KM 33(1990)
pag 189-214

Boekbesprekingen

Redactie:

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

telefoon 013 - 640319

Catrien C.J.H. Bijleveld

Exploratory linear dynamic systems analysis

DSWO-Press, Leiden, 1989, 158 pag, ISBN 90-6695-035-8, *f* 43.00.

Na de introduktie in hoofdstuk 1 worden in hoofdstuk 2 van dit boek enige niet-dynamische regressie technieken beschreven, zoals multiple regressie, redundantie analyse en 'reduced rank' regressie.

Hoofdstuk 3 behandelt een aantal technieken voor analyse in het tijddomein van tijdsafhankelijke gegevens, zoals Box-Jenkins technieken, covariantie struktuur analyse en dynamische faktor analyse. Ook spektraalanalyse komt aan de orde.

In hoofdstuk 4 wordt uitgebreid ingegaan op het lineaire dynamische model. De rol van de latente toestanden en overgangsmatrices wordt toegelicht. Het Kalman-filter wordt besproken en enkele bestaande technieken om toestand ruimte modellen te fitten worden geschetst.

In hoofdstuk 5 wordt het ALS algorithme gepresenteerd. Hierin worden de overgangsmatrices en de latente toestanden alternerend geschat; voor gegeven latente toestanden wordt een optimale oplossing voor de transitiematrices berekend, gegeven die optimale oplossing voor de overgangsmatrices, wordt eveneens een optimale oplossing voor de latente toestanden afgeleid, enzovoorts.
Deze twee substappen kunnen worden uitgebreid met een derde substap waarin, gegeven de oplossingen voor de overgangsmatrices en latente toestanden, optimale kwantifikaties voor niet-numerieke variabelen worden berekend.
Verder worden enkele mogelijke rotaties voor de oplossingen gegeven en wordt de mogelijkheid besproken om voor sommige input variabelen een vertraagde werking te modelleren.

Tenslotte geeft hoofdstuk 6 een groot aantal voorbeelden waarbij gebruik wordt gemaakt van de ontwikkelde technieken.

Tot voor kort werd analyse van lineaire dynamische systemen alleen toegepast in exacte wetenschappen. In dit boek wordt bovengenoemd onderwerp bestudeerd vanuit een sociaal wetenschappelijke achtergrond.
Vanwege de vele voorbeelden en de praktische toepasbaarheid van de afgeleide resultaten is het boek interessant voor methodologen op sociaal-wetenschappelijk, medisch en economisch gebied.

Kees van Montfort
Hoogovens Groep, IJmuiden

Prof. dr. B.B. van der Genugten

Inleiding tot de waarschijnlijkheidsrekening en mathematische statistiek deel 2

Stenfert Kroese, Leiden, 1988, xiv + 587 pag., ISBN 90-207-1669-7, *f* 89.50.

Dit lijvige boek is het tweede deel van een tweedelige inleiding tot de waarschijnlijkheidsrekening en mathematische statistiek (totaal ruim 1000 bladzijden).
De nummering van de hoofdstukken in dit deel loopt door:

6 Multivariate kansverdeling
7 Het lineaire model
8 Convergentie
9 Basisbegrippen in de verklarende statistiek
10 Statistische beslissingsmethoden
11 Schatting- en toetsingtheorie
12 Generalisaties van het lineaire model

Indien tijdens het studeren de aangegeven volgorde aangehouden wordt, is het boek zeer goed leesbaar, duidelijk in de bewijsvoering en konsekwent in opbouw. Zelfs bij het 'hier en daar' een paragraaf lezen is de draad van het betoog vrijwel altijd duidelijk. In dat geval geven de titels van de paragrafen vaak te weinig informatie over de inhoud.
De schrijver geeft vaak blijk van eigen inbreng in stukken die in het algemeen een gebruikelijke aanpak vertonen.

Het is verheugend dat het meetkundige aspekt bij het lineaire model niet achterwege is gelaten. Toch zou de dualiteit van de uitkomstenruimte en de variabelenruimte, zoals onder meer uitgedrukt is in het lemma op pag. 153, meer gebruikt moeten worden. Het is dan ook jammer dat de stelling van Markov ±250 pagina's later behandeld wordt dan de LS-methode.

Aan de keuze van onderwerpen is te merken, dat het boek geschreven is vanuit een studierichting econometrie. Overigens is dat niet storend. De architektuur van het boek doet denken aan die van een kollegediktaat. De toetsen zoals die van Spearman komen wel erg laat.

Het boek 2 is, in tegenstelling tot boek 1, als getypt manuscript gedrukt. De kleinere letters voor de voorbeelden geven naar mijn mening een onrustig beeld.

A.C. van Eijnsbergen
Landbouwuniversiteit Wageningen

Andrew F. Siegel

Statistics and data analysis, an introduction

John Wiley & Sons, New York, 1988, xxv + 518 pag., ISBN 0-471-87659-3, £ 28.35.

De auteur heeft zich te doel gesteld een boek te schrijven, welk een inleiding geeft voor de niet technische studenten in de methoden van de statistiek. De klassieke methoden van presenteren van de stof en de wat informelere methode van Exporatieve Data Analyse zijn verweven in de tekst.
De inleiding van het boek begint met 'Welcome to the world of statistics, the art of drawing conclusions from imperfect data'.

Ieder hoofdstuk is op dezelfde wijze ingedeeld. Na een zeer korte inleiding wordt de stof behandeld, deze wordt op ongeveer één pagina samengevat aan het eind van het hoofdstuk. Daarna worden er vragen gesteld, welke betrekking hebben op de begrippen die in het hoofdstuk zijn behandeld en een verzameling vraagstukken sluiten het geheel af.

Wat komt in het boek aan de orde. De negentien hoofdstukken bevatten de volgende onderwerpen

1. *Inleiding*, korte omschrijvingen van begrippen als populatie en steekproef, variabiliteit, robuuste methoden en modellen, wat is statistiek en waarom is het noodzakelijk iets van dit vakgebied te weten.

2. *Weergave van een dataset*, tabellen, histogram, stam- en blad diagram.
3. *Verschillende vormen van verdelingen*, symmetrische verdeling, normale verdeling, scheve verdelingen, uitschietende waarnemingen.
4. *Locatiematen*, mediaan (gevoeligheid voor uitschieters), rekenkundig gemiddelde en een vergelijking tussen beide, modus.
5. *Maten voor variatie*, range, standaard deviatie (met verwijzing naar de eigenschap van de normale verdeling), uitleg berekeningswijze (waarom delen door n-1), kwartielsafstand en een vergelijking tussen deze laatsten. Er wordt ook even aandacht besteed aan de invloed van rekenonnauwkeurigheid.
6. *Gegroepeerde waarnemingen*, histogram, gemiddelde, mediaan, standaard deviatie en kwartielsafstand als de waarnemingen in klassen zijn ingedeeld, dichotome waarnemingen.
7. *Cumulatieve verdeling en percentielen*, wat is een cumulatieve verdeling, hoe teken je die (uitleg aan de hand van plaatjes van deelresultaten), berekeningswijze voor percentielen, waarbij mediaan en kwartielen als speciale percentielen worden genoemd.
8. *Eenvoudige plaatjes*, vijf-getallen samenvatting (aantal waarnemingen, minimum, eerste kwartiel, mediaan, derde kwartiel, maximum), twee soorten box-and-whisker plaatjes, met de extremen als eindpunten van de snorharen en met de lengte snorharen 1.5 maal de kwartielsafstand.
9. *Transformaties*, uitgebreide beschrijving van mogelijke transformaties en de redenen, met behandeling van de transformaties x^2, $\sqrt{x}$, log(x), $1/\sqrt{x}$ en 1/x en voor proporties en percentages log(x/(1-x)) en $\sqrt{x}$-$\sqrt{(1-x)}$. Er wordt stilgestaan bij de vraag wanneer en waarom transformaties goed zouden zijn.
10. *Kansrekening*, standaard uitleg aan de hand van Venn diagrammen en kansbomen.
11. *Stochastische variabelen en normale verdeling*, in de inleiding wordt duidelijk gesteld dat stochastiche variabelen refereren aan een proces en dat de uitkomsten realisaties zijn van dat proces. Discrete stochasten, (standaard) normale verdeling, centrale limietstelling, wat te doen als de verdeling niet normaal is.
12. *Statistisch ondersteund beslissen*, populatie en steekproef, goede keuze van steekproefmethode, standaard fout van gemiddelde en percentage.
13. *Betrouwbaarheidsintervallen*, uitleg over de interpretatie, interval voor gemiddelde van normale verdeling, parametervrij interval voor de mediaan.
14. *Toetsen van hypothesen*, eerst afgeleid uit betrouwbaarheidsinterval, daarna pas via kritieke waarden, t-toets, z-toets, tekentoets.
15. *Vergelijken van twee groepen waarnemingen*, tekenen van histogrammen (visuele vergelijking), betrouwbaarheidsinterval voor gemiddeld verschil (toetsen wordt ook weer via deze intervallen behandeld) en gepaarde waarnemingen.
16. *Vergelijken van meerdere groepen waarnemingen*, eerste visueel (Box-and-whisker plots), daarna via variantie analyse.
17. *Bivariate waarnemingen*, inleiding via stimulus-respons experiment en tijdreeksen, spreidingsdiagrammen, smoothing via voortschrijdend gemiddelde en voortschrijdende medianen. Hier komen ook transformaties aan bod.
18. *Lineaire relaties*, correlatie coëfficiënt (veel spreidingsdiagrammen als voorbeeld getekend, met overeenkomstige r-waarde), lineaire regressie en residuen.
19. *Categorische waarnemingen*, chi-kwadraat toetsen voor onafhankelijkheid en gelijkheid van verdelingen.

Het boek is zeer ruim opgezet, er wordt uitgebreid aandacht besteed aan zaken als veronderstellingen bij toetsen, interpretatie van significantiewaarden en waarde van correlatie coëfficiënt.
Bij de behandeling van transformaties worden de verschillende schalen onder elkaar gezet om te kijken wat de getransformeerde grootheid doet met de afstanden.

