

HOOFDSTUK 5. MODELLEN, METEN EN ONDERZOEK

5.1 EMPIRISCHE CYCLI

Iedere vorm van empirische kennisvermeerdering bestaat uit een combinatie van waarnemen, daarover uitspraken doen, daaruit andere uitspraken afleiden en deze vervolgens toetsen aan de hand van waarneming. (De Groot, 1970, blz. 1-18, zet dit uitvoerig uiteen.) Kennisvermeerdering heeft dus een cyclisch karakter. In zo'n empirische cyclus zijn kennis en waarneming nauw aan elkaar verbonden, ook in de zin dat de kwaliteit van waarnemingen door kennis wordt bevorderd. Het volgende voorbeeld (dat ik ontleen aan Madge, 1953) moge dat illustreren. Twee personen staan op het strand en kijken naar de horizon. Een van hen merkt op: "Ik zie drie stippen in de verte". De ander, een kustbewoner, antwoordt: "Ik zie drie schepen".

Ook kennisvermeerdering als wetenschappelijke activiteit heeft dit cyclische karakter. De gedachte is nu, dat wanneer de activiteiten die onderdeel uitmaken van van deze cyclus aan bepaalde nauwkeurig omschreven eisen voldoen, dan ook het resultaat van die activiteiten - kennis - zo goed mogelijk is. De Groot (1970, blz. 29) formuleert de cyclus als volgt:

Fase 1. Observatie: Verzamelen en groeperen van empirisch feitenmateriaal.

Fase 2. Inductie: Formulering van hypothesen.

Fase 3. Deductie: Afleiding van speciale consequenties uit de hypothesen, in de vorm van toetsbare voorspellingen.

Fase 4. Toetsing: van de hypothesen, aan het al dan niet uitkomen van de voorspellingen in nieuw empirisch materiaal.

Fase 5. Evaluatie: van de uitkomsten van de toetsing, in verband met de gestelde hypothesen, c.q. theorie(en), en in verband met mogelijke nieuwe, aansluitende onderzoekingen.

5.2 SOORTEN ONDERZOEK

In het licht van de hierboven weergegeven empirische cyclus kan men

(fundamenteel) onderzoek noemen: alle aan bepaalde eisen (o.a. van falsificeerbaarheid en korrekte denkprocessen) voldoende bijdragen aan de op een bepaald wetenschapsterrein bestaande empirische cycli.

Omdat iedere fase van de empirische cyclus een verschillende doelstelling heeft, kan men ook verschillende soorten onderzoek onderscheiden. De Groot (1970, blz. 316) noemt:

1. Descriptief onderzoek.

Hierbij wordt, volgens een bepaalde expliciet aangegeven methodiek, geregistreerd wat zich voordoet op een bepaald gebied, zonder dat anders dan door de principes van die methodiek zelf op theorie- of hypothesevorming wordt vooruitgelopen.

2. Exploratief onderzoek.

Bij de opzet hiervan wordt uitdrukkelijk gedacht aan de vorming of uitwerking van een theorie of van afzonderlijke hypothesen. Het gaat niet zozeer, althans niet alleen, om verzameling en ordening van feiten, maar ook om verwachte en te vinden samenhangen, die voor een bepaald doel relevant worden geacht. Doordat een scherpe formulering van die samenhangen, als toetsbare hypothesen, ontbreekt, kunnen zij echter nog niet volgens de regelen van de kunst worden getoetst.

3. Toetsingsonderzoek.

Hierbij wordt een beperkt aantal (soms een), gewoonlijk aan een theorie ontleende, onderling samenhangende hypothesen aan empirisch materiaal getoetst. Meestal betreft het een steekproefonderzoek (i.t.t. een universumonderzoek) dat vaak onder experimentele condities wordt uitgevoerd. (Voor verklaring van enkele in deze beschrijving gebruikte termen: zie hieronder.)

