

Redactie: J.J. Dik

Correspondentie adres: Subfaculteit Wiskunde UvA,
Roetersstraat 15
1018 WB Amsterdam
(020) 522-3090

R.J. Muirhead

Aspects of Multivariate Statistical Theory

John Wiley & Sons, 1982, 673p, £31.00

De multivariate statistiek bevat veel deelgebieden en ieder deelgebied is zo uitgebreid dat er een apart boek over geschreven kan worden. Muirhead beperkt zich in zijn boek "Aspects of Multivariate Statistical Theory" tot enkele onderwerpen uit de klassieke multivariate statistiek. De technieken die hij beschouwt zijn alle gebaseerd op de multivariate normale verdeling en bij de behandeling legt hij de nadruk op de verdelingstheorie. Bij toetsingsproblemen wordt de meeste aandacht besteed aan de likelihood ratio toets; dit is jammer, want voor veel van de behandelde toetsingsproblemen zijn ook andere goede toetsen bekend. Niet alle onderwerpen uit de klassieke multivariate statistiek worden even uitvoerig behandeld (bijvoorbeeld factoranalyse wordt niet en discriminantanalyse wordt summier behandeld), maar de meest elementaire klassieke technieken worden helder en gedegen beschreven. Door de keuze van de onderwerpen en de uitvoerige theoretische behandeling ervan is dit boek een zeer goed leer- en referentieboek voor onderzoekers op het gebied van de (klassieke) multivariate statistiek. Voor toepassers is dit boek minder bruikbaar omdat veel technieken die in de praktijk vaak toegepast worden (bijvoorbeeld technieken op het gebied van de discrete multivariate analyse, meer-dimensionale schaaltechnieken, clusteranalyse, etc.) niet beschreven worden; bovendien is het boek te theoretisch opgezet.

Het boek is gebaseerd op het bekende boek van T.W. Anderson: "An Introduction to Multivariate Statistical Analysis" (1958, Wiley) en er worden min of meer dezelfde onderwerpen behandeld. Het punt waarop Muirheads boek duidelijk afwijkt, en op dit punt wijkt het ook af van de meeste multivariate statistiek boeken, is de toepassing van de zogenaamde zonale polynomen en hypergeometrische functies van matrix argumenten in de statistiek. Dit is een van de belangrijkste ontwikkelingen in de klassieke multivariate statistiek sinds het verschijnen van Andersons boek, en het neemt bij Muirhead een centrale plaats in. Met behulp van deze zonale polynomen en hypergeometrische functies is het mogelijk de niet-centrale verdelingen van toetsingsgrootheden, welke voorkomen bij de op de multivariate normale verdeling gebaseerde technieken, te bestuderen. Deze niet-

centrale verdelingen zijn weer nodig om het onderscheidend vermogen van de op deze grootheden gebaseerde toetsen te berekenen. Aangezien voor de meeste multivariate toetsingsproblemen geen "beste" toets bestaat is het interessant om verschillende toetsen op hun onderscheidend vermogen te vergelijken. Muirhead beperkt zich echter tot de niet-centrale verdelingen en schenkt aan zo'n vergelijking weinig aandacht, misschien omdat dit tot op heden slechts numeriek uitgevoerd is.

Het boek bevat elf hoofdstukken en een appendix. Hiervan kan de inhoud als volgt kort samengevat worden.

In het eerste hoofdstuk wordt de multivariaat normale verdeling geïntroduceerd, worden de belangrijkste eigenschappen ervan bewezen en worden een aantal verwante verdelingen zoals de (niet-)centrale χ^2 en F verdeling en kwadratische vormen van normale grootheden behandeld. Bovendien wordt de klasse van elliptische en spherische verdelingen geïntroduceerd. De verdelingen uit deze klasse hebben dichtheden met, net als de multivariate normale verdeling, elliptische of spherische niveaulijnen, maar kunnen dikkere of dunnere staarten hebben. Ze kunnen gebruikt worden bij onderzoek naar robuustheid van multivariate technieken waarbij men wil weten hoe gevoelig de techniek is voor de veronderstelling van normaliteit. Muirhead besteedt hieraan bij de in hoofdstuk 8 behandelde toetsingsproblemen nog enige aandacht.

In hoofdstuk twee wordt enig wiskundig gereedschap behandeld. Voor veel gebruikte variabelentransformaties wordt de Jacobiaan berekend. Verder wordt nog enige aandacht besteed aan de invariante Haar maat en ook aan Kronecker matrix-producten.

Hoofdstuk drie behandelt de meest aannemelijkheidschatters (met hun verdeling) voor de parameters van de multivariate normale verdeling. Tevens worden de Wishart en de multivariate bèta verdeling uitvoerig behandeld.

Hoofdstuk vier beschouwt het schatten van de parameters van de multivariate normale verdeling uit een beslissingstheoretisch standpunt. Dit gebeurt vrij beknopt; alleen ontoelaatbaarheid (inadmissibility) van meest aannemelijkheidschatters wordt aangetoond.

De gewone, multiple, en partiële correlatie coëfficiënten worden in hoofdstuk vijf behandeld. Voor ieder van deze drie wordt de verdeling van steekproef correlatie coëfficiënt berekend en tevens wordt een toets voor hypothesen over de populatie correlatie coëfficiënt beschreven.

Hoofdstuk zes geeft een vrij algemene maar beknopte inleiding in de invariantie theorie. Vervolgens wordt dit toegepast op de multiple correlatie coëfficiënt en Hotellings t^2 .

In hoofdstuk zeven worden de zonale polynomen en hypergeometrische functies van matrix argumenten geïntroduceerd en enkele eigenschappen afgeleid. Dit hoofdstuk legt de basis voor het bestuderen van de niet-centrale verdelingen welke

bij de toetsingsproblemen in de volgende hoofdstukken naar voren komen.

