

De relatie tussen studietijd, studiesucces en capaciteit

Dato N.M. de Gruijter *

Summary

A model is proposed in which - apart from measurement error - study succes is a function of ability $\times$ time spent studying. A procedure for the estimation of student abilities from study results and time spent is derived in case of more than one subject. An empirical example is given.

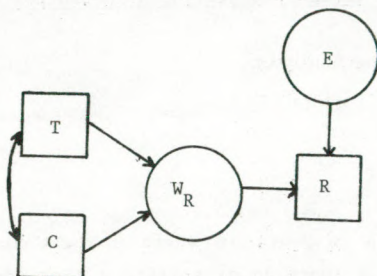
* Bureau Onderzoek van Onderwijs, Rijksuniversiteit leiden.

De auteur is dank verschuldigd aan Eric A. Bakker die het programmeerwerk verzorgde.

Inleiding

In het algemeen gaat een grotere studie-inspanning van een student samen met een hoger verwacht studieresultaat. Toch is de relatie tussen studietijd en succes in correlatieve termen niet hoog; de correlatie tussen succes en studietijd kan zelfs licht negatief zijn. Dit verschijnsel kan verklaard worden met een model waarin naast studietijd ook intellectuele capaciteit een determinant van studiesucces is (zie bijv. Tromp en Wilbrink, 1977). Studenten met een grotere tijdsinvestering zijn wellicht de zwakkere studenten terwijl de slimmeren minder tijd investeren en dezelfde resultaten behalen.

Een eenvoudig causaal model voor de relaties tussen studietijd (T), studieresultaat (R) en capaciteit (C) is weergegeven in Figuur 1.



Figuur 1. Een causaal model voor de relatie tussen studietijd en studieresultaat (voor een verklaring van de symbolen, zie de tekst).

In de figuur geven de rechte lijnen causale relaties aan, de gebogen lijn een samenhang; E is de meetfout in gemeten studieresultaat, W_R staat voor studieresultaat zonder meetfout. Er wordt in het schema vanuit gegaan dat T en C zonder meetfouten bepaald zijn. Het model wordt in dit artikel nader omschreven. Voor het geval de capaciteiten onbekend zijn, wordt een procedure gegeven voor het schatten van de parameters. De procedure wordt toegelicht aan de hand van een empirisch voorbeeld.

Het model

Het eenvoudigste model voor de relatie tussen studieresultaat x, studietijd en capaciteit is het lineaire model

$$(1) x_i = a + bc_i + dt_i + e_i$$

waarbij e_i de meetfout is en i student i aangeeft. Wij mogen aannemen dat d positief is, hetgeen betekent dat een extra inspanning Δt resulteert in een stijging van het verwachte studiesucces $\Delta x = d\Delta t$. De coëfficiënten b en d uit dit lineaire causale model kunnen geschat worden aan de hand van de variantie-covariantie-matrix van x , c en t .

Nu geeft het totaal aantal tentamenvragen een bovengrens aan de mogelijk te behalen score, en als een student al vrij veel van het betreffende vak afweet, zal een extra studie-inzet tot een geringere verhoging van het resultaat leiden dan wanneer hij er nog niet zoveel van weet. Model (1) lijkt door dit plafondeffect niet erg plausibel. Een plafondeffect kan vermeden worden door het gebruik van een score-transformatie die de scoreschaal bij hoge scores uitrekt. Indien mogelijk, kan men geschatte latente trekparameters gebruiken in plaats van de oorspronkelijke scores. Zo onderzocht Fischer (1976) veranderingen op de schaal van het additieve Rasch model.

Men kan zich ook afvragen of de veronderstelling van de additiviteit van de effecten c en t reëel is. Het additieve model houdt in dat twee studenten die evenveel geleerd hebben, evenveel erbij leren op de gehanteerde scoreschaal als zij dezelfde extra studietijd investeren, ongeacht hun capaciteiten. Een multiplicatief model (Van Naerssen, 1971; De Groot, 1979) waarbij capaciteit gedefinieerd wordt als leersnelheid ofwel als leerwinst per tijdseenheid, is te prefereren.

Het model met deze capaciteitsdefinitie wordt nu

$$(2) y_i = a + bc_i t_i + e_i \quad b > 0; c_i > 0 \quad i=1, \dots, N$$

waarbij y een transformatie van de totaalscore is en i student i aangeeft; formule (2) geeft dus het multiplicatief model bij getransformeerde toets-scores.

