

## 8.3 De profeet, de missionaris en de handelsreiziger

R.D. Gill en W. Mettrop

Profeet, missionaris en handelsreiziger, zo omschrijft Jan Hemelrijk de rol die respectievelijk Van Dantzig, Sittig en hijzelf speelden in de naoorlogse jaren van wederopbouw en vernieuwing bij de introductie van de statistiek en de statistische consultatie op het MC en verder in Nederland. Ons gesprek is er in de eerste plaats op gericht de drijfveren van deze mensen te achterhalen, en een beeld te krijgen van de context waarbinnen hun activiteiten zich afspeelden.

Hoe kwam het dat Van Dantzig, althans bij velen, als een profeet werd gezien?

“Voor de oorlog had je helemaal geen statistiek in Nederland. Je had wel Hamaker bij Philips, maar die was eigenlijk een scheikundige. In Leiden had je Nass, een bioloog, en in Wageningen aan de Landbouwhogeschool Prof. Van Uven. Het was allemaal normale verdelingen, regressie-analyse en variantie-analyse. Je moet niet vergeten, er waren ook helemaal geen boeken in die tijd, althans geen goede. *Kendall* kwam pas in 1946 uit, en *deel 2* pas twee jaar later.<sup>1</sup> Ik heb het zelf voor *Statistica* gerecenseerd.<sup>2</sup> Onmisbaar om je weg in de literatuur te vinden, maar vol fouten - overgeschreven uit die literatuur - en drukfouten. Ik heb een hele lijst aan Kendall opgestuurd en kreeg een vriendelijk briefje terug. Ja, je had alleen dat boek van Fisher, *Design of Experiments*.<sup>3</sup> Als je een praktisch persoon was, was dat allemaal onbegrijpelijk gebrabbel. Het enige wat je er voor de praktijk uit kon leren, was hoe je een 2x2 tabel kon analyseren. Hij was wel geniaal hoor, maar...”;

“er werd wat afgeknoeid in die tijd. Vreselijk was het. Allemaal receptjes overnemen uit slechte boeken. Dan kwam er één ster uit, soms twee sterren of drie. Ik zal je eens een voorbeeld vertellen...”.

1. [Kendall 1948]

2. Boekbespreking door Hemelrijk in *Statistica* 4 (1950) p.226.

3. [Fisher 1935].

Door een onverwachte wending in het gesprek komen we er niet achter wat dan wel het slechtste voorbeeld aller tijden van statistische analyse geweest is, misschien is dat maar beter ook. Voordat we het gesprek verder weergeven en daadwerkelijk over Van Dantzig beginnen is het wellicht nuttig uit te leggen hoe het kwam dat de statistiek pas zo laat in Nederland tot ontwikkeling kwam.

De statistiek is altijd een buitenbeentje geweest onder de verschillende onderdelen van de toegepaste wiskunde. Ze is, internationaal gezien, überhaupt pas heel laat ontstaan en is in de eerste plaats niet door wiskundigen zelf ontwikkeld, maar door demografen, economen en actuarissen, wier vakken zelf pas in de achttiende en negentiende eeuw ontstaan zijn, samen met de opkomst van grote kapitalistische en centralistische staten. We hebben het dan alleen over het prille begin van de statistiek. De volgende grote ontwikkeling begon pas tegen het einde van de negentiende eeuw. Terwijl de wiskunde zelf een periode van formalisering, abstractie en consolidatie doormaakte, waren het nu medische en biologische onderzoekers die de statistiek voor hun eigen doeleinden begonnen te ontwikkelen, helemaal buiten de 'officiële' wiskunde om. Vooral in Engeland, waar men de continentale formalistische tendenzen lange tijd heeft tegengehouden, is de theorie van toepasbare, maar betrekkelijk ongeformaliseerde statistiek enorm tot bloei gekomen. Dit gebeurde met name door toedoen van mensen als K. Pearson en, later, R.A. Fisher. Zoals bekend kan de statistiek gezien worden als een tak van de waarschijnlijkheidsrekening. Maar zelfs de waarschijnlijkheidsrekening, behalve de meest elementaire aspecten daarvan, is pas door het werk van A.N. Kolmogorov in de jaren dertig geaccepteerd en geïntegreerd gedeelte van de moderne wiskunde geworden - tot dan toe was zelfs het begrip 'random variable' geen mathematisch begrip! Wat de statistiek betreft heeft het nog twintig jaar geduurd voordat door het werk van J. Neyman, E. Pearson, zoon van de eerder genoemde Karl Pearson, A. Wald, en anderen, een echt *mathematische* statistiek is ontstaan; een ontwikkeling die het eerst in de Verenigde Staten gestalte heeft gekregen.

