STATISTISCHE AFDELING
Rapport S 50
door J. Hemelrijk en
H. Nieuwenhuys

November 1950

Clearanceproeven met honden.

1. Inleiding.

Vijf honden (Herta, Marja, Tilly, Tonny en Bruno) werden elk ingespoten met 30 cc van een oplossing (10 cc intraveneus en 20 cc subcutaan) die 5% creatine (CR) en 2% para-amino-hippuurzuur (PAH) bevatte.

Onder de Clearance van een stof op tijdstip t verstaat men:

$$\text{Cl.} = \frac{\text{concentratie in de urine} \times \text{hoeveelheid urine}}{\text{concentratie in het bloed}},$$

alles gemeten op tijdstip t.

Voor t werden drie tijdstippen I, II en III gekozen, waarvan het eerste drie kwartier na de inspuiting.

De hoeveelheid urine (~~diurese~~), afgescheiden door de nieren, (de diurese) werd als volgt gemeten: de nieren van de honden werden vlak voor de inspuiting leeggezogen; op tijdstip I gebeurde dit weer en de nu verkregen hoeveelheid was dan de hoeveelheid urine op tijdstip I. Evenzo voor tijdstippen II en III. De concentratie werd gemeten aan bloed- en urinemonsters. Dit procédé werd een aantal malen, met tussenpozen van enige dagen, herhaald.

Na verloop van een aantal proefdagen werd de proef in zoverre iets gewijzigd, dat bovendien een met het gewicht van de hond evenredige hoeveelheid van een zekere stof X peroraal werd toegediend. Verder bleef de proef ongewijzigd.

Het aantal waarnemingen op iedere der tijdstippen I, II en III verricht is in onderstaand staatje weergegeven:

Hond	Aantal waarnemingen zonder ~~met~~ X	Aantal waarnemingen met ~~zonder~~ X
Herta	9	4
Marja	9	4
Bruno	5	2
Tilly	5	4
Tonny	8	1

Bij Tilly waren enige waarnemingen mislukt.

Bij Tonny was een creatine-bepaling zonder toediening van X mislukt.

Het doel van de proeven was in hoofdzaak na te gaan of de stof X invloed uitoefent op de Clearance.

Wij hebben verder onderzocht of er afhankelijkheid bestond tussen de grootheden Cl(PAH), Cl(CR) en hun quotient $\frac{Cl(CR)}{Cl(PAH)}$, de filtratiefactor genaamd en tenslotte of er verband was tussen de diurese en de Clearances.

Men lese ter inleiding van de volgende paragrafen het als appendix bijgevoegde memorandum S 47 (M 6) over de gang van zaken bij het toetsen van hypothesen.

2. Statistische verwerking der waarnemingsresultaten.

A. Onderzoek van de invloed der stof X.

Teneinde na te gaan of de stof X de Clearances beïnvloedde werd voor ieder der honden het gemiddelde bepaald van de op tijdstip I bij de verschillende proeven verkregen uitkomsten voor Cl(PAH) en Cl(CR) van de dagen waarop de stof X niet was toegediend. Hetzelfde werd gedaan voor de dagen waarop de stof X wel was toegediend. De gehele bewerking werd ook voor de tijdstippen II en III uitgevoerd.

Noem de gemiddelden zonder toediening van X: z_i en de overeenkomstige gemiddelden met toediening van X: m_i.

Indien wij onderstellen, dat de Clearances bij iedere hond apart (voor PAH en CR afzonderlijk), zowel met als zonder de toediening van X, symmetrisch ten opsichte van hun gemiddelde verdeeld zijn[1], dan is de grootheid m_i-z_i, indien het middel X geen invloed heeft, symmetrisch ten opsichte van 0 verdeeld.

