

eerste concept voor artikel in Tijdschrift voor Onderwijsresearch

Ben Wilbrink

Cesuurbepaling bij dichotome criterium variabele.

Er doen zich in het onderwijs nogal eens situaties voor waarin een groep personen in tweeën gedeeld moet worden: selectie, het leggen van de grens voldoende-onvoldoende bij tentamens, verdelen over verschillende onderwijsvormen of studierichtingen, e.d. Criterium variabelen zijn vaak te dichotomiseren: voorbeelden zijn gebleken geschiktheid, beheersing van de stof tot een bepaald criterium, slagen voor een volgend studieonderdeel. Het cesuurprobleem is o.a. door Rorer et al (1966 a en b), Alf & Dorfman (1967), en Vastenhouw (1973) behandeld, zij het op zo op het eerste gezicht nogal uiteenlopende wijze. Het probleem is ook heel goed aan te pakken met behulp van eenvoudige besliskundige techniek (Lindley, 1971). De bedoeling is te demonstreren hoe via laatstgenoemde aanpak het werk van de eerder genoemde auteurs op eenvoudige wijze met elkaar in verband gebracht kan worden. Een en ander kan bijzonder verhelderend werken waar het gaat om een betere begripsvorming van de problematiek inherent aan cesuurbepaling. In het bijzonder is de implicatie van deze wijze van aanpakken dat cesuurbepaling zoals voorgesteld door Van Vliet (1975) geen optimale, zelfs geen behoorlijke, oplossing is.

De algemene situatie is dat op grond van een bepaald gegeven of een test of toets een voorspelling gedaan wordt m.b.t. een 'criterium, variabele: bijv. geschiktheid, beheersing van de stof, etc. De aanname is dat de criterium variabele een tweedeling is, of zinvol in tweeën gedeeld kan worden (zie bijv. Cronbach & Gleser 1965 blz. 49, 50); we nemen ook aan dat deze dichotome criterium variabele perfect betrouwbaar gemeten wordt, en dat ze een valide maat is voor hetgeen we bedoelen te 'voorspellen'. (Rekening houden met onbetrouwbaarheid en tekortkomingen in de validiteit van de criterium variabele zou een en ander aan complicaties meebrengen, maar niets aan de strekking van de te geven argumentatie veranderen).

Voor onbetrouwbaarheid van de variabele die als 'voorspeller' gebruikt wordt, zal niet gecorrigeerd worden, omdat in praktische toepassingen een dergelijke 'correctie voor attenuatie' geen zin heeft.

Een belangrijk gegeven dat nodig is, een eis waar in de praktijk ook nogal mee geschipperd zal moeten worden, is dat het verband tussen voorspeller en criterium bekend is. In de meeste gevallen zal de voorspeller niet op een continue, maar op een discrete schaal gescoord worden (cijfers, aantal items correct, lengte afgerond op hele centimeters, e.d.), in welk geval voor iedere mogelijke score X_i op de voorspeller bekend moet zijn hoe groot de proportie p_i is die op het criterium 'geschikt' scoort. Als voorbeeld nemen we dus voorspelling van 'geschiktheid'.

Tenslotte nemen we nog aan dat de proportie p_i hoger wordt bij hoger wordende score x op de voorspeller. Deze aanname is slechts bedoeld om complicaties, zoals besproken door Rorer et al. 1966 a, te vermijden. Dus: hoe hoger score op de voorspeller, des te groter de proportie 'geschikten'.

Wanneer personen die op de voorspeller beneden de cesuur, zeg x_c scoren afgewezen worden, en de overigen toegelaten, dan zijn er vier verschillende mogelijke uitkomsten: een geschikte kandidaat kan toegelaten of afgewezen zijn, en idem voor een ongeschikte kandidaat. (In feite wordt hiermee ook de voorspeller in tweeën gedeeld, wat niet altijd even zinvol hoeft te zijn). Aan iedere mogelijke uitkomst van een beslissing over een enkele persoon moeten we een utiliteit toekennen, als laatste vereiste stap om de cesuur die we zoeken te kunnen berekenen. De beste uitkomst is, stel, het toegelaten hebben van een geschikte kandidaat, zodat deze uitkomst de utiliteit 1 toegekend wordt. De slechtste uitkomst is, stel, het toegelaten hebben van een geschikte kandidaat: utiliteit 0. Stel dat het afgewezen hebben van een geschikte kandidaat een gebeurtenis is die we zouden hebben willen inruilen voor een lot dat ons een kans van $1/3$ geeft op een geschikte kandidaat die aangenomen kan worden, tegen een kans van $2/3$ op een ongeschikte kandidaat die we dan gedwongen zijn aan te nemen (zie voor een uiteenzetting over deze werkwijze om utiliteiten toe te kennen o.a. Raiffa 1968, of Lindley 1971). De utiliteit van het afgewezen hebben van een geschikte kandidaat is dan $1/3$. En stel dat het afgewezen hebben van een ongeschikte kandidaat een gebeurtenis is die we zouden hebben willen inruilen tegen een lot dat ons een kans van $9/10$ geeft op de beste uitkomst: een geschikte kandidaat die we kunnen toelaten, tegen het alternatief (kans $1/10$) genoodzaakt te zijn een ongeschikte kandidaat toe te laten (de slechtste uitkomst). De utiliteit van het afgewezen hebben van een geschikte kandidaat is dan $9/10$.

