handcock, M.S., Morris, M.

Relative distribution methods in the social sciences

1999, Statistics for social science and public policy,
Springer-Verlag, Berlin,
xiii+265 blz, DM 119.00, ISBN 0-387-98778-9.

Via de VVS
25% korting!

Ross, S.M.

Introduction to the mathematics of finance: options and other topics

1999, Cambridge University Press, Cambridge,
xv+184 blz, £ 21.95, ISBN 0-521-77043-2.

Waller, D.L.

Operations management. A supply chain approach

1999, International Thomson Business Press, London,
xvii+841 blz, £ 26.95, ISBN 1-86152-415-3.

Hodson, R.

Analyzing documentary accounts

1999, Quantitative applications in the social sciences,
Sage, Thousand Oaks, California,
96 blz, £ 8.99, ISBN 0-7619-1743-8.

Adér, H.J., Mellenbergh, G.J. (ed)

Research methodology in the social, behavioural and life sciences. Designs, models and methods

1999, Sage, Thousand Oaks, California,
416 blz, £ 18.99, ISBN 0-7619-5884-3.

Nocedal, J, Wright, S.J.

Numerical optimization

1999, Springer series in operations research,
Springer-Verlag, Berlin,
xx+636 blz, DM 129.00, ISBN 0-387-98793-2.

Via de VVS

25% korting!

Hox, J.J, de Leeuw, E.D. (ed)

Assumptions, robustness, and estimation methods in multivariate modelling

1999, TT-publikaties, Amsterdam,
224 blz, fl 60.00¹, ISBN 90-801073-6-0.

Wilkinson L.

The grammar of graphics

1999, Statistics and computing,
Springer-Verlag, Berlin,
xvii+408 blz, DM 139.00, ISBN 0-387-98774-6.

Via de VVS

25% korting!

Haller, G.

Chaos near resonance

1999, Springer texts in statistics,

Springer-Verlag, Berlin,

xvi+427 blz, DM 139.00, ISBN 0-387-98697-9.

Via de VVS

25% korting!

Härdle, W., Klinke, S., Müller, M.

XploRe. The interactive statistical computing environment

1999, Springer-Verlag, Berlin,

iv+526+CD-ROM blz, DM 170.69, ISBN 3-540-14767-5.

Via de VVS

25% korting!

Härdle, W., Klinke, S., Müller, M.

XploRe - learning guide

1999, Springer-Verlag, Berlin,

iv+526 blz, DM 79.00, ISBN 3-540-66207-3.

Via de VVS

25% korting!

De door Springer gepubliceerde boeken in bovenstaande lijst kunnen door individuele leden met een korting van 25% op de officiële prijs² aangeschaft worden via de Centrale Administratie van de VVS. Bestelformulieren zijn te verkrijgen via Internet, URL <http://www.vvs-or.nl/>, knop *Ledenservice*.

¹Voor VVS-leden fl. 35.00, te bestellen via j.hox@fss.uu.nl onder vermelding van uw postadres

²De prijzen zoals vermeld in de lijst van binnengekomen boeken kunnen foutief zijn. Een betrouwbare bron van informatie is de Springer website <http://www.springer.de/>.