DE EVOLUTIE VAN DE STATISTIEK EN HAAR TOEPASSINGEN

door H. FLORIN,
Katholieke Universiteit te Leuven.

De statistiek telt zowel warme vrienden als gezworen vijanden. Terwijl zij aan de ene kant zoveel gegevens levert, dat zonder haar het bestuur van een modern land ondenkbaar zou zijn, is het aan de andere kant ook waar, dat het gemakkelijk is, vertrokken beelden van de werkelijkheid te krijgen : de statistische leugens.

Ook worden aan de statistiek vaak vragen gesteld, waarvoor zij niet is gemaakt, bij voorbeeld als zij wordt gebruikt bij beroepskeuze met persoonlijkheidstests, e.d. zoals Gibbon al zegde : de uitkomsten van de statistiek zijn even waar in het algemeen als bedrieglijk in het bijzonder [1].

Maar waarvoor is de statistiek dan wel gemaakt ? Niet ieder heeft hierover dezelfde mening. Wij zijn hier allen vrienden van de statistiek, wij mogen ze allen beschouwen als een kostbare edelsteen, maar wij zien ze door verschillende facetten, en ik bied U van te voren mijn verontschuldigingen aan, als ik juist die facetten, die uw bijzondere liefde hebben zou vergeten hebben.

Inderdaad, wanneer we ons bezinnen over de definitie van Statistiek, zien we ons voor een moeilijke taak gesteld. Wij bemerken de meest uiteenlopende definities bij de verschillende schrijvers. Sommigen leggen de nadruk op : studie van massaverschijnselen — maar ook van zeldzame verschijnselen maakt men statistieken. Sommigen keren zich naar de wetmatigheid, die vele statistische verschijnselen kenmerkt — maar ook van onwetmatige verschijnselen maakt men statistieken. Anderen weer zeggen, dat statistiek zich bezighoudt met quantitatief meetbare feiten, die voor een belangrijk deel afhangen van een groot aantal oorzaken — maar b.v. gezien in het licht van de quantenstatistiek levert het begrip oorzaak hier moeilijkheden. Zelfs het woord « quantitatief » kan aangevallen worden : reeds wellicht de eerste statistische waarneming door mensen verricht was eerst qualitatief, n.l. de opvolging van dag en nacht.

Reeds aan dit voorbeeld zien we de rol van de statistiek vanaf de vroegste tijden : het vaststellen van de opvolging van dag en nacht, die van

---

[1] aangehaald in Keynes, A Treatise of Probability, Macmillan (1921).

zomer en winter had niet alleen wat men nu zou noemen « wetenschappelijke » waarde, het had een zeer groot economisch belang voor het vaststellen van zaai- en oogsttijden bij voorbeeld.

En zo is het gebleven : de statistiek is enerzijds een wetenschap, anderzijds en vooral een bescheiden en gewillige dienares van de mensen. Bescheiden echter is veel gezegd : evenals de wiskunde gaat zij wel met getallen om, maar welk een verschil. De zuivere wiskunde, ver boven elke toepassing verheven, noemt zich de scientiarum genetrix — en daar tegenover de statistiek — een uiterst bemoeizieke dame, die zich onbescheiden overal indringt, en alles van U wil weten : hoeveel kinderen U hebt, hoe dikwijls U vlees eet, of U een badkamer hebt.

Wij zijn hiermee afgedwaald van het zoeken naar een definitie voor statistiek — of hebben we misschien door de laatste opmerkingen beter dan door geleerde omschrijvingen haar aard blootgelegd. Hoe het zij, wij laten het bij de vaststelling, dat bijna iedere schrijver er een eigen standpunt op nahoudt, doordat hij de nadruk legt op een der vele aspekten, die het begrip omvat.

Over de *afleiding* van het *woord* statistiek is meer eenheid. Vrijwel alle schrijvers vinden de oorsprong in het Latijnse woord « status », in de betekenis van « staat », rijk, en wijzen naar Achenwall als eerste gebruiker. Deze schreef in 1749 een boek over staatswetenschap [2], waarin dit begrip reeds in vrij moderne betekenis gebruikt werd, al heeft sindsdien het quantitatieve aspekt het meer beschrijvende verdrongen. Onder Duitse invloed kwam het woord reeds na enkele tientallen jaren ook in Engeland in gebruik, waar het gemakkelijk ingang vond, omdat daar het woord « statist » al door Shakespeare in « Hamlet » was gebruikt rond 1600 [3]. In « Hamlet » zou men kunnen twijfelen aan de betekenis van die term bij Shakespeare, maar hij gebruikt hem ook in « Cymbeline » [4] en daar is wel duidelijk, dat « statist » voor Shakespeare nog « staatsman » betekende. Tegenwoordig is het verband met het begrip « staat » verloren gegaan : men denke aan begrippen als statistische mechanica.

Als men moet spreken over de evolutie van de statistiek dient men enigszins historisch te werk te gaan. Ik hoorde enige dagen geleden klagen, dat als iemand zich voorneemt een bepaald onderwerp te schilderen, hij vaak de fout maakt te beginnen bij de Grieken en te eindigen bij Napoleon. Ik zal dan trachten te beginnen vóór de Grieken en verder te komen dan Napo-

---

[2] G. Achenwall, Abriss der Staatswissenschaft der Europäischen Reiche, (1749).

[3] Hamelt, 5e akte, 2e toneel (ongeveer 1600).

[4] Cymbeline, 2e akte, 4e toneel, 16e vers (ongeveer 1610).

leon. Dit laatste zal gemakkelijk zijn, want toen begon pas de wetenschappelijke statistiek met Quetelet.

In de oudste tijden vinden we overal en in alle werelddelen, dat de ontluiking van de beschaving met statistiek gepaard ging : in Amerika b.v. de volkstellingen bij de Incas ; bij de Egyptenaren vindt men dat reeds ten tijde van de derde dynastie, dus bijna 3000 jaar voor Christus koning Snefru optekende, dat hij 7000 gevangenen en 200 000 runderen op de Nubiërs veroverd had.