Een en ander wordt samengevat in een beslissingstabel, tabel 1.

	y1: ongeschikt	Y2: geschikt
X1: afwijzen	u11 = 9/10	u12 = 1/3
X2: toelaten	u21 = 0	u22 = 1

TABEL 1. Beslissingstabel.

Nu kan de cesuur rechttoe-rechtaan bepaald worden, tenzij men zich ondertussen zou bedenken ten aanzien van de toegekende utiliteiten (dan moeten die eerst met de nieuwe inzichten in overeenstemming gebracht worden). Neem een willekeurige score x_i en bereken dan de totale utiliteit die verkregen wordt wanneer iedereen met deze score toegelaten zou worden. Doe hetzelfde voor het geval iedereen met score x_i afgewezen zou worden. Als de totale utiliteit voor iedereen met score x_i toelaten groter is dan voor iedereen met score x_i afwijzen, dan laten we iedereen toe. De gezochte cesuur ligt in dit geval bij een lagere x score. Doe in gedachten dezelfde berekeningen voor telkens de dichtst bij zijnde score x die kleiner is, totdat we een x_i vinden waarvoor geldt dat de totale utiliteit voor afwijzen van kandidaten met deze score groter is dan de totale utiliteit voor toelaten van alle kandidaten met deze score x_i : dan is dit de gezochte cesuur x_e . Hetzelfde anders gezegd:

Stel in de groep kandidaten met score x_i is de proportie geschikten p_i . Wordt deze groep toegelaten, dan is de utiliteit te berekenen als

$$(1) \quad p_i \cdot u_{22} + (1-p_i) \cdot u_{21} \quad (\text{utiliteit toelaten})$$

Wordt deze groep afgewezen, dan is de utiliteit

$$p_i \cdot u_{12} + (1-p_i) \cdot u_{11} \quad (\text{utiliteit afwijzen})$$

De cesuur leggen we bij die score x_i waar de utiliteit toelaten gelijk is aan de utiliteit afwijzen, dus waar

$$p_i \cdot u_{22} + (1-p_i) \cdot u_{21} = p_i \cdot u_{12} + (1-p_i) \cdot u_{11}$$

Als in (1) de utiliteiten uit tabel 1 ingevuld worden, dan kan de

waarde van p_i uitgerekend worden:

$$p_i \cdot 1 + (1 - p_i) \cdot 0 = p_i \cdot (1/3) + (1 - p_i) \cdot (9/10)$$

$$p_i - 1/3 p_i + 9/10 p_i = 9/10$$

$$p_i = 27/47$$

Omdat we voor iedere score x_i de bijbehorende proportie p_i kennen, kunnen we nu zonder meer de bij $p_i = 27/47$ behorende score x_i vinden, welke score de gezochte cesuur is. (meestal is er geen score x_i met exact de berekende waarde van p_i , dan nemen we als cesuur de score x_j waarvoor geldt $p_j < p_i < p_{j+1}$).

In vergelijking (1) is de enige onbekende p (we laten de index i voortaan weg). De oplossing voor de cesuur hangt, gegeven de utiliteiten en de overige data die bekend worden verondersteld, alleen af van p , we zullen p dan ook onze selectieparameter noemen. Uit (1) is p uit te drukken in de utiliteiten voor de verschillende mogelijke uitkomsten:

$$(2) \quad p \cdot u_{22} + (1 - p) \cdot u_{21} = p \cdot u_{12} + (1 - p) \cdot u_{11}$$

$$p(u_{22} - u_{21}) + u_{21} = p(u_{12} - u_{11}) + u_{11}$$

$$p(u_{22} - u_{21} - u_{12} + u_{11}) = u_{11} - u_{21}, \text{ waaruit tenslotte:}$$

