

SCHATTEN VAN STRUCTURELE EN
FUNCTIONELE RELATIES
Op zoek naar toepassingen van
de theorie *)

Ir. A.A.M. Jansen

Inhoud

	pag
Samenvatting	52
1. Inleiding	53
2. Voorbeelden	54
3. Schatten van relevante parameters	60
4. Waarvoor zijn de verkregen schattingen bruikbaar	65
5. Conclusies	70
6. Literatuur	71

SAMENVATTING

Aan de hand van een aantal praktische voorbeelden wordt duidelijk gemaakt dat men voor het schatten van een functionele of structurele relatie eerst zorgvuldig deze relatie en het bijpassende model voor de meetfouten moet definiëren. De opzet van het experiment en de analyse van de gegevens moeten daarop worden afgestemd. Tenslotte wordt ingegaan op het gebruik van een geschatte relatie voor predictiedoeleinden, en het verschil dat daarbij kan worden gemaakt tussen functionele en structurele relaties.

*) Inhoud van een voordracht, op 30-01-1981 gehouden voor de Medisch Biologische Sectie en de Landbouwkundige Sectie van de Vereniging voor Statistiek.

1. INLEIDING

Het begrip 'functionele relatie' wordt in de statistiek gehanteerd als tussen twee grootheden Y en X een strict monotone relatie bestaat. Veelal beperkt men zich tot lineaire relaties:

$$Y = \alpha + \beta X$$

Hoewel dat niet essentieel is zal deze beperking eenvoudigheidshalve ook hier worden verondersteld. Als X een stochastische variabele voorstelt, en Y dus ook, spreekt men van een 'structurele relatie' tussen X en Y.

De hier te bespreken statistische methodologie komt aan de orde als men de al dan niet stochastische grootheden X en Y gaat meten en daarbij meetfouten maakt die worden verondersteld stochastisch te zijn. Eenvoudigheidshalve zal steeds worden aangenomen, dat zulke meetfouten een verwachtingswaarde gelijk aan 0 en voor alle X dezelfde variantie hebben en verder onderling en van X en Y onafhankelijk verdeeld zijn. Tenslotte wordt verondersteld dat alle meetfouten, en in het geval van een structurele relatie ook X, simultaan normaal verdeeld zijn, tenzij anders vermeld. Deze beperkingen zijn niet essentieel voor het betoog, hoewel uiteraard de presentatie en de formulering van de resultaten daar wel van afhankelijk zijn.

Als alleen bij de meting van X of van Y meetfouten worden geïntroduceerd kan de relatie geschat worden met behulp van lineaire regressie-theorie. Als beide variabelen met meetfouten zijn behept, vergt het schatten van de relatie speciale hulpmiddelen, die te zamen kunnen worden aangeduid als de theorie van het schatten van functionele en structurele relaties. Een goed overzicht geeft Linssen (1980); deze auteur maakt tevens aannemelijk dat er met betrekking tot het schatten geen onderscheid tussen functionele en structurele relaties behoeft te worden gemaakt.

In het volgende zal niet worden ingegaan op de schattingsmethoden zelf. Wel zal aan de hand van een aantal praktijkvoorbeelden, waarop het model van toepassing is, worden aangeduid welke problemen zich voordoen bij de definitie van de meetfouten, en de consequenties daarvan voor

proefopzet en analyse. Verder zal worden getracht aannemelijk te maken dat men in de praktijk niet altijd geïnteresseerd is in de structurele en functionele relaties zelf, maar in functies die kunnen worden gebruikt voor het schatten van de ene grootheid op basis van metingen aan de andere grootheid. In het vrij vaak voorkomende geval van structurele relaties vereist dit het schatten van andere dan de modelrelaties zelf. Het verband tussen beide, en daaruit voortvloeiende eisen voor de proefopzet worden aan de orde gesteld.

2. VOORBEELDEN

Het is instructief te beginnen met de bespreking van een aantal praktische voorbeelden. Wellicht is het niet triviaal dat bij al deze voorbeelden functionele of structurele relaties in het geding zijn. Desondanks bieden de voorbeelden houvast bij de bestudering van de theorie en de praktische consequenties daarvan.

Voorbeeld 1: Relatie tussen de hoeveelheid progesteron in het bloed van dragende ooiën en het aantal lammeren dat zij dragen.

Bij het Nederlandse schapenras, de Texelaar, werpt het merendeel van de ooiën per dracht 1 of 2 lammeren. Tijdens de dracht is de hoeveelheid progesteron in het bloed van de moederdieren afhankelijk van het aantal vruchten. Het aantal vruchten, X , is voor een gegeven drachtige ooi gelijk aan 1 of 2. De gemiddelde hoeveelheid progesteron in het bloed, Y , is bij Texelaars gelijk aan μ_1 respectievelijk μ_2 , afhankelijk van het aantal vruchten. Er bestaat dus een functionele relatie tussen het werkelijke aantal vruchten en het desbetreffende populatiegemiddelde van de hoeveelheid progesteron in het bloed. De variatie van individuele ooiën rond deze waarden zou men, met enige goede wil, kunnen opvatten als meetfouten bij de waarneming van Y ; voor de gemeten hoeveelheid progesteron y geldt dan:

$$y = Y + e, \quad \text{met } e \sim \sigma_e X, \quad \text{en } Y = E y | X.$$

Het verdient echter de voorkeur in dit geval de stochastische bijdrage e te schrijven als de som van een echte meetfout e_m , met variantie σ_{em}^2 , en

een structurele afwijking e_s , met variantie σ_{es}^2 . Het deel van de meetfout, dat bij herhaalde metingen aan hetzelfde individu of object steeds dezelfde realisatie oplevert, die eigen is aan dat individu of object, zal steeds worden beschouwd als het structurele deel. Als men bovendien nog aan de beschrijving van het model toevoegt de frequenties π_1 en π_2 , waarmee worpen van 1 en 2 lammeren in de populatie voorkomen, spreekt men van een structurele relatie.

