KM 36(1991)
pag 25 - 46

MODELBOUW EN STATISTIEK

rede

uitgesproken tijdens het symposium
ter gelegenheid van het afscheid van
professor dr. G. J. Leppink
van de Rijksuniversiteit te Utrecht

door

P. C. Sander

Faculteit Technische Bedrijfskunde
Technische Universiteit Eindhoven
Postbus 513
5600 MB Eindhoven
tel. 040 - 47 36 55

1. INLEIDING

"Zeer gewaardeerde Toeschouwers, Statistiek is voor velen een geheimzinnig vak."

Zo begon op 24 februari 1964 professor Leppink zijn rede bij de aanvaarding van het ambt van gewoon hoogleraar in de mathematische statistiek. De rede had tot titel: MODELLEN. In die rede gaf professor Leppink, aan de hand van een aantal voorbeelden, aan, dat aan het toepassen van statistische technieken altijd een modelbouwfase voorafgaat. In die fase wordt een *model* opgesteld dat

a. goed genoeg aansluit bij de realiteit, en
b. statistische conclusies mogelijk maakt.

Dat is eenvoudig gezegd, maar hier zit natuurlijk een probleem, namelijk: wat is goed genoeg? Wat goed genoeg is hangt af van het doel en van de situatie. Ik zal een voorbeeld geven uit de inaugurele rede van professor Leppink. Stel u voor, wij zitten in de werkkamer van een statisticus en luisteren naar het gesprek dat deze voert met één van zijn cliënten, een bioloog in dit geval. Om een en ander zo goed mogelijk tot zijn recht te laten komen zal ik een stukje citeren. We beginnen met de bioloog. Ik lees:

> " Deze heeft bij een aantal bonen van een bepaald ras de lengte en het gewicht gemeten. Hij heeft zijn waarnemingsuitkomsten op aanschouwelijke wijze voorgesteld door het tekenen van een puntenwolk: elke boon is weergegeven door een punt, waarvan de horizontale coördinaat de lengte van de boon voorstelt en de verticale coördinaat het gewicht. Het verzoek aan de statisticus is, of deze de beste lijn door de punten wil uitrekenen. De statisticus vraagt, wat voor soort lijn door de punten moet worden getrokken. Verbaasd antwoordt de bioloog, dat hij nu juist is gekomen om dit van de statisticus te vernemen. Hierop pakt deze laatste een potlood en trekt met de hand een vloeiende lijn door de punten. Op de opmerking van de bioloog, dat hij dit zelf ook wel had kunnen doen, repliceert de statisticus, dat hij een groot aantal wiskundige functies kan bedenken, van welke de grafiek een goede aanpassing met de puntenzwerm oplevert. Deze functies behoeven echter geen enkele biologische betekenis te hebben en zijn derhalve onbruikbaar voor een zinvolle biologische interpretatie van de waarnemingsuitkomsten. Slechts als op biologische gronden mag worden aangenomen, dat de lijn behoort tot een bepaalde klasse functies, wat er meestal op neerkomt, dat de formule van de lijn op een paar parameters na bekend is, dan kunnen deze parameters worden berekend. Anders is de met de hand getrokken lijn even goed als een volgens de een of andere formule berekende kromme; het loont dan de moeite niet om een

berekening uit te voeren. Wat hier ontbreekt is het wiskundig model."
Einde citaat.

Dames en heren, professor Leppink was zijn tijd ver vooruit: het was in de 60-er jaren bepaald geen gebruik om de aandacht te vestigen op de constructie van modellen. Dit ondanks het feit dat Fisher (1922) al in de twintiger jaren had opgemerkt, dat de eerste stap bij het toepassen van statistiek is, de constructie van een laag-dimensionaal parametrisch model.

Zeker voor de handboeken voor toepassers geldt, dat deze in de 60-er jaren slechts bestonden uit een opsomming van een aantal modellen. De toepasser koos uit die collectie een recept, voerde de aangegeven rekenpartij uit, en trok op de voorgeschreven wijze een conclusie. Helaas moet ik vaststellen dat de leerboeken in de statistiek nog steeds weinig aandacht schenken aan het bouwen van geschikte modellen. De meeste boeken komen nog steeds niet verder dan het leren *herkennen* van de belangrijkste statistische modellen. Gezien de rol van de statisticus is dit een uiterst ongewenste situatie, want in de praktjk is herkennen niet genoeg. Juist de statisticus hoort zich met de modelbouw te bemoeien, want juist de statisticus heeft tot taak om de praktische bruikbaarheid van modellen te beoordelen in het licht van de aanwezige of nog te verwerven waarnemingen. Een aardig voorbeeld van de rol van de statistische facetten tijdens de modelbouwfase, is te vinden in een zeer lezenswaardig artikel van Douglas Miller (1986). Miller geeft in dat artikel voorbeelden van op het eerste oog heel aardige modellen voor software reliability; modellen die verschillende voorspellingen geven voor het aantal bugs in een software programma, maar die met behulp van de waar te nemen gebeurtenissen niet zijn te onderscheiden.

Tussen de waarnemingen en de conclusies zit nooit één model, maar altijd een *klasse* van modellen. Eerst moet worden bepaald welke die klasse is, vervolgens moet uit die klasse een geschikt model worden gekozen (bijvoorbeeld via het schatten van de modelparameters). Formeel zou je kunnen zeggen dat het de taak is van de betreffende onderzoeker om de *klasse* van modellen aan te geven, terwijl de statisticus uit die klasse het *beste* model moet distilleren. De praktijk is echter, dat de statisticus onmisbaar is bij het vertalen van het verbaal geformuleerde probleem in een mathematisch model. Bij zijn bijdrage aan het bouwen van het model, moet de statisticus de volgende twee punten goed in het oog houden:

1. hij moet zorgen dat de onderzoeker het model begrijpt, zodat de onderzoeker kan beoordelen of het model inderdaad zijn probleem weergeeft; en
2. hij moet zorgen dat het model geschikt is voor verdere verwerking, de in

het model opgenomen parameters moeten bijvoorbeeld voldoende nauwkeurig kunnen worden geschat.

