

Parametrische en parameter vrije methoden en hun toepassingen *)

door J. Hemelrijk
Mathematisch Centrum

S u m m a r y

Parametric and non-parametric methods and their applications

The rapid development of non-parametric methods in statistics raises the question of the difference between these methods and the classical parametric methods and of their different fields of application. This paper, read at a meeting of the "Vereniging voor Statistiek" on March 6, 1952, attempts to throw some light on this problem by means of two examples. A short exposition of an application of the method of moments of K. Pearson to the estimation of a population is given for a case, where this method is quite adequate and where it would be difficult to find an equally adequate parameterfree method. On the other hand a method of determination of histamine in human blood is described, which is based on an application of the method of m rankings of M. Friedman and M. G. Kendall. In this case the transition from the classical methods to parameterfree methods constitutes a real gain in rigour and efficiency.

1. Inleiding

Bij de ontwikkeling van nieuwe methoden in de statistiek wordt tegenwoordig veel aandacht besteed aan parameter vrije methoden. Daarbij beschouwt men meestal niet de door een experiment verkregen waarnemingen zelf, maar men vervangt deze waarnemingen, alle tezamen of groepsgewijs, door hun rangnummers bij rangschikking naar grootte. De theorie wordt dan gebaseerd op onderstellingen en hypothesen omtrent de waarschijnlijkheidsverdelingen van deze rangnummers. De naam „parameter vrij” (ook wel „verdelingsvrij”) wordt gebruikt, omdat de onderstellingen, die voor de opbouw van de theorie gebruikt worden, gewoonlijk geen eisen stellen aan de vorm van de waarschijnlijkheidsverdelingen der oorspronkelijke waarnemingen, behalve dan (gewoonlijk), dat deze verdelingen continu zijn (en soms dat zij gelijk zijn) en dat de waarnemingen onafhankelijk zijn. Ook de naam „rang-invariante methoden” wordt gebruikt¹⁾. Dit berust op het feit, dat bepaalde klassen van vervormingen van de indeling van de schaal der oorspronkelijke grootheden, waarbij de volgorde van de schaaldelen niet verandert, geen invloed hebben op de waarden, die aangenomen worden door de in deze theorie gebruikte statistische grootheden.

*) Voordracht gehouden op de Statistische Dag, 6 Maart 1952.

¹⁾ Zie D. van Dantzig [1], p. 271 e.v.

Het doel van de ontwikkeling van deze theorie ligt voor de hand. Aangezien men minder onderstellingen omtrent de oorspronkelijke waarschijnlijkheidsverdelingen maakt dan bij de „klassieke” parametrische methoden, bezitten zij een ruimer veld van toepassing dan de laatstgenoemden. Daarbij komt nog, dat de gedachtegang, waarop deze nieuwe theorie berust vaak eenvoudig en aanschouwelijk is, terwijl de wiskundige uitwerking zeer fraai, zij het vaak moeilijk is. Ten slotte is er ook nog een winst voor de gebruikers; het rekenwerk dat voor de toepassing nodig is, is vaak relatief gering.

Tegenover deze voordelen staan natuurlijk ook nadelen. Vergelijkt men het complex der momenteel beschikbare parameter-vrije methoden met de overige statistische theorie, dan ziet men reeds spoedig, dat er nog slechts weinig van deze nieuwe methoden ontwikkeld zijn en dat zij nog lang niet ieder gebied van de statistiek bestrijken. Maar dit wordt steeds beter. Verder zijn zij, indien er wél voldoende over de waarschijnlijkheidsverdelingen der onderzochte grootheden bekend is om er de klassieke methoden op toe te passen, gewoonlijk minder doeltreffend, daar zij van minder gegevens gebruik maken. Hoewel het verlies aan doeltreffendheid¹⁾ soms vrij gering is²⁾, kunnen de parameter-vrije methoden toch het best gezien worden als een aanvulling op de klassieke parametrische theorie, en wel voor die problemen, waarbij er twijfel aan bestaat of de onderstellingen, nodig voor toepassing van de klassieke theorie, vervuld zijn of waarbij bekend is dat zulks niet het geval is. Tegenover het — in dat geval trouwens denkbeeldige! — verlies aan doeltreffendheid, staat dan de grotere zekerheid van de conclusies.

Ter bevordering van een goed begrip van de draagwijdte en toepasbaarheid der beide theorieën, zullen wij van beide soorten methoden een typisch voorbeeld van toepassing bespreken. Als eerste voorbeeld hebben wij een toepassing van de momenten-methode gekozen, die zo goed aan het doel van het onderzoek, waarbij zij gebruikt is, beantwoordt, dat er geen reden is een methode met minder onderstellingen te gebruiken. Het tweede voorbeeld heeft betrekking op een fysiologisch experiment, waarbij moeilijkheden in de bepalingstechniek de toepassing van klassieke methoden zo goed als onmogelijk maken. Hierbij blijkt een kleine wijziging in de bepalingmethode, die er op gericht is, het waarnemingsmateriaal toegankelijk te maken voor toepassing van de methode der m rangschikkingen en van andere parameter-vrije methoden, een wezenlijke verbetering van het experiment als geheel te zijn.