Wij toetsen nu deze symmetrie ten opzichte van 0, door op m_i-z_i de symmetrietoets T_2 van J.Hemelrijk toe te passen (zie bijlage S 47 (M 10)). Het resultaat was niet significant (Overschrijdingskans $\sim$ 0,3 voor de Cl(PAH), terwijl

---

1) Wij weten niet in hoeverre deze onderstelling gerechtvaardigd is. Zij is echter zwakker dan de gewoonlijk gebruikte onderstelling van normaliteit. Het vermijden van iedere onderstelling, was, wegens de geringe uitgebreidheid van het materiaal en wegens het feit, dat het aantal met en zonder X verrichte metingen niet gelijk was, niet mogelijk.

voor Cl(CR) een nog grotere overschrijdingskans gevonden werd).
Vergelijk grafiek 1, waaruit men wel reeds ziet, dat de eventueel aanwezige asymmetrie t.o.v. 0 gering is.

Gewoonlijk wordt op een stelsel gegevens als het onderhavige de toets van Student toegepast voor de hypothese, dat het gemiddelde van deze verschillen gelijk aan 0 is (zie S 47 (M 8)). Daartoe moet men echter onderstellen, dat alle verschillen afkomstig zijn uit <u>dezelfde normale</u> verdeling, een onderstelling, die hier zeker niet vervuld is, daar m_i en z_i gemiddelden van voor iedere hond verschillende aantallen waarnemingen zijn. Ter contrôle werd de overschrijdingskans, behorende bij deze toets, toch berekend. Deze bedroeg 0,11 voor Cl(PAH) en 0,33 voor Cl(CR), zodat er ook op grond van deze toets geen reden bestaat te onderstellen, dat de stof X invloed op de Clearance uitoefent.

Ook werd op de reeks van op tijdstip I verkregen waarden der Cl(PAH) van elke hond (met en zonder X) de toets van Wilcoxon (zie bijlage S 47 (M 7)) toegepast (en wel op de waarnemingen zelf, niet op hun gemiddelden). Hetzelfde werd gedaan voor de tijdstippen II en III. Behoudens in één enkel geval vonden wij ook hier geen overschrijdingskansen kleiner dan 0,05. Eén uitzondering op een dergelijke serie proeven is echter te verwachten ook als de stof X geen uitwerking heeft.

Vervolgens is met behulp van bovengenoemde symmetrietoets nagegaan welke constante waarde men minstens bij één der groepen gegevens (n.l. die, waarbij X is toegediend) moest optellen om een significante uitkomst (met onbetrouwbaarheid 0,05) te verkrijgen. Deze constante blijkt voor de Cl(PAH) gelijk aan 12 te zijn, dit is dus van de orde van 10% van de waarde van Cl(PAH).

Indien dus alle uitkomsten van Cl(PAH) na toediening van X 12 groter waren geweest, dan zij nu waren, zou het resultaat significant zijn geweest met een onbetrouwbaarheidsdrempel 0,05. De overeenkomstige berekening voor Cl(CR) gaf voor deze verschuivingswaarde 3,5, d.i. iets minder dan 10% van de gemiddelde waarde van Cl(CR).

B. Onderzoek van het verband tussen Cl(CR) en Cl(PAH).

Gezien de resultaten van het onder A uitgevoerde onderzoek is in het vervolg geen onderscheid gemaakt tussen de uitkomsten met en zonder de stof X, daar deze (in het algemeen niet zonder meer aan te bevelen) methode het weinig uitgebreide materiaal vergroot.

Alvorens de afhankelijkheid van de Cl(CR) en de Cl(PAH) na te gaan is het volgende vooronderzoek verricht:

Om te zien of wij de waarnemingen bij de verschillende honden verricht, tezamen zouden nemen, werd (zie grafiek 2) voor Marja en Herta (waarvoor de meeste waarnemingen waren verricht) de Cl(CR) tegen de Cl(PAH) uitgezet, (de bij Herta behorende punten zijn daarbij aangegeven door een stip en de bij Marja behorende punten door een kruisje) en op de zo verkregen "puntenwolk" werd een toets voor twee steekproeven toegepast (zie S 47 (M 19)).

Hierbij werd een overschrijdingskans $< 0{,}001$ gevonden zodat het niet redelijk zou zijn, de bij beide honden behorende waarnemingen als waarnemingen uit één verdeling te beschouwen. Deze groepen gegevens werden daarom verder afzonderlijk onderzocht. Bij de andere honden was het aantal waarnemingen te gering om een apart onderzoek mogelijk te maken.

