

DE EENDIMENSIONALE STATISTICUS

Rede, in verkorte vorm uitgesproken
bij het aanvaarden van het ambt van
hoogleraar in de mathematische statistiek
in de Faculteit der Wiskunde en Informatica
van de Universiteit van Amsterdam
op woensdag 20 mei 1992

door

Dr. Chris A.J. Klaassen

1992

Universiteit van Amsterdam

Mijnheer de Rector Magnificus,

Leden van het College van Bestuur van deze Universiteit, en voorts Gij allen die door Uw aanwezigheid blij geeft van belangstelling voor de al dan niet ééndimensionale statistiek,

Zeer gewaardeerde Toehoorders,

Stochasticus versus Determinist

De beroemdste natuurkundige van deze eeuw, Einstein, schreef eens : "Du glaubst an den würfelnden Gott und ich an volle Gesetzlichkeit"; zie Born (1949, p. 122). Deze woorden "Jij gelooft in een God die dobbelt, en ik in volledige wetmatigheid" waren gericht aan Born. Born is onder andere bekend vanwege zijn interpretatie van het kwadraat van de golffunctie als kansdichtheid. Ik wil U geruststellen. De technische details van mijn betoog zijn niet van wezenlijk belang. Mijn bedoeling in deze rede is om een algemeen facet van de huidige statistiekbeoefening te belichten. Details kunnen het resulterende lichtbeeld verfraaien maar niet essentieel veranderen.

Welnu, in het gegeven citaat is duidelijk sprake van een botsing van twee wereldbeelden: het deterministische van Einstein en dat van Born, waarin het toeval een fundamentele rol speelt. Laten we ons nu eens een statisticus voor de geest halen. U kunt zich van zo iemand reeds enige voorstelling maken. U hebt er tenslotte al drie hun oratie horen geven en wellicht bent U zelf statisticus of, erger nog, is Uw partner het; zie Kendall en Stuart (1961, p. v). Voor de meeste situaties zal onze statisticus als benadering van de werkelijkheid een stochastisch model kiezen, dat wil zeggen een model waarin het toeval door middel van een kansmechanisme de loop der dingen bepaalt. Daarmee kiest hij de zijde van Born. Uiteraard is Einstein qua vermaardheid ideaal gezelschap, maar onze statisticus kan zich troosten met het feit dat ook Born Nobelprijswinnaar is.

De keuze tussen stochastische en deterministische modellen als benaderingen voor een stukje realiteit is meestal overigens niet erg moeilijk. Vaak ligt de juiste keuze voor de hand. Zelden is zij lastig te maken. Bijna steeds is er daarbij sprake van een inschatting en moet de vraag worden beantwoord of het toevalselement in het model verwaarloosd kan worden, zodat kan worden volstaan met een deterministisch model. Daarom kan enige beroepsdeformatie bij zulke keuzes een rol spelen. Tenslotte is voor de man met een hamer bijna elk probleem een spijker.

Schatten en Schatters

Laten we onze statisticus nu eens een schattingsprobleem voorleggen. De term schatten heeft in de statistiek nagenoeg dezelfde betekenis als in het dagelijks taalgebruik. Een aannemer schat de bouwkosten van een huis; in een taxatierapport wordt de schade aan een auto geschat; een arts schat de duur van een ziekteperiode van een patient; een bakker schat de hoeveelheid brood die hij moet bakken; enzovoorts. Binnen de statistiek gaat dit schatten op een

wetenschappelijke manier waarbij voor een goed gedefinieerde situatie een goedgedefinieerde grootheid geschat moet worden. Ik wil U enige voorbeelden noemen van te schatten grootheden in enkele van de zeer vele toepassingsgebieden van de statistiek:

- Sterrenkunde; de relatie tussen helderheid en snelheid van verre hemellichamen; zie Bhattacharya, Chernoff and Yang (1983).
- Biologie; de energiebehoefte van strandlopers tijdens hun overwintering in Afrika; zie Klaassen, Kersten and Ens (1990).
- Kankeronderzoek; de mate waarin zekere stoffen kankerverwekkend zijn; zie Freedman and Zeisel (1988).
- Epidemiologie; de kans op HIV-infectie via bloedtransfusies; zie Fusaro et alii (1989).
- Accountantscontrole; het totale bedrag aan verkeerd geboekte posten in een boekhouding; zie Panel on Nonstandard Mixtures of Distributions (1989).
- Criminologie; de waarschijnlijkheid dat een verdachte schuldig is bij identificatie via DNA-onderzoek; zie Berry (1991).
- Parapsychologie; de kans op een zes bij het werpen met een dobbelsteen door een paranormaal begaafd persoon; zie Utts (1991).
- Techniek; de kans op een melt down tijdens de voorziene bedrijfsperiode bij een te bouwen kerncentrale; zie Speed (1985).