¹⁾ Engels: "efficiency".

²⁾ Zie b.v. H. R. van der Vaart [2].

2. Schatting van een frequentieverdeling met behulp van de momentenmethode

In het boek „De Juiste Maat” van J. Sittig en Prof. Dr H. Freundenthal [3] wordt een onderzoek beschreven, dat verricht is met het doel een maatstelsel te ontwerpen voor het fabriceren van confectiekleding voor de in het Westen des lands winkelende Nederlandse vrouw ¹⁾. Hiertoe werden metingen verricht aan 5000 winkelende vrouwen. De vele problemen, die zich bij dit onderzoek voordeden, vindt men in het genoemde boek uitvoerig besproken en wij zullen er hier dan ook slechts één noemen. Bovendien is het probleem, dat wij op het oog hebben, reeds op één der vroegere Statistische Dagen door één der auteurs (J. Sittig [4]) ter sprake gebracht, zodat wij er kort over kunnen zijn.

De waarnemingen, tijdens het onderzoek verricht, zijn onderhevig aan verschillende soorten fouten: administratieve, systematische en toevallige. De eliminatie van de administratieve en systematische fouten zullen wij niet bespreken; wij zullen daarentegen onderstellen, dat deze geëlimineerd zijn en dat de metingen dus nog slechts toevallige fouten bevatten. Verder vatten wij één enkele der afmetingen in het oog en wel de rugbreedte. Hiervan zijn dus 5000 metingen verricht en het probleem is nu, om uit deze 5000 metingen een schatting te vinden van de frequentie-verdeling van de rugbreedte op de beschouwde collectie van de in het Westen des lands winkelende Nederlandse vrouwen. Bij het tot stand komen van deze schatting wenst men bovendien rekening te houden met de meetfouten, die dus zo goed mogelijk geëlimineerd dienen te worden. Daartoe zijn een aantal in meervoud verrichte metingen beschikbaar.

De gebruikte schattingsmethode berust nu op de volgende *onderstellingen*:

1e. De meetfouten zijn normaal verdeeld met gemiddelde 0 en met gelijke spreiding voor de verschillende meetsters, die de waarnemingen verrichtten.

2e. De gezochte frequentieverdeling kan voldoende nauwkeurig beschreven worden door een Pearson-verdeling.

3e. De meetfouten zijn onafhankelijk van de grootte der gemeten rugbreedte.

Ter toelichting: een Pearson-verdeling is een verdeling van een door K. Pearson [5] (zie ook [1], p. 136—166 en [6]) beschreven klasse van verdelingen. Binnen deze klasse wordt een verdeling gekarakteriseerd door zijn eerste 4 (gereduceerde) momenten: gemiddelde, spreiding, scheefheid en exces. De hier toegepaste methode is dus een typisch parametrische, gebaseerd op een klasse van verdelingen met 4 parameters. Zijn deze 4 momenten

¹⁾ Strikt genomen zou hier moeten staan: „de in het Westen des lands in de grootste steden winkelende vrouw”, daar de metingen in Amsterdam, 's-Gravenhage en Rotterdam zijn verricht.

bekend, dan wordt hierdoor precies één verdeling van de klasse aangewezen. Zijn van de 4 genoemde momenten schattingen bekend, dan wordt door deze 4 getallen een verdeling aangewezen (waarvan de eerste 4 momenten aan deze 4 getallen gelijk zijn), die als schatting van de gezochte verdeling kan gelden. Voorbeelden van Pearson-verdelingen zijn: de normale verdeling, de exponentiële verdeling, de Bêta- en Gammaverdelingen. De reden, dat niet ondersteld is, dat ook de rugbreedte zelf, evenals de meetfout, normaal verdeeld is, was, dat deze verdeling duidelijk scheef is (zie fig. 1).

Op grond van de bovenstaande onderstellingen kan men nu gemakkelijk tot een *schatting van de verdeling der rugbreedte* komen. Daartoe behoeft men slechts een schatting van de eerste 4 momenten te berekenen, waarbij met de meetfouten rekening gehouden is. De wijze, waarop dit gebeurt, is in [3] en [4] uitvoerig uiteengezet en wij volstaan daarom met een korte aanduiding voor de eerste twee momenten.

Noemen wij de rugbreedte \underline{x} ¹⁾ en de meetfout \underline{v} ²⁾ en de gevonden meetuitkomst \underline{y} , dan is dus

$$(1) \quad \underline{y} = \underline{x} + \underline{v}$$

en volgens de gemaakte onderstellingen zijn \underline{x} en \underline{v} onafhankelijk verdeeld. Derhalve geldt voor de mathematische verwachtingen („theoretische eerste momenten’’):

$$(2) \quad \mathcal{E}\underline{y} = \mathcal{E}\underline{x} + \mathcal{E}\underline{v}$$

of, daar volgens de eerste onderstelling $\mathcal{E}\underline{v} = 0$ is:

$$(2') \quad \mathcal{E}\underline{y} = \mathcal{E}\underline{x}.$$

Dit betekent, dat het eerste moment door de meetfout niet systematisch wordt beïnvloed, zodat het gemiddelde van de 5000 gevonden waarde zonder correctie voor de meetfouten als schatting van het eerste moment $\mathcal{E}\underline{x}$ van de rugbreedte kan dienen.

