Boekbesprekingen

Redactie

Alex J. Koning

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

Besproken boeken in Kwantitatieve Methoden nr. 64

Venables, W.N., Ripley, B.D.
Modern applied statistics with S-PLUS

Waller, D.L.
Operations management. A supply chain approach

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

Billingsley, P.
Convergence of probability measures, 2nd edition

Durrett, R.
Essentials of stochastic processes

Loader, C.
Local regression and likelihood

Singpurwalla, N., Wilson, S.P.
Statistical methods in software engineering. Reliability and risk

Grätzer, G.
First steps in LaTeX

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

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

Does, R.J.M.M., Roes, K.C.B., Trip, A.
Statistical process control in industry

Morrison Paul, C.J.
Cost structure and the measurement of economic performance. Productivity, utilization, cost economics, and related performance indicators.

Lange, K.
Numerical analysis for statisticians

Zhang, H., Singer, B.
Recursive partitioning in the health sciences

Hutcheson, G.D., Sofroniou, N.
The multivariate social scientist. Introductory statistics using generalized linear models

Gerbing, D.
Relevant business statistics using Excel

Eldredge, D.
An Excel companion for business statistics

Albright, S., Winston, W., Zappe, C.
Data analysis and decision making with Microsoft Excel

Utts, J.M.
Seeing through statistics

van der Kloot, W.A.
Meerdimensionale schaaltechnieken

Venables, W.N., Ripley, B.D.
Modern applied statistics with S-PLUS
1999, Statistics and computing,
Springer-Verlag, Berlin,
xii+501 blz, DM 129.00, ISBN 0-387-98825-4.

Deze derde editie van Modern Applied Statistics with S-PLUS is weer een fraaie uitgave en tevens vastlegging van de voortdurende on-line bijwerkingen die van dit boek beschikbaar zijn via diverse webpagina's waaronder `http://www.stats.ox.ac.uk/pub/MASS3`. Verrassend, maar helaas niet verhelderend voor een eventuele aanschaf, is het ontbreken van een opsomming waarin de derde editie vernieuwd of verbeterd is t.o.v. de vorige editie. Wel wordt duidelijk aangegeven dat het boek gebaseerd is op de S-engine versie 3 voor de S-PLUS Windows versies 3.x en 4.x en op de S-engine 4 voor de S-PLUS 5 versie onder Unix. Dat na januari 1999 S-PLUS 2000 voor Windows is verschenen doet aan de bruikbaarheid van het boek niets af.

Venables en Ripley staan garant voor een gedegen beschrijving hoe je met S-PLUS (moderne) statistiek kunt bedrijven. Ze geven duidelijk aan dat het boek geen tekstboek over de theorie van de behandelde statistische methoden is, maar ieder hoofdstuk begint wel met een beknopt overzicht van de theorie, met literatuur verwijzingen en met voorbeelden. Naarmate de methodes moderner zijn worden vanaf hoofdstuk 7 de literatuuroverzichten langer.

Het boek is bestemd voor mensen die bereid zijn de S-PLUS commando taal te leren want alleen in Appendix B wordt expliciet op de nieuwe grafische user interface van S-PLUS 4.x ingegaan. Via de commando taal is S-PLUS namelijk volledig te benutten alsmede de vele programma's die bereidwillig door velen via http://lib.stat.cmu.edu/S/ beschikbaar worden gesteld, zodat de nieuwste technieken relatief snel en eenvoudig onder handbereik komen. Starters worden kernachtig in de taal ingeleid in de eerste vier hoofdstukken (resp. Inleiding, de S-taal, grafische mogelijkheden en programmeren is S) en in Appendix A wordt het inrichten van de directories en de verschillende mogelijkheden van het opstarten van S-PLUS behandeld. In hoofdstuk 6 wordt de S-PLUS taal verder uitgelegd voor wat betreft de uniforme schrijfwijze van formules in allerlei modellen. In hoofdstuk 5 wordt de univariate statistiek behandeld en daarbij komt al een van de pijlers van de moderne statistiek n S-PLUS naar voren: de grafische weergave van gegevens als basis voor verdere analyses of t.b.v. de evaluatie van analyse resultaten. In de hoofdstukken 6 t/m 8 worden resp. de lineaire modellen, GLM's en niet-lineaire modellen behandeld. Hierbij worden zeer bruikbare details over numerieke berekeningen met voor- en nadelen gegeven. Naast de traditionele lineaire regressie technieken worden ook robuuste en resistente regressie (kort) uitgelegd met hun implementatie in S-PLUS. In hoofdstuk 9 komen de moderne regressie methoden aan de beurt o.a. additieve modellen met scatterplot smoothers, projection persuit, neurale netwerken. Vervolgens komen in hoofdstuk 10 de classificatie en regressieboom technieken aan de beurt, waarbij naast de S-PLUS

functies ook de Statlib bibliotheek rpart beschreven wordt. Multivariate analyses en patroonherkenning komen in hoofdstuk 11 aan de orde via grafische methoden, cluster en discriminant analyses en verschillende classificatie technieken. In hoofdstuk 12 wordt zeer uitgebreid op survival analyse ingegaan en in hoofdstuk 13 op tijdreeksanalyse. Als laatste wordt in hoofdstuk 14 kort de ruimtelijke statistiek behandeld via de Statlib bibliotheek spatial. De S-PLUS module S+SpatialStats wordt niet besproken (zie KM 58, Review of the statistical package S+SpatialStats door A. Stein, pag. 139-145).

Tenslotte dient te worden opgemerkt dat de auteurs er uitstekend in zijn geslaagd volledige en gedetailleerde informatie te combineren met een prettige en vlotte manier van schrijven. Door de enorme uitgebreidheid en openheid van S-PLUS blijven er natuurlijk altijd vragen bestaan. Dat zelfs dit boek niet allesomvattend is, is echter helemaal geen probleem, want in Appendix C geven de auteurs naast de nodige web pagina's ook hun eigen e-mail adres. Gezien het verkeer op de wereldwijde S-PLUS nieuwsgroep is het duidelijk dat men bijna altijd binnen 24 uur een antwoord van de auteurs krijgt!

dr A.L.M. Dekkers
Rijksinstituut voor Volksgezondheid en Milieu
Bilthoven

Waller, D.L.
Operations management. A supply chain approach
1999, International Thomson Busines Press, London,
xxii+841 blz, £ 26.95, ISBN 1-86152-415-3.

De schrijver van het boek, Derek L. Waller, doceert aan E.M. Lyon, een management school in Frankrijk. Zijn verdere ervaring is gebaseerd op docentschap aan universiteiten in de V.S. en deelname aan projecten in de industrie.

Wallers boek is bedoeld voor gebruik in het onderwijs. Aan het eind van elk hoofdstuk staan een aantal review- en discussievragen. Verder zijn er veel opgaven en casussen opgenomen. Een boek met antwoorden is apart aan te schaffen door docenten. Het boek kent een duidelijke opbouw. Het inleidende hoofdstuk waarin operations management neergezet wordt, wordt gevolgd door vier delen, te weten:

I Strategische beslissingen en bedrijfsactiviteiten

II Ontwerp in operations management

III Planning, organisatie en controle

IV Verdere analyse

Deel I en II zijn vooral kwalitatief van aard, de kwantitatieve methoden komen vooral in deel III en IV aan de orde. Aan het begin van ieder deel staat in een kort stukje de

inhoud van het betreffende deel beschreven en de samenhang tussen de hoofdstukken, verduidelijkt met een figuur. Aan het begin van elk hoofdstuk worden de doelstellingen met hun samenhang kort aangegeven. Elk hoofdstuk eindigt met een samenvatting met daarin de belangrijkste begrippen. In het hele boek staan veel verduidelijkende plaatjes, schema's en grafieken. Berekeningen worden gellustreerd aan de hand van uitgewerkte voorbeelden. De voorbeelden en casussen zijn duidelijk van de tekst onderscheiden door een andere achtergrondkleur. Achter in het boek is een uitgebreide verklarende woordenlijst opgenomen en een index.

De supply chain benadering die de titel belooft, blijkt de supply chain IN een bedrijf te zijn. De supply chain wordt in dit boek gedefinieerd door input - transformatie - output. Het gaat om de input van grondstoffen en/of halffabrikaten, en de output van eindproducten voor de klant. De voorbeelden behandelen vooral bedrijven met toeleveranciers, de bedrijven leveren over het algemeen direct aan de consument. Er wordt in het boek geen aandacht besteed aan een keten die bestaat uit meer dan twee bedrijven en de (afstemmings)problemen die daarbij kunnen optreden. Eenmaal wordt een voorbeeld uit de retail-voedingsindustrie gegeven, waarbij het gaat om vijf schakels. De schakels behoren weer tot twee bedrijven. Ik vind het jammer dat de supply chain niet wat ruimer is opgevat. Dat is namelijk een veel voorkomende situatie, waar veel over te zeggen valt.

De kwantitatieve methoden in het boek worden wat mager gepresenteerd. Hierna volgen enkele voorbeelden daarvan.

- In het hoofdstuk over locatiekeuze (H3) worden vooral de kwalitatieve aspecten die daarbij van belang zijn behandeld. De kwantitatieve benadering bestaat uit een viertal heuristische methoden. Locatie/allocatie met behulp van lineaire programmering was een goede aanvulling geweest met wat meer diepgang.

- Forecasting wordt in twee hoofdstukken behandeld (H9 en H21). Niet elk model wordt even duidelijk behandeld. De (Ordinairy) Least Squares methode wordt uitgewerkt met behulp van Microsoft Excel, maar niet uitgelegd. Ook deel IV (H21) waar forecasting wat uitgebreider wordt behandeld, bevat enkele formules, die zonder afleidingen gepresenteerd worden.

- Lineaire Programmering wordt genoemd bij een voorbeeld van een transportprobleem en wordt opgelost met de Solver uit Excel. In deel IV wordt LP uitgebreider beschreven. Daar staat dat LP-problemen worden opgelost door softwarepakketten "die gebruikt kunnen worden zonder de vervelende wiskundige gymnastiekoefeningen te begrijpen die nodig zijn om LP-problemen op te lossen"! Niet direct de aanpak die wij zouden prefereren! Problemen met twee variabelen worden grafisch aangepakt.

- Er is ook een hoofdstuk besteedt aan scheduling (H14). Daarin wordt een goede uitleg gegeven wanneer welke techniek gebruikt kan worden. Soms is de uitleg enigszins omslachtig. Een theoretisch kader ontbreekt.

In het boek zijn twee hoofdstukken gewijd aan statistiek. In vergelijking met de overige kwantitatieve onderwerpen heeft dat ruime aandacht. Voorraadbeheer en wachtrijen worden goed behandeld. De uitgelichte onderwerpen vormen slechts een greep uit het totale aanbod in het boek.