Uiteraard dient het schatten van de onbekende waarde van grootheden als in deze voorbeelden te geschieden op basis van data, ook wel (meet)gegevens, of waarnemingen genoemd. Hierbij zijn de adagia: "meten is weten" en "meer meten is meer weten". Laten we nu aannemen dat onze statisticus voor een van de genoemde voorbeelden met de bijbehorende data een stochastisch model gebouwd heeft. Vanzelfsprekend heeft hij dat gedaan in samenspraak met de deskundigen op het desbetreffende onderwerp. Na deze fase van de modelbouw is onze statisticus binnen het eigenlijke domein van de statistiek beland.

Er staan hem nu verscheidene schattingsmethoden ter beschikking om op basis van de waarnemingen tot schattingen te komen van bijvoorbeeld de energiebehoefte van strandlopers. Zulke schattingsmethoden worden meestal schatters genoemd. Onze statisticus wil natuurlijk de beste schatter kiezen. Er rijzen dan echter vragen: "In welke zin de *beste*? Bestaat de *beste* wel? Is het wel zo natuurlijk om naar de beste schatter te zoeken?" Dit zijn vragen die ik in deze inaugurale rede met U wil bestuderen.

Kernenergie en Amateurisme

Alvorens aan deze studie te beginnen, wil ik U nog even meenemen op een interessant maar gevaarlijk zijpadje. Bij een van de genoemde voorbeelden sprak ik U over het schatten van de kans op een "melt down" in een kerncentrale. Sinds Fermi een halve eeuw geleden, om precies te zijn op 2 december 1942, in Chicago de eerste kernreactor demonstreerde, zijn er zoveel kerncentrales in bedrijf genomen dat het mogelijk moet zijn de gevraagde kans redelijk nauwkeurig en betrouwbaar te schatten, mits de noodzakelijke gegevens ter beschikking worden gesteld.

In 1975 verscheen in opdracht van de regering van de Verenigde Staten van Amerika het zogenaamde Rasmussen rapport. Hierin werd een methode voor risico-analyse bij kernenergie ontwikkeld die veel kritiek heeft uitgelokt. Kansen op ketens van gebeurtenissen die tot een melt down leiden, worden daarbij geschat door de kansen op de individuele gebeurtenissen te schatten. Deze laatste kansen worden vaak niet met behulp van statistische methoden op basis van data geschat maar worden gegokt op basis van "expertise"; zeg maar met de natte vinger. Bovendien worden onbekende risicofactoren en de onberekenbare mens als risicofactor in deze methode niet meegenomen. Deze onwetenschappelijke benadering is in latere rapporten, zoals de Sizewell B Probabilistic Safety Study van Westinghouse uit 1982, niet vervangen door een betere.

In 1985 schreef de mathematisch statisticus Speed dan ook, dat de schrijvers van deze rapporten blijkbaar de subtiliteiten van de statistiek niet goed begrepen hadden en geen gebruik hadden gemaakt van duidelijke en gangbare statistische analysemethoden en voorts, dat vele statistici bereid en in staat zouden zijn aan een *wetenschappelijke* studie over dit onderwerp mee te werken. Zo'n studie zou dan leiden tot veel betrouwbaardere schatters en tot waarschijnlijk hogere schattingen van de kans op een melt down dan de 1 op 20.000 per jaar uit het Rasmussen rapport. Overigens betekent deze laatste kans bij de 423 kerncentrales die eind 1990 in gebruik waren en die ieder ongeveer 25 jaar in bedrijf zouden moeten zijn, een kans van 41% dat het bij minstens een van deze centrales tot een melt down komt.

Met U teruglopend over dit doodlopende zijpaadje is mijn conclusie dat de neiging van experts op vele vakgebieden om de kennis en expertise van echte statistici te negeren kan leiden tot ronduit foutieve en soms zelfs levensgevaarlijke conclusies.

Ordeningen en Schaatsen

Teruggekeerd bij het hoofdthema van deze rede "Is het wel zo natuurlijk om naar de beste schatter te zoeken?" zien we al snel dat het begrip ordening hier van cruciaal belang is. We zullen twee typen ordeningen definiëren: partiële en totale. Ik wil deze begrippen invoeren aan de hand van het wereldkampioenschap hardrijden op de schaats voor bijvoorbeeld heren.

