

## Waarneming, wiskunde en waarschijnlijkheid

**Citation for published version (APA):**

Doornbos, R. (1968). *Waarneming, wiskunde en waarschijnlijkheid*. Technische Hogeschool Eindhoven.

**Document status and date:**

Gepubliceerd: 01/01/1968

**Document Version:**

Uitgevers PDF, ook bekend als Version of Record

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

WAARNEMING, WISKUNDE  
EN  
WAARSCHIJNLIJKHEID

DR. R. DOORNBOS

WAARNEMING, WISKUNDE  
EN  
WAARSCHIJNLIJKHEID

REDE

UITGESPROKEN BIJ HET AANVAARDEN VAN  
HET AMBT VAN HOGLERAAR  
IN DE WISKUNDE  
AAN DE TECHNISCHE HOGESCHOOL  
TE EINDHOVEN  
OP VRIJDAG 11 OKTOBER 1968

DOOR

DR. R. DOORNBOS

*Mijne heren curatoren,  
Mijne heren hoogleraren en lectoren,  
Dames en heren leden van de wetenschappelijke,  
de technische en de administratieve staf,  
Dames en heren studenten,  
en voorts gij allen die door Uw aanwezigheid van  
Uw belangstelling blijke geeft,*

*Zeer gewaardeerde toehoorders,*

Het is mijn bedoeling U aan de hand van een aantal voorbeelden een indruk te geven van de problemen die zich voordoen bij het toepassen van de mathematische statistiek. Voor een beter begrip zal ik daarbij eerst het werkterrein van de statisticus afbakenen en tegelijkertijd de gedachtengang toelichten die aan de statistische werkwijze ten grondslag ligt. Wij kunnen de mathematische statistiek kenschetsen als de wetenschap die ons in staat stelt conclusies te trekken uit waarnemingsmateriaal en op grond van deze conclusies beslissingen te nemen in onzekere situaties. Niet alleen in de ervaringswetenschappen, maar ook, hoewel minder systematisch, in het dagelijks leven, gaan wij op een dergelijke wijze te werk. Wij doen waarnemingen, wij rangschikken deze waarnemingen en trachten hierin bepaalde systemen te ontdekken. De aldus opgemerkte wetten worden gegeneraliseerd en vormen de basis voor ons denken en handelen. Wij nemen zonder meer aan dat dergelijke vaste wetten bestaan en dat ze alleen maar opgespoord hoeven te worden om er ons voordeel mee te kunnen doen. Zoals L. E. J. Brouwer het uitdrukt in zijn boek „Over de grondslagen der Wiskunde”: „Den menschen is een vermogen eigen dat al hun wisselwerkingen met de natuur begeleidt, het vermogen nl. tot *wiskundig bekijken* van hun leven, tot het zien in de wereld van herhalingen van volgreesken, van causale systemen in de tijd”.

Worden de gevonden wetten in een wiskundige vorm gegoten, dan spreekt men van een wiskundig model. Dit maakt het mogelijk allerlei wiskundige bewerkingen toe te passen die kunnen leiden tot nieuwe wetten of liever: hypothesen. Het is duidelijk dat de geldigheid van deze wetten op deze manier nooit kan worden bewezen. De doel-

treffendheid van op deze wijze genomen besluiten staat of valt met de aansluiting van het uitgangsmodel met de werkelijkheid. De mathematicus kan zich formeel op het standpunt stellen dat hem dit niet aangaat: hij bewerkt waarnemingsmateriaal en stelt wetten op. De verantwoordelijkheid voor de praktische toepassing ligt geheel bij de waarnemer of onderzoeker. De statisticus zal zover niet kunnen gaan; hij zal samen met de waarnemer moeten nagaan of de voorwaarden voor de toepasbaarheid aanwezig zijn. De statisticus is dus niet uitsluitend beoefenaar van de wiskunde die Huxley als volgt beschreef: „Mathematics is that study which knows nothing of observation, nothing of experiment, nothing of induction, nothing of causation”. Vele natuurverschijnselen die met een wiskundig model kunnen worden beschreven, zijn volledig gedetermineerd. De tijdstippen waarop zon en maan ondergaan op vrijdag 11 oktober 1968 heeft men eeuwen van tevoren precies kunnen berekenen. Het weer op deze dag daarentegen moet bij de huidige stand van wetenschap als niet van tevoren berekenbaar en als alleen binnen zekere grenzen voorspelbaar worden beschouwd. Dit laatste verschijnsel is van stochastische aard: het bevat een deel dat zodanig variabel is, dat in een bepaald geval niet exact van tevoren bekend is wat de uitkomst zal zijn. Gedeeltelijk kan deze onzekerheid het gevolg zijn van onze gebrekkige kennis die, althans in principe, aangevuld zou kunnen worden. Zo is het wel zeker, dat het weer van morgen beter voorspeld zou kunnen worden als het aantal waarnemingsstations zeer sterk zou worden uitgebreid. Heeft men dan een model, waarin de belangrijkste grootheden zoals luchtdruk, temperatuur, vochtigheid, windsnelheid en windrichting op een groot aantal hoogten voorkomen, dan is het mogelijk om met behulp van de modernste rekenapparatuur verwachtingen op te stellen voor drie dagen die nauwkeuriger zijn dan de huidige voorspellingen voor één dag. Door de praktische grenzen van waarnemings- en rekenmogelijkheden zal er echter een belangrijk element van onzekerheid blijven bestaan.

