

## Arnoud den Boer

Afdeling Stochastische Operationele Research  
Universiteit Twente  
a.v.denboer@utwente.nl

### Onderzoek Veni-project

# Menselijke keuzes begrijpen en beïnvloeden

Aan Arnoud den Boer, tenure-tracker aan de Universiteit Twente, werd in 2014 een NWO Veni-beurs toegekend. In dit artikel beschrijft hij in het kort zijn onderzoek naar statistiek en optimalisatie van menselijke keuzes.

Discrete-keuzemodellen [1, 13] zijn wiskundige modellen die beschrijven hoe mensen kiezen uit een verzameling alternatieven met gegeven eigenschappen. Deze modellen worden in tal van toepassingsgebieden gebruikt om menselijk keuzegedrag te voorspellen en te bepalen hoe daar optimaal op gereageerd kan worden: marketing [7], landbouw [8], psychometrie [9], transportwetenschap [2], enzovoort. Vanwege de vele praktische toepassingen wordt er ook veel theoretisch onderzoek verricht naar schatten en optimaliseren met deze modellen. In 2000 ging zelfs de Nobelprijs voor Economie naar Daniel McFadden “for his development of theory and methods for analyzing discrete choice” (nobelprize.org).

Een veelgebruikt framework om menselijke keuzes te modelleren is het zogenaamde Random Utility Model (RUM) [10, 12]. Stel dat een persoon precies één item moet kiezen uit  $n \in \mathbb{N}$  alternatieven, gelabeld  $1, \dots, n$ . De eigenschappen van elk alternatief  $i$  worden gemodelleerd door een vector  $x_i \in \mathbb{R}^d$ , voor een integer  $d$ . Volgens het RUM is de kans dat

alternatief  $i$  wordt gekozen dan gelijk aan

$$\mathbb{P}(U_i > U_j, \forall j \neq i), \quad (1)$$

waarbij  $U_1, \dots, U_n$  stochastische variabelen

zijn die het nut (‘utility’) modelleren dat de persoon aan alternatieven  $1, \dots, n$  toekent. De stochasten  $U_1, \dots, U_n$  mogen onderling afhankelijk zijn, en hun verdeling hangt (uiteraard) ook af van de eigenschappen  $x_1, \dots, x_n$ . Men neemt ook aan dat hun verdelingen continu zijn, zodat er in (1) met kans één geen ‘tie’ voorkomt.

Het algemene RUM zoals hier geformuleerd blijkt in de praktijk nog te algemeen om



effectief statistiek en optimalisatie mee te bedrijven. Men neemt daarom vaak de toevlucht tot eenvoudigere modellen. Het meest populaire model is het zogenaamde *multinomial logit*-model (MNL-model) met lineaire nutsfuncties, dat veronderstelt dat

$$U_i = \beta^\top x_i + \epsilon_i, \quad i = 1, \dots, n. \quad (2)$$

Hier is  $\beta \in \mathbb{R}^d$  een (onbekende) parameter, en zijn  $\epsilon_1, \dots, \epsilon_n$  onderling onafhankelijke stochasten met Gumbel verdeling met gemiddelde nul. Door deze specifieke verdeling van  $\epsilon_i$  aan te nemen krijgen de kansen (1) een elegante gesloten vorm:

$$\begin{aligned} \mathbb{P}(U_i > U_j, \forall j \neq i) &= \frac{e^{\mathbb{E}[U_i]}}{\sum_{j=1}^n e^{\mathbb{E}[U_j]}} \\ &= \frac{e^{\beta^\top x_i}}{\sum_{j=1}^n e^{\beta^\top x_j}}. \end{aligned} \quad (3)$$

De parameter  $\beta$  is eenvoudig te schatten uit data met behulp van ‘maximum-likelihood estimation’ (MLE), wat de populariteit van MNL deels verklaart.