Intussen had in de jaren twintig en dertig de Engelsman R.A. Fisher zo ongeveer alle onderdelen van de moderne statistiek bedacht, maar dan veelal zonder wiskundige formalisering, of liever: zonder wiskundige *precisie*. Fisher was een geneticus van huis uit, ook een bekwaam wiskundige, maar dat was voor hem niet de hoofdzaak. Een groot gedeelte van zijn loopbaan bracht hij door aan het landbouwkundig onderzoeksinstituut Rothamstead, waar hij methoden ontwikkelde voor het ontwerpen en analyseren van veldproeven ter verbetering van teeltrassen en teeltmethoden. Volgelingen van de 'Wageningse School' in Nederland, dat wil zeggen leerlingen van Van Uven, zijn in de jaren dertig nog naar Rothamstead afgereisd om zich deze nieuwe statistiek eigen te maken. Hier werd in hoofdzaak statistiek gebruikt die was gebaseerd op de veronderstelling van een normaal verdeelde toevallige variatie in bijvoorbeeld de opbrengst van een proefveldje, in combinatie met de veronderstelling van een lineaire afhankelijkheid tussen de gemiddelde opbrengst en de factoren die de opbrengst beïnvloeden (teeltras, grondeigenschappen etc.). Met behulp van slim bedachte proefopzetten lukte het in een beperkte tijd en op een beperkte ruimte een aantal nieuwe variaties, kunstmesten en dergelijke, uit te proberen in een voornamelijk exploratief onderzoek om veelbelovende variaties en teeltwijzen op te

sporen. Bovendien slaagde men erin om een aantal niet te beheersen variabelen (bijvoorbeeld onbekende fertiliteitsgradiënten over een proefveld) onder zekere randvoorwaarden uit te schakelen, door willekeurige toewijzing van de verschillende behandelingen in een proef aan de verschillende proefveldjes. De theorie die hierbij te pas komt, maakt nog steeds een byzantijnse indruk met termen als grieks-latijnse vierkanten en 'confounded split-plot designs'. Temeer gold deze indruk destijds, toen de wiskundige achtergronden nog niet duidelijk begrepen waren. Hemelrijk vertelt hierover:

"Sommen van kwadraten werden op een geheimzinnige manier afgesplitst en op elkaar gedeeld, en daar kwam weer een F-verdeling uit. Pas later, eerst door een boekje van H.B. Mann en later door boeken als *Kendall deel 3*, leerden we dat het 'gewoon' een lineair model was, gewoon de kleinste kwadraten methode dus. Van Zwet heeft ooit eens een programma geschreven om een algemene variantie-analyse uit te voeren, ook voor niet uitgebalanceerde modellen en met ontbrekende waarnemingen. Maar dat lukte niet helemaal; de computers waren toen nog niet goed genoeg."

Fisher had zich niet beperkt tot deze bijzondere statistische methoden, maar in het algemeen de basis gelegd voor een groot deel van de belangrijke ideeën uit de moderne statistiek. Hij presenteerde zijn geniale vondsten echter op een sterk intuïtieve, individualistische en vaak dogmatische manier. Voor een kritisch denkend en modern gevormd wiskundige was zijn boek bepaald niet ideaal om er statistiek uit te leren. Bovendien was zeker niet alles wat hij uitvond van blijvende waarde. Zo is er zijn theorie van fiduciële inferentie.

"Daar heeft de Engelse school twintig jaar mee geknoeid. De theorie van fiduciële inferentie berustte gewoon op het door elkaar halen van wat een stochastische variabele is en wat niet."

Nu nog worden er incidentele pogingen ondernomen om deze gedachte van de meester, de enige die nog steeds niet in de mathematische statistiek geïncorporeerd is, te redden, hoewel de praktische betekenis ervan gering is.

De Profeet.

Het gesprek gaat verder over Van Dantzig. Hoe groot was de rol van Van Dantzig in de ontwikkeling van de statistiek in Nederland en hoe nauw is Hemelrijk daarbij betrokken geweest? Was men voldoende op de hoogte met ontwikkelingen in het buitenland?

"Van Dantzig had in de oorlog de statistiek ontdekt, terwijl hij ondergedoken was. Er waren toen helemaal geen boeken - alleen wat hopeloos verouderde Duitse tekstboeken, van Von Mises en anderen -, dus heeft hij maar alles zelf uitgevonden. Dat hebben we allemaal in het begin gedaan. Van de Vaart en ik [de twee medewerkers van Van Dantzigs afdeling statistiek in het nieuwe MC - *R.D.G.*] hebben bijvoorbeeld de principes van de Neyman-Pearson theorie 'ontdekt'. We waren er zelf op gekomen, dat je niet alleen de getoetste hypothese erbij moest betrekken, maar ook de alternatieven waartegen getoetst werd, om de 'optimale' toets te kunnen bepalen. Toen dat gedaan was, zei Van Dantzig, dat het wel heel mooi was, maar: waren we er wel zeker van dat het nog niet bestond? Toen ontdekten we helaas dat het er

inderdaad al was. We hebben er een pedagogisch artikel voor *Statistica* van gemaakt".<sup>4</sup>

"Toen hebben we de niet-parametrische statistiek ontdekt. Van de Vaart had het artikel van Mann en Whitney in de *Annals of Mathematical Statistics* over hun U-statistiek ontdekt. Dat was het! Precies wat we nodig hadden. Meteen gingen we het zelf toepassen bij de consultaties, er zelf varianten op bedenken, enzovoorts. We kwamen er pas later achter dat Wilcoxon het al eerder bedacht had, maar die publiceerde in zulke obscure tijdschriften".