Het boek is zeer basaal, met erg veel voorbeelden, het is daarom geschikt voor studenten die nog geen of weinig voorkennis op het gebied van de bedrijfseconomie hebben. Voor studenten met voorkennis is het boek misschien wel overduidelijk. Naar mijn mening valt het kwantitatieve niveau van het boek tegen. Er wordt goed aangegeven wanneer men welke kwantitatieve methoden gebruikt, maar met de behandelde stof mag bij de meeste onderwerpen niet verwacht worden dat studenten het in de praktijk gaan gebruiken. Ze weten hooguit waar het over gaat als iemand anders iets berekend heeft.

Het boek is een geschikt studieboek en past binnen de bedrijfseconomie of bedrijfskunde. Het kan gebruikt worden door studenten in die studierichtingen, of in andere richtingen waar enige kennis van die vakken vereist is. Voor de wat meer kwantitatief ingestelde student geeft dit boek wel een kader waarbinnen je verschillende kwantitatieve methoden gebruikt. De diepgang moet dan uit andere boeken geleerd worden. Volgens de auteur is het boek geschreven voor studenten die operations management en/of supply chain management vakken volgen in universitaire business- of engineering programma's.

K.G.J. Pauls-Worm
Leerstoelgroep Operationele Research en Logistiek
Wageningen Universiteit

Everitt, B.S.
Chance rules. An informal guide to probability, risk and statistics
1999, Springer-Verlag, Berlin,
202 blz, DM 49.00, ISBN 0-387-98768-1.

Via de VVS
25% korting!

Het boek begint met een korte uiteenzetting over de rol die het toeval (chance) in de geschiedenis heeft gespeeld. De onderwerpen varieren van de oorsprong van het dobbelen tot de houding van diverse volkeren uit de "oudheid" ten aanzien van het toeval. Verder wordt aandacht geschonken aan mannen als Gallileo, De Witt, Bernouilli, Pascal, Huygens en Graunt, die als eersten getracht hebben om het toeval een enigszins theoretische fundering te geven.

Na deze korte historische inleiding volgen enkele meer praktische hoofdstukken waarin de onderstaande onderwerpen, aan bod komen:

- tossing coins and having babies
- rolling dice
- gambling for fun: lotteries and football pools

- serious gambling: roulette, cards and horse racing
- birthdays and coincidences
- conditional probability and the Reverend Thomas Bayes
- puzzling probabilities

In deze hoofdstukken ligt de nadruk op het uitrekenen van kansen. Dit alles aan de hand van een inleidend vraagstuk of probleemstelling met een uitgebreide uitwerking van de oplossing. Begrippen als (on)afhankelijkheid, conditionele kansen, het combineren van kansen, permutaties, combinaties, coincidenties, sensiviteit, specificiteit en de stelling van Bayes worden behandeld.

Na de uitleg over het uitrekenen van bepaalde kansen, volgt in het tweede deel van het boek een uiteenzetting over hoe kansen in het dagelijks leven geinterpreteerd worden en hoe ze eigenlijk geinterpreteerd dienen te worden. De volgende hoofdstukken behandelen achtereenvolgens:

- taking risks
- statistics, statisticians and medicine
- alternative therapies-Panaceas or Placebos?
- chance, chaos and chromosomes

Het hoofdstuk "taking risks" gaat nader in op de relatie tussen het werkelijke risico en de houding ten aanzien van risico's die activiteiten als b.v. roken, vliegen, huwen en eetgewoonten met zich meebrengen. De volgende twee hoofdstukken besteden aandacht aan statistiek in de geneeskunde. Na een inleidend historisch overzicht wordt de clinical trial, randomisatie en de placebo besproken. In het hoofdstuk over alternatieve therapieen wordt geprobeerd het "succes" van behandelingen als acupunctuur en homeopathie te verklaren. Bijgeloof, het placebo effect en het onvermogen van de reguliere geneeskunde worden als de belangrijkste argumenten aangevoerd. Kritisch wetenschappelijk onderzoek zal dan ook weinig kunnen veranderen aan de status die deze alternatieve therapieen heden ten dage hebben. Het laatste hoofdstuk is wat meer filosofisch getint en gaat in op de tegenstelling "deterministisch versus stochastisch" wereldbeeld. Aan de hand van de evolutie-theorie, de chaos theorie, de genetica en Heisenberg's onzekerheids principe tracht de auteur te illustreren dat ons wereldbeeld allerminst deterministisch is. Een allesomvattende theorie, zoals sommige wereldberoemde fysici voorstaan moeten we dan ook maar vergeten.

Volgens de uitgever is het boek vooral bestemd voor gebruikers en studenten. Uit de wijze waarop de onderwerpen behandeld zijn, valt af te leiden dat het boek vooral gericht is op studenten en gebruikers met weinig statistische kennis. Het boek is verder

helder geschreven en vormt met z'n historische achtergronden, gevarieerde onderwerpen en leuke voorbeelden een mooie inleiding tot het vakgebied. Het geheel is afgewerkt met leuke illustraties en enkele fraaie foto's waaronder een markante afbeelding van sir R.A. Fisher. Een minder sterk punt van dit boek is dat het weinig diepgaand en te verhalend is om het voor de wat meer ervaren statisticus echt interessant te maken.

M.H.J. de Bruijne
Capaciteitsgroep Medische Statistiek
Leids Universitair Medisch Centrum

Billingsley, P.
Convergence of probability measures, 2nd edition
1999, Wiley series in probability and statistics,
John Wiley & Sons Ltd., Chichester,
ix+277 blz, £ 65.00, ISBN 0-471-19745-9.

Een lastig concept in de mathematische statistiek is zwakke convergentie. Hoewel dit in veel takken van de statistiek terugkomt, is een gedegen maattheoretisch begrip een hele opgave. De eerste keer dat ik over zwakke convergentie leerde, was dat aan de hand van de eerste druk van Billingsley's Convergence of Probability Measures. Vorig jaar, 31 jaar na dato, is de tweede editie verschenen.

Een van de centrale stellingen in de zwakke convergentie is de stelling van Donsker. Deze zegt dat, onder zekere voorwaarden, een stochastisch process convergeert naar een bijna zeker continue Brownse beweging. Donsker's stelling is een generalisatie van de centrale limiet stelling en ontleent zijn kracht aan het aanknopingspunt dat het biedt voor een heel scala aan limietverdelingen. Denk bijvoorbeeld aan goodness-of-fit testen; wat is de limietverdeling van de Kolmogorov-Smirnov test, of van de de Anderson-Darling statistic?

Het boek kent de volgende 5 hoofdstukken:

1. Weak convergence in metric spaces. Over het begrip convergentie in de reeële ruimte. Convergentie in kans en in verdeling.

2. The space C. Over de ruimte van continue functies op [0,1]. Inleiding in de Brownse beweging en de stelling van Donsker.

3. The space D. Over de ruimte van rechtcontinue functies met limiet bestaand van links, d.w.z. functies met jumps zoals de empirische verdelingsfunctie. Metrisering van D met behulp van de Skorohod metriek en toepassing van de stelling van Donsker in D.

4. Dependent variables. Over martingalen en ergodische processen.

5. Other modes of convergence. Over convergentie in kans.

Het blijft in het vage waarom deze nieuwe druk uitgebracht is. Het voorwoord meldt slechts dat bepaalde delen herschreven of ingekort zijn en dat hier en daar recente technieken (d.w.z. minder dan 30 jaar oud) gebruikt zijn. Verder heeft het boek meer weg van een lesboek dan zijn dertig oudere voorganger die meer weg had van een onderzoeksboek. Verder zijn aan de meeste secties nu opgaven toegevoegd.

De doelgroep van het werk staat nergens expliciet vermeld maar het boek zal mikken op mathematisch statistici en wiskundigen. Dit is geen pocket die je op je vrije zondagmiddag even ter hand neemt; zonder een gedegen kennis vooraf van maattheorie is het boek niet leesbaar. Overigens bevat het een appendix met daarin allerhande resultaten uit de analyse en de maattheorie. Het boek is goed opgebouwd en helder geschreven voor de doelgroep. De belofte op de achterzijde van het boek "[...] applications [from] engineering, economics, and population biology" vind ik optimistisch; ze dienen met een vergrootglas gezocht te worden. Wel bevat het boek voldoende voorbeelden uit de analyse en de getaltheorie.

Billingsley is een ijzersterk wiskundige. Hij slaagt erin zijn kennis op een effectieve manier over te brengen en schrikt niet terug voor ingewikkelde maattheoretische redenering. Het boek is goed uitgebalanceerd en prettig leesbaar. Al met al een waardig opvolger van het dertig jaar oudere broertje.

C.A. Bernaards
Faculteit Sociale Wetenschappen
Rijksuniversiteit Utrecht

Durrett, R.
Essentials of stochastic processes
1999, Springer texts in statistics,
Springer-Verlag, Berlin,
305 blz, DM 134.00, ISBN 0-387-98836-X.

Via de VVS
25% korting!

Dit boek geeft een inleiding in de theorie der stochastische processen. Het is bedoeld voor lezers die een inleiding in de kansrekening hebben gehad, maar die niet bekend zijn met maattheorie. Na de herhaling van een aantal basisbegrippen uit de kansrekening komen achtereenvolgens de volgende stochastische processen aan bod: Markovketens in diskrete tijd, martingalen, Poisson processen, Markovketens in continue tijd, vernieuwingsprocessen en Brownse beweging. In het hele boek ligt de nadruk op de toepasbaarheid van de gepresenteerde theorie. Een zeer groot gedeelte van de tekst bestaat dan ook uit voorbeelden en opgaven. De antwoorden op de helft van de opgaven zijn achterin het boek te vinden.

Door de nadruk op toepassingen komen niet alle stochastische processen even uitvoerig aan bod. Markovketens worden bijvoorbeeld vrij uitgebreid behandeld en nemen ongeveer de helft van het boek in beslag. De martingaaltheorie blijft echter beperkt tot de stoptijdstelling, een nuttig gereedschap voor het doen van concrete berekeningen. Theoretische resultaten probeert de auteur onder het kopje 'Why is this true'

eerst aannemelijk te maken zonder in detail te treden. Daarna volgt dan in de meeste gevallen een echt 'Proof' waarin de details worden ingevuld. Ondanks de bescheiden eisen wat betreft voorkennis wordt op een aantal plaatsen toch met meer geavanceerde begrippen gewerkt zoals het conditioneren op eventualiteiten met kans nul, stoptijden en de sterke Markov eigenschap. Dit leidt soms tot wat schimmige definities en redeneringen. De auteur gebruikt dan ook zelf af en toe termen als 'hand-waving' en 'leap of faith'. Concrete voorbeelden moeten dit soort onduidelijkheden steeds ophelderen.

Het onderhavige boek lijkt me een aardige inleiding in de theorie der stochastische processen voor lezers die vooral geinteresseerd zijn in toepassingen en minder in de wiskundige achtergronden. Een nadeel is de grote hoeveelheid andere tikfouten en slordigheden, waardoor zeker bij een eerste kennismaking de duidelijkheid niet wordt bevorderd. De auteur voorziet dit probleem al in het voorwoord en belooft daar dat hij ontvangen typos op zijn website (`http://www.math.cornell.edu/~durrett`) zal zetten. Tot nu toe is daar voor dit boek echter nog geen lijst te vinden. Misschien hadden auteur en uitgever iets meer tijd in dit probleem kunnen steken voordat het boek gedrukt werd. Voor deze prijs mag je dat toch eigenlijk best verwachten.