Voorbeeld 2: Relatie tussen het referentie vleespercentage en de spekdikte bij varkensarkassen.

De kwaliteit van varkensarkassen wordt ingevolge een internationale afspraak gewaardeerd met het percentage mager vlees, X . Dit kan vrijwel foutvrij worden vastgesteld, dat wil zeggen $e_m = 0$, maar dit is tijdrovend en bovendien destructief voor wat betreft de verkoopbaarheid van de arkassen. Daarom bestaan er classificatiemethoden die gebaseerd zijn op het meten van de spekdikte; deze methoden geven met het vleespercentage gecorreleerde resultaten. Voor een meting met zo een eenvoudige methode, y , kan men precies hetzelfde model hanteren als in voorbeeld 1:

$$y = Y + e, \text{ met } e = e_s + e_m \text{ en } Y = E y | X = \alpha + \beta X.$$

Als functionele relatie kan men opvatten de relatie tussen de conditionele verwachtingswaarde van de spekdikte, Y en vleespercentage, X ; het verschil met het vorige voorbeeld is dat X een continue variabele is. Ook hier geldt, dat als X een kansverdeling heeft, men kan spreken van een structurele relatie.

Voorbeeld 3: Meting van spermadichtheid.

Bij kunstmatige inseminatie van varkens is het van belang de dichtheid van spermatozoën in een ejaculaat vast te stellen om een optimale verdunning te kunnen kiezen. Dit kan geschieden met tijdrovende telkamer-methoden met behulp van een microscoop, maar ook met benaderende colorimetrische methoden die berusten op de absorptie van doervallend licht van een bepaalde golflengte. De met de telkamer verkregen resultaten geven in principe de ware dichtheid X weer, maar zijn behept met meetfouten, omdat men beperkt is ten aanzien van de per monster te besteden tijd. De colorimetrische methoden geven resultaten, y , die behoudens een

meetfout e een functie van de ware dichtheid zijn. Deze functie is de gezochte functionele relatie, of als X een verdeling heeft, de gezochte structurele relatie. De meetfout in y kan men ook hier samengesteld denken uit een echte meetfout e_m en uit een structurele afwijking e_s die specifiek is voor elk ejaculaat en bijvoorbeeld veroorzaakt kan zijn door andere substanties dan spermatozoën. Derhalve geldt:

$$y = Y + e, \text{ met } e = e_s + e_m \text{ en } Y = E y | X = \alpha + \beta X.$$

Het verschil met voorbeeld 2 is dat hier een meetfout wordt verondersteld bij het meten van X ; men meet x waarvoor geldt:

$$x = X + d \quad d \stackrel{\infty}{=} \sigma_d X$$

Omdat men bij het tellen alleen spermatozoën telt kan men hier aannemen dat d alleen echte meetfout voorstelt: $d = d_m$, met variantie σ_{dm}^2 . Het model dat hier van toepassing is wordt onder meer beschreven door Cox (1968) en Acton (1959), door hen aangeduid als 'model met variantiecomponenten', en door Johnston (1963).

Voorbeeld 4. Meting van de dichtheid van een asfaltlaag in een wegdek.

Een algemeen gehanteerde maat voor de kwaliteit van een asfaltlaag is de dichtheid. De meting daarvan wordt conventioneel voorgeschreven via een directe bepaling via het onderwatergewicht aan boorkernen die op een aantal plaatsen uit het wegdek worden geboord. Deze meting levert dus waarden x waarvoor het model van voorbeeld 3 in principe opgaat.

Een verschil is dat d hier, vooral bij constructies met relatief veel holle ruimte, een structurele component d_s kan bevatten die specifiek is voor een te meten boorkern en die wellicht een negatieve verwachtingswaarde heeft: bij het wegen onder water mag het water in principe niet de boorkern binnendringen. We zullen aannemen dat dit met behulp van een paraffinehuidje wordt voorkomen. Omdat deze conventionele methode tijdrovend is en bovendien de resultaten pas laat beschikbaar komen, tracht men de dichtheid in situ te bepalen met behulp van nucleaire methoden. Voor de daarmee verkregen metingen, y , geldt hetzelfde model als in voorbeeld 3. Omdat deze methoden de dichtheid indirect meten is er alle aanleiding te veronderstellen dat elke meetplaats zijn eigen specifieke

bijdrage aan de meetfout zou kunnen leveren, afhankelijk van de samenstelling van het asfaltmengsel ter plaatse van de meting. Zou men de boven gesuggereerde structurele afwijking in x toelaten dan is niet uitgesloten dat deze met de structurele afwijking in y gecorreleerd is. Evenals in de vorige voorbeelden kan ook hier een voor de hand liggende functionele of structurele relatie tussen Y en X worden gedefinieerd.

Voorbeeld 5. Relatie tussen gewicht en lengte van individuen.

Als men van een groep volwassen mensen hun gewicht en hun lengte meet vindt men een twee-dimensionale verdeling van uitkomsten, die men in het geval een normale verdeling verondersteld mag worden, kan karakteriseren met gemiddelden, varianties en covariantie van lengte en gewicht. Deze verdeling is een eigenschap van de groep mensen die men beschouwt, en heeft niet van doen met meetfouten (die gemakshalve verwaarloosd worden in dit voorbeeld). Er is ook geen enkele hiërarchie in de variabelen, zoals in de voorgaande voorbeelden, en bijgevolg is er ook geen zinvolle functionele of structurele relatie te definiëren (Cox, 1968; Ehrenberg, 1975, ch 14).