Bij het lezen krijg je het gevoel dat de auteur je mee wil nemen voor een onderzoekstocht door een spannend verhaal te vertellen. De tekst is ook in een dergelijke vorm geschreven.
Regelmatig worden er in de tekst vragen gesteld, die dan direct daarop worden beantwoord:

> 'why is the five-number summary useful to us? Because ....'
>
> 'is it fair to transform the data values? Can't transformations be used to manipulate the data in order to sunstantiate a lie?' .....
>
> '... in particular, you begin to wonder about 57.1% you have now. Is it significant different from 50%?'

Alhoewel het beperken van de stof het grootste probleem is voor een enthousiaste verteller, en er dus wel goed overdachte redenen zullen zijn waarom bepaalde onderwerpen niet aan bod komen, mis ik toch een aantal zaken. Bijvoorbeeld het onderscheidingsvermogen en plaatjes van ß-risico bij toetsen. Het begrip type II fout wordt wel aangedragen, maar daar wordt verder niets mee gedaan. Aan de hand van een enkel plaatje zou dit onderwerp duidelijk gemaakt kunnen worden. Tevens ontbreekt een toets op of een betrouwbaarheidsinterval voor σ, terwijl het quotient van twee varianties wel wordt behandeld bij variantie analyse en de chi-kwadraatgrootheid aan bod komt bij de categorische waarnemingen.

De tekst is in twee kleuren gedrukt, zwart en blauw, waarbij de blauwe kleur is gebruikt voor de aandachtspunten en de hoofdstuk- en paragraafkoppen. De talrijke plaatjes zijn ook in blauw en zwart, soms met een grijze achtergrond. Dit maakt de tekst extra overzichtelijk. Er worden zeer veel getalsvoorbeelden gebruikt om de aangedragen stof te illustreren.
Zowel de inhoudsopgave als de index (drie kolommen per pagina) bestaan uit 13 pagina's, wat opzoeken van begrippen eenvoudig maakt.

Het lijkt mij een uitstekend boek voor studenten, maar ook voor anderen die een opstap willen maken in de beginselen van de statistiek. Het is uitstekend geschikt voor zelfstudie. Maar ook mensen die een tekst voorbereiden over de behandelde onderwerpen kunnen er hun voordeel mee doen. Fijn boek!

Arend Oosterhoorn
Van Doorne's Transmissie bv

D. Stoyan, W.S. Kendall and J. Mecke

Stochastic Geometry and Its Applications

John Wiley & Sons, Chichester, 1987, 345 pag., ISBN 0-471-90519-4, $ 57.50.

Dit boek is een vertaling van een oorspronkelijke in het duits verschenen boek van de eerste en derde auteur [1], aangevuld met nieuw materiaal. Er wordt sterk de nadruk gelegd op toepassingen, terwijl daarnaast gestreefd wordt naar een redelijk volledige prestatie van de benodigde wiskunde. Het is duidelijk dat in een dergelijk opzet noch de wiskunde, noch de toepassing volledig self-contained behandeld kunnen worden, althans niet binnen het kader van een nog overzichtelijk aantal pagina's. Dat heeft tot gevolg dat bij de toepassingen vaak de buitenwiskundige overwegingen, die tot keuze van het wiskundig model leiden, achterwege blijven, danwel vervangen worden door heuristische overwegingen.

Een dergelijke hybride opzet maakt het schrijven van een boek tot een hachelijke onderneming: het gevaar dat geen enkele kategorie van lezer nog enige informatie in het werk kan vinden is

groot. Naar mijn mening zijn de auteurs desondanks zeer wel geslaagd in hun opzet. Het boek verschaft aan de wiskundige toegang tot de meest diverse toepassingen van de stochastische meetkunde. De zeer uitvoerige bibliografie verleent daarbij belangrijke steun. Ik kan minder goed beoordelen of de behandeling van de wiskundige theorie toegankelijk is voor de niet-wiskundige toepasser; wel lijkt mij de presentatie van de theorie helder en zeer geschikt al inleiding in een voorbereiding op een uitvoeriger behandeling.

De samenwerking tussen Wiley en Akademie Verlag heeft geleid tot een behoorlijke uitgave met wat te veel drukfouten naar mijn smaak; bij een vorige gelegenheid heb ik al opgemerkt dat het wenselijk zou zijn als Wiley in het kader van deze samenwerking Akademie Verlag fatsoenlijk papier zou leveren; het moet toch mogelijk zijn een dergelijk boek op minder grauw papier te vervaardigen. Afgezien van deze weinig belangrijke bezwaren beveel ik het boek van harte aan.

[1] Stoyan, D. en Mecke, J. (1983). *Stochastische Geometrie: Eine Einführung.* Akademie Verlag, Berlin

C.L. Scheffer
TU Delft

T.C. Gard

Introduction to stochastic differential equations

Marcel Dekker Inc., New York en Bazel, 1988, xi + 234 p., ISBN 0-8247-7776-X, $ 78.00.

De literatuur over stochastische differentiaal vergelijkingen - zeker voor zover het inleidende boeken betreft - groeit onstuimig. De behoefte aan een zo pijnloos mogelijke kennismaking met dit weerbarstige en toch voor toepassers zo belangrijke vakgebied, verleidt steeds meer auteurs ertoe hun versie van een eenvoudige inleiding in dit gebied aan het wiskundige en/of niet-wiskundige publiek aan te bieden, en wat minstens zo belangrijk is: brengt uitgevers ertoe om zulke boeken in steeds grotere hoeveelheden op de markt te brengen. Als er dan weer zo'n nieuwe tekst verschijnt dan is enige konsumentenonderzoek wel op zijn plaats.

Het hier besproken boek van gard kan qua inhoud, intentie en omvang vergeleken worden met o.a.

[1] Arnold, L. (1974). *Stochastic differential equations, Theory and applications.* John Wiley and Sons

[2] Schuss, Z. (1980). *Theory and applications of stochastic differential equations.* John Wiley and Sons.

Het boek van Gard en [1] en [2] hebben de beperking tot stochastische integralen met betrekking tot Brownse beweging, redelijk veel aandacht aan expliciete toepassingen gemeen. Het is dan ook niet verbazend dat deze boeken nogal op elkaar lijken. Onder deze omstandigheden moeten buiten-wiskundige overwegingen een rol spelen bij het bepalen van een voorkeur. Dan lijkt me het volgende gegeven van doorslaggevende betekenis: het boek van Gard kost US $ 78.00, [2] kost US $ 32.50 en van [1] weet ik de huidige prijs niet, maar mijn exemplaar kostte destijds *f* 66.10.

Men begrijpe mij goed: ik vindt het besproken werk zeker niet slecht; in tegendeel, er zijn vele aardige vondsten in de presentatie te vinden. ik begrijp alleen niet goed voor wie een dergelijk 'kapitaal' boek bedoeld is. Moet men een dergelijk boek voorschrijven bij een kollege stochastische analyse? De prijs is dan een te grote barriere. Moet een bibliotheek dit boek aanschaffen?

Bij het grote aanbod boeken op dit terrein op uiteenlopende nivo's zou ik ook geneigd zijn daarover niet positief te adviseren, gelet op de steeds krapper wordende bibliotheekbudgetten.

Samenvattend moet ik zeggen dat het best de moeite waard is om sommige paragrafen van dit boek te bestuderen, maar dat het aan te bevelen is om dit te doen uit een toevallig langslopend exemplaar. Aanschaf lijkt me niet een verstandige investering.

C.L. Scheffer
TU Delft

Janos Galambos

Advanced probability theory

Marcel Dekker Inc., New York and Basel, 1988, vii + 405 pag., ISBN 0-8247-7873-1, $ 119.50.

Dit boek is een gedegen tekst aangaande de waarschijnlijkheidsrekening als wiskundige theorie. Het boek is 'self-contained' en bevat volledige bewijzen van behandelde stellingen. Het potentiële lezerspubliek bestaat uit docenten en studenten aan universiteiten, die zich willen verdiepen in de fundamentele aspekten van de waarschijnlijkheidsrekening zoals deze wordt toegepast in onder andere de statistiek, fysica, chemie, biologie, econometrie, etc. Ook voor onderzoekers met belangstelling voor de wiskundige finesses van de kansrekening, kan dit boek als een naslagwerk en referentiebron fungeren.

Hoofdstuk 1 behandelt de basis begrippen van de kansrekening.
De verwachting van stochastische grootheden en allerlei konvergentie eigenschappen van rijen stochastische grootheden komen in Hoofdstuk 2 aan de orde.
Integraal transformaties, zoals de Fourier transformatie (karakteristieke funkties) en de Laplace transformatie, vormen het onderwerp van Hoofdstuk 3.
In hoofdstuk 4 wordt een uiteenzetting gegeven van de eigenschappen (o.a. limietverdelingen) van rijen stochasten. In de hoofdstukken 5-8 komen wat meer gespecialiseerde onderwerpen aan de orde: oneindig deelbare kansverdelingen (Hoofdstuk 5), onafhankelijke en gelijk verdeelde stochastische grootheden (Hoofdstuk 6), de stelling van Randon-Nikodym in verband met de voorwaardelijke verwachtingsoperator en martingalen (Hoofdstuk 7), stochastische processen (Hoofdstuk 8).

Ieder hoofdstuk is voorzien van opgaven. De aanwijzingen welke ten aanzien van deze opgaven achter in het boek zijn opgenomen, maken het boek tevens geschikt voor zelfstudie.

D.A. Overdijk
Faculteit Wiskunde & Informatica
TU Eindhoven

Terry E. Dielman

Pooled cross-sectional and time series data analysis

Marcel Dekker Inc., New York and Basel, 1989, vii + 249 pag., ISBN 0-8247-7864-2, $ 79.75.

Het boek is als overzicht en praktische handleiding bijzonder interessant, wanneer men zich wil gaan bezighouden met paneldata. De onderzoeker die beschikking heeft over gegevens van meerdere objekten op verschillende tijdstippen, zou aan de hand van dit boek analyses uit kunnen voeren.
Hiervoor behoeft hij/zij niet het gehele boek te bestuderen, maar kan er begonnen worden met paragraaf 8.3. In deze paragraaf worden richtlijnen gegeven om tot de keuze van een analysemodel te komen. Dielman beperkt zich daarbij tot de methoden die in de voorafgaande hoofdstukken zijn besproken. In de terminologie van Dielman zijn dat 'Separate Regressions and Classical Pooling', 'Seemingly Unrelated Regressions', 'Error Components Models', 'Random Coefficients Regression', 'Time and Cross-sectionally Varying Parameter Models' en 'Models for Discrete Data'.
Aan de hand van een aantal vragen kan men vervolgens het hoofdstuk bekijken, dat het meest geschikte model behandeld.