Verder volgt uit (1) en de onderstelling van onafhankelijkheid van \underline{x} en \underline{v} voor de spreidingskwadraten (varianties):

$$(3) \quad \sigma^2(\underline{y}) = \sigma^2(\underline{x}) + \sigma^2(\underline{v}),$$

of

¹⁾ De onderstreping duidt op het stochastische karakter van deze grootheid, d.w.z. op het feit, dat x een grootheid voorstelt, die op de beschouwde collectie een frequentieverdeling bezit.

²⁾ Ook v is stochastisch: de meetfout bezit een waarschijnlijkheidsverdeling.

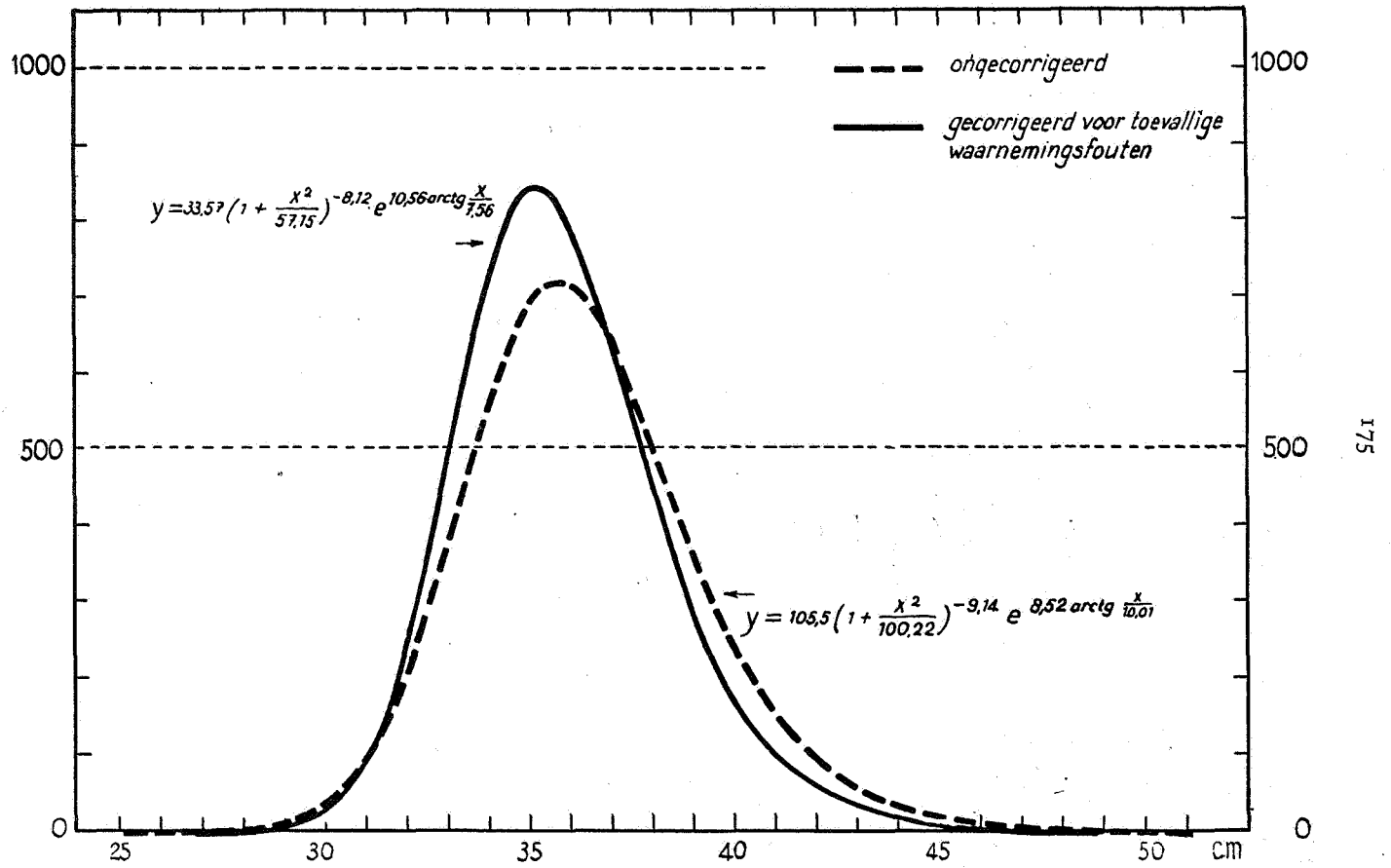


Fig. 1. Schatting van de frequentieverdeling der rugbreedte met en zonder correctie voor meetfouten.

$$(3') \quad \sigma^2(\underline{x}) = \sigma^2(\underline{y}) - \sigma^2(\underline{v}).$$

Uit de 5000 waarnemingen kan men een schatting van $\sigma^2(\underline{y})$ berekenen en uit de meervoudig verrichte metingen een schatting voor $\sigma^2(\underline{v})$. Het verschil is een schatting voor $\sigma^2(\underline{x})$ en is dus een schatting van de tweede parameter. Op analoge wijze kan men schattingen verkrijgen van de overige twee parameters, de scheefheid en het exces van \underline{x} .

