

Boekbesprekingen

Redactie:

Arend D. Oosterhoorn
Van Doorne's Transmissie bv
Postbus 500
5000 AH Tilburg

telefoon 013 - 640319

K.O. Bowman en L.R. Shenton

Properties of estimators for the gamma distribution

Marcel Dekker, New York, 1988, xvi + 268 pag., ISBN 0-8247-7556-2, \$ 78.00.

Het materiaal van dit boek is gepresenteerd in vijf hoofdstukken en drie appendices. Het eerste hoofdstuk (Basic formulas, 44 pag.) geeft een overzicht van eigenschappen van de gamma verdeling; zowel theoretische als numerieke aspecten komen aan bod. Het tweede hoofdstuk (Maximum likelihood estimators for the two parameter gamma distribution, 55 pag.) geeft methoden voor een nauwkeurige berekening van momenten (en benadering van de verdeling) van de ML-schatters; het derde hoofdstuk (Approximate distribution of percentage points, 27 pag.) richt zich op het schatten van de verdeling van de percentielen. In het vierde hoofdstuk (Simulation and estimation problems associated with the 3-parameter gamma density, 33 pag.) wordt het probleem van het tweede hoofdstuk uitgebreid door de introductie van de derde (drempel-)parameter. De benadering behandelt in het vijfde hoofdstuk (The distributions of the standard deviation, skewness and kurtosis in random samples from a gamma distribution, 41 pag.) hebben betrekking op de gamma verdeling met twee parameters. De appendices zijn

- a) Series into sums associated with the exponential (19 pag.)
- b) Alternative summation algorithms: Levin and modified Borel (10 pag.)
- c) Miscellaneous formulas (20 pag.)

De gepresenteerde resultaten worden toegelicht met voorbeelden en er zijn ook opgaven. Ondanks dit lijkt het boek te zware kost voor een doorsnee statisticus: het boek is duidelijk bedoeld voor specialisten op het gebied van numerieke aspecten van de gamma verdeling.

V. Fidler
RU Groningen

Damaraju Raghavarao

Exploring Statistics

Marcel Dekker Inc., New York, 1988, ix + 276 pag., ISBN 0-8247-7952-5, \$ 47.50.

Het algemene beeld:

Het boek is bedoeld voor beroepsbeoefenaars, die zich door middel van zelfstudie een aantal veel gebruikte onderzoekstechnieken eigen willen maken. Daarnaast is het gericht op studenten, die een eerste elementaire cursus (toegepaste) statistiek moeten volgen. Om het verschil in doelgroepen te overbruggen zijn de meeste hoofdstukken voorzien van oefenopgaven en een appendix, waar de mathematische achtergrond van de onderzoekstechniek wordt behandeld.

In de tekst staan veel voorbeelden, die uit een groot aantal disciplines gekozen zijn, waardoor de toegankelijkheid van de stof voor de met statistiek onbekende beroepsbeoefenaar vergroot wordt. Ook de afwezigheid van wiskundige formules en afleidingen in de hoofdtekst doet dit. Een bezwaar is, dat hierdoor een receptmatige aanpak ontstaat, die nog versterkt wordt door het uitsluitend weergeven van de benodigde instructies en uitkomsten van het softwarepakket Minitab.

De basis van het boek zijn de onderzoekstechnieken regressieanalyse, variantieanalyse en het gebruik van de chi-kwadraat toets. De hiervoor benodigde voorkennis, zoals het trekken van steekproeven, het verwerken van data en het toetsen van hypothesen vormt de inleiding.

De inhoud per hoofdstuk:

In hoofdstuk 1 een beschouwing over 'Wat zijn data?', gevolgd door in hoofdstuk 2 een overzicht van de mogelijke manieren om steekproeven te trekken uit populaties. Het geeft de doe-het-zelver een brede kijk op de mogelijkheden, maar laat een aantal essentiële zaken, zoals het bepalen van de steekproefomvang en hoe stratificeren, onbesproken. In hoofdstuk 3 komen manieren om data te verwerken aan de orde. Naast de gebruikelijke technieken, zoals histogrammen, worden ook recente wijzen besproken, zoals 'stem-and-leaf display'. Tevens vindt een bespreking plaats van de bekende maten van een verdeling. In het appendix wordt ingegaan op de wiskundige achtergrond van een aantal besproken maten en worden deze voorzien van rekenvoorbeelden. In hoofdstuk 4 volgt een duidelijke uiteenzetting over de begrippen kans en schatter, gevold door het toetsen van hypothesen. In het appendix een goede en noodzakelijke wiskundige aanvulling op het begrip kans, de veel gebruikte kansdichtheden en wat misplaatst het bepalen van de steekproefomvang. In de hoofdstukken 5, 6 en 7 het eigenlijke werk: onderzoekstechnieken voor de samenhang tussen variabelen. Receptmatig, dat wil zeggen het aanreiken van een aantal voorbeelden uit veel disciplines, verduidelijken het toepassingsterrein, de uitkomst van Minitab laten de resultaten zien van de gebruikte techniek en tot slot een interpretatie van de uitkomsten, worden behandeld:

- regressie-analyse, enkelvoudig en meervoudig
- variantie-analyse, inclusief een korte uiteenzetting van 'randomised block design'
- het gebruik van de chi-kwadraat toets.

De beschrijving in ieder hoofdstuk van de technieken is duidelijk, er ontbreekt echte een heldere formulering van de voorwaarden van de verschillende methoden. Ook hier weer de noodzakelijke mathematische achtergronden in de appendices.

Konklusie:

Voor de doe-het-zelver is het boek slechts ten dele bruikbaar, of hij moet in staat zijn zich door de appendices te worstelen. Beperkt hij zicht tot de hoofdtekst, dan is hij na lezing in staat aan te geven wanneer welke techniek gebruikt kan worden en weet hij globaal wat er gedaan moet worden om de gewenste resultaten te verkrijgen. Daar hij het 'hoe' van de techniek en berekeningswijze niet of in onvoldoende mate gezien heeft, is hij op vakmensen aangewezen voor de uitvoering.

