

	enkelvoudig ('univariaat')	meervoudig ('multivariaat')
formule:	$y = \alpha + \beta_1 x_1$	$y = \alpha + \beta_1 x_1 + \dots + \beta_n x_n$
Symbol	Betekenis	
y	afhankelijke variabele = gemiddelde van de uitkomstvariabele, hier het cholesterolgehalte	
α	snijpunt met de y -as, hier de waarde van het cholesterolgehalte indien $x_1 = 0$, in het univariate model	
β_1	de regressiecoëfficiënt behorende bij de primaire determinant: geeft de verandering in y (hier het cholesterolgehalte) aan, per eenheid verandering in x_1 (hier lichamelijke activiteit als dichotome variabele). In het multivariate model is deze coëfficiënt gecorrigeerd voor alle in het model opgenomen covariabelen (potentiële confounders)	
x_1	onafhankelijke variabele, primaire determinant	
β_i	de voor alle andere variabelen gecorrigeerde regressiecoëfficiënten behorende bij de desbetreffende confounders	
x_i	onafhankelijke variabelen, verschillende confounders	

Schema 2 Het lineaire regressiemodel

Dit wordt aangeduid met de term effect-modificatie. Daarvoor moet er dan wel sprake zijn van een 'wezenlijk' verschil in effect tussen de strata. Als je beide groepen bij elkaar neemt kom je uit op een verschil ergens daartussen in. Maar hiermee doe je de werkelijkheid geen recht: voor jongeren zit je dan te laag, en voor ouderen lijkt er een effect te zijn dat er in werkelijkheid niet is.

In ons voorbeeld is geen sprake van effect-modificatie, dus kunnen we met een gerust hart overstappen naar een model. Omdat de uitkomst een continue variabele is, is het lineaire regressiemodel een geschikt model. In zijn meest eenvoudige vorm, het univariate model, ziet de vergelijking er als volgt uit: $y = \alpha + \beta_1 x_1$. Zie schema 2 voor een toelichting. Waar alles om draait is de regressiecoëfficiënt β_1 van de primaire determinant. Deze is te interpreteren als de (gemiddelde) verandering in de y -variabele per eenheid verandering in de x_1 -variabele. Vinden we bijvoorbeeld een regressiecoëfficiënt van 0,5 mmHg/kg voor het verband tussen gewicht en bloeddruk, dan betekent dit dat bij personen die 1 kg zwaarder zijn de bloeddruk gemiddeld 0,5 mmHg hoger is; en bij degenen die 10 kg zwaarder zijn, is de bloeddruk gemiddeld 5 mmHg hoger. In ons voorbeeld is de primaire determinant een indeling in twee categorieën van lichamelijke activiteit. Indien we x_1 de waarde 0 geven bij mensen die lichamenlijk actief zijn en de waarde 1 voor wie niet-lichamelijk actief is, dan geeft de regressiecoëfficiënt β_1 het verschil in cholesterolgehalte aan tussen lichamenlijk niet- en wel-actieven. In een model met alleen deze indicatorvariabele voor activiteit erin krijgt β_1 de waarde 0,8 mmol/L: het 'ruwe' verschil.

Als we nu willen corrigeren voor leeftijd dan voegen we de variabele leeftijd aan het model toe. Dat mag als continue variabele, maar ook, zoals in ons voorbeeld, als een variabele met slechts twee categorieën. Deze variabele krijgt zelf ook een β -coëfficiënt (die in dit geval 1,8 mmol/L zal zijn), maar die is voor de vraagstelling minder interessant. We zijn immers in de eerste plaats geïnteresseerd in het 'zuivere' effect van lichamenlijke activiteit en kijken daarom met name hoe de β_1 verandert door rekening te houden met leeftijd. Deze coëfficiënt zal nu geen 0,8 mmol/L zijn, maar 0,2 mmol/L. En mochten er nog meer variabelen zijn waarvoor we als potentiële confounder willen (en mogen) corrigeren, dan voegen we deze aan het model toe en kijken wederom hoe de β_1 -coëfficiënt verandert.