De situatie is echter verschillend als men geen volwassen mensen meet, maar mensen op een bepaald tijdstip gedurende hun groeiperiode. Een functionele relatie waarin men geïnteresseerd zou kunnen zijn is de relatie tussen de gemiddelde lengte, X , en het gemiddelde gewicht, Y , in de groeiende populatie, een soort groeicurve dus. In het algemeen zal dit niet 'het gemiddelde' zijn van de groeicurven van de afzonderlijke individuen! De structurele meetfouten d_s en e_s bij deze relatie, met varianties $\sigma_{d_s}^2$ en $\sigma_{e_s}^2$ respectievelijk, kan men op zinvolle wijze definiëren als de afwijkingen van de individuen ten opzichte van het gemiddelde van de sub-populatie van even oude individuen. Merk op dat de afwijkingen ten aanzien van lengte en ten aanzien van gewicht nu onderling gecorreleerd zijn, en dat de covariantiematrix van de leeftijd zou kunnen afhangen. Een bijzonderheid in dit voorbeeld is dat een derde variabele, de leeftijd, exact kan worden vastgesteld; omdat de leeftijd een strict monotone relatie met X en Y heeft kan deze fungeren als een zogenaamde instrumentele variabele bij het schatten van de functionele relatie tussen lengte en gewicht (Kendall en Stuart, 1961).

Voorbeeld 6. De allometrische relatie.

Als men van volwassen individuen van verschillende soorten twee maten x en y beschouwt (lengte en gewicht zoals in het vorige voorbeeld, maar in het algemeen: één maat voor het geheel en één maat voor een deel, of maten van twee delen) stelt men vaak dat de soortgemiddelden X' en Y' samenhangen volgens de zogenaamde allometrische relatie:

$$Y' = \alpha X'^{\beta} \quad \text{of} \quad \log Y' = \log \alpha + \beta \log X'$$

Deze relatie geldt niet exact: de soortgemiddelden liggen er omheen gespreid, zonder dat men overigens kan aangeven hoe de afwijkingen d_{s*} ten opzichte van $X = \log X'$ en e_{s*} ten opzichte van $Y = \log Y'$ zich onderling verhouden: er is geen enkele reden een verdeling daarvoor te veronderstellen. Wel geldt dat waarnemingen y en x binnen elke soort op een in principe vast te stellen wijze gespreid zijn rond de voor die soort geldende waarden van X' en Y' .

Een overzicht van de in de voorbeelden gehanteerde modellen wordt gegeven in de volgende tabel. Alleen de structurele bijdragen d en e , die aangegeven zijn met de index s , resp. s^* , zijn onderling gecorreleerd. Verder geldt:

$$\begin{aligned} \text{var } d_s &= \sigma_{ds}^2, \quad \text{var } d_m = \sigma_{dm}^2, \quad \text{var } d = \sigma_{ds}^2 + \sigma_{dm}^2, \\ \text{var } e_s &= \sigma_{es}^2, \quad \text{var } e_m = \sigma_{em}^2, \quad \text{var } e = \sigma_{es}^2 + \sigma_{em}^2. \end{aligned}$$

voor- beeld	x		y	
	naam	model	naam	model
1	aantal lammeren	$x = X$	progesteron in bloed	$y = \alpha + \beta X + e_s + e_m$
2	vleespercentage	$x = X$	spekdikte	$y = \alpha + \beta X + e_s + e_m$
3	spermatozoën- concentratie	$x = X + d_m$	colorimeter- bepaling	$y = \alpha + \beta X + e_s + e_m$
4	dichtheid asfalt	$x = X + d_s + d_m$	nucleaire meting	$y = \alpha + \beta X + e_s + e_m$
5	gewicht	$x = X + d_s$	lengte	$y = \alpha + \beta X + e_s$
6	maat voor het geheel	$x = X + d_{s*} + d_s$	maat voor een deel	$y = \alpha + \beta X + e_{s*} + e_s$

In alle voorbeelden bevat y een structurele term als deel van de meetfout; de voorbeelden verschillen voornamelijk met betrekking tot de modelveronderstellingen voor x . Het al of niet opnemen van een structurele component d_s in het model voor x is afhankelijk van hoe de functionele relatie gedefinieerd wordt.

In het algemeen dient te worden opgemerkt dat men aan de verdeling van (x, y) -waarden, noch aan het al dan niet aanwezig zijn van stochastische componenten daarin, kan aflezen wat een relevant model voor een eventuele functionele relatie is. Als X normaal verdeeld is geldt in de voorbeelden 2 t/m 5 dat (x, y) bivariaat normaal verdeeld is. In feite kan elke bivariate normale verdeling worden gegenereerd via deze modellen door de keuze van de parameters, en ook omgekeerd kan men bij elke bivariate normale verdeling een of meer passende structurele relaties bedenken. Het hangt van de probleemstelling af of het zinvol is bij een bivariate verdeling van x - en y -waarden aan een structurele relatie te denken en hoe die dan zou moeten worden gedefinieerd.