Zeër uiteenlopende hbo opleidingen, die enerzijds met een beperkt aantal beschikbare uren zitten en anderzijds de noodzaak inzien studenten uit te rusten met een aantal veelgebruikte onderzoekstechnieken kunnen dit boek goed gebruiken: de noodzakelijke voorkennis wordt precies in voldoende mate aangebracht, zodat zoveel mogelijk tijd besteed kan worden aan het aanleren van de gewenste beroepsbekwaamheden.

Het blijft echter in het engels geschreven.

T. Severijn
Hogeschool Midden-Brabant
Tilburg

G.E.P. Box & N.R. Draper

Empirical model-building and response surfaces

John Wiley & Sons Ltd., Chichester, 1987, xiv + 669 pag., ISBN 0-471-81033-9, £ 43.25.

Dit mooi uitgevoerde, lijvige boek beschrijft de overwegingen en methoden voor het opzetten, uitvoeren en analyseren van (de resultaten van) experimenten met betrekking tot responsie oppervlakken.

Box en Draper (en in eerste instantie ook Stu Hunter) hebben in een goed leesbaar verhaal hun ervaring opgeschreven, welke zij hebben opgedaan in hun jarenlange werk op het gebied van de statistische consultatie. Het boek bevat niet veel technische verhalen en bewijzen, wel veel praktische overwegingen en datasets, aan de hand waarvan de theorie wordt behandeld of verder wordt toegelicht. Voor het lezen van recente ontwikkelingen op het gebied van de respons oppervlak technieken is dit boek niet het juiste, literatuurverwijzingen in de tekst zijn over het algemeen oud (ver) voor 1980) en zijn vaak artikelen geschreven door de auteurs zelf.

In hoofdstuk 1 wordt een algemene inleiding gegeven, met een opsomming van de (noodzakelijke) overdenkingen voordat een experiment wordt uitgevoerd. Keuze van de variabelen, de niveau's waarop deze worden ingesteld en de keuze van het ontwerp komen daar aan bod.

Ook staat daarin het plaatje van het iteratieve aspect van het leren door experimenteren, zoals dat ook in het boek van Box, Hunter en Hunter (1978) staat.

Hoofdstuk 2 geeft een motivering voor het benaderen van een willekeurige functie door een eerste of een tweede graads polynoom. In hoofdstuk 3 wordt de kleinste kwadraten methode behandeld. De methode wordt beschreven aan de hand van matrix algebra, zodat enige voorkennis op dat gebied verondersteld mag worden. In een appendix wordt wel enige aandacht besteed aan matrix algebra (rekenregels en begrippen als inverse, getransponeerde en determinant, in totaal een kleine acht pagina's).

In dit hoofdstuk wordt ook gekeken naar de gevolgen voor de parameterschattingen indien het 'werkelijke' model een tweede graads polynoom is en het geanalyseerde model een eerste graads polynoom.

Daarvoor wordt de alias-matrix beschreven. Deze aanpak komt veel voor in het boek. De auteurs staan vaak stil bij het aspect van onzuiverheid.

Factoriële proefopzetten voor twee niveau's, het vormen van blokken en construeren van fractionele opzetten wordt behandeld in de hoofdstukken 4 en 5. Het hier behandelde staat grotendeels ook al in het eerder genoemde boek.

Hoofdstuk 6 beschrijft de theorie van de steilste helling voor het bereiken van een proces verbetering waarbij ook een betrouwbaarheids uitspraak wordt gedaan over de gevonden richting.

Het aanpassen van een tweede orde model wordt in hoofdstuk 7 behandeld aan de hand van een voorbeeld, waarbij ook aandacht wordt besteed aan het gevolg van transformeren van de respons variabele. De analyse wordt uitgevoerd met de oorspronkelijke en met de getransformeerde (logarithme) data.

In hoofdstuk 8 wordt doorgegaan op het probleem van het transformeren van de waarnemingen. Hier komen transformaties van zowel de respons- als de predictor variabele(n) aan de orde met overwegingen van reden en gevolg (veranderen van de foutstructuur). De Box-Cox transformatie wordt hier behandeld.

Hoofdstukken 9, 10 en 11 zijn geheel gewijd aan het onderzoek van ridge systemen en de kanonieke presentatie van het model. Als voorbeeld worden de resultaten van een Central Composite Design gebruikt, welk ontwerp pas in hoofdstuk 15 nader wordt uitgelegd.

Er worden twee kanonieke reducties aangegeven, de A kanonieke reductie (alleen roteren van het assenstelsel) en de B kanonieke reductie (roteren van het assenstelsel en verschuiven van de oorsprong naar het stationaire punt).

De connectie tussen empirische en theoretische modellen wordt in hoofdstuk 12 aangegeven aan de hand van een voorbeeld, afkomstig uit de chemie. Vanuit zuiver fysische overwegingen wordt een model opgebouwd welke de relatie aangeeft tussen de respons variabele en de predictor variabelen.

Hoofdstuk 13 behandelt overwegingen van variantie, bias en lack of fit met betrekking tot de keuze van het proefontwerp. Heel duidelijk wordt beschreven hoe onzuiverheid en variantie elkaar in de haren kunnen zitten. Tevens wordt het concept van de 'integrated mean squared error' in dit hoofdstuk behandeld. Er wordt aan de hand van enkele voorbeelden gekeken hoe het ontwerp eruit moet zien indien minimale onzuiverheid wordt verlangd. Daarbij wordt uitgegaan van een situatie dat de werkelijke relatie exact gerepresenteerd wordt door een polynoom welk één graad hoger is dan die van het model. Dit hoofdstuk is het meest technische in wiskundige zin.

Hoofdstuk 14 begint met de opsomming van 14 punten waaraan een proefontwerp voor een respons oppervlak volgens de auteurs moet voldoen, zoals de mogelijkheid voor het testen van lack of fit, het kunnen schatten van de experimentele fout en het bevatten van een minimaal aantal punten.