De grafieken 3 en 4 geven de waarnemingsuitkomsten weer voor de honden Herta en Marja apart, waarbij op de horizontale as Cl(PAH) en op de verticale as Cl(CR) is uitgezet. Men ziet direct, dat er een sterke afhankelijkheid tussen Cl(CR) en Cl(PAH) bestaat. Op het oog lijkt het verband tussen deze twee grootheden in beide gevallen, afgezien van stochastische afwijkingen, ongeveer lineair. Gewoonlijk voert men nu de "filtratiefactor", d.i. het quotiënt van de twee grootheden Cl(CR) en Cl(PAH), in met het doel een grootheid te verkrijgen, die niet meer afhankelijk is van Cl(CR) of Cl(PAH). Dit betekent dus, dat men verwacht, dat in het puntendiagram dat verkregen wordt door Cl(CR) of Cl(PAH) tegen de filtratiefactor uit te zetten, de filtratiefactor bij stijgende Cl geen systematische stijging of daling zal vertonen.

De filtratiefactor zal echter slechts dan aan deze eis voldoen, indien het in de grafieken 3 en 4 duidelijk zichtbare verband tussen Cl(CR) en Cl(PAH) door een rechte lijn <u>door de oorsprong</u> weergegeven kan worden. Deze laatstgenoemde eigenschap werd daarom getoetst op een in de Appendix beschreven wijze, voor de gegevens van Herta en Marja apart en voor de gegevens behorende bij de overige honden tezamen. In alle drie gevallen werden overschrijdingskansen $\gg 0,05$ gevonden, zodat er geen reden is de getoetste hypothese te verwerpen.

Op grond hiervan mag men dus verwachten, dat de filtratiefactor inderdaad geen afhankelijkheid van Cl(CR) of Cl(PAH) meer zal vertonen. Ter contrôle zijn in de grafieken 5, 6, 7 en 8 de door de filtratiefactor aangenomen waarden uitgezet tegen die van Cl(PAH) en Cl(CR), voor Herta en Marja afzonderlijk. Deze puntendiagrammen wijzen inderdaad geenszins op afhankelijkheid, hetgeen bovendien nog gecontroleerd is door toepassing van een onafhankelijkheidstoets met behulp van de methode der dubbele dichotomie (zie bijlage S 47 (M 15)) en door toepassing van de hoektoets voor onafhankelijkheid (zie bijlage S 47 (M 11)).

Bij deze 8 toetsingen werd één overschrijdingskans $< 0,05$ gevonden en wel bij toepassing van de hoektoets op grafiek 7. De gevonden overschrijdingskans was ongeveer 0,02, terwijl met de methode der dubbele dichotomie voor dit geval een grote overschrijdingskans gevonden werd. Eén kleine overschrijdingskans is bij een serie toetsingen als de onderhavige echter zeer goed mogelijk, ook als de getoetste hypothese van onafhankelijkheid juist is. Wij constateren derhalve, dat het onderzochte materiaal geen aanwijzingen in de richting van afhankelijkheid tussen de filtratiefactor en Cl(CR) of Cl(PAH) geeft.

Opmerking: In de grafieken 5, 6, 7 en 8 zijn de waarnemingsparen, die zonder resp. met toediening van de stof X zijn verricht, met punten resp. kruisjes aangegeven. Hiervoor zijn de aanduidingen "boven-" resp. "onder de streep" gebruikt. Aangezien er op het oog geen systematisch verschil tussen de ligging der punten en kruisjes te zien is, is deze onderscheiding bij bovenvermelde toetsingen niet gemaakt.