3. SCHATTEN VAN DE RELEVANTE PARAMETERS

Uit alle gegeven voorbeelden komt duidelijk naar voren dat wat men als functionele of structurele relatie wil beschouwen, gedefinieerd moet worden en dat dit op een praktisch relevante wijze moet gebeuren. Het is niet voldoende te volstaan met de constatering dat zowel y als x niet foutvrij gemeten worden. De parameters van zorgvuldig gedefinieerde functionele en structurele relaties kunnen op basis van één waarnemingspaar (x, y) voor iedere (X, Y) in de steekproef, in principe worden geschat met behulp van de methoden die door Linssen (1980) worden beschreven, mits de verhouding tussen σ_d^2 en σ_e^2 bekend is. Hier zal niet op deze schattingsmethoden op zichzelf worden ingegaan; wel is het nuttig te constateren dat deze niet verschillen voor functionele en structurele relaties omdat men in het laatste geval de relatie het best kan schatten uitgaande van de realisaties van (X, Y) als gegeven. Er zij verder slechts vermeld dat de schattingsprocedures voor voorbeeld 1 en 2 resulteren in de regressie van y op X , en dat men in voorbeeld 5 door gebruik te maken van de leeftijd als instrumentele variabele met behulp van multivariate regressie op leeftijd de covariantiematrix van d en e kan schatten. Voorbeeld 6 tenslotte kan niet worden aangepakt met de theorie: een 'vrije hand' curve is even goed als elke andere voorgestelde curve, zolang de verdeling van de afwijkingen van Y en X ten opzichte van de allometrische relatie niet kan worden gespecificeerd. Beschrijvende technieken, zoals componentanalyse, zijn wellicht bruikbaar in zulke gevallen als objectieve methode, maar deze leveren geen schatting op van een goed gedefinieerde functionele relatie.

Het probleem bij de schatting is echter de juiste vaststelling van de verhouding tussen σ_d^2 en σ_e^2 . Zodra deze verhouding bekend is, is de theorie toepasbaar. Als deze verhouding foutief wordt gespecificeerd is het gevolg daarvan dat een niet bedoelde relatie wordt geschat, dus dat een systematische afwijking wordt geïntroduceerd. Als men in voorbeeld 2 de verhouding $\sigma_d^2/\sigma_e^2 \neq 0$ kiest, schat men niet de gezochte functionele relatie. Ditzelfde geldt in voorbeeld 3 en 4 als men σ_d^2/σ_e^2 baseert op de verschillen tussen waarnemingen verkregen aan dezelfde monsters en dus in feite daarvoor neemt $\sigma_{dm}^2/\sigma_{em}^2$: het structurele deel σ_{es}^2 van de variantie σ_e^2 wordt dan niet in rekening gebracht. Ook in voorbeeld 5 wordt

een verkeerde relatie geschat als men verzuimt de correlatie tussen lengte en gewicht in het model op te nemen.

Het is dus van groot belang dat de opgegeven specificatie met betrekking tot de verdeling van (e, d) juist is. Een onjuiste specificatie leidt tot systematische afwijkingen. Maar hoe komt men aan de juiste informatie? Het enige realistische antwoord moet zijn dat men de covariantiematrix van (e, d) moet meten. Dit is eenvoudig als geen structurele variabiliteit aanwezig is: men kan dan de nodige informatie putten uit herhaalde waarnemingen bij iedere waarde van (Y, X). Maar in het algemeen moet men ervan uitgaan dat zulke structurele variabiliteit wel aanwezig is, en daarnaar handelen (zie ook Johnston, 1963). Het is niet eenvoudig te bewijzen dat geen structurele variabiliteit aanwezig is. Daarom zou men zich tevreden kunnen stellen met de specificatie dat de met de structurele afwijkingen geassocieerde variantiecomponenten verwaarloosbaar klein zijn. Het is moeilijk aan te geven hoe groot 'verwaarloosbaar' klein precies is; uiteraard dient dit af te hangen van de gevoeligheid van de schatter voor β en de variantie daarvan voor afwijkingen van de juiste waarde van σ_d^2/σ_e^2 . Een indruk daarover kan men krijgen op basis van de door Patefield (1978) gegeven formules, maar dit wordt hier niet nader uitgewerkt.

De voorbeelden

Stel:

$$\begin{aligned} \text{var } y &= \sigma_Y^2 + \sigma_e^2 \\ \sigma_e^2 &= \sigma_{es}^2 + \sigma_{em}^2 \\ \sigma_{em}^2 &= \gamma_m(\sigma_Y^2 + \sigma_{es}^2) \\ \sigma_{es}^2 &= \gamma_s(\sigma_Y^2 + \sigma_{es}^2) \\ \text{var } x &= \sigma_X^2 + \sigma_d^2 \\ \sigma_d^2 &= \delta \sigma_X^2 \end{aligned}$$

De grootheden γ_m , γ_s en δ zijn niet afhankelijk van de gekozen schaal voor x en voor y en zijn daarom rechtstreeks vergelijkbaar.

In de voorbeelden 3 en 4 bleken de geschatte waarden van γ_m en van γ_s voor verschillende meetmethoden uit te komen op de volgende

combinaties van waarden; ter vergelijking is ook opgenomen de relatieve grootte δ van de meetfouten in x .

	voorbeeld 3		voorbeeld 4	
	methode 1	methode 2	instrument 1	instrument 2
γ_m	.02	.01	.12	.15
γ_s	.08	.11	.23	.39
δ	.02		.20	

Het is duidelijk dat in deze voorbeelden γ_s groot is ten opzichte van γ_m en δ . Bij verwaarlozing van de structurele afwijkingen zou de verhouding σ_d^2/σ_e^2 een factor 3 à 12 zijn misgeschat op basis van herhaalde waarnemingen. Op grond van deze bevindingen werd voor deze voorbeelden de oorspronkelijke aanname, dat het gebruikelijke model voor functionele of structurele relaties van toepassing zou zijn, verlaten. Voor de analyse van de gegevens van voorbeeld 3, waaraan bovenstaande resultaten zijn ontleend, zij verwezen naar Jansen (1980). Bij voorbeeld 4 was de proefopzet en de analyse vergelijkbaar met die in voorbeeld 3.