Geïnspireerd door deze ideeën en genoodzaakt door een bepaalde consultatie, heeft Hemelrijk bijvoorbeeld voor zijn promotieonderzoek een niet-parametrische symmetrietoets ontworpen, die achteraf erg veel op die van Wilcoxon bleek te lijken.

Het speciale enthousiasme voor deze nieuwe, niet-Fisherianse aanpak van de statistiek kwam met name voort uit de kritische inslag van Van Dantzig en zijn volgelingen, en uit hun roeping om de wiskunde, i.c. de statistiek, zo korrekt mogelijk toe te passen. In de niet-parametrische statistiek vermijdt men zoveel mogelijk vooronderstellingen van parametrische aard betreffende de kansverdelingen van de waarnemingen, bijvoorbeeld de veronderstelling dat ze een normale verdeling hebben.

"Dat was vaak een totaal ongefundeerde veronderstelling. Bovendien kon je het helemaal niet nagaan. Goede toetsen voor normaliteit waren er nog niet".

Door methoden te gebruiken waarin bijvoorbeeld alleen gebruik wordt gemaakt van de volgorde van grootte van de waarnemingen, kan men belangrijke hypothesen toetsen terwijl men alleen de veronderstelling van onafhankelijkheid van de waarnemingen hoeft te maken. De methoden zijn in eerste instantie vrij eenvoudig en helemaal exact. Later werd ook nog ontdekt dat ze niet eens inefficiënter hoeven te zijn dan methoden gebaseerd op een vooronderstelde parametrische verdeling van de waarnemingen, die bovendien volkomen onbetrouwbaar zijn, als die veronderstelling niet deugt. Een nadeel is dat alleen betrekkelijk eenvoudige statistische problemen met niet-parametrische methoden kunnen worden opgelost.

"Het is opgehouden doordat er geen niet-parametrische variantie-analyse te doen is".

Van Dantzig en zijn medewerkers hebben nadrukkelijk gewezen op het begrip *mathematisch model*.<sup>5</sup> Een dergelijk model is meer dan alleen een beschrijving van de fysische wereld: het wordt mede gevormd door de doelstellingen van de statisticus. Zijn handelings- of beslissingsregels liggen vervolgens besloten in dit model. Natuurlijk zijn er willekeurig veel modellen mogelijk van eenzelfde situatie. Hemelrijk en Van Dantzig legden zeer de nadruk op deze keuzemogelijkheid; het doel van het onderzoek bepaalt of een model meer of minder adequaat is. Verschillende doelstellingen zorgen voor verschillende *scholen* in de statistiek. Voor mij [R.D.G.] ligt hier een dubbele tegenspraak. Vanwaar die felle bestrijding van andere scholen door Hemelrijk en Van Dantzig, in het bijzonder van de Bayesiaanse en fiduciële;<sup>6</sup>

4. [Hemelrijk/Vaart 1950]

5. [Dantzig 1947; 1953; 1954ab] [Hemelrijk 1954; 1958; 1967; 1978].

6. Het ene model is het andere niet. In Bayeriaanse modellen is het subjectieve karakter van de keuze van de a priori verdelingen gevaarlijk en doet schade aan de wetenschappelijkheid. De fiduciële theorie werkt met niet interpreteerbare begrippen. [J.H]

vanwaar, als ze toch het bestaan van meerdere modellen erkennen? En bovendien: waarom die nadruk op het onderscheid tussen model en werkelijkheid,<sup>7</sup> terwijl men telkens op zoek was naar modellen die zo 'perfekt' waren, dat ze niet van de werkelijkheid te onderscheiden waren? Van Dantzig en Hemelrijk bedreven immers een zodanige statistiek, dat dat onderscheid niet eens gemaakt kon worden; veronderstellingen die werkelijk beperkingen aan het model opleggen, zoals die aangaande lineariteit en normale verdelingen, werden juist vermeden! Sterker nog: ze wilden met behulp van hun modellen dingen uit de werkelijkheid *bewijzen*,<sup>8</sup> hetgeen een overschrijding betekent van de grens tussen model en werkelijkheid. Hemelrijk over deze laatste tegenstrijdigheid:

"Maar dat was het juist, dat vonden we zo mooi als model en werkelijkheid dicht bij elkaar stonden. Dan waren de conclusies onaanvechtbaar, betrouwbaar. Daarom heb ik ook nooit zoveel gehouden van modellen met meer dan twee of drie factoren. Dan vind je altijd wat, en je moet zoveel veronderstellen. Als *screening* is het goed, maar daarna moet je een nieuw onderzoek doen, helemaal gericht op die één of twee potentieel belangrijke en interessante effecten die je gevonden hebt".