Aansluitend bij het bestaande taalgebruik zal in het vervolg vaak worden gesproken over een model (enkelvoud), als een klasse van modellen wordt bedoeld.

Het zal duidelijk zijn dat een goede wisselwerking tussen onderzoeker en statisticus essentieel is voor een zinvol resultaat. De onderzoeker moet enige kennis hebben van de statistiek, en de statisticus moet tenminste een globale kennis hebben van het terrein van de onderzoeker. De combinatie van de vakkennis van beide moet de basis zijn voor een goed model, maar een garantie wordt niet gegeven. Zoals Freedman vorige maand tijdens het 2e Wereldcongres in Uppsala vaststelde: gewoonlijk zijn de veronderstellingen in een statistisch model moeilijk te verifiëren en wordt daar ook weinig aandacht aan geschonken. De kracht van de uitspraken over de empirie volgt dus niet uit de soliditeit van de veronderstellingen.

> *...most statisticians do not seem to become involved* **deeply enough** *in subject matter areas to understand the scientific problems in their contexts.*
>
> *R. Gnanadesikan*

In feite heeft de beschikbaarheid van statistische softwarepakketten op ieders PC de onverwijdelijke konsekwentie, dat *iedere* gebruiker van statistiek, en niet alleen de statisticus, getraind moet worden in modelbouw. Gebeurt dit niet, dan is het onvermijdelijk dat problemen worden gewrongen in standaardmodellen i.p.v. dat de modellen worden aangepast aan de problemen.

> *... unwillingness of many statisticians to rethink the conventional statistical formulations whenever they are inappropriate in a new formulation.*
>
> *J.W. Tukey*

Dames en heren, het begrip 'model' is al heel oud. Lehmann (1990) geeft een aardig voorbeeld (stammend uit de tijd van de slag bij Nieuwpoort) van de erkenning van het nut van modellen. Het betreft Galileo en de Rooms Katholieke kerk. Galileo mocht van kardinaal Bellarmino best de aarde om de zon laten draaien, schrijft Lehmann, als Galileo dit maar bracht als een wiskundig model, slechts bedoeld om de berekeningen eenvoudiger te maken. De kardinaal had er alleen bezwaar tegen dat Galileo beweerde dat de aarde ook *echt* om de zon draaide.

Ondanks het feit dat wiskundige modellen al heel oud zijn, is de aandacht voor modelbouw als onderzoekobject pas de laatste jaren in allerlei publicaties terug te vinden, en ook, opnieuw, in intreeredes. Zo luidt de titel van de rede die collega Van Harten op 11 mei vorig jaar uitsprak aan de Universiteit Twente: MODELLEN PASSEN BIJ HET MANAGEMENT. En net als professor Leppink, besteedt Van Harten zowel aandacht aan het herkennen van situaties waarin bekende modellen met succes kunnen worden toegepast, als aan het ontwikkelen van adequate modellen. Beide facetten vang ik in het ene woord: MODELBOUW.

Wie wil lezen wat Erich Lehmann (1990) en Sir David Cox (1990) vandaag de dag schrijven over de modelbouwfase, moet het mei-nummer van Statistical Science van dit jaar daar maar eens op naslaan. Het betreft heel lezenswaardige artikelen.

Al met al stel ik vast: professor Leppink is niet langer de uitzondering; het belang van modelbouw wordt inmiddels breed ingezien. Ook het onderwijs haakt hier op in. Ik geef twee voorbeelden uit mijn eigen universiteit:
1. de Faculteit Bedrijfskunde verzorgt al in het eerste jaar van de opleiding een college Modelbouw van 9 maal 2 uur;
2. de Faculteit Wiskunde en Informatica heeft een modellen-colloquium als essentieel onderdeel van de na-doctoraalopleiding Wiskunde voor de Industrie.

Dames en heren, tot zover de inleiding. In de resterende tijd wil ik twee dingen doen. Allereerst wil ik ingaan op het begrip MODEL:
- wat is een model?
- zijn er verschillende typen modellen?
- is het zinvol je bij het bouwen van modellen te realiseren dat er verschillende typen zijn?

Daarna wil ik bij wijze van voorbeeld aan een bekende dataset uit de literatuur demonstreren dat het spelen met modellen (en wat is modelbouw anders?) een heel boeiende bezigheid is.

2. MODELBOUW

2.1. Wat is een model?

Als ik het over een model heb, dan bedoel ik een denkconstructie met of zonder realiteitsgehalte, die kan dienen om inzicht te krijgen en/of om te kunnen optimaliseren.

Daar de problemen waar een statisticus mee wordt geconfronteerd empirisch van aard zijn, bedoel ik met modellen: abstracte weergaven van concrete verschijnselen. Dat is dus anders dan voor een wiskundige, voor een wiskundige is een model juist een concrete interpretatie van een abstract systeem.

Voorbeeld. Een wiskundige kan de Booleaanse algebra toelichten via een model bestaande uit electronische schakelingen. Het model dient dan eerder als demonstratiemodel dan als model om de theorie verder te helpen.

In de empirische wetenschappen in een model een vereenvoudiging van de realiteit. De ene keer is het model in hoofdzaak bedoeld als hulpmiddel om het inzicht in de complexe realiteit te vergroten, de andere keer is het model bedoeld als hulpmiddel om verstandige beslissingen te kunnen nemen. M.a.w. sommige modellen zijn voor theoretische beschouwingen interessant, andere modellen zijn vanuit praktisch standpunt zinvol: zij beïnvloeden de kwaliteit van beslissingen. In beide gevallen geldt, dat modellen functioneren binnen een theoretische context, zij vormen een schakel tussen empirie en theorie.