Uit de vergelijking is duidelijk dat in dit model verondersteld wordt dat bij de aanvang van de voorbereiding op het tentamen alle studenten dezelfde voorkennis hebben: als $t_i=0$, dan is de verwachte score gelijk aan de constante a . In de praktijk blijkt de onderzoeker veelal wegens problemen van organisatorische of financiële aard niet in staat alle relevante gegevens te verzamelen: of de tijdbestedingsgegevens of capaciteitsschattingen ontbreken. In de volgende paragraaf wordt - voor het geval meerdere vakken getentamineerd zijn en c de relevante capaciteit voor al deze vakken is - een procedure geschetst om de onbekende c 's te schatten.

De schatting van de capaciteitsparameters

Eerst wordt een transformatie van de totaalscores gekozen. Een bruikbare transformatie van de totaalscores x_i bij 0-1 scores voor de toetsvragen is

$$(3) y_i = \ln \frac{x_i}{n - x_i} \quad x_i \neq 0; x_i \neq n,$$

waarbij n het aantal vragen in de toets is. De transformatie is tevens op een lineaire transformatie na een benadering van de meest aannemelijke schatter van de persoonsparameter in het Rasch model onder de veronderstelling dat de itemparameters normaal verdeeld zijn (Wright & Stone, 1979, p. 21); in dit artikel wordt overigens niet verondersteld dat het Rasch model past. De y_i kunnen gestandaardiseerd worden en in het vervolg zullen wij ervan uitgaan dat dat gebeurd is.

Als c onbekend is, kan men de volgende schatting van c_i in formule 2 gebruiken:

$$(4) \hat{c}_i = y_i / t_i$$

met $a=0$ en $b=1$. Het is duidelijk dat het model in het geval van één tentamen ontoetsbaar is; bij één tentamen is het altijd mogelijk om een perfecte pas van het model te verkrijgen.

Veronderstel nu dat men beschikt over resultaten van N personen op k vakken die een beroep doen op dezelfde capaciteit.

Wij herschrijven (2) nu tot

$$(5) y_{ij} = a_j + b_j c_i t_{ij} + e_{ij} \quad i=1, \dots, N; \quad j=1, \dots, k.$$

Weer wordt aangenomen dat de beginkennis voor alle studenten hetzelfde is, nl. $a_j (j=1, \dots, k)$. Dat betekent dat verondersteld wordt dat er geen transfer is van het ene vak op het andere; die mogelijkheid zou namelijk inhouden dat er individuele verschillen in het beginniveau op een vak kunnen ontstaan door verschillen in beheersing van andere vakken.

De c_i liggen op een ratio schaal, d.w.z. $c_i^* = k c_i$ en $b_j^* = k^{-1} b_j$ voldoen even goed als c_i en b_j . Wij kunnen de schaal van de c_i vastleggen door de eis dat de standaardafwijking van de c_i 's gelijk is aan één.

De c_i - en de a_j en b_j - moeten geschat worden uit de y_{ij} en de t_{ij} . De kleinste kwadraatoplossing voor deze parameters wordt gegeven door de schattingen die

$$(6) S = \sum_{ij} e_{ij}^2 = \sum_i \sum_j (y_{ij} - a_j - b_j c_i t_{ij})^2$$

minimaliseren. Gegeven een schatting c_i ($i=1, \dots, N$) kunnen de a_j en b_j gevonden worden door minimalisatie met betrekking tot a_j en b_j van

$$(7) S_j = \sum_i (y_{ij} - a_j - b_j z_{ij})^2 \quad j=1, \dots, k,$$

waarbij

$$(8) z_{ij} = \hat{c}_i t_{ij}.$$

Dit minimalisatieprobleem is een bekend probleem; b_j is de regressiecoëfficiënt van de regressie van y_{ij} op z_{ij} voor tentamen j :

$$(9) b_j = r_{yz(j)} / s_{z(j)} \quad j=1, \dots, k.$$

Voor a_j vindt men

$$(10) a_j = -b_j \bar{z}(j) \quad j=1, \dots, k.$$

Gegeven een schatting van de a_j en b_j vindt men een schatting voor c_i door differentiatie van S naar c_i en het op nul stellen van het resultaat.

Men krijgt de volgende schatting

$$(11) c_i = \frac{\sum_j b_j t_{ij} (y_{ij} - a_j)}{\sum_j b_j^2 t_{ij}^2} \quad i=1, \dots, N.$$

Deze schatting is analoog aan de kleinste-kwadratenschatting van factorscores in het één-factormodel uit de factoranalyse.