H. van Zanten
Centrum voor Wiskunde en Informatica

Loader, C.
Local regression and likelihood
1999, Statistics and computing,
Springer-Verlag, Berlin,
305 blz, DM 129.00, ISBN 0-387-98775-4.

Lokale regressie is een niet-parametrische techniek waarbij een te schatten functie lokaal benaderd wordt met behulp van een polynoom van doorgaans lage graad en met onbekende, te schatten, coefficienten. De techniek is vooral bekend als niet-parametrische uitbreiding van het klassieke lineaire model en implementaties als Loess zijn beschikbaar in diverse statistische pakketen.

Het idee is echter breder toepasbaar en Clive Loader introduceert in 'Local Regression and Likelihood' lokale regressie methoden als niet-parametrische uitbreiding van een aantal lineaire technieken. Hij realiseert binnen deze brede opzet een prettig leesbaar boek van, toch, beperkte omvang. Het boek wordt zinvol aangevuld met `LOCFIT`, te downloaden van `http://cm.bell-labs.com/stat/project/locfit/`. `LOCFIT` is een mooi verzorgd programma voor het ontwikkelen van lokale regressie en likelihood modellen, en implementeert de schattingsalgoritmen en het relevante diagnostische gereedschap. `LOCFIT` is beschikbaar als stand-alone programma onder UNIX en Windows en als bibliotheek onder R en S-plus.

De kern van het boek wordt gevormd door de hoofdstukken over lokale regressie als niet-parametrische uitbreiding van, dan wel alternatief voor, een aantal bekende statistische technieken: het klassieke lineaire model, het gegeneraliseerd lineaire model, dichtheid-schatters, overlevingsduur modellen en discriminatie en classificatie. De opzet van deze hoofdstukken is vrij uniform: introductie van de lokale likelihood vergelijkingen, een voorbeeld (uitgewerkt met `LOCFIT`), praktische aspecten van model selectie met een nadruk op AIC, kruis-validatie, vrijheidgraden en 'Influence' functies, gevolgd door 'enige' theorie. Bij de theorie ligt de nadruk meer op intuitie dan op preciese wiskunde.

De 'consistente' opzet van deze hoofdstukken werkt goed in didactische zin. Er zijn echter ook onderwerpen die niet binnen ieder van de behandelde modellen al even uitvoerig onderzocht zijn. Deze onderwerpen komen in aparte hoofdstukken aan de orde. Specifiek binnen de context van de lineaire modellen zijn er hoofdstukken over het schatten van de varianties, de constructie van betrouwbaarheidsintervallen over adaptieve bandbreedte selectie en over uitbreiding van de modellen, bijvoorbeeld om betere robuustheid te bereiken. Binnen de context van de dichtheidsschatters is er een hoofdstuk met specifieke aandacht voor automatische bandbreedte selectie (klassiek vs. 'plug-in').

Een tweede kern van het boek wordt gevormd door overwegingen betreffende implementatie van de methoden. Dit betreft zowel de praktische omgang met het programma `LOCFIT` als de achtergrond van de implementatie van de belangrijkste algorithmen. Tenslotte zijn er twee hoofdstukken over 'speciale onderwerpen': een historische overzicht van lokale regressiemethoden met nadruk op ontwikkelingen in de actuariële wetenschap in het begin van de twintigste eeuw en een theoretisch hoodstuk over convergentie snelheden.

'Local Regression and Likelihood' is een boek met wat schoonheidsfoutjes: type fouten, de paginering in de inhoudsopgave klopt niet geheel, en zie ook de boven geciteerde web pagina. Ik ben echter vol bewondering. Het boek introduceert een reeks nuttige, niet-parametrische, modellen vanuit een consistente optiek, levert een krachtig hulpmiddel in de vorm van `LOCFIT` voor het stoeien met dergelijke modellen, en koppelt dit aan een intuitieve blik op de achterliggende theorie. Wat kan een toepasser zich nog meer wensen!

drs D. Denteneer
Philips Research Laboratories Eindhoven

Singpurwalla, N., Wilson, S.P.
Statistical methods in software engineering. Reliability and risk
1999, Springer series in statistics,
Springer-Verlag, Berlin,
310 blz, DM 139.00, ISBN 0-387-98823-8.

Sinds het begin van de zeventiger jaren hebben een behoorlijk aantal statistici onderzoek verricht naar aspecten die met betrouwbaarheid van software te maken hebben,

of wellicht te maken zouden kunnen hebben. Eén van de actieve onderzoekers op dit gebied is Nozer Singpurwalla, en in 1994 publiceerde hij samen met Simon Wilson een overzichts artikel over modelleren voor *software reliability*. Dit heeft nu vervolg gekregen in de vorm van een boek. Alhoewel het boek enige aardige aspecten heeft, en er nog maar weinig boeken verschenen zijn op dit specifieke gebied, heb ik moeite met het inzien van het nut van dit werk, los van het feit dat ik vraagtekens zet bij elk werk waarin het woord *'synergism'* gebruikt wordt in het voorwoord.

De kern van dit werk ligt in het gebruiken van veelal redelijk eenvoudige statistische modellen, vaak *point* of *counting processes* met soms wat Bayesiaanse elementen toegevoegd, voor het beschrijven van software betrouwbaarheid. Dit kan door bijvoorbeeld aan te nemen dat software een bepaald (mogelijk random) aantal bugs heeft, en dan te kijken hoe vaak deze tot fouten zullen leiden, met enige aannamen voor het verwijderen van bugs. Allerlei combinaties van aannamen hebben geleid tot veel artikelen, soms in statistische tijdschriften, soms in computer of software tijdschriften. Echter, het is typisch voor dit soort werk dat statistici hebben geprobeerd hun bestaande methoden toe te passen op een nieuw gebied, en daarbij meer aandacht besteed lijken te hebben aan wetenschappelijke output dan het daadwerkelijk oplossen van problemen waarmee de software wereld dagelijks geconfronteerd wordt. Het grote practische probleem is hoe software getest moet worden, en hoe test resultaten en overige kennis over het software pakket tezamen vertaald kunnen worden naar 'betrouwbaarheid'. Ondanks een hoofdstuk over *'optimal testing'*, waarbij een aantal eenvoudige vragen tot interessante wiskunde leiden, zal men in dit boek, en in dit soort methodes in het algemeen, zelden duidelijke antwoorden aantreffen, behalve dan voor uitermate eenvoudige software systemen. Lastige practische problemen voor software testen hebben bijvoorbeeld te maken met erg diverse inputs, of zelfs inputs buiten het gedefinieerde input domein, gecompliceerde verbanden tussen software acties, en practische beperkingen zoals deadlines voor release en beperkte test budgetten. Ook is het tellen van bugs erg discutabel, omdat bugs meestal niet duidelijk gedefinieerd zijn (zie Frankl, *et al.* (1998) voor een goede discussie). Dit soort zaken komen niet aan de orde in dit boek. Natuurlijk, in dit soort werk is steeds een claim van practische relevantie door wat voorbeelden te presenteren die gebruik maken van 'echte data'; het feit dat het meestal gaat om één twintig jaar oude data set die in de literatuur steeds weer aangehaald wordt, zonder verdere aandacht voor de context, kan zogewenst gebruikt worden ter ondersteuning of ter ontkrachting van deze claim.

Wat doet het boek dan wel? Het slaagt er in bijdragen gepubliceerd in uiteenlopende tijdschriften bij elkaar te scharen - alhoewel het zeker geen volledig overzicht is, de belangrijkste aspecten duidelijk te introduceren en toe te lichten, en aan te tonen wat met redelijk toegankelijke statistische methoden mogelijk is; het is aan de lezer om in te zien wat niet mogelijk is. In het voorwoord worden de doelen en doelgroepen van het boek genoemd, en het is duidelijk uit de vele doelen en de grote variëteit van lezers voor wie dit boek geschikt geacht wordt, dat het risico bestond een boek te creëren waarop eigenlijk niemand zat te wachten. Het boek is niet echt een boeiende wetenschappelijke bijdrage voor de statistiek literatuur, en levert geen practisch

nuttige methoden voor *software engineers*. Het geeft wel wat enigszins interessante voorbeelden van sommige modellen voor processen, en gerelateerde statistische methoden voor schatten en het nemen van beslissingen. Ook kan het gebruikt worden als inleiding in een vakgebied waarbij goede publicatie mogelijkheden bestaan, hetgeen academici wellicht aan zal spreken. Jammer dat het boek geringe practische waarde heeft voor *software engineering*. Redmill (1999) heeft een interessant artikel geschreven waarin relevante vraagstukken uit de *software engineering* besproken worden, het boek van Singpurwalla en Wilson biedt weinig tot geen practisch bruikbare oplossingen.

Verwijzingen:

[1] P.G. Frankl, R.G. Hamlet, B. Littlewood, L. Strigini (1998). Evaluating testing methods by delivered reliability. *IEEE Transactions on Software Engineering*, 24, 586-601.

[2] F. Redmill (1999). Why systems go up in smoke. *The Computer Bulletin* (September), 26-28.

dr F.P.A Coolen
Department of Mathematical Sciences
University of Durham

Grätzer, G.
First steps in LaTeX
1999, Birkhäuser, Basel,
136 blz, DM 44.00, ISBN 3-7643-4132-7.

De grootschalige beschikbaarheid van pc's heeft het werk van statistici sterk beïnvloed: niet alleen kan men tegenwoordig berekeningen uitvoeren die tien jaar geleden ondenkbaar waren, ook worden de gevonden resultaten veelal zelf gedocumenteerd. Wetenschappers typen tegenwoordig zelf hun papers en boeken en de tijd dat proefschriften door een secretaresse werden getypt ligt ook ver achter ons. Een van de betere pakketten om teksten te zetten die formule's bevatten is LaTeX. Eigenlijk is LaTeX een mark-up taal die beschrijft wat welke tekst voorstelt, en er zijn verschillende programma's die een LaTeX-tekst transformeren in iets dat kan worden bekeken op het scherm of kan worden afgedrukt (een bespreking van een aantal verschillende LaTeX-programma's kan worden gevonden in Koning (2000)). Het boekje van Grätzer geeft een korte inleiding in LaTeX, zonder op verschillende programma's in te gaan.

In het dunne boekje komen de volgende onderwerpen aan de orde.