In de praktijk bestaat de modelbouwfase altijd uit het bouwen van meerdere modellen voor dezelfde situatie. Elk model geeft slechts een partiële interpretatie vanuit een bepaald gezichtspunt. De verschillende modellen kunnen verschillende doeleinden dienen. Of een model goed of fout is, hangt dus (mede) af van de aspecten waarin men is geïnteresseerd. Alle modellen hebben echter gemeen dat er een structuur is aangebracht, waarvan de bruikbaarheid moet worden aangetoond.

Voorbeeld. Op een landkaart een stad aangeven met een cirkeltje, is prima als de kaart slechts dient om de stad te localiseren. Het cirkeltje is minder geslaagd, als de kaart ook moet dienen om in de stad de weg te kunnen vinden.

Modellen zijn, naar mijn mening, uiterst geschikt voor het verkrijgen van inzicht in nog betrekkelijk onbekende onderzoeksgebieden. Juist daar kunnen eenvoudige modellen helpen om inzicht in de problematiek te krijgen. Veel

moeilijker is het om complexe operationele situaties zo te modelleren dat optimalisatie mogelijk is.

Voorbeeld. De produktiebesturing van een niet–triviaal produktieproces is veel te gecompliseerd om met enige kans op succes volledig modelmatig tot een optimale besturing te kunnen komen.

2.2. Klassificeren

Een van de eerste stappen op weg naar een theorie is gewoonlijk klassificeren en ordenen. Alleen al hieruit zou de conclusie kunnen worden getrokken dat er nog nauwelijks kan worden gesproken over een modelbouw*theorie*. Immers, er is wel enige literatuur over klassificeren van typen modellen, maar het is volstrekt duidelijk dat de belangstelling nog maar pril is, en dat de gedachten nog niet goed zijn uitgekristalliseerd. Lehmann en Cox gaan beide, in het al eerder genoemde nummer van Statistical Science, in op klassificatie van modellen, en het tekent de situatie, dat hun literatuurlijsten niet één gemeenschappelijke verwijzing bevatten. Ze komen overigens wel globaal tot dezelfde indeling.

Voor we gaan klassificeren, dienen we ons natuurlijk eerst af te vragen wat daarvan het doel is. Naar mijn mening is het doel: modellen en problemen koppelen. Als een bepaald model nuttig is gebleken bij de aanpak van een bepaald probleem, dan kunnen we hopen dat een model uit dezelfde klasse ook nuttig is voor een soortgelijk probleem.

Wat mij verbaast, is dat Lehmann en Cox beide, aansluitend op de bestaande literatuur, alleen ingaan op het klassificeren van *modellen*. Daar een model pas interessant is door de relatie met een probleemstelling, lijkt mij dat je moet beginnen met het klassificeren van *problemen*. Dit sluit ook aan bij de wijze waarop in de praktijk problemen worden aanpakt. Je zorgt dat je eerst het probleem goed kent, daarna ga je nadenken over geschikte modellen. Alléén modellen klassificeren sluit overigens goed aan bij de leerboeken over statistiek, want ook die hanteren altijd een indeling in hoofdstukken, die is gebaseerd op modellen en niet op problemen.

Problemen zijn op diverse manieren te klassificeren. Bij het al eerder genoemde college Modelbouw in Eindhoven wordt de volgende indeling van problemen gehanteerd, gebaseerd op contrasten:

1. *Deterministisch versus stochastisch.* Het onderscheid spreek voor zich.

2. *Statisch versus dynamisch.* Bij een dynamisch probleem treedt volgtijdelijkheid op, bij een statisch probleem niet. Een dynamisch systeem is meer dan een aantal situaties na elkaar, er is ook een zekere afhankelijkheid tussen situaties.

3. *Verklarend versus besturend.* Het onderscheid wordt bepaald door de positie van de onderzoeker. Wil de onderzoeker het gedrag van een systeem bestuderen/beschrijven zonder in het systeem in te grijpen, dan spreken we van een verklaringsprobleem. Van een besturingsprobleem is sprake als de onderzoeker het systeem op grond van bepaalde normen wil beïnvloeden.

Laten we nu eens kijken wat in de literatuur wordt opgemerkt over het klassificeren van *modellen.*

Ik begin met een pessimistische uitspraak. Neyman zegt (zie Lehmann (1990)): om een realistische verklarende theorie te kunnen ontwikkelen is een diepgaande kennis nodig van de achtergrond van het probleem. Het pessimistische is, dat dit suggereert dat elk model van de grond af moet worden opgebouwd, zodat van een theorie niet veel kan komen.

Voorbeeld. Een aardig voorbeeld van modelbouw zonder een voldoende diepgaande kennis van de achtergrond van het probleem is te vinden in Freudenthal (1966). Het betreft de wijze waarop de Schot John Arbuthnot, de schepper van de figuur 'John Bull', het bestaan Gods bewees. De redenering was de volgende.
Ik citeer:

> "Volgens statistieken zijn in Londen in 82 jaren achter elkaar meer jongens dan meisjes geboren. Indien de geboorte van jongen of meisje een kanskwestie was zoals 'kruis en munt', dan was dat onbegrijpelijk. Want een serie van 82 keer achter elkaar 'kruis' bezit een waarschijnlijkheid van $(\frac{1}{2})^{82} = 2 \cdot 10^{-26}$. Van iets, wat een zo geringe waarschijnlijkheid bezit, mag men gevoeglijk zeggen, dat het onmogelijk is. En hieruit volgt de onmogelijkheid van de veronderstelling, dat het geslacht over het menselijke nakroost wordt verdeeld volgens het toeval van kruis of munt. Dus is er een goddelijke voorzienigheid, die ingrijpt om de verhouding der geslachten te regelen."