In hoofdstuk acht komen enkele standaard toetsingsproblemen over de parameters van de multivariate normale verdeling aan de orde (o.a. toets op gelijkheid van een aantal covariantie matrices, sphericiteitstoets, toets op covariantie matrix en toets op verwachting en covariantie matrix).

De nadruk ligt op de likelihood ratio toets en helaas komen andere toetsingsgrootheden nauwelijks aan de orde. Voor de likelihood ratio toetsingsgrootheid worden de verdelingen onder nulhypothese en alternatief bestudeerd.

Hoofdstuk negen besteedt aandacht aan principale componentenanalyse en in het bijzonder aan de (asymptotische) verdeling van de bij deze techniek berekende eigenwaarden en eigenvectoren van de steekproef covariantie matrix.

In hoofdstuk tien komt het multivariate lineaire model uitvoerig aan de orde. De nadruk ligt nu niet totaal op de (niet-)centrale verdeling van de likelihood ratio toetsingsgrootheid, maar ook worden verdelingen van andere toetsingsgrootheden beschouwd. Echter aan het vergelijken van het onderscheidend vermogen van op deze grootheden gebaseerde toetsen schenkt Muirhead weinig aandacht. In dit hoofdstuk wordt tevens kort ingegaan op discriminantanalyse en het toetsen van gelijkheid van een aantal normaal verdeelde populaties.

Het laatste hoofdstuk behandelt het toetsen van onafhankelijkheid tussen een aantal groepen van normale variabelen en behandelt tevens canonische correlatie. Bij het eerste probleem wordt weer alleen de likelihood ratio toets bekeken en bij de exploratieve canonische correlatie ligt de nadruk op de verdeling van de steekproef canonische correlaties.

Tot slot geeft Muirhead in een appendix de nodige matrixtheorie.

Bert Schriever, C.W.I.

J. Whitehead

The Design and Analysis of Sequential Clinical Trials

Ellis Horwood Limited, 1983, 272p, \$44.50.

Dit boek is een handleiding voor het opzetten en analyseren van sequentiële klinische trials met als voornaamste doelgroep medische statistici.

Geschreven in een even leesbare stijl als het bekende boek van Armitage (1975) verwerkt dit boek recente ontwikkelingen op het gebied van klinische sequentiële trials zodat het ook gezien kan worden als een aanvulling op en uitbreiding van Armitage's handleiding.

Met uitzondering van hoofdstukken 1 en 2 is de volgorde van presentatie in elk hoofdstuk steeds: inleiding, "kookboek" formulering, uitvoerige voorbeelden, theorie.

Bovendien wordt elk hoofdstuk afgesloten met een paragraaf interessante en vaak uitgebreide bibliografische notities en opmerkingen.

In tegenstelling tot het boek van Armitage bevat het boek van Whitehead grotendeels voldoende theoretische verantwoording van de gebruikte procedures. De presentatie van de theorie echter is nogal kompakt en vereist een behoorlijke scholing in de mathematische statistiek.

Het boek bevat 8 hoofdstukken met de volgende inhoud:

Hoofdstuk 1 is een algemene inleiding.

Hoofdstuk 2 bevat een handzaam overzicht van de gangbare methoden om gestratificeerd te randomiseren.

Hoofdstukken 3, 4 en 5 vormen de kern van het boek.

Door zich te beperken tot asymptotische sequentiële toetsen met een driehoekig akseptatie gebied waarbij steeds de efficiënt score statistic Z tegen de Fischer informatie V wordt geplot en door $\alpha = \beta$ te kiezen geeft Whitehead hier een uniforme en overzichtelijke inleiding in het opzetten en analyseren van sequentiële trials voor het vergelijken van twee behandelingen. Deze driehoek "toets" is een speciaal geval van de door Anderson (1960) geïntroduceerde gemodificeerde sequential probability ratio toets.

Aandacht wordt besteed aan crossovertrials met een kwantitatieve of dichotome respons en aan trials met een kwantitatieve, ordinale of dichotome respons waarbij vergelijking "tussen personen" plaats vindt. Ook de aanpak van survival data wordt besproken. Alle belangrijke aspecten van sequentiële procedures voor vergelijkende trials komen aan de orde: een- of tweezijdige toetsen, keuze van inspektie intervallen, interim analyses, truncatie, "overshoot" correctie, inferentiële versus besliskundige aanpak, punt- en intervalschattingen, overrunning, zeer vroeg stoppen, het berekenen van overschrijdingskansen.

Naast de driehoektoets komt ook de klassieke probability ratio toets aan de orde.

De keuze van $\alpha = \beta$ lijkt geen zwaarwegende beperking van de mogelijkheden in te houden. Deze beperking is vermoedelijk mede ingegeven door het gekompliceerde karakter van de driehoektoets als $\alpha \neq \beta$. Voor de probability ratio toets wordt wel een korte bespreking van het geval $\alpha \neq \beta$ gegeven.

Aangezien de inferentiële statistiek grotendeels op asymptotische methoden en resultaten steunt kan de introductie van asymptotische sequentiële technieken niet als een beperking gezien worden. Bovendien zijn tabellen voor het hanteren van de "exacte" sequentiële methoden grotendeels met behulp van theoretische benaderingen en computer simulaties tot stand gekomen.

Hoofdstukken 6, 7 en 8 presenteren mogelijkheden om de "schrane" standaard sequentiële analyses aan te vullen wanneer er sprake is van prognostische variabelen en meer dan twee behandelingen. Daartoe worden partiële sequentiële

procedures geïntroduceerd: modificaties van de door Armitage geïntroduceerde restricted opzetten.

Bij Whitehead gaat het om een aanpassing van het akseptatie gebied zodanig dat men bij een groot verschil tussen de behandelingen nog steeds tot een snelle beslissing kan komen maar bij een klein verschil meer waarnemingen doet dan bij een niet partiële aanpak. De voorgestelde uitvoerige analyses steunen gedeeltelijk op deze meer informatieve steekproef aantallen.