4. Interpretatief-theoretische studies.

Kenmerkend is hierbij, dat een of meer meetresultaten-verzamelingen (die niet als onderdeel van ditzelfde onderzoek vervaardigd zijn) in onderling verband gebracht worden door ze proberenderwijs af te leiden uit een hypothese of theorie, opvatting of visie, die door de onderzoeker op het gegeven materiaal van toepassing wordt geacht. Een mogelijk resultaat is: een alternatieve theorie of hypothese.

5. Instrumenteel-nomologisch onderzoek.

Kenmerkend is, dat de onderzoeker binnen een nauw terrein van een of meer hypothesen rond een bepaald begrip aan een meetinstrument, een meetprocedure werkt.

Het hierboven gegeven onderscheid tussen verschillende soorten onderzoek heeft in beginsel alleen betrekking op fundamenteel (empirisch) wetenschappelijk onderzoek, omdat alleen daar voor "alle" op een bepaald terrein werkzame onderzoekers een gemeenschappelijke empirische cyclus geldt.

Op het terrein van niet fundamenteel onderzoek zijn te onderscheiden onderzoeken waarbij de methodische en technische eisen van falsificeerbaarheid en korrekte denkprocessen zijn nageleefd, d.i. toegepast-wetenschappelijk onderzoek, en overig onderzoek. Zoals gezegd is niet fundamenteel wetenschappelijk onderzoek dikwijls gericht op het kunnen sturen van verschijnselen in de ervaringswereld. Betreft het onderzoek naar menselijk gedrag, dan kan gesproken worden van evaluatie- of beleidsgericht onderzoek. Zulk onderzoek kan allerlei verschillende doelen hebben. Zo kan men een gang van zaken willen beoordelen, maar ook de subjecten, die bij die gang van zaken betrokken zijn. Afhankelijk van die doelen kan men van de verschillende soorten onderzoek een ad-hoc indeling maken. Bijvoorbeeld, in het onderwijs kan men studietoetsen (zie De Groot, Van Naerssen e.a., 1969) afnemen, die gericht kunnen zijn op het beoordelen van studenten, maar ook op het beoordelen van de onderwijsactiviteiten als geheel (dus o.a. met aandacht voor de effectiviteit van de docenten). Beide soorten evaluatie-onderzoek kan men ook weer voor verschillende doelen gebruiken b.v. om studenten (docenten) informatie te geven over de doelmatigheid van hun inspanningen ("feed-back"), of om beslissingen te nemen over studenten (of docenten) ("selectie").

Ad-hoc indelingen kan men niet alleen maken aan de hand van de doelstelling van onderzoeken maar ook aan de hand van de gebruikte methode. (B.v. Adams, 1975, blz. 48.)

Ook al is evaluatie-onderzoek (per definitie) niet in de eerste plaats gericht op empirische kennisvermeerdering, dit ontslaat onderzoekers niet van de verplichting zo goed mogelijk aan de eisen die met dat doel samenhangen te voldoen. De Groot (1967) houdt een krachtig pleidooi voor een betere wetenschappelijke basis van beoordelingen in het onderwijs, in het bijzonder van selectiebeslissingen. N.m.m. geldt dit betoog mutatis mutandis in even sterke mate voor juridische beslissingen, i.h.b. indien deze beslissingen een selectiekarakter hebben.

Een onderscheid dat door sommigen zowel gebruikt wordt t.a.v. fundamenteel als niet fundamenteel onderzoek is de tweedeling "kwalitatief" vs. "kwantitatief" onderzoek. Dit onderscheid heeft soms betrekking op de gehanteerde meetschalen, soms op de grootte van de in een onderzoek gehanteerde steekproef en soms op nog iets anders. In de meeste gevallen is erg onduidelijk naar welke begrippen met de termen "kwalitatief" en "kwantitatief" in dit verband wordt verwezen.