De voorbeelden tonen duidelijk dat het nodig is het gestelde model te verifiëren. Dit kan alleen gebeuren als voldoende herhalingselementen in de proefopzet zijn ingebouwd om dit mogelijk te maken (waardoor de schattingsprocedures voor enkelvoudige waarnemingen bij iedere (X, Y)-combinatie plaats moeten maken voor andere procedures). Er zijn een aantal mogelijkheden om de extra benodigde informatie te verkrijgen:

- Men kan metingen herhalen bij een aantal realisaties van X en van Y.
- Men kan aan dezelfde individuen metingen verrichten die van gelijke soort zijn als de metingen x en y , maar niet identiek.

Gelijksoortige metingen zijn per definitie metingen waarvoor op een of andere wijze aannemelijk gemaakt kan worden dat de meetfouten geen structurele component bevatten (Jöreskog, 1971). In voorbeeld 3 zijn als gelijksoortig beschouwd metingen bij verschillende verdunningen met elk van de methoden.

- Men kan instrumentele variabelen proberen te vinden; dat zijn variabelen die een strict monotone relatie met X hebben, en zonder fout gemeten kunnen worden. Een goed voorbeeld is de leeftijd in voorbeeld 5.

Als men per individu of object over n onafhankelijk herhalingswaarnemingen beschikt kan een bivariate variantie-analyse worden berekend van de volgende vorm (in de situatie dat x geen structurele afwijking bevat):

verwachtingswaarde van de gemiddelde gereduceerde kwadraten en produkten			
	EMS_y	EMP_{xy}	EMS_x
Objecten	$\sigma_{em}^2 + n\sigma_{es}^2 + n\beta^2\sigma_X^2$	$\rho_m \sigma_{dm} \sigma_{em} + n\beta\sigma_X^2$	$\sigma_{dm}^2 + n\sigma_X^2$
Metingen binnen objecten	σ_{em}^2	$\rho_m \sigma_{dm} \sigma_{em}$	σ_{dm}^2

Hierin is ρ_m de correlatie tussen de ware meetfouten in y en x , en σ_X^2 staat voor de variantie van X in de populatie, of voor gemiddelde kwadratische afwijking van de in het experiment gerealiseerde X -waarden. Door nu de berekende gemiddelde gereduceerde kwadraten en produkten gelijk te stellen aan de verwachtingswaarden daarvan kan men consistente schattingen voor alle parameters berekenen. Deze schattingen zijn niet zuiver (Acton, 1959, p. 138-170). De precisie van de schattingen als functie van het aantal objecten en het aantal metingen per object, kan in principe worden benaderd met behulp van de eigenschappen van centrale of niet-centrale Wishart-matrices; daarmee kan men desgewenst de efficiëntie van verschillende keuzen vergelijken.

Voor bijpassende schattingsmethoden voor de andere mogelijkheden, die veelal op maximum likelihood zijn gebaseerd, zij verwezen naar de literatuur (Cox (1968), Johnston (1963), Jöreskog (1970 en 1971), Kendall en Stuart (1961)). Of een model zonder structurele meetfouten acceptabel is zou men kunnen nagaan met likelihood ratio toetsen,

tenminste als men normale verdelingen kan veronderstellen. Voor het hier gegeven voorbeeld 3 is relevant in dit opzicht Jansen (1980) en Theobald en Mallinson (1978); het herhalingsselement bestaat in dit voorbeeld uit gelijksoortige metingen, en de daarvoor benodigde schattingsmethoden zijn gebaseerd op factoranalyse.

Als bij analyse blijkt, of reeds op voorhand duidelijk is, dat de meetfout in x geen structurele component bevat, en dat de variantie van het structurele deel van de meetfout in y groot is t.o.v. de echte meetfouten in x en y , zal de schatting van de structurele of functionele relatie vrijwel samenvallen met de gewone regressie van (aselect uit de verdeling van $y|X$ getrokken) waarden van y op x ; een verdere beperkende voorwaarde daarvoor is nog dat de meetfout in x klein moet zijn ten opzichte van het traject van X -waarden in de steekproef. Dit is het geval in voorbeeld 3, en in feite ook in voorbeeld 2 waar op voorhand duidelijk was dat de natuurlijk ook hier aanwezige meetfouten in x en y verwaarloosbaar klein zijn ten opzichte van de structurele bijdrage in de meetfout van y .

Ten aanzien van de proefopzet kan het volgende worden opgemerkt. Voor het verkrijgen van de grootste precisie kan men voor het schatten van een lineaire relatie het best zo extreem mogelijk waarden van X kiezen. Dit geldt in principe ook voor situaties waarin X niet direct kan worden vastgesteld, maar dan zal het buitengewoon moeilijk zijn dit zo in te richten dat gegarandeerd blijft dat y aselect getrokken is uit de verdeling van $y|X$. In de praktijk zal men een dergelijke garantie alleen hebben bij aselect trekken uit de verdeling van (x, y) . Truncatie op basis van de gemakkelijk te verkrijgen waarden van y is in ieder geval funest; truncatie op basis van x is voor praktische doeleinden misschien toelaatbaar als x geen structurele component bevat, maar nauwelijks uitvoerbaar als x de 'moeilijke' meting is.