In figuur 1 (overgenomen uit [3], p. 151) ziet men de schatting van de gezochte frequentieverdeling der rugbreedte. De gestippelde kromme stelt de schatting voor, die verkregen zou zijn, indien men met de meetfouten geen rekening gehouden had, terwijl de getrokken lijn de gecorrigeerde schatting weergeeft. Men ziet, dat het verschil aanzienlijk is.

Indien wij ons nu afvragen, of in dit geval de toegepaste klassieke methode aan zijn doel beantwoordt en of de gebruikte veronderstellingen niet wat ver gaan, dan kunnen wij hierover het volgende opmerken. Enerzijds zijn de onderstellingen zeker niet exact vervuld (de verschillende meetsters zullen b.v. hoogstwaarschijnlijk een verschillende spreiding in hun meetfouten bezitten). Anderzijds zal de verdeling van de, door alle meetsters tezamen gemaakte, meetfouten wellicht desondanks niet sterk van een normale verdeling afwijken. Het doel van het onderzoek is bovendien het vaststellen van een maatsysteem, dat als richtsnoer mot dienen voor de fabrikant. Daar een dergelijke richtsnoer toch nooit precies gevolgd wordt, is een globale correctie van de meetfouten voldoende. En tenslotte het belangrijkste argument: er is nauwelijks een parameter vrije schattingsmethode aan te geven, die even goed aan het doel van dit onderzoek beantwoordt als de momentenmethode. Vooral voor de eliminatie van de meetfouten is de momentenmethode wel de aangewezen. Een mislukte poging om een parameter vrije methode voor deze momentenmethode in de plaats te stellen, vermeld op p. 385 e.v. van [3], moge als bevestiging van deze uitspraak gelden. Wij komen daarom tot de conclusie, dat de toegepaste parametrische methode voor dit geval geheel adequaat is.

3. *Parameter vrije analyse van experimenten omtrent het histaminegehalte van bloed*

3.1 *Een gangbare methode*

De situatie is precies omgekeerd bij het probleem, dat wij nu zullen bespreken. Het betreft hier de bepaling van het histaminegehalte in bloed en in het bijzonder de vergelijking van het gehalte van verschillende bloedsera ¹⁾.

¹⁾ Dit voorbeeld is ontleend aan een rapport van het M.C. [7], dat geschreven werd in opdracht van de Stichting voor Biochemisch-Klinisch Onderzoek.

Een gangbare methode ter bepaling van dit gehalte vindt men uitvoerig beschreven in [8], hoofdstuk III en [9], hoofdstuk V. Zeer in het kort kan deze methode als volgt geschetst worden. Nadat het bloedmonster van de proefpersoon is afgenomen wordt het aan een aantal chemische bewerkingen onderworpen (tezamen de „opwerking” genoemd), die ten doel hebben door verwijdering van andere stoffen uit het bloed e.d. het monster voor de uiteindelijke bepaling geschikt te maken. Deze bepaling zelf geschiedt dan met behulp van de contractie, die een stukje caviadarm (darm van een marmot), opgehangen in een oplossing van fysiologisch zout, ondergaat indien men aan deze oplossing histamine toevoegt. Na de contractie wordt de vloeistof, waarin de darm is opgehangen, weggenomen, waarna weer een oplossing van fysiologisch zout voor deze vloeistof in de plaats wordt gebracht. De darm ontspant zich dan en de proef kan worden herhaald. De contractie wordt met een schrijvende wijzer op een draaiende trommel geregistreerd. Eén opgewerkt bloedmonster stelt de onderzoeker in staat viermaal een dergelijke contractie te meten. Tussen deze vier contracties door worden dan histamineoplossingen van verschillende bekende gehalten gebruikt, zodat men afwisselend contracties behorend bij bekende en bij het onbekende gehalte registreert. Na iedere contractie wordt de draaiende trommel stilgezet; deze momenten zijn kenbaar aan verticale strepen op de trommel, die boven de

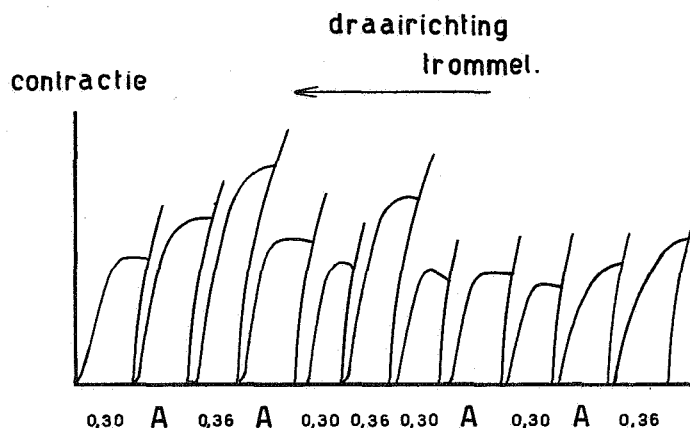


Fig. 2. Contractiecurve bij een histaminebepaling. A = onbekende oplossing, de getallen geven de bekende gehalten aan.

contractiecurve uitsteken. In figuur 2 is een dergelijke contractiecurve geschetst.