De bijbel ook vermeldt statistische gegevens, maar het was in die tijden, niet zonder gevaar, een volkstelling te organiseren : de bijbel verhaalt, dat David als straf voor een volkstelling moest kiezen tussen 3 jaar hongersnood, 3 maanden vluchten voor vijanden of 3 dagen pest. Hij koos het laatste, maar was niet genezen van zijn passie voor de statistiek, want hij telde de doden : 70 000.

Onder keizer Augustus was een volkstelling er de oorzaak van dat de geboorte van Christus te Bethlehem plaatsgreep. Gouraud meent bij de Romeinen sporen van sterftetafels te vinden. In moderne tijdens vindt men sterftelijsten in Londen vanaf 1592, deze werden gebruikt in een eerste poging tot het verkrijgen van sterftetafels door Graunt.

Dan komt het belangrijke werk van de Amsterdamse burgemeester en wiskundige Van Hudden en vooral in 1671 het beroemde werk van Raadpensionaris Jan de Witt over sterftetafels en lijfrenten, een werk dat tevergeefs werd gezocht door Leibniz en door Jacob en de oudere Johan Bernouilli, maar dat men nu weer kent.

Toen pas kwam het meer bekende werk van Halley over sterftetafels in 1693, dat vaak ten onrechte het eerste op dit gebied wordt genoemd. Bij hem vindt men zelfs reeds lijfrenten op 2 levens.

In de 18e eeuw vinden we verder werk op dit gebied. Namen als de Moivre (1718) : actuarieel werk, Euler (1760) : sterfte, annuïteiten, levensverzekeringen moeten hier wel in de eerste plaats genoemd worden zonder de Bernoullis te vergeten [Jacob, grootvader en kleinzoon Johan, Nicolaas, en in dit verband misschien vooral Daniël (1760)] met zijn werk over geboorten van jongens en meisjes, en ten slotte rond 1800 Laplace b.v. over de duur van huwelijken en over pokken.

Naast al dit werk, dat sterftetafels en dergelijke problemen betrof, ontwikkelde zich in de 17e en 18e eeuw een zeer intensieve belangstelling voor de waarschijnlijkheidsleer. Hierin werd vooral een rol gespeeld, behalve door de reeds genoemde persoonlijkheden, door Pascal en Fermet (1654), d'Alembert (rond 1760), Lagrange (1770). En niet te vergeten Bayes (1764-1765) met zijn fameus theorema over waarschijnlijkheden van oorzaken.

Deze waarschijnlijkheidsleer had oorspronkelijk betrekking op weinig belangrijke zaken als kansspelen, om te beginnen met de bekende speler, de Chevalier de Méré, die Pascal en Fermat problemen leverde, maar meer en meer werd de waarschijnlijkheidsrekening de wetenschap, die de statistiek haar wiskundige grondslag zou geven. Dit gebeurde enerzijds door de Moivre, Jacob Bernoulli, Laplace en Gauss, die de twee voornaamste frequentiefuncties : de normale verdeling en de Bernoulliverdeling leverden en zelfs hun samenhang, en anderzijds de theorie van de kleinste kwadraten door Legendre, Gauss en Laplace.

De twee lijnen, die we bespraken : de oude statistiek, die bestond uit een louter registreren van feiten, en : de toepassing van wiskundige hulpmiddelen, die de moderne statistiek kenmerkt, convergeren in een figuur, een nationale figuur : de grote Belg Adolphe Quetelet (1796-1874), die door velen, zelfs door de kritische Keynes [5] de vader van de moderne statistische methode wordt genoemd.

Quetelet was een veelzijdig man : bekwaam meetkundige, statisticus, sterrekundige, meteorologe, climatoloog, onderzoeker van het aardmagnetisme en dichter.

Aan zijn stuwkracht is onder meer te danken de oprichting van het Belgisch Observatorium.

Deze man was getroffen door de wetmatigheden, die in de natuur optreden en die door statistisch onderzoek tot hun recht komen ; hij was de uitvinder van de uitdrukking « gemiddelde mens » (homme moyen, want hij drukte zich in het Frans uit, al scheen deze Gentenaar ook Nederlands te kennen, althans hij vertaalde o.m. gedichten van Tollens [6]). De uitdrukking « gemiddelde mens » heeft in de loop der tijden veel stof doen opwaaien. Is de gemiddelde mens een ideaal wezen of een monster ? De moderne statisticus zal zich wellicht verbazen over de passie, waarmee men zich over deze zaken opwond : gemiddelden zijn ons zo vertrouwd geworden, dat we weten, hoe veel en hoe weinig wij er van verwachten kunnen.

Quetelet onderzocht niet alleen physische, maar ook morele eigenschappen. Vooral dit laatste was nieuw en hij werd getroffen door het feit, dat er niet alleen een regelmaat bestaat in de gemiddelde lengte, het gemiddelde gewicht, maar ook in morele eigenschappen, b.v. in het voorkomen van misdaden. Hij sprak van een « schrikbarende wetmatigheid » [7]. Hij zag een

---

[5] Keynes, *l.c.*, p. 334 : « the parent of modern statistical method ».

[6] E. Mailly, Adolphe Quetelet. Biographie, p. 15. Brussel 1874.

[7] J. Belfroid, Les fastes de la mathématique belge. Etudes classiques VIII - 1939, p. 374.

sterke parallel in de regelmaat van alles, wat betrekking heeft op de mens, met de wetten van de physica. Een zijner werken hierover noemde hij dan ook « Essai de physique sociale » (1835).