Het toevallige element in een fysisch verschijnsel als het weer heeft bovendien nog een meer fundamentele grondslag. Volgens de onzekerheidsrelatie van Heisenberg, geformuleerd in 1927, kan men van een deeltje, bijv. een zuurstofmolecuul, niet tegelijkertijd met volstrekte nauwkeurigheid de plaats en de snelheid bepalen. Deze onnauwkeurigheid is echter zo klein, dat de praktische betekenis ervan in ons voorbeeld, waar wij te maken hebben met zeer grote aantallen deeltjes, te verwaarlozen is. Probeert men het menselijke gedrag te meten en te voorspellen, dan stuit men op soortgelijke problemen. Goudriaan voert in zijn boek „Economie in Zestien Bladzijden

(Inleiding tot de analytische economie)" drie axioma's in, waarvan het eerste luidt: „Een handeling van één mens, niet gebonden door externe dwang of vaststaande traditie, is onbepaald". De onzekerheidsrelatie van Heisenberg heeft hier een analogon dat ook deze onbepaaldheid onherroepelijk maakt. Probeert men door vragen de plannen of meningen van individuen te weten te komen dan worden deze meningen daardoor beïnvloed. Zowel bij fysische als bij economische verschijnselen beheerst dus het toeval bij de kleinste eenheden het wereldbeeld. Dit betekent niet dat over deze zaken geen enkele uitspraak meer gedaan kan worden. Meteorologen, markt- en opinie-onderzoekers bewijzen voortdurend het tegendeel. Om U duidelijk te maken op welke methode deze uitspraken berusten zal ik hem toelichten aan de hand van het klassieke model dat de statistici hiervoor gebruiken: de vaas (ook wel genoemd: urn) met zwarte en witte ballen.

Dit is een model dat kan worden toegepast in alle situaties, waar twee uitkomsten mogelijk zijn: een pasgeborene is een jongen of een meisje, een Amerikaanse kiezer kiest democratisch of republikeins, een pakje margarine bevat het vermelde gewicht of het is te licht, de temperatuur bereikt op een bepaalde dag zomerse waarden of het blijft kouder. Gaat men uit van bekende aantallen zwarte en witte ballen in de vaas, dan kan met behulp van de waarschijnlijkheidsrekening bepaald worden hoe groot de kansen zijn op verschillende uitkomsten in een steekproef van een zeker aantal ballen uit de vaas. Weet men de verhouding tussen jongens- en meisjesgeboorten, dan is gemakkelijk te berekenen hoe groot de kans is, dat er in een gezin met vier kinderen precies twee jongens en twee meisjes zijn.

In een plaatselijk dagblad kon men onlangs naar aanleiding van de Amerikaanse presidentsverkiezingen het volgende lezen: „De techniek van de opiniepeilers is tamelijk eenvoudig. George Gallup bijvoorbeeld baseert zich op de waarschijnlijkheidstheorie, die ongeveer leert dat als men honderdduizend bonen, de helft witte en de helft bruine, in een Keulse pot stort en men er vervolgens 100 uithaalt, men bijna altijd vijftig witte en vijftig bruine bonen heeft met een kans van drie procent om zich daarin te vergissen". U ziet dat het vaasmodel ook in de pers de nodige aandacht krijgt. Ik hoop echter dat U met mij twijfelt aan de uitspraak dat op de honderd bonen „bijna altijd" vijftig witte voorkomen. De kans op deze uitkomst blijkt namelijk niet zoals men ons wil doen geloven 97 procent, maar slechts ongeveer 8 procent te bedragen. Bovendien helpt het de opiniepeiler weinig als hij alleen maar weet wat een trekking op kan

leveren als hij weet wat er in de vaas zit. Zijn probleem is het omgekeerde. Hij heeft een greep in de vaas gedaan, de bonen of de ballen geteld en moet nu zeggen hoe de verhouding tussen wit en zwart in de vaas is, of in statistische taal: men doet waarnemingen aan een *steekproef* en trekt daaruit conclusies omtrent de *populatie*. Daarbij komt dan nog de complicatie dat de inhoud van de vaas, in dit geval de mening van de kiezer, nog kan veranderen tussen het moment van de steekproef en de dag van de verkiezing.