Vervolgens bevat het boek naast dit praktische element ook een duidelijk theoretisch element, wat van nut is voor een breed overzicht van de mogelijkheden. Een aantal methoden is naar onze mening dusdanig uitgebreid dat een daadwerkelijke schatting vrijwel onmogelijk is. Het idee zou dan meer gezien moeten worden als een theoretische uitwerking van een model, dat, voordat zo'n idee praktisch toegepast kan worden, eerst door een aantal aannames vereenvoudigd moet worden.

Het is dus wellicht ook de ingewikkeldheid van de modellen, die ervoor gezorgd heeft dat in een aantal formules wat vergissingen zijn geslopen.

Het zou handig zijn geweest wanneer door het boek heen een nog meer eenduidige notatie zou zijn gevoerd, zodat een gamma en een delta niet zowel als parameter als ook als storingsterm gebruikt worden.
Verder hadden wij in plaats van de nu geplaatste appendices, die voor bekenden met econometrie vrij eenvoudig zijn, liever iets meer ingewikkelde modelafleidingen, zoals de stap van formule 4.14 naar 4.19, wat uitgewerkt gezien.

Tot slot zijn de bovengenoemde opmerkingen niet bedoeld om het boek te ontraden, omdat dit werk naar ons idee voor eenieder die zich met paneldata wil gaan bezighouden van veel nut kan zijn. Wanneer men dit boek als uitgangspunt neemt, kunnen van hieruit veel referenties gevonden worden. Wanneer de kennis over dit onderwerp niet zo groot is, vormt dit boek een goede inleiding daartoe, er wel van uitgaande dat er een zekere econometrische basiskennis bestaat.
En ook voor onderzoekers die snel met een databestand aan de slag willen, is dit boek een geschikte handleiding, waarbij dan wel enige referenties gezocht moeten worden, die een wat bredere uitwerking van de methoden geven.

Wanneer er bij een aantal methoden iets meer aanwijzingen en voorwaarden gegeven waren, zou het boek, hoewel dikker, nog meer waarde als naslagwerk krijgen.

G.R. Klab en G.J. Zijlstra
GAK, Amsterdam

D.R. Cox & E.J. Snell

Analysis of Binary Data, second edition

Chapman and Hall, London, 1989, xi + 236 pag., ISBN 0-412-30620-4, £ 20.00.

In dit boek wordt op een systematische wijze de analyse van binaire waarnemingen behandeld. Elke waarneming heeft maar twee mogelijke uitkomsten, bijvoorbeeld succes en mislukking, en het doel van de analyse is het nagaan hoe de kans op succes afhangt van verklarende variabelen en van andere strukturen in het materiaal.

De stof wordt gepresenteerd in 5 hoofdstukken en 4 appendices.
Hoofdstuk 1, "Binary response variables' (25 blz.), introduceert de problematiek door middel van voorbeelden en geeft een diskussie van diverse mogelijkheden voor het modelleren van de succeskans als funktie van verklarende variabelen.
Hoofdstuk 2, 'Special logistic analysis' (70 blz.), heeft als onderdelen onder andere Simple regression, 2x2 Contingency table, Matched pairs, Several 2x2 contingency tables, Multiple regression, Residuals and diagnostics, Cross-over designs, Rasch model en Binary time series.
In het derde hoofdstuk, 'Some complications' (20 blz.), worden de volgende onderwerpen behandeld: overdispersie en quasi-likelihood, empirical-Bayes methoden, meetfouten in het waarnemen van binaire responsies.
In hoofdstuk 4, 'Some related approaches' (23 blz.), wordt het gebruik van logistische modellen in analyse van case-control onderzoek besproken. Daarnaast worden er verschillende theoretische formuleringen besproken die alle tot dezelfde logistische regressie-analyse leiden.
De onderwerpen van hoofdstuk 5, 'More complex responses' (18 blz.), zijn: Paired preferences, Nominal data, Ordinal data, Multivariate data en Mixed binary and continuous responses.
(1974)

Elk hoofdstuk eindigt met de bespreking van relevante literatuur (Bibliographic notes). De appendices zijn: 'Theoretical backgrounds (21 blz.), 'Choice of explanatory variables in multiple regression' (7 blz.), 'Review of Computational aspects' (3 blz.) en 'Further results and exercises (58 opgaven op 17 blz.).

Het boek is de tweede uitgave van het gelijknamige boek van Cox uit 1970. Cox en Snell hebben de originele tekst grondig aangepast en deze verjongingskuur is geslaagd. De bruikbaarheid van de aanpassingen hebben de auteurs ook in de praktijk getoetst, zoals de deelnemers aan de door Cox en Snell gegevens kursus in september 1987 in Groningen mochten ervaren. De voornaamste wijzigingen betreffen het toevoegen van nieuw materiaal (zowel theoretische resultaten als voorbeelden van praktische toepassingen) en het verplaatsen van de nadruk op de methode van de meeste aannemelijkheid ten koste van de methode van de gewogen kleinste kwadraten.

Ook de appendices zijn nieuw en in plaats van het destijds tamelijk komplete geannoteerde literatuuroverzicht, beperken de auteurs zich nu tot het noemen van de belangrijkste bronnen. Ondanks deze beperking is de nieuwe literatuurlijst met zijn 15 bladzijden omvangrijker dan in de eerste uitgave. Nog één verandering wil ik expliciet noemen. In het onderwerpregister van de eerste uitgave komt de kurieuze ingang 'Impatient author, evidence of'. Die is nu verdwenen.

De presentatie van theoretische en praktische aspekten van statistische analyse en binaire responsies is in deze monografie heel goed uitgebalanceerd.
Cox verstaat deze kunst als geen ander. Het boek kan ik warm aanbevelen aan eenieder met belangstelling voor binaire data.

V. Fidler
RU Groningen

M.E. Johnson

Multivariate Statistical Simulation

John Wiley & Sons, New York, 1987, ix + 230 pag., ISBN 0-471-82290-6. £ 33.75.

Het nabootsen van steekproeftrekking in Monte Carlo studies is bijna zo oud als de statistiek zelf (zie bijvoorbeeld Teichroew (1965)). De scherpe prijsdaling van het rekenen in de laatste decennia maakt het bovendien relatief eenvoudig om MC studies op grote schaal uit te voeren. Tegelijkertijd is de belangstelling voor simulatie onder statistici toegenomen, onder andere door ontwikkelingen op het gebied van exploratieve data-analyse, robuuste methoden en toepassing van bootstrap. De potentiële vraag naar rekenkapaciteit in de statistiek heeft daardoor dezelfde astronomische omvang aangenomen als in de getaltheorie of de theoretische scheikunde.

In de ontwikkeling past vanzelfsprekend een toegenomen aandacht voor het genereren van pseudo-stochastische grootheden. Daarbij is echter relatief weinig gedaan aan het genereren van *multivariate* data. Dit is jammer, omdat juist in de multivariate analyse veel interessante problemen, bij het ontbreken van analytische oplossingen, door simulatie onderzocht kunnen worden. Het onderzoek van Boomsma (1983) naar de robuustheid van LISREL is hiervan een goed voorbeeld. Het is daarom verheugend dat Mark Johnson (Los Alamos National Laboratory, New Mexico, USA) een monografie op dit gebied heeft geschreven.

Het boek behandelt de simulatie van trekking uit kontinue multivariate verdelingen. In de inleiding wordt aan de hand van drie uitgebreide voorbeelden het nut van Monte Carlo-onderzoek nog eens toegelicht. Na een kort overzicht van algemene methoden voor het genereren van univariate en multivariate verdelingen, volgt een aantal hoofdstukken waarin specifieke families van verdelingen worden behandeld. Achtereenvolgens zijn dit multivariate normale verdelingen en mengsels daarvan (hoofdstuk 4); verdelingen die door eenvoudige transformaties uit de normale verdeling kunnen worden afgeleid, zoals de lognormale verdeling, de arcsinh-normale en de logit-normale verdeling (hoofdstuk 5); verdelingen met elliptische kontouroppervlakken van de dichtheid, waarbij met name Pearson Type II en VII verdelingen aandacht krijgen (hoofdstuk 6); uniforme en niet-uniforme verdelingen op de n-sfeer (hoofdstuk 7).
In hoofdstuk 8 wordt de toepassing besproken van een stelling van Khintchine over het ontbinden van stochastische grootheden met een eentoppige verdeling rond de oorsprong.
Hoofdstuk 9 bespreekt het genereren van multivariate Burr, Pareto en logistische verdelingen.

De waarde van een boek als dit is vooral dat informatie, die verspreid in tijdschriften is te vinden, op systematische wijze bijeen wordt gebracht. Alleen al de literatuurlijst is om die reden interessant. Aardig vindt ik verder, dat de vorm van uni- en bivariate verdelingen uitgebreid met behulp van grafieken worden geïllustreerd.

Er zijn echter ook zwakke kanten. De algemene inleiding op het konstrueren van afhankelijkheid tussen te genereren grootheden is wel zeer summier; de keuze van onderwerpen die daarop volgen, krijgt daardoor enigszins een ad hoc karakter.
Verder ontbreekt vrijwel alle informatie over de efficiëncy van beschreven algoritmen. Bij het opzetten van een grotere Monte Carlo studie is dat nu juist een punt van ernstige overweging. Wie niet voldoende heeft aan de spaarzame en vage evaluatieve opmerkingen over algoritmen, zal zelf in de oorspronkelijke literatuur moeten zoeken.

Samenvattend kan gesteld worden dat 'Multivariate Statistical Simulation' een boek is, dat - ondanks tekortkomingen- veel recente informatie over het genereren van multivariate data bevat. Het is daarom interessant zowel voor degenen die een multivariate Monte Carlo studie willen uitvoeren, als voor de ontwikkelaars van nieuwe methoden op dit gebied.

referenties:

Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and non-normality*. Dissertatie RU Groningen

Teichroew, D. (1965). A history of distribution sampling prior to the area of the computer and its relevance to simulation, *J. Amer. Statist. Assoc.*, **60**, 27-49.