Er wordt hem tegenwoordig wel eens verweten, te veel de nadruk te hebben gelegd op het constant zijn van dergelijke gemiddelden, en vooral zijn enthousiasme voor deze ontdekking en voor alle perspectieven, die de statistiek biedt nemen sommigen hem kwalijk. Men kan dit begrijpen van een man als Keynes, die zelfs uitdrukkingen als Wet van de Grote Getallen (de uitdrukking schijnt afkomstig te zijn van Poisson) te romantisch vindt. Keynes verwijt hem bijna zijn dichterschap : hij vreest, dat de statistiek nog niet salonfähig is en is beducht voor een smaakje van alchemie, van astrologie [8]. Maar men moet zich in de tijd van Quetelet verplaatsen en hem dankbaar zijn voor zijn enthousiasme, waardoor deze nieuwe denkrichting bekend is geworden. Het is normaal en wellicht wenselijk, dat een ontdekker van nieuwe principes in zwart-wit ziet en desnoods overdrijft. De wetenschap kent hiervan vele voorbeelden, denken wij maar aan Freud bij het ontstaan van de psycho-analyse.

Aan de geestdrift van Quetelet danken wij onder meer zijn sterftetafel voor Brussel in 1825, en voor geheel België op grond van de volkstelling van 1830.

In 1833 heeft zijn invloed de instelling ten gevolge gehad van een sectie voor statistisch onderzoek van de Britse Associatie voor wetenschappelijk onderzoek. Wij zien dan reeds in 1834 de oprichting tot stand komen van de Royal Statistical Society, en in 1853 kwam onder de leiding van Quetelet het eerste Internationale Congres voor de Statistiek tot stand, dat werd gevolgd door vele andere in de meest verschillende plaatsen.

De verwijten van overdrijving, die men soms aan Quetelet richt zijn daarenboven onverdiend. Hij was de eerste, die wees op de moeilijkheid, objectieve statistische gegevens te krijgen, en er op aandrong, af te zien van zulke, die vervalst kunnen zijn door tegenzin of fraude [9]. Dit is meer, dan men van menig tegenwoordig statisticus kan zeggen.

Hij wordt door Keynes [10] ook de eerste genoemd, die er de aandacht op vestigde, dat de foutenkromme slechts op een beperkt aantal phenomenen kan worden toegepast, dus geenszins een natuurwet is.

---

8) Keynes, *l.c.*, p. 335.

9) Mailly, *l.c.*, p. 22.

10) Keynes, *l.c.*, p. 208.

Men heeft hier de gewone moeilijkheid, die optreedt bij de toepassing van de wiskunde op de praktijk : de wiskunde bedient zich van scherpe formuleringen, uitgaande van nauwkeurige hypothesen, zij dankt haar nauwkeurigheid aan haar armoede. De werkelijkheid is veel komplexer, zij kan er dus geen volledig beeld van geven. Poincaré heeft eens gezegd : de wiskunde is de kunst, aan dingen die verschillen, dezelfde naam te geven. Dit is een zeer diep woord. Denken wij aan het begrip getal : oorspronkelijk : natuurlijk getal, later ook gebroken getal, nog later, n.l. 450 voor Christus ook irrationeel getal ; rond 1600 ook negatief getal, nog in de 17e eeuw ook komplex getal, in de 19e eeuw ook quaternionen en transfiniete getallen.

Ook in de statistiek worden zeer verschillende dingen in hetzelfde wiskundige schema gewrongen. Dit brengt altijd een vereenvoudiging mee, en het is de grote kunst, die zo te kiezen, dat de uitkomst bruikbaar is, dus een voldoende benadering van de werkelijkheid. Bijzonder moeilijk is dit, wanneer men zo ingewikkelde phenomenen wil beschrijven, als voorkomen bij levende wezens.

Quetelet heeft, met de gemiddelden en met de normale verdeling een eerste poging hiertoe gedaan, en was zich, zoals wij boven aantoonden, bewust van de begrenzingen van de methode.

Het is hier wellicht reeds de plaats, om te vermelden, hoe men in deze geest verder is gegaan tegenwoordig met de verfijnde steekproefmethodes, waar men grotere nauwkeurigheid bereikt door stratificatie, dus door deelmonsters te nemen met meer verwante eigenschappen, door ratio-estimates die het mogelijk maken met vroeger verkregen inlichtingen rekening te houden, door cluster-sampling, die meer rationele indeling van de steekproeven tot stand brengt, door de multistage sampling, die deze processen nog verfijnt.

Speciaal in België, zoals het past in het land van Quetelet, is in de laatste jaren op dit gebied veel verricht : nadat een werk van Directeur-Generaal Dufrasne de aandacht had gevestigd op het steekproefonderzoek, werd een wet gestemd, die de statistische diensten machtigt tot het inrichten van steekproeven en nog onlangs gaf het Nationaal Instituut voor de Statistiek een studie uit van een steekproefonderzoek toegepast bij de Belgische statistiek betreffende de landbouwarbeidskrachten.

Maar laten wij terugkomen op de kritische houding, waarvan Quetelet reeds in zijn tijd blijk gaf ten opzichte van de normale frequentieverdeling, die toch een eeuw lang bijna onbeperkte heerseres is gebleven. Bekend is de uitspraak van Lippman, die door Poincaré gemeengoed is geworden : « Iedereen gelooft in de foutenkromme, de experimentatoren, omdat zij denken, dat

zij een wiskundig theorema is, en de wiskundigen, omdat zij denken, dat zij een experimenteel feit is ».

En inderdaad, sinds de Moivre een bijzonder geval bewees is op het gebied van de theorie een grote weg afgelegd : men denke aan het centrale limiettheorema voor gemiddelden.