Andere toepassingen

Soms is de onderzoeksvraag iets anders, namelijk: "Hoe kunnen we de afhankelijke variabele zo goed mogelijk verklaren met behulp van een spaarzaam, maar toch voldoende volledig model?" In dat geval is er niet één primaire determinant die de aparte rol van blootstelling vervult, maar voegen we variabelen aan het model toe op geleide van andere criteria. Het is zelfs mogelijk dit via een 'automatische procedure' te laten doen, maar daar moet je wel mee oppassen. Het statistisch programma selecteert dan alleen die variabelen die nog een 'statistisch significante' bijdrage leveren aan de voorspelling van de uitkomst. Het helpt om daarbij van te voren te bepalen wat je een 'relevant' effect vindt, maar het risico blijft bestaan dat je bepaalde qua grootte 'relevante' effecten niet zult vinden, omdat ze (net) niet statistisch significant zijn.

Bij deze onderzoeksvraag wil men soms het relatieve gewicht van de diverse determinanten weergeven. Omdat echter variabelen verschillende eenheden hebben (leeftijd in jaren, gewicht in kg, alcoholgebruik in eenheden per tijds-eenheid, geslacht als dichotome variabele enzovoort) zijn de gewone regressiecoëfficiënten daarvoor niet bruikbaar. Om die reden kiest een auteur wel voor vermelding van zogenaamde gestandaardiseerde β -coëfficiënten. Deze geven de gemiddelde verandering in de afhankelijke variabele, in eenheden van standaarddeviaties $SD(y)$, per SD -eenheid verandering in x (al dan niet gecorrigeerd voor confounders). Een voorbeeld hiervan is te vinden in TSG 2006;84:13-21 in tabel 3 op pagina 18. Het voordeel van deze weergave is dat de β -coëfficiënten onderling qua 'effectsterkte' vergelijkbaar zijn. Daar staat natuurlijk wel tegenover dat de β -coëfficiënt niet direct interpreteerbaar is als "het gemiddelde verschil in de uitkomst, gegeven een eenheid verschil in de determinant".

Bevorderen interpreteerbaarheid

Als we ons beperken tot het gebruik van het model om het effect van de primaire determinant te corrigeren voor confounders, dan adviseren wij auteurs om zelf te bepalen voor welke confounders men wil corrigeren en deze consequent aan het model toe te voegen, of ze nu 'statistisch significant' zijn of niet. Geef daarbij altijd zowel de 'ruwe' als de gecorrigeerde verschillen, zodat de lezer kan zien wat voor effect de correctie heeft gehad en of dat qua richting passend is bij hetgeen hij al vooraf had beredeneerd. Voor de juiste interpretatie van de β -coëfficiënten is het verder van belang dat de auteur goed

aangeeft welke eenheden de primaire determinant en de confounders hebben. Als we bijvoorbeeld leeftijd niet in twee categorieën maar als continue variabele in het model hadden gestopt, dan was de β -coëfficiënt voor leeftijd veel kleiner dan 1,8 geweest omdat dan de eenheid mmol/L per jaar was geweest. Bij dichotome variabelen is het essentieel dat aangegeven wordt wat de referentie- ($x_1 = 0$) en wat de index-categorie ($x_1 = 1$) is. Zo wordt de 'gemakkelijk' te interpreteren β -coëfficiënt ook voor iedereen echt 'interpreteerbaar'.

CORRESPONDENTIEADRES

Dr. Willy-Anne H.J. van Stiphout, vrijgevestigd sociaal geneeskundige en docent epidemiologie, www.decijfersdebaas.nl of e-mail: stip.info@xs4all.nl

A LETTER FROM MELBOURNE

Internationaal vergelijken, maar hoe?

Ooit wilde ik 'international comparative health systems research' studeren. Ik meende dat, wanneer ik die kennis maar eenmaal onder de knie had, ik de onrechtvaardigheid in de wereld een beetje zou kunnen rechtzetten.

Ik vertrok naar de University of California in Berkeley. Daar kon je ICHSR studeren. Tot mijn verbijstering bleek het echter te gaan om een soort slimme epidemiologie, waarmee uitsluitend bevestigd werd dat er verschillen waren tussen landen.