Nu, 280 jaar later, hoef ik niet meer uit te leggen wat er aan het model schort.

Positiever dan de vorige opmerking van Neyman, is het volgende. Neyman stelt dat er twee typen modellen zijn:

a. *empirische modellen*: kies uit een handige, flexibele, verzameling a priori gegeven modellen, een model dat het beste past bij de waarnemingen.

Voorbeeld. Kies uit de klasse van Weibullverdelingen degene die het beste past bij de waarnemingen. Of, kies uit een klasse regressiemodellen een model met een zinvol aantal parameters.

De eerste stap is natuurlijk de specificatie van de handige, flexibele verzameling modellen waaruit moet worden geselecteerd. Waarom zou je bijvoorbeeld regressiemodellen kiezen, en niet ARMA–modellen? Welke klasse modellen kies je als je bent geïnteresseerd in het faalgedrag van een bepaald type dieselmotor?

b. *verklarende modellen*: deze modellen zijn bedoeld om een verklaring te geven voor het mechanisme dat achter de waargenomen fenomenen verscholen gaat (Denk bijvoorbeeld aan Darwin's erfelijkheidstheorie).

Het essentiële verschil tussen empirische en verklarende modellen komt ruwweg op het volgende neer:
– empirische modellen zijn bedoeld voor aktie, zij zijn specifiek voor een afgebakend probleem,
– verklarende modellen zijn bedoeld om inzicht te krijgen, zij worden liefst zo algemeen mogelijk geformuleerd.

Voorbeeld. Als een bierfabrikant tijdig wil weten hoeveel bier of hij in de zomer van 1991 kan verkopen, dan heeft hij behoefte aan een voorspelmodel dat voldoende accuraat voorspelt. Of in dat model het koopgedrag van de klant op een sociaal–wetenschappelijk verantwoorde wijze wordt verwerkt, is niet relevant.

Voorbeeld. Als bij een medisch probleem een discriminant–analyse aantoont dat 3 * de leeftijd + .7 * de systolische bloeddruk de relevante patiëntengroepen uitstekent scheidt, dan geeft dit een uiterst bruikbaar model.

Voorbeeld. Fysische wetten vallen onder de verklarende modellen.

Het checken van de validiteit gaat bij de empirische modellen vanzelfsprekend geheel anders dan bij verklarende modellen. Voor empirische modellen behoeft slechts te worden nagegaan of het model past bij de waarnemingen.

Voorbeeld. Hebben we te maken met een Weibullverdeling? Een aanpassingstoets kan zinvol zijn.

Voorbeeld. Op grond van de waarnemingen wordt besloten niet het model y=ax + b te hanteren, maar het model y=a logx + b. .

In empirische modellen worden alleen die aspecten opgenomen, die van praktisch belang zijn. Gewoonlijk passen verschillende modellen heel redelijk. Dan is het verstandig om een model te kiezen dat goed hanteerbaar is.

Een verklarend model moet aan veel strengere eisen voldoen dan een empirisch model. Een verklarend model moet theorie-vormend werken, het moet leiden tot nieuwe, toetsbare hypothesen.

Voorbeeld. De relativiteitstheorie. Experimenten bevestigden de theorie: licht wordt door massa afgebogen.

Het zal duidelijk zijn dat empirische modellen passen bij besturingsproblemen, en verklarende modellen bij verklaringsproblemen. Het zal ook duidelijk zijn, dat een bepaald type model in de ene situatie een verklarend model levert, en in de andere situatie een empirisch model.

Cox(1990) voegt aan de empirische en de verklarende modellen nog toe de zg. *indirecte modellen.* Dat zijn situaties, waarin gegevens/waarnemingen met een bekende structuur worden gebruikt om een nieuw voorgestelde analysemethoden te bestuderen; dit in afwijking van de gebruikelijker situatie, waarin bekende methoden worden gebruikt om gegevens met een onbekende structuur te analyseren.

De gepresenteerde klassificaties van problemen en modellen leidt tot de conclusie, dat de klassen voor praktisch gebruik onhandelbaar groot zijn. Tijdens de modelbouwfase, en bij de keuze van de analysetechniek, zijn de gegeven klassificatie van geringe betekenis. Het valt kennelijk ook niet mee om een indeling te maken die bij elk probleem automatisch een geschikt model genereert. Als dit eenvoudig was, dan zou elk statistisch pakket een expertsysteem bevatten. Mij is een goed expertsysteem voor de gebruiker van statistiek echter niet bekend. Juist in 1990 is het toepasselijk om hoop te ontlenen aan de volgende uitspraak:

Zulke dingen zijn echter moeilijk en zullen zo ineens niet lukken. Het gelukken is soms 't resultaat van een hele serie mislukkingen.

Vincent van Gogh

Het deel over modellen zal ik nu afronden met enkele algemene opmerkingen.

1. Het is gewenst dat modellen kunnen worden getoetst.
2. Een gevoeligheidsanalyse is essentieel, in het bijzonder voor niet–toetsbare veronderstellingen.
3. De randvoorwaarden waarbinnen modellen gelden, moeten worden aangegeven.In het bijzonder als voor de praktijk een uiterst simpel model is gewenst (een soort vuistregel), dan moet goed bekend zijn wat de grenzen zijn van de bruikbaarheid.
4. Modelbouw is slechts te leren m.b.v. echte cases (en daarbij helpt natuurlijk dat het rekenwerk tegenwoordig geen probleem meer is). Die cases dienen in het onderwijs te worden gepresenteerd door degene(n) die de cases van dichtbij hebben meegemaakt.
5. Of een statistische analyse waardevol is, wordt meer bepaald door de zinvolheid van de gebruikte modellen, dan door de geavanceerdheid van de gekozen rekenpartijen. Binnen de statistiek loert altijd het gevaar, dat een complexe analyse is gebaseerd op onduidelijke, om niet te zeggen onjuiste veronderstellingen. Het probleem is, dat het in het onderwijs verleidelijk is deze opmerking in het eerste college statistiek één keer te maken en vervolgens de Banach–ruimte in te stappen.