De experimentatoren vonden inderdaad vele voorbeelden van toepasbaarheid, te beginnen met de verdeling van de lengte van de mensen. In vele gevallen zijn er zeker afwijkingen, in strenge zin natuurlijk altijd, maar zeer dikwijls zijn die afwijkingen niet zo groot, dat zij in de praktijk opwegen tegen het voordeel van bruikbaarheid van formules en tabellen. Dit wordt met een overstelpend materiaal bewezen door zovele onderzoekingen, vooral in het gebied van landbouw en industrie. Maar men gaf zich er meer en meer rekenschap van, dat de normale verdeling niet alles verklaart. Van 1837 reeds dateert b.v. de ontdekking van de Poissonverdeling, al werd die pas herontdekt door von Bortkiewicz omstreeks 1914 onder de naam van « wet van de kleine getallen ». In plaats van alleen het rekenkundig gemiddelde te beschouwen, onderzochten geleerden als Galton ook het geometrisch gemiddelde, wat Gumbel er toe bracht de lognormale verdeling naar Galton te noemen. Die geometrische gemiddelden hebben in de statistiek heel wat moeilijkheden gebracht. Zo bij de berekening van een gewogen indexcijfer, waar Irving Fisher ten onrechte invariantie er van eiste. Wagemann [11] heeft, om hem te weerleggen, een Oosters sprookje moeten bedenken : Een wijze tovenaar komt daar tot sombere voorspellingen over toekomstige opbrengsten, zijn toverfomules gebruiken geometrische gemiddelden.

Maar een wijze schaapherder met alleen maar gezond verstand stelt de koning gerust met gewogen rekenkundige gemiddelden.

Tegenwoordig zijn deze strijdpunten wel opgeklaard, en teruggebracht tot een transformatie van coördinaten : na een logarithmische transformatie is immers een meetkundig gemiddelde overgegaan in een rekenkundig. Dit neemt niet weg, dat deze transformatie met haar uitbreiding heeft gevoerd tot zeer uitgebreide toepassingen, in de biologie, pharmacie en geneeskunde ; wij wijzen vooral op de biologische assay, wat men misschien zou kunnen vertalen door doseringswetenschap, waarin statistisch de invloed van verschillende dosissen van stimuli op verschillende subjecten wordt gemeten, b.v. geneesmiddelen, rattenverdelgingsmiddelen, enz.

Men gebruikt in dit en verwante gebieden, behalve de logarithmische, tal van andere transformaties. Wij noemen de logistische transformatie bij groeikrommen, met de z.g. logits, ingevoerd door Berkson ; maar vooral de

---

11) E. Wagemann, Narrenspiegel der Statistiek, Hamburg 1935, p. 99.

zeer algemene *probit*transformatie, die willekeurige verdelingen terugbrengt tot de normale frequentiekromme. De naam probit, die probability unit betekent, is afkomstig van Bliss.

Maar reeds veel eerder had men de afwijkingen van de normale frekwentiekromme systematisch bestudeerd. Dit gebeurde bijzonder succesvol door Karl Pearson met zijn kriteria b_1 en b_2 voor de scheefheid en de welving van frequentiekrommen, die, wat wiskundigen als Wolfowitz er ook van zeggen, nog steeds van groot nut zijn. Pearson kwam met behulp er van tot zijn beroemde klassificatie van frequentiekrommen in een aantal typen, die de grote meerderheid omvatten van de in de praktijk voorkomende gevallen. Zijn theorie gaat slechts uit, en dat is het geniale er van, van een bijna simplistische discussie van de oplossingen van een zeer eenvoudige differentiaalvergelijking, waarbij hij steunt op momenten. Fisher [12] merkte wel op, dat de momentenmethode reeds teruggaat op Bessel en Gauss, maar dit doet weinig af aan de verdiensten van Pearson. Elderton [13] liet zien, met voorbeelden uit het actuariaat, hoe dikwijls ook in onverwachte maar *reële* gevallen, deze typen hun waarde behouden, b.v. bij J- en U-vormige frequentiekrommen.

Wij hebben terloops twee grote namen genoemd : Galton en Pearson, die wij niet zo kort kunnen voorbijgaan. Francis Galton (1822-1911) was geen wiskundige, maar anthropoloog en ook weerkundige. Men noemt hem de stichter van de eugenetica [14]. Door deze aktiviteiten ontstond zijn belangstelling voor de statistiek.

Aan hem dankt men de experimentele studie van de binomische verdeling, waartoe hij met Pearson toestellen samenstelde. Van hem is ook de naam « ogive » voor wat wij nu de cumulatieve normale frequentiekromme noemen.

Ook de eigenaardige term « regressie » is van hem afkomstig : Hij konstateerde, dat, terwijl lange vaders gemiddeld lange zoons hebben, de kleinzoons minder van het gemiddelde afwijken, en noemde deze teruggang naar het gemiddelde « regressie ».

Het is wel merkwaardig, dat nu de regressie in de statistiek eerder het omgekeerde betekent.

Men dankt aan hem ook de correlatiecoëfficiënt, evenals de percentielen, inzonderheid de quartielen als middel voor de schatting van de standaarddeviatie. Pas veel later zou men ontdekken, dat deze hiertoe minder geschikt zijn wegens hun grote spreiding.

---

[12] R.A. Fisher, Annals of Eugenetics, 7, 4, 303 (1937).

[13] W.P. Elderton, Frequency Curves and Correlation, Cambridge, versch. drukken.

[14] Encycl. Britan., artikel over Galton.

Eerst in 1946 zou Mosteller vaststellen [15], dat hiertoe het 7e en 93e percentiel het meest geschikt zijn. Merken wij op, dat nu trouwens nog andere methoden zijn uitgewerkt, die met vrij geringe vermindering van de efficiëntie, zeer vlug de standaarddeviatie doen schatten. Zo de methode van de range, die bij het gebruik van Shewartkaarten bij de fabriekscontrole zulke goede diensten kan bewijzen. De range is, zoals men weet het verschil tussen de grootste en kleinste in een steekproef verkregen waarde. Uit onderzoekingen o.a. van Egon Pearson blijkt, dat men met behulp hiervan, bij oordeelkundig gebruik, zeer gemakkelijk de standaarddeviatie kan berekenen. Weliswaar wordt bij al deze methoden verondersteld, dat men een normale verdeling onderzoekt.

Galton stierf in 1911 en vestigde bij testament aan het Londense University College een leerstoel in nationale eugenetica. In deze leerstoel werd benoemd zijn medewerker Karl Pearson, die tot dan professor was in de toegepaste wiskunde aan dezelfde universiteit. Het Galton Institute, dat nog steeds een der belangrijkste centra van aktiviteit is op het gebied van de statistiek, staat tegenwoordig onder leiding van zijn zoon Egon Pearson.