Uit deze contractiecurve wordt het histaminegehalte van het opgewerkte

serum nu visueel geschat. In figuur 2 zou men b.v. vermoedelijk de schatting 0,32 nemen.

Deze methode bezit, zoals men uit bovenstaande korte beschrijving wel reeds vermoeden kan, enkele aspecten, die de statistische verwerking van de verkregen waarnemingsuitkomsten ernstig belemmeren.

1e. De bepaling bevat twee subjectieve elementen; de keuze der bekende gehalten en de volgorde der contracties zijn niet aan voorschriften onderworpen en ook de uiteindelijke bepaling van de schatting op grond van de verkregen contractiecurve is tot op zekere hoogte subjectief. Het is bijna altijd erg moeilijk, de invloed van dergelijke subjectieve elementen op de betrouwbaarheid der eindconclusies na te gaan.

2e. Indien een contractiecurve al te onregelmatig verloopt, zodat het redelijkerwijze ondoenlijk is er een schatting voor het gezochte histaminegehalte uit af te leiden, wordt de bepaling als mislukt beschouwd. Dit komt nog al eens voor en dergelijke „gaten” in een waarnemingsreeks kunnen in bepaalde gevallen bij de statistische analyse moeilijkheden veroorzaken.

3e. Het ernstigste bezwaar echter is, dat de gevoeligheid van een stukje caviadarm voor histamine aan vrij snelle en vaak sterke wisselingen onderhevig is. Deze gevoeligheid loopt ook voor verschillende stukjes caviadarm en voor stukjes van verschillende caviadarmen sterk uiteen. Dit heeft tengevolge, dat de spreiding van de waarnemingsfout zelfs niet als bij benadering constant kan worden beschouwd. Hierdoor wordt, zoals bekend, toepassing van een geheel complex van parametrische methoden, zoals de variantie- en regressieanalyse onverantwoord. Dit heeft tengevolge, dat bepaalde vragen ternauwernood meer met behulp van een statistische analyse kunnen worden beantwoord. Wij zullen dit in de volgende paragraaf aan een voorbeeld toelichten.

3.2 Een speciaal probleem

Eén der vragen, die zich bij het onderzoek van het histaminegehalte van bloed voordoen, is, of dit gehalte bij één proefpersoon op korte termijn aan wisselingen onderhevig is, waarbij wij in dit geval wisselingen van onregelmatig karakter bedoelen, niet speciaal stijgingen of dalingen met de tijd. Neemt men nu b.v. 10 bloedmonsters snel na elkaar af en schat men de gehalten op de boven beschreven wijze, dan verkrijgt men 10 getallen als schattingen. Indien men nu kon beschikken over een schatting van de nauwkeurigheid van deze bepalingen en indien men mocht aannemen, dat de waarnemingsfouten onderling onafhankelijk en normaal verdeeld waren en alle dezelfde spreiding bezaten, zou men b.v. met de door R. A. Fisher gegeven z -toets [10] de hypothese kunnen toetsen, dat de 10 monsters alle

hetzelfde histaminegehalte bezitten. Dit is nu echter niet mogelijk en uit de 10 schattingen verkrijgt men dus eigenlijk in het geheel geen gegevens over het al of niet gelijk zijn van de gehalten der monsters.

3.3 Een gewijzigde methode

Wij zullen nu een kleine methodologische wijziging beschrijven, die erop gericht is het waarnemingsmateriaal toegankelijk te maken voor parameter-vrije methoden zoals b.v. de methode der m rangschikkingen van M. Friedman [11] (men vindt deze methode ook beschreven in [12], [13] en [14]). Hoewel het chemische en biologische deel van de bepaling en het gehele instrumentarium onveranderd worden gehandhaafd en de ingevoerde wijziging dus slechts een ondergeschikte plaats in de zeer bewerkelijke bepalingmethode inneemt, stelt zij ons toch in staat het in 3.2 genoemde probleem en andere daarop gelijkende te onderzoeken zonder van niet vervulde onderstellingen gebruik te maken. Wij zullen de gewijzigde methode aan de hand van dit probleem toelichten. Het is dan niet moeilijk in te zien, dat dezelfde wijziging het eveneens mogelijk maakt andere problemen van soortgelijke aard, zoals het opsporen van een niveau-verloop (Engels: "trend") of van een verschil van niveau bij twee of meer groepen proefpersonen, op soortgelijke wijze op te lossen.