De analyse van prognostische factoren, gefit met weglating van de faktor behandeling, is namelijk asymptotisch onafhankelijk van de analyse van de behandelingen alleen gebaseerd op Z en V . Na afloop van de trial wordt dus een standaard analyse (vaste steekproefgrootte) van de prognostische factoren met weglating van de faktor behandeling uitgevoerd. Daarnaast worden behandelings-effecten, gecorrigeerd voor prognostische factoren, met behulp van Z en V geanalyseerd. Bij een grote steekproef zijn deze twee analyses onafhankelijk.

Ook interim analyses van prognostische variabelen, in het bijzonder met betrekking tot selectie en inklusie, worden besproken.

De benodigde theorie is geformuleerd in termen van een generalized linear model en het gebruik van GLIM wordt gepropageerd.

Bij het vergelijken van meer dan twee behandelingen beperkt Whitehead zich tot eliminatie procedures gebaseerd op paarsgewijze vergelijking van de behandelingen met partiële sequentiële procedures. Dit gedeelte van het boek is nogal schetsmatig en exploratief, hetgeen echter ook van toepassing is op de stand van zaken bij het onderzoek.

Tenslotte worden bij de partiële procedures ook herhaalde significantie toetsen en groeps sequentiële opzetten gepresenteerd als een geplande gestructureerde mogelijkheid voor tussentijds "spieken" en vroegtijdig stoppen.

Illustratie van de analyse van trials met prognostische factoren gebeurt aan de hand van 2 gesimuleerde data bestanden vanwege het ontbreken van geschikt gepubliceerd materiaal.

Deze gesimuleerde data zijn, ten gerieve van eigen analyses door de lezer, in de appendix opgenomen. Verder bevat de appendix de tabellen die nodig zijn voor het gebruik van de gepresenteerde technieken.

Deze tabellen konden, mede als gevolg van de gehanteerde restricties, uitvoeriger zijn dan de tabellen in Armitage zodat men een veel grotere keuze kan maken uit geschikte opzetten bij de planning. Het boek is nagenoeg vrij van storende fouten en onzorgvuldigheden en helder en goed leesbaar geschreven.

Sterk aanbevolen voor elke statisticus die frekwent bij de analyse van klinische trials betrokken is. Het werk is ook uitstekend geschikt als leerboek voor degenen die zich willen inwerken op het gebied van klinische trials en sequentiële procedures voor de analyse van klinische trials.

Helaas heeft de auteur geen vraagstukken opgenomen.

Literatuur

- ARMITAGE, P. (1975) Sequential Medical Trials, Blackwell Scientific Publications.
- ANDERSON, T.W. (1960) A modification of the sequential probability ratio test to reduce the sample size, Ann. Math. Statist. 31, 165-197.

A. Volovics, R.U. Limburg

*Lothar Sachs*Applied statistics, a handbook of techniques

Springer-Verlag; New York, Heidelberg, Berlin, 1982, 38 + 706p, 59 figuren, f143,95

Dit boek is gebaseerd op de nieuwe gereviseerde 5e editie van Sachs' Angewandte Statistik, (1978). De auteur schrijft in zijn inleiding dat hij dit boek heeft geschreven als een leerboek met veel verwijzingen voor onderzoekers met een geringe wiskundige achtergrond.

Het boek blijkt encyclopedisch van aard te zijn. Het beschrijft de klassieke toetsen en veel verdelingsvrije toetsen en geeft de bijbehorende tabellen (76 grote en kleine tabellen). Men kan zich afvragen of met het vele gebruik van elektronische zakrekenapparaten sommige tabellen nog nodig zijn, zoals die van $\log x$, $e^{-\lambda}$ en $2n \log n$. De auteur heeft zijn dossierkast met artikelen en literatuur verwijzingen in dit boek verwerkt, maar of hij de artikelen zelf wel allemaal begrepen heeft wordt niet geheel duidelijk.

Als naslagwerk voor een statistisch geschoold toepasser is dit boek te gebruiken, maar als leerboek is het ongeschikt en zelfs af te raden. Er wordt bijna nergens onderscheid gemaakt tussen exacte toetsen en benaderingen daarvan. Achterin is er een kleine verzameling opgaven; de opgaven zijn vaak slecht gesteld en vele antwoorden deugen niet. Zo wordt in het boek de (steekproef)-standaardafwijking gedefinieerd als $s = \sqrt{\sum (x_i - \bar{x})^2 / (n-1)}$ maar in de antwoorden van de desbetreffende opgaven 1.7 t/m 1.10 en 5.4 waar s berekend moet worden, wordt de waarde gegeven van $\sqrt{\sum (x_i - \bar{x})^2 / n}$. Verder zijn ook de antwoorden van opgaven 3.2, 3.3, 4.2 en 5.3 fout.

Sachs haspelt begrippen als "standard deviation" voor de populatie en standaardafwijking door elkaar (zie pag. 69). Op pag. 70 lezen we bij s^2 : "The use of "n-1" in (1.32a) gives us an unbiased estimator of σ^2 and therefore an unbiased estimate of s^2 ". Verder wordt het betrouwbaarheidsinterval voor σ^2 correct gegeven met behulp van s^2 en χ_{n-1}^2 (pag. 259), plompverloren (en nog wel omkaderd) gevolgd door: "The estimate for σ^2 is obtained according to $\hat{\sigma}^2 = s^2 (n-1) / \chi_{n-1;0.5}^2$ (3.17a)".

De argeloze lezer wordt op pag. 96 geleerd om bij onafhankelijke steekproeven de verwachting en bijbehorende standaardafwijking te schatten bij vermenig-

vuldiging en deling van de steekproefgemiddelden, zonder dat er bij staat dat dit slechts bij benadering geldt.