Het hoofdstuk is geheel gewijd aan variantie-optimale ontwerpen, waarbij (dus) ook de de A, E, D en G optimaliteit wordt behandeld. Aan het eind staat nogmaals een opmerking over het verwaarlozen van de bias bij de keuze van het ontwerp, waar de auteurs toch wel vaak aandacht aan besteden en voor waarschuwen om het niet uit het oog te verliezen.

Het boek wordt besloten met hoofdstuk 15, waarin praktische overwegingen voor de keuze van het ontwerp voor respons oppervlakken worden beschreven als de mogelijkheid voor het toetsen op lack of fit en het uitvoeren van het experiment in blokken. Hier komen uitgebreid de Central Composite Designs aan de orde met de overwegingen voor de keuze van de sterpunten en het aantal waarnemingen in het centrum. Afsluitend worden een aantal tweede orde ontwerpen aangegeven waarbij slechts drie niveau's nodig zijn (Box-Behnken) en ontwerpen met een gering aantal waarnemingen (Hartley, Westlake, Plackett en Burman, Koshal).

Een zeer groot pluspunt van het boek is de grote hoeveelheid dataset in de opgaven, welke vrijwel allemaal zijn gekozen uit (oude) publicaties, dus 'real life' problemen beschrijven. De onderwerpen komen veelal uit de chemie of aanverwante gebieden.

De datasets nodigen uit om de aangeboden stof toe te passen en daarmee wordt de effectiviteit van het lezen verhoogd. In totaal staan er 124 opgaven in het boek, waarvan de beknopte antwoorden achterin zijn opgenomen, waarbij een kleine 40-tal datasets zijn beschreven.

Ook de behandelde theorie wordt geheel verduidelijkt door het gebruik van resultaten van experimenten.

Aan literatuurverwijzingen is geen gebrek, meer dan 900 referenties zijn vermeld, waarbij ook recente (tot 1985). Bij een groot deel van deze referenties is de naam van de auteurs terug te vinden.

Al met al een zeer waardevol boek, goed geschreven, keurig verzorgd en vol met praktische tips.

referenties

Box, G.E.P., W.G. Hunter & J.S. Hunter, (1978) *Statistics for Experimenters*, John Wiley & Sons, New York

Arend Oosterhoorn

Van Doorne's Transmissie bv

B.D. Ripley

Statistical inference for spatial processes

Cambridge University Press, Cambridge, 1988, vi + 148 pag., ISBN 0-521-35234-7, £ 20.00.

Dit boek is de bekroonde inzending van een door de Cambridge University uitgeschreven prijsvraag met als onderwerp 'Spatial and geometrical aspects of probability and statistics'. In de inleiding worden de problemen opgesomd die optreden bij de analyse van ruimtelijke processen. De omvang van de problemen en het specifieke karakter ten opzicht van klassieke- en tijdreeksmethoden is in veel gevallen pas recent bekend geworden. Het zijn deze specifieke aspecten die in de overige hoofdstukken worden uitgewerkt. Het is raadzaam Ripley's 'Spatial statistics' (Wiley, 1981) bij de hand te hebben als algemeen overzichtswerk.

Hoofdstuk 2 behandelt 'Likelihood analysis for spatial Gaussian processes'. De MLE van parameters van kovariantiefuncties blijkt niet zondermeer optimaal doordat de likelihood functies lokale maxima vertonen. Numerieke maximalisatie-algorithmen die unimodaliteit veronderstellen kunnen dus niet worden gebruikt.

Hoofdstuk 3 bevat nieuwe en verbeterde procedures voor 'Edge corrections for spatial point processes', uitgesplitst naar inter-punt, nearest-neighbour en empty-space statistics.

Een belangrijke aanvulling op 'Spatial statistics' vormt hoofdstuk 4: 'Parameter estimation for Gibbsian point processes'. Gibbs processen spelen een rol bij het modelleren van interacties en context. Paarsgewijze interactie tussen punten wordt beschreven met een potentiaal functie. Ook hier blijkt MLE van de parameters niet optimaal. Monte Carlo methoden zijn een goed alternatief. Als speciaal geval wordt het Strauss-proces uitgewerkt.

De hoofdstukken 5 en 6 concentreren zich op de analyse van digitale beelden.

Hoofdstuk 5 ('Modelling spatial images') geeft een overzicht van statistische procedures voor beeldverbetering, segmentatie e.d. Verbetering van een door 'verontreiniging' van een onbekend beeld S ontstaan beeld Z loopt via Bayes regel $P(S|Z) = P(Z|S)P(S)$. Modellen voor S en procedures voor MAP-schattingen ('maximum a posteriori') van $P(S|Z)$ zoals simulated annealing en Besag's ICM-algoritme worden uitgewerkt. Alternatieve loss-functies en kausale Markov-random-velden worden kort besproken.

Hoofdstuk 6 'Summarizing binary images' geeft een helder overzicht van mathematisch morfologische operaties voor gebruik in binaire beelden. Voor erosie, dilatie, opening en sluiting worden probabilistische definities opgesteld, voorzien van correcties voor randeffecten. Deze operaties als functie van de straal van een schijfvormig structurerend element blijken verwante patronen goed van elkaar te onderscheiden in tegenstelling tot de gebruikelijke tweede orde statistics. Er wordt kort ingegaan op operaties in grijswaardenbeelden.

Afgezien van fundamentele resultaten worden in dit boek concrete aanwijzingen gegeven met betrekking tot numerieke procedures en implementatie aspecten. Een bijzonder waardevol boek dat bange gevoelens rechtvaardigt over wat ons nog te wachten staat op het gebied van ruimtelijke statistiek.

R.J. Bijlsma
Rijksinstituut voor Natuurbeheer

C. Chatfield

Problem solving, a statisticians guide

Chapman and Hall, Londen, xvi + 261 pag. ISBN 0-412-28670-X (gebonden), ISBN 0-412-28680-7 (paperback), £ 14.95 (paperback)

Dit boek is niet geschreven om de statistische methoden te leren. Het is bedoeld om aan te geven hoe je die technieken, waarover al zoveel boeken zijn geschreven, in de praktijk toe te passen met alle problemen van dien.