In het in 3.2 genoemde geval beschikken wij over 10 bloedmonsters. Deze worden op de gebruikelijke in [8] en [9] beschreven wijze opgewekt, zodat wij 10 oplossingen van onbekend histaminegehalte verkrijgen, die ieder voldoende vloeistof bevatten om 4 maal een contractie te verkrijgen. De hypothese, die wij wensen te toetsen is, dat deze oplossingen, na de opwerking, alle hetzelfde histaminegehalte bezitten¹⁾. Wij geven de oplossingen de nummers 1, 2, . . . , 10 en kiezen vervolgens uit alle 10! permutaties van deze rij nummers achtereenvolgens 4 permutaties door loting uit, waarbij wij er zorg voor dragen, dat bij iedere loting alle permutaties gelijke kans bezitten gekozen te worden. Dit kan b.v. gemakkelijk worden uitgevoerd met behulp van een aselechte getallenrij (tabel van "random numbers"²⁾), zie b.v. [15] en vgl. ook [16]. Laat b.v. de eerste permutatie zijn:

9 2 5 7 4 10 1 3 8 6.

¹⁾ De eventuele „opwerkingsfout” wordt hier dus bij het eventueel aanwezige verschil der histaminegehalten opgeteld. Deze opwerkingsfout kan echter ook apart worden onderzocht, door een groot bloedmonster in delen te verdelen en deze apart op te werken, waarna de hier beschreven methode kan worden toegepast. Op die wijze kan nagegaan worden, of de opwerkingsfout alleen reeds een merkbaar verschil in histaminegehalten veroorzaakt.

²⁾ De term „aselechte getallenrij” is afkomstig van Prof. Dr D: van Dantzig.

We kiezen nu een oplossing met bekend histaminegehalte, en wel zo mogelijk met een gehalte, dat niet veel verschilt van het gemiddelde gehalte van de onbekende oplossingen. Dit kan, zo nodig, geschieden met behulp van een voorafgaande ruwe bepaling van het histaminegehalte van een extra bloedmonster.

Vervolgens registreren wij op dezelfde wijze als bij de gangbare methode een contractiecurve, waarbij nu echter de histamine-oplossingen in de hieronder aangegeven volgorde worden gebruikt. Eerst worden enkele (b.v. 3) contracties met de bekende oplossing (*B*) geregistreerd, vervolgens één contractie van onbekende oplossing no. 9, daarna weer *B* (éénmaal), onbekende no. 2, *B*, onbekende no. 5, etc., tot de hele rij onbekende oplossingen éénmaal in de door de eerste permutatie aangegeven volgorde aan de beurt is geweest. De reeks contracties wordt besloten met een contractie met *B*. Het is in het algemeen mogelijk een dergelijke curve te verkrijgen met behulp van één stukje caviadarm. Daar iedere oplossing van onbekend gehalte voldoende is voor 4 contracties, kunnen wij dit proces 4 maal uitvoeren, desgewenst met 4 verschillende bekende oplossingen (voor iedere reeks één) en zo nodig met voor verschillende reeksen verschillende caviadarmen en wij gebruiken daarbij de door de 4 permutaties gegeven volgorde.

Voor toepassing van de methode der *m* rangschikkingen behoeven wij slechts te beschikken over een aantal rijen rangnummers, die ieder een schatting van de volgorde naar grootte van de onbekende histaminegehalten vertegenwoordigen. Numerieke schattingen van de histaminegehalten zijn overbodig en deze zullen wij daarom voor het moment buiten beschouwing laten. Uit ieder der 4 verkregen contractiereeksen leiden wij nu een dergelijke rij rangnummers af. Wij meten daartoe met een passer de in een dergelijke reeks

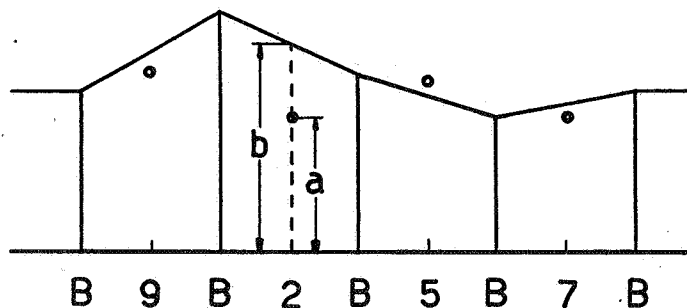


Fig. 3. Grafische voorstelling van een deel van een reeks contracties. De contracties, behorende bij de bekende oplossing, zijn door verticale lijnen weergegeven, waarvan de toppen door rechte lijnstukken verbonden zijn. De contracties, behorende bij de onbekende oplossingen, zijn door stippen aangegeven. Relatieve hoogte = a/b .

verkregen contracties op en zetten deze op gelijke afstanden verticaal uit (zie figuur 3; in deze figuur is de eerste contractiereeks voorgesteld).

