

# DE 'PROBABILISTIC REVOLUTION' IN DE WETENSCHAPPEN

IDA H. STAMHUIS\*

## *Inleiding*

The normal or bell-shaped curve at first represented the probability of observational error in astronomy, then of nature's 'errors' from *l'homme moyen* in sociology, then of anarchic individual gas molecules exhibiting orderly collective properties. Eventually the normal curve came to represent the distribution of almost everything.<sup>1</sup>

De introductie van de klokvormige curve in uiteenlopende wetenschapsgebieden is kenmerkend voor de plaats die de kanstheorie tussen ongeveer 1800 en 1930 heeft verkregen. Stapsgewijs werd het toepassingsgebied van de waarschijnlijkheidsrekening uitgebreid op grond van succesvolle analogieredeneringen. Tijdens dit proces kregen kanstheoretische begrippen een andere interpretatie en ontstonden nieuwe begrippen en technieken. Deze gebeurtenissen brachten zowel voor de kanstheorie als voor de betreffende wetenschappen grote veranderingen met zich mee. De eerste wetenschap waarin de kanstheorie werd gebruikt was de astronomie. Daarna kreeg ze een nieuwe toepassing in de 'wetenschap der samenleving', ofwel een kennisgebied dat wij nu zouden omschrijven als een combinatie van sociologie, economie en sociale geneeskunde; daarna ongeveer tegelijkertijd in de biologie en in de fysica en na de tweede wereldoorlog in de psychologie. Het valt op dat de twee vakgebieden natuurkunde en sterrenkunde, die toch zeer verwant zijn, in deze rij worden gescheiden door het andersoortige vakgebied van de 'wetenschap der samenleving'. Dit gegeven vraagt om een nadere behandeling. In het nu volgende artikel wordt de chronologie van de introductie van de kanstheorie nader geanalyseerd. Ten slotte wordt ingegaan op de te onderscheiden stadia van de 'probabilistic revolution'.<sup>2</sup>

## *Waarschijnlijkheidsrekening in de sterrenkunde*

De waarschijnlijkheidsrekening is niet in het vakgebied van de sterrenkunde ontstaan. Ze kwam op in de zeventiende eeuw als een theorie van dobbelen en kaartspelen, en werd bijna tegelijkertijd toegepast in de levensverzekeringstheorie, door onder anderen onze landgenoot Johan de Witt (1625-1672). In deze theorie speelde niet het begrip kans, maar het begrip verwachting de centrale rol. De formule voor de normale verdeling werd voor

\* Vrije Universiteit, Amsterdam, Faculteit der Natuurkunde en Sterrenkunde, Vakgroep Algemene Vorming, De Boelelaan 1081, 1081 HV Amsterdam.

1. G. Gigerenzer, Z. Swijtink, T. Porter, L. Daston, J. Beatty & L. Krüger, *The empire of chance. How probability changed science and everyday life* (Cambridge, 1989) xiv.

2. Dit is de titel van het tweedelige werk: L. Krüger, L.J. Daston & M. Heidelberger, eds., *The probabilistic revolution. Volume 1: Ideas in history*, en L. Krüger, G. Gigerenzer & M.S. Morgan, eds.: *The probabilistic revolution. Volume 2: Ideas in the sciences*. Beiden (Cambridge (Massachusetts)/Londen, 1987).

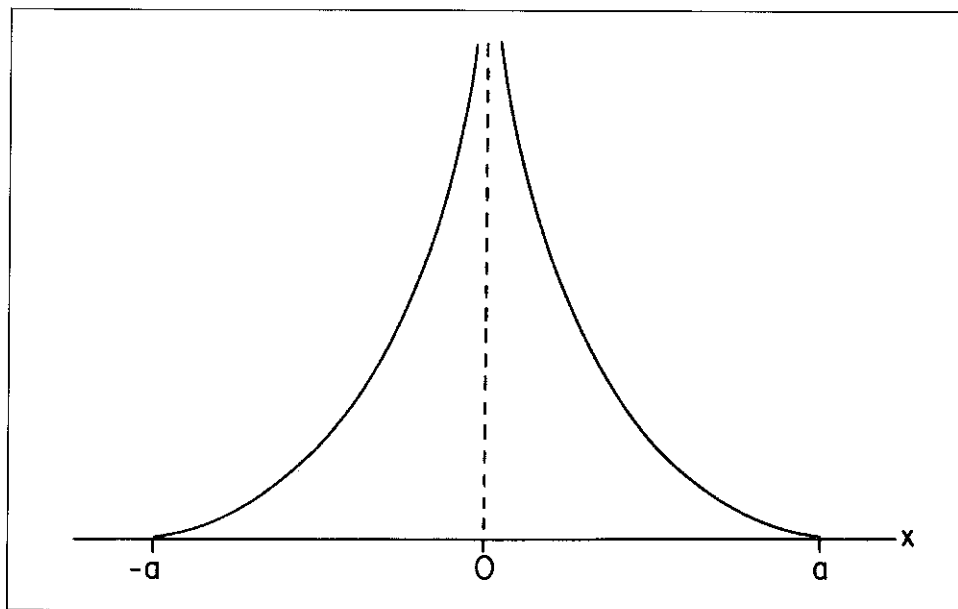


fig. 1. Een suggestie voor de foutenkromme, gedaan door Laplace in een artikel van 1777.  
P.S. de Laplace, 'Recherches sur le milieu qu'il faut choisir entre les résultats de plusieurs observations', 228-256 in: C.C. Gillispie, *Mémoires inédits ou anonymes de Laplace sur la théorie des erreurs, les polynômes de Legendre, et la philosophie des probabilités*, *Revue d'histoire des sciences* 32 (1979): 223-279.

het eerst geformuleerd door de wiskundige A. de Moivre (1667-1754) en wel als een limietverdeling van de zogenaamde binomiale verdeling.