Karl Pearson (1857-1936) was een wiskundige van Cambridge ; hij is meer dan 50 jaar professor geweest aan het Londense University College.

Wie Pearson zegt, denkt allereerst aan het belangrijke tijdschrift « Biometrika », dat in 1901 werd gesticht ten gevolge van moeilijkheden, die waren gerezen in zake publicaties in de Philosophical Transactions van de Royal Society, door controversen tussen Pearson en Bateson over het Mendelisme, controversen, waaraan de statistiek ten slotte veel te danken heeft gehad.

Maar Pearson deed meer : met zijn χ^2-verdeling gaf hij ook een middel aan de hand om te zien, in hoeverre de experimenten overeenstemmen met een vooropgestelde theorie : de zogenaamde goodness of fit.

Wij kunnen niet alle punten noemen, waarop hij werkte, wij willen volstaan met het in herinnering brengen van de vele uiterst bruikbare tabellarische werken van zijn hand.

Karl Pearson had vele leerlingen, die de statistiek verder brachten. Wij denken aan Yule, lector in Cambridge, wiens boek zo lange jaren toonaangevend was, en waaruit het monumentale werk van M.G. Kendall, nu Kendall en Stuart, is gegroeid.

Wij denken ook aan de chemicus Gosset, met zijn beroemde t-verdeling of « Student »-verdeling, die zo belangrijk is geworden bij het vaststellen van betrouwbaarheidsintervallen, van significantietoetsen en, zoals hij zelf reeds vond, als toets voor het vaststellen van correlatie. Hij brak het gebied open voor de behandeling van kleine steekproeven. Zijn intuïtie bracht hem steeds

---

[15] F. Mosteller, Ann. Math. Stat. 17, 4, 377 (1946).

op het juiste spoor, ondanks zijn beperkte mathematische ondergrond. Dikwijls pas veel later vond men wiskundig bevredigende bewijzen.

Het is wel algemeen bekend, hoe de eigenaardige naam « Student » is ontstaan : Gosset was verbonden aan de Guiness-brouwerij in Dublin en een publicatieverbod moest er voor zorgen, dat alle vruchten van zijn wijsheid alleen in dit bier terechtkwamen. Vandaar zijn statistische publicaties onder de schuilnaam Student. Student bleef zijn leven lang zowel de statistiek als het bier trouw, later als brouwmeester van de Londense brouwerij Park Royal. Het werd de eerste brouwerij met een statistische afdeling.

Student was dus een verbindingselement tussen zeer verschillende zaken, maar niet alleen op dit gebied. Hij was een even groot vriend zowel van Pearson als van Ronald Aylmer Fisher (1890-1962). En dit wil heel wat zeggen, als men weet, dat deze twee grote geleerden helaas een deel van hun kostbare tijd hebben besteed aan het uitwisselen van minder vriendelijke opmerkingen.

Niemand zal betwisten, dat Fisher de grote coryphee is van de statistiek. Zijn werk is van een ontzagwekkende omvang. Hij bracht nieuwe methoden en synthetiseerde oude.

Ofschoon Fisher een wiskundige opleiding had gehad, met schitterende successen in Cambridge, heeft hij, althans in zijn publicaties, zelden behoefte gehad aan grondig ingaan op theoretische wiskundige detailmoeilijkheden. Hij kon zich dit veroorloven door zijn feilloze intuïtie. Hij gaf de grote lijn en gaf ze juist. Maar sommige zijner werken brachten degenen, die alle wiskundige moeilijkheden, die onbehandeld bleven tot klaarheid wilden brengen, tot wanhoop. In zijn « Statistical Methods for Research Workers » met zijn ontelbare herdrukken in het bijzonder ontbreken niet alleen bewijzen, maar zelfs neemt hij niet de moeite te verwijzen naar zijn eigen artikelen, waarin hij soms bewijzen heeft gebracht.

Zoals Bliss [16] opmerkt, zal de wiskundige telkens stoten op beweringen, die ogenschijnlijk geen andere steun hebben dan de intuïtie van Fisher, maar na ijverig bladzijde na bladzijde te hebben volgeschreven met berekeningen zal hij enkel maar tot de conclusie komen : « Fisher had gelijk ».

Een zijner verdiensten is het gebruik van meetkundige meerdimensionale beelden om in de statistiek voorkomende funkties te behandelen.

Men meent soms de bijzondere geaardheid van de wetenschappelijke methode van Fisher te kunnen verklaren uit de manier, waarop hij de wiskunde leerde. Zijn gezichtsvermogen was zo slecht, dat hem reeds in zijn

---

[16] C.I. Bliss, Biometrica 18, 4, 438 (1962).

jonge jaren was verboden bij kunstlicht te werken. Hij leerde dan ook een groot deel van de wiskunde langs mondelinge weg.

Vanaf het begin van zijn loopbaan had hij bijzondere belangstelling voor de genetica, en in overeenstemming met zijn strijdvaardig karakter nam hij reeds onmiddellijk scherp positie tegen de, onder de invloed van Bateson, toen in Cambridge gangbare meningen. Hij bracht een verzoening tot stand tussen het Mendeliaans en het biologisch standpunt van die tijd. Hij gaf een statistische theorie van de evolutie. In een woord hij bracht dit gebied veel verder door zijn tweezijdige begaafdheid : zijn inzicht zowel op biologisch gebied als op dat der wiskunde.