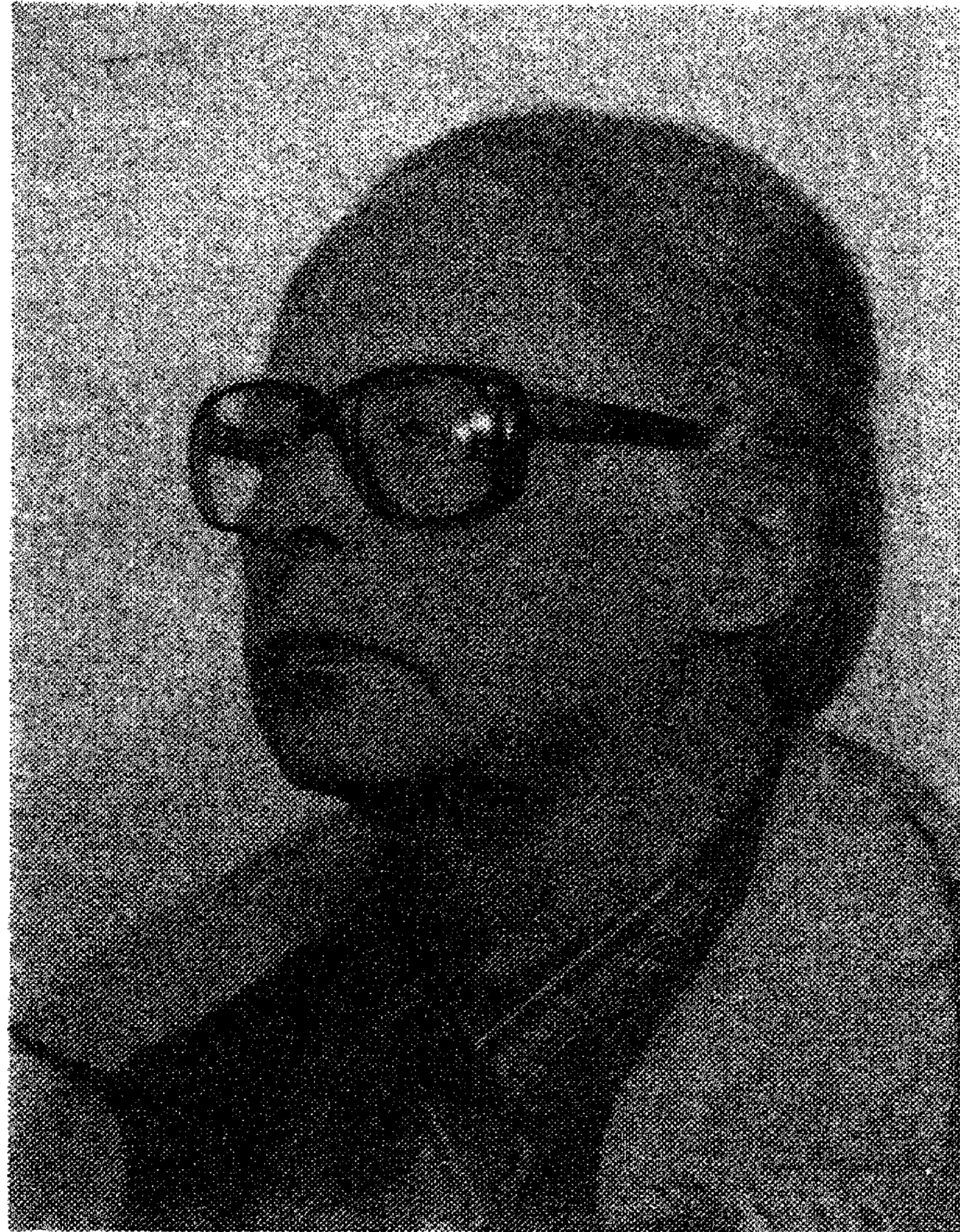
"ja, dat idee van een wiskundig model, dat is vreselijk belangrijk. Dat idee had Von Mises nog niet.<sup>9</sup> Toch moet het wel heel ver terug gaan. Wie zou het woord het eerst gebruikt hebben? Dat zal wel gauw tweehonderd jaar geleden zijn".

Op deze abstracte zaken komen we later terug als we de praktische aspecten van de statistische consultatie bekijken.

7. Als men wiskundig model en werkelijkheid niet onderscheidt komt men in moeilijkheden, getuige bijv. de klassieke griekse paradoxen. Het model moet wel zo realistisch mogelijk zijn, maar het blijft een vereenvoudigde weergave of zo men wil afspiegeling. [J.H]

8. Dit is een misvatting. Met een model, ook met de hele logica en wiskunde, kan men nooit iets over de werkelijkheid bewijzen. Daarvoor zijn waarnemingen nodig en wiskundige modellen kunnen alleen een kritisch en vaak machtig hulpmiddel zijn om uit die waarnemingen tot verantwoorde conclusies te komen. [J.H]

9. De Wiener Kreis-leden Von Mises en Reichenbach waren de belangrijkste twee van de eerder bedoelde Duitse auteurs. Van Dantzig refereert in [Dantzig 1941] onder meer aan [Mises 1931; 1939] en [Reichenbach 1935].