3. EEN MODEL VOOR HET FAALGEDRAG VAN REPAREERBARE SYSTEMEN

Zoals de heren Ascher en Feingold (1984) in hun boek Repairable Systems Reliability met veel nadruk betogen, is er nog nauwelijks een begin gemaakt met de bestudering van repareerbare systemen. Dit is verrassend, omdat vrijwel alle technisch systemen, en zeker de kostbare, repareerbaar zijn.

Tot op de dag van vandaag beginnen de meeste verhalen over repareerbare systemen met de veronderstelling, dat het systeem na reparatie letterlijk weer als nieuw is. Nu wil ik niet beweren dat aan deze veronderstelling nooit is voldaan, i.h.b. voor electronische systemen is het nog wel eens een redelijke veronderstelling, maar ik beweer wel dat de hoofdreden voor de stroom artikelen over zulke systemen is, dat het makkelijk is om artikelen over zulke

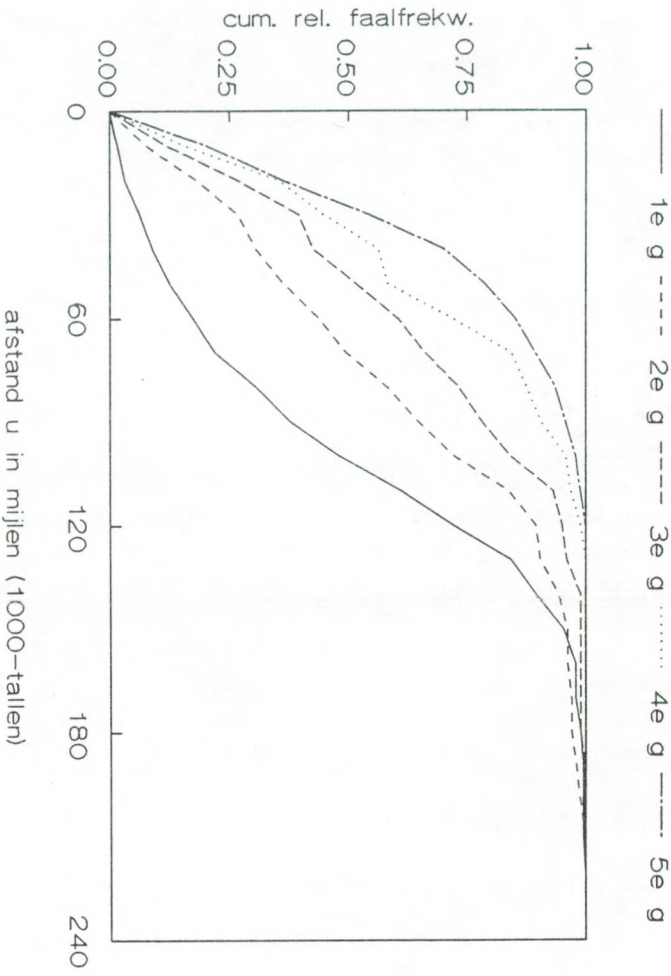
systemen te schrijven. Over de toepasbaarheid laten de auteurs zich nooit uit, laat staan dat ze met cases komen.

Naast het "na reparatie weer zo goed als nieuw" model (te modelleren door een vernieuwingsproces), is er nog een model dat in de literatuur meer dan incidenteel aandacht krijgt, en dat is het "na reparatie even slecht als vlak voor het falen" model, ook wel "zo-slecht-als-oud" model genoemd (te modelleren met een niet-homogeen Poissonproces).

Deze modellen zijn beide eenvoudig, en beide niet erg algemeen. Een logische vraag is: hoe modelleer je systemen, die na reparatie beter zijn dan vlak voor het falen, maar slechter dan een nieuw systeem? In zo'n algemener model willen we natuurlijk ook kenmerken kunnen opnemen als:
- onderhoudspolitiek,
- afhankelijkheid tussen de componenten,
- mate van belasting, enz..

Het is ook verrassend, dat in de literatuur over repareerbare systemen verklarende variabelen zelden een rol spelen. Dit is verrassend, omdat Prentice, Williams en Peterson al in 1981 Cox's Proportional Hazards model hebben gegeneraliseerd tot het faalgedrag van repareerbare systemen. Natuurlijk is een beperking van het Proportional Hazards model dat de covariabelen multiplicatief werken op de basis hazardfunktie, dus dat voor verschillende waarden van de covariabelen de hazardfunkties in de tijd een constante verhouding hebben, maar je zou toch verwachten dat het Proportional Hazards model bij de bestudering van repareerbare modellen met enige regelmaat opduikt. Dit is echter niet het geval.

Een oud en veel aangehaald voorbeeld van een repareerbaar systeem is te vinden in Davis (1952). Het betreft faalgegevens (gebaseerd op ernstige storingen) van busmotoren. De gegevens worden vollediger gepresenteerd in Downton (1969).
Zoals duidelijk in figuur 1 is te zien, is er sprake van veroudering, want de tijd tussen twee faalgebeurtenissen wordt in stochastische zin steeds kleiner. (Nauwkeuriger informatie dan de gegevens die in deze grafiek zijn weergegeven, hebben we niet kunnen achterhalen. Alleen de frekwentietabel waarop de grafiek is gebaseerd staat in de literatuur. In het bijzonder zijn geen individuele gegevens per motor beschikbaar. Een ander hinderlijk punt is dat de groepen per faalgebeurtenis (1-ste, 2-de, 3-de, enz) niet evengroot zijn, en dat niet bekend is wat hier de reden van is; mogelijk is er enige vertekening door select censureren. Ondanks deze onvolkomenheden zullen we deze gegevens toch gebruiken om een stukje modelbouw te demonstreren, en wel omdat dit de enige



Figuur 1. Cumulatieve relatieve frekwentiepolygoon voor de faalgebeurtenissen per gap (= periode tussen twee belangrijke storingen).

gegevensverzameling in de literatuur is die geschikt is om iets over veroudering te zeggen (In de terminologie van Cox (1990) hebben we te maken met een indirect model).