Een mijlpaal in zijn leven is ook voor de ontwikkeling van de statistiek een mijlpaal geweest : zijn benoeming in 1911 door Russell tot statisticus aan het grote en uitgestrekte Engelse centrum voor landbouwwetenschap : Rothamsted, waar o.m. graanvelden staan, waarop sinds 1843 elk jaar graan is gezaaid, geoogst met de meest wetenschappelijke methoden, op ieder perceel met een eigen welomschreven vaste bemesting. Dit was een ideaal werkterrein voor Fisher, die er de proefveldtechniek tot de volmaakte hoogte bracht, die wij thans kennen. Dit heeft ten gevolge gehad, dat vele zijner statistische ontdekkingen, die nu hun toepassingen vinden in de meest verscheidene gebieden, het eerst de landbouw ten goede zijn gekomen. Wij denken in de eerste plaats aan de analyse van de variantie en aan het plannen van experimenten. Nu in alle gebieden gebruikelijke woorden, zoals « blok » herinneren aan de landbouwkundige oorsprong van deze methoden.

Een *hoofdgedachte* van Fisher is de volgende : om te bewijzen, dat b.v. een graansoort verbeterd is, moet men bij de statistische berekeningen *niet* uitgaan van het nieuwe verbeterde gemiddelde, maar van de *nulhypothese,* die inhoudt, dat niets veranderd is. Blijken dan de verkregen experimentele waarden onwaarschijnlijk ver hiervan te liggen, dan is er reden, de nulhypothese te verwerpen. Het gaat dus bij hem om bewijzen uit het ongerijmde. Fisher neemt tot ergernis van Lancelot Hogben, zeer scherp positie [17] door te verklaren : de nulhypothese wordt door de experimenten nooit bewezen of bevestigd, maar *wel* mogelijkerwijze weerlegd. Die nulhypothesen worden gebruikt in verband met *significantietests,* die aangeven, hoe groot een afwijking moet zijn, om afwijzing van de nulhypothese aanvaardbaar te maken, en ook in de *analyse van de variantie,* waar bepaalde schattingen van de variantie als zij zeer groot worden eveneens de nulhypothese onwaarschijnlijk maken.

---

[17] R.A. Fisher, The Design of Experiments, 5th Ed. (1949), p. 16.
L. Hogben, Statistical Theory. An examination of the contemporary crisis in statistical theory from a behaviourist viewpoint, p. 333 en 500 (1957).

Een *tweede principe* door Fisher doorgevoerd is het variëren van meer faktoren tegelijkertijd : hij varieert b.v. zowel het stikstof- als het kalium- en het phosphorgehalte van de bodem. Dit in tegenspraak met de toen gangbare uitspraak van Russell, die voorhield : Stel de natuur maar één vraag tegelijk ! Zo ontstonden de *factoriële experimenten*. De gedachte was volgens Yates [18] niet geheel onbekend, maar pas door het werk van Fisher werd zij bruikbaar.

Deze principes, verder aangevuld en uitgebreid, gaven een revolutie op landbouwkundig gebied en bleken later bruikbaar in de industrie, de openbare gezondheid, de geneeskunde, opvoedkunde, biologie, psychologie, kortom voor allerlei wetenschappelijke experimenten.

Wij herinneren aan de random-blocks, met alle verfijningen er van, zoals latijnse kwadraten, confounding en zoveel andere proefschema's, dat er tegenwoordig gehele boeken mee gevuld zijn.

Vooral zijn z-toets, gekoppeld aan zijn eigenlijk meetkundig begrip van vrijheidsgraden maakten de analyse van de variantie tot een uiterst doelmatige werkmethode. Zijn z-verdeling werd in de praktijk vervangen door een meer bruikbare verdeling, die door Snedecor [19] ter ere van Fisher F-verdeling werd genoemd. Fisher heeft echter zelf die letter nooit willen gebruiken, maar omschreef ze als e^{2z} en noemde ze variance ratio [20].

Door zijn definitieve afleidingen van de frequentiefuncties voor de variantie, de z, de χ^2, de regressiecoëfficiënten, de t-verdeling, de gewone, de partiële en de multiepele correlatiecoëfficiënt werd een stevige basis geschapen, waarop voortgebouwd kon worden, zodat men nu b.v. kan berekenen, hoeveel herhalingen nodig zijn om met een analyse van de variantie bepaalde verschillen te kunnen vaststellen.

Zeer belangrijk zijn de toepassingen van de correlatierekening in de economie, waar Tinbergen, de grondlegger van de econometrie, zijn boeken over « explanation » schreef, nu gevolgd door zovele anderen, die pogen de economische verschijnselen te verklaren en zelfs te voorspellen.

Ook de psychologen gebruiken deze methoden. Zij trachten hun experimenten door een aantal factoren te verklaren. Hun methoden stellen bijzondere eisen, aangezien bij hen deze « factoren » niet te voren bekend zijn, maar uit de verwerking van het experimenteel materiaal moeten worden opgespoord.

Het woord « factor » moet met omzichtigheid gebruikt worden, aangezien het verschillende betekenissen heeft. Het is een voorbeeld van de verwar-

---

[18] F. Yates, Biometrics, 18, 4, p. 445 (1962).

[19] G.W. Snedecor, Analysis of Variance and Covariance, Ames (1934).

[20]) Fisher and Yates, Statistical Tables, b.v. 4e druk (1953).

ringen die kunnen ontstaan in een zo snel groeiende tak van wetenschap, die een trefpunt is van beoefenaars van zeer verschillende disciplines. Statistici, economisten, psychologen geven ieder een andere betekenis aan het woord als de twee betekenissen die in de wiskunde voorkomen, en elk van die betekenissen vindt men terug in de statistiek : n.l. respektievelijk :

in de statistiek : komponenten in de analyse van de variantie bij factoriële experimenten ; in de economie : factoren, die een rol spelen bij berekeningen van indexcijfers ; in de psychologie de factoren in de regressieproblemen van de z.g. factoranalyse ; in de wiskunde de elementaire betekenis van het woord factor en ten slotte in verband met differentieproducten, b.v. in factoriële momenten.

Behalve de significantietoetsen is zeer belangrijk de zogenaamde *puntestimatie :* hierbij wordt op grond van de proeven getracht, voor een bepaalde parameter, b.v. een gemiddelde, de beste aanvaardbare waarde aan te wijzen. Ook in dit gebied bracht Fisher klaarheid door zijn eisen te stellen : consistentie, efficiëntie, sufficiëntie en onvertekendheid of unbiassedness, een woord, waarvan de Engelsen zelf niet weten, hoe ze het spellen moeten [21]. Zijn *maximum likelihood* methode bracht hier schitterende uitkomsten.