1. Typing text
2. Typing math
3. Formula's and user-defined commands
4. The anatomy of an article

5. An AMS article
6. Working with LaTeX

A. Math symbol tables

B. Text symbol tables

C. LaTeX and the Internet

Het boekje is bedoeld voor '... the mathematician, physicist, engineer, scientist, or technical typist who needs to quickly learn how to write and typeset articles containing mathematical formulas.' Het voldoet inderdaad voor deze doelgroepen, waarbij de auteur er wel van uit gaat dat men een werkend LaTeX-programma op de computer heeft staan. Aangezien met name het zetten van formule's veel aandacht krijgt in het boekje, bespreekt de auteur ook uitbreidingen op LaTeX die door de American Mathematical Society zijn geïntroduceerd. In de marge van de tekst geeft hij aan wanneer hij deze uitbreidingen bespreekt. Uitbreidingen in de vorm van zogenaamde 'style'-files die behulpzaam zijn bij het veranderen van de layout van de tekst, komen nauwelijks aan bod: de auteur beperkt zich hoofdzakelijk tot dingen die LaTeX van zichzelf al kan.

Uiteraard is het van belang om het boek te vergelijken met de standaardwerken op LaTeX-gebied: Lamport (1994) en Goossens, Mittelbach en Samarin (1994). Duidelijk is dat het boek van Grätzer veel minder onderwerpen behandelt, en dat dat de onderwerpen die worden behandeld in het algemeen minder diepgaand worden behandeld. Het beperkt zich tot de LaTeX-syntax, terwijl allerlei hulpprogramma's (BibTeX), en achtergrondinformatie (fontselectie, grafisch materiaal, PostScript) niet of nauwelijks wordt besproken.

Een gemis in het boekje is de stiefmoederlijke behandeling van grafieken, plaatjes en tabellen in LaTeX. Het is niet altijd even eenvoudig om deze op te nemen in een LaTeX document, en vaak wel noodzakelijk.

Het boekje van Grätzer is een aardige inleiding in LaTeX, maar bepaalde zaken missen. Eenieder die serieus LaTeX wilt gaan gebruiken om een proefschrift of boek te zetten, zal meer baat hebben bij de boeken van Lamport en Goossens et al. De vele voorbeelden alsmede de compactheid maakt het boekje toch de moeite van het lezen waard.

Verwijzingen:

[1] Lamport, L. (1994) *LaTeX A Document Preparation System User's Guide and Reference Manual* Reading: Addison-Wesley.

[2] Goossens, M., F. Mittelbach en A. Samarin (1994) *The LaTeX Companion* Reading: Addison-Wesley.

[3] Koning, R.H. (2000) 'A Comparison of Different LaTeX-programs', *Journal of Applied Econometrics*, te verschijnen.

dr R.H. Koning
Vakgroep Econometrie
Rijksuniversiteit Groningen

Headayat, A.S., Sloane, N.J.A., Stufken, J.
Orthogonal arrays. Theory and applications
1999, Springer series in statistics,
Springer-Verlag, Berlin,
440 blz, DM 126.00, ISBN 0-387-98766-5.

Via de VVS
25% korting!

An orthogonal array of strength t is loosely speaking a rectangular table of symbols such that any t-tuple appears equally often. The relation with e.g. classical factorial designs or Taguchi designs in statistics is immediate by interpreting the rows of the table as coding for the runs of a design. It interesting to note that orthogonal arrays are also directly connected to coding theory. In this case we interpret the rows of the table as code words. We thus have an object that links discrete mathematics, coding theory and statistics. As a result, useful results on orthogonal arrays are scattered over the literature (and hence, we see names like Fisher, Bose and Rao in the discrete mathematics literature). The present book combines the results of the different fields in a superb way. The first 10 chapters of this book give an encyclopaedic, yet very readable overview of the mathematical theory of orthogonal arrays. Several constructions for orthogonal arrays are treated in detail. In some chapters (e.g., chapter 7 on Hadamard matrices) the link with statistical design is already made. Statistical aspects of orthogonal arrays can be found in chapter 11. This chapter assumes a working knowledge of the theory of statistical designs. Finally, chapter 12 contains an extensive list of tables on orthogonal arrays. Each chapter has many exercises that help understanding the mathematical details.

I heartily recommend this book to statisticians working on or using experimental designs. They will find the state of the art on orthogonal arrays (especially if we include the web sites of the authors) which can help to build designs for non-standard situations or determine that some designs simply do not exist.

Dr A. Di Bucchianico
Capaciteitsgroep Product- en Proceskwaliteit
Technische Universiteit Eindhoven

Cowell, R.G., Dawid, A.P., Lauritzen, S.L., Spiegelhalter, D.J.
Probabilistic networks and expert systems
1999, Statistics for engineering and information science,
Springer-Verlag, Berlin,
340 blz, DM 139.00, ISBN 0-387-98767-3.

Via de VVS
25% korting!

Grafisch modelleren via Bayesiaanse netwerken heeft een enorme ontwikkeling ondergaan sinds het begin van de jaren negentig, en heeft bijgedragen tot grootschalige toepassingen van statistiek. De auteurs van dit boek hebben in belangrijke mate bijgedragen aan deze ontwikkeling, en geven in dit boek een goed samenhangende presentatie van Bayesiaanse netwerken in het algemeen, en hun bijdragen aan dit gebied in het bijzonder. Lezers die al wat werk van de auteurs kennen, zullen bekend zijn met

veel van de theorie, en nagenoeg alle voorbeelden. Toch is het prettig dit werk mooi gestructureerd bij elkaar te zien. Het boek is ook geschikt als een eerste kennismaking met dit onderwerp, en de auteurs hebben duidelijk aangegeven welke hoofdstukken wellicht in eerste instantie overgeslagen kunnen worden.

De eerste twee hoofdstukken zijn inleidingen (over het boek en veelal bekende aspecten van kanstheorie). Het derde hoofdstuk behandelt het opbouwen en het gebruik van grafische modellen. Ik vind het lastig dat soms termen en concepten gebruikt worden die pas in latere hoofdstukken formeel geintroduceerd worden, maar het is een verdedigbare keuze van de auteurs die voorkomt dat het boek begint met zuiver theoretische hoofdstukken. In hoofdstukken 4 en 5 wordt onderliggende theorie besproken (grafen theorie en relevante Markov eigenschappen), en het is prettig dit in hetzelfde boek te zien, alhoewel deze hoofdstukken niet iedereen vrolijk zullen maken. Het zesde hoofdstuk bespreekt discrete netwerken, het zevende hoofdstuk gaat over 'normale verdeling'-netwerken en mengsels van discrete en 'normale-verdeling'-netwerken. Deze twee hoofdstukken vormen het belangrijkste gedeelte van het boek, en zijn prettig leesbaar. Met name discrete netwerken zijn goed begrijpbaar voor nieuwelingen op dit gebied, en zelfs lezers met wat minder goede statistische achtergrond zullen de basis concepten begrijpen. Veel toepassingen maken ook gebruik van dit soort netwerken.

In hoofdstuk 8 wordt getoond hoe beslissingen impliciet in de modellen opgenomen kunnen worden ('decision networks' of 'influence diagrams'). Tot zoverre gaat men ervan uit dat kansen bekend zijn, en men geinteresseerd is in conditionele of marginale kansen. In hoofdstuk 9 wordt de uitbreiding besproken die nodig is om te kunnen leren over de kansen, in het geval nieuwe data beschikbaar komen. De modellen worden dan een stuk complexer, voor mij had dit wat meer aandacht mogen hebben. In de laatste twee hoofdstukken worden een aantal mogelijkheden voor model diagnose besproken, en tevens 'structural learning', het leren van conditionele onafhankelijkheids structuren in data-sets. Deze hoofdstukken zijn korte, maar goede, introducties in deze belangrijke aspecten.

Er zijn drie nuttige appendices, met name het laatste dat een overzicht geeft van informatie en software beschikbaar via het internet (zowel gratis als commercieel). Natuurlijk zal dit soort informatie snel kunnen verouderen, maar het is prettig de software te kunnen vinden om deze methodes inderdaad ook te kunnen gebruiken.

De auteurs geven duidelijk aan dat ze niet streven naar een compleet overzicht van de ontwikkelingen op dit gebied, maar ze geven zeer nuttige overzichten van ontwikkelingen, met referenties, aan het eind van ieder hoofdstuk. Vergeleken met twee andere standaard werken op dit gebied (Jensen (1996) en Lauritzen (1996)) vind ik dit boek een vooruitgang, het leest wat prettiger en neemt de sterke punten van die boeken over. Het boek kan uitstekend dienen als basis voor een cursus voor ouderejaars studenten statistiek of wiskunde, maar de docent zal zelf moeten zorgen voor meer voorbeelden en opgaven. In dit verband mogen we opmerken dat HUGIN (commercieel software pakket - genoemd in de appendix - voor deze methoden dat ik prefereer boven de genoemde alternatieven, maar wel vrij prijzig) sinds kort een (verbeterde) gratis ver-

sie voor onderwijs beschikbaar stelt via hun web-site (`http://www.hugin.dk/`), met goede functionaliteit en de mogelijkheid netwerken tot 200 'states' (bijvoorbeeld honderd binaire variabelen) in het geheugen op te slaan. Deze 'HUGIN-Lite' versie (windows 95/98 en NT) is voor iedereen nuttig, zowel voor eerste studie van dit soort methodes als voor menige toepassing, maar heeft helaas (nog) geen goede gebruikers handleiding.

Samenvattend, een goed boek, dat zeker in de statistiek bibliotheek moet staan, en waar menigeen veel plezier aan kan beleven. Op dit moment ken ik geen beter boek op dit gebied. Gezien de snelle ontwikkelingen verwacht ik alternatieven binnen afzienbare tijd, waarbij ik met name een tekstboek op prijs zou stellen, bijvoorkeur gekoppeld aan (voor studenten) gratis beschikbare software. Ik vind ook dat Springer de prijzen van dit soort boeken wel wat aantrekkelijker zou kunnen maken, of wellicht goedkope paperback versies uit zou kunnen brengen (het is zinloos dit over enige jaren te doen, zoals soms het geval is, vanwege de snelle ontwikkelingen).

Verwijzingen:

[1] Jensen, F.V. (1996). *An Introduction to Bayesian Networks.* University College London Press, London.[1]

[2] Lauritzen, S.L. (1996). *Graphical Models.* Clarendon Press, Oxford.

dr F.P.A Coolen
Department of Mathematical Sciences
University of Durham

Does, R.J.M.M., Roes, K.C.B., Trip, A.
Statistical process control in industry
1999, Mathematical modelling: theory and practice,
Kluwer, Dordrecht,
248 blz, fl 195.00, ISBN 0-7923-5570-9.

Readership: quality managers, engineers, students in operational research, students in engineering, students in quantitative economics.

During the past decades the interest in quality management has increased very fast. Certification on the basis of ISO norms has become a must in many production facilities and industries. Without an ISO certificate many suppliers are excluded from taking part in business.

The present book is devoted to Statistical Process Control (SPC), which forms an important element in Total Quality Management. There are many books and papers about SPC and the technical and statistical backgrounds for using SPC. There are few books that give a detailed outline about how to introduce and implement SPC step-by-step on the shop floor. The authors present an excellent book of the second type.