Dit omgekeerde probleem: het trekken van conclusies uit een steekproef, uit onvolledige waarnemingen, is het terrein van de wiskundige statistiek. Het is duidelijk dat gevolgtrekkingen over het algemene geval uit het speciale geval risico's van een verkeerd oordeel inhouden en dat uitspraken die op deze wijze tot stand komen, onzekerheden bevatten. De statistische methoden waarmee we gevolgtrekkingen uit waarnemingen maken, worden gewoonlijk in twee groepen verdeeld: het schatten van parameters en het toetsen van hypothesen. Voor een voorbeeld van de eerste methode nemen we nog eens de vaas met witte en zwarte ballen. Dit model bevat als enige onbekende parameter het percentage witte ballen. De meest voor de hand liggende schatting hiervoor is het percentage witte ballen in de steekproef. Het vervelende is echter dat deze schatting niet precies gelijk is aan de onbekende parameter. Hieraan kan worden tegemoetgekomen door twee grenzen aan te geven waartussen het percentage zich waarschijnlijk bevindt. Men noemt dit gebied een betrouwbaarheidsinterval voor de parameter. Naarmate dit interval, het verschil tussen onder- en bovengrens, groter is, is de kans dat de onbekende parameter erin ligt uiteraard groter. De informatie is dan echter geringer, terwijl een uitspraak gebaseerd op een kleiner interval weer een lagere betrouwbaarheid heeft. Het is precies als met een weersvoorspelling, die betrouwbaarder kan zijn naarmate ze vager is. Volgens het eerder geciteerde krante-artikel verklaarde Gallup dat hij de uitslag van de presidentsverkiezingen met een marge van anderhalf procent kon voorspellen. De betrouwbaarheid van deze uitspraak, dat is de kans dat ze niet wordt gelogenstraft, werd hierbij niet vermeld. De andere groep van statistische methoden, het toetsen van hypothesen, komt bijvoorbeeld aan de orde, wanneer de uitslagen van twee opinie-onderzoeken die zijn gehouden met een tussenpoos van enkele maanden, met elkaar worden vergeleken. Men wil dan weten of er in deze tussentijd iets is veranderd en toetst dan merkwaardigerwijze de hypothese: de opinie is precies gelijk gebleven. Deze hypothese wordt de nulhypothese genoemd. Is de uitkomst van de beide steekproeven hiermee in strijd, dan wordt de nulhypothese verwor-

pen. Het toetsen van een hypothese komt ook in de toepassingen van de statistiek in de industrie zeer veel voor, zoals ik U aan de hand van een ander voorbeeld hoop aan te tonen. Denken wij ons een machine die elke seconde een pakje margarine vult. Op ieder pakje staat een gewicht van 250 gram aangegeven, dat wil zeggen dat volgens de warenwet elk pakje minstens 250 gram moet bevatten. Daar het onmogelijk is om tijdens de fabricage kleine gewichtsverschillen te vermijden kan het voorkomen van te lichte pakjes alleen worden vermeden door gemiddeld een zeker overwicht te geven. Uit dit overwicht en uit de bekende onnauwkeurigheid van het fabricageproces worden grenzen berekend waartussen de gewichten kunnen variëren. De man die de machine bedient, kan de hoeveelheid margarine per pakje regelen door aan een wiel te draaien. Van tijd tot tijd neemt hij een pakje van de lopende band, legt het op een weegschaal en controleert het gewicht. Ligt dat tussen de eerdergenoemde grenzen, dan reageert hij verder niet. Is het pakje te zwaar of te licht dan wordt de machine bijgesteld. Zonder het te weten toetst de gewichtscontroleur „aan de lopende band” een hypothese. Deze hypothese is: gemiddeld hebben de pakjes het juiste gewicht, dat wil zeggen 250 gram plus het gewenste overwicht.

Drie factoren bepalen de afstand tussen de grenzen, waarbinnen niet wordt ingegrepen of, in statistische taal, waarbinnen de hypothese niet wordt verworpen. Dit zijn in de eerste plaats toevallige gewichtsvariatiës, bijv. tengevolge van kleine spelingen in de apparatuur. Verder zijn er nog twee risico's die we lopen. Het eerste risico treedt op wanneer het proces goed loopt. Het gemiddelde gewicht heeft de juiste waarde, maar de toevallige schommelingen resulteren in een pakje dat precies buiten de grens valt. Er wordt dan ten onrechte bijgesteld en het vulproces wordt juist in de war gestuurd. Nu is dit gemakkelijk te verhelpen door de grenzen wat ruimer te nemen. Maar U ziet al wat er dan gebeurt: een afwijking moet nu al heel groot zijn om ontdekt te worden en allerlei onregelmatigheden treden niet of te laat aan het licht. Het tweede risico, nl. dat van niet ingrijpen als het wel nodig is, is nu groter geworden. U ziet dat er een duidelijke analogie bestaat met de situatie die zich voordeed bij het betrouwbaarheidsinterval: er moet een compromis worden gezocht tussen te vaak ingrijpen of te veel foute uitspraken aan de ene kant en te weinig ingrijpen en nietszeggende uitspraken aan de andere kant. Beide soorten risico's kunnen tegelijkertijd worden verkleind door de steekproef groter te nemen, als de mogelijkheid daartoe ten minste tot onze beschikking staat. Men zal dus op de een of andere wijze de consequenties van deze risico's en de kosten van de steekproef tegen

elkaar af moeten wegen.