*J. Hemelrijk*

De handelsreiziger.

Het ging Van Dantzig en Hemelrijk er dus om de statistiek te gebruiken, de wereld ermee te beïnvloeden. Op het MC werd dit ideaal verwezenlijkt door intensieve aandacht voor de statistische consultatie. Voor Hemelrijk was dit praktische aspect van het werk ongetwijfeld zeer belangrijk. Gold dat volgens hem ook voor Van Dantzig?

“Minder, hij was door en door theoreticus. Daarom vulden wij elkaar zo goed aan. Ik was een heel praktisch persoon”.

Van Dantzig, de theoreticus, had wel een leidraad opgesteld voor de statistische consultatie, die voor zijn medewerkers goud waard was.

“Dat waren de tien zogenaamde gulden regels van de consultatie. De eerste was het belangrijkste: Je moest er achter zien te komen wat het eigenlijke doel van het onderzoek was. Vaak wist men dat zelf niet! Onze klanten werden dikwijls door hun professor gestuurd met een of andere kant en klare, ingewikkelde proefopzet die vaak zelfs al uitgevoerd was ook! Dan ontdekte je soms dat de hele boel beter de prullemand in kon, omdat ze iets over het hoofd gezien hadden, en moest je ze met lege handen terugsturen. Ja, het kerkhof van de medische onderzoeken werden we genoemd, hoorde ik later van Rümke.<sup>10</sup> Dan moest je er nog achter zien te komen hoe de metingen precies gedaan werden - meestal gingen we ook ter plaatse kijken, dan kwam je er pas goed achter hoe het allemaal in elkaar zat. En ook hoe groot de nauwkeurigheid van de metingen was. En of er duplo's verricht waren,

10. Dr.Chr. L. Rümke werd in 1961 lector farmacologie en medische statistiek aan de Vrije Universiteit, in 1976 gewoon hoogleraar. Al in de jaren vijftig bouwde hij een hechte samenwerking op met de Afdeling Mathematische Statistiek van het MC. Zie bijvoorbeeld [Rümke 1958].

waarmee je dat kon bepalen; en zo niet, of dat alsnog kon. En wie de analisten waren! Dat is zeker ook heel belangrijk. Ik herinner me een mooi voorbeeld. Dat ging over een onderzoek naar het cholesterol-niveau in het bloed. In die tijd had nog niemand gehoord van cholesterol. De onderzoeker, een inmiddels zeer vooraanstaand medicus, kreeg het idee dat het cholesterol-niveau iets te maken had met het humeur of de gemoedstoestand. Hij had namelijk bij een echtpaar geconstateerd dat de cholesterol-niveaus van beide partners vrijwel parallel liepen, terwijl ze verschillende diëten volgden in een zeer goed gecontroleerde proef. Ik opperde de suggestie, dat dit kwam doordat de chemische analyse van de regelmatige bloedprikjes van de twee tegelijkertijd door dezelfde analist gemaakt werd. Inderdaad, het bleek dat de parallele curves te wijten waren aan schommelingen in de zogenaamde 'standaard' die elke keer opnieuw bepaald moest worden. De medicus heeft zijn humeuren-theorie moeten laten vallen".

Het feit dat de theoreticus Van Dantzig, een man zonder duidelijke aanleg voor of ervaring met de praktische statistiek, deze lijst van tien gulden regels opstelde - elementaire, maar daarom ook zo vaak vergeten, waarna zijn jonge medewerkers ze serieus in de praktijk toepasten, dit feit toont nog eens duidelijk zijn gaven. De aanpak is tekenend voor de gedrevenheid, de toewijding en het zelfvertrouwen van deze mensen, die de statistiek welswaar pas ontdekt hadden, maar die toch al overtuigd waren van haar belang voor de maatschappij. Waar kwam nu Hemelrijks gevoel voor de statistische consultatie vandaan?

"Ja, dat kreeg ik in de oorlog denk ik, in de illegaliteit. Toen moest je doortastend zijn. Beslissen wat er gedaan moest worden en het dan uitvoeren ook. Precies wat je in de consultatie nodig hebt. Tenslotte had ik er gewoon aanleg voor, een instinct".

Na de oorlog had Hemelrijk een afgebroken studie wiskunde te voltooien. Verder had hij een jong gezin, en een bijbaantje Kriterion, de net door oud-collegas uit de illegaliteit opgezette bioscoop.<sup>11</sup> Hij werd assistent bij Van Dantzig aan de Universiteit van Amsterdam, en toen hij afgestudeerd was - "negen tentamens in negen maanden" - medewerker, samen met Van de Vaart, bij Van Dantzigs nieuwe afdeling in het MC. Later werd hij hoofd Consultatie van de Statistische Afdeling.