$$(3) \quad p = \frac{u_{11} - u_{21}}{(u_{22} - u_{21}) - (u_{12} - u_{11})}.$$

Deze selectieparameter p is gelijk aan de selectie parameter P door Alf & Dorfman 1967 onder onnodig restrictieve aannamen afgeleid, en is ook gelijk aan de selectieparameter s_p van Vastenhouw 1973. Rorer et al. 1966 a & b leiden een selectieparameter V af:

$$(4) \quad v = (u_{21} - u_{11}) / (u_{12} - u_{11})$$

Uit (3) en (4) is eenvoudig af te leiden dat

$$(5) \quad p = V / (V + 1).$$

$$(6) \quad V = p / (1 - p).$$

De uitbreiding die in het voorgaande gegeven is aan het werk van auteurs als Rorer et al 1966, Alf & Dorfman 1967 en Vastenhouw 1967 is niet alleen dat hun werk met elkaar in verband is gebracht, maar vooral dat gebruik gemaakt wordt van een duidelijk omschreven (Lindley 1971)

procedure voor het toekennen van utiliteiten. Utiliteiten komen nu niet meer zomaar uit de lucht vallen.

In plaats van utiliteiten wordt vaak ook met verliezen gewerkt. Het eerder gegeven voorbeeld zullen we verder uitwerken door verliezen te specificeren. Deze werkwijze is alleen een beetje omslachtiger dan het direct werken met utiliteiten, omdat verliezen gedefinieerd zijn aan utiliteiten. Lindley 1971 volgend definiëren we:

l_{ij} = het verlies van beslissing x_i bij gebeurtenis y_j

$$(7) \quad l_{ij} = (\max_i u_{ij}) - u_{ij} \quad [\text{max over alle } i \text{ te nemen}]$$

Uit tabel 1 kunnen we dan m.b.v. (7) de verliestabel 2 afleiden:

	y1: ongeschikt	Y2: geschikt
X1: afwijzen	$l_{11} = 0$	$l_{12} = 2/3$
X2: toelaten	$l_{21} = 9/10$	$l_{22} = 0$

TABEL 2. Verliestabel.

De cesuur ligt dan weer bij die score x waarvoor geldt dat het verlies dat volgt uit het toelaten van iedereen met die score x gelijk is aan het verlies dat volgt uit het afwijzen van iedereen met die score x . Immers. uit onze aanname dat de proportie p groter wordt bij hogere score x volgt dat voor kleinere waarden van x het verlies bij toelaten groter is dan bij afwijzen, en bij grotere waarden van x het verlies bij toelaten kleiner is dan bij afwijzen.

We lossen weer p op uit

$$(8) \quad p \cdot l_{22} + (1-p) \cdot l_{21} = p \cdot l_{12} + (1-p) \cdot l_{11}$$

Waaruit we vinden, na invullen van de verliezen, dat

$p = 27/47$, dezelfde oplossing die eerder verkregen werd.

Het hoeft geen verdere toelichting dat in het algemeen de cesuur bij een andere utiliteiten keuze anders zal komen te liggen. Wel is het hier goed om er nog eens duidelijk op te wijzen dat de keuze van utiliteiten geheel anders kan uitvallen, al naar gelang de bepaalde belanghebbende die de utiliteiten vaststelt. Bij een bepaalde studietoets zou de keuze van utiliteiten waarschijnlijk anders uitvallen wanneer ze gedaan wordt door de betrokken docent, door de betrokken studenten, of misschien door een commissie van beroep. Bij toelating tot het wetenschappelijk onderwijs zal de toekenning van utiliteiten aanmerkelijk verschillen tussen de kandidaten, belanghebbenden binnen de betrokken faculteit, en beleidsorganen die uitgebreider belangen moeten behartigen. Om een en ander duidelijker te laten overkomen, zal een concreet voorbeeld gegeven worden. Vanzelfsprekend wordt een iets gesimplificeerde situatie als start genomen. Stel gemiddeld eindexamencijfer wordt als 'voorspeller' gebruikt, en al dan niet afstuderen in de gekozen studierichting is het criterium voor 'geschiktheid'. (het criterium is weinig betrouwbaar, en over validiteiten kunnen slechts gissingen gedaan worden; tot de simplificatie hoort dan ook de veronderstelling dat betrouwbaarheid en validiteiten geen problemen opleveren). Uit vroegere (of onder experimentele condities verkregen) waarnemingen weten we voor ieder gemiddeld eindexamencijfer (5,5, 6, 6,5, 7, 10) de proportie kandidaten p die in de gekozen richting ook afstudeerde.