In ons voorbeeld is het duidelijk welke aktie moet worden ondernomen als een van de kritieke grenzen wordt overschreden. Is het pakje te zwaar, dan wordt de kraan wat dichtgedraaid, is het te licht dan wordt de hoeveelheid per pakje iets vergroot. Hoe groot deze korrekcie moet zijn, laten we nu maar in het midden. Dit komt, omdat de mogelijkheden voor het geval dat de getoetste hypothese niet juist is, hier zeer overzichtelijk zijn: het gemiddelde gewicht is ofwel te hoog, ofwel te laag. Bij zeer veel praktische problemen is de situatie veel minder duidelijk.

Ik wil trachten dit ook weer met een aantal voorbeelden aan te tonen. Vanaf mijn kamer in het hoofdgebouw van de T.H. kijk ik uit op de parkeerplaats naast het auditorium. Op een woensdag om 15,00 uur heb ik waargenomen dat er van de 180 plaatsen 160 waren bezet. Ik wil nu de hypothese toetsen dat elke automobilist bij aankomst een willekeurige keuze doet uit alle beschikbare plaatsen, geheel onafhankelijk van de ligging van deze plaatsen. Ik bereken daartoe de kans dat in zo'n geval juist die 20 plaatsen zouden zijn overgebleven, die op de bewuste woensdag nog vrij waren. Die kans is zo klein dat deze configuratie zich hoogstwaarschijnlijk nog nooit zou hebben voorgedaan, als sinds het ontstaan van de aarde elke seconde door loting 20 uit de 180 plaatsen waren gekozen. Het is wél gebeurd en dus is de hypothese van de toevallige verdeling fout. Die veronderstelling was ook niet erg aannemelijk en dus is alles in orde. Het is alleen vervelend dat deze conclusie op deze wijze uit elke waargenomen situatie kan worden getrokken.

Nu heb ik echter ook opgemerkt dat 15 van de 20 lege plaatsen zich bevonden in de achterste twee rijen, die maar 45 van de 180 parkeerplaatsen bevatten. Dus 75% van de lege plaatsen bevindt zich op 25% van het parkeerterrein. Dat kan toch geen toeval meer zijn? Als ik de statisticus hierover raadpleeg, berekent hij dat de kans hierop kleiner is dan één op de vijf millioen. Niet zo klein als daarnet, maar toch klein genoeg om de willekeurige verdeling voor goed vaarwel te zeggen? Hoewel het verwerpen van de willekeurige verdeling gezien de ongunstige ligging van de bewuste rijen en de bekende afkeer van lichamelijke inspanning van de gemiddelde automobilist verleidelijk is, moeten wij toch nog een paar dingen overwegen.

In de eerste plaats zou het argument tegen de willekeurige verdeling nog sterker zijn geweest, als niet 15 van de 20, maar 16, 17, 18, 19 of alle 20 lege plaatsen zich in de laatste rijen hadden bevonden. Ook in die gevallen zou ik de hypothese hebben willen verwerpen en daarom moet ik de kansen op deze waarnemingen optellen bij de al gevonden

kans op precies 15 om te voorkomen dat een voorbarig oordeel wordt geveld.

Er is echter nog een meer fundamenteel bezwaar tegen de gevolgde redenering aan te voeren. Immers, de bewuste rijen zijn mij tijdens het waarnemen pas als bijzonder opgevallen. Waren de lege plaatsen voornamelijk in een ander tweetal rijen voorgekomen, dan was er misschien wel een redenering te vinden om juist die keuze te verklaren. Een dergelijke verklaring achteraf zou ook gevonden kunnen worden wanneer de lege plaatsen voorkwamen in één rij, voor het midden van het terrein, voor de schaduwzijde of de zonzijde enz. Hieruit kunnen we concluderen dat de gevallen die tot verwerping van de nulhypothese leiden, vóór het doen van de waarneming moeten worden gekozen. Eventueel kan men zich daarbij door voorafgaande oriënterende uitkomsten laten leiden, maar het bewijs dat de nulhypothese onjuist is, kan dan pas in tweede instantie door nieuwe onafhankelijke waarnemingen worden geleverd. Dit voorbeeld was niet van groot praktisch belang, maar soortgelijke moeilijkheden doen zich, misschien iets minder duidelijk, voor in vele situaties die wél praktische betekenis hebben.

