

Wat zijn de sterke punten van dit boek in vergelijking met andere leerboeken over multiniveau-analyse (Bryk en Raudenbush 1992; Hox 1994; Goldstein 1995)? Voor wie de rekenalgoritmen zelf wil programmeren of er inzicht in wil hebben, is dit boek een must (met de kanttekening dat het zich vooral, maar niet uitsluitend, richt op Fisher scoring). Het is ook aan te raden voor wie de wiskundig statistische achtergrond van de multiniveau-analyse wil kennen; wat dat betreft is het uitgebreider, en didactisch beter dan de concurrenten. Allerlei formules, bijvoorbeeld van standaardfouten van geschatte varianties, kunnen de gebruiker van pas komen en zijn moeilijk elders te vinden. Ook het verband met de meer algemene statistische literatuur wordt in de 'Bibliographical notes' aan het eind van elk hoofdstuk goed aangegeven (al mis ik de dwarsverbanden met de econometrische literatuur). De uitvoerige, zeer diverse en soms complexe voorbeelden zijn een schatkamer waarin meer aan te treffen is dan in andere boeken, die zich concentreren op een standaardbehandeling volgens de beschikbare software. De prijs van dit alles is duidelijk: voor het ontsluiten van de schatkamer is enige wiskundige rijpheid en statistische bagage nodig, of op zijn minst het vermogen om zich niet te laten afschrikken door het bos der formules.

## Referenties

- Bryk, A.S., en S.W. Raudenbush (1992), *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage, Newbury Park etc.
- Goldstein, H. (1995), *Multilevel Statistical Methods* (2nd ed.). Edward Arnold, Londen.
- Hox, J.J. (1994), *Applied Multilevel Analysis*. TT Publikaties, Amsterdam.
- Longford, N.T. (1993), *VARCL: Software for variance Component Analysis of Data with Nested Random Effects*. Manual. Groningen, iec ProGAMMA.

Tom A.B. Snijders

A.S. Bryk en S.W. Raudenbush, **Hierarchical Linear Models: applications and data analysis methods**. Sage Publications, Newbury Park, Londen/New Delhi 1992, XVI + 265 p.

Bryk en Raudenbush kunnen met Goldstein en Longford gezien worden als de grondleggers van het statistisch multilevel-model. Niet omdat zij als eersten de theorie formuleerden, maar meer nog door hun werk aan het toepasbaar maken van de theorie en aan de verfijning ervan. Het boek van Bryk en Raudenbush kan beter dan Goldsteins *Multilevel Statistical Models of Longfords Random Coefficient Models* gezien worden als inleidend.

De basis van het boek bestaat na een kort historisch overzicht en een heldere motivering uit drie kernhoofdstukken: een algehele inleiding in de logica van hiërarchisch lineaire modellen (HLM) (hoofdstuk 2); schatten en toetsen bij HLM (hoofdstuk 3); en de adequaatheid van HLM (hoofdstuk 9). De overige hoofdstukken zijn verfijningen, uitbreidingen en illustraties.

De logica van HLM start zeer basaal met de principes van lineaire regressie (wat is een intercept, een helling en een residu) en drie pagina's verderop zitten we al midden in HLM, geholpen door een groot aantal grafieken. Vervolgens wordt uitgelegd hoe HLM gezien kan worden als een algemenere (namelijk met minder restricties) formulering van het lineaire model zoals dat gebruikt wordt in variantieanalyse, covariantieanalyse en regressieanalyse. Het bijzondere zit hem in het modelleren van effecten op meerdere niveaus (leerlingen en

klassen, wijkbewoners en buurten, herhaalde metingen en individuen) tegelijkertijd. De auteurs kiezen ervoor – om begrijpelijke didactische redenen – het model steeds uiteen te leggen in submodellen voor de diverse niveaus van de hiërarchie. Naar mijn smaak wordt niet expliciet genoeg gewezen op het type steekproef dat aanleiding geeft tot het gebruik van HLM (namelijk de meertrapssteekproef) en wordt het probleem van de locatie van het intercept bij niet-parallelle regressielijnen wel heel erg uitgemolken (in de paragrafen over centren van variabelen) zonder dat de lezer nu meteen weet wat hij met deze kennis moet. Voor de lezer is de voornaamste opgave consequent de notatie te blijven volgen.