Stel student A is in de gelegenheid zelf de utiliteiten te kiezen die in de toelatingsbeslissing een rol zullen spelen. Stel dat hij, als hij werkelijk geschikt is (wat hij op dit moment nog niet weet), en toegelaten zou worden, dit als de beste uitkomst kiest: $u_{22} = 1$, stel dat hij verwacht dat hij een aantal jaren onnodig zou verspelen die hij in een andere studie beter gebruikt zou hebben, wanneer hij ongeschikt zou zijn en toch toegelaten was: $u_{21} = 0$. Bij afwijzing kiest hij een andere studie, en of hij geschikt dan wel ongeschikt zou zijn voor zijn eerste studiekeuze is hem onverschillig: $u_{11} = u_{12}$. Niettemin zou hij een aanbod om alsnog te kunnen studeren in de richting van zijn eerste keuze accepteren als zijn kans op succes 7/10 of hoger zou zijn (waarover zijn studieadviseur hem misschien kan inlichten): $u_{11} = u_{12} = 7/10$.

Student B krijgt ook de gelegenheid om voor zijn eigen toelatingsbeslissing de utiliteiten te kiezen. Ook zij kiest, wat voor de hand ligt nietwaar, $u_{22}=1$. Zij vindt echter een onterechte afwijzing erger dan een onterechte toelating (en een terechte afwijzing is erg jammer, maar als het zo eenmaal ligt, tja): $u_{12}=0$. Hoewel die onterechte toelating niet zoveel beter is dan een onterechte afwijzing, zij zou dit

resultaat dan ook wel willen inruilen tegen een kans $1/10$ op de beste, tegen $9/10$ op de slechtste uitkomst (gesteld even in gedachten dat er een instantie is die dit soort kansen aan zou kunnen bieden, dan zou zij er gebruik van maken als de kansen zo, eventueel beter, lagen): $u_{21} = 1/10$. De laatste vraag is dan tegen welke kans op de beste uitkomst (en 1 min die kans op de slechtste uitkomst) zij een terechte afwijzing nog zou willen inruilen: De studie van eerste keuze is erg aantrekkelijk, dus ook met een kleine kans op succes zou zij al genoeg nemen, maar lager dan $2/10$ wil ze toch niet gaan: $u_{11} = 1/5$. Er volgt nog een klein gewetensonderzoekje of een terechte afwijzing wel twee keer zo aantrekkelijk is als een ten onrechte toegelaten zijn, en na ampele overweging, en na zich nog eens goed gerealiseerd te hebben dat het er niet om gaat of ze liever geschikt dan ongeschikt is maar gegeven geschiktheid of ongeschiktheid als vertrekpunt moet nemen, besluit ze het bij de eerstgemaakte keuzen te houden.

Student A en B zijn het met elkaar dus in het geheel niet eens over de wijze waarop je 'geschiktheid' voor de studie van eerste keuze moet inschatten wanneer je nooit zult weten of je inderdaad geschikt was of niet (omdat je niet tot de studie toegelaten werd).

Faculteit C, die probeert de toelatingsregeling zoveel mogelijk in haar voordeel te beïnvloeden, voert uitvoerig intern beraad over de keuze van utiliteiten. Zij besluiten dat het toegelaten hebben van een geschikte student weliswaar erg goed is, maar dat het afwijzen van een ongeschikte student nog beter is, gezien de kosten en frustraties die de 'cooling out' van ongeschikte studenten met zich meebrengt voor de faculteit: $u_{11} = 1$, en $u_{21} = 0$. Men realiseert zich dat ook andere overwegingen mogelijk zinvol zijn en tot andere utiliteitstoekenning kunnen leiden, maar houdt het bij deze eerste keuze. Het toegelaten hebben van een geschikte student is ook erg prettig: $u_{22} = 9/10$, wat gewijzigd wordt in $u_{22} = 1$ onder pressie van een aantal docenten die aan $u_{22} = 9/10$ in dit geval geen zinvolle betekenis kunnen geven. Tenslotte, $u_{12} = 7/10$, d.w.z. het afwijzen van een geschikte student wil men wel inruilen tegen een kans $7/10$ op toelaten van een geschikte student tegen een kans $3/10$ op het toelaten van een ongeschikte student.