Als voorbeeld nemen we het vergelijkend warenonderzoek zoals het hier te lande wordt verricht door de Nederlandse Consumentenbond. Van allerlei zaken zoals pakken celstof luiers, koekepannen, foto-toestellen en flessen sinaasappelsap worden één of meer exemplaren van verschillende merken gekocht en op één of meer eigenschappen onderzocht. Een bepaald onderzoek had betrekking op het vitamine C-gehalte van sinaasappelsap. Het doel van een dergelijk onderzoek is tweeledig: het nagaan of er sprake is van verschillen tussen de merken en het op overzichtelijke wijze meedelen van deze verschillen aan de consument. De statistische complicaties bij een dergelijk onderzoek zijn van tweeërlei aard. In de eerste plaats is er zoals steeds de stap van steekproef naar populatie, in ons geval van fles naar merk. Waarschijnlijk is van elk merk het vitamine C-gehalte niet constant. Dit hangt o.a. af van de soort sinaasappels, van de rijpheid, van de omstandigheden tijdens het transport, van de duur van het transport; verder van de temperatuur tijdens de bereiding, van de invloed van licht en temperatuur in de winkel en van de tijd dat de fles in de winkel staat. Bovendien kan de bepaling van het vitamine C-gehalte niet met absolute nauwkeurigheid worden uitgevoerd. Er kunnen verschillen optreden tussen twee bepalingen bij hetzelfde monster, verricht door verschillende analisten of door één analist op twee verschillende dagen. Kortom, er zijn allerlei oorzaken aan te wijzen, die er toe kunnen leiden dat het vitamine C-gehalte bij twee flessen

van één merk uiteen kan lopen. En wij zullen alleen mogen concluderen dat het ene merk beter is dan het andere, als de verschillen zo duidelijk zijn, dat ze boven de variaties binnen de merken zelf kunnen worden geconstateerd. Dit betekent dat de steekproeven moeten bestaan uit voldoende aantallen flessen, voldoende gespreid in de tijd en over verschillende winkels om de spreiding binnen een merk te kunnen waarnemen. Wat bovendien de lezer van het verslag interesseert, is niet het vitamine C-gehalte in de steekproef, maar in de flessen die hij in de toekomst zal gaan kopen. Men moet dus aannemen, wil het onderzoek enige zin hebben, dat het beeld dat de merken in het verleden hebben vertoond, gedurende zekere tijd zal blijven bestaan. Hieruit blijkt nog eens hoe nodig het is de variaties die in de tijd optreden, in de steekproef toe te laten. De eisen die de mathematische theorie stelt, zijn veel strenger. Alle methoden zijn slechts toepasbaar als de steekproef kan worden beschouwd als door loting verkregen uit de gehele populatie. Als een gedeelte van de populatie in de toekomst ligt, kan hieraan uiteraard niet worden voldaan.

Zo zit er dus in deze toepassing een element van voorspellen, waarbij de theorie ons niet kan helpen. Bij de gewichtsbeheersing van margarine is het al niet anders: de controleur regelt niet bij omdat het gewicht fout is, maar omdat het bij niet bijregelen fout *you zijn*; hij voorkomt afwijkingen. Men moet voorspellingen, ook als ze op statistische methoden zijn gebaseerd, dus steeds met een zekere mate van voorzichtigheid hanteren. Zeker kan van de statisticus niet geëist worden, dat hij uit een verkeerde voorspelling dezelfde consequentie trekt als professor Bosbijtelt in het verhaal „Eind goed al goed” van Belcampo. Als zijn voorspelling van het aantal zelfmoorden in een bepaald jaar op oudejaarsavond één te hoog dreigt te blijven, vindt zijn huishoudster hem de volgende morgen dood aan zijn bureau. Naast hem ligt een papier waarop de volgende woorden:

#### Laatste wil

Ik wil dat in de grote bladen van ons land de volgende advertentie wordt opgenomen:

Op dinsdag 31 december 's-avonds 1 minuut voor 12 heeft Johannes Bosbijtelt, hoogleeraar aan de Universiteit van Amsterdam, zacht en kalm de hand aan zichzelf geslagen. Geen bezoek. Geen bloemen.

In het verslag van het onderzoek van sinaasappelsap is een gemakkelijker weg gekozen. Daarin wordt alleen gerapporteerd wat in de onderzochte flessen is aangetroffen, waarbij niets wordt vermeld

over de grootte van de steekproef en over de wijze waarop de steekproef is genomen. Er wordt dus aan de lezer overgelaten om de generalisatie van steekproef naar populatie zelf te maken, waarvoor alle gegevens echter ontbreken.