Eén van de belangrijkste sterrenkundige problemen in de achttiende eeuw was het bepalen van de complexe banen van een aantal planeten, in het bijzonder van Jupiter en Saturnus. Verder stonden de bepaling van de beweging van de maan en het wiskundig representeren van deze beweging volop in de belangstelling. Ten slotte was er behoefte aan nauwkeurige metingen van de vorm en de afmetingen van de aarde.<sup>3</sup> Men had bij metingen in het kader van deze problematiek vastgesteld dat toevallige meetfouten een bepaalde regelmatigheid vertonen en men zocht naar de exacte vorm van deze regelmaat. Verschillende vooraanstaande geleerden aan het eind van de achttiende eeuw hebben zich met dit probleem beziggehouden, zoals Daniel Bernoulli (1700-1782), P.S. de Laplace (1749-1829) en C.F. Gauss (1777-1855). Er werden voor de verdeling van toevallige meetfouten verschillende suggesties gedaan, die later weer werden verworpen (zie fig. 1). Uiteindelijk werd duidelijk dat de kansverdeling van meetfouten dezelfde was als de kansverdeling van de centrale limietstelling (de som van  $n$  onafhankelijke willekeurige toevallige grootheden convergeert altijd naar dezelfde limietverdeling). Bovendien bleek dat het eveneens deze verdeling was

3. S.M. Stigler, *The history of statistics. The measurement of uncertainty before 1900*, (London, 1986) 11-158.

die de hoofdrol speelt bij de kleinste-kwadratenmethode,<sup>4</sup> omdat geldt dat, in het geval de waarnemingsfouten normaal zijn verdeeld, de kleinste-kwadratenmethode de beste methode is om een werkelijke waarde te schatten.

Onder andere door deze ontwikkelingen kon in de sterrenkunde een begin worden gemaakt met de belangrijke ontwikkeling die ze in de negentiende eeuw heeft doorgemaakt.<sup>5</sup> De kleinste-kwadratenmethode werd spoedig in verschillende takken van de natuurwetenschappen toegepast. De waarschijnlijkheidsrekening was niet langer vooral een theorie van dobbelen en kaartspelen die ook voor de levensverzekeringstheorie kon worden gebruikt; ze werd daarnaast bekend als de theorie van meetfouten, waarin de foutenkromme, die later de normale verdeling zou worden genoemd, de centrale rol speelde.

#### *Van sterrenkunde naar de wetenschap der samenleving: Quetelet en de 'physique social'*

In Duitsland was in de zeventiende eeuw een vakgebied aan de universiteit ontstaan waarin aan de hand van bepaalde criteria kenmerken van een staat werden behandeld. In de achttiende eeuw werd dit vakgebied 'Statistik' genoemd. Deze behandeling moest plaatsvinden op grond van feitelijke informatie, die niet persé uit getalsmatige gegevens behoefde te bestaan. Deze Duitse traditie van statistiek werd in 1798 onder de naam 'Statistics' door een publikatie van de Schot Sir John Sinclair (1754-1835), getiteld 'Statistical Account of Scotland' in Groot-Brittannië geïntroduceerd. Dit vakgebied werd toen geïntegreerd in de traditie van politieke rekenkunde, waarin op basis van kwantitatieve gegevens en met behulp van kansen informatie over een staat werd geleverd.

Het voorgaande had tot gevolg dat men in de jaren dertig van de negentiende eeuw in Groot-Brittannië onder statistische gegevens feiten verstond die informatie over en inzicht in de samenleving gaven. Die feiten moesten bij voorkeur getalsmatig zijn en in tabellen kunnen worden geordend.<sup>6</sup>

Zo was de situatie bij het optreden van de Belg L.A.J. Quetelet (1796-1874).<sup>7</sup> Deze wiskundige van opleiding begon zich in het begin van de jaren twintig te interesseren voor astronomie. In dat kader maakte hij in 1823 een studiereis naar Parijs. Daar had men binnen de astronomie de waarschijnlijkheidsrekening verder ontwikkeld. Bovendien had men er zich beziggehouden met de popularisering van de kanstheorie en met de mogelijke toepassing op maatschappelijke en juridische onderwerpen.<sup>8</sup> Zo had men opgemerkt dat vele gebeurtenissen die ogenschijnlijk door irrationele factoren werden bepaald, toch een, naar het de mensen toen voorkwam, verbazingwekkende orde vertoonden.

Na zijn bezoek aan Parijs legde Quetelet een hartstochtelijke interesse aan de dag voor zowel waarschijnlijkheidsrekening als statistiek. Hij was al in 1829 van mening dat de vergelijkende statistiek, waarmee hij bedoelde een verzameling van numerieke gegevens betreffende de maatschappij, een belangrijk toepassingsgebied vormde voor de waarschijnlijkheidsrekening.