Ik had daar drie problemen mee. In de eerste plaats ging het voornamelijk om 'international comparative health care systems research'. Men bekeek, bijvoorbeeld, of aantallen gezondheidsprofessionals en hun verdeling over landelijke en verstedelijkte gebieden een invloed had op gezondheidsuitkomsten. Ten tweede had de epidemiologische benadering weinig van doen met culturele aspecten van vormgeving en implementatie van gezondheidssystemen. Ook medisch antropologen konden mij daar maar ten dele raden: zij keken met name naar cultuurinvloeden op individueel gezondheidsgedrag, en niet naar het systeem als geheel. En ten derde had het geloof in mij postgevat dat het niet het zorgsysteem was dat bepaalde of er gezondheidsverschillen bestonden, maar andere zaken zoals inkomensverdeling, onderwijsniveau, fysische en sociale geografie, handelsverhoudingen, enzovoort. Ik werd niet gelukkig in Berkeley. En hoewel de 'nieuwe Max Weber' - Manuel Castells - mij maar blééf proberen uit te leggen dat mijn ideeën over systeem-determinanten van gezondheid interessant en vernieuwend waren koos ik voor de studie van gezondheidsbeleid in Nederland. Het duurde overigens ook nog tien jaar voordat Manuel zélf het eerste deel van zijn opus magnus afleverde¹ en mij opnieuw probeerde te verleiden groots en meeslepend te denken.

Tegen die tijd was ik secretaris-generaal van de Association of Schools of Public Health in the European Region met andere zorgen. Het aantal Schools explodeerde: overheden hadden

ontdekt dat er moest worden geïnvesteerd in de volksgezondheid en dat daarvoor speciaal opgeleide professionals handig waren. Daarnaast probeerden landen uit midden- en oost-Europa na de val van de Muur een inhaalslag te maken van sanitaire en hygiënische wetenschappen naar een (post-) moderne public health. Ik geloofde dat het aardig zou zijn om een verzameling minimum-kwaliteitseisen voor opleidingen in de volksgezondheid te bedenken. Och, had ik toen maar het soort 'international comparative health systems research' kunnen studeren dat mij had kunnen laten begrijpen waarom die tientallen opleidingen er zo'n moeite mee hadden zich te voegen naar mijn 'Gouden Standaard'...

Het gevolg van mijn ideële inspanningen was wel dat ik werd uitgenodigd om in Denemarken een School of Public Health op te zetten. En pas daar daagde mij de crux van de moeite die wetenschappers hebben met internationale gezondheidssysteemvergelijkingen.

Als je naar het buitenland gaat verhuizen is de vraag of je daar kunt 'aarden'. Nou, dacht ik, dat zal voor Leeuw in Denemarken allerm minst een probleem zijn. Was ik niet al tientallen malen in Kopenhagen geweest? Was Denemarken niet een verzorgingsstaat, net als Nederland? Stond Denemarken niet, net als het Vaderland, bekend als een tolerant land dat haar rol in de wereldgemeenschap consciëntieus vervult?

In mijn meer gedeprimeerde momenten betitelde ik de Deense verzorgingsstaat als Stalinistisch. Als vreemdeling heb ik me meer dan eens gediscrimineerd en buitengesloten gevoeld. En dan dat gezondheidszorgsysteem!

Het is een deel van de Deense cultuur om verantwoordelijkheid voor gezondheid (szorg) op het laagste niveau te leggen, het liefst bij het individu. Als dat individu zich kapot wil roken, eten en drinken dan gaat de Staat uitleggen dat dat niet verstandig is. Dat dat ook voor de zorg geldt ontdekte ik pas toen ik eens zelf in het ziekenhuis moest worden opgenomen. Vergeleken met Nederland vond ik het systeem onhandig, te ver gedecentraliseerd en ongecoördineerd. Decentrale verantwoordelijkheid voor planning en financiering van voorzieningen lijkt zinnig, maar wat heb ik in dat Deense ziekenhuis gewenst dat ik in een Nederlands bed lag. Begrip en vergelijking van