Aangezien c_i gedefinieerd is als een positief getal wordt een negatieve schatting vervangen door een klein positief getal. Na de iteratie worden de c_i 's gedeeld door hun standaardafwijking. De nieuwe waarden c_i zijn de uiteindelijke c-waarden in de desbetreffende iteratie.

De twee schattingsfasen, die van de capaciteiten en die van de a_j en b_j , kunnen alternerend toegepast worden tot de capaciteitsschattingen geconvergeerd zijn. Een mogelijke startwaarde voor de capaciteiten is $c_i=1$ voor alle waarden van i .

De niet door het model verklaarde variantie is gelijk aan

$$(12) S/N = \sum_j S_j / N = \sum_j (1 - r_{yz(j)}^2).$$

De correlaties $r_{yz(j)}$ geven dus een goede indruk van de voorspellende waarde van het model.

Een toepassing

Binnen het kader van een onderzoek naar het nieuwe eerstejaars curriculum rechten van de Rijksuniversiteit Leiden in het studiejaar 1979-1980 (Crombag e.a., 1980) werden o.a. tijdbestedingsgegevens verzameld voor een deelgroep van studenten. Aan het eind van het eerste semester waren er 190 studenten waarvan tijdbestedingsgegevens en de resultaten van drie meerkeuze-tentamens bekend waren; bij deze groep studenten kwamen er geen perfecte scores noch scores gelijk aan 0 voor (zie formule 3). De tentamens Rechtssociologie (RE) en Ideeëngeschiedenis (ID) hadden elk 26 vragen, het tentamen Strafrecht (ST) 76 vragen. Enkele gegevens m.b.t. de vakken zijn vermeld in de Tabellen 1 en 2.

Tabel 1. Enkele gegevens m.b.t. de vakken in het onderzoek.

	$\bar{t}^*$	s_t
ID	4.59	2.21
RE	4.30	2.21
ST	15.98	5.22

* uitgedrukt in tientallen uren

Tabel 2. Intercorrelaties

	t_{ID}	t_{RE}	t_{ST}	y_{ID}	y_{RE}	y_{ST}
t_{ID}	-					
t_{RE}	.71	-				
t_{ST}	.59	.65	-			
y_{ID}	.24	.09	.19	-		
y_{RE}	.11	.13	.24	.59	-	
y_{ST}	.14	.05	.25	.63	.50	-

Uit Tabel 2 blijkt dat de correlaties tussen studietijd en resultaat vrij laag zijn. Verder blijkt dat de studenten een min of meer consistent studiegedrag vertonen in termen van voorbereidingstijd. De correlaties tussen de (getransformeerde) resultaten zijn als gevolg daarvan ook redelijk hoog. Dat betekent dat de studieresultaten ook goed te beschrijven zijn aan de hand van het éénfactor model

$$(13) y_{ij} = d_j f_i + e_{ij} \quad j=1, \dots, k \quad i=1, \dots, N$$

waarbij de d_j de factorladingen zijn. Bij een traditionele analyse van studieresultaten, een analyse waarbij geen rekening gehouden wordt met de tijdsbesteding van de studenten, ligt een analyse volgens model (13) voor de hand. De factorladingen zijn gemakkelijk te berekenen aangezien drie variabelen altijd in termen van één gemeenschappelijke factor beschreven kunnen worden met ladingen

$$(14) d_j = \sqrt{\frac{r_{jj'} r_{jj''}}{r_{j'j''}}}.$$

Deze ladingen kan men, tezamen met de coëfficiënten van de eerste principale component, vinden in Tabel 3. De coëfficiënten van de eerste principale component zijn hier interessant aangezien de coëfficiënten van de eerste principale component de kleinste kwadratenoplossing geven van de ladingen d_j uit het model dat in formule (13) gegeven wordt bij een gelijktijdige schatting van de d 's en de f 's.

Tabel 3. Correlaties tussen resultaat en onderliggende factor.

	één-factor oplossing	eerste principale component	model- correlaties
ID	.86	.88	.85
RE	.68	.82	.79
ST	.73	.84	.85

De correlaties $r_{yz(j)}$ uit model (5) (met $z_{ij}=c_{it_{ij}}$) staan in de derde kolom van Tabel 3. Deze correlaties zijn redelijk hoog, ook in vergelijking met de coëfficiënten van de eerste principale component; bovendien moeten wij er rekening mee houden dat de betrouwbaarheden van de toetsen verre van perfect zijn. De correlaties tussen capaciteit uit model (5) en studietijd zijn -0.22 voor ID, -0.28 voor RE en -0.16 voor ST, terwijl de correlaties tussen capaciteit en resultaat boven de 0.65 liggen. Dat wil zeggen, studenten met een hoge capaciteit kunnen in minder tijd betere resultaten behalen dan hun zwakkere medestudenten en dat doen zij ook.