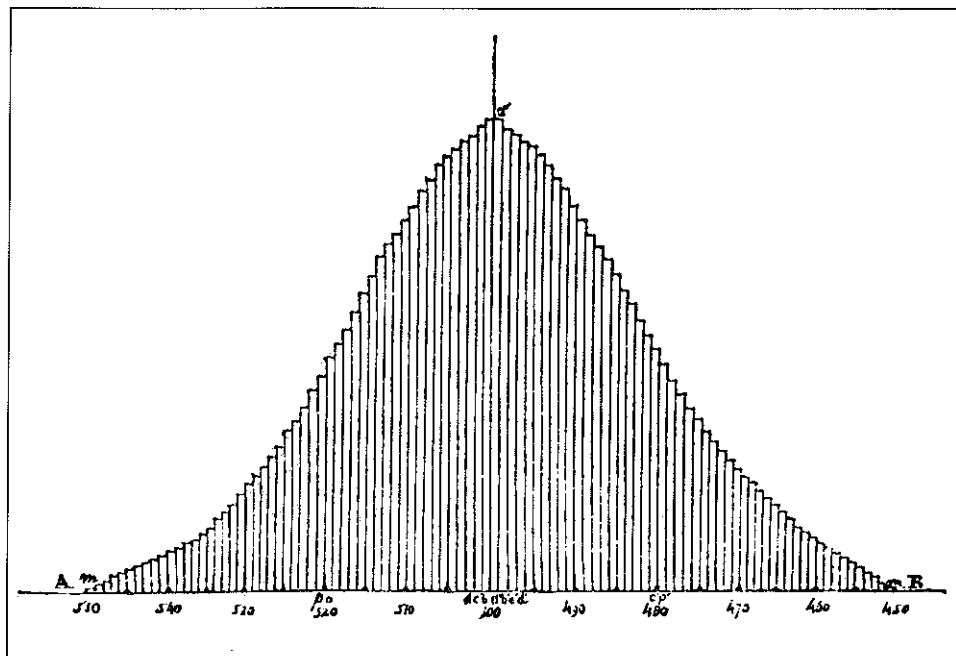
4. Bij deze methode gaat het om een grootheid die een geparparameteriseerde functie is van een aantal variabelen. De grootheid wordt voor verschillende waarden van de variabelen gemeten, waarbij het aantal metingen groter is dan het aantal parameters. De optimale waarde van de parameters wordt bepaald door het minimaliseren van de som van de kwadraten van de verschillen tussen gemeten en voorspelde waarden van de betreffende grootheid.

5. Zie hierover bijvoorbeeld het artikel van E. Dekker, 'Een procesverbaal van verhoor' elders in dit nummer.

6. V.L. Hiltz, 'Aliis extensum, or, the origins of the Statistical Society of London' *Isis* 69 (1978) 21-44.

7. O.a. in T.M. Porter, *The rise of statistical thinking 1820-1900* (Princeton, 1986) 41-55.

8. Gigerenzer (n.1), *The empire of chance*, 32-34.



Een benadering door Quetelet van de normale verdeling met behulp van de binomiale verdeling.  
A. Quetelet, *Lettres à S.A.R. le Duc Régnant de Saxe-Cobourg et Gotha, sur la théorie des probabilités, appliquée aux sciences morales et politiques* (Brussel, 1846) 396.

Met behulp daarvan zouden filosoof en staatsman in staat zijn de fysieke en morele toestand van de mens te bestuderen.<sup>9</sup>

Onder die mens verstond Quetelet niet een specifiek individu, maar een statistische abstractie: *l'homme moyen*, ofwel de gemiddelde mens, die voor alle te onderscheiden eigenschappen een gemiddelde waarde bezat. Voor Quetelet was dat het ideaaltypemens, het 'zwaartepunt' van de maatschappij. Deze mens moest worden bestudeerd in een nieuwe wetenschap die hij *physique social* noemde. Elke samenleving kon door een gemiddelde mens worden gekarakteriseerd. Quetelet was onder andere bijzonder geïnteresseerd in de neiging tot misdaad van die mens. Hij liet zien dat in elk gebied of land elk jaar evenveel misdaden, evenveel moorden en ook een vrijwel gelijk aantal zelfmoorden werden gepleegd. Het percentage daders leverde een getal op dat hij beschouwde als de neiging tot misdaad van de gemiddelde mens van het desbetreffende gebied.

Quetelets toepassing van kansentheorie en statistiek in een 'exacte' sociale wetenschap kreeg grote aanhang. Hij liet zien dat de verdeling, die tot dan toe de kansverdeling van meetfouten was geweest, heel goed kon worden gebruikt voor de beschrijving van de variabiliteit in de natuur. De waarschijnlijkheidsrekening kreeg er zo een belangrijk toepassingsgebied bij en de klokvormige curve werd van een nieuwe betekenis voorzien. Voor hem zelf was trouwens geen sprake van een nieuwe visie, omdat voor hem de afwijkingen van de

9. L.A.J. Quetelet, *Recherches statistiques sur le royaume des Pays Bas* (Brussel, 1829).

gemiddelde mens in essentie fouten waren, want het waren afwijkingen van de werkelijke waarde. Voor anderen echter hebben zijn ideeën zeer innoverend gewerkt.