The main part of the book is devoted to the philosophy behind SPC. In Chapter 2 the authors answer basic questions such as: Why using SPC? and, What is the place of

[1] Finn Jensen werkt momenteel aan een tweede editie van dit boek.

SPC in a Quality Policy? In Chapter 3 they discuss the 4 phases in the implementation of SPC (Awareness, Pilot Projects, Integral Implementation, Total Quality by means of SPC). They also stress the importance of SPC-steering committees and the composition of SPC-teams. The core Chapter 5 gives a detailed overview of a plan of action for introducing SPC. The authors present a 10-step programme that ultimately yields a controlled process. Each of the steps is treated in a uniform way and provides information about Goal?, Result?, Method?, Guidelines? of the step. Each section concludes with an illustration. Chapter 10 discusses progress control of SPC teams and briefly goes into assuring SPC points.

In the book, the statistical and technical treatment (chapters 6-9) is rather limited and the derivation of formulaes is largely left aside. The authors merely stress interpretation and motivation of the results. Also, the list of control charts is limited to a number of basic, but very important charts. In my opinion it is a pity that the words "confidence interval" or "confidence level" do not appear in the text.

In chapter 11 the authors briefly discuss their ideas about software for SPC. The final chapter 12 is devoted to some comments about SPC competition and self evaluation.

As a conclusion, the book deserves its place in the office of anyone who is interested in TQM in general and espcially in SPC . The presentation (lay-out, presentation of graphs and sharts, ...) in the book is very good (only page 162 contains a typographical error). The ideas offered in the book can be of use elsewhere and the treatment of SPC can be attractive to a large audience. Readers interested only in the technical treatment of SPC and in an extensive list of all possible control charts will be disappointed. The chapter devoted to software (Chapter 11) could be omitted, or, could be expanded by a discussion about existing software packages. The list of references is acceptable and the subject-index is useful. The value of the book could increase by ending each chapter with some bibliographical comments and suggestions for "further reading".

prof. dr E. Omey
EHSAL

Morrison Paul, C.J.
Cost structure and the measurement of economic performance. Productivity, utilization, cost economics, and related performance indicators.
1999, Kluwer, Dordrecht,
384 blz, fl 270.00, ISBN 0-7923-8403-2.

This book is about the measurement and understanding of productivity change. Empirically, the productivity change of an economic entity such as a firm is measured by the growth rate of its output minus the growth rate of its input, or, equivalently, by an output quantity index divided by an input quantity index. If output is related to all inputs, we speak of total factor productivity (TFP) change. If output is related

to a single input factor, we speak of single factor productivity change. A wellknown example of the latter type is labor productivity change. Given appropriate data, measures of productivity change can be calculated not only for single firms, but also for aggregates like industries, regions, and countries. They are among the most important measures of economic performance. Their history dates back to Solow (1957) – hence the frequently occurring name 'Solow residual' – or even Tinbergen (1942).

In order to appreciate the contribution and importance of the book under review, I will give a brief, hopefully not too technical, sketch of the classical position. Against that backdrop the contents of the book will be reviewed. For ease of presentation the discussion will be cast in terms of a single firm.

Let the technology that governs the firm's production process at time period t be characterized by the cost function $C(w, y, t)$. This function provides the minimum cost for producing the output quantities y (an M-dimensional vector) when the input prices are given by w (an N-dimensional vector). Let the actual cost incurred by the firm producing $y(t)$ be $w(t) \cdot x(t)$, where $x(t)$ denotes the vector of input quantities at period t and $\cdot$ denotes the inner product.

One of the classical key assumptions is that the firm is a cost minimizer, that is, actual cost is assumed to be equal to minimum cost, or, in terms of the formulas just introduced:

$$w(t) \cdot x(t) = C(w(t), y(t), t). \tag{1}$$

Assuming that all functions are continuous, we can differentiate with respect to time t to obtain growth rates. The lefthand side of (1) yields

$$\frac{d \ln w(t) \cdot x(t)}{dt} = \sum_n s_n(t) \frac{d \ln w_n(t)}{dt} + \sum_n s_n(t) \frac{d \ln x_n(t)}{dt}, \tag{2}$$

where the $s_n(t)$'s denote the actual cost shares of the inputs. The righthand side of (1) yields

$$\frac{d \ln C(w(t), y(t), t)}{dt} = \sum_n \frac{\partial \ln C(.)}{\partial \ln w_n} \frac{d \ln w_n(t)}{dt} + \sum_m \frac{\partial \ln C(.)}{\partial \ln y_m} \frac{d \ln y_m(t)}{dt} + \frac{\partial \ln C(.)}{\partial t}. \tag{3}$$

The assumption of cost minimization combined with Shephard's Lemma leads us to conclude that

$$\frac{\partial \ln C(.)}{\partial \ln w_n} = s_n(t), \tag{4}$$

that is, the partial derivatives occurring in the first term at the righthand side of expression (3) are equal to the actual cost shares of the inputs.

The second assumption is that the output prices are proportional to marginal cost, that is

$$p_m(t) \propto \frac{\partial C(.)}{\partial y_m}. \tag{5}$$

Simple manipulations with this relation lead to the following expression:

$$\frac{\partial \ln C(.)}{\partial \ln y_m} = u_m(t) \sum_m \frac{\partial \ln C(.)}{\partial \ln y_m}, \tag{6}$$

where the $u_m(t)$'s are the actual revenue shares of the outputs.

The third and final assumption is that the technology exhibits constant returns to scale. This implies that the second factor at the righthand side of expression (6) is equal to 1. Thus, the second and third assumption together imply that the partial derivatives occurring in the second term at the righthand side of expression (3) are equal to the actual revenue shares of the outputs.

We are now in the position to put the various pieces together. By construction the righthand sides of expressions (2) and (3) are identically equal. Thus, substituting the two results just obtained, we arrive at the following expression:

$$\sum_m u_m(t) \frac{d \ln y_m(t)}{dt} - \sum_m s_n(t) \frac{d \ln x_m(t)}{dt} = -\frac{\partial \ln C(.)}{\partial t}. \tag{7}$$

The lefthand side of this expression measures TFP change (in growth rate form). The righthand side is the minimum cost decrease associated with the mere passage of time and is conventionally interpreted as a measure of technological change. Thus, under the stated assumptions, TFP change is equal to technological change.

The assumptions amount to the following. First, the firm is seen as cost efficient, that is, technically efficient – the firm acts at the frontier of the current technology – as well as allocatively efficient –the input quantities have the optimal mix. Second, the firm is assumed to act in a competitive environment. Third, the technology is assumed to exhibit constant returns to scale.

In reality, these classical assumptions are almost nowhere satisfied. Due to various rigidities (e.g. fixity of capital in the short run) and/or regulatory constraints (e.g. with respect to environmental effects of certain inputs or outputs) firms almost never reach the high goal of full cost efficiency. Furthermore, markets are seldom perfectly competitive. And, finally, technologies do have regions with differing scale properties.

Now the relevance of Morrison's book can be clearly seen. It presents us with an extensive discussion of the complexities that arise when the classical assumptions are dispensed with. How are we to interpret the empirical measure of TFP in more complex situations? Must in those situations the empirical measure of TFP be replaced by more adequate performance indicators? Those are the key questions discussed in this book.

Of course, more complex situations ask for more complex modelling, and here necessarily some econometrics comes into the picture. However, Morrison's overall

approach is to discuss concepts and ways of modelling rather than detailed econometric technicalities. The level of technicality of the book almost nowhere transcends that of this review. Morrison cares about the economics of performance measures rather than their econometrics.

The structure of the book is clear. After a chapter on the traditional model there follow chapters on capital utilization, scale economies, external effects other than technological change, market structure, regulation, and (in-)efficiency. The three final chapters are of more technical nature, respectively discussing the main elements of production theory, the problems of data construction, and some basic econometric issues.

Concluding, we have here a well-written book by an author who knows the subject from within. The book guides into a rich literature and opens up many avenues for further research.

dr B.M. Balk
Department of Statistical Methods
Statistics Netherlands

Lange, K.
Numerical analysis for statisticians
1999, Statistics and computing,
Springer-Verlag, Berlin,
360 blz, DM 144.00, ISBN 0-387-94979-8.

"This book [...] was written out of frustration" is the very first sentence of the Preface. The frustration originates from the fact that no textbook was the perfect match for the course this book was originally designed for. Admittedly, I do not know of any other textbook that discusses numerical analysis focused specifically at statistics. This book elaborates on the mathematics behind random number generators and all kinds of statistical methods. If you are looking for prefab algorithms to be put into your computer, you are looking in the wrong spot.

The book covers 24 chapters divided over 356 pages: Recurrence relations, Power series expansions, Continued fraction integration, Asymptotic expansions, Solution of nonlinear equations, Vector and matrix norms, Linear regression and matrix inversion, Eigenvalues and eigenvectors, Splines, The EM algorithm, Newton's method and scoring, Variations on the EM theme, Convergence of opimization algorithms, Constrained algorithms, Concrete Hilbert spaces, Quadrature methods, The Fourier transform, The Finite Fourier transform, Wavelets, Generating random deviates, Independent Monte Carlo, Bootstrap calculations, Finite-State Markov chains, Markov chain Monte Carlo. Although some of these chapters may have a mathematical flavour to its title, the author succeeds very well in finding statistical examples of all subjects he covers. Mathematical results and derivations are placed in a statistical context.

A disadvantage of this book is that it lacks a proper introduction. Not only does it not contain an introductory chapter, also the chapters itself are mostly not properly introduced. For example, Chapter 1 on recurrence relations, only contains nine examples

of such relations under the denominator "commonly employed techniques". Why precisely these examples are chosen remains unclear. After reading other chapters it turns out that some of the examples from Chapter 1 are used elsewhere in the text.

The power of the book is undoubtedly that the author is very succesfull in explaining statistical mathematics. All main steps are included in derivations of results. The book is well organized and the text is explained in a clear way. At some points Lange is really driving; for example chapters 2 and 20 discuss the numerical calculation of cumulative probabilities and the generation of random numbers respectively from all common distributions. They are an absolute delight to read.

The book aims at advanced Master's students and professional statisticians. Presumably most statisticians are willing to step over the short introductions. They will find this book an asset to their shelves because its caleidoscope of treated subjects ensures that there is something for everyone. Notably, each chapter contains a number of exercises which look challenging and solvable –some of them are mathematical, others invite the reader to write explicit algorithms. But almost all have a statistical interpretation. Teachers will find it a valuable source for finding problems (solutions are not included). Students will find this book readable although I can imagine that only a small portion is discussed in class since there is so much to choose from.

Summarizing, this book is one of a kind. It provides the reader with tools for solving numerical problems. An accesible text, useful for anyone involved in computational statistics.

C.A. Bernaards
Faculteit Sociale Wetenschappen
Rijksuniversiteit Utrecht

Zhang, H., Singer, B.
Recursive partitioning in the health sciences
1999, Statistics for biology and health,
Springer-Verlag, Berlin,
230 blz, DM 119.00, ISBN 0-387-98671-5.