Wat voor mensen kwamen daar om advies, werd er op de een of andere manier aan reclame gedaan? Werd er gericht gezocht in bepaalde toepassingsgebieden?

"Nee, aan reclame deden we helemaal niet. Wel overal cursussen geven en lezingen houden. En later kwamen er artikelen in medische tijdschriften. In het begin hadden we niet veel klanten, de mensen leken nogal schuw, maar toen er een paar door ons goed geholpen waren, kwamen ze meteen in grotere aantallen tegelijk, als vanzelf. Ook werden ze gestuurd door hun profs en later zelfs door de redacteurs van medische tijdschriften. Het waren voornamelijk medische onderzoekers, uit de industrie kwamen er heel weinig. We hielpen ze allemaal gratis. De maatschappij had het nodig, we waren verschrikkelijk idealistisch. Zelf wilde ik eerst mijn eigen naam niet op rapporten hebben, alleen de naam van het instituut. Van Dantzig niet hoor, die vond dat dat onzin was en daar had hij ook wel gelijk aan."

11. Zie [Dolfsma 1985].

Enthousiasme.

Op de afdeling moet een sfeer van enthousiasme en toewijding geheerst hebben.

“De consultatie, ja, dat vond ik het belangrijkste van wat op de afdeling gebeurde. Van Dantzig was een gedreven man en een stuwende kracht voor anderen. Zelf zat hij dag en nacht te werken en o.a. zijn theorie van collectieve kenmerken te ontwikkelen. Ook al bemoeide hij zich weinig met consultaties, toch luisterde hij kritisch naar de plannen van zijn medewerkers en las hun rapporten. Het kritisch-zijn was een van zijn sterkste punten. Altijd vroeg hij: 'Maar is dat wel zo?'. Hij eiste ook heel veel van je. Hij stimuleerde ook door je dreigend aan te kijken en te vragen: 'Hoe is het nu met je proefschrift?'.

“Als hij ergens een lezing moest houden, arriveerde hij soms drie kwartier van te voren om hem ter plekke nog te voltooien, toch waren het uitstekende lezingen. Alles wat hij nodig had, bedacht hij zelf. Hij las haast nooit wat, dat heb ik van hem overgenomen. Dat gold ook voor zijn colleges. Soms werkte hij eraan in een café, de avond te voren, tot na middernacht. Ik moest er dan bij blijven; hij had iemand nodig om tegen te praten. Ik zat tenslotte te knikbollen met 'Ja, professor; nee, professor...'. Hij ging ook wel 'ns door terwijl om hem heen de kelniers alles opruimden, de stoelen omgekeerd op de tafels gelegd werden,... maar daar trok hij zich niets van aan. Ik werd er maar zenuwachtig van. Hij sliep dus haast nooit, en als hij het wel deed sliep hij heel slecht. Bij lezingen van anderen dommelde hij altijd. Hij zat dan op de eerste rij. Eerst lette hij heel goed op totdat hij een vraag bedacht had, dan sliep hij in. Hij moest altijd een vraag hebben. Bij het applaus werd hij weer wakker om zijn vraag te stellen. Ondanks deze eigenaardigheden was hij beslist niet geïsoleerd. Hij correspondeerde veel en is ook een jaar in Amerika gebleven, dat was in 1948 of 1949. Wel is het zeker dat hij zich doodgewerkt heeft, hij is betrekkelijk jong gestorven. Hierdoor komt het dat maar een handjevol mensen bij hem zelf gepromoveerd is. Van Dantzigs arbeid heeft zeker succes gehad. Hij heeft eigenhandig de statistiek geïntroduceerd in Nederland als een volwaardig wiskundig vak en haar bijzonder gestimuleerd, onder meer dus door de consultatie aan het MC op te zetten. Hij heeft goede medewerkers aangetrokken en ze tot grote daden gestimuleerd. Zo heeft hij Runnenburg en Kesten, die theoretische natuurkunde wilden doen, maar bij hem voor een tentamen waren gekomen, omgepraat om tot de statistiek over te gaan. Die twee, die alles samen deden, werden ook wel Kestenburg genoemd. Van Zwet, die al net zo gedreven was als Van Dantzig zelf, studeerde in Leiden maar kwam colleges statistiek volgen bij Van Dantzig in Amsterdam. Nederlands onderzoek in de mathematische statistiek is vanaf dat begin blijvend van internationale betekenis geweest. Op de afdeling werd enthousiast en heel hard gewerkt. De sfeer trok ook zeer bekwame mensen aan. Er werd veel over het werk gepraat. Wel werden de medewerkers gewoon in het diepe gegooid, ze moesten zichzelf zien te redden, want er was geen tijd om ze intensief te begeleiden. Maar haast iedereen heeft het gered. Ook voor de wat minder begaafden was er genoeg nuttig werk te doen, tot het instandhouden van een kaartsysteem van de mathematisch statistische literatuur aan toe... Terpstra hield niet van dat kaartsysteem. Hij heeft een keer negen maanden aan iets gewerkt, en had iets moois opgelost. Hij zat met zijn rug naar de kaartenbakken. Toen het klaar was gingen we toch even kijken in het systeem - en ja hoor, precies achter hem zat het er al die tijd in. Maanden werk voor niets!”