Ondanks Quetelets ideeën, leenden de toenmalige gegevens, die bestonden uit tabellen gebaseerd op volledige tellingen, zich in de praktijk niet zo gemakkelijk voor een kans-theoretische analyse. Aan de 'probabilisering' van de statistische praktijk heeft Quetelet geen bijdrage geleverd. Wel heeft hij door zijn aandeel in de oprichting in 1834 van de London Statistical Society, en nog meer door de op zijn initiatief gehouden internationale statistische congressen er aan meegewerkt, dat de statistiek in de negentiende eeuw zeer populair werd. Zijn ideeën over de rol van kans en statistiek hebben zoveel invloed gehad dat men dit geheel van gedachten tegenwoordig 'Queteletismus' noemt.<sup>10</sup> Uit het volgende zal blijken dat Quetelet personen uit zeer verschillende vakgebieden heeft beïnvloed bij het toepassen van kans-theoretische principes.

Statistiek werd in de vorige eeuw beschouwd als een vak dat zocht naar de hogere orde in de samenleving.<sup>11</sup> De statistiek hield zich niet bezig met het gedrag van individuen; dat gedrag was te variabel en te inconsistent. De statistiek spoorde de *wetten* op die een samenleving bepaalden. Hiermee bedoelde men de constante percentages misdaden en zelfmoorden, huwelijken, geboorten en gestorvenen. Om deze visie 'wiskundig' te rechtvaardigen had men het vaak over de wet van de grote aantallen. De Franse wiskundige S.D. Poisson (1781-1840) had een eigen versie van de wet van de wiskundige Jakob Bernoulli (1654-1705) geformuleerd en bewezen: een oneindige herhaling van gebeurtenissen leidt tot een stabiel gemiddelde ook al fluctueren de onderliggende kansen. Statistici uit de negentiende eeuw verwezen veelvuldig naar deze wet. Het feit, dat gebeurtenissen die veelvuldig voorkomen en niet te zeer afhankelijk van elkaar zijn, de neiging hebben om van jaar tot jaar in ongeveer constante aantallen op te treden, beschouwden zij als een manifestatie van deze wet.

Tot het interessegebied van de 'wetenschap der samenleving' behoorde ook kennis over de hygiënische toestanden waarin de mensen leefden. Er bestonden, onder andere in Groot-Brittannië en in Nederland, groepen artsen die zich om die reden voor statistiek interesseerden.<sup>12</sup> Deze zogenoemde hygiënisten maakten gebruik van statistische gegevens ter onderbouwing van hun visie op de medisch-hygiënische aspecten van de samenleving en zij vroegen zich af wat de betekenis van die gegevens was voor de volksgezondheid. Sterftegegevens waarvoor traditioneel altijd belangstelling had bestaan, speelden in deze visie een centrale rol. Uitsplitsing van deze gegevens naar verschillende doodsoorzaken moest relevante aanvullende informatie geven. Discussies over internationale unificatie van de registratie van doodsoorzaken zijn meerdere malen gevoerd op de internationale statistische congressen, overigens met wisselend succes.

10. Zie bijvoorbeeld in I. Hacking, *The taming of chance* (Cambridge, 1990) 113 en 125-127.

11. In Nederland was er al vroeg belangstelling voor statistiek. Vanaf 1802 werd er een vak aan de universiteit te Leiden gedoceerd dat statistiek heette. In 1857 werd een statistische vereniging opgericht, die door economisch geïnteresseerde juristen werd gedomineerd. De Duitse traditie van de statistiek speelde binnen deze omvangrijke vereniging tot ver in de negentiende eeuw een rol en de ideeën van Quetelet over een sociale fysica kregen er nauwelijks voet aan de grond. Zie verder I.H. Stamhuis, *Cijfers en aequaties' en kennis der staatskrachten. Statistiek in Nederland in de negentiende eeuw* (Amsterdam, 1989).

12. Voor Groot-Brittannië zie J.M. Eyler, *Victorian social medicine. The ideas and methods of William Farr* (Baltimore, 1979). Voor Nederland E.S. Houwaart, *De hygiënist, artsen, staat en volksgezondheid in Nederland 1840-1890* (Groningen, 1991).



Een apparaat, genaamd de 'quincunx', ontworpen door Galton om het optreden van de normale verdeling, die hij in het onderschrift de 'law of dispersion' noemde, te demonstreren. (Tisley & Spiller, 172 Frampton Rd.)

### Van de 'wetenschap van de samenleving' naar de genetica en de fysica

Vervolgens werd de statistiek opgenomen in vakgebieden waarin grote aantallen verschijnselen werden bestudeerd. De negentiende-eeuwse statistiek als wetenschap van de samenleving had laten zien dat een model kan worden gemaakt van talrijke autonome gebeurtenissen, waardoor statistische wetmatigheden aan het licht komen. De sociale statistiek heeft zo als voorbeeld gediend voor de toepassing van statistiek in bepaalde takken van de fysica en van de biologie.