Aansluitend bij het Proportional Hazards model, spreek ik over hazardfunktie en bedoel ik daarmee wat Ascher en Feingold de Force of Mortality noemen, d.w.z.

$$h(t) = \frac{f(t)}{R(t)}$$

zodat h(t)dt de instantane kans op falen geeft.

We starten op tijdstip t=0 met een nieuw systeem. We veronderstellen dat de tijd die aan reparatie wordt besteed geen rol speelt (zoals bijvoorbeeld bij een kapotte pomp veelal het geval is: de pomp wordt vervangen door een andere pomp, en de reparatie geschiedt op een andere plaats en op een andere tijd). Dit levert een tijdas op, waarop reparatie geen tijd kost. Stel de faalmomenten op deze tijdas zijn t_1, t_2, ... We voorzien de hazardfunktie h van een index i, i= 0, 1, 2, ..., die het aantal uitgevoerde reparaties aangeeft, en we laten de tijdas voor h_i bij nul beginnen, door af te spreken dat h_0 wordt gedefinieerd op $[0,\infty)$, en dat

$h_i(t-t_i)$ is gedefinieerd voor $t \in [t_i, t_{i+1})$ $\qquad i = 1, 2, 3, \ldots$

Dan zijn de volgende relaties eenvoudig in te zien.

i) voor het "na reparatie weer zo goed als nieuw" model geldt

$$h_i(t-t_i) = h_0(t-t_i) \qquad t_i \le t < t_{i+1} \quad \text{en } i = 1, 2, \ldots$$

ii) voor het "na reparatie even slecht als vlak voor het falen" model geldt

$$h_i(t-t_i) = h_0(t) \qquad t_i \le t < t_{i+1} \quad \text{en } i = 1, 2, \ldots$$

We zoeken een model dat het faalgedrag van de busmotoren redelijk weergeeft.

We kijken eerst naar het Proportional Hazards model. Als we het eenvoudig

houden, stelt dit model dat voor de hazardfunktie geldt:

$$h(t;z) = h_0(t)\, g(z)$$

met h_0 de baseline hazardfunktie en g een positieve funktie. Cox kiest

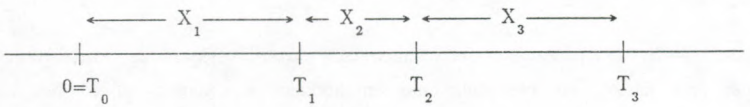
$$g(z) = \exp(\beta' z)$$

De covariabelen zijn samengevat in de vektor z. Voor de busmotoren kan bijvoorbeeld het aantal uitgevoerde reparaties als covariabele worden gezien.

Het Proportional Hazards model is wiskundig gezien bijzonder elegant, maar de structuur voor de hazardfunktie, met de multiplicatieve invloed van de covariabelen is heel speciaal.

Het model dat ik nu zal bespreken, is niet multiplicatief maar additief van aard. Voor een uitvoeriger beschrijving van het model verwijs ik naar Pijnenburg (1990), en op termijn naar Pijnenburgs proefschrift (1992).

Het additieve model is gebaseerd op de volgende beschouwing. Modelleer een repareerbaar systeem als een seriesysteem bestaande uit de onderling onafhankelijke componenten C_1 en C_2. Hierin representeert C_1 het repareerbare systeem onder standaardcondities, terwijl C_2 een imaginaire component voorstelt, die de invloed van verklarende variabelen representeert. T_i is het tijdstip van de i-de faalgebeurtenis, en $X_i = T_i - T_{i-1}$ is de i-de levensduur. (Zie figuur 2).



Figuur 2. Het faalproces.

Als een nieuw systeem wordt opgestart op tijdstip 0, dan geldt:
- C_1 heeft een hazardfunktie h_1
- C_2 heeft een constante hazardfunktie α_0

Op het tijdstip T_i (i = 1, 2, ...) worden beide componenten vervangen:
- C_1 door een aan C_1 identieke component
- C_2 door een component met een constante hazardfunktie α_i

De levensduren van alle componenten zijn onderling onafhankelijk verondersteld.

De gedachte is dus dat de hazardfunktie α van de tweede component alleen sprongsgewijs kan veranderen, en alleen op de momenten dat het systeem faalt. Mogelijke invloedsfaktoren zijn bijvoorbeeld:
- het aantal koude starts,
- het aantal reparaties.

De hazardfunktie van het systeem kan nu worden geschreven als

$$h(t, z_i) = h_1(t - t_i) + \alpha_i \qquad t \in [t_i, t_{i+1})$$

In figuur 3 worden voor het additieve model drie mogelijke hazardfunkties gepresenteerd.

We zullen nu het essentiële verschil tussen het additieve en het Proportional Hazards model bekijken, en dit toespitsen op één verklarende variabele.