Het is geschreven voor een lezer met een behoorlijke basisopleiding in de statistiek, maar zich nog onzeker voelt in het aanpakken van deze 'real-file' problemen. Nu weet uw resercent niet precies welke groep statistici daar dan mee wordt bedoeld, maar uit gesprekken met collega's zou je op kunnen maken dat die groep bestaat uit alle statistici die in de 'praktijk' werkzaam zijn en aan die groep kan ik het boek ook van harte aanbevelen.

Het boek bestaat uit drie duidelijk te scheiden delen. In deel I worden algemene principes beschreven voor het aanpakken van (statistische) problemen, daarover direkt meer.

In het tweede deel worden vele oefeningen gegeven. Deze zijn in de vorm van het beschrijven een probleem, met de vraag 'hoe zou je dat aanpakken'. Nadat een aantal problemen op deze wijze zijn geformuleerd, wordt door de schrijver zelf een oplossing aangedragen en wordt enigszins uitgewijd over mogelijke alternatieven of mogelijk optredende valkuilen. Soms wordt een dergelijke behandeling afgesloten met een moraal ('don't always do what you are asked to do', bij een probleem waarbij de onderzoeker bij de statisticus komt met de opmerking 'doe even een ANOVA', maar waar de auteur meerdere mogelijkheden behandelt (geen ANOVA)).

De onderwerpen van de problemen variëren nogal. Beschrijvende statistiek (interpreteren van grafieken uit de media over misdaad aantallen) regressie en correlatie, grote datasets, resultaten van proefopzetten (Latijns vierkant, ongebalanceerd twee-weg schema), tijdreeksanalyse, gebruik van kansrekening in de rechtszaal, (Califonia 1968, people versus Collins) verzamelen van waarnemingen (moraal: 'be on the look-out for biased samples').

In deel III (in het boek appendix A) wordt in 51 pagina's een overzicht gegeven van een vijftiental onderwerpen van de statistiek. Beschrijvende statistiek, kansregels, kansverdelingen, schatten, toetsen, regressie analyse, variantie analyse, lineaire model (matrix notatie), gegeneraliseerd lineair model, steekproeven, proefopzetten, klinische proeven, multivariatie analyse (PCA en FA), tijdreeksen en kwaliteitskontrolle en bedrijfszekerheid passeren op een verbazingwekkend vlot leesbare wijze de revue.

Na het lezen van deze pagina's had ik echt het gevoel van 'is het echt zo eenvoudig'. Uiteraard worden er niet meer dan de basis ideeën behandeld, maar toch. Het is uitermate knap en overzichtelijk geschreven.

Dan nu nog het eerste deel. Dit deel bestaat uit twaalf hoofdstukken, waarvan de meeste slechts een aantal pagina's lang zijn, en een samenvatting, getiteld 'how to be an effective statistician'. De ingrediënten voor het beantwoorden van deze vraag staan in die twaalf hoofdstukken voorzover het niet het statistisch-technisch werk betreft.

De auteur behandelt in die hoofdstukken de eigenschappen welke voor het goed uitvoeren van de taak onmisbaar zijn en dus ook onderwezen zouden moeten worden. Ieder hoofdstuk wordt afgesloten met aanbevolen literatuur voor verdere informatie. De tekst is ook gebaseerd op de colleges die door de auteur zijn verzorgd voor 'final year statistics undergraduates' (hebben we daar een nederlands equivalent voor?).

Na een algemene inleiding worden in hoofdstuk 2 een aantal stappen geformuleerd waaruit, vanuit de statisticus gereedeneerd, een onderzoek is opgebouwd. Om een indruk te geven: 'wees zeker dat je weet waar het over gaat', 'verzamel de waarnemingen', 'beoordeel de kwaliteit van de waarnemingen', 'voer een eerste analyse uit', 'interpreteer en communiceer over de resultaten'.

Hoofdstuk 3 gaat over het formuleren van het probleem (stel veel vragen en luister goed) en hoofdstuk 4 over het verzamelen van gegevens (een van de belangrijkste taken binnen de statistiek).

De drie volgende hoofdstukken handelen over het analyseren van waarnemingen. Na in hoofdstuk 5 een algemene strategie uitgezet te hebben voor het analyseren van waarnemingen breekt de auteur in hoofdstuk 6 een lans voor wat hij noemt 'initial data analysis' (IDA). Hiermee wordt bedoeld het berekenen van beschrijvende maten, het maken van plaatjes e.d. met als doel de algemene structuur van de data te weerspiegelen, om gevoel te krijgen voor de data. Onderwerpen als foutieve data en uitschieters (en het verschil daartussen), beschrijvende maten, grafieken, tabellen en transformaties worden besproken.

Hij onderscheidt IDA en EDA (explorative data analyse, Tukey 1977) met het argument dat IDA bedoeld is voor *data beschrijving* en het een aanzet voor *model formulering* (cursief van auteur). Naar de mening van de auteur heeft Tukey een bijdrage geleverd aan de data analyse methodieken, maar heeft hij nagelaten de gereedschappen onder te brengen in het algemeen gebruik. De discussie daarover is gevoerd in Chatfield (1985).

In het laatste hoofdstuk over data analyse, 'the 'definitive' analysis', met als ondertitel 'I know - how to do a t-test, but not when', worden de formelere methoden behandeld en van aanbevelingen voorzien.

Hierna volgen twee hoofdstukken over het gebruik van hulpmiddelen, de computer (hoofdstuk 8) en de bibliotheek (hoofdstuk 9). Deze hoofdstukken bevatten vrij algemene opmerkingen en een opsomming van enkele statistische pakketten en tijdschriften. In de appendix B wordt nog een iets gedetailleerdere bespreking gewijd aan MINITAB en GLIM.

Hoofdstuk 10 en 11 gaan over communicatie. Het eerste over consultatie en participatie in projecten. Hij geeft adviezen welke volgens hem algemeen bekend maar toch vaak vergeten zijn ('know the 'customer' and his needs', 'be prepared to ask probing questions', 'try to get involved at the planning stage').