Inconsequenties zijn er in §1.3.8.7 bij de schatting van scheefheid en kurtosis met s die alleen in deze paragraaf gedefinieerd wordt met noemer n , maar bij de samenvatting van de formules op pag. 106 ineens, nu ten onrechte, met noemer $n-1$ verschijnt. Op pag. 186 moet bij de Poisson-benadering "For $\lambda \approx 0,5$ " zijn: "For accurate approximation".

De schrijver huldigt op pag. 261 het standpunt: "If the hypothesis (H_0) is rejected then the confidence interval (CI) is computed....," en "Since H_0 is rejected at the 10% level, we specify the 90% CI by....". Bij de verdelingsvrije toets op gespreidheid van Siegel en Tukey (1960) op pag. 286 wordt slechts een heel klein tabelletje vermeld van kritieke waarden maar er wordt niet aangegeven dat voor de Siegel-Tukey grootheid geldt $I_1 = U + \frac{1}{2}n(n+1)$, waarbij U de (Wilcoxon)-Mann-Whitney grootheid is, waarvan de kritieke waarden zeer uitgebreid getabelleerd zijn in tabel 63. De bekende verdelingsvrije toets op gespreidheid van Mood (1954) met uitgebreide tabellen door Laubscher, Steffens en De Lange (1968) in *Technometrics* 10, 497-508 wordt niet vermeld!

Ten onrechte suggereert Sachs op pag. 291 dat de Kolmogoroff-Smirnoff toets ook toegepast mag worden voor discrete verdelingen zonder te vermelden dat deze toets dan conservatief is. In §6.1.1 behandelt Sachs de $k \times 2$ tabellen: de binomiale homogeniteits toets. Het is wel een blunder om in deze paragraaf als voorbeeld $k=3$ te behandelen met een zogenaamde 3×2 tabel die echter een 2×3 tabel blijkt te zijn, nl. 2 onafhankelijke steekproeven met drie klassen. Dit geval wordt pas in §6.2 bij de $r \times c$ tabellen besproken. Opvallend is verder dat in het geheel niet over het log-lineaire model gesproken wordt.

Samenvattend kan gezegd worden dat dit boek als een handboek voor statistische technieken gebruikt kan worden door een ervaren statisticus, maar het boek moet afgeraden worden voor slechts elementair statistisch geschoolde toepassers.

L.R. Verdooren, L.H.W.

A. Buijs

Kwantitatieve toepassingen: een aantal toepassingen van wiskunde en statistiek in de bedrijfskunde

Stenfert Kroese, Leiden, 1983, x en 249 pp.

De auteur geeft aan dat dit "boekje" (: zijn eigen terminologie) is bedoeld om een aantal toepassingen van wiskundige en statistische methoden in de bedrijfskunde te behandelen; en wel op het nivo van een eerste kennismaking.

Het resultaat is een boek waarin 8 inleidingen worden gegeven over de volgende onderwerpen:

- * inleiding met "samenvatting" van onderwerpen uit de kansrekening (31 pp).
- * lineaire programmering (48 pp).
- * netwerken (23 pp).
- * simulatie (31 pp).
- * beslissingsproblemen algemeen (25 pp).
- * voorraadproblemen (37 pp).
- * wachttijdproblemen (23 pp).
- * markovketens (27 pp).

Deze 8 inleidingen worden vrijwel geheel losstaand van elkaar besproken, hetgeen de auteur ook zo bedoeld heeft, zodat men een onderwerp los van de andere onderwerpen kan bestuderen. Alleen in het hoofdstuk over simulatie worden toepassingen besproken die de voorgaande onderwerpen betreffen.

Het aardige van het boekje is dat de onderwerpen uitsluitend toepassingsgericht aan de orde worden gesteld en telkens aan de hand van uitgewerkte voorbeelden worden besproken. Het eerste hoofdstuk, de onderwerpen uit de kansrekening, had beter achterwege kunnen blijven. Iemand die daar te weinig van weet, moet zich die kennis eerst op een goede manier eigen maken. Dit hoofdstuk is daar niet voor geschikt. Het is te summier en veel te fragmentarisch. Voor hen die de statistiek wel voldoende beheersen, is het overbodig.

Voor de andere hoofdstukken geldt wat ik reeds eerder stelde. Zij die zich voor het eerst willen oriënteren in deze onderwerpen krijgen op een prettige manier een goed beeld van de probleemstellingen en de oplossingsmethoden; prettig vanwege de vele voorbeelden.

Elk hoofdstuk wordt telkens besloten met een aantal opgaven. Ik denk dat het voor de geïnteresseerde lezer een gemis is dat er achter in het boek geen antwoorden van de opgaven vermeld staan. Ook vind ik het lijstje boeken dat is opgenomen ter aanbeveling voor verdere studie wel erg beperkt. Bijv.: van de leergang beslis-kunde van het Mathematisch Centrum wordt alleen naar deeltje 4 verwezen.

Maar het zal inmiddels duidelijk zijn dat ik vind dat dit (Nederlandstalige) boek zeker aanbevolen kan worden aan hen die eens willen kennismaken met deze besliskundige onderwerpen.

E. Vogelvang, V.U.

G. Kallianpur, P.R. Krishnaiah & J.K. Ghosh (eds.)

Statistics and Probability: Essays in Honor of C.R. Rao

North-Holland Publishing Company-Amsterdam, New York, Oxford, 1982, XII + 720p.
f310,--.

De welbekende statisticus C.R. Rao is de auteur van verschillende boeken, zoals o.a. "Linear statistical inference and its applications", "Advanced statistical methods in Biometric research" en "Generalized inverse of matrices and its applications". Verder heeft hij meer dan 200 artikelen geschreven.

Op 10 september 1980 werd C.R. Rao zestig jaar. Ter ere van deze zestigste verjaardag van Rao is dit te bespreken boek, een verzameling van artikelen, geschreven.