Dit boek gaat over een veelvoud aan beschikbare recursive partitioning technieken. Deze technieken worden, in het boek, toegepast op real-life onderzoeksproblemen uit de gezondheidszorg. Recursive partitioning vormt de basis voor twee bekende klasse van niet parametrische methoden: CART (classification and regression trees) en MARS (multivariate adaptive regression splines). Enkele onderwerpen die de revue passeren zijn: classificatie bomen voor binaire data, risicofactor analyse m.b.v. op bomen gebaseerde stratificatie, survival trees, regression trees, analyse van longitudionale data m.b.v. regression trees etc.

Het boek is volgens de schrijvers geschreven voor drie verschillende groepen van lezers: 1) Biomedische wetenschappers die op het gebied van de gezondheidzorg werken, 2) Statistici die recursive partitioning (willen) gebruiken om op een effectieve en inzichtelijke manier problemen van hun klanten op te lossen 3) meer methodologisch

en theoretisch georiënteerde statistici. Gezien het stevige theoretische karakter van het boek zullen delen van het boek voor de eerste groep niet zo geschikt zijn. Door het ruime aanbod van relevante voorbeelden kan het boek, voor de eerste groep lezers, dienen om een inzicht te krijgen wat er mogelijk is m.b.v. recursive partitioning. Afhankelijk van de software zal men dan samen met een statisticus deze technieken kunnen toepassen.

De voorbeelden zijn goed uitgewerkt en leesbaar. De theoretische stukken in het boek wisselen nogal in leesbaarheid, zeker voor de minder statistisch of wiskundig geschoolde lezer. De korte beschrijvingen van standaard analysetechnieken uit de statistiek als logistische regressie, survival analyse en longitudinale analysetechnieken zijn duidelijk.

Sterk aan het boek is de enorme voorraad aan uitgewerkte voorbeelden.

Wat gemist wordt in het boek is een relatie tussen de "standaardtechnieken" en modellen en de bijbehorende recursive partitioning technieken. Zo wordt een heel hoofdstuk gewijd aan logistische regressie. Het hoofdstuk daarop gaat over boomconstructie van binaire responses. Interessant is de vraag wat de relatie tussen logistische regressie is en de technieken die gebaseerd zijn op boomconstructie. Wat zijn de voordelen van de recursive partitioning technieken boven de klassieke logistische regressie? Hierop wordt niet ingegaan en dat is, m.i. een gemiste kans.

Daarnaast is het jammer dat er niet uitgebreider wordt ingegaan op recursive partitioning technieken voor polychotome data. Dit type data is de laatste jaren enorm in de belangstelling komen te staan, zeker in het gezondheidsonderzoek, en er is duidelijk een behoefte aan literatuur en software om problemen met polychotome responses op te lossen.

Kortom, voor onderzoekers, en dan met name voor statistici, op het gebied van de gezondheidszorg en de biostatistiek is dit boek een aanrader. Het bevat veel nieuwe ideeën om complexe data aan te kunnen pakken. Voor de wat minder theoretisch statistisch geschoolde onderzoek is dit boek harde kost.

dr C. Oudshoorn
Afdeling Statistiek
TNO - Preventie en Gezondheid

Hutcheson, G.D., Sofroniou, N.
The multivariate social scientist. Introductory statistics using generalized linear models
1999, Sage, Thousand Oaks, California,
288 blz, £ 17.99, ISBN 0-7619-5201-2.

De commissie onderwijs vraagt om een cursus multivariate methoden voor doctoraal studenten sociale wetenschappen, die het propedeusevak 'inleiding in de statistiek' met goed gevolg hebben afgerond. Na afloop van de te houden vervolgcursus dient de

student zelfstandig zowel continue als categorale data multivariaat te kunnen analyseren. Aan de cursus zijn zeven colleges en een totale belasting van zestig studie-uren toebedeeld.

Het beste antwoord op een dergelijke vraag is een verzoek om meer studiebelastingsuren. Zestig uur is net voldoende om regressie-analyse naar behoren te behandelen, terwijl er in een dergelijke cursus ook plaats zou moeten zijn voor logistische regressie, loglineaire analyse en afhankelijk van de specifieke groep variantie- en factoranalyse. Waarschijnlijk kiest de onderwijscommissie in de daaruit voortvloeiende discussie voor het standpunt dat in de cursus het toepassen centraal moet staan en cursisten niet de details van de technieken hoeven te beheersen. Voor het juist toepassen van iedere statistische techniek is echter ook enig rekenkundig en meetkundig begrip noodzakelijk.

Het één na beste antwoord voor de geschetste vraag is het boek 'The Multivariate Social Scientist' van Hutcheson en Sofronou. Volgens het voorwoord is dit boek bedoeld voor een cursus aan postgraduate students en onderzoekers, maar de ondertitel 'Introductory Statistics Using Generalized Lineair Models' duidt op een publiek met slechts elementaire kennis van statistiek. Dit sluit ook beter aan op de inhoud van het boek. In afzonderlijke hoofdstukken behandelen de schrijvers OLS regressie, logistische regressie, loglineaire analyse en principale componenten analyse (PCA). Samen met een inleidend, een afsluitend hoofdstuk en een hoofdstuk over het opschonen van gegevens geeft dit een cursus van zeven bijeenkomsten van twee uur elk.

OLS regressie, logistische regressie en loglineaire analyse vormen het hart van het boek. Het eerste hoofdstuk legt een verbinding tussen de technieken en vat ze op als drie variaties van een algemeen lineair model, zonder dat de cursist op dat moment de mathematische modellen van de drie technieken kent. Dit is een lastige opgave voor zowel de auteurs als de cursisten. De aansluitende uitleg over OLS regressie is kort maar compleet. Het schatten van parameters en hun betrouwbaarheid, goodness of fit, multicollineairiteit, dummycodering en modelselectie komen allemaal in nog geen zestig pagina's helder aan de orde. Van de verschillende statistieken wordt duidelijk de betekenis aangegeven, maar door de beperkte ruimte zijn berekeningen en afleidingen achterwege gelaten. Slechts zelden komen de beperkingen van een maat aan bod, een kritische beschouwing van een gekozen model met een controle op schendingen van assumpties, wordt nergens voorgesteld. Het hoofdstuk over logistische regressie behandelt de meeste aan de techniek verbonden begrippen: odds, de logit-functie, de log-likelihood, de Wald statistic. Toch blijft het niveau gelijk aan de handleiding van menig statistisch pakket. Wat de cursist in zo'n handleiding niet zal vinden is de heldere inleiding in het nut en gebruik van verschillende dummycoderingen. Dit is de aardigste paragraaf van dit hoofdstuk.

Het hoofdstuk over loglineaire modellen leidt de techniek goed in. Het nesten van modellen en de opname van lineaire factoren komen duidelijk aan bod. Voor sommige cursisten is het misschien wel teleurstellend als blijkt dat logische regressieproblemen ook met loglineaire modellen kunnen worden opgelost. De uitleg is van deze analogie echter niet altijd even helder.

PCA is een toegift, en wordt gepresenteerd als een methode om problemen met multicollineaire data op te lossen. Dit is een welkome afwisseling op het veel gehoorde maar minder praktische 'collect more data' verhaal. Het is echter storend dat de auteurs over factor analyse spreken, waar ze PCA bedoelen en voorbij gaan aan de fundamentele verschillen in de modellen, die aan de twee technieken ten grondslag liggen. Wordt PCA zo vaak gezien als een vorm van factoranalyse, omdat het binnen SPSS met het commando FACTOR wordt uitgevoerd? Ook de uitleg is nodeloos ingewikkeld. PCA opgevat als regressie-analyse op een onbekende te voorspellen variabele, de latente trek, zou daarentegen een uitleg opleveren die naadloos op het hoofdstuk over OLS regressie aansluit.

Kunnen de cursisten aan het einde zelfstandig hun data multivariaat analyseren? Met behulp van dit boek weten ze de procedures binnen SPSS en GLIM te vinden en kunnen ze de resultaten interpreteren. Ze zullen niet de techniek kunnen afstemmen op hun gegevens. Daarvoor hebben ze meer begrip nodig van wat er in de procedures met de data gebeurt, en behandelt het boek te oppervlakkig de verschillende technieken.

drs G.L.M. Hilkhuysen
Instituut voor Extramuraal Geneeskundig Onderzoek
Vrije Universiteit Amsterdam

Gerbing, D.
Relevant business statistics using Excel
1999, South-Western College Publishing, Cincinnati, Ohio,
598 blz, £ 33.99, ISBN 0-324-00358-7.

De meeste bedrijfsmanagers zullen vast met enige schrik terugdenken aan het eerste handboek statistiek dat ze ooit onder ogen kregen (als ze het al ooit onder ogen kregen). Dikke pillen met moeilijk te doorgronden tekst, waarvan in eerste instantie de relevantie moeilijk is in te schatten. De meeste studenten zien statistiek als een kwaad waarvan de noodzaak niet altijd wordt ingezien. Statistiek is echter breder dan kansberekening, verdelingen en het toetsen van hypotheses; het wordt gebruikt om planningen te maken, prognoses op te stellen, op basis van statistische gegevens worden - soms vergaande - beslissingen voor de (nabije) toekomst gemaakt. Niet alleen statistici en onderzoekers maken gebruik van statistische technieken, iedereen doet het, dus zeker ook de manager. Gerbing heeft zich als doel gesteld dat duidelijk te maken en een aantal technieken toe te lichten met behulp van de gebruikersvriendelijke mogelijkheden van Excel.

Volgens Gerbing is dit boek geschreven als hulpmiddel voor de manager om effectieve beslissingen te kunnen nemen, tenslotte wordt deze de hele dag gebombardeerd met min of meer onvoorspelbaar lijkende gebeurtenissen. Het boek moet een bijdrage leveren deze gebeurtenissen meer voorspelbaar te maken en is daarom niet zozeer technisch als wel praktisch bruikbaar. Dat betekent dat de opzet van het boek beperkt is tot meer 'alledaagse' situaties waarin de manager terecht kan komen en niet dieper ingaat

op de (mathematische) materie dan strikt noodzakelijk is (deze worden echter niet uit het oog verloren). Het boek pretendeert dan ook niet een statistisch handboek te zijn; het is meer een hulpmiddel in de dagelijkse praktijk. Verder is het boek sterk gericht op concrete situaties: aan de hand van concrete situaties wordt en probleem duidelijk gemaakt, belicht en worden oplossingen aangedragen.

Als onderzoeker denk ik in eerste instantie niet aan Excel wanneer het gaat om statistische methoden; ik gebruik Excel met name als opslagmedium en om plaatjes en grafieken te maken die inzichtelijk zijn en eenvoudig kunnen worden ingelezen in andere toepassingen. Ook gebruik ik Excel regelmatig om te rekenen, maar veel verder dan dat gaat mijn gebruik niet. Met veel nieuwsgierigheid ben ik dan ook in het boek gedoken.