Hoofdstuk 3 bevat de statistische theorie betreffende schatten en toetsen bij HLM. Voor de ongeoefende sociale wetenschapper bevat dit hoofdstuk de taaiste kost, namelijk de nodige matrices, inversen en wat al niet meer. Misschien schrikt dit hoofdstuk menigeen zodanig af dat ze het boek terzijde zullen leggen, maar het aardige is dat wat in dit hoofdstuk moeizaam wordt uitgelegd in de daarop volgende hoofdstukken helder wordt geïllustreerd. Maar zelfs zonder de gehele matrixnotatie te kunnen volgen zal de lezer begrijpen:

- 1) dat de twee soorten variantie hebben (steekproevenvariantie en parametervariantie);
- 2) dat het vervelend is dat het design meestal opgebalanceerd is en dat daar statistisch het nodige aan te verhapstukken valt (niet evenveel leerlingen per school bijvoorbeeld);
- 3) dat daarbij betrouwbaarheidsgewichten worden gebruikt;
- 4) dat de regressiecoëfficiënten bijbehorende standaardfouten hebben waarvan de grootte afhangt van op welk niveau het submodel gesitueerd is;
- 5) dat voor de variantiecomponenten (hoeveel variantie zit er tussen wijken?) andere toetsen nodig zijn dan voor de regressiecoëfficiënten;
- 6) dat we op moeten passen als er maar weinig scholen, buurten en dergelijke zijn.

Hoofdstuk 4 doet hoofdstukken 2 en 3 nog eens dunnetjes over door een en ander te bespreken aan de hand van een voorbeeld en de lezer die in de matrices verdwaald was geraakt, zal hier snel de draad weer kunnen oppakken.

Hoofdstuk 9 ten slotte is te ver achterin het boek geplaatst: het behandelt namelijk de assumpties van HLM, hoe je kunt nagaan of en in welke mate die geschonden worden en wat je daar dan vervolgens aan kunt doen. Bovendien worden problemen met betrekking tot misspecificatie uitvoerig behandeld. Ongelukkigerwijze wordt misspecificatie alleen opgevat in termen van vergeten variabelen en niet als: is er een belangrijk niveau in de hiërarchie over het hoofd gezien?

De hoofdstukken 5 tot en met 7 zijn vervolgens toepassingen van het twee-niveaumodel op: organisatieonderzoek (leden binnen organisaties, leerlingen binnen scholen); het meten van veranderingen (de tijdstippen waarop gemeten is, worden hier gezien als genest binnen het individu); en meta-analyse (onderzoeken in de review worden gezien als steekproefelementen uit een populatie en de primaire onderzoekseenheden worden gezien als genest binnen elk onderzoek). De lezer doet er wel verstandig aan andere boeken over het meten van veranderingen, respectievelijk meta-analyse te raadplegen alvorens met HLM in deze situaties aan de slag te gaan.

Hoofdstuk 8 bevat uitbreiding van het twee niveaumodel naar drie niveaus (leerlingen in klassen in scholen, tijdstippen binnen individuen in organisaties etcetera). De prachtige paragraaf 'A measurement model at level 1' over hoe je uit percepties van individuen variabelen op organisatieniveau kunt destilleren, had wel een belangrijkere plaats in het boek mogen krijgen.

Voor de statistische 'die-hards' is er tenslotte nog een hoofdstuk 10 met technische details.

Het boek is didactisch zeer adequaat opgezet met 'advance organizers', vele en heldere voorbeelden en samenvattingen aan het eind van elk hoofdstuk. In combinatie met het software-pakket HLM3 is het bijzonder leerzaam, daar datasets behandeld in het boek door de

gebruikers van dat pakket ook geanalyseerd kunnen worden.