Bij de wedstrijden voor dit wereldkampioenschap wordt geschaatst over vier afstanden te weten de 500 en 1500 meter en de 5 en 10 kilometer. Na afloop van de wedstrijden kunnen de deelnemers gedeeltelijk geordend worden door af te spreken dat schaatser A beter is dan schaatser B, als A op alle vier afstanden een betere tijd heeft gerealiseerd dan B. Het is duidelijk dat niet bij elk tweetal schaatsers op deze manier de beste van de twee bepaald zal kunnen worden. Het kan immers voorkomen dat A op twee afstanden betere en op de andere twee afstanden slechtere tijden heeft bereikt dan B. Wel is het zo dat, als A beter is dan B en B op zijn beurt beter is dan C, dan is A ook beter dan C. Dit natuurlijke begrip van beter leidt dus tot een gedeeltelijke ofwel partieel geordende verzameling schaatsers.

Slechts zeer zelden is er een schaatser die in deze partiële ordening beter is dan elke andere schaatser, en die we dus de beste kunnen noemen. Hij zou dan immers winnaar op alle vier afstanden zijn. Om in ieder geval één wereldkampioen schaatsen te kunnen aanwijzen heeft men een puntensysteem ingevoerd. Hierbij worden de tijden op de verschillende afstanden

omgerekend tot gemiddelde tijden voor 500 meter, worden de resulterende vier tijden opgeteld en wordt het resultaat het puntentotaal genoemd. Nu kunnen we A beter noemen dan B als het puntentotaal van A kleiner is dan dat van B. Het komt zeer zelden voor dat twee schaatsers een zelfde puntentotaal bereiken. Meestal leidt deze procedure dan ook tot een totale ordening van de deelnemende schaatsers in het zogenaamde puntenklassement. De wereldkampioen is dan degene met het laagste puntentotaal in dit klassement met de extra regel, dat een schaatser die drie afstanden gewonnen heeft, wereldkampioen wordt ongeacht zijn puntentotaal.

We hebben dus gezien, dat het onderling vergelijken van schaatsers via hun tijden op de vier afstanden slechts tot een partiële ordening leidt en dat het puntenklassement een totale ordening geeft. Overigens is het nuttig op te merken dat de totale ordening via het puntenklassement in overeenstemming is met de partiële ordening via de vier afstanden. Als immers schaatser A op elk van de vier afstanden beter is dan B dan is ook het puntentotaal van A lager dan dat van B. De formele definities luiden nu als volgt. Een verzameling heet *partieel geordend* als er een relatie "beter" is gegeven tussen sommige en niet noodzakelijk alle elementen van de verzameling waarvoor het volgende geldt: (1) Als A beter is dan B, dan is B niet beter dan A en zijn A en B verschillend. (2) Als A beter is dan B en B beter dan C, dan is A beter dan C. Een verzameling heet *totaal* of ook wel *lineair geordend* als er een relatie "beter" is gegeven waarvoor geldt: (1) Voor elk tweetal verschillende elementen A en B van de verzameling is òfwel A beter dan B òfwel B beter dan A. (2) Als A beter is dan B en B beter dan C, dan is A beter dan C.

Geordende verzamelingen geven via het lemma van Zorn en het daarmee equivalente keuzeaxioma aanleiding tot een grondslagendiscussie in de wiskunde. Het is in dit verband aardig om te verwijzen naar de rede "Intuitionisme en Formalisme" van de vermaarde wiskundige Brouwer uitgesproken bij zijn "aanvaarding van het ambt van buitengewoon hoogleraar in de wiskunde aan de Universiteit van Amsterdam op maandag 14 oktober 1912", zie Brouwer (1912).

Ordeningen en Schatters

Zoals schaatsers bij een wereldkampioenschap met elkaar vergeleken worden op grond van hun prestaties op de vier afstanden, zo moeten ook schatters bij een zeker statistisch probleem met elkaar worden vergeleken op grond van hun prestaties onder verschillende omstandigheden. Laten we hier even over nadenken aan de hand van het voorbeeld van het schatten van de energiebehoefte van strandlopers. Met deze energiebehoefte bedoelen we een theoretische grootheid waar de waarnemingen, dat wil zeggen de gemeten energiebehoeften van een aantal strandlopers, omheen schommelen. Deze theoretische energiebehoefte zullen we de parameter noemen en de onbekende waarde van deze parameter moet geschat worden.

Het is duidelijk dat als de waarde van de parameter groot is en dus de theoretische energiebehoefte groot is, dan zullen ook de waarnemingen in het algemeen grote waarden te zien geven. Omdat de schatting op deze waarnemingen gebaseerd is, zal de prestatie van de corresponderende schatter ook afhangen van de theoretische energiebehoefte; algemener geformuleerd: de prestatie van een schatter hangt af van de waarde van de parameter. Inderdaad moeten schatters dus met elkaar worden vergeleken op grond van hun prestaties

onder verschillende omstandigheden, namelijk onder de verschillende waarden van de parameter. Dit betekent dat schatters van de theoretische energiebehoefte van strandlopers met elkaar dienen te worden vergeleken op grond van hun prestaties onder alle in principe mogelijke waarden van de theoretische energiebehoefte, waarvan er een de onbekende werkelijke waarde is.