Het andere hoofdstuk behandelt het schrijven van een rapport.

De laatste boodschap staat in hoofdstuk 12 'Numeracy'. Naar een oud chinees spreekwoord: 'Velen kijken, doch weinigen zien', moet een statisticus de eigenschap ontwikkelen om patronen en wetmatigheden te zien in data door ernaar te kijken.

Een statisticus moet kritisch staan ten opzicht van andermans uitspraken, zoals 'the possibility of a nuclear meltdown is only one in 10000 years (remember Chernobyl!)'.

Afgesloten wordt met de reeds genoemde samenvatting.

Echt aan te bevelen, niet alleen aan niet zo ervaren beroepsbeoefenaren maar ook aan de ervaren statistici. Al is het alleen maar om jezelf na het lezen een schouderklopje te geven omdat je het, volgens de Chatfield normen althans, nog niet zo slecht doet.

Een compliment aan de schrijver voor het opschrijven van zo veel in zo weinig woorden.

referenties

Chatfield, C., (1985) The initial examination of data (with discussion). *J. R. Stat. Soc. A*, **148**, 214-53

Tukey, J.W. (1977) *Exploratory Data Analysis*, Addison-Wesley, Reading, Mass.

Arend Oosterhoorn

Van Doorne's Transmissie bv

Search Problems

John Wiley & Sons, Chicester, 1987, 284 pag., ISBN 0-471-90825-8, f 29.50.

Dit boek geeft een overzicht van de vele resultaten die de afgelopen jaren behaald zijn met betrekking tot het oplossen van zoekproblemen, en die, zoals de auteurs zelf opmerken, behaald zijn door onderzoekers uit zeer uiteenlopende disciplines, die vaak niet van elkaars werk op de hoogte lijken te zijn.

Na een kort inleidend hoofdstuk illustreren de auteurs het zoekmodel, dat centraal staat in het eerste deel van het boek, aan de hand van het volgende voorbeeld:

Stel, men test door middel van een bloedonderzoek de leden van een populatie op het al of niet aanwezig zijn van een bepaalde ziekte.

Wanneer de mogelijkheid bestaat het bloed van meerdere mensen tegelijkertijd te onderzoeken, bestaat er wellicht een meer efficiënte wijze van testen dan het afzonderlijk testen van ieder lid van de populatie. Immers, als een 'gezamenlijke test' een negatief resultaat geeft, staat vast dat iemand van de onderzochte personen besmet is.

In het algemeen hebben we te maken met een eindige verzameling N (de zoekruimte) en zijn we op zoek naar een element i van N met een bepaalde eigenschap. Om i te vinden, beschikken we over de mogelijkheid alle elementen van N aan een test te onderwerpen. Onder een (non)sequentiele strategie t verstaat men een reeks testen $t_1, \dots, t_m$, waarbij het uitvoeren van test t_k (geen) gebruik wordt gemaakt van de uitkomsten van de vorige testen. Onder de zoektijd $l_t(i)$ van t voor i verstaat men het aantal tests uit de reeks t , dat men moet uitvoeren alvorens te kunnen concluderen dat men i gevonden heeft. De zoektijd van een strategie t kan nu worden gedefinieerd als het maximum van $l_t(i)$ over alle i in N ; een optimale strategie als een strategie met minimale zoektijd.

In hoofdstuk 3 worden zoekprocessen behandeld waarbij de beschikbare tests foutvrij zijn en slechts ja-nee uitkomsten hebben (binaire tests). Er worden worst-case resultaten afgeleid voor zowel sequentiële als non-sequentiële strategieën. Tevens wordt de situatie behandeld waarin een a priori kansverdeling is gegeven voor de elementen van de zoekruimte. Diverse resultaten blijken nauw verwant te zijn met bekende resultaten uit de informatie theorie.

Hoofdstuk 4 behandelt op analoge wijze als hoofdstuk 3 de situatie waarin er een totale ordening van zoekruimte is gegeven en alleen tests zijn toegestaan, die deze ordening respecteren.

Sorteer- en weegproblemen komen aan de orde in de hoofdstukken 5 en 6. Het gaat daarbij om problemen als het sorteren van een verzameling elementen op een bepaald kenmerk, het samenvoegen van meerdere, reeds gesorteerde lijsten, het door middel van wegen met behulp van een balans bepalen van één valse munt in een verzameling munten, etc.

Tenslotte worden in hoofdstuk 7 enkele speciale zoekproblemen behandeld, die niet passen binnen de in de voorafgaande hoofdstukken gehanteerde modellen.

In het tweede deel van het boek worden zoekproblemen behandeld waarbij de tests niet langer foutvrij zijn. In hoofdstuk 8 wordt de theorie die voor deze problemen is ontwikkeld geïllustreerd aan de hand van methoden voor nulpuntsbepaling van één-dimensionale functies met behulp van metingen van funktiewaarden die aan een bepaalde (stochastische) ruis onderhevig zijn (de methode van Robbins & Monro).

In hoofdstuk 9 worden de eerder gevonden resultaten uitgebreid voor de situatie van niet-binaire tests.

In het laatste deel van het boek worden zoekproblemen behandeld die met de mogelijkheid van inspectie kunnen worden opgelost. Een voorbeeld van dit soort problemen is het volgende probleem:

Een persoon is op zoek naar zijn agenda, die in één van de n kamers van zijn huis aanwezig is. Hij moet een beslissing nemen ten aanzien van de volgorde waarin hij de kamers zal doorzoeken. Het is mogelijk dat hij de kamer inspekt waar de agenda zich bevindt, zonder deze te vinden. Als bepaalde gegevens, zoals de kans dat de agenda zich in kamer k bevindt ($k = 1, \dots, n$), de kosten van de j -de inspectie van kamer k en de kansen $q(j,k)$ dat de agenda tijdens de j -de inspectie van kamer k wordt gevonden, bekend zijn, kan een strategie worden afgeleid met minimale verwachte zoekkosten.