Voor het gebruik van methoden uit waarschijnlijkheidsrekening en statistiek in de biologie was F. Galton (1822-1911) de pionier. Hij en zijn navolgers, die de biometrici werden genoemd, deden onderzoek naar evolutie, erfelijkheid en eugenetica.<sup>13</sup>

Galton bouwde voort op het werk van Quetelet. Galton, die was geïnteresseerd in erfelijkheid bij de mens, liet zien dat vele menselijke kenmerken normaal zijn verdeeld. Zijn visie op het gemiddelde was tegengesteld aan die van Quetelet: 'the rarity of commanding ability, and the vast abundance of mediocrity, is no accident, but follows of necessity,

13. Zie hiervoor o.a. het boek D.A. MacKenzie, *Statistics in Britain. 1865-1930, The social construction of scientific knowledge* (Edinburgh, 1981).

from the very nature of these things'.<sup>14</sup> De gemiddelde mens was geen ideaalbeeld maar een middelmatig persoon; bij Galton ging het om de zeldzame personen die uitblonken. Hij was ervan overtuigd dat erfelijkheid hierbij een belangrijke rol speelde. Hij was het daarom eens met zijn neef Darwin dat de afwijking van het gemiddelde, ofwel de variatie, voor de evolutie van cruciaal belang moest zijn geweest.

Galton verschilde echter radicaal met Darwin van mening over de aard van deze variatie. Volgens Darwin waren het de minutieuze afwijkingen van het gemiddelde, die door selectie over een voldoende aantal generaties voor een definitieve verschuiving van het populatiegemiddelde konden zorgen. De grote discontinue variaties, zogenaamde 'sports', waren niet relevant voor de evolutie, omdat waarden van eigenschappen volgens de erfelijkheidstheorie van Darwin uitmiddelen, en 'sports' kwamen te sporadisch voor om in dit proces een blijvende invloed te kunnen uitoefenen. Galton was ervan overtuigd dat het wel de 'sports' waren, die verantwoordelijk waren voor evolutie. Hij baseerde dat op de door hem geformuleerde 'Ancestral Law', die impliceert dat erfelijk materiaal van generatie op generatie ongemengd aan het nageslacht wordt doorgegeven. Op de continue variatie was volgens hem daarentegen de regressie naar het gemiddelde van toepassing, hetgeen betekende dat deze variatie geen blijvende verschuiving kon veroorzaken.

Galtons leerling K. Pearson (1857-1936) was het met Darwin eens, en de minutieuze afwijkingen die Darwin als verantwoordelijk voor de evolutie beschouwde, konden volgens Pearson worden geanalyseerd met behulp van de normale verdeling, die hij dan ook grondig bestudeerde. Volgens hem en andere biometrici waren de problemen met betrekking tot erfelijkheid en evolutie in hoge mate statistisch van aard. Pearson formuleerde het als volgt: deze vraagstellingen zijn 'in the first place statistical, in the second place statistical and only in the third place biological'.<sup>15</sup> En de bioloog W.F.R. Weldon (1860-1906) meende dat 'the questions raised by the Darwinian hypothesis are purely statistical, and the statistical method is the only one at present obvious by which that hypothesis can be experimentally checked'.<sup>16</sup> Pearson richtte het tijdschrift *Biometrika* op, dat zich ontwikkelde tot een medium waarin deze visie kon worden uitgewerkt. Uit de inhoud van dit tijdschrift blijkt dat de huidige mathematische statistiek haar wortels heeft in het werk van de biometrici.

De plaats van statistische methoden in biologisch onderzoek was binnen de biologie niet onbesproken. De visie van de biometrici op het belang van statistiek voor erfelijkheids-onderzoek riep verzet op in het kamp van de vroege mendelianen. De meesten waren van opleiding bioloog en vonden de biometrische benadering niet relevant. De biometrici beperkten zich met hun statistische benadering louter tot correlaties tussen groepen, terwijl de mendelianen causale verklaringen op individueel niveau noodzakelijk achtten. Zij vonden dat de fysiologische mechanismen die voor erfelijkheid verantwoordelijk waren, moesten worden blootgelegd. C.B. Davenport (1866-1944) bracht deze visie als volgt onder woorden: 'We are biologists, not applied statisticians'.<sup>17</sup>

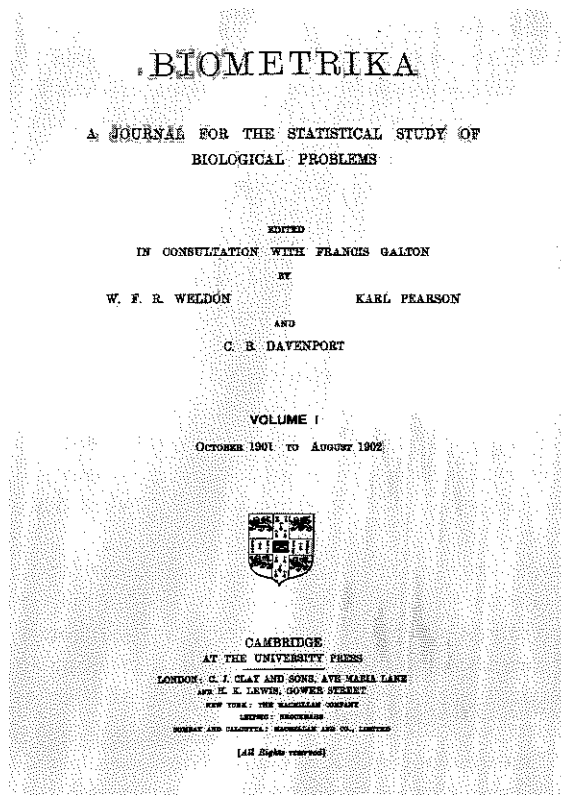
De mendelianen dachten inzicht in erfelijkheidskenmerken te verkrijgen door twee soorten

14. F. Galton, *Heredity Genius*, (London, 1869) 35, gec. in Gigerenzer (n.1), *The empire of chance*, 55.

15. Pearson to Galton, February 12, 1897, gec. in Gigerenzer (n.1), *The empire of chance*, 147.

16. W.F.R. Weldon, 'Attempt to measure the death-rate due to the selective destruction of *Carcinus moenam* with respect to a particular dimension', *Proceedings of the Royal Society* 58 (1895), 360-379, gec. in Gigerenzer (n.1), *The empire of chance*, 144.