De prestatie van een schaatser op een zekere afstand geven we weer met zijn tijd op die afstand. Wat we met de prestatie van een schatter onder een zekere parameterwaarde bedoelen, moeten we nog preciseren. Laten we daarvoor een bepaalde schatter beschouwen onder een zekere parameterwaarde. Deze schatter geeft met behulp van de waarnemingen een schatting van de parameterwaarde. Als we het hele experiment zouden herhalen, zouden we andere waarnemingen krijgen die bij deze zelfde schatter een andere schatting van de parameterwaarde zouden opleveren. Zouden we het experiment blijven herhalen dan zouden we steeds andere schattingen vinden. Zoals de waarnemingen van de energiebehoefte van strandlopers variëren rond de theoretische energiebehoefte, dus rond de parameterwaarde, zo zouden dan al deze schattingen variëren rond dezelfde parameterwaarde. Deze variatie van de schattingen is een eigenschap van de schatter die in de terminologie van de statistiek volledig wordt beschreven door de *verdeling van de schatter*. De prestatie van de schatter onder de parameterwaarde wordt dus beschreven door de verdeling van de schatter onder deze parameterwaarde.

Een vollediger naam voor verdeling is kansverdeling. De kansverdeling van een schatter onder een parameterwaarde beschrijft dus de variatie van de bijbehorende schattingen rond deze parameterwaarde bij een eindeloze rij herhalingen van het experiment. Laat ik een poging wagen dit begrip kansverdeling iets nader te preciseren. Daarbij zal de terminologie sterker wiskundig getint zijn dan die welke ik tot nu toe heb kunnen gebruiken.

Wanneer we in de bovengenoemde rij van herhalingen van het experiment telkens weer nieuwe schattingen voor dezelfde parameterwaarde vinden, kunnen we de fractie van de eerste n schattingen beschouwen die lager uitvallen dan een zeker getal, zeg x . Omdat we een eindeloze rij herhalingen van het experiment ter beschikking hebben, kunnen we bestuderen hoe deze fractie zich gedraagt als we door de rij lopen, dat wil zeggen als we n , het aantal schattingen dat we bekijken, steeds maar groter nemen. Intuïtief zien we in dat deze fractie steeds dichter bij een vaste waarde komt te liggen. Deze waarde hangt van x af en deze noemen we de kans dat de schatting lager zal uitvallen dan x bij uitvoering van ons experiment. De resulterende functie van x beschrijft nu de kansverdeling van de schatter. Deze kansverdeling geeft dus aan hoe de waarden verdeeld zijn die de schattingen aannemen in een eindeloze rij herhalingen van het experiment.

Zoals gezegd, wordt de prestatie van de schatter onder een bepaalde waarde van de parameter beschreven door de kansverdeling van de schatter onder deze parameterwaarde. Samenvattend kunnen we het volgende zeggen. Zoals schaatser bij een wereldkampioenschap partieel geordend worden door hun tijden op elk van de vier afstanden onderling te vergelijken, zo kunnen schatters partieel geordend worden door hun kansverdelingen onder elk van de mogelijke parameterwaarden onderling te vergelijken.

Ordeningen en Kansverdelingen

Het vergelijken van twee tijden op één schaatsafstand kan op een natuurlijke manier: de kortere tijd is de betere. Merk op, dat we hier weer een totale ordening tegen komen, net als bij het puntenklassement. Echter, hoe de kansverdelingen van twee schatters onder een waarde van de parameter met elkaar vergeleken moeten worden, is niet zonder meer duidelijk. In feite is er slechts één natuurlijke eigenschap die een ordeningsrelatie voor kansverdelingen van schatters moet bezitten: de kansverdeling van een schatter onder een parameterwaarde moet beter zijn naarmate deze verdeling sterker geconcentreerd is rond deze waarde. Dit betekent immers dat we de voorkeur geven aan de ene schatter boven de andere als de schattingen van de ene in het algemeen dichter bij de parameterwaarde komen te liggen dan die van de andere. We zijn dus op zoek naar een operationalisering van het begrip concentratie van een kansverdeling.