Het MNL-model negeert mogelijke correlaties tussen de verschillende  $U_i$ . Dat dit onrealistisch kan zijn wordt geïllustreerd door de befaamde ‘red bus – blue bus paradox’. Stel dat er drie manieren zijn om van A naar B te reizen: (i) met de auto, (ii) met een blauwe bus, en (iii) met een rode bus. Zij  $U_1, U_2, U_3$  de bijbehorende nutsvariabelen. De kleur van een bus is irrelevant voor de voorkeuren van reizigers, en dus is  $U_2 = U_3$  met kans één. Als nu de rode bus ineens niet meer beschikbaar is verwacht je dat er niet meer of minder mensen met de auto gaan: ieder die eerst de rode bus nam kiest nu voor de blauwe bus. In het MNL-model verandert de kans dat iemand voor de auto kiest echter wél, namelijk van  $\exp(\mathbb{E}[U_1]) / (\exp(\mathbb{E}[U_1]) + 2 \exp(\mathbb{E}[U_2]))$  naar  $\exp(\mathbb{E}[U_1]) / (\exp(\mathbb{E}[U_1]) + \exp(\mathbb{E}[U_2]))$ . Dit fenomeen maakt MNL in sommige situaties een onrealistisch model.

Het ‘mixed multinomial logit model’ (MMNL-model) is een uitbreiding van MNL die veronderstelt dat  $\beta$  in (3) zelf ook een stochastische variabele is. Deze multidimensionale stochast  $\beta$  modelleert mogelijke heterogeniteit van de populatie. Als  $F$  de cumulatieve verdelingsfunctie is van  $\beta$ , dan is de kans dat een persoon kiest voor optie  $i$  gelijk aan

$$\begin{aligned} \mathbb{P}(U_i > U_j, \forall j \neq i) &= \int_{\mathbb{R}^d} \frac{e^{\beta^\top x_i}}{\sum_{j=1}^n e^{\beta^\top x_j}} dF(\beta). \end{aligned} \quad (4)$$

Een van de resultaten van reeds genoemde

Nobelprijswinnaar McFadden is dat hij, samen met Kenneth Train, heeft laten zien dat ieder RUM arbitrair goed benaderd kan worden door een MMNL-model [11]. De klasse van MMNL-modellen is daarmee veel rijker dan de MNL-modellen, en doordat het correlaties tussen  $U_1, \dots, U_n$  modelleert heeft het niet het hierboven beschreven nadeel van MNL. Het schatten van de verdeling van  $\beta$  kan numeriek erg uitdagend zijn, evenals optimalisatieproblemen die gebruikmaken van MMNL-modellen.

### Multi-armed bandit-problemen

*Multi-armed bandit*-problemen (MAB-problemen) zijn dynamische beslissingsproblemen waarbij de te optimaliseren doelfunctie onbekend is, maar gaandeweg geleerd wordt uit binnenkomende data. Het cruciale verschil met statische problemen is dat de beschikbare dataset groeit, doordat genomen beslissingen nieuwe data genereren. Deze nieuwe data kan helpen om de doelfunctie beter te leren en in de toekomst betere beslissingen te nemen. Er ontstaat zo een wisselwerking tussen statistiek en optimalisatie: je zoekt naar beslissingen die niet alleen de verwachte kortetermijnopbrengst maximaliseren, maar ook bijdragen aan het zo goed en snel mogelijk leren van de doelfunctie. Deze twee doelen zijn meestal conflicterend: kortetermijnopbrengst maximaliseren zorgt er vaak voor dat beslissingen snel convergeren naar de beslissing die het beste lijkt op grond van de data, terwijl het leren van de doelfunctie juist sneller gaat als er *spreiding* in de beslissingen is. Het vinden van beslissingsregels die de juiste balans vinden in deze zogenaamde *exploration–exploitation trade-off* is een belangrijk terugkerend vraagstuk in MAB-problemen.

MAB-problemen zijn momenteel een *hot topic* in de operations research, en worden ook veelvuldig bestudeerd in de statistiek, econometrie en computer science. Dat komt door de vele maatschappelijke en commerciële toepassingen van MAB, maar ook door de wiskundige uitdagingen die het biedt (MABs zijn in de regel niet exact oplosbaar).

### Veni-onderzoek: MABs met keuzemodellen

In mijn Veni-onderzoek verbind ik discrete-keuze modellen met multi-armed bandit-problemen.

Dit is gemotiveerd door dynamische beslissingsproblemen uit de praktijk die met menselijk keuzegedrag te maken hebben. Twee voorbeelden uit de (online) retail zijn assortimentsoptimalisatie (welke producten moet

een retailer in zijn schap zetten, of op zijn website tonen, om de verwachte opbrengst te maximaliseren) en prijssoptimalisatie (bij welke verkoopprijzen is de verwachte opbrengst het grootst). Dit zijn allebei MAB-problemen waarbij menselijk keuzegedrag een belangrijke rol speelt. Bij assortimentsoptimalisatie gaat het om het bepalen van de *alternatieven* waaruit personen kunnen kiezen, bij prijssoptimalisatie om het bepalen van *eigenschappen* van de alternatieven. Ik bekijk problemen waarbij zowel de alternatieven als hun eigenschappen geoptimaliseerd kunnen worden.