Een maatstaf voor het succes van Van Dantzig is volgens Hemelrijk het feit dat op een gegeven ogenblik, op een enkele uitzondering na, alle hoogleraren statistiek in Nederland - en het waren er inmiddels best veel, en van niveau - eerder medewerker aan Van Dantzigs afdeling waren geweest. En dat, terwijl voor de oorlog de statistiek



als onderdeel van de wiskunde academisch gewoon niet bestond. Van Dantzig heeft Nederland overtuigd van het feit dat het een belangrijk vak was en dat er leerstoelen moesten komen. Ook een aantal hoogleraren in de medische statistiek en psychometrie heeft een tijd op het MC doorgebracht, evenals econometristen, onder wie Theil. Van Dantzig was een gedreven man, maar waardoor?

“In de eerste plaats door eerezucht, ook al moet je dat niet teveel benadrukken, of op een vriendelijker manier zeggen. En verder door maatschappelijk besef en nieuwsgierigheid. Hij was ontzettend nieuwsgierig”.

Was Hemelrijk zelf ook zo gedreven?

“Nou niet zo, hoewel, ik was *wel* gedreven voor de consultatie. Ik gunde mijzelf geen rust voordat ik een consultatie tot een goed einde had gebracht. Dat liet me niet los. Ik was een soort adjudant van Van Dantzig, en ook een handelsreiziger. We zijn overal naar toe geweest. Later heb ik me daarnaast helemaal ingezet voor het onderwijs. Mijn ideeën over de onvoorspelbaarheidskarakterisering van een aselector en tenslotte over het equivalentie principe voor wiskundige modellen, daar ben ik best tevreden over”.

Uit het soort klanten van de Consultatie-Afdeling en uit de statistische achtergrond van Van Dantzig en Hemelrijk is goed te begrijpen waarom ze vaak voor bepaalde statistische methoden kozen. In medisch onderzoek is het meestal heel belangrijk geen valse positieve resultaten abusievelijk toch aan te kondigen; en het is zeker zo belangrijk dat alle veronderstellingen van een statistische analyse degelijk onderbouwd zijn. Ook past de speciale aandacht voor de grondslagen van het vak binnen dit plaatje.

“Die grondslagen, die had je nodig om een brug te slaan naar de praktijk. Bij de zuivere wiskunde heb je geen grondslagen nodig, alleen wat axioma's. Maar als je wiskunde wilt toepassen, dan heb je een soort woordenboek nodig om begrippen uit de praktijk te vertalen tot begrippen binnen het model. Over iets dat in de praktijk voor komt, maar niet in het model, kun je niks zeggen. Dat is ook de speciale betekenis van het aselecten. Stel A,B en C zitten wel in het model en de rest nog niet; door aselecteren verdwijnt dan juist die rest en wordt volledig in het model opgenomen, opgelost als het ware.”

Hemelrijk en zijn collega's waren niet alleen bedreven in het onberispelijk toepassen van de statistiek, ze waren er ook heel goed in om hun opdrachtgevers te overtuigen van het belang van de uitkomsten. Een bijzonder voorbeeld hiervan is de halve meter extra die ze de politici en ingenieurs aangepraat hebben voor de hoogte van de dijken in het Deltaplan.

“Ze wilden het peil op 4,50 meter hebben. Maar wij hadden de extrapolatie naar 1 op 10.000 uitgevoerd en kwamen uit op 5,10 meter. Nou, die 10 centimeter deden er niet zoveel toe. We zaten in die commissie en vergaderden in de Treveszaal op het Binnenhof. We praatten en praatten. Er zaten allemaal oude rotten van politici en zo bij, die vonden dat ze het zelf wel zouden uitmaken hoe hoog de dijken zouden worden, ze konden hun oren niet geloven. Wat verbeeldde die wiskundigen zich toch! Er was niet veel statistiek bij, het was een ontzettend verre extrapolatie. Maar we waren blijkbaar zo overtuigend, dat het er toch doorkwam. Van Dantzig wou eigenlijk anderhalve meter erbij, dat vond hij wel zo veilig denk ik, maar dat was

onhaalbaar".<sup>12</sup>

Hier zien we ook het instinct voor wat haalbaar is en ook de overredingskracht van Hemelrijk duidelijk naar voren komen. Soms was het puur bluffen, zoals de keer dat Hemelrijk plotseling in een moeilijk lopende vergadering met zakenlieden verklaarde: "Maar dat is een wiskundige variabele!", waarop alle kritiek op zijn voorstel verstomde.