Statistische Beslissingstheorie

De klassieke benadering hiervoor is die van de statistische beslissingstheorie van Wald (1950). Daarin wordt één enkel getal toegekend aan de kansverdeling van de schatter onder de parameterwaarde. Dit getal wordt het risico van de schatter onder de parameterwaarde genoemd en dit risico is inderdaad groter naarmate de verdeling van de schatter minder geconcentreerd is rond die parameterwaarde. Het risico van een schatter onder de parameterwaarde kan dus opgevat worden als een maat voor het niet-geconcentreerd zijn van de verdeling van de schatter onder deze waarde.

De verdelingen van schatters onder één parameterwaarde kunnen nu totaal geordend worden door de corresponderende risico's te vergelijken. Dus de verdeling van schatter A is beter dan die van schatter B als het risico van schatter A kleiner is dan dat van schatter B. We merken hier nog bij op dat verschillende schatters hetzelfde risico kunnen hebben en dat zij, wat dat betreft, dus equivalent kunnen zijn.

Als U mij zou toestaan nog iets technischer te worden, dan zou ik U vertellen, dat het risico van een schatter onder een parameterwaarde gedefinieerd wordt via een verliesfunctie en de kansverdeling van de schatter onder de parameterwaarde. Deze verliesfunctie geeft aan hoe groot het verlies wordt geacht als de schatting op een bepaalde afstand van de parameterwaarde terecht komt. In het kiezen van de verliesfunctie is een grote vrijheid, maar veruit de meest gemaakte keuze is die van een kwadratische functie.

Laten we nu eens kijken hoe het volledige plaatje voor het ordenen en vergelijken van schatters voor een schattingsprobleem er in de statistische beslissingstheorie uitziet. Voor elke parameterwaarde hebben we een totale ordening van de schatters via hun risico's, zoals we voor elke afstand een totale ordening van de schaatsers hebben via hun tijden. Verder krijgen we een partiële ordening van de schatters door te stellen dat schatter A beter is dan schatter B als onder *elke* parameterwaarde het risico van schatter A kleiner is dan dat van schatter B. Ook dit is weer analoog aan de situatie bij het wereldkampioenschap schaatsen. Om nu een kampioen onder de schatters aan te wijzen kunnen we verscheidene wegen bewandelen. Eén daarvan is de risico's onder de verschillende parameterwaarden optellen en zo een

puntenklassement maken. Na herleiding van alle vier schaatstijden tot 500 meter tijden is dit wat er in de schaatswereld gebeurt. In de wereld van de statistiek produceren variaties op deze methode kampioenen die Bayes-schatters worden genoemd.

Een andere weg wordt ingeslagen door voor elke schatter het grootste ofwel maximale risico onder alle parameterwaarden te bepalen en door vervolgens die schatter als de beste te kiezen waarvoor dit maximale risico het kleinste ofwel minimaal is. Dit levert de zogenaamde minimaxschatter op. Voor het schaatskampioenschap zou dit betekenen dat voor elke schaatser de vier tijden worden omgerekend tot 500 meter tijden, dat van deze vier 500 meter tijden de grootste wordt bepaald en dat diegene minimaxkampioen wordt voor wie deze maximale tijd het kleinst is. Hoewel statistici heel content kunnen zijn met een minimaxschatter, zal een voorstel om het wereldkampioenschap schaatsen te verrijden onder dit minimaxregime slechts hilariteit opwekken, omdat het bijna haaks staat op de vigerende regel dat de winnaar van drie afstanden kampioen is.

De daarstraks genoemde keuze van een kwadratische functie als verliesfunctie gaat niet zelden samen met een eis op de schatters, namelijk dat deze zuiver moeten zijn, of zoals de Vlamingen onder ons dat zo expressief zeggen: onvertekend. Voor zuivere schatters reduceert het risico tot de zogenaamde variantie en een zuivere schatter heeft dus de beste verdeling onder een parameterwaarde als zijn variantie onder deze parameterwaarde minimaal is. Het aardige en verrassende is nu dat er statistische problemen zijn waarvoor er een zuivere schatter bestaat die minimale variantie heeft onder alle parameterwaarden tegelijkertijd. Hierbij worden alle schatters van deelname uitgesloten die niet zuiver zijn, en dit is natuurlijk wel een ernstige smet op het blazoen van deze schatterkampioen. Zoals Wolfowitz (1965, p. 249) in een zelfde metafoor al opmerkte, leidt het uitsluiten van deelnemers aan wedstrijden niet tot een *echte* kampioen.