Het boek bestaat uit 66 artikelen geschreven door statistici en waarschijnlijkheidstheoretici uit de gehele wereld. De onderwerpen zijn zeer uiteenlopend van aard. Enkele bijdragen, die een indruk geven van de gevarieerdheid van de inhoud van het boek, zijn:

- ANDERSON, T.W. & G.P.H. STYAN, Cochran's theorem, rank additivity and tripotent matrices (pp).
 BARTLETT, M.S., The development of population models (13 pp).
 BUNKE, O. & M. MÖHNER, Minimax estimators in linear models under normalized quadratic loss functions (7 pp).
 DRYGAS, H., Asymptotic confidence-intervals for the variance in the regression model (9 pp).
 DUGUÉ, D., On the way towards a non-parametric multidimensional test of normality (9 pp).
 HANNAN, E.J., Fitting multivariate ARMA models (11 pp).
 KHATRI, E.G., A theorem on quadratic forms for normal variables (9 pp).
 PRAKASA, B.L.S., Maximum probability estimation of diffusion processes (17 pp).
 RAMACHANDRAN, B., An integral equation in probability theory and its applications (9 pp).
 ROZANOV, Yu.A., On multidirectional Markov processes (5 pp).
 WATSON, G.S., The estimation of Palaeomagnetic pole positions (11 pp).

Opgenomen is een inhoudsopgave van de artikelen naar alfabetische volgorde van de auteur(s). Achterin het boek staat een lijst met boeken en artikelen door C.R. Rao geschreven.

P. van der Laan, L.H.W.

A.R.G. Heesterman

Matrices and Simplex Algorithms

D. Reidel Publishing Company, Dordrecht

De doelgroep die een schrijver bij het schrijven van een boek voor ogen heeft, is in hoge mate bepalend voor de uiteindelijke inhoud. Voor dit boek heeft de schrijver gekozen voor economie-studenten. Het merkwaardige is echter, dat hij zich de consequenties van deze keus niet of in ieder geval onvoldoende heeft gerealiseerd; zo ontbreken in het boek economische voorbeelden vrijwel geheel. Ook wordt er nauwelijks aandacht besteed aan het formuleren van de mathematische modellen die opgelost worden.

Anderzijds wordt de nodige wiskundige voorkennis uitgebreid behandeld en is de uitleg van de gegeven algoritmen zeer uitvoerig. Daarenboven wordt van ieder algoritme een computerprogramma in Algol 60 gegeven; de schrijver beoogt daarmee twee doelen: eerstens ondersteuning van de uitleg van het algoritme en tweedens de mogelijkheid de algoritmen op een computer te gebruiken. Het eerste doel komt mij voor economie-studenten onwerkelijk voor. Het doorwerken van de programma-teksten van soms meer dan tien bladzijden zal zelfs voor meer computer geöriënteerde studenten een zware opgave zijn en veelal niet bijdragen tot een beter begrip van het betreffende algoritme. Het tweede doel vereist een ruime kennis

van zaken van een programmeur die de programma's zou moeten aanpassen aan een andere computer dan de schrijver ter beschikking had. Het boek bevat geen aanwijzingen of gegevens voor een dergelijke aanpassing.

Voor wat betreft de uitvoering van het boek moet ik enigszins kritisch zijn; de tekst is niet gezet, maar rechtstreeks van een manuscript in schrijfmachine-schrift vermenigvuldigd, wat de duidelijkheid en overzichtelijkheid niet ten goede komt. Zo is het begin van een sectie van een hoofdstuk moeizaam te vinden, door het ontbreken van een groter lettertype voor de titels van de secties.

Aangezien men mag aannemen dat het samenstellen van een leerboek van 800 bladzijden veel tijd en inspanning kost, is het jammer dat het resultaat niet geestdriftig maakt.

A. Volgenant, UvA

A. Araujo & E. Giné

The central limit theorem for real and Banach valued random variables

John Wiley & Sons, 1980, ISBN 0-471-05304-X.

Het boek begint met een hoofdstuk over zwakke convergentie en "tightness" van kansmaten op metrische ruimten, i.h.b. op Banachruimten. Het tweede hoofdstuk is gewijd aan de centrale limiet stelling in het reële geval. De gebruikelijke versies van deze stelling worden opgevoerd als speciale gevallen van het volgende zeer algemene resultaat:

Theorem (general central limit theorem, real case).

Let $\{X_{nk}\}$, $k = 1, \dots, k_n$, $n = 1, \dots$, be an infinitesimal system and $\{a_n\}$ a sequence of constants. Then $\{L(S_n - a_n)\}_{n=1}^{\infty}$ converges weakly if and only if:

(i) There exists $\sigma \in \mathbb{R}$ such that

$$\lim_{\delta \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} \sum_k E(X_{nk\delta} - EX_{nk\delta})^2 = \sigma^2.$$

(ii) There exists a Lévy measure μ such that

$$w - \lim_{n \rightarrow \infty} \sum_k L(X_{nk}) \{ \{|x| > \delta\} \} = \mu \{ \{|x| > \delta\} \}$$

for every $\delta > 0$ such that $\mu[-\delta, \delta] = 0$.