5.3 OPZET VAN EEN ONDERZOEK

De soorten onderzoek, die in de vorige paragraaf zijn onderscheiden hebben allerlei verschillende doelstellingen, of anders allerlei verschillende methodieken. Dat neemt niet weg, dat men bij de opzet van een onderzoek bepaalde beslissingen toch altijd wel moet nemen. Aan Verduijn (1970, blz. 99) ontleen ik het volgende basisschema voor het verloop van een onderzoek (enigszins gewijzigd i.v.m. verschillende terminologie).

HET VERLOOP VAN EEN ONDERZOEK

1. Voorbereidingsfase

1. Literatuurstudie
2. Maken onderzoeksplan:
 1. Ontwerpen van het onderzoek (principiele beslissingen over):
 1. probleemstelling + hypothesen
 2. begripsvorming + indicering
 3. typen van onderzoek
 4. waarnemingseenheden
 5. waarnemingstechnieken
 6. ordening + analyse
 2. Operationele uitwerking ontwerp + organisatorische opzet

2. Uitvoeringsfase

- Het doen van de metingen

3. Uitwerkingsfase

1. De verwerking van de meetresultaten, b.v. het onderbrengen in een nadere classificatie, sorteren, tellen en tabelleren
2. Het analyseren en interpreteren
3. Het rapporteren

De beslissingen die uit het hier geschetste verloop voortvloeien kunnen onderscheiden worden in

1. Principiële, d.w.z. methodische beslissingen. Dit betreft onder meer de vragen:
 - Hoe luidt de probleemstelling?
 - Welk type van onderzoek?
 - Welke wijze van meten?
 - Welke wijze van verwerken en interpreteren?
2. Operationele, d.w.z. technische beslissingen. De omzetting van de methodische beslissingen in praktische termen van onderzoek.
3. Organisatorische beslissingen.

Het onderscheiden van deze drie soorten beslissingen houdt niet in, dat ze alle gescheiden van elkaar genomen zouden moeten worden. De bedoeling is slechts aan te geven, welke aspecten in ieder geval aan de orde moeten komen. (Vgl. Verduyn, t.a.p. blz. 98-100)

5.4 ONDERZOEK EN METEN

Onderdeel van ieder (empirisch) onderzoek vormt dus het meten. Met het uitvoeren van de meting kan soms het gehele onderzoek reeds praktisch verricht zijn maar soms vertegenwoordigt het slechts een zeer klein deel van een onderzoek.

Enkele manieren om meetresultaten te verkrijgen zijn (vgl. b.v. Webb e.a. 1970):

- observatie, al dan niet met behulp van instrumenten
- bestudering van uitspraken, b.v.
 - inhoudsanalyse
 - secundaire analyse
 - soms worden speciale technieken gebruikt om uitspraken te verkrijgen b.v. enquetes en vrije interviews

Is het voor interpretatief-theoretisch onderzoek zo, dat de meetresultaten niet door meting, verricht door de onderzoeker of het onderzoekteam zelf verkregen worden, ook voor andere soorten onderzoek is dit soms niet noodzakelijk. Wanneer men kan beschikken over betrouwbare en voor het specifieke doel van het onderzoek valide meetresultaten behoeven er geen metingen te worden verricht. Een voorbeeld vormt secundaire analyse, waarbij men de meetresultaten die in een ander onderzoek zijn verzameld, opnieuw, al dan niet op andere wijze, verwerkt.

Een manier om uitspraken te verkrijgen die voor onderzoekers op juridisch terrein van bijzonder belang is, wordt gevormd door de officiële (justitiële e.d.) statistieken. Geldigheid en betrouwbaarheid van zulke uitspraken zijn dikwijls gering. Voor sommige problemen die rijzen wanneer men ze niettemin wenst te gebruiken wordt verwezen naar Janswoude en De Mulder (1980).