17. C.B. Davenport, 'A discussion of the methods and results of dr. Heron's critique', *Eugenics Record Office Bulletin* 11 (1914), 3-24, m.n. 7, gec. in Gigerenzer (n.1), *The empire of chance*, 148.



Titelblad van het eerste deel van *Biometrika*.

organismen te kruisen die met betrekking tot een beperkt aantal erfelijke kenmerken verschillend waren. Zij bestudeerden aanvankelijk en in tegenstelling tot de biometrici alleen discontinu variërende kenmerken, dat zijn kenmerken die slechts een beperkt aantal waarden kunnen aannemen. Door kruisingen verkregen zij inzicht in het erfelijke gedrag van die specifieke kenmerken en probeerden de uitkomsten van hun kruisingen met de wetten van Mendel te verklaren. Hun experimenten leverden gegevens op die statistisch dienden te worden bewerkt. Zo verwachtten zij op grond van de mendeliaanse kruisingswetten in de tweede generatie bepaalde percentages voor verschillende groepen en hadden ze criteria nodig om te beslissen of experimentele en theoretische percentages voldoende overeenstemden. Voor de statistische interpretatie van hun experimentele resultaten waren ze volkomen afhankelijk van de biometrici. Door Pearson was al in 1900 de chi-kwadraat test op 'goodness of fit' gepubliceerd. Met deze test kon worden nagegaan of experimentele aantallen voldoende overeenkwamen met de aantallen die op grond van de theorie te verwachten waren. De chi-kwadraat test werd in 1912 door een zekere biometricus J.A. Harris toegepast op een mendeliaanse hypothese met de opmerking dat de mendelianen hiervan ook gebruik zouden moeten maken. Volgens *The empire of chance* werd het gebruik van deze toets, waarvan de mendeliaan en statisticus R.A. Fisher (1869-1962) in 1925 aantoonde dat deze adequaat was in het geval van een mendeliaanse hypothese, voor-

lopig nog geen gemeengoed in kringen van mendelianen.<sup>18</sup> Het zou tot in de jaren dertig duren vóór significantie-toetsen in de genetica waren ingeburgerd. Het toepassen van kans en statistiek in de biologie is dus controversieel geweest. Biometrici en mendelianen groeven zich zozeer in hun stellingen in, dat er eerst een nieuwe generatie onderzoekers moest komen voordat ingezien werd dat de oude tegenstellingen inmiddels achterhaald waren. Er was toen een nieuwe wetenschap genetica ontstaan, die experimenteel en kwantitatief was en waarin statistische methoden een belangrijke rol speelden. De statistiek had zich tijdens dit proces zo zeer ontwikkeld en geëmancipeerd dat de tegenwoordige mathematische statistiek voor haar ontstaan naar deze episode teruggrijpt.

Statistische noties deden hun intrede in de natuurkunde in het raamwerk van de kinetische gastheorie. Deze theorie, welke beoogde de macroscopische eigenschappen van een gas af te leiden uit de bewegingen en botsingen van gasmoleculen, maakte snel opgang in de jaren rond 1860. Een klassieke deterministische beschrijving van dergelijke processen was veel te complex om uitvoerbaar te zijn. Wanneer men deze processen zou gaan behandelen als statistische konden ze met behulp van kansrekening worden beschreven.<sup>19</sup> Maar dat vergde wel een andere wijze van denken in de fysica, die sinds de wetenschappelijke revolutie was gebaseerd op de deterministische mechanische wetten van Isaac Newton (1642-1727). Een bespreking door J.F.W. Herschel (1792-1871) van Quetelets werk uit 1846 over waarschijnlijkheidstheorie en de mogelijke toepassing in sociale en politieke wetenschappen speelde hierbij een sleutelrol.<sup>20</sup> Onder invloed van deze publikatie raakte de Engelse fysicus J.C. Maxwell (1831-1879) er namelijk van overtuigd dat moleculaire snelheden verdeeld waren volgens de kansverdeling waarmee in de sociale statistiek allerlei biologische en menselijke kenmerken werden beschreven. Tijdens een bijeenkomst in 1873 van de British Association of the Advancement of Science zei Maxwell hierover het volgende:

The modern atomists have therefore adopted a method, which is, I believe, new in the department of mathematical physics, though it has long been in use in the section of Statistics. ... the smallest portion of matter which we can subject to experiment consists of millions of molecules; so that we are obliged ... to adopt the statistical method of dealing with large groups of molecules. ... we meet a kind of regularity, the regularity of averages, which we can depend upon quite sufficiently for all practical purposes, but which can make no claim to that character of absolute precision which belongs to the laws of abstract dynamics.<sup>21</sup>

Met kans en statistiek konden zo processen worden beschreven die in principe wel deterministisch waren, maar te veelvuldig waren om als zodanig te worden behandeld. De fysicus L. Boltzmann (1844-1906) liet zich in 1872 in dezelfde zin uit. Kanstheoretische argumenten bleken ook in staat licht te werpen op een probleem dat direct voortkwam uit het werk aan de kinetische gastheorie. Maxwells statistische verdeling van de snelheden der moleculen had betrekking op een gas in evenwichtstoestand. Zijn afleiding verklaarde niet hoe het gas in die toestand terecht kwam. Volgens de thermodynamica correspondeert de verandering van een systeem in de richting van de evenwichts-

18. Gigerenzer (n.1), *The empire of chance*, 142-152.

19. Porter (n.7), *The rise of statistical thinking* 111-128 en 194-219 en Gigerenzer (n.1), *The empire of chance*, 163-202.