(iii) If $\delta > 0$ satisfies $\mu[-\delta, \delta] = 0$, there exists $a_\delta \in \mathbb{R}$ such that

$$\lim_{n \rightarrow \infty} (ES_{n,\delta} - a_n) = a_\delta.$$

Then if this is the case, we have

$$w - \lim_{n \rightarrow \infty} L(S_n - a_n) = N(a_\delta, \sigma^2) * c_\delta \text{ Pois} \mu.$$

[Hier is $x_{nk\delta} = x_{nk} \cdot \chi(|x| \leq \delta)$, $S_{n,\delta} = \sum_j x_{nj\delta}$ en

$$c_{\delta} \text{Pois} \mu = w - \lim_{\varepsilon \rightarrow 0} \delta_{\varepsilon, \delta} * \text{Pois}(\mu | |x| > \delta),$$

waarbij $c_{\varepsilon, \delta} = \int_{\varepsilon < |x| \leq \delta} x \, d\mu(x).$]

In hoofdstuk 3 wordt na een inleiding over vectorwaardige stochasten allereerst een generalizatie gegeven van bovenstaande stelling voor stochasten met waarden in algemene Banachruimten. Vervolgens worden bruikbaarder varianten afgeleid voor Banachruimten met speciale eigenschappen. Het in dit verband relevante begrip is het type, resp. cotype van de ruimte. [E heeft type p , $1 \leq p \leq 2$, desd als er een constante c bestaat zó dat $E \|\sum_{i=1}^n x_i\|^p \leq c \sum_{i=1}^n E \|x_i\|^p$ voor elk eindig onafhankelijk stelsel E-waardige stochasten $\{x_i\}_{i=1}^n$ met $EX_i = 0$, $i = 1, \dots, n$. Voor de definitie van cotype p , $p \geq 2$, vervange men $\leq$ door $\geq$.] Er geldt nu:

- (a) Een Banachruimte E heeft type 2 desd als voor elke E-waardige stochast X met $EX = 0$ en $E \|X\|^2 < \infty$ de centrale limietstelling geldt (waarmee we hier bedoelen dat $L(\frac{S_n}{\sqrt{n}})$ zwak convergeert als $S_n = \sum_{i=1}^n x_i$, $\{x_i\}$ onafhankelijk en $L(x_i) = L(X)$, $i = 1, \dots, n$).
- (b) Een Banachruimte E heeft cotype 2 desd als elke E-waardige stochast X waarvoor de centrale limietstelling geldt voldoet aan $EX = 0$, $E \|X\|^2 < \infty$.

Het is bekend dat E desd type 2 en cotype 2 heeft als E isomorf is met een Hilbertruimte. Dit levert dus een probabilistische karakterisering van Hilbertruimten.

Het boek is in het algemeen helder geschreven. Er zijn weinig drukfouten en veel opgaven. Vooral het derde hoofdstuk bevat veel merendeels zeer recente informatie en maakt dit boek tot een waardevol naslagwerk. Toch is het jammer dat de auteurs geen weerstand hebben kunnen bieden aan de verleiding elk resultaat in zijn meest algemene vorm te presenteren. Hierdoor valt het de niet-gespecialiseerde lezer niet gemakkelijk het zicht op de hoofdzaken te behouden. Zo zijn b.v. de hierboven vermelde karakterisering van type 2 en cotype 2 slechts met moeite te vinden. ((a) staat verborgen in een algemenere bewering en (b) is een opgave). Kortom, een nuttig boek voor specialisten, maar als inleiding minder geschikt.

J.P. Brans (ed.)

Operational Research '81: Proceedings of the Ninth IFORS International Conference on Operational Research, Hamburg, Germany, July 20-24, 1981
North-Holland, Amsterdam, 1981, xx+984p. f 250,--

Wie midden in de zomer van 1981 naar Hamburg is gegaan om deel te nemen aan het negende wereldcongres van IFORS (International Federation of Operational Research Societies), en wie deze driejaarlijkse congressen onafgebroken vanaf de eerste keer te Oxford in 1957 heeft bezocht, heeft nu een mooi rijtje Proceedings in de kast staan.

De eerste, "Proceedings of the First International Conference on Operational Research" (Oxford, 1957) is een aardig boekje met 29 papers en een foto van de deelnemers, waarop we de jongensachtige gezichten van A.R. v.d. Burg, W. Monhemius, A.H. Schaafsma en J. Sittig kunnen herkennen.

De jongste, "Operational Research '81", is een kloek boek van duizend bladzijden met 69 papers, die geselecteerd zijn uit 350 op het congres gepresenteerde papers. Alvorens tot een bespreking van de Proceedings over te gaan, eerst iets over het congres zelf.

Het IFORS IX-congres was een groter evenement dan ooit tevoren. Dit kwam voornamelijk omdat het congres "opengegoid" was voor "contributed papers" van een ieder die wilde, naast de klassieke "national contributions" ingediend door de nationale verenigingen. De organisatoren waren zelfs verder gegaan en hadden een 40-tal ondervoorzitters benoemd die op uitnodiging hadden samengesteld:

- a) 7 mainstream sessions (zie tabel 1),
- b) 13 technical sessions (zie tabel 2),
- c) 14 workshops (zie tabel 3),
- d) HORSE, Hamburg Operational Research Software Exhibition, met "live" demonstraties van OR standaard software van 24 deelnemende bedrijven (georganiseerd door H.J. Zimmermann).

De 350 papers waren ingedeeld in 120 zittingen van elk anderhalf uur. Voor elk paper werd exact een half uur ter beschikking gesteld. Er werden maximaal 16 zittingen simultaan gehouden.

Wat vinden we nu van het congres in de Proceedings terug?

- a), b), van de mainstream en technical sessions ieder twee geselecteerde papers;
- c) van de workshops 19 pagina's verslagen;
- d) van HORSE 3 pagina's verslag.

Bovendien zijn alle 36 "national contributions" opgenomen. Van de "contributed papers" is niets opgenomen. In totaal is $69/350 = 20\%$ van de papers opgenomen, ingedeeld in 13 hoofdstukken (zie tabel 4).

Door de eenzijdige selectie (alle national contributions, geen contributed papers) is het onwaarschijnlijk dat het de top-20% is. Weliswaar zijn alle

papers beoordeeld ("refereed"), maar dit heeft kennelijk bij de national contributions niet geleid tot afwijzingen. Nu zijn de "national contributions", als het goed is, door hun "national society" geselecteerd, maar ik heb in Nederland drie keer meegemaakt dat er voor de "national contribution" maar één kandidaat was. Er zijn dus goede en slechte papers bij, maar wij zullen hier geen individuele papers bespreken.