4. WAARVOOR ZIJN DE VERKREGEN SCHATTINGEN BRUIKBAAR?

In voorbeeld 5 (de relatie tussen gewicht en lengte in een populatie groeiende individuen) is het duidelijk dat de te verkrijgen relatie dient ter beschrijving van de relatieve verandering van gemiddeld gewicht en gemiddelde lengte in de populatie. In de voorbeelden 2, 3 en 4 zal de verkregen relatie meestal bestemd zijn om te worden gebruikt voor het vaststellen van de waarde van X op basis van uitsluitend meten van y aan nieuwe individuen of objecten. Dit kan men doen, ongeacht of men met een structurele dan wel functionele relatie te doen heeft, door in de geschatte relatie:

$$Y = \hat{\alpha} + \hat{\beta} X$$

voor Y een nieuw gemeten waarde y in te vullen en vervolgens X daaruit op te lossen:

$$\hat{X} = \frac{y - \hat{\alpha}}{\hat{\beta}}$$

Invullen van het voor y gestelde model levert op

$$\hat{X} = \frac{\alpha - \hat{\alpha}}{\hat{\beta}} + \frac{\hat{\beta}}{\hat{\beta}} X + \frac{e}{\hat{\beta}}$$

De schattingen $\hat{\alpha}$ en $\hat{\beta}$ moeten voor zulk gebruik van de relatie worden opgevat als gegeven constanten, en dus bevat $\hat{X}$ een systematische fout ter grootte van

$$\frac{\alpha - \hat{\alpha}}{\hat{\beta}} + \frac{\hat{\beta} - \beta}{\hat{\beta}} X$$

De relatie is voor dit doel slechts bruikbaar als deze systematische fout verwaarloosbaar is, dat wil zeggen als men voor praktische doeleinden met voldoende benadering kan stellen dat $\alpha = \hat{\alpha}$ en $\beta = \hat{\beta}$. Dan kan men $\hat{X}$ beschouwen als een zuivere schatter voor X , met variantie

$$\text{var } \hat{X} = \sigma_e^2 / \beta^2$$

Omdat in alle genoemde voorbeelden σ_e^2 overwegend bestaat uit structurele variantie is de precisie van de zo verkregen vaststelling van X veel geringer dan men op grond van de variantie van herhaalde metingen van y bij hetzelfde individu of object zou verwachten. Afgezien van deze structurele variantie is de situatie geheel symmetrisch ten aanzien van X en Y : men kan dezelfde relatie ook gebruiken voor het vaststellen van Y op basis van een waarneming x . Tenslotte zij opgemerkt dat var $\hat{X}$, inclusief de bijdrage van de structurele afwijkingen, de relevante grootte is voor het vergelijken van de precisie van verschillende meetmethoden. In de situatie dat X niet stochastisch is, terwijl y een vector-variabele is, is de theorie ontwikkeld door Lieftinck-Koeijers, Van Oorschot en Sieben (1980) relevant, na eventuele aanpassing voor de situatie waarin X niet instelbaar is en meetfouten worden gemaakt bij het meten van X .

In het geval men te doen heeft met een structurele relatie kan men een betere schatter voor X bij een gegeven nieuwe waarde van y berekenen door gebruik te maken van de frequentieverdeling van de X -waarden in de populatie, $f_1(X)$; vgl. Lindley en Smith (1972). Men kiest als schatter voor X de waarde met de grootste aannemelijkheid, gegeven de waarde van y ; dus die X waarvoor het maximum van de zogenaamde a posteriori kansdichtheid $f(X|y)$ van X bereikt wordt:

$$f(X|y) = \frac{f_2(y|X) f_1(X)}{f_3(y)}$$

Hierin is $f_2(y|X)$ de kansdichtheid van y bij een gegeven waarde van X : deze kan worden berekend op basis van de schattingen voor de structurele relatie en σ_e^2 , mits deze voldoende precies zijn en mits de verdeling van (y, X) normaal is. Verder is $f_3(y)$ de marginale kansdichtheid van y , maar deze is voor een gegeven nieuwe y een constante, en dus niet relevant voor het vinden van het maximum van $f(X|y)$. Omdat de teller in het rechterlid gelijk is aan de kansdichtheid van de bivariate normale verdeling van y en X is onmiddellijk duidelijk dat het maximum daarvan voor gegeven y ligt ter plaatse van de regressie van X op y ; deze is tevens gelijk aan de regressie van x op y , omdat $E(x|X, y) = X|y$. Dus:

$$\hat{X}_0 = EX|y = Ex|y$$

Men kan ook schrijven, als σ_X^2 de variantie van X in de populatie voorstelt:

$$\hat{X}_0 = EX|y = \frac{\sigma_e^2}{\beta^2 \sigma_X^2 + \sigma_e^2} \cdot \mu_X + \frac{\beta^2 \sigma_X^2}{\beta^2 \sigma_X^2 + \sigma_e^2} \cdot \frac{y - \alpha}{\beta} = g_1 \mu_X + g_2 \cdot \frac{y - \alpha}{\beta}$$

Hieraan ziet men dat $\hat{X}_0$ zou kunnen worden opgevat als een gewogen gemiddelde van de a priori schatting $\hat{X}_1 = \mu_X$ en van eerder genoemde, alleen op de gegeven y gebaseerde, schatter $\hat{X}_2 = \frac{y - \alpha}{\beta}$. De gewichten zijn omgekeerd evenredig met de variantie van de a priori verdeling van X en van de variantie van $\hat{X}_2$: respectievelijk σ_X^2 en $\text{var } \hat{X}_2 = \sigma_e^2 / \beta^2$. De variantie van $\hat{X}_0$ kan men rechtstreeks berekenen op basis van de verdeling van $X|y$; men vindt dan:

$$\text{var } \hat{X}_0 = \frac{\sigma_e^2}{\beta^2} \cdot g_2 \quad \text{en dit is kleiner dan } \sigma_e^2 / \beta^2.$$