20. L.A.J. Quetelet, *Lettres à S.A.R. le duc régnant de Saxe Coburg et Gotha sur la théorie des probabilités, appliquée aux sciences Morales et Politiques* (Brussel, 1846), besproken in [J.F.W. Herschel], 'Quetelet on probabilities', *Edinburgh Review* 92 (1850) 1-57.

toestand met een irreversibele, ofwel niet-omkeerbare, toename van de entropie. Hier rees een levensgroot probleem. Immers, hoe kon men dit irreversibele gedrag van gassen, gecodificeerd in de tweede hoofdwet van de thermodynamica, rijmen met het strikt reversibele karakter van de bewegingen en botsingen van de gasmoleculen?

In een klassiek artikel, gepubliceerd in 1877, ontkende Boltzmann de absolute geldigheid van de tweede hoofdwet. Hij definieerde de entropie van een gas als een functie van de waarschijnlijkheid van de macroscopische toestand van dat gas. Voor elke microscopische toestand gold dat deze werd bepaald door de eigenschappen van de afzonderlijke moleculen. De waarschijnlijkheid van het optreden van een bepaalde macroscopische toestand van het gas was evenredig aan het aantal microscopische toestanden dat deze macroscopische toestand opleverde. De evenwichtstoestand verschijnt dan, niet als het noodzakelijke eindpunt van de ontwikkeling van het gas, maar als de meest waarschijnlijke toestand ervan. Afname van de entropie is niet onmogelijk, maar hoogst onwaarschijnlijk.

Nog fundamenteeler lijkt de rol van het kansbegrip in de quantumtheorie, die in de jaren twintig van onze eeuw werd ontwikkeld. Deze theorie beperkt zich tot uitspraken over de waarschijnlijkheid van de resultaten van metingen aan deeltjes. Volgens de meest gebruikelijke interpretatie van de theorie is er hier, anders dan in de statistische mechanica van Boltzmann, niet simpelweg sprake van gebrek aan informatie over de toestand van het systeem. De quantummechanische meetuitkomsten worden niet gedetermineerd door de toestand van het systeem voorafgaand aan het moment van meting. Het kansbegrip is hier irreducibel ofwel niet te elimineren.

Statistiek en kans hebben de fysica dus op een diepgaand niveau beïnvloed. Hun toepassing bleef niet beperkt tot de interpretatie van meetfouten, hoe belangrijk dat voor de ontwikkeling van een wetenschap ook kan zijn. Statistiek en kans werden geïntegreerd in de fysische theorie, en daartoe was inspiratie opgedaan bij de 'wetenschap der samenleving'.

Over de introductie van kans en statistiek in de psychologie zal ik kort zijn, omdat vóór 1940 statistiek in dit vakgebied een ondergeschikte rol speelde.<sup>21</sup> Een essentieel verschil met de tot nu toe beschreven wetenschappen is dat op het moment dat kansmodellen hun intrede deden in de psychologie, er intussen een apart vakgebied (mathematische) statistiek bestond.

De eerste periode in de opname van statistiek in de psychologie was tussen 1940 en 1955. Psychologen ontdekten toen het werk van Fisher, en met name zijn theorie over het toetsen van hypothesen en zijn variantie-analyse. Deze theorieën werden spoedig in cursussen onderwezen en in tekstboeken opgenomen. In deze periode werd statistiek gebruikt als instrument. In de jaren zestig vond er een radicale verandering plaats. Onderzoekers in de cognitieve psychologie gingen de menselijke geest beschouwen als een entiteit die statistisch te werk gaat. Bijvoorbeeld zou de geest, veelal onbewust, de omgeving benaderen aan de hand van het toetsen van hypothesen en het toekennen van intuïtieve kansen aan allerlei mogelijke alternatieven. Kortom kans en statistiek werden de instrumenten om het object van onderzoek in de psychologie te analyseren.

21. Geciteerd in Porter (n.7), *The rise of statistical thinking*, 111-112.

22. Zie G. Gigerenzer, 'The probabilistic revolution in psychology – an overview', en de vier artikelen in deel 2 van *The probabilistic revolution*, die over psychologie gaan, Krüger (n.2), *The probabilistic revolution*, vol.2, 7-9 en 11-100.