Het boek is helder van opzet en kent een min of meer uniforme indeling van de hoofdstukken. Allereerst wordt de relevantie van een probleem of vraag belicht onder het motto "Why do we care?". Het tweede deel gaat in op de toepasbaarheid in een zakelijke omgeving; hoe kan een bepaalde statistische methode (of methoden) worden toegepast om een probleem op te lossen. Daarna volgt een toelichting (statistical illustration), waarin een voorbeeld stap voor stap wordt aangepakt. Terwijl in de meeste boeken opgaven aan het einde van ieder hoofdstuk zijn opgenomen, treffen we nu door het hoofdstuk heen opgaven aan. Het is een kwestie van persoonlijke voorkeur, maar ik vind dat wel handig. De opgaven zijn onderverdeeld in drie categorien, de soorten opgaven zijn door middel van symbolen eenvoudig van elkaar te onderscheiden. Dat is gemakkelijk voor de lezer. De bedoeling van deze aanpak is dat de lezer uitstijgt boven passief begrip en leert de methoden te beheersen. Bij het boek wordt een diskette geleverd met bestanden die in het boek worden gebruikt, de lezer kan met behulp van deze diskette het boek volgen. In het boek wordt regelmatig verwezen naar bestanden op de diskette. Het boek is in eerste instantie geschreven als lesboek (het komt ook voort uit de lessen die Gerbing geeft), naast het boek zelf is er een instructor's manual, een test bank en een presentatie in Power point verkrijgbaar.

De hoofdstukken bouwen wel (deels) voort op elkaar, dat betekent dat indien het boek wordt gebruikt als naslagwerk, soms moet worden teruggegrepen op voorgaande hoofdstukken. De eerste drie hoofdstukken bevatten basisbeginselen die van belang zijn voor het begrip van statistiek in het algemeen, zij gaan in op soorten variabelen en mogelijke verdelingen. In het vierde hoofdstuk wordt een sprongetje gemaakt en wordt ingegaan op generaliseerbaarheid (en betrouwbaarheid) van de data. Daarna komen een aantal toetsen aan bod; allereerst de zin en onzin van toetsen, het vergelijken van gemiddelden, analyse van proporties en regressieanalyse.

Het boek is duidelijk geschreven vanuit de mogelijkheden die Excel biedt. Die zijn aanmerkelijk uitgebreider dan ik tot voor kort dacht, maar het legt uiteraard wel zijn beperkingen aan de mogelijkheden: Excel is tenslotte geen statistisch analyseprogramma. Niettemin blijkt Excel een aantal heel aardige mogelijkheden te bevatten, die zonder al te veel technische kennis te hanteren zijn. Dat maakt dit boek heel duidelijk: de voorbeelden zijn concreet, er wordt duidelijk, middels de diverse menus en pop-up schermen die Excel biedt, aangegeven wat de mogelijkheden zijn en hoe een vraagstuk

kan worden aangepakt. Verder wordt er duidelijk ingegaan op de interpretatie van de resultaten van een toets.

Een andere beperking - en daar wordt wat minder uitgebreid op ingegaan - is dat de mogelijkheden om data te bewerken in Excel beperkt zijn. Er wordt impliciet vanuit gegaan dat je goede data hebt, waarop direct kan worden geanalyseerd. Misschien is dat in bedrijfstoepassingen wel vaak zo, in de onderzoekspraktijk is dat zeker niet altijd zo. Hoe ga je om met ontbrekende gegevens, wat doe je met uitbijters etc. De mogelijkheden van Excel om dit soort problemen het hoofd te bieden zijn beperkt en het boek besteedt daar eigenlijk geen aandacht aan.

Tot slot kent Excel nog een beperking: het kan slechts databestanden van beperkte omvang aan: het aantal observaties (records) is hierin minder een beperking dan het aantal variabelen.

Krachtig punt van Excel - en daarop wordt uitgebreid ingegaan - ligt in de grafische mogelijkheden. Het is heel eenvoudig om met Excel grafieken die goed te lezen zijn te maken. Het boek geeft hiervan veel en duidelijke voorbeelden.

Samenvattend is het een helder en duidelijk boek, dat sterk gericht is op de praktijk. Het boek is geschreven als lesboek, maar kan ook daarbuiten worden gebruikt. Het boek kent een heldere opbouw en geeft duidelijke toepassingen en neemt ons stap voor stap mee in de zoektocht naar een antwoord op een vraag. Het boek is niet bedoeld voor onderzoekers, waarschijnlijk omdat ervan wordt uitgegaan dat zij krachtiger analysetools tot hun beschikking hebben. Het is duidelijk bedoeld voor meer algemene zakelijke toepassingen en pretendeert ook niet meer dan dat. Het boek is duidelijk om Excel en zijn mogelijkheden heen geschreven en dat bepaalt dan ook de beperkingen van het boek. Excel is tenslotte geen statistisch pakket, maar een spreadsheet met uitgebreide mogelijkheden. Het grote voordeel van Excel ligt meer de gebruikersvriendelijkheid en de grafische mogelijkheden die het pakket biedt, daar sluit het boek goed bij aan.

R. Doornbos
ITM Research

Eldredge, D.
An Excel companion for business statistics
1999, South-Western College Publishing, Cincinnati, Ohio,
656 blz, £ 13.99, ISBN 0-538-89088-6.

An important aspect of teaching a basic applied statistics course is the choice of software. The most straightforward choice is to use dedicated statistical software. However, in many practical situations, data is collected in the form of spreadsheets. Since spreadsheets may have options to perform statistical calculations, there is something to be said to use spreadsheet software when teaching statistics courses. The author of the book under review has been teaching business statistics courses using Excel for several years and claims that this has been successful. Unfortunately, he does not

provide details. Since the actual use of software is discussed marginally in most statistics textbooks, the author decided to write a book full of details how to use Excel for statistical analyses. Thus the present book should be used together with a statistics textbook. The results of this is a carefully written book with detailed instructions on using Excel and useful advice on obtaining good computer habits.

Let us now turn to details of the contents. The book mainly discusses the use of Excel 5 through 7, with notes on the slight deviations necessary for the use of Excel 97 (for mysterious reasons, not called Excel 8 by Microsoft). Add-ins for Excel like e.g. Xlstat or Winstat are not discussed. As said before, the book is carefully written. Nevertheless, I have several minor points of criticism, mainly on statistical methodology rather than software issues.

A general criticism is that normality is assumed everywhere, but it is not shown how to test normality with Excel.

On page 10, the useful autosequence feature should be mentioned under point 3 as a way to quickly enter sequences with patterns. The reference to Figure 13.13 on page 62 should be to Figure 3.13. In order to obtain cumulative probabilities (which Excel only offers for few of the standard distributions), the moving sum feature should be used rather than computations for each cell separately. It is regrettable that the sometimes sloppy Excel terminology for statistical objects is not corrected (e.g., EXPONDIST returns the exponential distribution). On page 82 I would speak about a discrete distribution to clearly distinguish from continuous distributions. The Central Limit Theorem is stated incorrectly on page 87. On page 86 a normal approximation is applied to a binomial distribution without checking whether a Poisson approximation is appropriate. On page 95 it is not correct to use the standard normal distribution instead of the Student T distribution. I do not understand why a two-sided test is used on the same page, since the context clearly indicates a one-sided alternative. The use of terminology like "highly significant" is doubtful (given a prescribed type I error, a result is significant or not). The calculation of a minimum sample size to obtain a prescribed accuracy on a mean form a normal distribution uses an estimated variance, and hence is not correct. Likewise, on page 133 an estimated proportion is used to calculate a minimum sample size rather than the conservative choice $1/2$. In Table 5.7 the test for the hypothesis $H_0 : \mu_1 - \mu_2 = \delta \neq 0$ is lacking, while Excel does offer this option (similarly, on page 149 only $H_0 : \sigma_1^2/\sigma_2^2 = 1$ is discussed). Amazingly, Excel has no standard procedure available for testing $H_0 : \sigma^2 = \sigma_0^2$. In chapter 8, the Excel terminology on analysis of variance is sloppy. The χ^2-tests for contingency tables are based on approximations, which require certain restrictions on the cell frequencies. These restrictions are not checked on page 169. In the chapter on regression analysis, the interpretation of high values of R^2 on page 193 is misleading. In chapter 12 control charts are discussed, CUSUM charts are lacking. However, the companion web site to the book by Hawkins and Olwell on CUSUM charts (`http://www.stat.umn.edu/users/cusum/`) contains extensive Excel sheets. The 3σ limits used for np and p-charts may be poor approximations.

I conclude that this is a useful book for anyone who wishes detailed instructions on how to use Excel for basic statistical calculations. However, for teaching purposes it has the slight drawback of containing several minor inaccuracies regarding statistical methodology.

Dr A. Di Bucchianico
Capaciteitsgroep Product- en Proceskwaliteit
Technische Universiteit Eindhoven

Albright, S., Winston, W., Zappe, C.
Data analysis and decision making with Microsoft Excel
1999, South Western, Belmont, California,
850 blz, £ 23.99, ISBN 0-534-26124-8.

Readership: students, decision makers, teachers of statistics, econometrics and operations research.

In this book the authors give an integrated approach to the transformation of data into useful and important information. They discuss various topics and diverse methods, but the underlying goal in each case is the same: to equip the reader with tools that he can be applied as a decision maker. In bookshops one can find many books about statistics, Excel, operations management, simulation, etc. Usually these are taught in separate courses. There are fewer books that integrate different methods and applications in one book. In this work Albright, Winston and Zappe start from realistic examples and problems and gradually introduce the miscellaneous (statistical or other) concepts needed to answer the questions. Since real problems can not be treated manually, the authors offer a spreadsheet-driven approach.

There are many (non-spreadsheet) statistical packages available, but the authors decided to use Microsoft Excel and to provide the necessary tools and add-ins to do so. In our opinion, this is a good and democratic approach: most students and teachers have easily access to Excel. Moreover, it is not necessary first to go through a list of manuals of statistical packages and learn about the special features of each of these non- spreadsheet-based packages. Virtually everyone has basic knowledge about the usual features of Excel. Fewer people are acquainted with the more powerful features such as the SOLVER add-in or the DATA ANALYSIS add-in. In this book the authors show in detail how these add-ins can be used.

It is well-known that the standard Microsoft Excel has some disturbing features: graphs are not nice and should be adapted, tables don't look as you want them and you have to adapt them, etc. Together with the book the authors provide a CD-ROM containing some extra useful tools. In this discussion, we will highlight three of them.

With the StatPro add-in, the reader can perform in a very elegant way most of the usual statistical tests and calculations such as: describe data by measures and by graphs; test hypothesis and construct confidence intervals; take samples; construct control charts for variables and for attributes; perform (multiple) regression; analyse

time series, etc. The provided software is user-friendly and, in an interactive way, the user has to demonstrate what he/she wants the software to do. We used StatPro to perform some regression analysis, and we are very satisfied with the performance of the package. The time series component however is rather limited.