Ron Visser
RU Leiden

G.V. Dyke

Comparative experiments with field crops (2nd edition)

Charles Griffin/Oxfort University Press, London/New York, 1988, xiv + 262 pag.,
ISBN 0-85264-282-2, £ 27.50.

Het boek valt uiteen in twee delen. Het eerste deel 'How to do and interpret field experiments' gaat in op praktische zaken rond het opzetten en uitvoeren van experimenten. Het tweede deel bevat een simpele beschrijving van de statistische technieken die gebruikt worden bij landbouwkundig onderzoek. De doelgroep is de onderzoeker die zelf de experimenten uitvoert. Het boek is zeer expliciet niet bedoeld voor statistici, of zoals Dyke het zelf verwoordt: 'None of it, at least of all the chapters on statistics, is written for statisticians' (Preface, p. vi).
De tweede editie verschilt van de eerste in de hoofdstukken 3 (over het uitleggen van proeven in bijzondere omstandigheden, bijvoorbeeld de tropen), 9 (historie), 15 (statistics for applied research in the tropics, geschreven door G.M.F. Grundy) en 17 (Patterns of yield).

Deel 1: Opzetten en uitvoeren van experimenten.
De onderwerpen die in dit deel behandeld worden lopen uiteen van de definitie van een experiment tot aanwijzingen omtrent zaaien en oogsten en interpretatie van de resultaten. Voor de in West-Europa werkzame onderzoekers en statistici geldt, dat de meeste zaken bekend of al opgelost zijn of anders specialisten met de nodige kennis in de buurt. Voor onderzoekers en statistici werkzaam in minder ontwikkelde landen, kunnen de beschreven oplossingen en overwegingen behulpzaam zijn, al was het maar als ondersteuning van de eigen ideeen.
De behandelde zaken kunnen zeer basaal zijn ('Don't use 2 cubical dice to randomize 11 treatments, coded 2, 3, ... 12, using the total score; 6 is five times more likely to be thrown than 2 or 12') en zeer praktisch (hoe maken we een rechthoekig proefveld? Gebruik van kettingen van vaste lengte, spiegels, etc.; hoe zorgen we ervoor, dat markeringen over verschillende jaren heen gehandhaafd blijven, zonder een obstakel te zijn). De stijl van het boek is informeel. Dit betekent, dat de (statistische) overwegingen om tot een bepaalde keuze te komen, losjes geformuleerd staan. Dit boek kan derhalve niet tippen op dit front aan het boek van D.R. Cox (Planning of experiments). Niet zelden wordt ook de eigen mening van Dyke gegeven of de gangbare aanpak beschreven. Dit kan voor iemand die geïsoleerd moet werken, een prettige schrijfvorm zijn.

Deel 2: Statistiek in veldexperimenten.
Het eerste hoofdstuk in dit deel gaat over kontrasten, vrijheidsgraden en (het nut van) randomisatie. Opmerkelijk is het feit, dat voor de geometrische interpretatie van regressie en variantie-analyse een hoofdstuk ingeruimd is. Regressie-analyse en covariantie-analyse worden

vervolgens reken-voorschrift-achtig en zeer kort besproken, waarbij de bespreking beperkt blijft tot het schatten van de parameters, zoals te verwachten zonder matrix-rekening. Hoofdstuk 'Statitics for applied research in the tropics' gaat in op het werk van de statisticus/wetenschapper in de onderontwikkelde landen. Dit hoofdstuk was voor mij een van de interessantste om te lezen als anekdote voor de omstandigheden waaronder daar gewerkt moet worden en wegens zijn praktische aard. Duidelijk is, dat hoogstandjes met behulp van computers niet relevant zijn, maar simpele, snelle en duidelijke technieken vereist zijn. De voortdurende afweging van de praktische bruikbaarheid van resultaten en het streven naar zo ekonomisch mogelijk gebruik van de schaars ter beschikking staande middelen, spelen de hoofdrol.
De laatste twee hoofdstukken gaan in op patronen in het veld. Residuen-onderzoek en het gebruik van covariabelen krijgen aandacht. Modernere ontwikkelingen, zoals 'Nearest neighbour analysis' blijven buiten beschouwing ('As a throughly interested party (and one who finds Wilkinson et al.'s arguments difficult to follow) I dare not offer a judgement.').

Konklusie:

Dit boek kan dienen als een aanvulling op andere, wanneer een makkelijk leesbaar en praktisch opgezet boek gezocht wordt. Het boek kan gebruikt worden om ideeën op te doen omtrent proefopzet en analyse van resultaten, maar is zeer zeker niet bruikbaar als kookboek. Voor de statisticus, werkzaam in een van de ontwikkelde landen, biedt het boek weinig nieuws (zoals ook opgemerkt in het voorwoord van het boek). Voor onderzoekers en statistici in de onderontwikkelde landen kan het dienen als een 'gesprekspartner'.

Peter van Ewijk
Duphar bv

P.J.M. Wijngaard

Scheduling models in farm management: a new appraoch

IMAG, Wageningen, 1989, 244 pag., *f* 30.00.

In dit boek worden modellen voor de operationele planning ter ondersteuning van de landbouwbedrijfsvoering beschreven. Deze modellen worden gebruikt om verschillende bewerkingen, zoals oogsten, zaaien, cultiveren, enzovoort, in de juiste volgorde en op het goede moment uit te voeren. Hierbij is steeds het doel de kosten te minimaliseren.

De drie meest gebruikte modellen voor landbouwkundige planning, gebaseerd op respektievelijk dynamische programmering, lineaire programmering en simulatie, worden besproken. Ze zijn retrospektief onderzocht op verschillende aggregatie en relaxatie nivo's. Een stochastisch dynamisch programmeringsmodel is eveneens behandeld.

Het dynamische programmeringsmodel is gekozen als basis voor verder onderzoek. Een nadeel van dit model is de lange rekenduur. 'Hill-climbing' kan een 'goede' oplossing snel vinden (weinig rekentijd), maar er is een grote kans dat deze oplossing niet optimaal is. Hill-climbing, bijgestaan door een heuristieke evaluatie funktie om het verschil tussen de gevonden oplossing en een optimale oplossing zo klein mogelijk te houden, is gekozen om de optimale oplossing in een dynamisch programmeringsmodel te vinden. Dit algoritme wordt gebruikt voor een nieuw landbouwplanningsmodel, dat FLOS (= Farm Labour and Operations Scheduling) heet.
FLOS is stochastisch en maakt gebruik van weersvoorspellingen en bodem- en graanvochtigheidsvoorspellingen om toekomstige werkzaamheden te berekenen.

Het is op diverse manieren, onder andere met een uitgebreid praktijk voorbeeld, getest. De praktijktest had betrekking op de oogst van een akkerbouwbedrijf, dat onderdeel van de Rijksdienst voor de IJselmeer Polders (RIJP) is.

Naar aanleiding van het FLOS-algoritme is het Decision Support System SCHEMA (= SCHeduling Multiple Activities) ontworpen. Het programma, dat ontwikkeld is voor het gebruik op een PC en volledig menu-gestuurd is, wordt uitvoerig behandeld.

In dit boek worden akkerbouw, operationele research, stochastiek, meteorologie, bodemscheikunde en kunstmatige intelligentie geïntegreerd. Fundamenteel-theoretische noviteiten op operationele research of stochastiek-gebied komen echter niet aan de orde. Het boek illustreert onder andere op gestruktureerde wijze het gebruik van operationele research technieken voor landbouwkundige toepassingen en toont welke stadia hierbij moeten worden doorlopen.

Kees van Montfort
Hoogovensgroep, IJmuiden

J. Bert Keats & Norma Faris Hubele (eds.)

Statistical Process Control in Automated manufacturing

Marcel Dekker Inc., New York, xv + 294 pag., ISBN 0-8247-7889-8, $ 83.50.

Het boek is deel 15 uit de serie 'Quality and Reliability' van Marcel Dekker, onder editie van Edward Schilling.
In het voorwoord geeft Schilling aan dat het doel van de serie is, bij te dragen aan de effektieve integratie van de verschillende elementen in de moderne benadering van kwaliteit: statistiek, quality engineering, management en motivatie. De aangegeven doelgroep van de serie omvat vrijwel elke volwassen die kan lezen.

Het betreffende deel heeft als doelgroep de 'quality professional' en de toegepast statisticus.

Het boek is een verzameling artikelen, ondergebracht in 6 sekties:

I	Key Issues and Implementation Strategies
II	Time Series Applications in Statistical Process Control
III	Innovative Techniques for Statistical Process Control
IV	Statsitical Databases for Process Control
V	Knowledge-Based Systems in Process Control
VI	Real-Time Machine Tool Control

Een groot deel van de artikelen is afkomstig van 'Statistical Process Control : Keeping Pace with Automated Manufacturing, a National Symposium', gehouden 6-7 november 1986.

Het grootste deel (in pagina's: 200, sekties I t/m III) is statistisch getint. Sektie VI valt uit de toon (het betreft 'Deterministic Control').

De basisgedachte die doorklinkt in de meeste artikelen is, dat door de verdergaande automatisering van produktie-processen (inklusief het meten) grote hoeveelheden gegevens on-line beschikbaar komen en geanalyseerd kunnen worden. Doordat er dan allerlei correlaties (in de tijd en tussen meerdere variabelen) optreden, zijn de klassieke quality control methoden niet meer bruikbaar. Als toegepast statisticus heb ik moeite met deze sterk data-gerichte benadering (in

tegenstelling tot probleem-gericht). De in het voorwoord als vanzelfsprekend geuite vooronderstelling '... data is information ...', kan ik dan ook niet onderschrijven.

Desalniettemin zijn naar mijn idee klassieke quality control technieken (Shewart Control Charts op basis i.i.d. variabelen) in een fabrikage-omgeving vaak ontoereikend, en het is met die achtergrond aardig alternatieve modellen te zien (sekties II en III).

Sektie V, met name 'An Expert System Tool for Real-Time Control', laat zien dat SPC een gebied is waar Expert Systemen goed toegepast kunnen worden. Het geschetste systeem sluit wat gedachtengang goed aan op de vorm waarin wij zelf bezig zijn SPC te introduceren.

Het gepresenteerde is allesbehalve 'self-contained', zodat voor eventuele toepassing van modellen en systemen teruggevallen moet worden op de verwijzingen. Dit maakt het als leerboek of naslagwerk ongeschikt.