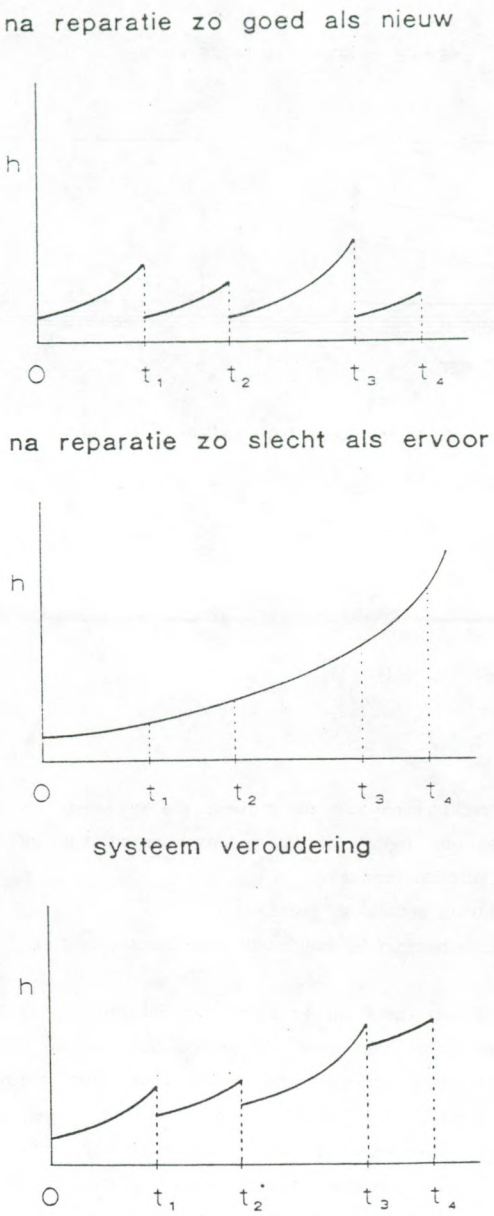
Uit de gedaante van het additieve model volgt, dat voor twee verschillende waarden van de covariabele, de bijbehorende hazardfunkties een constant verschil hebben. Het Proportional Hazards model geeft in dat geval juist een constant quotiënt. In figuur 4 zijn de hazardfunkties weergegeven voor een lineaire baseline hazardfunktie.

We zullen nu nagaan met welk van de twee modellen het faalgedrag van de busmotoren het beste is te beschrijven.

Helaas geeft de literatuur geen informatie over verklarende variabelen. We zullen ons model dus eenvoudig moeten houden: we kunnen alleen het aantal reparaties als verklarende variabele gebruiken.

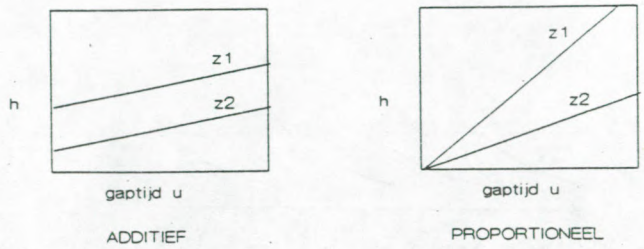
Via de cumulatieve hazardfunktie hebben we een mogelijkheid om de veronderstelling van additiviteit te verifiëren. Definieer de cumulatieve hazardfunktie H als de integraal van de hazardfunktie h, dan volgt direct dat het quotiënt $H(u;z_i)/u$ constant is, en als z het aantal reparaties aanduidt, dan liggen de grafieken van $H(u;z_i/u$, voor i = 1, 2, ... op onderling gelijke afstanden. Het schatten van de funktie $H(u;z_i)$ gebeurt op de standaard manier, namelijk door

$$\hat{H}(u,z_i) = -\ln(\hat{R}(u;z_i)$$



Figuur 3. Drie verschillende hazardfunkties voor het additieve hazards model.

LINEAIRE BASELINE HAZARDFUNKTIE



Figuur 4. Hazardfunkties voor het additieve en voor het proportionele hazards model.