Blijkbaar geldt met betrekking tot de verdeling van $y|X$:

$$EX_0|X = g_1 \mu_X + g_2 X = X - g_1 (X - \mu_X)$$

$$\text{var}(\hat{X}_0|X) = g_2^2 \sigma_e^2 / \beta^2$$

zodat $\hat{X}_0$ geen zuivere schatter is voor de aktuele waarde van X die de gevonden waarde van y heeft geproduceerd. Verder geldt dan

$$\text{MSE}(\hat{X}_0|X) = g_1^2 (X - \mu_X)^2 + g_2^2 \sigma_e^2 / \beta^2$$

en deze is gemiddeld over de verdeling van X gelijk aan de rechtstreeks berekende waarde van $\text{var } \hat{X}_0$. Tot waarden van X die minder dan ongeveer $\sigma_X \sqrt{2}$ afwijken van μ_X is $\text{MSE}(\hat{X}_0|X) < \text{var}(\hat{X}_2|X)$. Blijkbaar is $\hat{X}_0$ in deze zin niet uniform beter dan $\hat{X}_2$.

Bij de afleiding van deze schatter is gebruik gemaakt van de aanname dat f_2 en f_1 exact of tenminste voor praktische doeleinden voldoende precies bekend zijn. Als alleen f_1 precies bekend is kan men samen met een benadering voor f_2 de aannemelijkheidsfunctie f berekenen en op basis daarvan $EX|y$ schatten. Als f_1 niet bekend is en op basis van de steekproef moet worden geschat lijkt het voor de hand te liggen direct de regressie van x op y te berekenen als predictierelatie. Dit is in de

praktijk ook de gebruikelijke procedure. Een essentiële voorwaarde voor de geldigheid hiervan is dat men in de steekproef een aselechte trekking heeft uit de populatie van (x, y) -waarden. In de literatuur stuit men echter op bezwaren tegen de hier gesuggereerde procedure; men zie bijvoorbeeld de discussie over 'empirical Bayes methods' in Barnett (1973). Tenslotte zij opgemerkt dat in situaties waarin een ijking tot stand komt via een aantal tussenstappen men slechts in één van die stappen de informatie over $f_1(X)$ kan inbrengen, dan wel een predictierelatie kan schatten; bij alle tussenstappen moet men de structurele relaties zelf hanteren.

Als men over de verdeling van X niets concreets wil veronderstellen is wellicht de oplossing gegeven door Lwin en Maritz (1980) relevant: deze auteurs laten als schatter voor X op basis van nieuwe metingen y slechts lineaire combinaties van de gerealiseerde x -waarden in de eerste steekproef toe.

De voorbeelden

In voorbeeld 2 wordt de functionele of structurele relatie gegeven door de regressie van spekdikte op vleespercentage, maar de predictierelatie door de regressie van vleespercentage op spekdikte. Een dergelijke dualiteit geldt algemeen bij bivariate normale verdelingen, als men één van de regressies als een structurele relatie kan opvatten.

Voor het schatten van de functionele relatie is het niet relevant of X stochastisch is of niet: het onderscheid tussen functionele en structurele relaties is hiervoor van geen enkel nut. Voor het schatten van de predictierelatie is er wel een fundamenteel verschil tussen stochastische en niet-stochastische X . Zo is het in voorbeeld 4 niet zonder meer evident dat voor de X -waarden een verdeling gespecificeerd kan worden. In dat geval zou men dus ook voor predictiedoeleinden gebruik moeten maken van de functionele relatie (of eventueel de regressie van y op x als benadering daarvoor).

De geconstateerde dualiteit geldt in principe ook in situaties waarin de verdelingen niet normaal zijn. Uiteraard dienen de schattingsprocedures te worden aangepast en zullen de predictierelaties geen regressies zijn in het algemeen. Zo wordt in voorbeeld 1 de structurele relatie gegeven door de regressie van y op X , maar de predictierelatie

wordt gegeven door aan een nieuwe waarneming y de X -waarde toe te voegen die hoort bij de grootste van de zogenaamde discriminantscores:

$$\{\pi_1 f(y|X=1), \pi_2 f(y|X=2)\}$$

waarin f de conditionele kansdichtheid van $y|X$ voorstelt.

Men schat het aantal lammeren dus gelijk aan 1 als

$$\pi_1 f(y|X=1) > \pi_2 f(y|X=2)$$

dus als

$$y < \frac{\mu_1 + \mu_2}{2} - \frac{\sigma_e^2}{\mu_2 - \mu_1} \ln \frac{\pi_2}{\pi_1}.$$

en gelijk aan 2 als y groter is dan de uitdrukking in het rechterlid. Als men het teken $<$ vervangt door het gelijktteken krijgt men de zogenaamde lineaire discriminerende functie, die in dit geval met maar één y -variabele een eenvoudige vorm aanneemt ($y = \text{constante}$).

De gevolgde redenering blijft in dit voorbeeld, en ook in voorbeeld 2, volledig geldig als y een vectorvariabele voorstelt.