De boektechnische uitvoering is abominabel, tenminste voor een boek dat een kwartje per bladzijde kost. Een luxe kft omhult een verzameling verkleind tikwerk, door de auteurs "camera-ready" ingeleverd, met de verschillende lettertypen, regelafstanden, alinea-indelingen, en gradaties van zwarting van dien. Uitgevers zouden moeten wachten met het overhevelen van de lasten van het persklaar maken naar de auteurs, totdat alle auteurs over behoorlijke tekstverwerkende systemen beschikken.

Laatste kritische opmerking: het Frans ware af te schaffen als officiële tweede taal van IFORS. Slechts 3 van de 69 papers zijn in het Frans, maar alle titelbladen worden ontsierd door tweetalige titels en abstracts. Goed Engels is al moeilijk genoeg voor de meeste mensen in de wereld.

C.B. Tilanus, T.H.E.

Tabel 1. Mainstream sessions

Onderwerp

1. Multinational enterprises
2. Modelling for decision making in organizations
3. World societal problems
4. OR in developing countries
5. Regional planning and infrastructure
6. Locational models
7. Energy planning models

Georganiseerd door

- H. von Falkenhausen
B. Roy and J.L. Lemoigne

R. Tomlinson
A.N. Elshafei
F. Hanssmann
D. Erlenkotter
P. Meier

Tabel 2. Technical sessions

Onderwerp

1. OR in China
2. Decision analysis
3. Multiple objective programming
4. Redundancy in mathematical programming
5. Integer programming and combinatorial optimization
6. Linear programming and non-linear programming
7. Air and railroad transport
8. Planning and urban transportation
9. Production management
10. Computational complexity and heuristics
11. Graphs and networks
12. Planning in financial institutions
13. Simulation

Georganiseerd door

- Ku Chi Fa
R.L. Keeney
J. Spronk
S. Zionts and J. Telgen
M. Gondran

U. Eckhardt

T. Hasegawa
O. Saatcioglu and Ph. Vincke
L.F. Gelders
J.K. Lenstra
F. Maffioli
J.M. Williamson
R.G. Sargent

Tabel 3. Workshops

<u>Onderwerp</u>	<u>Georganiseerd door</u>
1. Multinational enterprises	H. von Falkenhausen
2. Modelling for decision making in organizations	B. Roy and J.L. Lemoigne
3. World societal problems: The challenge of change	R. Tomlinson
4. Introduction and promotion of OR	C. Carlsson, I. Pappas and A.N. Elshafei
5. O.R. in developing countries	A.N. Elshafei
6. Energy and economy modelling for developing and industrializing countries	J.E. Samouilidis
7. Regional planning and infrastructure	F. Hanssmann
8. Locational models	D. Erlenkotter
9. Resource allocation models	J. Weglarz
10. Energy planning models	P. Meier
11. Financial planning in an insurance business	M. Goovaerts
12. Validation and assessment of models	S. Gass
13. Military OR	J.D. Kettelle
14. New results in forecasting	E.F. Kasper

Tabel 4. Indeling van de papers in de Proceedings

<u>Onderwerp</u>	<u>Aantal papers</u>
1. The history, evolution and future of OR	3
2. OR in organizations	7
3. OR in developing countries	3
4. OR and management in various countries	4
5. OR in the public sector	8
6. Decision analysis and multiple objective programming	8
7. Mathematical programming and optimization	8
8. Locational problems	5
9. Transport and traffic	7
10. Graphs, networks and games	5
11. Production management	4
12. Finance	3
13. Simulation, information and modelling	4
	<u>69</u>

R.J. Serfling

Approximation Theorems of Mathematical Statistics

John Wiley & Sons, New York, 1980

Is voor een statistisch probleem geen exacte oplossing voorhanden, dan wordt veelal een beroep gedaan op asymptotische methoden. Het hier te bespreken boek behandelt een groot aantal limietstellingen die te pas komen in de schattings- en toetsingstheorie, en voorziet derhalve in een behoefte. De opzet van het boek is als volgt:

Na een bespreking van de basisconvergentiebegrippen en stellingen in hoofdstuk 1, volgt in hoofdstuk 2 een systematische behandeling van de asymptotiek m.b.t. steekproefgrootheden: de empirische verdelingsfunctie, steekproefmomenten, kwantielen, order statistics etc. Bahadur's representatie van kwantielen krijgt hier ook een plaats. Na een hoofdstuk over transformaties van de steekproefgrootheden, komen in hoofdstuk 4 maximum likelihood schatters en likelihood ratio

toetsen aan de orde. Uiteraard besteedt Serfling ruime aandacht aan U-statistics. In dit hoofdstuk (5) wordt ook de projectie-methode ten tonele gevoerd; in de hoofdstukken over L- en R- schatters keert dit onderwerp terug. Veel statistics kunnen geschreven worden als $T(F_n)$, waarbij T een functionaal is en F_n de empirische verdelingsfunctie gebaseerd op n waarnemingen. Differentieerbaarheid van T is het sleutelwoord in hoofdstuk 6 bij de bespreking van de asymptotische verdeling van $T(F_n) - T(F)$. In de hoofdstukken 7, 8 en 9 vindt een behandeling van M-, L- en R-schatters plaats. Het boek wordt afgesloten met het 10e hoofdstuk dat als onderwerp heeft asymptotische efficiency.