Bovendien staan een aantal artikelen erg ver van de praktijk (o.a. betreffende Kalman-filters).
Storend vond ik verder dat in een aantal artikelen uitgangs-veronderstellingen wordt gedaan die de essentie van SPC lijken te missen: In 'A Multivariate Framework for Statistical Process Control' bijvoorbeeld, wordt de univariate situatie (signaleren met control chart(s) op één variabele) gegeneraliseerd tot de multivariate door uit te gaan van een toestand waarin het proces (statistisch) beheerst is én capable (afstand tussen specifikatiegrenzen is groter dan 6x de processpreiding).
Het is echter een belangrijk doel van de SPC het proces in zo'n toestand te brengen! Ik kan me dan ook niet aan de indruk onttrekken, dat hier en daar veronderstellingen gedaan zijn vanuit 'mathematical convenience' in plaats van praktische relevantie.

Ondanks het feit dat het een verzameling artikelen betreft, sluit de stijl goed op elkaar aan. Het boek ziet er ook goed verzorgd uit. Er zit wel enige overlap tussen artikelen, en de notatie (met name met betrekking tot ARIMA-modellen en Kalman-filter) is niet konsekwent, maar dat wordt nergens storend. (Zet?)foutjes in formules komen sporadisch voor.

Resumerend lijkt het me, met name voor de quality professional, een aardig boek om eens kennis te maken met alternatieven voor de traditionele Control Charts en ontwikkelingen met betrekking tot Expert Systems. De benadering (sterk data-gericht én hier en daar diskutabele ten aanzien van de invulling van SPC) spreekt me echter niet aan.

Kit Roes
Philips Components, Nijmegen

A.I. Khuri & J.A. Cornell

Response Surfaces, Design and Analyses

Marcel Dekker, Inc., New York and Basel, Milwaukee, 1987, xii + 405 pag., ISBN 0-8247-7653-4, $ 54.00.

In dit boek behandelen de auteurs verschillende aspekten van regressie-analyse met vrijwel uitsluitend kwantitatieve verklarende variabelen en van de konsekwenties voor het ontwerp van bijbehorende experimenten. Een en ander gebeurt op nogal wisselende nivo's van uitleg, veelal op dat van een degelijk leerboek, soms op een hinderlijk omslachtige of haast onnozele wijze, maar ook nogal eens in de vorm van de mededeling van onderzoeksresultaten zonder volledige bewijsvoering. Die laatste mist men in het ene geval niet erg, maar in andere gevallen zou men die toch wel verwacht hebben bij een op dat moment meer geavanceerd nivo van presentatie.

Het boek is bedoeld als een aanvulling op Surface Methodology door R.H. Myers (1976) en op Empirical Model Building and Response Surfaces door G.E.P. Box en N.R. Draper (1987).

Hoofdstuk 1 is een globale inleiding to het onderwerp, gekoppeld aan regressie-analyse, waarbij de regressiefunktie een veelterm in de verklarende variabelen of faktoren is, onder andere te bestuderen door contouren of nivokrommen, elk de verzameling van punten in de faktorruimte met een bepaalde funktiewaarde of responsie omvattend.
Het doel is tevens het soms sequent zoeken van de optimale kombinatie van faktoren, waarbij de responsie maximaal of minimaal is.

Hoofdstuk 2 bevat eerst een (overbodige) opsomming van stukjes matrixrekening (met een onwerkbare definitie voor de rang van een matrix) en kleinste-kwadraten theorie ten behoeve van schatting van regressie-coëfficiënten en bijbehorende t-toets of F-toetsen bij normaliteit der verstoringen.
Het feit dat een t-toets eenzijdig wordt uitgevoerd naar aanleiding van het gekonstateerde teken van de geschatte regressie-coefficient, dat de som van kwadraten voor gebrek aan aanpassing (b.v. aanwezigheid van tweede-graadstermen boven een eerste-graadsmodel), na het niet-verwerpen van zijn afwezigheid door een toets, zonder waarschuwing wordt samengevoegd met die voor puur toeval teneinde een nieuwe residuele variantie-schatting te verkrijgen, en dat een aangepaste waarde, dat is een schatting van de verwachting in een bepaald punt van de faktorruimte, als voorspelling of voorspeller wordt betiteld terwijl hun varianties wezenlijk verschillen, wekt dan wel verwondering.
Vervolgens wordt elke faktor gecentreerd en door schaalverandering van hetzelfde tweede moment voorzien zodat proefschema-kriteria als orthogonaliteit en roteerbaarheid kunnen worden geïntroduceerd als eigenschappen van de proefplanmomenten-matrix X'X/n.

In hoofdstuk 3 (eerste-graadsmodellen) en hoofdstuk 4 (tweede-graadsmodellen) komen deze kriteria aan bod. Het bewijs, dat orthogonaliteit (een diagonale momenten-matrix) bij eerste-graadsmodellen volgt uit het minimaliseren van de variantie van elke aangepaste waarde bij begrensde tweede momenten, is (voorspelbaar) onjuist: orthogonaliteit volgt echter wel (met een eenvoudig meetkundig argument) uit het minimaliseren van de varianties van de regressie-coëfficiënt-schatters bij begrensd tweede moment voor elke faktor. Bovendien worden die schatters dan tevens ongecorreleerd.
Als proefschema's van eerste-graadsmodellen verschijnen de (frakties van) 2^k faktoriële proeven, simplex-schema's, schema's volgens Plackett-Burman, alle eventueel met toegevoegde waarnemingen in het centrum van zulk een schema.

Bij tweede-graadsmodellen vereist roteerbaarheid, dat is gelijkheid van de variantie voor alle aangepaste waarden op een bol in de faktorruimte om het centrum van het schema, dat van het proefschema alle oneven momenten nul zijn, de tweede momenten alle gelijk aan elkaar en elk zuiver vierde moment driemaal zo groot als elk gemengd even moment van de vierde graad (elk paar faktoren tot de tweede macht).
De volledige bewijsvoering hiervoor ontbreekt en er wordt geen melding gemaakt van de betrekkelijkheid van de eis van roteerbaarheid in verband met de schaalkeuze van de faktoren. Het enige argument dat men zou kunnen aanvoeren is, dat de contouren van gelijke variantie voor aangepaste waarden in de ongeschaalde faktorruimte convex, want (hyper)ellips(oid)en zijn.
Bij invoering van zogenaamde orthogonale veeltermen kan een roteerbaar schema ook orthogonaal zijn.
Als proefschema's van tweede-graadsmodellen worden onder andere beschouwd de 3^k faktoriële en de centrale samengesteld schema's bestaande uit een fraktioneel 2^k faktorieel deel, een axiaal deel en een groep centrale punten. Door de keuze van het aantal centrale punten kan men roteerbaarheid bewerkstelligen, danwel de variabiliteit in de varianties van de aangepaste waarden in de meetpunten minimaliseren.

Verder komen aan de orde schema's volgens Box en Behnken, equiradiale schema's (d.w.z. de meetpunten liggen op tenminste twee bollen met het centrum van het schema als middelpunt en bij elk zo'n bol equidistant, b.v. een regelmatige n-hoek en een centraal punt), cylindrisch, danwel

asymmetrische roteerbare schema's, verzadigde schema's volgens Box en Draper, schema's met gelijkmatige verdeling over bolschillen en hybridische schema's.
Tenslotte wordt blokindeling orthogonaal met de interessante effekten behandeld.

Hoofdstuk 5 gaat over het vinden van een optimale faktorkombinatie. Bij eerste-graadsfunkties wordt gewerkt met de voorzichtig te hanteren methode van de steilste helling en het inrichten van telkens nieuwe experimenten in die richting, af te sluiten met experimenten gericht op tweede-graadsfunkties.
Bij tweede-graadsfunkties komt het aanwijzen van een stationair punt en het overgaan op de canonische vorm van de regressiefunktie tergend langzaam aan de orde. Over een betrouwbaarheidsgebied wordt voor het optimum niet gesproken. Als het optimum ver buiten het experimenteergebied ligt, zoekt men naar (voorwaardelijk) extremen of bollen rond het centrum van het experimenteergebied. Over de samenhang tussen zulke extreme waarden en de daarmee verbonden Lagrange-multiplikatoren wordt een viertal resultaten medegedeeld, die alleen gelden onder de niet-vermelde voorwaarde, dat de kwadratische vorm positief-definiet is.

Tenslotte wordt een alternatieve sequente procedure voor het bepalen van een extreem van een kontinue niet-gespecificeerde responsfunktie volgens Nelder en Mead uiteengezet, waarbij men telkens een hoekpunt van een simplex al experimenterend vervangt door een nieuw punt in een meer belovende richting. In het samenvattend stroomdiagram komen twee fouten voor.

In hoofdstuk 6 wordt de aandacht gericht op proefschema's waarbij onder de veronderstelling van een eerste-graadsmodel niet slechts de over het interessante gebied gemiddelde variantie geminimaliseerd wordt, maar ook (of tezamen met) het gemiddelde kwadraat van de onzuiverheid als gevolg van de mogelijke aanwezigheid van tweede-graadstermen, dus ook minimalisatie van het gemiddeld tweede moment van de onderhavige schattingsfouten. Het is dan wel curieus (pagina 215) over een in dat opzicht optimaal proefschema te spreken, als men de verhouding tussen variantie en kwadraat van onzuiverheid kiest. Insgelijks wordt aandacht besteed aan de veronderstelling van een tweede-graadsmodel bij mogelijke aanwezigheid van derde-graadstermen: in één geval gaat het om de keuze van de straal van een 2^k faktorieel schema, in een ander geval om die van een centraal samengesteld roteerbaar schema. Daarnaast worden twee alternatieve schattingsprocedures voor de kleinste-kwadratenmethode beschouwd, weer met het oog op minimalisatie van het bewuste gemiddelde tweede moment. De ene is een lineaire reductie van de kleinste-kwadraten-regressie-coëfficiënten bij het uitgebreidere alternatieve model, de andere de vermenigvuldiging van de regressie-coëfficiënten bij het veronderstelde model, met konstanten zo dat gemiddelde varianties van aangepaste waarden, danwel het gemiddeld kwadraat van de onzuiverheid, minimaal is. Ook al omdat enige kennis van de onbekende regressie-coefficienten nodig is, kunnen de voordelen van deze alternatieve procedures als akademisch en niet van praktisch belang worden aangemerkt.