De hoogte van de coëfficiënten van de eerste principale component geeft beperkingen bij het onderzoek van alternatieve modellen.

Elk model van de vorm

$$(15) y_{ij} = a_j + d_j c_i + z_{ij} + e_{ij},$$

waarbij z_{ij} in het model nader gespecificeerd moet worden, verklaart minstens evenveel variantie als de eerste principale component. Het is waarschijnlijk dat 'c' in het model dat in formule (15) gegeven wordt veel van de gemeenschappelijke variantie tussen de resultaten opneemt, ook de variantie van gemeenschappelijke factoren die niet in het model gespecificeerd zijn. In een dergelijk geval reflecteert 'c' niet de factor 'capaciteit' uit het model. Om deze reden is er geen analyse van mogelijke alternatieve modellen van de vorm (15) opgenomen.

De hoogte van de modelcorrelaties in Tabel 3 is ongetwijfeld geflatteerd door het feit dat de capaciteiten waarop zij gebaseerd zijn, uit hetzelfde materiaal berekend zijn. Om de invloed van dit effect enigszins na te gaan werden

de modelcorrelaties $r_{yz(j)}$ ook berekend met capaciteiten, geschat m.b.v. de resultaten en tijden van de twee overige vakken. Deze correlaties waren 0.68 voor ID, 0.59 voor RE en .060 voor ST. Zij geven een ondergrens aan de voorspellende waarde van het model aangezien zij gebaseerd zijn op de gegevens van slechts twee vakken.

Discussie

Het blijkt mogelijk om studieresultaat redelijk goed te beschrijven in termen van een eenvoudig model voor de relatie tussen studietijd, capaciteit en studieresultaat (zie voor een meer uitgewerkt model Van Naerssen, 1976). Een verdere modeltest zou verricht kunnen worden in nader onderzoek met herhaalde metingen, bijv. met een pretest en een posttest.

Voor in de praktijk t.b.v. studenten bruikbare resultaten zou men ruim voor de tentamendata - dus voordat men over de resultaten y_{ij} beschikt - over capaciteitsschattingen moeten beschikken. Dat betekent dat de capaciteiten uit het model d.m.v. externe capaciteitsmetingen redelijk geschat moeten kunnen worden (zie ook Bryk, e.a., 1980). Hiervoor zou men tests moeten afnemen waarvan men aanneemt dat zij aan de model-capaciteit gerelateerd zijn en vervolgens zou men de gevonden capaciteitsparameters uit het model hierop moeten regresseren. Indien men tests vindt die de modelcapaciteit goed voorspellen, dan kunnen deze tests in latere toepassingen gebruikt worden.

Literatuur

- Bryk, A.S., Strenio, J.F. en Weisberg, H.I. A method for estimating treatment effect when individuals are growing. Journal of Educational Statistics, 1980, 5, 5-34.
- Crombag, H.F.M., De Gruijter, D.N.M., Van der Ende, P. en Vos, P. De nieuwe propaedeuse in de Faculteit der Rechtsgeleerdheid: verslag over het eerste semester. Rapport No. 20. Leiden: Bureau Onderzoek van Onderwijs
- Fischer, G.H. Some probabilistic models for measuring change. In: D.N.M. de Gruijter and L.J.Th. van der Kamp Advances in psychological and educational measurement. London: Wiley, 1976.
- De Groot, A.D. Studielast en normstudent: ontwerp van een akkoord-theorie 1. Algemeen model. Tijdschrift voor Onderwijsresearch, 1979, 4, 257-274.
- Naerssen, R.F. van, Een model voor tentamens (vervolg). Nederlands Tijdschrift voor de Psychologie (DPO), 1971, 26, 551-559.
- Naerssen, R.F. van, Het derde tentamenmodel met een toepassing. Tijdschrift voor Onderwijsresearch, 1976, 1, 161-171.
- Tromp, D. en Wilbrink, B. Het meten van studietijd. In: Congresboek Onderwijsresearchdagen 1977, Stichting Onderwijsresearch, 1977.
- Wright, B.D. and Stone, M.H. Best test design. Chicago: Mesa Press, 1979.