In figuur 3 is tevens het verdere procédé geschetst, waarmee wij een schatting van de volgorde naar grootte der onbekende gehalten verkrijgen. De punten, die de bij B behorende contracties aangeven worden door een gebroken lijn verbonden en de relatieve hoogte a/b van ieder der contracties behorende bij de onbekende oplossingen (in de figuur alleen aangegeven bij no. 2) worden berekend. De volgorde naar grootte van deze relatieve hoogten nemen wij als schatting van de volgorde naar grootte van de onbekende histaminegehalten.

Op deze wijze worden dus 4 rijen rangnummers verkregen. Indien nu de hypothese H_0 , die wij in dit geval willen toetsen en die inhoudt, dat de onbekende histaminegehalten alle gelijk zijn, vervuld is, geldt voor ieder van deze rijen rangnummers, dat zij op volkomen toevallige wijze tot stand zijn gekomen en dat dus alle permutaties van de rij rangnummers 1, 2, . . . , 10 even waarschijnlijk zijn. Dit laatste is echter juist de hypothese, die getoetst kan worden met behulp van de methode der m rangschikkingen (met in dit geval $m = 4$). Deze toetsingsmethode zullen wij hier niet beschrijven, daar dit reeds eerder uitvoerig is geschied ([11], [12], [13] en [14]). Wij geven slechts de volgende korte aanduiding.

Bij ieder der bloedmonsters behoren nu dus 4 rangnummers (uit ieder der contractiereeksen één; dit zijn dus niet de gelote volgordenummers, maar de naar aanleiding van de relatieve hoogte a/b bepaalde rangnummers naar grootte). Voor ieder der monsters wordt de som van deze 4 rangnummers bepaald (deze sommen worden gewoonlijk kolomtotalen genoemd). Indien H_0 vervuld is, zullen deze sommen in het algemeen niet sterk uiteenlopen. Indien de onbekende histaminegehalten verschillend zijn zal dit echter vaak wel het geval zijn. De spreiding van de kolomtotalen wordt daarom als toetsingsgrootte voor H_0 gebruikt, waarbij grote waarden kritiek zijn, d.w.z. tot verwerping van H_0 aanleiding geven.

3.4 Enkele opmerkingen over de gewijzigde methode

Een uitvoerige behandeling van de eigenschappen van deze nieuwe methode en van de wijze, waarop ook andere hypothesen dan H_0 kunnen worden getoetst, zou ons te ver voeren. Wij zullen daarom volstaan met enkele opmerkingen.

Daarbij kunnen wij dan in de eerste plaats opmerken, dat wij nu niet, zoals bij de vroegere methode, slechts één rij van schattingen der 10 onbekende gehalten hebben verkregen, maar 4 onafhankelijke schattingen van de volgorde naar grootte. Hierdoor wordt de mogelijkheid geschapen de

hypothese H_0 te toetsen, waarvan in 3.2 was opgemerkt, dat dit oorspronkelijk niet mogelijk was. Desgewenst kunnen wij (al is dit voor de statistische verwerking overbodig) óók nog numerieke schattingen van de gehalten geven, analoog aan de volgens de gangbare methode verkregene, door de 4 contracties, behorende bij een bloedmonster, tezamen met de aangrenzende B-contracties te beschouwen en het gehalte hieruit te schatten.

De in 3.1 genoemde bezwaren zijn grotendeels opgeheven. De subjectieve elementen zijn nl. uit de bepalingen vrijwel volledig geëlimineerd. Alleen de keuze der 4 bekende gehalten (één voor iedere contractiereeks) is daarvan nog over, maar deze is nu voor alle 10 monsters dezelfde. De mogelijkheid van mislukkingen is nog aanwezig (soms contraheert de darm niet gewoon, maar loopt de wijzer langzaam weg). Dit zou, bij het hier beschreven probleem, het aantal van 4 rangschikkingen kunnen verminderen of ons kunnen noodzaken één of meer monsters buiten beschouwing te laten. Dit is echter bij sommige andere toetsen, die hier niet beschreven zijn, niet nodig. Het belangrijkste bezwaar, als derde genoemd in 3.1, is nu echter geheel opgeheven. De wisselende gevoeligheid van de caviadarm heeft op de bepalingen een veel geringere invloed dan tevoren, daar iedere onbekende oplossing 4 maal op geheel verschillende tijdstippen wordt gebruikt. Bovendien garandeert de loting van de 4 volgorden van bepaling ons, dat deze wisselende gevoeligheid de geldigheid van de gebruikte toetsingsmethode niet meer aantast. Wij kunnen ons anderzijds niet verhelen, dat deze gewijzigde methode moeilijker uit te voeren is dan de tot nu toe gebezigde. Dit is in de praktijk ook gebleken. De uitvoering vereist van de onderzoeker een grotere concentratie en snelheid van handelen. Dit bezwaar lijkt ons echter van minder belang dan de winst aan betrouwbaarheid van de conclusies, de winst aan mogelijkheden voor de statistische analyse en aan onderscheidingsvermogen van de analyse. Ten slotte staat de statistische analyse van waarnemingsresultaten niet los van het experiment. Een experiment is gericht op het trekken van conclusies en alles, wat aan deze conclusies voorafgaat, maakt deel uit van het experiment, ook de statistische analyse.