Hoofdstuk 7 gaat over meervoudige responsies, dat wil zeggen, meer dan één te verklaren eigenschap per meetpunt met voor elke eigenschap in principe verschillende regressiefunkties, bijvoorbeeld veeltermen van verschillende graad.
Schatting van de regressie-coefficienten vereist hetzij een voorafgaande schatting van de covariantie-matrix tussen eigenschappen, hetzij minimalisering van een determinantkriterium.
Vooraf moet men met behulp van eigenwaardenonderzoek nagaan of er lineaire afhankelijkheden tussen de eigenschappen voorkomen, en zo ja, welke dat zijn, opdat vervolgens overbodige eigenschappen geschrapt kunnen worden.
Voor een optimaal proefschema bij niet-lineaire regressiemodellen streeft men naar maximalisatie van de determinant van de informatie-matrix door additionele waarnemingen.
Bij lineaire modellen zoekt men het in D-optimaliteit en wel met een sequente uitbreiding van een beginschema met telkens een nieuw waarnemingspunt volgens de methode Fedorov, waarbij het spoor van de covariantie-matrix van aangepaste waarden, voorvermenigvuldigd met de inverse van de (geschatte) covariantie-matrix tussen de eigenschappen wordt gemaximaliseerd over het experimenteergebied, terwijl dit maximum convergeert naar het totaal aantal te schatten parameters.

Bij toetsing op gebrek aan aanpassing bij lineaire regressiemodellen komt men bij Roy's toets met de grootste eigenwaarde in MANOVA terecht; het vaststellen van eigenschap(pen) met gebrek aan aanpassing, geschiedt door net zo'n toets op alle deelverzamelingen van eigenschappen uit te voeren, met dezelfde kritieke waarde als bij het geheel.
Het vinden van een optimale kombinatie van de faktoren voor alle eigenschappen tezamen vereist kompromissen. Bij twee eigenschappen, beide met een kwadratische regressiefunktie, is het mogelijk informatie te verkrijgen over stationaire punten van één primaire eigenschap, gegeven de waarde van een sekundaire. Door het uitzetten van de tweede eigenschap tegen de optimale waarde van elk der faktoren, komt men tot optimale kondities voor de eerste eigenschap. Men kan ook aan elke eigenschap een waardering of een gewicht toekennen en vervolgens voor lineaire kombinatie met die gewichten als coëfficiënten het optimum bepalen. Verder kan men met een geschikte afstandsdefinitie het punt in de faktorruimte bepalen waarvoor de meervoudige responsie het dichtst bij de vektor van afzonderlijke extremen ligt. De Mahalanobis-afstand is een voorbeeld van zo'n afstand, waarbij de inverse van de covariantie-matrix van een meervoudige responsievektor wordt gebruikt.
Omdat de afzonderlijke optima zelf ook stochastisch zijn, wordt ook wel eerst een rechthoekig betrouwbaarheidsgebied voor de vektor van extremen geconstrueerd en bij elk punt in de faktorruimte het maximum van de bewuste afstand over dit betrouwbaarheidgebied bepaald. Dat maximum wordt tenslotte geminimaliseerd over het relevante deel van de faktorruimte.

In hoofdstuk 8 komen niet-lineaire regressiemodellen aan de orde, in het bijzonder de kleinste-kwadratenschatting en bijbehorende toetsingmethoden, gevolgd door lokaal D-optimale proefplankonstruktie met evenveel waarnemingen als parameters door middel van maximalisatie van de informatie-matrix. Daarbij komt een modifikatie bij belangstelling voor een deel van de parameters en bij partieel niet-lineaire modellen.
Tenslotte wordt melding gemaakt van parametervrije proefplannen bij niet-lineaire modellen in één verklarende variabele op basis van de benadering van de regressiefunktie door een Lagrange-interpolatie-veelterm waarvan de graad r hoger is dan het aantal parameters. De keuze van de r+1 interpolatiepunten is doorslaggevend voor de kwaliteit van zo'n benadering.
Bij de keuze van zogenaamde Chebychev-punten wordt de bovengrens van de benaderingsfout, een funktie van de gekozen graad r, geminimaliseerd en daarmee wordt Lagrange-interpolatie in de praktijk haast even efficiënt als die met de best benaderende veelterm, die moeilijk uitvoerbaar is. Het proefplan richt zich nu op de Lagrange-benadering van de onbekende regressiefunktie, een zogenaamde partieel niet-lineair model. De vektor van partiële afgeleiden bestaat uit lineaire kombinaties van de Lagrange-veeltermwaarden met de waarden van de partiële afgeleiden van de onbekende regressiefunktie in de interpolatiepunten als coëfficiënten. Maximalisatie van de verwachting van de determinant van de hieruit volgende informatie-matrix, onafhankelijk van de parameterwaarden, wordt nu nagestreefd. Voor de wijze waarop dit dient te gebeuren, wordt verwezen naar computerprogrammatuur.

Hoofdstuk 9 is gewijd aan ontwerp en analyse van mengselproeven, dat wil zeggen, de waarde van elke verklarende variabele de fraktie in een mengsel is, dus per waarneming met som 1. De experimenteerruimte bestaat bij q faktoren uit een (q-1)-simplex. Het proefplan zal bestaan uit hetzij een simplexrooster, hetzij een simplex-centroïde. Bij een veelterm als responsiefunktie ontstaat door gebruik van de afhankelijkheid een vereenvoudigde symmetrische veelterm, waarin termen van zuiver even graad niet meer voorkomen.
Als men benedengrenzen aan de variabelen wil of moet stellen, wordt het expermenteergebied weer een simplex; door lineaire transformatie van de faktoren ontstaan pseudo-faktoren zoals in het geval zonder die grenzen. Worden er bovendien bovengrenzen gesteld, eventueel tezamen met ondergrenzen, dan wordt het experimenteer gebied de doorsnee van twee simplexen en dient men de hoekpunten daarvan vast te stellen, want men experimenteert tenminste in dergelijke hoekpunten, maar ook in zwaartepunten van randen of van zijvlakken en dergelijke.

In het laatste hoofdstuk 10 worden reeds frekwent gebruikte optimaliteitskriteria, zoals A-, D-, G- of F-optimaliteit, eindelijk gedefinieerd.
Vervolgen wordt aangegeven hoe bij een eerste-graadsmodel een proefschema kan worden geconstrueerd, waarbij de F-toets op alle regressie-coefficienten nul zijn, ook bij niet-normaliteit

der verstoringen (ten naaste bij) korrekt is. Als één of twee waarnemingen in een fraai bedoeld schema zouden ontbreken, ontstaat een aantasting van D-optimaliteit; zulk een aantasting kan men a priori minimaliseren. Wat vermeld wordt over de wijze van aanvullen van ontbrekende waarnemingen in een mooi schema en de konsekwenties daarvan, is verward en onjuist.
Als $Y=(Y_1'; Y_2')'$, waarin Y_2 de ontbrekende waarnemingen voorstelt, en de overeenkomstige $X=(X_1'; X_2')'$, zal men de (moeilijker) berekening van $b=(X_1'X_1)^{-1}X_1'Y_1$ kunnen vervangen door eerst $\hat{Y}_2$ te bepalen zodat $X_2(X'X)^{-1}X'(Y_1';\hat{Y}_2')'=\hat{Y}_2$, dat zijn lineaire vergelijkingen in de weinige elementen van $\hat{Y}_2$, en vervolgens b berekenen als $(X'X)^{-1}X'(Y_1';\hat{Y}_2')'$.

Maar daarmee wordt de optimaliteit van het schema niet gered. Met de opsomming van enige onderwerpen die nader onderzoek verdienen wordt het hoofdstuk in het boek besloten

Konklusie:

In dit boek komt veel interessants ter sprake, maar het laat ook wel het een en ander te wensen over.

L.C.A. Corsten

A.C. Cohen en B.J. Whitten

Parameter estimation in reliability and life span models

Marcel Dekker, Inc., New York, 1988, xv + 394 pag., ISBN 0-8247-7980-0, $ 83.50.

Dit boek behandelt het schatten van de parameters van een aantal scheve verdelingen uit zowel volledige als gecensureerde steekproeven.
Scheve verdelingen worden vaak als model gebruikt bij onderzoek naar levensduur, betrouwbaarheid, etcetera, maar vinden bijvoorbeeld ook toepassing in de economie om de verdeling van inkomens te modelleren.

Van de volgende verdelingen worden de parameters geschat:
Weibull (3-parameter), exponentieel (2-par), lognormaal (3-par), invers normaal (3-par), gamma (3-par), Rayleigh (2-par), Pareto (2-par), gegeneraliseerde gamma (4-par) en de extreme waardenverdeling (2-par).
Alle verdelingen worden uitgebreid besproken. Aan sommige verdelingen wordt een heel hoofdstuk besteed waarin zowel volledige als gecensureerde steekproeven ter sprake komen. Andere verdelingen (Weibull, exponentieel, extreme waarden) worden tezamen behandeld, maar in een hoofdstuk worden dan de volledige steekproeven behandeld, terwijl in een ander hoofdstuk het schatten van de parameters uit gecensureerde steekproeven de aandacht krijgt.

Telkens worden de momentenschatters (ME) en de maximum likelihoodschatters (MLE) besproken. Zolang er geen drempelparameter in het spel is, of de waarde ervan wordt bekend verondersteld, geven deze traditionele methoden goede resultaten. Maar als de drempel onbekend is, ontstaan er problemen die in het boek steeds duidelijk worden vermeld. De ME vertonen dan grote varianties en de MLE kunnen moeilijk of soms helemaal niet worden uitgerekend.
Als alternatief worden gemodificeerde momentenschatters (MME) voor volledige steekproeven, en gemodificeerde maximum likelihoodschatters (MMLE) voor gecensureerde steekproeven geïntroduceerd.

In vergelijking met de berekening van de ME wordt bij de MME het derde steekproefmoment vervangen door een funktie van de kleinste waarneming. De MME leveren, net als de ME, zuivere schatters voor de verwachting en variantie van de verdeling op, maar de varianties van de MME

zijn te vergelijken met die van de MLE. De MME kunnen altijd, en met behulp van de vele tabellen en grafieken in het boek, in vele gevallen eenvoudig met de hand worden uitgerekend. Volgens de schrijvers verdienen de MME de voorkeur boven de ME en de MLE voor de beschreven verdelingen indien de coefficient van scheefheid groter is dan 1.