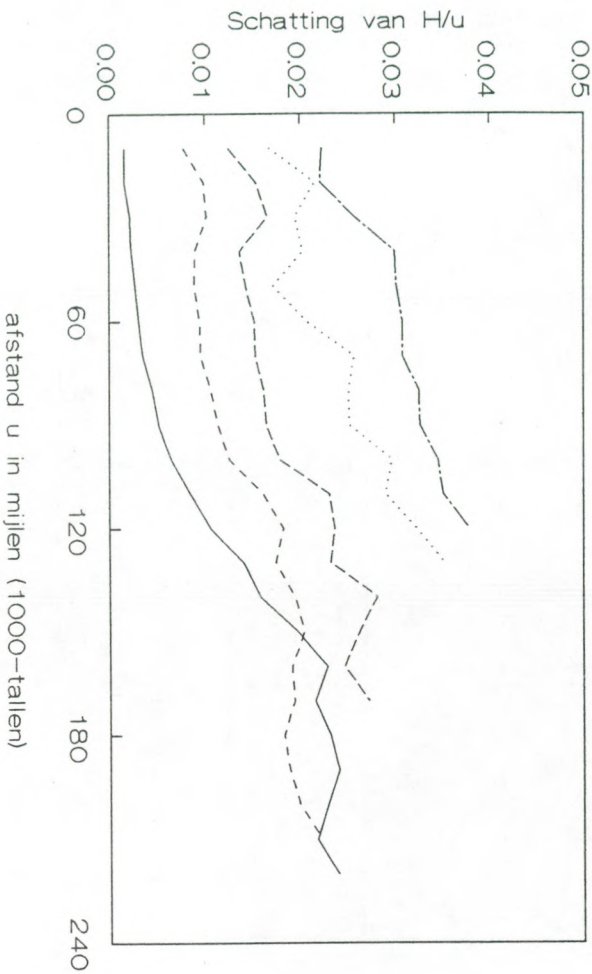
met bijvoorbeeld

$$\hat{R}(u;z_i) = (n_i(u)+1)/(n_i(0)+1)$$

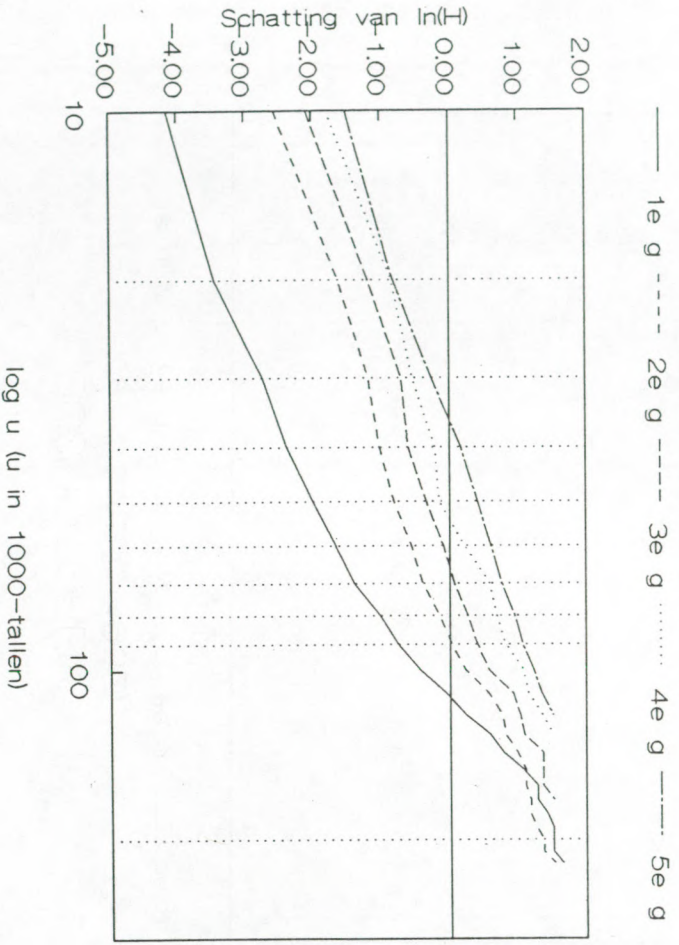
Dit leidt tot figuur 5.

Afgezien van de rechterkant van de grafiek, die gebaseerd is op heel weinig waarnemingen, lijkt de figuur aardig in overeenstemming met wat bij een additief model mag worden verwacht:
- de 5 krommen zijn bij benadering parallel,
- de krommen zijn bij benadering equidistant (in verticale richting).

Op analoge wijze is een check uit te voeren op het Proportional Hazards model. Zoals eenvoudig is in te zien moet nu gelden dat ln(H(u)) uitgezet tegen u parallelle en equidistante krommen moet opleveren. Nog aardiger wordt de grafiek als we ln(H(u)) niet uitzetten tegen u maar tegen ln u (immers, trekkingen uit een Weibullverdeling leveren op deze wijze uitgezet een rechte lijn). Uit figuur 6 lezen we nu af, dat het Proportional Hazards model de faalgegevens minder goed lijkt te beschrijven als het additieve model: de gegevens tot het eerste falen wijken duidelijk af van de rest.



Figuur 5. Het additieve hazards model toegepast op het faalgedrag van busmotoren.



Figuur 6. Het Proportional Hazards model toegepast op het faalgedrag van busmotoren.

Conclusie

1. Kijkend naar de fysische eigenschappen van systeemveroudering, heeft het additieve model eigenschappen die voldoende interessant zijn voor de praktijk om bestudering wenselijk te maken.
2. In het voorbeeld van de busmotoren lijkt het additieve model, in de primitieve opzet met alleen het aantal storingen als covariabele, niet alleen fysisch aantrekkelijker dan het Proportional Hazards model, maar lijkt het ook beter aan te sluiten bij de waarnemingen.

4. AFSLUITING

Dames en heren, hiermee kom ik aan het einde van mijn bijdrage. Ik heb gepoogd uw aandacht te vestigen op MODELBOUW, een gebied dat naar mijn mening nog nauwelijks is ontgonnen, maar tevens een gebied waar de statisticus zich voortdurend op beweegt. Een gebied ook, waarvan het belang door Professor Leppink al ruim 25 jaar geleden werd onderstreept.

LITERATUUR

Aranda Ordaz, F.J. (1980). *Transformations to additivity for binary data.* Unpublished Ph.D. thesis, Imperial College, London.

Ascher, H. & Feingold, H. (1984). *Repairable systems reliability.* Marcel Dekker, New York.

Cox, D.R. (1990). Role of models in statistical analysis. *Statistical Science*, **5**, 169–174.

Davis, D.J. (1952). An analysis of some failure data. *J. Amer. Stat. Assoc.*, **47**, 113–150.

Downton, F. (1969). An integral equation approach to equipment failure. *J. Roy. Stat. Soc. B*, **31**, 335–349.

Elandt–Johnson, R.C. (1980). Some prior and posterior distributions in survival analysis: A critical insight on relationships derived from cross–sectional data. *J. Roy. Stat. Soc. B*, **42**, 96–106.

Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans. Roy. Soc. London Ser. A*, **222**, 309–368.

Freudenthal, H. (1966). *Waarschijnlijkheid en statistiek* (3^e *druk*). De Erven F. Bohn N.V., Haarlem.

Lehmann, E.L. (1990). Model specification: the views of Fisher and Meyman, and later developments. *Statistical Science*, **5**, 160–168.

Miller, D.R. (1986). Exponential Order Statistics Models of Software Reliability Growth. *IEEE Transactions on Software Engineering*, **12**, 12–24.

Pijnenburg, M. (1990 of 1991). Additive hazards models in repairable systems reliability. *Reliability Engineering and Systems Safety* (geaccepteerd).

Prentice, R.L., Williams, B.J. & Peterson, A.V. (1981). On the Regression Analysis of Multivariate Failure Time Data. *Biometrics*, **68**, 373–379.