Faculteit D daarentegen zegt direct: $u_{22}=1$ en $u_{21}=0$, en met afgewezen studenten, of ze nu geschikt gebleken zouden zijn of niet, hebben we nooit meer iets te maken, dus $u_{11}=u_{12}$ maar we moeten wel een voorzichtig = conservatief) toelatingsbeleid voeren, dus de utiliteit van afwijzingen zouden we op $4/5$ kunnen stellen, d.w.z. voor iedere afgewezen kandidaat zouden we best iemand anders wel willen

toelaten, maar dan moet zijn kans op succes $4/5$ (of beter) zijn, in ieder geval niet minder dan $4/5$. Dus: $u_{11}=u_{12}=4/5$.

De overheid, die in dit geval de universiteitsraden naast zich vindt, voert een geheel andere argumentatie, omdat zij het selectiebeleid t.a.v. een bepaalde faculteit niet los kunnen zien van het selectiebeleid t.a.v. andere faculteiten. Of anders gezegd: het afwijzen van geschikte kandidaten door faculteit C of D is niet ongunstig (voor andere faculteiten), en het toelaten van ongeschikte studenten tot studie C of D is niet zo heel erg, wanneer het zo is dat wanneer deze studenten afgewezen zouden zijn, zij in een andere studie wél toegelaten, maar eveneens ongeschikt zouden blijken. De argumentatie zoals hier beschreven is uiteraard te grof, maar we zullen het er mee moeten doen om wille van de plaatsruimte. Stel dus dat de overheid komt tot toekenning van de volgende utiliteiten voor selectie voor een bepaalde (in vergelijking tot andere studies nogal kostbare) opleiding: $u_{22}=1$, $u_{21}=0$, $u_{11}=u_{12}=1/3$.

We kunnen nu alles in één tabel zetten:

		ongeschikt	geschikt
STUDENT A	afwijzen	$7/10$	$7/10$
	toelaten	0	1
--			
STUDENT B	afwijzen	$1/5$	0
	toelaten	$1/10$	1
--			
FACULTEIT C	afwijzen	1	$7/10$
	toelaten	0	1
--			
FACULTEIT D	afwijzen	$4/5$	$4/5$
	toelaten	0	1
--			
OVERHEID	afwijzen	$1/3$	$1/3$
	toelaten	0	1

TABEL 3. Beslissingstabellen voor diverse belanghebbenden.

Met behulp van formule (3) kunnen we voor iedere belanghebbende

(instantie) de selectieparameter p berekenen, en vinden:

voor student A: $p = 7/10$

student B: $p = 1/11$

faculteit C: $p = 10/13$

faculteit D: $p = 4/5$ (zoals uit de argumentatie al
opgemaakt kon worden voor de
keuze $u_{11}=u_{12}=4/5$)

overheid: $p = 1/3$.

Dat wil zeggen, student A zou er vrede mee hebben wanneer voor deze bepaalde studie (maar niet noodzakelijk ook voor andere universitaire studies) de toelatingsdrempel gelegd zou worden op dat eindexamen cijfer gemiddelde waarvoor gebleken is dat 7/10 van de kandidaten met dat cijfer deze studie succesvol afsluiten. Wanneer in werkelijkheid een lagere cesuur gehanteerd zou worden, zou hij uit eigen beweging liever een andere studie kiezen wanneer zijn eindexamencijfer beneden de eerstgenoemde grens ligt.

Met betrekking tot student B geldt dat zij een selectieparameter heeft gekozen die zo ligt dat iedereen toegelaten zou moeten worden, c.q. zij wil onder bijna iedere omstandigheid in de gelegenheid zijn om deze studie te kunnen gaan doen. Ook de aftestgrens die de overheid in dit voorbeeld heeft bepaald, betekent in de meeste gevallen dat alle kandidaten toegelaten worden.

Een aantekening m.b.t. numerus fixus situaties: bij een numerus fixus hebben we te maken met een vastgesteld aantal kandidaten dat toegelaten wordt. De betrokken faculteit zal in haar eigen belang dan vrijwel altijd de voorkeur geven aan het toelaten van de bovenste groep die het quotum precies vult. Het vaststellen van utiliteiten kan dan alleen nog zin hebben om te controleren of ondanks de numerus fixus de cesuur toch niet hoger gesteld moet worden dan met het aantal beschikbare plaatsen zou overeenkomen (= onderzoekje naar de noodzaak van een numerus fixus dus eigenlijk). Voor allerlei andere toelatingsregelingen, loting, gewogen loting, combinaties met een 'vrijwillig' toelatingsexamen e.d., is een utiliteitsafweging zonder meer gewenst. Met name ook in al die gevallen waar ideologische motieven (opvattingen over verdeling van kennis en macht, e.d.) een belangrijke plaats in de afweging van argumenten krijgen toebedeeld.