Bij de MMLE wordt de afgeleide van de loglikelihoodfunktie naar de drempelparameter vervangen door een funktie van de kleinste waarneming. Over de eigenschappen van de MMLE wordt niet veel gezegd.

Ieder hoofdstuk wordt afgesloten met voorbeelden waarin de verschillende schatters worden uitgerekend en vergeleken. Het aantal voorbeelden is aan de magere kant.

In de appendix worden naast uitgebreide tabellen van de verdelingsfunkties van de gestandaardiseerde Weibull, gamma, lognormale, invers normale en gegeneraliseerde gammaverdeling, ook FORTRAN listings gegeven van procedures voor het berekenen van ME, MLE en MME voor de parameters van Weibull, lognormale, invers normale en gammaverdeling.

Door het hele boek heen staan vele verwijzingen en de lijst met referenties op het eind is dan ook vrij uitgebreid.
Drie bekende boeken op het gebied van 'Reliability and life time data' zijn:

1. Mann, Schafer en Singpurwalla, *Methods for statistical analysis of reliability and life time data*
2. Bain, *Statistical analysis of reliability and life-testing models: Theory and methods*
3. Lawless, *Statistical models and methods for lifetime data*

Dit boek hoort in het rijtje thuis, maar beperkt zicht tot het bepalen van puntschatters voor de parameters, terwijl in [1, 2 en 3] ook andere onderwerpen, zoals bijvoorbeeld betrouwbaarheidsintervallen en toetsen ter sprake komen.
In tegenstelling tot [1, 2 en 3] besteden Cohen en Whitten echter bijna alle aandacht aan het geval dat er een onbekende drempelparameter aanwezig is. Zij introduceren daarvoor de MME en MMLE.
Het boek is uitermate geschikt voor diegenen die geïnteresseerd zijn in het schatten van parameters voor scheve verdelingen met een onbekende drempelparameter. Het is specialistisch, maar goed leesbaar.

T. Smeulders
Fokker Aircraft BV

Paul E. Green, Frank J. Carmone, Jr. & Scott M. Smith

Multidimensional Scaling: concepts and applications

Allyn and Bacon, Boston, 1989, 407 pag., ISBN 0-205-11657-4, *f* 93.00.

Dit boek is een vernieuwde editie van een boek dat bijna 20 jaar geleden geschreven is door Green & Carmone (Multidimensional Scaling and Related Techniques in Marketing Analysis. Allyn & Bacon (1970)). Scott Smith is voornamelijk verantwoordelijk geweest voor de software toepassingen, die nu in de vernieuwde versie uitgebreid aan bod komen (Scott Smith heeft het PC-MDS pakket ontwikkeld). Alle in het boek besproken software toepassingen krijgt men door middel van twee diskettes aangeleverd.

Het doel van het boek is, volgens de auteurs, om (markt)onderzoekers, docenten en studenten kennis te laten maken met c.q. op de hoogte te laten blijven van meerdimensionale schaaltechnieken (hierna: MDS) en gerelateerde technieken. Het boek is volgens de auteurs niet-mathematisch van aard en gemakkelijk te begrijpen. De auteurs hebben niet de intentie om een 'state of the art' van het toegepaste MDS-veld te geven.

Het boek is onderverdeeld in delen.
Deel I is een inleidend deel. Zowel basis koncepten van MDS worden besproken als het gebruik van deze concepten in marketing.
Deel II, 'analytic techniques', bestaat uit 5 sekties. De belangrijkste onderwerpen die aan bod komen, zijn:

1) meettheorie
2) klassifikatie van data- en schalingstechnieken
3) analyse van gelijkenissen
4) analyse van preferenties
5) dimensie reducerende technieken (zoals principale componenten analyse en cluster analyse).

In deel III is een aantal artikelen opgenomen die de literatuur op het gebied van MDS, cluster analyse en correspondentie analyse bespreken. De sekties zijn geclusterd rond MDS, cluster analyse en correspondentie analyse. Opgenomen zijn artikelen van L.G. Cooper (A Review of Multidimensional Scaling in Marketing research), J.D. Carrol and P. Arabie (Multidimensional Scaling), G. Punj and D.W. Stewart (Cluster Analysis in marketing Research: Review and Suggestions for Application), P. Arabie, J.D. Carrol, W. DeSarbo and J. Wind (Overlapping Clustering: A New method for Product Positioning), D.L. Hoffman and G.R. Franke (Correspondence Analysis: Grafical Representation of Categorical Data in Marketing Research), J.D. Carrol, P.E. Green and C.M. Schaffer (Comparing Interpoint Distances in Correspondence Analysis: A Clarification).
Het laatste deel van het boek (deel IV) heeft betrekking op software toepassingen en interpretatie. De volgende programma's en algoritmen worden behandeld: MDPREF, INDSCAL, PREFMAP, PROFIT, Howard-Harris Cluster Analysis, Correspondence Analysis: A mathematical Algorithm en KYST.

Paul Green heeft de afgelopen 20 jaar enorm veel gepubliceerd op het gebied van 'marketing' en 'marketing research'.
De voorbeelden die in het boek gebruikt worden, hebben dan ook allemaal betrekking op de marktkunde en het marktonderzoek. Mijns inziens had dit specifieke toepassingsveld best tot uitdrukking in de titel van het boek mogen komen. Dit geldt ook met betrekking tot de 'gerelateerde technieken'. De toevoeging 20 jaar geleden '...and related techniques in marketing analysis' is ook nu nog aktueel, daar ongeveer 1/3 van het boek hieraan gewijd is.

Een ander algemeen punt van kritiek is, dat de Gifi-modellen (bv. HOMALS en PRINCALS) niet genoemd of behandeld worden. Met name HOMALS (homogeniteitsanalyse of dual scaling) en PRINCALS (kan gezien worden als intern vektor ontvouwingsmodel) zijn de afgelopen 15 jaar op talloze onderzoeksterreinen toegepast, waaronder ook dat van marketing.

Tot slot wat meer specifieke opmerkingen. De eerste opmerking betreft de niet-mathematische en eenvoudig te begrijpen wijze waarop het boek geschreven is. Voor de delen I, II en IV gaat dit op, maar deel III kan nauwelijks gelezen worden zonder enige kennis van matrix algebra. Met name de sekties over 'overlapping clustering' en 'correspondence analysis' maken in ruime mate gebruik van matrix notatie.
Daarnaast bevat het boek een aantal slordigheden. Bijvoorbeeld bij de bespreking van het INDSCAL-model laat men na het ALSCAL-programma (welk een niet-metrische uitbreiding van het INDSCAL-model als optie bevat) te bespreken.
Verder is de notatie vaak inconsistent.

Ondanks deze kritiek blijft dit boek zeker de moeite waard om te bestuderen, zeker voor onderzoekers werkzaam op het marketing en marktonderzoek terrein. In de eerste plaats vanwege de wijze waarop de relevantie van MDS voor het probleemgebied marketing duidelijk gemaakt wordt, in de tweede plaats vanwege de PC-software besprekingen die in het boek aan de orde komen.

Drs Marco Vriens
Faculteit der Economische Wetenschappen
R.U. Groningen

Suddhendu Biswas

Stochastic Processes in Demography and Applications

John Wiley & Sons, New York, 1988, 359 blz. ISBN 0-470-21046, £ 20.15

Dit is een boek met veel formules en weinig woorden. Hoewel de hoeveelheid tekst zeer beperkt is, bevat deze zeer veel taalfouten. In het bijzonder het gebruik van lidwoorden en de interpunktie laat veel te wensen over. De afleidingen worden in extenso weergegeven en zijn over het algemeen gemakkelijk te volgen. Er wordt echter bijzonder weinig aandacht besteed aan de interpretatie van de formules. Tal van relaties worden afgeleid (zoals tussen verschillende overlevingstafelfunkties en tussen verschillende vruchtbaarheidsmaten) zonder dat wordt aangegeven waarom de afgeleide vergelijkingen interessant zouden zijn en hoe de uitkomsten kunnen worden geïnterpreteerd. Het boek lijkt daardoor meer geschikt als naslagwerk dan als studieboek.

Na een uitgebreid inleidend hoofdstuk over stochastische processen (88 pagina's) volgen hoofdstukken over sterfte, geboorte en bevolkingsgroei (respektievelijk 73, 37 en 96 pagina's). Deze hoofdstukken bevatten over het algemeen oud en bekend materiaal (de meeste resultaten zijn bekend sinds de jaren vijftig en zestig of nog eerder), maar dan besproken uitgaande van een stochastische benadering.
Daarna volgt een hoofdstuk over 'competing risks' (42 pagina's) met meer recent materiaal (het lange inleidende hoofdstuk is vooral als achtergrond hiervoor bedoeld) en een kort slothoofdstuk over dataproblemen (13 pagina's).
Er wordt geen aandacht besteed aan migratie.

Het hoofdstuk over sterfte behandelt onder andere sterftematen, standaardisatie, overlevingstafels (met bijzondere aandacht voor het schatten van overlevingstafelsfunkties op grond van steekproefdata) en curve fitting.
Het hoofdstuk over vruchtbaarheid bespreekt de relatie tussen verschillende vruchtbaarheidsmaten en vruchtbaarheidsmodellen van Dandekar, Brass en Perrin en Ships.
In het hoofdstuk over bevolkingsgroei wordt aandacht besteed aan het logistische groeimodel, bevolkingsprognoses (uitgaande van konstante vruchtbaarheid en sterfte), eigenschappen van quasi-stabiele bevolkingen en modellen voor geboorte- en sterfteprocessen.
Het hoofdstuk over 'competing risks' behandelt de relatie tussen bruto en nettokansen, afhankelijke en onafhankelijke risiko's, gecensureerde data en parametrische overlevingsmodellen.
Het slothoofdstuk handelt over Indiase demografie en besteedt vooral aandacht aan methoden om geboorte- en sterftecijfers af te leiden uit volkstellingsgegevens.