Kleeft aan deze restrictie tot zuivere schatters het bezwaar dat de klasse van toegelaten schatters op enigszins kunstmatige wijze ingeperkt wordt, een ernstiger bezwaar tegen de geschetste benadering is in mijn ogen de exclusieve aandacht voor de variantie als het standaard criterium voor het vergelijken van de concentratie van kansverdelingen van schatters. Dit laatste bezwaar geldt overigens voor de gehele opzet van de klassieke statistische beslissingstheorie, omdat daarin steeds één bepaalde verliesfunctie wordt gekozen en omdat met behulp van de corresponderende risico's de kansverdelingen van schatters worden vergeleken en geordend. De resulterende ordening onder een parameterwaarde hebben we totaal en ook wel lineair genoemd, maar deze kunnen we met goed recht ook het predikaat ééndimensionaal opplakken. Immers de schatters worden op deze manier op een rijtje gezet en wanneer we dit visualiseren, kunnen we voor ons geestesoog een rechte lijn tekenen, die maar één dimensie heeft en dus ééndimensionaal is, en op deze rechte lijn kunnen we dan de schatters of beter gezegd hun risico's door middel van punten weergeven.

Kortom: de benadering met verliesfuncties van de klassieke statistische beslissingstheorie is in wezen ééndimensionaal.

Asymptotiek

De meeste statistische schattingsproblemen zijn moeilijk te analyseren. Daarom worden ze vaak benaderd met behulp van eenvoudiger schattingsproblemen. Dit wordt meestal gedaan door het aantal n van de waarnemingen waarop de schatters kunnen worden gebaseerd, in gedachte steeds maar groter te laten worden. Het resultaat is een schattingsprobleem dat in veel gevallen een goede benadering is voor het originele schattingsprobleem en bovendien buitengewoon eenvoudig is. Het wiskundig probleem is nu relaties aan te geven tussen de beide schattingsproblemen, het originele en het benaderende.

De twee belangrijkste stellingen op dit terrein, de minimaxstelling en de convolutistelling, zijn te danken aan de mathematisch statistici Hájek (1970, 1972) en Le Cam (1953, 1972, 1986). Het zijn stellingen die aangeven welke de beste prestaties zijn, waarop voor schatters kan worden gehoopt. Of er schatters zijn die inderdaad deze prestaties leveren, is voor zogenaamde semiparametrische en nietparametrische modellen vaak een open vraag. Dit is dan ook in het huidige statistiekonderzoek een belangrijk terrein; zie Bickel et alii (1992).

De tijd ontbreekt hier om op alle technische aspecten van deze stellingen in te gaan. Laat ik volstaan met enkele opmerkingen. Zoals de naam al suggereert, is de aanpak bij de minimaxstelling dezelfde als bij de minimaxschatters uit de klassieke statistische beslissingstheorie. Daarom gaat ook de minimaxstelling gebukt onder haar ééndimensionaal karakter. De convolutistelling echter heeft dit nadeel niet. Zij is niet ééndimensionaal van aard en is gerelateerd aan een natuurlijke partiële ordening voor kansverdelingen geïntroduceerd door Bickel en Lehmann (1979); zie ook Klaassen (1985). Deze partiële ordening zullen we de spreidingsordening noemen. De enige minder prettige eigenschap van de convolutistelling is, dat zij aan de schatters restricties oplegt; zie Van der Vaart (1989).

de Ééndimensionale Statisticus

In de ideale benadering voor de statistische analyse van een schattingsprobleem, die ik U hier heb pogen te schetsen, worden schatters met elkaar vergeleken en daarmee partieel geordend door onder elk van de parameterwaarden hun kansverdelingen te vergelijken. Verder worden de kansverdelingen die bij één parameterwaarde horen, zelf weer partieel geordend volgens hun concentratie rond deze parameterwaarde. Hierbij behoort de resulterende ordening van kansverdelingen niet totaal ofwel ééndimensionaal te zijn, omdat kansverdelingen nu eenmaal vele verschillende facetten en dus dimensies hebben. In sommige situaties is de spreidingsordening een natuurlijke keuze. Of de gekozen ordening voor kansverdelingen nu ééndimensionaal is of niet, de resulterende ordening van schatters is partieel.

Bij zulke ordeningen kan het voorkomen dat er een beste schatter is, dat wil zeggen een schatter die in de gegeven partiële ordening beter is dan elke andere. Dit behoeft echter niet het geval te zijn. Wel zijn er dan in het algemeen schatters die niet verbeterd kunnen worden. Een schatter kan niet verbeterd worden als er geen enkele schatter is die beter is. Dit impliceert niet dat deze schatter dan de beste is. Vaak zal het zo zijn, dat er een hele groep van schatters is die niet verbeterd kunnen worden. Dit kan natuurlijk alleen optreden als deze schatters onderling niet vergelijkbaar zijn via de gegeven partiële ordening.

Ik geef toe dat het plaatje dat een partiële ordening oplevert, bij lange niet zo overzichtelijk is als een ééndimensionale ordening. Dit verklaart naar mijn mening de neiging van de meeste mensen en dus ook statistici om "de dingen op een rijtje te zetten", dat wil zeggen ééndimensionaal te ordenen. Verder zal een statisticus die "ze niet allemaal op een rijtje heeft", niet serieus worden genomen.