Dit vormt het onderwerp van hoofdstuk 11. In hoofdstuk 12 wordt de situatie behandeld waarin er een bovengrens aan het aantal inspecties is en men een strategie zoekt waarvoor de kans op succes maximaal is.

In hoofdstuk 13 worden nog enkele uitbreidingen van de in hoofdstukken 11 en 12 afgeleide resultaten besproken.

Het onderhavige boek is de Engelse vertaling van het uit 1979 daterende boek van Ahlswede en Wegener, getiteld 'Suchprobleme' (Teubner, Stuttgart). In vergelijking met de oorspronkelijke editie is de Engelstalige versie uitgebreid met een hoofdstuk 'Further Reading', waarin artikelen en boeken worden genoemd, die van na 1979 dateren.

Het boek is helder geschreven en vereist geen uitgebreide achtergrondkennis van de lezer. De auteurs pogen met dit boek de uitgebreide literatuur op het gebied van zoekproblemen op een zodanige manier samen te vatten, dat de lezer een vlug begrip en overzicht krijgt van het soort vragen en problemen die in dit gebied een rol spelen. Naar mijn mening zijn ze daarin uitstekend geslaagd.

P.J.M. van Laarhoven
Nederlandse Philips Bedrijven bv

Brian D. Ripley

Stochastic Simulation

John Wiley & Sons, New York, 1987, xi + 237 pag., ISBN 0-471-81884-4, \$ 32.95.

Simulatie is een veel toegepaste techniek, zowel bij het oplossen van praktische problemen als bij het bestuderen van theoretische problemen waarvoor geen analytische oplossing bekend is.

Het boek van Ripley is bedoeld als veelomvattende gids voor de huidige stand van zaken op simulatie gebied. Het is geschreven voor statistici, onderzoekers op het gebied van de operations research en diegenen die simulatie als hulpmiddel gebruiken bij het analyseren en oplossen van hun problemen.

De lezer heeft volgens de inleiding slechts een elementaire kennis van wiskunde en statistiek nodig om het boek te kunnen volgen. Dit is echter zeker niet het geval. Het boek vraagt een behoorlijke scholing in de mathematische statistiek.

Het boek bestaat in totaal zeven hoofdstukken. Het eerste hoofdstuk gaat zeer kort en globaal in op de mogelijkheden van simulatie.

De volgende twee hoofdstukken zijn gewijd aan het genereren van pseudo-toevalsgetallen en stochastische variabelen. Een hoewel er een aantal mooie grafische weergaven zijn van wat er allemaal fout kan gaan bij het genereren van toevalsvariabelen, is de tekst voor de geïnteresseerde gebruiker of student van simulatie te ingewikkeld.

Daar staat tegenover, dat de tekst voor de vakman weer te oppervlakkig is. Dit maakt duidelijk wat het probleem is bij dit boek. Het is enerzijds geen inzichtelijk studieboek voor een nadere kennismaking met simulatie en anderzijds ook geen naslagwerk voor gebruik bij het opzetten en

uitvoeren van simulatie-experimenten.

In hoofdstuk 4 worden processen en multivariate verdelingen behandeld.

Hoofdstuk 5 behandelt variantiereducerende technieken, maar ook hier blijft de kracht van wat er met deze technieken kan worden bereikt, verborgen onder mathematische formulering.

Hoofdstuk 6 is gewijd aan de analyse van simulatieresultaten.

Hoofdstuk 7 zou een overzicht van een aantal toepassingen moeten zijn. Ook hier is weer dezelfde kritiek van toepassing als op de voorgaande hoofdstukken.

Een voordeel van het boek is de uitgebreide verwijzing naar de literatuur op het gebied van simulatie. Samenvattend mag worden gekonkludeerd dat het boek ondanks een paar pluspunten niet moet worden aanbevolen.

Voor diegenen die simulatie willen toepassen, is het onbruikbaar, niet alleen door de gebruikte formuleringen, maar ook door het bijna geheel ontbreken van relevante toepassingen en de daarmee samenhangende problemen.

Voor academici die onderzoek doen op het gebied van simulatie is het boek te breed georiënteerd en heeft het te weinig diepgang voor elk van de behandelde onderdelen.

Als inleidend studieboek op het gebied van simulatie is het niet geschikt omdat het een te groot beroep doet op de wiskundige en statistische kennis van de student.

W.J.H. van Groenendaal

Faculteit der Economische Wetenschappen KUB

S.K. Srinivasan

Point Process Models of Cavity Radiation and Detection

Charles Griffin & Co, Londen, 1988, ii + 179 pag., ISBN 0-85264-291-1, £ 19.75.

In dit boek behandelt de schrijver stochastische modellen in quantum optica, in het bijzonder puntproces modellen.

Het gaat hier om modellen voor de evolutie van populaties fotonen, zoals die bijvoorbeeld optreden in de beschrijving van de fysische verschijnselen bij laserstraling.

De eerste vier hoofdstukken van het boek gaan over de voor deze toepassing relevante stochastische processen: Markov processen en puntprocessen.

Voor een groot deel betreft het hier een herpublicatie van gedeelten uit eerder verschenen boeken van de auteur zelf. Helaas wordt met name het hoofdstuk over puntprocessen ontsierd door een aantal hinderlijke drukfouten, waardoor bijvoorbeeld paragraaf 3.2 over moment maten zeer moeilijk te begrijpen wordt.

Hoofdstuk 5 behandelt de fysische achtergrond en vereist een behoorlijke kennis van quantumfysica. Volgens de auteur zijn de laatste hoofdstukken over toepassingen van stochastische populatieprocessen in de Optica ook toegankelijk voor lezers die de noodzakelijke fysica achtergrond niet bezitten.

J.A.M. v.d. Weide

TU Delft

J.W. Srivastava (ed.)

Probability and statistics, essay in honor of Franklin A. Graybill

North-Holland, Amsterdam, 1988, xiv + 332 pag., ISBN 0-444-70547-4, f 195,=.