Joop de Beer
CBS

Samprot Chatterjee en Ali S. Hadi

Sensitivity Analysis in Linear Regression

John Wiley & Sons, New York, 1988, xiv + 315 pag., ISBN 0-471-82216-7, £ 25.95.

Dat ongebreideld gebruik van regressietechnieken gevaarlijk is, en hoe 'regression diagnostics' kunnen worden gebruikt om het al te stom gebruik van die technieken in te tomen, is in de 70-er jaren bekend geworden.
De boeken van Belsley, Kuh en Welsch (1980) en Cook en Weisberg (1982) markeerden de afsluiting van de beginfase van het onderzoek naar regression diagnostics.

Het hier besproken boek van Chatterjee en Hadi geeft een recente weergave van de stand van zaken wat betreft gevoeligheid van resultaten van lineaire regressiemodellen voor verschillende fouten in de modelspecifikatie. Daarbij blijft het bijna geheel binnen het kader van de kleinste kwadraten-benadering en de mogelijkheden die matrixbenaderingen daarbij bieden. Dit kan worden geïllustreerd door de grote rol die gespeeld wordt door de 'hat matrix', in dit boek 'prediction matrix' genoemd; dit is het woord met de meeste verwijzingen in de index.

Na het inleidende hoofdstuk 1 is hoofdstuk 2 geheel aan de prediction matrix gewijd.
Hoofdstuk 3 gaat over misspecifikatie, met name het opnemen van te veel of te weinig regressoren.
Hoofdstuk 4 is het langste van het boek (het beslaat meer dan 1/3 van het aantal bladzijden) en betreft het effekt van afzonderlijke observaties op de modelschatting. Het behandelt adekwaat en enigszins encyclopedisch vele maatstaven en grafieken ter bestudering van het effekt van het weglaten van afzonderlijke observaties. Dat encyclopedische heeft zijn voor en tegens: het maakt het boek minder spannend dan de genoemde twee boeken en dan bijvoorbeeld Atkinson (1985), maar wat betreft dit onderwerp wordt er heel veel behandeld en worden allerlei maatjes en plaatjes op verhelderende wijze met elkaar in verband gebracht.
Het effekt van de afzonderlijke observaties op collineariteitsproblemen is een relevant maar meestal onderbelicht onderwerp dat hier helder wordt besproken.
In hoofdstuk 5 worden de technieken uit hoofdstuk 4 gegeneraliseerd naar de effekten van meerdere observaties tegelijk. Uit encyclopedische overwegingen is dit nuttig; de praktische waarde van deze generalisatie is echter beperkt, omdat het kontroleren van het effekt van alle k-tallen observaties bij een totaal van n observaties leidt tot het nagaan van alle $n!/(n-k)!k!$ mogelijkheden. Tenzij k klein is, is dit aantal zo groot dat deze benadering uit de oogpunten van de hoeveelheid berekeningen en van kanskapitalisatie minder aantrekkelijk is (het behandelde voorbeeld is met $k=2$).
De benadering via de robuuste regressie (in de geest van bijvoorbeeld Rousseeuw en Leroy (1987)) vind ik vanuit praktisch standpunt veel aantrekkelijker, maar valt niet binnen het kader van dit boek.

Hoofdstuk 6 is gewijd aan het effekt van afzonderlijke observaties op afzonderlijke parameterschattingen; hier worden enkele recent door de auteurs ontwikkelde technieken behandeld, die goed bruikbare regression diagnostics zijn bij modelselektie. Voorbeelden van hier gebruikte begrippen zijn 'partial leverage', de leverage van een observatie voor het toevoegen van een regressor, en de toename van de residuele kwadratensom ten gevolge van het tegelijkertijd weglaten van een observatie en een regressor.

Hoofdstuk 7 gaat over meetfouten in de regressoren, vooral over een interessante en goed bruikbare perturbatie-benadering om bovengrenzen aan te geven voor het effekt van dergelijke meetfouten op de parameterschattingen.
In hoofdstuk 8 tenslotte wordt modelgevoeligheid voor gegeneraliseerde lineaire modellen (GLMs) behandeld. Het is een tamelijk flauw hoofdstuk, want het bevat een introduktie tot GLMs

en een uitgebreid voorbeeld met een vergelijking van meerdere modelspecifikaties voor de betreffende dataset. Er wordt geen overzicht gegeven van bestaande diagnostische technieken voor GLMs.

De uitvoerige behandeling van het effekt van weglating van observaties op parameterschattingen en de in het voorafgaande vermelde behandeling van een aantal minder bekende technieken maken dit boek tot een nuttige aanvulling op de bestaande literatuur over regression diagnostics. Aanbevolen voor de bibliotheek, maar de zakelijke en weinig stimulerende stijl en het ontbreken van enkele onderwerpen zoals transformaties en autokorrelatie maken, dat ik het niet zou aanraden als enige tekst bij een kursus.

Referenties:

A.C. Atkinson (1985). *Plots, transformation and regression: An introduction to graphical methods of diagnostic regression analysis.* Oxford: Clarendon Press.

D.A. Belsley, E. Kuh en R.E. Welsch (1980). *Regression diagnostics: Identifying influential data and sources of collinearity.* New York: John Wiley & Sons.

R.D. Cook en S. Weisberg (1980). *Residuals and influence in regression.* London: Chapman & Hall.

P.J. Rousseeuw an A.M. Leroy (1987). *Robust regression and outlier detection.* New York: John Wiley & Sons.

Tom A.B. Snijders
RU Groningen
RU Utrecht

E. Casillo

Extreme Value Theory in Engineering

Academic Press Inc., Boston, 1988, 389 xv + pag., ISBN 0-12-163475-2, $ 59.95.

Inhoud

De titel mag geïnterpreteerd worden als een poging mathematische waarschijnlijkheidsrekenaars en ingenieurs belangstelling en begrip voor elkaars werk bij te brengen. Enerzijds bevat het boek real live datasets en programmatuur opdat mathematici in indruk krijgen zich met ook praktisch relevante zaken bezig te houden; anderzijds vindt de wiskundig geschoolde ingenieur een overzichtelijke presentatie van de stand van zaken met betrekking tot asymptotisch gedrag van order statistics in steekproeven van i.i.d. data en in hoofdstuk 8 van i.d. data.
Slechts incidenteel wordt aandacht aan bewijzen geschonken. Peak over threshold modellen krijgen geen aandacht; dit onderwerp zou bovenaan dienen te staan bij uitbreiding van de stof.

Voor personen met meer oog voor de schoonheid der wiskunde worden de volgende twee redelijk recente boeken genoemd, namelijk R.-D. Reiss (1988). *Approximate Distributions of Order Statistics (with Applications to Nonparameric Statistics)*. Springer-Verlag, New York ..., p xii + 355 en J. Galambos (1987). *The Asymptotic Theory of Extreme Order Statistics, 2nd edition*. Malabar, Florida: Krieger.

Een recensie (door J.F. Lawless) van Castillo's boek is te vinden in Short Book Reviews ISI, vol. 9 No. 2 - August 1989, p. 27.

Het boek is geschikt als studieboek en als verwijsboek; het lijkt vooral nuttig voor wiskundig ingenieurs die zich met veiligheid, betrouwbaarheid van systemen, en dergelijke bezig (moeten) houden.

M.A.J. van Montfort
Lanbouwuniversiteit Wageningen

D.M. Bates & D.G. Watts

Nonlinear Regression Analysis and Its Applications.

John Wiley & Sons, New York, 1988, xiv + 365 pag., ISBN 0-471-81643-4, £ 34.50.

Het eerste hoofdstuk geeft een overzicht van lineaire regressie, waarbij de nadruk ligt op de geometrische interpretatie van de kleinste kwadraten methode en op een efficiënte methode om deze op de computer te implementeren (via QR decompositie van de design matrix).
Het tweede hoofdstuk bouwt hierop voort door behandeling van de iteratieve toepassing van het kleinste kwadraten principe bij niet-lineaire modellen. Ook hier wordt veel aandacht besteed aan de geometrische interpretatie.
Hierna volgt een luchtiger hoofdstuk met praktische overwegingen omtrent het vinden van startwaarden, parameter transformaties, alternatieve schattingsmethoden, convergentiecriteria, etcetera.
Vervolgens komen aan de orde multivariatie niet-lineaire regressie, compartimentsmodellen, betrouwbaarheidsgebieden en andere afbeeldingen van de parameterschattingen met hun onzekerheid (profile t plots e.d.). Het boek sluit af met een hoofdstuk over niet-lineariteitsmaten, het specialisme van de auteurs. Het is ruim voorzien van verhelderende illustraties en voorbeelden, en bevat een interessante collectie data. In de Appendix wordt pseudocode voor de besproken algoritmes gegeven. Bestaande programmatuur voor niet-lineaire regressie blijft onbehandeld.

De auteurs hebben het boek geschreven met het oog op de toepassing (met name in de experimentele setting) en schuwen het geven van praktische, op ervaring gebaseerde tips niet.
De volgorde van onderwerpen is soms wat onhandig waardoor nogal eens verwezen moet worden naar latere hoofdstukken. Desondanks is het een uitstekend leesbaar boek. De lezer wordt bekend verondersteld met lineaire regressie (niveau Draper en Smith, 1981) en met matrix algebra. Het benodigde wiskunde-niveau ligt wat hoger dan wat nodig is voor het boek van Draper en Smith.

Een zo grondig boek, dat behandeling van bijvoorbeeld multivariatie niet-lineaire regressie met missende waarnemingen niet uit de weg gaat, had mijns inziens ook wat ruimte kunnen geven aan andere dan kleinste kwadraten methoden c.q. de normale verdeling. De brug tussen het niet-lineaire en het gegeneraliseerd lineaire circuit is blijkbaar nog niet geslagen.
Het onderwerp keuze tussen modellen, dat hier wordt afgedaan met het extra kwadratensom principe, had misschien eveneens wat ruimere aandacht verdiend.

Deze kleine punten van kritiek nemen niet weg dat ik het boek van harte aan kan bevelen aan mensen die behoefte hebben aan een up to date behandeling van het onderwerp.

Referentie

Draper, N.R. and Smith, H. (1981). *Applied Regression Analysis. Second Edition.* John Wiley & Sons, New York.

Janneke Hoekstra
Centrum voor Wiskundige Methoden
Rijksinstituut voor Volksgezondheid en Milieuhygiëne