Wat het boek allemaal niet doet (logistische HLM, kruisclassificaties, multivariate toepassingen, etc.) is te vinden in de boeken van Longford en/of Goldstein. Het boek behandelt echter precies wat het software-pakket HLM kan. Het niveau van het boek schat ik in op: de beste inleiding voor sociale wetenschappers die met multilevel-modellen aan de slag moeten en geschikt voor postdoctoraal onderwijs (wellicht) voorzich in M&T specialiserende postdoctoraal studenten.

Roel J. Bosker

R. Prosser, J. Rasbash, H. Goldstein, **ML3 Software for Three-level Analysis, Users' Guide for V.2**. Institute of Education, University of London, januari 1991, 142 p. (1).

R. Prosser, J. Rasbash, H. Goldstein, **Data Analysis with ML3**, Institute of Education, University of London, juli 1991, 121 p. (2); **ML3 V2.2 Manual Supplement, New and Amended commands**. Multilevel Models Project, Institute of Education, University of London, november 1991, 10 p. (3).

G. Woodhouse (red.), **A Guide to ML3 for New Users, second edition**. Multilevel Models Project, Institute of Education, University of London, juli 1993, 124 p. (4).

J. Rasbash, **ML3-E Version 2.3 Manual Supplement**. Multilevel Models Project, Institute of Education, University of London, juli 1993, 11 p. (5).

De uitdaging om samen met Rob Eisinga een bijdrage te leveren aan dit themanummer van de Sociologische Gids over multiniveau-analyse noodzaakte me op korte termijn enig idee te krijgen van de theorie van de multiniveau-analyse, een programma te kiezen en er snel in thuis te raken. Het moest een logistische analyse worden op een groot databestand ( $n = 244.050$ ). Na het lezen van Bryk en Raudenbush's *Hierarchical Linear Models* (1992) had ik de beschikking over de manuals van HLM, VARCL en ML3 (hierboven 1) en van het laatste programma tevens de hierboven genoemde *Guide to ML3 for New Users* (4). Hoewel gewaarschuwd dat de 'ease of learning' van de drie programma's respectievelijk 'easy, easy and slow' is (*Newsletter 4* (2)), viel de keuze toch op ML3. Met HLM zijn blijkens dezelfde *Newsletter* geen binaire data te analyseren en in de *Guide to ML3*, naar de titel doet vermoeden begaan met problemen van 'new users', was in ieder geval een voorbeeld opgenomen van een logistische analyse.

Toen het artikel er bijna lag, werd ik door de redactie verzocht mijn ervaringen met ML3 en de programmadocumentatie op schrift te stellen. Na enige aarzeling te hebben toegezegd, stond ik voor het probleem: wat hoort precies tot de programmadocumentatie? Want het multiniveau denken, zo had ik al snel door, strekt zich uit over vele gebieden (logistische analyse, multivariate respons, primaire netwerken, herhaalde metingen, crosslevel-analyse, covariantiestructuuranalyse, etc.) en ik wist niet of ik over de relevante documentatie beschikte. Het verzoek aan het Institute of Education te Londen om de programmadocumentatie van ML3 toe te zenden, werd beantwoord met toezending van de vijf bovengenoemde documenten. Ik heb de indruk dat voor bepaalde toepassingen er meer nuttige documenten zijn. Zo zou ik zelf voor de logistische analyse met binaire respons hiermee hebben kunnen volstaan, maar in een later stadium kreeg ik *A Guide to ML3 Macros, Multilevel Binary Response Logistic Models* (juni 1993) onder ogen dat wel gedeeltelijk, maar niet helemaal overlappende informatie bevat. Men moet er ook rekening mee houden dat sommige toepassingen in deze documenten niet aan bod komen. Als wij bijvoorbeeld een polytome responsvariabele hadden gehad (op de programmadiskette zijn de macro's daarvoor voorhanden), dan hadden we niet genoeg gehad aan deze documenten. In het algemeen is het voor de gebruiker lastig er-