De redakteur heeft een aantal artikelen gebundeld ter ere van Frank Graybill. Een grote variëteit in de onderwerpen van de bijeengebrachte bijdragen reflecteert het brede scala van het werk van Graybill.

De artikelen zijn allemaal gerefereerd.

Na een kort overzicht van het leven (anderhalve pagina) en de publikaties van Graybill volgen de volgende 22 artikelen.

Admissibility in Linear Models with Polyhedral Covariance Structure

A.-M. Azzam, D. Birkes and J. Seely

Simulated Power Comparisons of the Asymptotic and Non-asymptotic Goodman and Kruskal Tau Tests for Sparse R by C Tables

K.J. Berry and P.W. Mielke Jr.

Schemes Exhibiting Hurst Behaviour

D.C. Boes

The Extension of Grubb's Estimators for Classes of Incomplete Block Designs

R.A. Bradley and D.A. Rollier

Applications of Innovation Representations in Time Series Analysis

P.J. Brockwell and R.A. Davis

Karl Pearson at Gresham College

J.L. Folks

On an Intermittent Stochastic Process

J. Gani and P. Todorovic

On the Unification of Bias Reduction and Numerical Approximation

H.L. Gray

Invariant Inference for Variance Components

D.A. Harville

An Argument for Adaptive Robust Estimation

R.V. Hogg, G.K. Bril, S.M. Han and L. Yuh

Distribution Free Procedures for Some Ratio Problems

H.K. Iyer and L. Kaiser

Prediction Intervals in balance One-Factor Random Model

S. Jeyaratnam and S. Panchapakesan

Goodness-of-Fit and g-Sample Omnibus Tests Based on Coverages

P.W. Mielke, Jr.

Statistical Sampling in Auditing: A Review

J. Neter and J. Godfrey

Sums of Independent Random Variables with the Extreme Terms Excluded

W.E. Pruitt

Uniform Rates of Convergence to Extreme Value Distributions

S.I. Resnick

Best Linear Unbiased Estimation in Mixed Models of the Analysis of Variance.

S.R. Searle

High Resolution Frequency Analysis of Almost Periodic Process

M.M. Siddiqui and C.C. Wang

On a General Theory of Sampling Using Experimental design Concepts II. relation with Arrays

J.N. Srivastava

β -Expectation Tolerance Limits for Balanced One-Way Random-Effects Model

C.M. Wang

Linear Calibration When the Coefficient of Variation is Constant

Y.C. Yao, D.F. Vecchia and H.K. Iyer

Generalization for Second-Order and Multivariate Normal Linear Models

J.S. Williams and M.K. Zaatar

T.P. Hutchinson

Road accident statistics

Rumsby Scientific Publishing, Adelaide, 1987, xxvi + 202 pag., ISBN 0-7316-0653-1, £ 28.00.

In dit boek worden statistieken van verkeersongevallen van de gehele wereld besproken. Zowel de collectie van de data als de manier van bewerken komen aan bod. Het is niet zo dat bij alle vergelijking alle landen betrokken worden, maar de meeste landen komen toch wel in het boek voor. Het bevat 175 tabellen en 18 plaatjes, waarbij de laatste voornamelijk afdrucken zijn van formulieren welke in de diverse landen worden gebruikt voor registratie van de ongevallen. Vele bevatten daarbij allerlei verschillende situatieschetsen om de ongevallen te kunnen classificeren.

Na een inleiding wordt in hoofdstuk 2 een overzicht gegeven van het aantal ongelukken met dodelijke afloop in 144 landen. Daarbij zijn de jaartallen per land nogal verschillend. Soms gaat het over de jaren 1976 en 1977 (Zaïre) en soms over 1984 en 1985 (Luxemburg). In het volgende hoofdstuk wordt de manier van verzamelen van de waarnemingen besproken welke via de politie beschikbaar komen. In dit hoofdstuk staan nogal wat van de genoemde formulieren afgedrukt. In het volgende hoofdstuk wordt besproken wat de politie nogal eens vergeet te melden bij deze registratie en hoofdstuk 5 gaat over de datacollectie van verkeersongevallen in de derde wereld.

In hoofdstuk 6 worden zes jaarboeken vergeleken (Finland, Groot-Brittannië, Japan, Nieuw Zeeland, Noorwegen en Saskatchewan (Canada)) op de manier waarop diverse variabelen in deze boeken worden vastgelegd.

De volgende hoofdstukken bevatten een veelheid aan tabellen. Effectiviteit van gordels (tabellen voor vijf Amerikaanse staten, Noorwegen, Groot-Brittannië, Zwitserland en Finland), effectiviteit van helmen (weer vier Amerikaanse staten, Noorwegen en Frankrijk), classificatie van letsel, oorzaken van dodelijke ongevallen en vele anderen, veelal uit verschillende landen.

Ook wordt in deze laatste hoofdstukken nog een vergelijking gemaakt tussen de verschillende bronnen van informatie, politie, ziekenhuizen en verzekeringsmaatschappijen.

De tabellen staan gewoon tussen de tekst, welke in twee kolommen in afgedrukt, geschreven. De tabellen zijn niet omlind en de kop van de tabel is in hetzelfde lettertype als de tekst, waardoor het niet altijd duidelijk is waar de tekst door gaat of waar een tabel begint.

Het boek is echt een samendraai van allemaal versnipperde informatie. Ik maak mij sterk dat er geen variabele is waarover een volledig beeld te verkrijgen is voor alle in het boek behandelde landen.

Tevens is het boek uitgegeven in A4 formaat, wat het nogal onpraktisch maakt voor een boekenkast.

Voor al door de onoverzichtelijkheid en de incompleetheit is het naar mijn mening geen boek voor algemeen gebruik.

***** CWI publicaties *****

A.W. van der Vaart

Statistical estimation in large parameter spaces

CWI, Amsterdam, tract 44, 1988, 205 pag., ISBN 90-6196-329-X, f 31.60.

S.A. van de Geer

Regression analysis and emperical processes

CWI, Amsterdam, tract 45, 1988, 161 pag., ISBN 90-6196-330-3, f 25.30.