Een ander aspect betreffende de bruikbaarheid van de functionele relaties en de predictierelaties is hun stabiliteit. Relaties zijn stabiel wanneer ze onder verschillende omstandigheden hetzelfde blijven. In de praktijk blijkt dit veelal niet het geval, waardoor de bruikbaarheid van eenmaal vastgestelde relaties vaak gering is. In voorbeeld 1 bleek dat de verdeling van de hoeveelheden progesteron in verschillende jaren belangrijk verschilde, door een niet bekende oorzaak, waardoor de relatie niet praktisch bruikbaar was. In voorbeeld 2 bleken zowel de functionele als de predictierelaties te verschillen voor varkenskarkassen uit de verschillende EEG-landen, wat een belangrijke hinderpaal is bij het vaststellen van een uniform EEG-classificatiesysteem. In voorbeeld 4 bleken de resultaten van de nucleaire metingen gevoelig voor de samenstelling van het asfaltmengsel; dit beperkt het gebruik van de nucleaire meetapparaten vooralsnog tot vergelijkende metingen binnen eenzelfde wegconstructie, omdat de apparaten in feite voor elke constructie opnieuw zouden moeten worden geijkt. Voorbeeld 6 is een typisch voorbeeld van een niet stabiele relatie; het daar gehanteerde model kan

ook in andere situaties worden gebruikt voor het beschrijven van eventuele instabiliteit.

5. CONCLUSIES

De begrippen 'functionele en structurele relatie' worden in de statistische literatuur veelal uitsluitend gehanteerd in gevallen waarin de relaties behoudens echte meetfouten kunnen worden waargenomen. Het blijkt vruchtbaar de begrippen uit te breiden voor situaties waarin naast afwijkingen door echte meetfouten ook structurele afwijkingen van de relaties kunnen voorkomen. Structurele afwijkingen zijn afwijkingen die eigen zijn aan het waargenomen individu of object, en die zich bij herhaalde meting aan dat individu of object identiek realiseren. Het ligt voor de hand, althans bij een normale verdeling van deze eigen bijdragen, de relatie te definiëren op basis van de verwachtingswaarde over die eigen bijdragen. Door de proefopzet zo te kiezen dat alle parameters van dergelijke relaties kunnen worden geschat, kan men nagaan of structurele afwijkingen praktisch belangrijk zijn. In de besproken praktische voorbeelden was dit het geval; gesuggereerd wordt dat men nooit een model kan hanteren waarin alleen echte meetfouten worden verondersteld, zonder zich van de geldigheid daarvan empirisch te overtuigen.

Ten aanzien van het schatten is er geen verschil tussen functionele en structurele relaties. Men kan beide ook gebruiken voor predictiedoeleinden: uit een nieuwe waarneming y kan men de bijbehorende waarde van X schatten, en omgekeerd, uit een nieuwe waarneming x kan men met dezelfde relatie Y schatten. Deze schatters bevatten bij herhaald gebruik van de relatie steeds dezelfde systematische afwijking als gevolg van de onnauwkeurigheid in de schatting van de relatie; het is dus van praktisch belang de relatie alleen te gebruiken als deze voor het doel voldoende nauwkeurig is vastgesteld.

In het geval van een structurele relatie dient men informatie over de verdeling van X , hetzij in de vorm van een gegeven a priori verdeling, hetzij de in de steekproef zelf aanwezige informatie, te benutten voor predictiedoeleinden. Het resultaat daarvan is dat men voor predictie een geheel andere relatie moet hanteren dan de structurele relatie.

In het geval van een bivariate normale verdeling van x en y is de predictierelatie voor X op basis van y gelijk aan de regressie van x op y , en omgekeerd is de predictierelatie voor Y op basis van x gelijk aan de regressie van y op x , ongeacht de modelstructuur. Voor het schatten van de predictierelatie is het noodzakelijk dat de steekproef van (x, y) -waarden aselekt getrokken wordt uit de bivariate verdeling van x en y . Dit zal in de praktijk ook nodig zijn voor het schatten van de structurele relatie, om te kunnen garanderen dat de waarden van y bij een gegeven waarde van X aselekt getrokken zijn uit de relevante conditionele verdeling.

Een belangrijke praktische beperking van het gebruik zowel van de functionele als van de predictierelaties kan zijn dat deze vaak niet stabiel zijn in de tijd of voor verschillende situaties.

6. LITERATUUR

- Acton, F.S. (1959). Analysis of straight-line data. New York: Wiley.
- Barnett, V. (1973). Comparative statistical inference. London: Wiley.
- Cox, D.R. (1968). Regression Methods. Journal of the Royal Statistical Society A 131, 265-279.
- Ehrenberg, A.S.C. (1975). Data reduction: Analysing and interpreting statistical data. London: Wiley.
- Jansen, A.A.M. (1980). Comparative calibration and congeneric measurements. Biometrics 36, 729-734.
- Johnston, J. (1963). Econometric methods. New York: McGraw-Hill (p.148-162)
- Jöreskog, K.G. (1970). A general method for analysis of covariance structures. Biometrika 57, 239-251.
- Jöreskog, K.G. (1971). Statistical analysis of sets of congeneric
- Kendall, M.G. and A. Stuart (1961). The advanced theory of statistics. Vol. 2. London: Griffin.
- Lieftinck-Koeijers, C.A.J., H.J.L. van Oorschot en J.W. Sieben (1980). Meest aannemelijke schatters in het multiresponsie ijkingsprobleem. Delft: Rapport THD-A-ST 80/06, Technische Hogeschool.

- Lindley, D.V. and A.F.M. Smith (1972). Bayes estimates for the linear model (with discussion). Journal of the Royal Statistical Society, B34, 1-41.
- Linssen, H.N. (1980). Functional relationships and minimum sum estimation. Eindhoven: Thesis Technische Hogeschool.
- Lwin, T. and J.S. Maritz (1980). A note on the problem of statistical calibration. Applied Statistics 29, 135-141.
- Theobald, C.M. and J.R. Mallinson (1978). Comparative calibration, linear structural relationships and congeneric measurements. Biometrics 35, 39-45.