5.5 ONDERZOEK EN MODELLEN

Bezieet men het in hoofdstuk 2 weergegeven (aan Lave en March ontleende) model van het modelbouwproces, dan valt de grote overeenkomst met de empirische cyclus in wetenschappen op. Gesteld kan worden dat de modelmatige benadering van Lave en March inhoudt dat in de loop van een onderzoek tenminste een gehele empirische cyclus wordt doorlopen. Dit geeft, vergeleken met de situatie dat een onderzoek niet meer dan een fase van een empirische cyclus beslaat, een grotere flexibiliteit aan onderzoekers, die met name op nieuwe, of nog weinig ontwikkelde terreinen van wetenschap voordelig kan zijn. (In hoofdstuk 13, wanneer o.m. de opvattingen van Kuhn (1970) behandeld worden, wordt hierop nader ingegaan.)

Het vervaardigen van een model volgens Lave en March kan dus een hypothese opleveren, of (als wij onder een theorie een verzameling onderling samenhangende en empirisch getoetste hypothesen verstaan) een theorie, maar het resultaat kan ook van aanzienlijk mindere - en een enkele keer ook van verstrekkender - betekenis zijn. Het begrip "model" is dus veel ruimer, heeft een grotere omvang (dus minder inhoud) dan "theorie" en "hypothese".

5.6 HET BELANG VAN STATISTIEK

Volgens Wijvekate (1979, blz. 16) is het de taak van de statistiek er voor te zorgen dat representatieve steekproeven uit goed gedefinieerde (statistische) verzamelingen worden getrokken, dat deze feitelijke gegevens doeltreffend worden samengevat en verwerkt en dat hieruit zo nauwkeu-

rig mogelijk konklusies omtrent de verzameling worden getrokken met beperkte en bekende risico's. Hieruit blijkt dat de statistiek in twee fasen uiteen valt. De eerste fase is die van de vastlegging en verwerking van de feiten; dit is de "beschrijvende statistiek".

Om uit het resultaat hiervan konklusies te kunnen trekken, heeft men de kansrekening nodig, waardoor onder meer de rol van het toeval bij de totstandkoming van steekproeven wordt ontsluitend.

Men onderscheidt (vgl. V.d. Ende, 1971, blz. XIII) "steekproef"- en "populatieonderzoek". ("Populatie" en "universum" worden, vgl. De Groot, 1970, blz. 194, dikwijls door elkaar gebruikt.)

De "populatie" is de verzameling eenheden die voor het onderzoek van belang is en waarover men door middel van het onderzoek een uitspraak wil doen.

Die eenheden kunnen in gedragswetenschappelijk onderzoek b.v. mensen zijn, en in natuurwetenschappelijk b.v. voorwerpen. Ook echter kan men eigenschappen van die eenheden, of relaties daartussen de populatie laten uitmaken. (Sommigen verkiezen in dit laatste geval de term "universum".)

Lang niet altijd beschikt men over waarnemingen m.b.t. alle eenheden uit de populatie, maar slechts van een deel daaruit: een "steekproef". Voor het trekken van geldige konklusies, of het maken van de juiste schattingen m.b.t. de populatie is het handig als een steekproef representatief (of "aselekt") is.

Een steekproef wordt t.a.v. een bepaalde variabele "representatief" genoemd, wanneer die variabele niet onder- of oververtegenwoordigd is t.o.v. de populatie.

De tweede fase van de statistiek, de methoden en technieken die men gebruikt om konklusies te trekken uit het beschrijvend materiaal wordt "verklarende" of "inferentiële" (ook wel "generaliserende" of "inductieve") statistiek genoemd.

Een voorbeeld van toepassing van verklarende statistiek, en wel van een toetsing van een hypothese, geeft Wijvekate (1979, blz. 207). De toetsingsprocedure betreft de vraag, of het plaatsen op een bepaalde weg van verkeersborden die een beperking van de maximumsnelheid voorschrijven het aantal ongevallen heeft verminderd. Men beschikt dan b.v. over meetresultaten van het aantal ongevallen voor en het aantal na de invoering van de snelheidslimiet. Stel, dat de hypothese die men wil bevestigen luidt, dat het aantal ongevallen is verminderd. Daartoe tracht men de "nulhypothese", nl. dat het aantal ongevallen ongewijzigd is gebleven, te verwerpen.