Als derde geval blijft dan nog de *interval-estimatie.* Men weet n.l. niet precies waar de juiste theoretische waarde van de gezochte parameter (b.v. het gemiddelde) ligt. Maar men wil dan een interval aangeven, waarvan men kan bewijzen, dat die waarde een kans van b.v. 95 % heeft, er binnen te liggen.

Fisher spreekt dan van *fiducial interval.*

In dit gebied zijn grote theoretische moeilijkheden aanwezig, die verband houden met het ontgaan van de hypothese van Bayes. Deze hypothese houdt in, dat onbekende oorzaken als even waarschijnlijk zouden moeten worden beschouwd. Fisher heeft duidelijk de onhoudbaarheid van deze hypothese aangetoond door coördinatentransformaties toe te passen. Het is trouwens de reden, waarom hij, om niet van [waarschijnlijkheden van mogelijke waarden van de parameter] te moeten spreken, niet het woord « probability », maar « likelihood » gebruikt [22]. Is hiermee echter deze zaak voldoende opgelost ? In deze moeilijke kwestie heeft Fisher veel kritiek gehad, vooral van Neyman en Egon Pearson, die het begrip likelihood in dit gebied ontgaan en het fiducial interval vervangen door het « conficence interval », waarin duidelijk geen sprake is van waarschijnlijkheden van niet bestaande parameterwaarden, dus van niet bestaande oorzaken.

---

21) M.G. Kendall, The Advanced Theory of Statistics, II (1946),p. 4.

22) Men zou dit kunnen vertalen met : niet « waarschijnlijkheid », maar « voor-de-hand-liggendheid ».

Deze problemen hebben een hevige strijd ten gevolge gehad, die nog niet geheel geluwd is en nog ingewikkelder werd door het in het strijdperk treden van Jeffreys met een derde theorie. Hogben [23] (hij spreekt van crisis) schreef er een boek van 510 bladzijden over en de discussie duurt voort. De moeilijkheid is steeds de toepasbaarheid van de wiskundige waarschijnlijkheid in de werkelijkheid : welke hypothesen acht men aanvaardbaar ? Soms, b.v. Kendall en Stuart, The Advanced Theory of Statistics II (1961), p. 157, tracht men op logische gronden de fiducial theorie aan te vallen.

Maar ten slotte blijft bij de beoordeling van de aanvaardbaarheid van dergelijke conclusies op grond van statistische onderzoekingen een subjectief element aanwezig. In die richting wijst het werk van de Finetti en Savage [24], met zijn z.g. personalistisch standpunt.

Gelukkig geven het fiducial en het confidence-standpunt in het algemeen dezelfde uitkomsten, zodat de strijd grotendeels op theoretisch gebied blijft. Een uitzondering vormt het probleem van het vergelijken van twee steekproeven, waarvan de onbekende standaarddeviaties verschillend moeten worden geacht. Daar geeft de methode van Fisher-Behrens, die op fiduciële grondslag berust andere uitkomsten dan die van Welch, die op de gedachtengang van Pearson en Neyman steunt, maar die, evenmin als die van Scheffé, niet ten volle bevredigt. Nog steeds wordt in dit gebied verder gewerkt.

Stappen wij nu af van het gebied van de interval-estimation en wenden we ons weer tot het gebied van toetsen en vergelijken van hypothesen. Ook hier is door Pearson en Neyman grote vooruitgang gebracht.

Zij voerden uitdrukkingen in als fouten van de eerste soort (α), n.l. verwerpen van een hypothese, als deze juist is, en fouten van de tweede soort (β), n.l. aanvaarden van een hypothese, als deze onjuist is. Deze begrippen hebben in de kwaliteitscontrole een zeer grote rol gekregen onder de naam van respectievelijk : risico van de producent en van de koper.

Van zeer groot belang is hun begrip « power » of macht van een toets geworden : d.i. de kans geen fout van de tweede soort te maken ; het is het complement van wat men in de kwaliteitskontrole OC-kromme (operational characteristic) is gaan noemen.

Hun pogingen, de beste kritieke zones aan te wijzen : die zones waarin de onderzochte hypothese moet worden verworpen, voerde hen tot een uitbreiding van het maximum likelihood principe : de likelihood ratio.

Dit principe voerde Wald weer tot zijn sequential analyse : een methode van kwaliteitsonderzoek, waar de omvang van de steekproef niet van te voren

---

23) L. Hogben, *l.c.*

24) L.J. Savage, The Foundations of Statistics, New York, 1954.

vast staat, maar zeer klein kan worden gemaakt, indien de te onderzoeken partij van zeer goede of van zeer slechte kwaliteit is. Het kan wellicht verduidelijkt worden met het voorbeeld van een examen : het examen van een zeer goede of zeer slechte student kan kort zijn, omdat de uitslag spoedig duidelijk wordt.

In deze sequentiële procedures komt wellicht meer dan in andere gebieden van de statistiek naar voren, dat de statisticus zich steeds voor een beslissing ziet geplaatst : hij heeft te kiezen tussen het risico een verkeerd besluit te nemen (immers een steekproef kan nooit zekerheid geven) en de kosten, die door verder voortgaan met de steekproef worden meegebracht.

Het kan dan ook niet verwonderen, dat Wald er toe kwam, een gehele theorie te ontwerpen, de decisietheorie, over het nemen van besluiten. Hij en vele anderen, b.v. [25] Lehmann beschouwen deze als de juiste grondslag voor de gehele statistiek. Niet allen zijn het hiermee eens : Kendall en Stuart [26] b.v. verzetten zich sterk tegen dit standpunt.

De weg tot dit gebied was geopend door de speltheorie, zoals deze was ontwikkeld in het belangrijke werk van von Neumann en Morgenstern, Theory of Games and Economic Behaviour, Princeton (1944) en latere uitgaven.