Zoals we hebben gezien is er dan ook slechts één klassiek resultaat dat in onze ideale benadering past, de convolutiestelling. Deze heeft echter als nadelen een stelling te zijn over een benadering van het eigenlijke schattingsprobleem en restricties op te leggen aan de schatters. Andere resultaten, met name uit de statistische beslissingstheorie, voldoen zeker niet, omdat daarbij de kansverdelingen ééndimensionaal geordend worden via hun risico's. Zelf heb ik enkele resultaten afgeleid die betrekking hebben op de spreidingsordening, maar naar mijn overtuiging kan en moet er nog zeer veel onderzoek in deze richting gedaan worden; zie Klaassen (1984, 1989a, 1989b).

Voorlopig echter ligt de nadruk binnen de theoretische statistiek nog op ééndimensionale ordeningen en bij de meeste toepassingen van de statistiek is de aandacht bijna exclusief op varianties gefixeerd. De ééndimensionale statisticus zal dus nog wel een lang leven beschoren zijn. Dit is geen ramp, want elke ontdekking van een ééndimensionale ordening is prachtig, maar vaak zal een partiële ordening natuurlijker zijn en een vollediger beeld geven. Ééndimensionaal is hier dan ook bijna een epitheton ornans en het is zeker niet zo kritisch en denigrerend bedoeld als in "One Dimensional Man" van Marcuse (1964). Die studie over de ideologie van de industriële, kapitalistische samenleving gaf aan het eind van de jaren zestig door haar maatschappij-kritiek een basis aan de studentenbeweging. Deze kritiek lijkt heden ten dage grotendeels verstomd. Ik mag slechts hopen dat mijn kritiek op een ééndimensionale benadering van de statistiek niet dezelfde weg gaat.

Andere Dimensies

Aan het eind van mijn oratie gekomen dank ik het College van Bestuur voor mijn benoeming en het in mij gestelde vertrouwen, en wil ik graag ook al diegenen bedanken die op een of andere wijze mijn ontwikkeling tot en als statisticus hebben beïnvloed, zoals J. Oosterhoff, J. Hemelrijk, P.J. Bickel, J.A. Wellner, Y. Ritov en vele anderen. Enkelen van hen wil ik speciaal noemen. In de eerste plaats mijn vrouw en kinderen.

Lieve Angeli, Paul, Linda en Angeli, zeer dankbaar ben ik jullie voor de oneindig vele dimensies buiten de wiskunde waarmee jullie mijn leven verrijken.

Verder wil ik mijn ouders noemen, die door hun zeer werkzame leven mij het enig juiste voorbeeld voor studie en onderzoek hebben gegeven. Ik ben heel gelukkig dat mijn vader deze ceremonie kan meemaken.

Hooggeleerde Van Zwet, Beste Willem, levendig kan ik me herinneren, hoe enthousiast je was, toen je me de consequenties liet zien van een ongelijkheid die ik gevonden had. In zekere zin is het onderwerp van deze rede een vervolg daarop. Tijdens mijn promotie-onderzoek onder jouw leiding en ook tijdens de jaren daarna in Leiden heb ik jouw snelle

scherpzinnigheid en feilloze intuïtie van nabij mogen meemaken. Ik ben je zeer erkentelijk voor de steun die ik in mijn loopbaan steeds van je heb gekregen. Ik kan slechts pogen een waardig leerling te zijn.

Hooggeleerde Does, Beste Ronald, jouw optimisme en blijmoedigheid maakten het bijzonder plezierig en stimulerend om met jou een kamer te delen op het Mathematisch Centrum. Ik ben erg blij dat je in staat bent gesteld jouw bijdrage te leveren aan een volwaardige statistiekopleiding aan de Universiteit van Amsterdam. Ik zou echter niet meer een kamer met je willen delen, want jouw bestuurlijke energie leidt geregeld tot een te hectische werkomgeving.

Leden van de Faculteit der Wiskunde en Informatica en al die leden van de Faculteit der Economische Wetenschappen en Econometrie die nauw betrokken zijn bij de Vrije Studierichting Statistiek. Vanuit een aanstelling waarin de nadruk lag op onderwijs en onderzoek, ben ik aan de Universiteit van Amsterdam benoemd in een positie waarin het accent veel sterker op de bestuurlijke activiteiten ligt. Veel hulp en advies hierbij heb ik tot nog toe van U allen mogen ontvangen. Ik verheug me op een vruchtbare voortzetting van onze samenwerking en heb daar alle vertrouwen in.