Er is echter bij een onderzoek als dit nog een tweede moeilijkheid, waarvoor de statistische theorie nog géén bevredigende oplossing biedt. De vraag is namelijk hoe dit probleem kan worden geformuleerd als de toetsing van een hypothese tegen één of meer andere hypothesen. Over de nul-hypothese kan weinig verwarring bestaan: alle merken hebben een gelijk vitamine C-gehalte. Er bestaan dus alleen toevallige variaties van fles tot fles, maar geen systematische verschillen tussen de merken. De moeilijkheid komt na het verwerpen van deze nulhypothese. In de klassieke toetsingstheorie luidt de conclusie dan: de merken zijn niet alle gelijk. Dat is natuurlijk een antwoord waarvan de geïnteresseerde consument de zin terecht in twijfel zal trekken. Want op een enkele scepticus na zal iedereen er wel bij voorbaat van overtuigd zijn geweest, dat de merken behalve verschillen in naam, kleur van het etiket, vorm van de fles en prijs ook wel verschillen in kwaliteit zullen hebben. Men verwacht van een warenonderzoek dat het antwoord geeft op vragen als: is er één merk dat duidelijk beter is dan alle andere, is er misschien een merk dat veel goedkoper is en maar iets minder goed dan het beste, is er een groepje dat boven de anderen uitsteekt, of zijn ze allemaal goed op een kleine uitzondering na, zijn er twee, drie of meer groepen aan te wijzen etc. Er bestaat dus een keuze tussen zeer veel mogelijke beslissingen en het is nog veel moeilijker dan bij het parkeerprobleem om hier van tevoren een zinvolle structuur in te brengen. Ook hier is het onjuist zich bij de toetsing te laten leiden door de uitkomst van het experiment. Want ook al zouden de negenentwintig merken waarom het ging, nauwelijks uiteen lopen, dan nog is er in de steekproef een groot verschil tussen het merk dat toevallig als beste en dat toevallig als slechtste van de 29 uit de bus komt.

Dit probleem, dat in de statistiek bekend staat onder de naam *multiple comparisons*, komt zeer veel voor. Het heeft de laatste decennia al zeer veel aandacht gekregen, waarbij vooral de namen van Duncan, Scheffé en Tukey genoemd moeten worden. Een volledig bevredigende methode is tot op heden echter niet beschikbaar.

Om ons voorbeeld af te ronden moet ik U nog vermelden hoe in dit concrete geval de resultaten werden gepubliceerd: de gevonden waarden werden gerangschikt in vier groepen, namelijk 40 tot 60 mg, 20 tot 40 mg, 10 tot 20 mg en minder dan 10 mg vitamine C per 100 ml. Deze vier groepen werden respectievelijk met goed, redelijk,

matig en slecht aangeduid. Bij twee merken werd een hoger gehalte gevonden dan in sinaasappels voorkomt, hetgeen wijst op een (verboden) toevoeging van kunstmatig vitamine C. De vragen die zich hierbij voordoen, zijn: zijn deze groepen van tevoren gekozen of op grond van de resultaten en welke zekerheid bestaat er dat de slechtste uit de goede groep werkelijk beter is dan de beste uit de redelijke groep?

Soortgelijke problemen doen zich overal voor waar een selectie moet worden gedaan, uit een reeks van dingen, of het nu tarwerassen, soorten veevoer, parfums voor toiletzeep, geneesmiddelen of productieprocessen betreft. Een boeiend terrein dus voor de statisticus die zich zowel voor de theorie als voor de toepassing van zijn vak interesseert. Het zal echter verre van eenvoudig zijn een synthese te vinden tussen wat men de Engelse school in de statistiek zou kunnen noemen, waarvan Fisher de grondlegger is, en de Amerikaanse van Neyman en Wald. Volgens de eerste is het trekken van conclusies uit waarnemingen iets fundamenteel anders dan het nemen van beslissingen, volgens de laatsten zijn veel problemen fundamenteel meervoudige beslissingsproblemen, die alleen traditioneel als toetsingsproblemen zijn geformuleerd. De Engelsman Kendall gaat in een herdenkingsartikel, gewijd aan Fisher, zover te zeggen: „Some difference of attitude is inevitable between countries where what a man does is more important than what he thinks, and those where what he thinks is more important than what he does”.

U ziet dat men zich zelfs in de statistiek bezint op de zedelijke en maatschappelijke aspecten van de wetenschapsbeoefening! Ik ben geneigd, afgezien van deze filosofische achtergronden, de besluitkundige formulering als de meest natuurlijke te beschouwen. Hiermee is echter nog geen oplossing bereikt, daar de zuiver wiskundige moeilijkheden hiervoor tot dusver nog te groot zijn gebleven.