Formeel beschouw ik het volgende type MAB-probleem: aan  $N \in \mathbb{N}$  achtereenvolgende personen  $n = 1, \dots, N$  wordt een verzameling alternatieven  $S(n) \in \mathcal{S}$  met eigenschappen  $\mathbf{X}(n) \in \mathcal{X}$  ter keuze aangeboden. Hier is  $\mathcal{S}$  een collectie niet-lege deelverzamelingen van alle  $J \in \mathbb{N}$  alternatieven  $\{1, \dots, J\}$  die aangeboden kunnen worden;  $\mathbf{X}(n)$  is een  $d \times J$  matrix waarvan de  $j$ -de kolom  $\mathbf{X}_j(n)$  eigenschappen van alternatief  $j$  modelleert, en is  $\mathcal{X} \subset \mathbb{R}^{d \times J}$  een collectie van toegestane eigenschapmatrices.

Het keuzegedrag van personen wordt gemodelleerd met een MMNL-model: de kans dat een persoon kiest voor alternatief  $j$  uit beschikbare alternatieven  $S$  met eigenschappen  $\mathbf{X}$  is gelijk aan

$$\int_{\mathbb{R}^d} \frac{e^{\mathbf{x}_j^\top \beta}}{\sum_{k \in S} e^{\mathbf{x}_k^\top \beta}} dF(\beta), \quad j \in S. \quad (5)$$

Hier is  $F$  de (onbekende) cumulatieve verdelingsfunctie van  $\beta$ . De keuze voor  $j$  geeft vervolgens een opbrengst  $r(j, S, \mathbf{X})$ , waar  $r : \{1, \dots, J\} \times \mathcal{S} \times \mathcal{X} \rightarrow \mathbb{R}$  een bekende continue functie is.

Zij  $j(n)$  de keuze van persoon  $n$ , en zij  $\mathbf{F}$  een verzameling van verdelingsfuncties die de onbekende  $F$  bevat. Het doel nu is om een reeks  $(S(n), \mathbf{X}(n))$  ( $n = 1, \dots, N$ ) alternatieven en eigenschappen te vinden die de worst-case cumulatieve verwachte opbrengst

$$\min_{F \in \mathbf{F}} \mathbb{E} \left[ \sum_{n=1}^N r(j(n), S(n), \mathbf{X}(n)) \right] \quad (6)$$

maximaliseert. Elk paar  $S(n), \mathbf{X}(n)$  mag afhangen van observaties in het verleden (dat wil zeggen van  $j(m), S(m), \mathbf{X}(m)$  voor  $m < n$ ), maar niet van toekomstige observaties.

Dit is een technisch uitdagend probleem waarvoor in het algemeen geen exacte oplossing

sing bestaat. Om er toch grip op te krijgen onderzoek ik de volgende twee deelproblemen:

1. *Karakteriseer de kwaliteit van schatters van  $F$  in eindige samples.* Goede beslisregels in MAB-problemen hebben een goede balans tussen *exploration* en *exploitation*: zij nemen niet altijd de beslissing die het beste lijkt op grond van een schatting  $\hat{F}_n$  van  $F$ , maar wijken daar soms van af om  $F$  sneller te leren. Om te bepalen hoeveel of hoe vaak je dit moet doen is een goed begrip nodig van de kwaliteit van  $\hat{F}_n$ ; in het bijzonder begrip hoe de verwachte schattingsfout  $\mathbb{E}[|\hat{F}_n - F|]$ , voor een zorgvuldig gekozen norm, afhangt van de eerdere beslissingen  $\{(S(m), \mathbf{X}(m)) \mid m < n\}$ . Voor MLE in het MNL-model heb ik zulke resultaten afgeleid [5] en toegepast op dynamisch prijzen [3–4, 6]; voor het MMNL-model zijn zulke resultaten niet bekend.