F.C. Drost

Asymptotics for generalized chi-square goodness-of-fit tests

CWI, Amsterdam, tract 48, 1988, iv + 104 pag., ISBN 90-6196-348-6, f 16.50.

P.J.M. van Laarhoven

Theoretical and computational aspects of simulated annealing

CWI, Amsterdam, tract 51, 1988, viii + 168 pag., ISBN 90-6196-352-4, f 25.30.

K.C.P. Machielsen

Numerical solution of optimal control problems with state constraints by sequential quadratic programming in function space

CWI, Amsterdam, tract 53, 1989, 214 pag., ISBN 90-6196-, f .

L.C.R.J. Willenborg

Computational aspects of survey data processing

CWI, Amsterdam, tract 54, 1989, iv + 154 pag., ISBN 90-6196-356-7, f .

M.C.M. de Gunst

A random model for plant cell population growth

CWI, Amsterdam, tract 58, 1989, ii + 152 pag., ISBN 90-6196-365-6, f 24.50.

P. de Jong

Central limit theorems for generalized multilinear forms

CWI, Amsterdam, tract 61, 1989, 84 pag., ISBN 90-6196-369-9, f 14.10.

J. Mayer (editor)

Bringing jobs to people, employment promotion at regional and local levels

International Labour Organisation, Geneva, 1988, ISBN 92-2-106325-9, 211 pag., Sfr 25.00

Dit boek bevat een verzameling artikelen uit *International Labour Review* van de laatste tien jaren. Er worden diverse aanpakken belicht voor het ontwikkelen van regionale of lokale werkgelegenheid. Het standpunt wordt uitgedragen dat het veel beter is het werk naar de mensen te brengen dan van de werknemers te verwachten dat zij zullen verhuizen om een arbeidsplaats te vinden. Vanwege allerlei beperkingen kunnen deze arbeidsplaatsen niet altijd de vorm aannemen van een betaalde baan in de moderne sectoren van de economie. Het past echter in de strategie van de ILO dat een ieder toegang moet hebben tot land, gereedschappen, onderwijs, kredietverstrekking en dergelijke om uitvoering te geven aan het recht op een inkomengenererende baan.

H. Haan & J.-B. Célestin (editor)

Urban informal sector information: needs and methods

International Labour Organisation, Geneva, 1989, ISBN 92-2-107004-2, 44 pag., Sfr 12.50.

Tijdens het inter-regionale seminar over arbeidsmarkt informatie in Kopenhagen in oktober 1986 is een aanbeveling gedaan om te onderzoeken op welke wijze de informatie verbeterd kan worden omtrent de stedelijke informele sector. Dit betreft het meten van de informele werkgelegenheid, vaak bestaande uit vrouwen, schoolverlaters, ouderen en ongeschoolde migranten, veranderingen in beloningen en het formuleren van een beleid ter ondersteuning van deze sector.

Deze informele sector bestaat uit kleine economische eenheden of enkelingen welke ten hoogste in kleine mate gebruik maken van de technologische ontwikkelingen en daardoor slechts een bescheiden produktiviteitsniveau halen.

Deze monografie is een eerste aanzet voor het uitvoeren van deze aanbevelingen. Het bestaat uit vier delen. Het eerste deel geeft een overzicht van het verrichte onderzoek, het tweede deel inventariseert de beschikbare methoden om de informatie te verzamelen. Het derde deel analyseert de behoefte aan informatie en in het laatste deel worden suggesties geformuleerd voor een geïntegreerd systeem van informatieverzameling.

L. Richter

Upgrading labour market information in developing countries: problems, progress and prospects

International Labour Organisation, Geneve, 1989, ISBN 92-2-106453-0, 62 pag., Sfr 12.50

In 1977 initieerde de ILO een programma om de ontwikkelingslanden te helpen hun systeem van arbeidsmarktinformatie te verbeteren. Deze studie geeft een overzicht van wat is bereikt en wat nog moet worden gedaan. Het beschrijft de beperkingen van diverse gebruikte methoden voor het genereren van arbeidsmarkt informatie.

Ook dit werk is een uitvloeisel van de conferentie in Kopenhagen in 1986.

The cost of social security

International Labour Organisation, Geneve, 1988, ISBN 92-2-006423-5 163 pag., Sfr 50.00

In drie talen (Engels, Frans en Spaans) wordt verslag gedaan van het twaalfde internationale onderzoek naar de kosten van de sociale zekerheid.

De tabellen bevatten de resultaten van vragenlijsten, welke in mei 1985 aan alle lidstaten van de ILO zijn toegestuurd. De overheden werden gevraagd informatie te verschaffen omtrent de financiële aspecten van hun sociaal zekerheids systeem voor de kalenderjaren 1981-83. Deze tabellen maakt een vergelijking mogelijk tussen de verschillende landen, zowel voor wat betreft de inkomsten als de uitgaven de sociale zekerheid betreffende. Je moet dan nog wel de wisselkoersen weten, want de gegevens staan weergegeven in miljoenen van de geldeenheid van het betreffende land.

Statistical sources and methods, volume 4: employment, unemployment, wages and hours of work (administrative records and related sources)

International Labour Organisation, Geneve, 1989, ISBN 92-2-106406-9, 180 pag., Sfr 32.50.

Dit werk bevat methodologische beschrijvingen van de werkgelegenheidsgegevens, werkeloosheidsgegevens, lonen en aantal werkuren van 68 landen. Eerder werd deze informatie gepubliceerd in de 1986 editie van het *Year Book of Labour Statistics* en in de in 1987 uitgekomen nummers van *Bulletin of Labour Statistics*.

Het bevat de bronnen waarvan gebruik wordt gemaakt voor het samenstellen van de gegevens in de verschillende landen en de methoden van bewerken van de data.

De beschrijvingen bevatten onder andere de titel van de publicaties, de organisatie welke verantwoordelijk is voor de publicaties, de dekking van de cijfers, (geografisch, industrieel, beroepsgroep), definities (wat wordt verstaan onder werkeloosheid), referentieperiode en methode van verzamelen van gegevens.