In de statistiek is het nog minder zinvol dan in andere onderdelen van de wiskunde een strenge scheiding te maken tussen zuivere en toegepaste wetenschap of zelfs tussen theorie en praktijk. Een mathematicus kan zich alleen dan statisticus noemen, als hij belangstelling heeft voor reële problemen en bereid is zijn kennis op deze problemen toe te passen. Hij moet hierbij soms veronderstellingen durven maken waarvan de juistheid niet vaststaat. Van Dantzig, begaafd wiskundige en tevens de man die het meest heeft bijgedragen tot de ontwikkeling van de statistiek in Nederland, zei het in 1954 in zijn lezing „De verantwoordelijkheden van de Statisticus” als volgt: „De statisticus heeft echter ook een maatschappelijk geweten, dat hem gebiedt onderzoekers op velerlei gebied althans zo goed mogelijk te

helpen, zij het zonder volstrekte zekerheid, dat zijn conclusies juist zijn. Het maatschappelijk geweten zegt hem dus, niet al te kritisch te zijn, maar genoeg te nemen met een redelijke mate van zekerheid. Hij moet een compromis sluiten tussen zijn wetenschappelijk en zijn maatschappelijk geweten; de statisticus is dus in de kern gespleten: hij lijdt aan kernsplitsing”.

Bestaat deze splitsing alleen bij de afzonderlijke beoefenaars van de statistiek, dan blijft het best mogelijke contact tussen de beoefening van de wiskunde en die van haar toepassing wel gewaarborgd. Gevaarlijker wordt de situatie, wanneer er gescheiden groepen beoefenaars ontstaan van theoretische en van toegepaste statistiek, wier wegen steeds meer divergeren. Sinds de tijden van Karl Pearson, Sir Ronald Fisher en „Student”, die evenzeer onderzoekers waren in de toepassingsgebieden van de statistiek als scheppers van nieuwe methoden, is er inderdaad veel veranderd. Er is een statistische literatuur ontstaan, waarvan alleen al de notatie en de formulering aan veel practisch ingestelde statistici de moed en de lust ontnemen ervan kennis te nemen. Niet iedereen wil daarbij zover gaan als een goede vriend van mij, een bekende Engelse statisticus, die mij eens verzekerde dat naar zijn mening alle exemplaren van de *Journal of the Royal Statistical Society, Series B* het best konden worden gebruikt om jaarlijks een vreugdevuur op Trafalgar Square te stoken. In de eerste plaats immers kan men betogen dat het hoogste doel van de wetenschap is: het vermeerderen van kennis en niet het dienen van een direct aanwijsbaar practisch doel. En de wiskunde is vanouds een wetenschap waarin dit ideaal kan worden uitgeleefd. Veel wiskundig onderzoek wordt nog steeds om haarzelfs wil bedreven, ook al betreft het maattheorie die via de waarschijnlijkheidsrekening aan de statistiek ten grondslag ligt. En zo kunnen er dus in tijdschriften, die in de titel de naam statistiek voeren, artikelen voorkomen waarvan de toepassingen niet onmiddellijk voor de hand liggen. Misschien wel omdat de auteurs met de befaamde Engelse wiskundige Hardy vinden, dat „it is what is commonplace and dull that counts for practical life”. Verder is het vaak moeilijk te voorspellen welke ideeën binnen afzienbare tijd practisch nut zullen afwerpen, zodat een doeltreffende scheiding tussen artikelen van alleen theoretisch belang en de onmiddellijk toepasbare niet is te trekken. Uit de 270 artikelen, die het vorig jaar verschenen in de *Annals of Mathematical Statistics*, een van de vijf of zes tijdschriften die de statisticus geacht wordt te lezen, zal hij dus de voor hem belangrijke moeten zien te vinden, zonder alles van Abrahamson tot Zyskind en van „Een tekentoets voor lineariteit”

tot „Een perfect meetbare ruimte die geen Lusin-ruimte is”, uitvoerig te kunnen bestuderen.

Ik heb mij deze kleine uitweiding aan het slot van mijn betoog veroorloofd, omdat de benoeming van een statisticus tot hoogleraar in de wiskunde naast de sinds 1960 bestaande leerstoel in de Toegepaste Statistiek tot de misvatting zou kunnen leiden, dat van nu af aan hier in Eindhoven twee soorten statistiek zouden worden beoefend die zeer weinig met elkaar gemeen zouden hebben. Dit nu ligt allerm minst in mijn bedoeling. Ik geloof dat er juist in de statistiek nog steeds plaats is voor theoretisch werk, dat nauw samengaat met praktische onderzoekingen en waar zou hiervoor een betere plaats te vinden zijn dan aan een Technische Hogeschool?

Aan het einde van mijn beschouwingen gekomen wil ik allereerst mijn eerbiedige dank betuigen aan Hare Majesteit de Koningin die mij tot Hoogleraar aan deze Technische Hogeschool heeft willen benoemen.

*Mijne heren curatoren,*

Het is mij een voorrecht U mijn oprechte dank te betuigen voor het vertrouwen dat U in mij hebt gesteld door mij voor deze benoeming voor te dragen. Ik verzeker U dat ik zal trachten Uw vertrouwen niet te beschamen.

*Mijn heren leden van de senaat,*

Het verheugt mij zeer in Uw kring te zijn opgenomen. Het is U uit mijn betoog duidelijk geworden dat ik ook voor de wiskundig statisticus contact met de praktijk uiterst belangrijk acht. Naar ik hoop zal ik in de toekomst de gelegenheid hebben met velen van U tot een vruchtbare samenwerking te komen.