Vanwege de voorliefde voor exacte, begrijpelijke en volstrekt betrouwbare statistische analyses werd met veel vindingrijkheid gezocht naar eenvoudige methoden om in een ingewikkelde materie duidelijke resultaten te verkrijgen - gedeeltelijk lag dit ook aan de bescheiden rekenmogelijkheden van die tijd, toen computers meisjes waren en geen elektronica. Men deelde de data in deelgroepen in; bekeek het effect van factor C binnen de categoriën gevormd door de factoren A en B... Hemelrijk bedacht een methode om de resultaten van een aantal onafhankelijke toetsen van dezelfde hypothese samen te voegen: de gecombineerde toets kon best significant zijn, ook al waren de onderdelen dat niet.

"Achteraf bleek dit veel op een methode van Fisher te lijken. Ook legden wij veel nadruk op het onderscheiden van detectie en bewijs. Een voorbeeld: een medicus kwam met een bijzonder ingewikkelde proefopzet aanzetten waarmee hij zeven factoren tegelijkertijd wilde onderzoeken; hij beschikte over veertig ratten van een bepaald ras, die elk aan verschillende combinaties van de niveaus van de zeven factoren onderworpen zouden worden. Hij had hier net genoeg tijd voor. Gelukkig was het experiment nog niet gestart. Eerst overtuigden we hem ervan dat wat hij wilde doen onmogelijk was. Er waren zoveel mogelijke effecten, zoveel mogelijke interacties, dat met zijn proefopzet niets te *bewijzen* zou zijn. Toen hebben we hem gevraagd wat hij het liefst zou willen bewijzen, welk vermoeden hij het belangrijkste vond om aan te tonen. Na enig nadenken kon hij dat vaststellen. Toen hebben we hem aangeraden om met precies tien van die ratten - twee keer vijf - dat eerst te onderzoeken. Dat heeft hij gedaan en enige tijd later kwam hij terug met resultaten die inderdaad in de goede richting wezen. Wij berekenden dat, als het echte effect inderdaad net zo groot was, hij precies dertig ratten nodig had om dat aan te tonen! Dus wij adviseerden hem de resultaten van de 'pilot study' weg te gooien en de rest van zijn materiaal te besteden aan een poging om het gedetecteerde verschil te bewijzen. Op deze manier lukte het ook. Als hij meteen zijn veertig ratten had gebruikt om *alles* te onderzoeken, was dit niet gelukt. Zijn professor was verbaasd. Moest hij daarvoor nu het advies van een statisticus inwinnen?"

Elders in dit boek wordt verslag gedaan van een interview met een van Hemelrijks belangrijkste collaboranten, Sittig, een 'missionaris', zoals Van Dantzig een profeet was. Ook in dit gesprek met Hemelrijk komt deze samenwerking ter sprake. Zoals Hemelrijk de adjudant was van Van Dantzig op het gebied van de statistiek, zo was Sittig de adjudant van Van Ettinger op het gebied van de kwaliteitscontrole. Hemelrijk

12. Het dijkhoogte-onderzoek voor het Deltaplan heeft een aardig staartje gekregen. Een onlangs door het CWI uitgevoerd herhaalonderzoek, waarbij een dertig jaar langere waarnemingsreeks en veel ingewikkelder methoden beschikbaar waren, leverde toch een schatting op van een meter lager. Wel is een extra parameter in het model nodig geweest en is het resulterende betrouwbaarheidsinterval veel breder. Als men dus aan de veilige kant wil zitten is het resultaat hetzelfde.

werd wiskundig adviseur bij Sittigs adviesbureau, het AKB. Sittig was ook een gedreven persoon, met zeer ambitieuze plannen maar ook met het doorzettingsvermogen en de kracht om ze uit te (laten) voeren. Weer trad Hemelrijk op als de temperende pragmaticus. Het lijkt een heel plezierige samenwerking geweest te zijn. Gezien ook de geheel andere klantenkring van AKB en MC was er van concurrentie helemaal geen sprake.

Wanneer Hemelrijk Van Dantzig een profeet noemt kunnen we hem gelijk geven. Van Dantzig voorzag de tijd dat er voor wiskundigen andere loopbanen zouden zijn dan die van schoolmeester of - voor enkelen - directeur van een verzekeringsmaatschappij of - voor de zeer gelukkigen - hoogleraar. Hij heeft ingezien dat statistiek en besliskunde voor de maatschappij hard nodig waren, en heeft met enorme inzet bereikt dat anderen zijn visie deelden.

Noemt Hemelrijk zichzelf echter een 'adjutant van Van Dantzig', dan is hij te bescheiden. Natuurlijk was hij, als rechterhand van Van Dantzig, de eerste om door hem meegesleurd te worden. Maar hun samenwerking moet zeker eenzelfde stimulans voor Van Dantzig betekend hebben. Samen immers hebben zij ervoor gezorgd dat het wetenschappelijk onderzoek in de statistiek in Nederland op een zeer aanzienlijk peil is komen te staan, terwijl het met name Hemelrijk is geweest die tevens de praktische beoefening van de statistiek met een enorme toewijding heeft gepropageerd en geïntroduceerd.