Twee soorten risico zijn daaraan verbonden. "Fouten van de eerste soort" zijn: de nulhypothese ten onrechte afwijzen. Een "fout van de tweede soort" is: de nulhypothese ten onrechte aanvaarden. (Omdat niet wordt verondersteld dat het aantal verkeersongevallen ook kan zijn toegenomen, spreekt men in dit geval van "eenzijdig toetsen", i.t.t. "tweezijdig toetsen".)

Niet iedere verandering, hoe klein ook, leidt tot de konklusie, dat de nulhypothese verworpen moet worden. Aan de hand van het soort toetsingsprobleem en de beschikbare gegevens, wordt een bepaalde toetsingsgrootte gekozen en berekend. Aan de hand hiervan kan worden nagegaan, in hoeverre het nieuwe aantal significant afwijkt van de "norm" (= gelijk blijven). Hanteert men een "significantiegebied" van 5% dan luidt de bewering: de kans, dat de gevonden afwijking toevallig is (dus dat de nulhypothese ten onrechte verworpen wordt), is kleiner dan 5%.

Tot slot van dit summiere voorbeeld zij opgemerkt, dat met de constatering, dat het aantal ongevallen significant minder is geworden nog niet is vastgesteld, dat die afname nu ook is veroorzaakt door de snelheidsbeperking. De aanwezigheid van een oorzakelijk verband is in het algemeen moeilijk aan te tonen. (Vgl. Dessaur, 1971, blz. 70.)

Kendall en Stuart (1969, blz. 2) maken het onderscheid tussen verklarende en beschrijvende statistiek niet. Zij definiëren:

Statistiek is de tak van wetenschappelijke methode die zich bezighoudt met data verkregen door tellen of meten van eigenschappen van populaties van verschijnselen in de ervaringswereld.

Betrekkelijk nieuw en nog enigszins omstreden is de Bayesiaanse (genoemd naar Bayes) statistiek, waarbij a priori (= vooraf) opvattingen over waarschijnlijkheden worden geformuleerd, die m.b.v. data in a-posteriori (= achteraf) opvattingen worden omgerekend. Ook voor hypothese toetsing kan deze benadering worden gebruikt: de Bayesiaanse statisticus zal weer a-priori opvattingen over de juistheid van de hypothese formuleren, en de beschikbare data gebruiken om deze opvattingen bij te stellen. Soms beschikt men helaas niet over zinvolle a-priori opvattingen, hetgeen door sommigen als een zwakte van de methode wordt beschouwd. (Zie Phillips, 1973, blz. 5 en Kendall en Stuart, 1969, blz. 198)

Deelt men de definitie van "statistiek" van Kendall en Stuart dan kan geconcludeerd worden, dat bij het vervaardigen van modellen volgens Lave en March, en bij het verrichten van (empirisch) onderzoek de statistiek vrijwel altijd een rol speelt.

Het gaat buiten het bestek van dit boek, een overzicht te willen geven van statistische technieken. Men zie daarvoor de aangehaalde literatuur.

N.m.m. (en in aansluiting op Muilwijk, 1978) dienen juristen, die zich bezighouden met modelbouw en/of empirisch onderzoek, begrip te hebben van de grondbeginselen van de statistiek (begrip te hebben van meten en meetschalen, frequentie- en kansverdelingen, karakteristieke maten van verdelingen, associatiematen, iets van multivariate analyse, voorts steekproeftechniek, schatting, toetsing en zekere prognose-methodieken.) Wanneer voor een bepaald onderzoek een meer precieze kennis nodig is van een bepaalde techniek, dan is het althans mogelijk deze uit een handboek te verkrijgen, dan wel een statisticus te raadplegen.

Tenslotte merk ik op dat tegenwoordig geen enkele jurist meer zonder de globale statistische kennis kan, die nodig is voor het lezen van onderzoeksrapporten en andere publikaties met statistisch materiaal.