Al met al een flink pakket voor één éénsemestercursus, hoewel het opmerkelijk is dat aan contiguity en asymptotic expansions vrijwel geen aandacht besteed wordt. Een dergelijk boek kan men hanteren om thuis te raken in de asymptotische methoden in de statistiek, of als naslagwerk. Voor het eerste gebruik is een flinke dosis zelfwerkzaamheid geboden, wil men althans de bewijzen van de stellingen 'af' maken: bijna alle opgaven van het boek zijn van het type: check details of ..., complete details of ..., supply details of ..., prove lemma ..., verify ..., corresponderend met veelvuldig voorkomende zinsneden in de tekst als: It is straightforward (Problem ...), it follows (Problem ...), thus (check), we have (Problem ...) enz. Voor het gebruik van het boek als naslagwerk zijn de onnauwkeurigheden in het boek storend. Zo dient bij de bekende stelling van Pitman over de gelijknamige efficiency een eis van monotonie of continuïteit van de powerfuncties voor vaste n opgenomen te worden. Een makkelijk verkeerd te interpreteren bewering treffen we aan in de voorbeelden 7.1.2. D en E: het getrimd gemiddelde wordt gerelateerd aan een M-schatter met ψ -functie die 0 is buiten $[-k, k]$ en x op $[-k, k]$ terwijl het Winsorized gemiddelde gekoppeld wordt aan een M-schatter met ψ -functie, die -k is op $(-\infty, k]$, x op $[-k, k]$ en k op $[k, \infty)$. Nu is de influence curve van het getrimde gemiddelde juist gelijk aan de influence curve van de laatstgenoemde M-schatter (bij geschikte keuze van het trimmingspercentage en k), en niet aan die van de eerstgenoemde! Het boek bevat een uitgebreide lijst van referenties (ruim 350).

Wilbert Kallenberg, VU.

R. Startup & E.T. Whittaker
Introducing social statistics

George Allen & Unwin, London, 1982, £ 4,95 (plm. f 28,-)

Op zich houdt de verschijning van dit n-de, niet onmisbare boek over dit onderwerp weinig nieuws waarde in. Toch verdient "introducing social statistics" wel even aandacht. Want het is zeldzaam compact met zijn 183 pagina's leertekst, zeldzaam helder en doordacht in de formuleringen, en het is ook als leerboek

compleet, met oorspronkelijke voorbeelden, begrippenlijstjes, een index van drie pagina's, vier of vijf opgaven bij elk hoofdstuk en uitwerkingen daarvan, en de nodige tabellen.

Het boek omvat elf hoofdstukken, en begint met een introductie over methodologie en datatypes die acht pagina's telt. Dan volgen drie hoofdstukken beschrijvende statistiek en kansrekening. Dit gedeelte omvat 61 p. en geeft als bijzonderheden daarin bijv. toch nog Pearson's scheefheidscoëfficiënt, twee opvattingen van het begrip kans, conditionele kansen, een klein beetje combinatoriek en de Poisson verdeling. In het hoofdstuk 'kansrekening' bevindt zich een lang rekenkundig betoog dat naar de conclusie voert dat verwachting van een binomiale verdeling np is. Dit bewijs is misschien leerzaam om na te rekenen, maar het lijkt toch een beetje verdwaald. De schrijvers laten in het volgende terecht na op dezelfde manier te bewijzen dat de variantie van de binomiale verdeling ook te bepalen is. Deze tekortkomingen worden echter wel goedgemaakt door het stuk dat ze als afsluiting van dit hoofdstuk aan de Poisson verdeling wijden. Hoofdstuk 5 over steekproeven en schatten is zeer expliciet voor wat betreft bijv. vrijheidsgraden en zuiverheid. Het noemt verschillende steekproefmethoden en de correctie voor eindige populaties. Dan volgen de hypothesetoetsingen met één en twee steekproeven, waarbij het begrip onderscheidend vermogen echter een beetje onder tafel verdwijnt. Aan type II fouten worden slechts 4 zinnen besteed. De eerste non-parametrische toetsen verschijnen in deze hoofdstukken, en een vermelding van de F-toets op gelijkheid van varianties in twee steekproeven leidt in op hoofdstuk 8. Dat geeft in 10 p. een introductie tot de variantieanalyse. Hoofdstuk 9 bespreekt nominale variabelen en behandelt de χ^2 -toets, Fisher's exact, Phi-kwadraat, Tschuprow's T, de contingentiecoëfficiënt en toetsen op gelijkheid van verdelingen. Als besluit van de inferentiële statistiek geeft hoofdstuk 10, na een behandeling van regressie en correlatie (in die volgorde!) een voorbeeld van een toepassing daarvan dat zeer direct de methodologische en theoretische problemen van het gebruik van modellen illustreert. Ook niet-lineaire modellen worden hier even aangedragen, maar een expliciete vermelding van problemen als curvilineariteit en heteroscedasticiteit ontbreekt. Een laatste hoofdstuk van 18 p., dat bronnen voor statistische gegevens behandelt, is teveel op Engeland georiënteerd en teveel gericht op sociaal economische en statistisch-jaarboek-achtige gegevens. Standaard meetinstrumenten worden overgeslagen en bovendien wordt in dit hoofdstuk alleen op een wat wrange, onuitgesproken wijze duidelijk hoe moeilijk het soms kan zijn recente gegevens te verkrijgen. Als dit hoofdstuk buiten beschouwing blijft lijkt het boek zonder meer aan te bevelen voor gebruik in het HBO. En wie voor universitaire studenten een voorkeur heeft voor gecomprimeerde en efficiënte teksten kan overwegen dit boek eens in te kijken en te gebruiken als leertekst, bijv. voor studenten die buiten een regulier programma vallen. De term multivariaat verschijnt in dit boek niet. Daarmee blijft dit werk een heldere en makkelijk te volgen introductie tot het bedrijven van statistiek. Meestal is het zeer volledig, maar op andere punten komt het niet helemaal tegemoet aan de eisen die de hedendaagse praktijk aan de basiskennis van wetenschappelijke onderzoekers stelt. Het is echter als inleiding tot de statistiek zeer geslaagd.

J.M. Bothenius Lohman, RUG.