Op grond van dit werk kwam Wald tot toepassing van de *minimax*-theorie, die veronderstelt, dat de tegenspeler altijd de voor ons nadeligste oplossing zal kiezen. Men zoekt dan welke tegenstrategie het minst nadelig is. In de statistiek is onze tegenspeler de natuur, en het is wat pessimistisch te veronderstellen, dat deze altijd ons nadeel zoekt. Het betekent immers, dat men nooit onder een boom zou mogen lopen, uit vrees te worden gedood door zijn val [27].

Men komt er daarom toe, rekening te houden, met wat men al weet over het gedrag van de natuur bij vroegere gelegenheden. Dit voert tot wat men Bayes-oplossingen is gaan noemen [28].

De principes van deze besluittheorie tracht men ten grondslag te leggen aan nieuwere statistische onderzoekingen.

Wij zouden nog over grote domeinen moeten spreken : de tijdreeksen met auto- en cross-correlatie, met al het werk van economisten en van Norbert Wiener met zijn Fouriertheorieën. Dit zou ons weer voeren tot de communicatie en informatietheorie in het algemeen en tot de school van Shannon met zijn algebraïsche of meer nog combinatorische methoden.

Ook hebben we nog geen woord gezegd over de niet-parametrische methoden, die trachten statisitsche gegevens te verwerken en vooral toetsen

---

25) E.L. Lehmann, Testing Statistical Hypotheses, New York, 1959, p. VII.
26) *l.c.*, II, 620.
27) Kendall & Buckland, A Dictionary of Statistical Terms, 1st Ed. (1957), p. 182.
28) A. Wald, Statistical Decision Functions (1950).

te geven, *zonder* een bepaalde verdelingsfunctie aan te nemen, dus ongeveer het tegendeel van de klassieke school, die soms zonder veel grond, de verdeling als normaal aannam. Zeer belangrijk is deze niet-parametrische methode in gevallen waar de gegevens wel volgens een schaal gerangschikt kunnen worden, maar zonder dat een vaste eenheid kan worden aangewezen, zoals b.v. bij het beoordelen van studenten, waar niemand zal durven beweren, dat de afstand tussen de waarderingscijfers 2 en 3 gelijk zou zijn aan die tussen b.v. 9 en 10. Nog sterker indien zelfs een rangschikking onmogelijk is, b.v. bij diagnoses, waar wordt ingedeeld in verschillende ziektesoorten, wat men *nominale* schalen noemt.

Hier zijn opmerkelijke resultaten behaald, die in sommige gevallen in efficiëntie weinig onderdoen voor de klassieke, wel parametrische toets, b.v. in het twee-steekproevenprobleem, waar de Kolmogorov-Smirnov-test en de U-test, die men ook Mann-Whitney-test en ook Wilcoxontest noemt, weinig in efficiëntie onderdoen voor de klassieke *t*-test.

Ten slotte noemen wij nog een methode, waarin de statistiek wel een heel bijzondere triomf viert : de Monte-Carlomethode. De naam « Monte Carlo » stamt uit de oorlogsjaren. Het was een codenaam voor deze wijze van rekenen, die voor de strategie van belang was, die zelfs kan gebruikt worden bij berekeningen in verband met atoomdesintegratie. In deze methode roept de statistiek niet de wiskunde te hulp, maar worden de rollen omgekeerd : wanneer bepaalde berekeningen te langdurig of onmogelijk zijn, vervangt men ze door statistische berekeningen.

Een zeer eenvoudig voorbeeld zou het volgende zijn : Stel men wil het oppervlak kennen van een vlakke figuur met zeer ingewikkelde rand. Men kiest nu willekeurig punten in een vierkant, dat de figuur omvat en dat natuurlijk een bekend oppervlak heeft. Door te tellen, hoeveel punten binnen de figuur liggen, hoeveel er buiten, kan men het oppervlak van de figuur bepalen

Een zeer oud voorbeeld van toepassing van de Monte-Carlomethode, lang voor deze naam er voor bestond, is het naaldprobleem van Buffon. Hierbij laat men een naald met lengte 1 vallen op een papier, waarop evenwijdige lijnen met afstand 1 getekend zijn. De kans, dat een lijn door de naald gesneden wordt is $2/\pi$, zodat men door tellen van de uitslagen een benadering van π kan krijgen. Lazzerini moet met deze methode de waarde 3,1415928 hebben gevonden, dus de eerste 7 cijfers correct[29].

---

[29] Verdere bijzonderheden over dergelijke experimenten in L. Hogben, *l.c.*, p. 54. Dr. Hirschberg was zo vriendelijk mij mede te delen, dat zeer goede kritiek op zo « nauwkeurige » uitkomsten te vinden is bij J.M. Hamersley en K.W. Morton, Journ. Roy. Stat. Soc. B. 16 (1954), 75.

Ik heb getracht het gebied van de statistiek met haar ontelbare vrijheidsgraden ééndimensionaal in de tijd af te beelden. Zoals U weet kan zulk een afbeelding niet gelijkvormig zijn, nog minder congruent ; dit althans meen ik empirisch wel bewezen te hebben.

Laat ik besluiten met een uitspraak die ik, naar ik meen bij Fisher, een keer gelezen heb : In de statistiek gebruiken we soms een kanon om een mus te schieten, en schieten er dan nog naast.

H. FLORIN.

Behalve de reeds genoemde werken werden geraadpleegd :

I. Todhunter, A. History of the Mathematical Theory of Probability, New York, 1949.

R.A. Fisher, Contributions to Mathematical Statistics, New York 1950 met biografie door P.C. Mahalanobis.

M.E. Smith, A Source Book in Mathematics, I, II, New York z.j.

J.O. Irwin, Biometric Method, Past, Present and Future, Biometrics 15, 3 (1959) 363.

G.U. Yule and M.G. Kendall, An Introduction to the Theory of Statistics, 13e druk, 1948.