Dames en Heren Studenten, der traditie getrouw gelden de laatste woorden U. De statistiek is een uitzonderlijk breed toepasbare wetenschap met zeer veel interessante en uitdagende aspecten die uiterst theoretisch, wiskundig of praktisch van aard zijn. Ik hoop U vandaag duidelijk gemaakt te hebben dat daar nog vele dimensies aan toegevoegd kunnen worden. Ik zie het als mijn belangrijkste taak U te inspireren tot enthousiasme voor de statistiek.

Ik heb gezegd.

References

- Berry, D.A. (1991), Inferences using DNA profiling in forensic identification and paternity cases, *Statist. Sc.* 6, 175-205.
- Bhattacharya, P.K., H. Chernoff and S.S. Yang (1983), Nonparametric estimation of the slope of a truncated regression, *Ann. Statist.* 11, 505-514.
- Bickel, P.J. and E.L. Lehmann (1979), Descriptive statistics for nonparametric models IV. Spread, *Contributions to Statistics*, J. Jureckova (ed.), Academia, Prague, 33-40.
- Bickel, P.J., C.A.J. Klaassen, Y. Ritov and J.A. Wellner (1992), *Efficient and Adaptive Estimation for Semiparametric Models*, The John Hopkins University Press, Baltimore.
- Born, M. (1949), *Natural Philosophy of Cause and Chance*, Oxford University Press, London.
- Brouwer, L.E.J. (1912), *Intuitionisme en Formalisme*, Clausen, Amsterdam.
- Freedman, D.A. and H. Zeisel (1988), From mouse-to-man: the quantitative assessment of cancer risks, *Statist. Sc.* 3, 3-56.
- Fusaro, R.E., N.P. Jewell, W.W. Hauck, D.C. Heilbron, J.D. Kalbfleisch, J.M. Neuhaus and M.A. Ashby (1989), An annotated bibliography of quantitative methodology relating to the AIDS epidemic, *Statist. Sc.* 4, 264-281.
- Hájek, J. (1970), A characterization of limiting distributions of regular estimates, *Z. Wahrsch. Verw. Gebiete* 14, 323-330.
- Hájek, J. (1972), Local asymptotic minimax and admissibility in estimation, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, L.M. Le Cam, J. Neyman, E.L. Scott (eds.), University of California Press, Berkeley, 1, 175-194.
- Kendall, M.G. and A. Stuart (1961), *The advanced Theory of Statistics*, Vol. 2, Griffin, London.
- Klaassen, C.A.J. (1984), Location estimators and spread, *Ann. Statist.* 12, 311-321.
- Klaassen, C.A.J. (1985), Strong unimodality, *Adv. Appl. Prob.* 17, 905-907.
- Klaassen, C.A.J. (1989a), Estimators and spread, *Ann. Statist.* 17, 859-867.
- Klaassen, C.A.J. (1989b), The asymptotic spread of estimators, *J. Statist. Plann. Inference* 23, 267-285.
- Klaassen, M., M. Kersten and B.J. Ens (1990), Energetic requirements for maintenance and premigratory body mass gain of waders wintering in Africa, *Ardea* 78, 209-220.
- Le Cam, L. (1953), On some asymptotic properties of maximum likelihood estimates and related Bayes estimates, *Univ. of California Publ. in Statistics* 1, 277-330.
- Le Cam, L. (1972), Limits of experiments, *Proc. 6th Berkeley Symp. Math. Statist. Probab.* 1, 245-261.
- Le Cam, L. (1986), *Asymptotic Methods in Statistical Decision Theory*, Springer-Verlag, New York.
- Marcuse, H. (1964), *One Dimensional Man. Studies in the Ideology of Advanced Industrial Society*, Beacon Press, Boston.
- Panel on Nonstandard Mixtures of Distributions (1989), Statistical models and analysis in auditing, *Statist. Sc.* 4, 2-33.

- Speed, T.P. (1985), Probabilistic risk assessment in the nuclear industry: WASH-1400 and beyond, *Proc. Berkeley Conf. Honor Jerzy Neyman and Jack Kiefer*, vol. 1, L.M. Le Cam and R.A. Olshen (eds.), Wadsworth, Belmont, 173-200.
- Van der Vaart, A.W. (1989), On the asymptotic information bound, *Ann. Statist.* 17, 1487-1500.
- Utts, J. (1991), Replication and meta-analysis in parapsychology, *Statist. Sc.* 6, 363-403.
- Wald, A. (1950), *Statistical Decision Functions*, Wiley, New York.
- Wolfowitz, J. (1965), Asymptotic efficiency of the maximum likelihood estimator, *Theory Probab. Appl.* 10, 247-260.