With the BestFit module, the reader can enter data (or import data from an Excel sheet) and fit models to the data. There is a large choice of available distributions and one can choose the criterion (Chi-square, Anderson-Darling or Kolmogorov-Smirnov) to select the best fit. The book is not divided into "theory" and "exercises". From the first page, the reader should use his computer and start performing the tasks described in the book. Following the guidelines of the work, the reader develops spreadsheet skills and gets acquainted with the (statistical) terminology and with the practical use of the method that is discussed.

From the operational research point of view, the book has a good introduction to optimization theory (linear programming) and its applications, mainly in the financial and economical arena. The construction of linear models in different real companies highlights the importance of linear programming tools in economic world. The last chapter shows a wide variety of simulation models applied in managerial field. The authors use @Risk as the foremost simulation tool in the Excel spreadsheet, showing its suitability in order to teach to students in an easy way.

The CD-ROM includes some tutorials, it contains all examples of the book and teachers can use the PowerPoint presentation files for all examples of the book. Also it contains the solution files for all the problems and cases of the book.

To conclude, this book is well written, it is suitable for self-study, it contains a lot of nice teaching material, it confronts the students with real problems, and it provides the information and the software to attack the problem with analytic methods. If it is properly used, it will reverse the negative attitude about statistics and quantitative methods of many students.

prof. dr J. Faulin
Public University of Navarra
prof. dr E. Omey
EHSAL Brussel

Utts, J.M.
Seeing through statistics
1999, South Western, Belmont, California,
504 blz, £ 21.99, ISBN 0-534-35786-5.

Most introductory statistics text books that I know put a lot of emphasis on computational aspects of statistics (i.e., calculation of test statistics) and relatively little emphasis on practical issues. The present book is a refreshing exception to this rule. Professor Utts extremely inspiringly describes how to collect data and perform statististical analyses. Many examples (mainly surveys on everyday life matters) are given. I highly value the large number of examples in which errors are shown like errors in choosing

the right population to sample from. Such examples are needed to appreciate the difficulties arising when using statistics in practice (and turns statistics into a challenging subject for students). The mathematical details are kept to a minimum and put in sections called *For Those Who Like Formulas*. The statistical techniques covered in this book are mainly estimation and hypothesis testing in the context of samples from a binomial or normal distribution. In the latter case, the variance is always assumed to be known, so that the Student em t distribution is not covered. Since the examples have a bias towards sampling and the scope of the statistical techniques is rather limited, I think that this book is mainly useful for an introductory statistics course in the social sciences or for use at high schools.

Dr A. Di Bucchianico
Capaciteitsgroep Product- en Proceskwaliteit
Technische Universiteit Eindhoven

van der Kloot, W.A.
Meerdimensionale schaaltechnieken
1999, Lemma, Utrecht,
456 blz, fl 99.50, ISBN 90-5189-692-1.

'Statistiek is leuk'. Dat blijkt opnieuw bij het lezen van 'Meerdimensionale Schaaltechnieken voor gelijkenis en keuze data'. In ruim vierhonderd pagina's krijgen de lezers op een boeiende manier een goed beeld van MDS en ontvouwingstechnieken.

Na een uitleg van metrische MDS staat de auteur stil bij een aantal fundamentele statistische begrippen zoals data, afstand en transformaties; begrippen die ook voor andere statistische technieken een essentiële betekenis hebben. Voortbouwend op verworven kennis volgt een uitleg over niet-metrische MDS, de interpretatie van MDS oplossingen, MDS van asymetrische tabellen en van drieweg/tweemodale data. Dit alles gebeurt in de eerste negen hoofdstukken met een duidelijke lijn en structuur in een heldere schrijfstijl. Voor de uitleg maakt de auteur gebruik van een ruimtelijke representatie van data: er wordt naar hartelust geroteerd, getransleerd, assen worden opgerekt of ingekrompen, kortom getransformeerd. In deze hoofdstukken zit naar mijn mening de kracht van het boek.

Vanaf hoofdstuk 10 wordt het minder, omdat de lijn vervaagt. De auteur probeert de indeling van keuzedata en preferentierangordeningen als rode draad te hanteren, maar het resultaat is een wat droge opsomming van technieken. En ofschoon de creativiteit bij de toepassingen indrukwekkend is, dreigt er even een rommelig kookboek voor de analyse van genoemde soorten gegevens te ontstaan.

Een tweede kritische noot betreft de gebruikte software. De auteur gebruikt SPSS en heeft de uitvoer van dit pakket inclusief figuren zo overgenomen, vermoedelijk met het idee om de lezer vertrouwd te maken met procedures als ALSCAL, INDSCAL, HOMALS en ANACOR. Maar SPSS levert figuren van een belabberde kwaliteit. De schoonheid van MDS, die juist in de figuren zit, komt zo onvoldoende tot uitdrukking. Dat geldt zeker voor de versie die de auteur heeft gebruikt. Het resultaat is dat

verschillende figuren uit lettertekens zijn opgebouwd. En ofschoon de lezers precies kunnen zien wat ze waar kunnen vinden, maakt de uitvoer het boek beperkt houdbaar: andere software of een nieuwe versie en het ziet er allemaal anders uit. En dat terwijl de deugdelijke uitleg van de technieken de tand van de tijd zonder moeite moet kunnen doorstaan.

Lezers zullen er dan ook in ruime mate te vinden zijn. Niet voor niets luidt de ondertitel 'Ruimtelijke modellen voor psychologie, marktonderzoek en andere wetenschappen.' Die 'andere wetenschappen' worden op de achterkant gespecificeerd: "politicologie, marketing, sociologie, pedagogiek en voor (sociale) geografie, economie, biologie en taalkunde." Dus zelfs als met iedere gebruikte wiskundige formule het publiek halveert, moet er een breed publiek te vinden zijn. Het lukt de auteur overigens om zonder gebruik van veel rekenkunde (voor sommigen in deze vakgebieden wellicht rekenkunsten) te verhelderen wat er bij MDS met de data gebeurt.

Dezelfde achterkant draagt het boek voor als studieboek na inleidingen in statistiek en methoden en technieken. Ook ik zie als publiek de doctoraalstudent uit genoemde vakgebieden met uitzondering van hen, die zich in methoden en technieken specialiseren. Voor deze laatste groep wordt de stof te oppervlakkig behandeld.

Wil een docent uit één van de genoemde vakgebieden het boek voor een cursus gebruiken, dan adviseer ik om eerst naar toepassingen in het betreffende vakgebied te zoeken. Van der Kloot gebruikt steeds dezelfde data, welke hij zo heeft gekozen dat de data het begrip van de techniek ten goede komt. Zeer educatief en verhelderend, maar kan ik me moeilijk voorstellen dat iemand uit een tabel met afstanden tussen Nederlandse steden een kaart van Nederland wil vervaardigen of op zoek is naar een schaal om het recht op auteurschap van wetenschappelijke publicaties vast te stellen. Tijdens de cursus zal de docent de toepassingsmogelijkheden van MDS in het specifieke vakgebied moeten belichten. Het is mijn ervaring dat veel studenten juist problemen hebben met het vertalen van een probleemstelling naar de toepassing van een bepaalde techniek. Wellicht heeft de auteur met jarenlange ervaring in het toepassen van MDS suggesties.

drs G.L.M. Hilkhuysen
Instituut voor Extramuraal Geneeskundig Onderzoek
Vrije Universiteit Amsterdam

Binnengekomen boeken

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

Arnold, B.C., Castillo, E., Sarabia, J.M.
Conditional specification of statistical models
1999, Springer texts in statistics,
Springer-Verlag, Berlin,
xvi+411 blz, DM 149.00, ISBN 0-387-98761-4.

Via de VVS
25% korting!

Vapnik, V.N.
The nature of statistical learning theory
2000, Statistics for engineering and information science,
Springer-Verlag, Berlin,
300 blz, DM 139.00, ISBN 0-387-98780-0.

Via de VVS
25% korting!

Rosenblatt, M.
Gaussian and Non-Gaussian linear time series and random fields
2000, Springer series in statistics,
Springer-Verlag, Berlin,
255 blz, DM 139.00, ISBN 0-387-98917-X.

Via de VVS
25% korting!

Houtman, D.
Een blinde vlek voor cultuur. Sociologen over cultureel conservatisme, klassen en moderniteit
2000, van Gorcum & Comp. BV, Assen,
180 blz, fl 35.00, ISBN 90-232-3544-4.

Haight, J.
Taking chances. Winning with probability
2000, Oxford University Press, Oxford,
344 blz, £ 8.99, ISBN 0-19-850291.

Cortina, J.M., Nouri, H.
Effect size for Anova designs
2000, Quantitative applications in the social sciences,
Sage, Thousand Oaks, California,
74 blz, £ 8.99, ISBN 0-7619-1550-8.

Paul, W., Baschnagel, J.
Stochastic processes. From physics to finance
2000, Springer-Verlag, Berlin,
xiv+234 blz, DM 98.00, ISBN 3-540-66560-9.

Via de VVS
25% korting!

Fielding, J., Gilbert, N.
Understanding social statistics
2000, Sage, Thousand Oaks, California,
336 blz, £ 16.99, ISBN 0-8039-7983-5.

van de Geer, S.A.
Empirical processes in M-estimation
2000, Cambridge series in statistical and probabilistic mathematics,
Cambridge University Press, Cambridge,
xii+286 blz, £ 35.00, ISBN 0-521-65002-X.

Mathsoft Incorporated (ed)
S-PLUS 2000: student version/Windows. Modern statistics and advanced graphics
1999, Springer-Verlag, Berlin,
xii+557+CD-ROM blz, DM 214.00, ISBN 3-540-14813-2.

Via de VVS
25% korting!

Thorisson, H.
Coupling, stationarity, and regeneration
2000, Springer-Verlag, Berlin,
xiv+516 blz, DM 159.00, ISBN 0-387-98779-7.

Via de VVS
25% korting!

Chen, M.-H., Ibrahim, J.G., Shao, Q.-M.
Monte Carlo methods in Bayesian computation
2000, Springer texts in statistics,
Springer-Verlag, Berlin,
400 blz, DM 159.00, ISBN 0-387-98935-8.

Via de VVS
25% korting!

Good, P.
Permutation tests. A practical guide to resampling methods for testing hypotheses
2000, Springer series in statistics,
Springer-Verlag, Berlin,
xvi+270 blz, DM 139.00, ISBN 0-387-98898-X.

Via de VVS
25% korting!

De door Springer gepubliceerde boeken in bovenstaande lijst kunnen door individuele leden met een korting van 25% op de officiële prijs[2] aangeschaft worden via de Centrale Administratie van de VVS. Bestelformulieren zijn te verkrijgen via Internet, URL `http://www.vvs-or.nl/`, knop *Ledenservice*.

[2]De prijzen zoals vermeld in de lijst van binnengekomen boeken kunnen foutief zijn. Een betrouwbaarder bron van informatie is de Springer website `http://www.springer.de/`.