Het grotere onderscheidingsvermogen van de gewijzigde methode in vergelijking met de oorspronkelijk gebruikte bleek uit een in [7] beschreven proef-experiment, uitgevoerd door Dr Strengers en Dr Goosen van de Stichting voor Biochemisch-Klinisch Onderzoek, waarbij aan een aantal porties van één groot bloedmonster door toevoeging van verschillende bekende hoeveelheden histamine een stijgend histaminegehalte werd gegeven. Deze porties werden vervolgens gehalveerd en de twee verkregen reeksen monsters werden volgens de beide methoden behandeld. Daar nu de volgorde naar grootte van de histaminegehalten bekend was, kon bij beide

methoden, door toetsing van de bovengenoemde hypothese H_0 nagegaan worden, of de kunstmatig aangebrachte stijging duidelijk werd teruggevonden. Bij de gangbare methode werd daartoe de methode der rangcorrelatie van Kendall (o.a. beschreven in [12]) toegepast. Deze gaf een éézijdige overschrijdingskans 0,035. Bij de gewijzigde methode werd een door G. Elfring en J. H. Whitlock [17] beschreven toets toegepast, die eveneens op de methode der rangcorrelatie berust. De daarbij gevonden overschrijdingskans was $< 0,0001$.

Waar wij bij het in § 2 behandelde voorbeeld van toepassing van de momentenmethode konden wijzen op het typisch parametrische karakter van het probleem, dat daar ter sprake was, zijn hier de voordelen van de parameter vrije methode zo duidelijk, dat de conclusie ons gerechtvaardigd lijkt, dat deze nieuwe methoden, naast de parametrische, een eigen terrein van praktische toepassingen bestrijken.

Wij nemen de gelegenheid te baat de onderzoekers van de Stichting voor Biochemisch-Klinisch Onderzoek, in het bijzonder Dr Strengers en Dr Gooszen te bedanken voor hun medewerking bij het tot stand komen van deze publicatie.

Literatuur

- [1] D. van Dantzig, *Kadercursus mathematische Statistiek*, Mathematisch Centrum, Amsterdam 1947—1950.
- [2] H. R. van der Vaart, *Some remarks on the power function of Wilcoxon's test for the problem of two samples*, I and II, Proc. Kon. Ned. Ak. v. Wet. 53 (1950), pp. 494—506 and 507—520; Indag. Math. 12 (1950) pp. 146—158 and 159—172.
- [3] J. Sittig en Prof. Dr H. Freudenthal, *De Juiste Maat*, Bijenkorf-Publicatie no. 4, Stafleu, Leiden 1951.
- [4] J. Sittig, *Het probleem der waarnemingsfouten bij een massaal onderzoek*, Voordracht gehouden op de Statistische Dag op 26 Januari 1949, Statistica 3 (1949) p. 49—68.
- [5] K. Pearson, *On the probable error of frequency-constants*, Biometrika 2 (1903) p. 273; 9 (1913) p. 1; 13 (1920) p. 113.
- [6] M. A. Bricas, *Le système de courbes de Pearson et le schéma d'urne de Polya*, Gauthier-Villars, Paris 1949.
- [7] J. Hemelrijk en Mej. C. van Eeden, *De vergelijking van twee methoden ter bepaling van het histaminegehalte van bloed*, Rapport S 75 van de Statistische Afdeling van het Mathematisch Centrum, 1952.
- [8] Th. Strengers, *Een vergelijkend onderzoek over de histaminespiegel van het bloed bij mens, paard en rund*, Diss., Utrecht 1946.
- [9] J. A. H. Gooszen, *Onderzoekingen over de histaminestofwisseling bij schizofrenen*, Diss., Utrecht 1951.
- [10] R. A. Fisher, *On a distribution yielding the error functions of several well-known statistics*, Proc. Int. Math. Congress. Toronto 1924, pp. 805—813.
- [11] M. Friedman, *The use of ranks to avoid the assumption of normality implicit in the analysis of variance*, Jrn. Am. Stat. Ass. 32 (1937) pp. 675—699.

- [12] M. G. Kendall, *Rank correlation methods*, Griffin and Co., London, 1948.
- [13] J. Hemelrijk, *Rangcorrelatie en de schattingsproef van Varangot*, *Statistica* 5 (1950), pp. 216—225.
- [14] Ph. van Elteren, *Methode van de m rangschikkingen*, Hoofdstuk II van de cursus „Parametervrije Methoden”, Rapport S 59, Mathematisch Centrum, Amsterdam 1951.
- [15] M. G. Kendall and B. Babington Smith, *Tables of random sampling numbers*, Cambridge Un. Press, 1939.
- [16] J. H. Enters, *Het gebruik van toevalscijfers*, *Statistica* 5 (1951) pp. 81—96.
- [17] G. Elfring and J. H. Whitlock, *A simple trend test with application to erythrocyte size data*, *Biometrics* 6 (1950), pp. 282—288.