### Slot

In de negentiende en het begin van de twintigste eeuw namen kans en statistiek bezit van zoveel en ook zo verschillende wetenschappen en andere kennisgebieden, dat van een 'probabilistic revolution' kan worden gesproken.<sup>23</sup> I. Hacking onderscheidt vier stadia die elk voor zich een belangrijke conceptuele verschuiving laten zien en die gezamenlijk de 'probabilistic revolution' vormen.<sup>24</sup> De eerste speelde zich af tussen 1820 en 1840 en bestond uit de exponentiële toename van officiële getalsmatige informatie over de samenleving. Het was voor het eerst mogelijk hierin op grote schaal regelmatigigheden te ontdekken die met behulp van het begrip kans konden worden beschreven. Tijdens het tweede stadium (1835-1875) raakte men ervan overtuigd dat de regelmaat in deze getallen een algemeen geldige wetmatigheid weergaf, die men de wet van de grote aantallen noemde. Quetelets werk speelde in deze fase een belangrijke rol. Essentieel is dat tijdens deze fase waarschijnlijkheidswetten niet als een bedreiging van een deterministische kijk op de wereld werden beschouwd. Men was ervan overtuigd dat in principe het optreden van een kansverdeling door onderliggende oorzaken kon worden verklaard. Tijdens de derde fase (1875-1895) werd dat niet langer belangrijk gevonden. Galton had bijvoorbeeld opgemerkt dat ouders die een bepaalde eigenschap met een extreme waarde bezaten, in het algemeen kinderen kregen bij wie deze eigenschap een minder bijzondere waarde aannam. Hij verklaarde dit met de opmerking dat regressie naar het gemiddelde een logisch gevolg is van het feit dat het desbetreffende kenmerk normaal is verdeeld. Een verdere causale beschouwing werd niet nodig gevonden. De vierde fase (1892-1936) was die van het eigenlijke indeterminisme en heeft zich afgespeeld in de fysica. Deze fase was het resultaat van het ontstaan van de quantummechanica.

De stadia die Hacking onderscheidt geven enige belangrijke momenten weer van de zogenaamde 'probabilistic revolution' in de wetenschappen. Maar Hacking beschrijft vooral een revolutie op conceptueel niveau. Dit is één aspect van de betekenis van kans en statistiek zoals die in het voorafgaande is beschreven. De introductie van de Gauss-kromme in de sterrenkunde als de kansverdeling van meetfouten, de toepassing ervan in de 'wetenschap der samenleving' als de verdeling van de normale biologische variatie, waarbij de term 'normale verdeling' in zwang kwam, zijn voorbeelden van deze conceptuele veranderingen. Later gaf in de kinetische gastheorie de beschouwde kromme de verdeling van snelheden van moleculen weer, waardoor microscopische en macroscopische eigenschappen van gassen met elkaar in verband konden worden gebracht. Een conceptuele verandering vond plaats bij de introductie van kans en statistiek in de biologie. Daar werd aanvankelijk de variatie in plaats van het gemiddelde als het belangrijkste statistische begrip beschouwd. Ten slotte kreeg in de quantummechanica het begrip waarschijnlijkheid een essentieel andere betekenis dan het tot dan toe had gehad. Het was niet langer een uitdrukking van een gebrek aan kennis. Het kansbegrip was irreducibel geworden.

23. Op de verschillende visies op de kwestie wanneer in de wetenschappen sprake is van een revolutie zal ik hier niet ingaan. In het eerste deel van *The probabilistic revolution* is aan deze kwestie een aantal pagina's gewijd, gevuld met artikelen van resp. T.S. Kuhn, I.B. Cohen en I. Hacking: Krüger (n.2), *The probabilistic revolution*, vol.1, 7-55.

24. I. Hacking, 'Was there a probabilistic revolution 1800-1930?', Krüger (n.2), *The probabilistic revolution*, vol.1, 45-55.

Dit artikel heeft zich echter niet tot veranderingen van concepties beperkt. Duidelijk is geworden dat de behandelde wetenschappen zelf door de opname van kans en statistiek ingrijpend zijn veranderd. Verder is aan de orde geweest dat er vele nieuwe statistische technieken werden ontwikkeld, die eerst nog een betekenis hadden die tot het betreffende vakgebied was beperkt, maar later algemeen bleken te kunnen worden toegepast. Bijvoorbeeld in het onderzoek naar evolutie en erfelijkheid beschreef regressie naar het gemiddelde in eerste instantie erfelijkheid, later ontwikkelde regressie-analyse zich tot een algemene statistische techniek die velerlei toepassing had.

De introductie van kans en statistiek in de wetenschappen, zoals in dit artikel is beschreven, kan niet alleen aan de hand van conceptuele veranderingen worden gekarakteriseerd. De veranderingen in een breed scala van wetenschappen vormden daarnaast een factor van niet te onderschatten betekenis, terwijl de statistiek zich zelf tot een niet weg te denken wetenschap heeft ontwikkeld.

#### *Dankwoord*

Dr. A. de Knecht-van Eekelen dank ik voor nuttig commentaar, dr. F.H. van Lunteren voor hulp bij het schrijven van het gedeelte over de introductie van kanstheorie in de natuurkunde.

#### SUMMARY

##### *The probabilistic revolution in the sciences*

In this article is discussed how in the nineteenth century probability entered various sciences successively : astronomy, 'science of society', biology, physics and after the second World War psychology. To bridge the gaps between the sciences successful reasonings of analogy were used. During this process probabilistic concepts got other meanings. For example, the bell-shaped curve represented observational error in astronomy, then biological variation in the science of society and thereafter anarchic gas molecules in physics. The sciences changed radically as well. During the interaction with a science, new statistical techniques were developed which later became general applicable statistical tools. The stages of the probabilistic revolution, according to I. Hacking, which describe a revolution on a conceptual level, are mentioned. However, the probabilistic revolution in the sciences as described in this article proved to be too complex to characterise only in conceptual terms .