2. *Ontwerp asymptotisch optimale beslisregels.* Omdat het niet mogelijk is een beslisregel te vinden die (6) maximaliseert, zoek ik een asymptotisch optimale benadering. In het bijzonder zoek ik een serie functies  $\pi_n : (S \times \mathcal{X} \times \{1, \dots, J\})^{n-1} \rightarrow (S \times \mathcal{X})$ ,  $n \in \mathbb{N}$ , zodat, als  $(S(n), \mathbf{X}(n)) = \pi(\{(S(m), \mathbf{X}(m), j(m)) \mid m < n\})$  voor alle  $n$ , dan de *groeisnelheid* van (6) in  $n$  zo laag mogelijk is. Om te bepalen of de gevonden beslisregel asymptotisch optimaal is, ga ik ook de laagst mogelijke groeisnelheid van (6) die een willekeurige beslisregel  $(\pi_n)_{n \in \mathbb{N}}$  kan behalen, karakteriseren.

Het oplossen van deze twee uitdagingen zal de twee vakgebieden van discrete-keuzetheorie en multi-armed bandit-problemen met elkaar verbinden, en hopelijk tot veel nieuw boeiend en relevant onderzoek leiden. Daarnaast zal het voor heel concrete problemen uit de praktijk, zoals het assortimentsoptimalisatieprobleem, een belangrijke bijdrage leveren.

## Biografie

Arnaud den Boer (1982) studeerde wiskunde aan de Universiteit Utrecht en deed vervolgens de post-master Mathematics for Industry aan het Stan Ackermans Instituut van de Technische Universiteit Eindhoven. Bij het Centrum Wiskunde & Informatica schreef hij zijn proefschrift *Dynamic Pricing and Learning*, in 2015 bekroond met de Gijs de Leve-prijs. Na postdoc-posities aan de Technische Universiteit Eindhoven en de Universiteit van Amsterdam is hij sinds november 2013 als tenure-tracker verbonden aan de Universiteit Twente, vakgroep Stochastic Operations Research. Zijn onderzoeksinteresses zijn multi-armed bandit-problemen, discrete-keuze modellen, statistische inferentie in wachttijdmodellen, en de interactie tussen modellering, statistiek en optimalisatie in stochastische beslissingsproblemen.

## Referenties

- 1 M. Ben-Akiva en S. Lerman, *Discrete Choice Analysis. Theory and Application to Travel Demand*, The MIT Press, Cambridge, MA, 1985.
- 2 E. Ben-Elia en Y. Shiftan, Which road do I take? A learning-based model of route-choice behavior with real-time information, *Transportation Research Part A: Policy and Practice* 44(4) (2010), 249–264.
- 3 A.V. den Boer, Dynamic pricing with multiple products and partially specified demand distribution, *Mathematics of Operations Research* 39(3) (2014), 863–888.
- 4 A.V. den Boer en B. Zwart, Simultaneously learning and optimizing using controlled variance pricing, *Management Science* 60(3) (2014), 770–783.
- 5 A.V. den Boer en B. Zwart, Mean square convergence rates for maximum quasi-likelihood estimators, *Stochastic Systems* 4(2) (2014), 375–403.
- 6 A.V. den Boer en B. Zwart, Dynamic pricing and learning with finite inventories, te verschijnen in *Operations Research* (2015).
- 7 P.K. Chintagunta en J.P. Dubé, Estimating a stockkeeping-unit-level brand choice model that combines household panel data and store data, *Journal of Marketing Research* 42(3) (2005), 368–379.
- 8 S. Colombo, N. Hanley en J. Louviere, Modeling preference heterogeneity in stated choice data: an analysis for public goods generated by agriculture, *Agricultural Economics* 40(3) (2009), 307–322.
- 9 T.R. Johnson, Discrete choice models for ordinal response variables: a generalization of the stereotype model, *Psychometrika* 72(4) (2007), 489–504.
- 10 D. McFadden, Economic choices, *The American Economic Review* 91(3) (2001), 351–378.
- 11 D. McFadden en K. Train, Mixed MNL models for discrete response, *Journal of Applied Econometrics* 15(1) (2000), 447–470.
- 12 P. Suppes, D.H. Krantz, R.D. Luce en A. Tversky, *Foundations of Measurement*, Volume II, Academic Press, San Diego, CA, 1989.
- 13 K. Train, *Discrete Choice Methods with Simulation*, Cambridge University Press, second edition, 2009.