*Mijne heren hoogleraren van de onderafdeling der wiskunde,*

Voor de wijze waarop ik in Uw midden ben ontvangen, zeg ik U hartelijk dank. Deel uit te mogen maken van Uw uitgelezen groep is een eer die tot dankbaarheid, maar ook tot bescheidenheid stemt. In de mathematische statistiek zien we ons gesteld voor vele wiskundige problemen. Ik vertrouw er op te kunnen rekenen op Uw steun, waar mijn kennis te kort schiet bij de oplossing van deze problemen.

*Hooggeleerde Hamaker,*

Onze eerste ontmoeting zal ongeveer vijftien jaar geleden hebben plaatsgevonden. Ik was als assistent van de Statistische Afdeling van het Mathematisch Centrum aanwezig bij een bespreking, waar U ons het probleem voorlegde van het experimenteren met mengsels. Dit was een van de eerste boeiende praktijkproblemen waarmee ik in aanraking kwam, maar ik vrees dat ik toen tot de discussie weinig heb kunnen bijdragen. Het verheugt mij zeer na zoveel tijd nog met U, één van de pioniers in Nederland op het gebied van de toegepaste statistiek, te mogen samenwerken. Dat U mij reeds bij een aantal van de vele wetenschappelijke onderzoekers, met wie U in contact staat hebt geïntroduceerd stel ik op hoge prijs.

*Hooggeleerde Benders,*

U bent degene die mij tot dusver het meest wegwijs hebt gemaakt binnen de onderafdeling der wiskunde; hiervoor dank ik U zeer. Uw vakgebied, de operations research, heeft met de statistiek vele bindingen. Ik ben verheugd nauw met U te kunnen samenwerken.

*Dames en heren medewerkers van de onderafdeling der wiskunde,*

De prettige wijze waarop ik het afgelopen jaar reeds met een aantal van U heb mogen samenwerken aan onze gemeenschappelijke onderwijstaak, heeft mij hier snel thuis doen voelen en geeft mij de beste verwachtingen voor de toekomst.

*Hooggeleerde Hemelrijk,*

Toen ik in het najaar van 1952 bij het Mathematisch Centrum kwam solliciteren had ik een eerste gesprek met U en met Uw leermeester, Prof. Van Dantzig. Ik prijs mij zeer gelukkig toen de kans te hebben gekregen onder Uw beider leiding te mogen werken. Ik ben er trots op nog een leerling van Van Dantzig te zijn, maar ik geloof dat U tot mijn vorming als statisticus het meest hebt bijgedragen. Mijn belangstelling ging immers vooral uit naar toepassingen en aan mijn eerste schreden in deze richting hebt U op zeer stimulerende wijze leiding gegeven.

Aan de tijd dat ik op het Mathematisch Centrum heb gewerkt, denk ik mede hierdoor met voldoening en dankbaarheid terug.

*Dames en heren medewerkers van Unilever N.V. te Rotterdam*

Toen ik ongeveer 12 jaar geleden onder leiding van de heer A. J. de Jong in Vlaardingen als statisticus in Uw bedrijf kwam werken, was er voor mij zeer veel te leren. Voor de zeer vele stimulerende contacten en discussies in binnen- en buitenland kan ik niet dankbaar genoeg zijn. Zij hebben mijn inzicht in de mogelijkheden tot toepassingen van de wiskunde bijzonder verruimd. Dat ik tot de realisering van deze mogelijkheden heb mogen bijdragen, stemt mij tot voldoening. De wijze waarop U, *hooggeachte De Boer*, gedurende de laatste jaren aan mijn werk leiding hebt gegeven, heb ik bijzonder gewaardeerd. Dat de banden met Uw bedrijf niet geheel zijn verbroken, stel ik op hoge prijs.

*Dames en heren studenten,*

U zult allen in Uw toekomstig werk herhaaldelijk waarnemingsresultaten moeten analyseren. U komt daarbij vaak zeer ver alleen gewapend met gezond verstand, maar met enige kennis van statistische methoden zult U beslist verder kunnen komen.

Statistiek is, vooral bij een eerste kennismaking, niet het gemakkelijkste vak om te leren; het is evenmin het gemakkelijkste vak om te onderwijzen. Voor velen van U is de kennismaking tijdens Uw studie bovendien van korte duur, te kort om veel technieken te kunnen leren, maar naar ik hoop voldoende lang om U met de statistische gedachtengang vertrouwd te maken. Voor de wiskundestudenten onder U zijn er nog vele keuzemogelijkheden. Graag zal ik U behulpzaam zijn bij het kiezen van een afstudeerrichting. Een enkele onder U heeft nu reeds de mathematische statistiek gekozen. Ik kan diegenen, wier belangstelling in deze richting ligt, verzekeren naar vermogen te zullen bijstaan.

Ik heb gezegd.