Onder A is opgemerkt, dat de waarnemingsuitkomsten voor de honden Herta en Marja sterk verschilden. Nu gebleken is, dat de filtratiefactor voor ieder der honden apart een in ieder geval veel kleinere afhankelijkheid van Cl(CR) en Cl(PAH) vertoont dan deze twee grootheden onderling, is het de moeite waard na te gaan of deze filtratiefactor voor de honden onderling systematisch verschilt. Met behulp van de toets van Wilcoxon (zie S 47 (M 7)) is daarom de hypothese getoetst, dat de waarden van de filtratiefactor, die bij Herta en Marja gevonden zijn steekproeven uit dezelfde collectie vormen. Daar deze toets een overschrijdingskans $< 0{,}01$ geeft, concluderen wij, dat ook de filtratiefactor voor de verschillende honden systematisch verschilt. Uit de waarnemingen blijkt, dat de filtratiefactor bij Marja systematisch hoger ligt dan bij Herta.

C. Onderzoek van een eventuele afhankelijkheid van Clearance en diurese.

In de grafieken 9 en 10 werd, voor de honden Herta resp. Marja, de diurese uitgezet tegen Cl(PAH) en vervolgens werden (zonder de stippen en kruisjes te onderscheiden) de boven reeds genoemde toetsen voor onafhankelijkheid toegepast. Deze leverden overschrijdingskansen, die alle groter waren dan 0,10, zodat er geen reden is de onafhankelijkheid van Cl(PAH) en de diurese te verwerpen. Een zelfde resultaat werd verkregen bij toetsing van de onafhankelijkheid van Cl(CR) en diurese.

3. Conclusies.

1. Uit het ons verschafte waarnemingsmateriaal bleek geen effect van toediening der stof X op de Clearance. Weliswaar was er een verhoging, maar deze was niet zo groot, dat zij niet aan het toeval kan worden toegeschreven. Was de verhoging ongeveer 10% van de gemiddelde Clearance-waarde groter geweest, dan zou men met een onbetrouwbaarheidsdrempel 0,05 tot een beïnvloeding van de Clearance door de stof X hebben kunnen besluiten.

2. Er bleek een sterk verband te bestaan tussen Cl(CR) en Cl(PAH), dat, voor zoverre het beperkte materiaal betreft, afgezien van stochastische afwijkingen zeer wel weergegeven kan worden door een rechte lijn, die door de oorsprong gaat.

3. In overeenstemming met 2 werd geen afhankelijkheid gevonden tussen de filtratiefactor en Cl(CR) of Cl(PAH).

4. De Clearance is voor verschillende honden duidelijk verschillend. Hetzelfde geldt voor de filtratiefactor.

5. Er werd geen afhankelijkheid gevonden tussen Clearance en diurese.

Opmerking: Daar de uitgebreidheid van het materiaal gering was, is het zeer wel mogelijk dat eventueel aanwezige effecten niet konden worden ontdekt. Dit is te meer het geval, daar, wegens punt 3 van de conclusie, bij de meeste der toegepaste toetsingen de gegevens der verschillende honden niet konden worden samengenomen.

4. Appendix.

De memoranda S 47, M 6, 7, 8, 10, 11, 15 en 19 aan dit rapport toegevoegd, beschrijven zeer in het kort de toegepaste toetsingsmethoden.

Hieronder volgt nog de beschrijving van de toets, die toegepast is, om de aard van het verband tussen Cl(CR) en Cl(PAH) na te gaan. Wij geven Cl(PAH) aan met x en Cl(CR) met y en toetsen dan de hypothese, dat x en y een simultane waarschijnlijkheidsverdeling bezitten, die door een schaalverandering van de y-as omgezet kan worden in een verdeling, die symmetrisch is ten opsichte van de lijn x=y door de oorsprong (x=0, y=0).

De waarnemingen $(x_1, y_1), \ldots, (x_n, y_n)$ (zie b.v. grafiek 3) worden als een "puntenwolk" in een vlak uitgezet. Vervolgens trekken wij een lijn door de oorsprong (in grafiek 3 de lijn AB), sodanig dat ter weerszijden daarvan evenveel punten liggen. Daarna trekken wij een lijn CD, die een hoek met de x-as maakt, die het supplement is van de hoek, die AB met deze as maakt. Wij zorgen verder, dat ter weerszijden van CD een gelijk aantal punten van de puntenwolk liggen. Op deze wijze wordt de puntenwolk in 4 delen verdeeld en de te toetsen hypothese houdt in, dat de verdelingen der paren kenmerken:

onder- en boven AB

en:

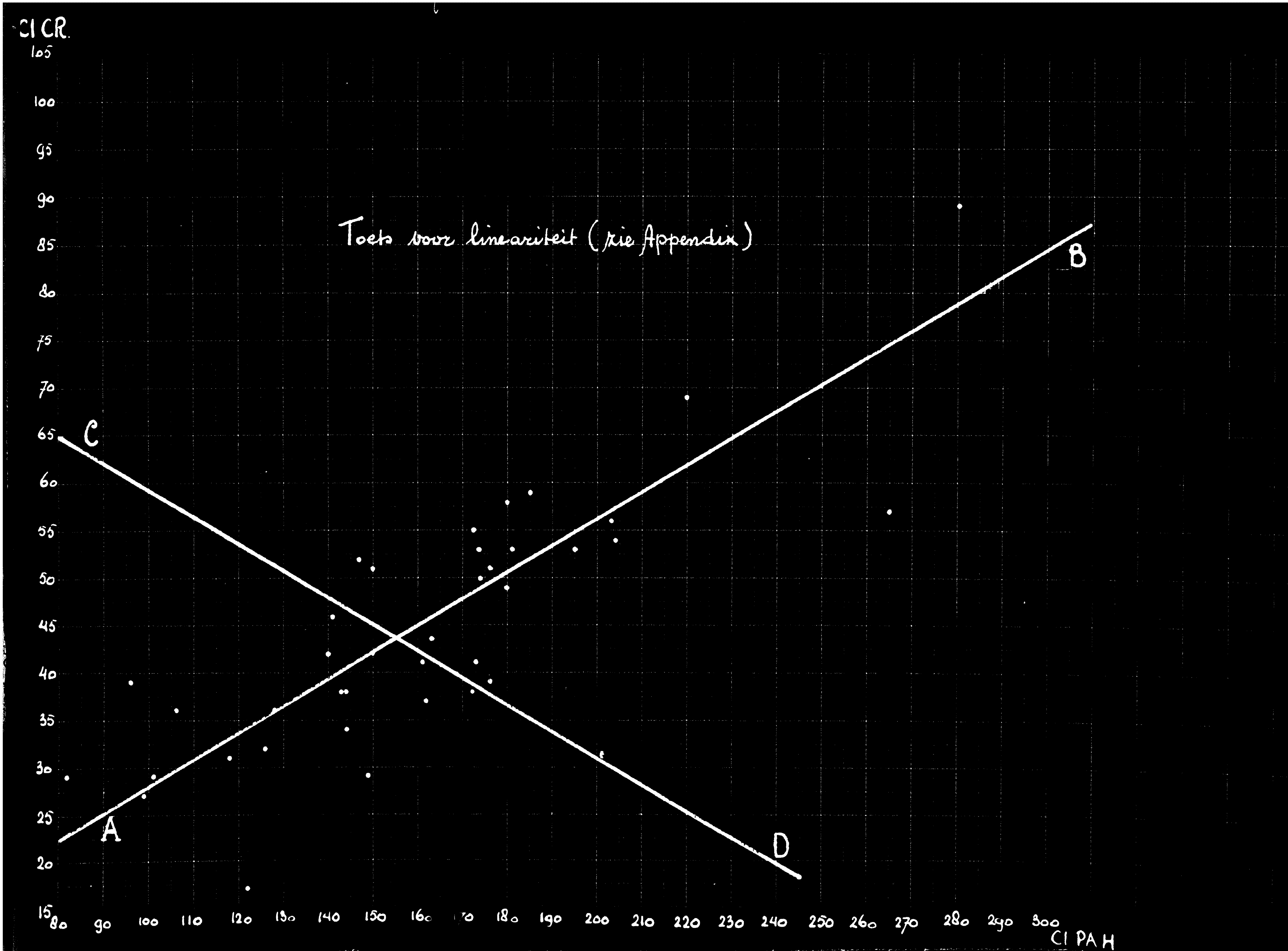
onder- en boven CD

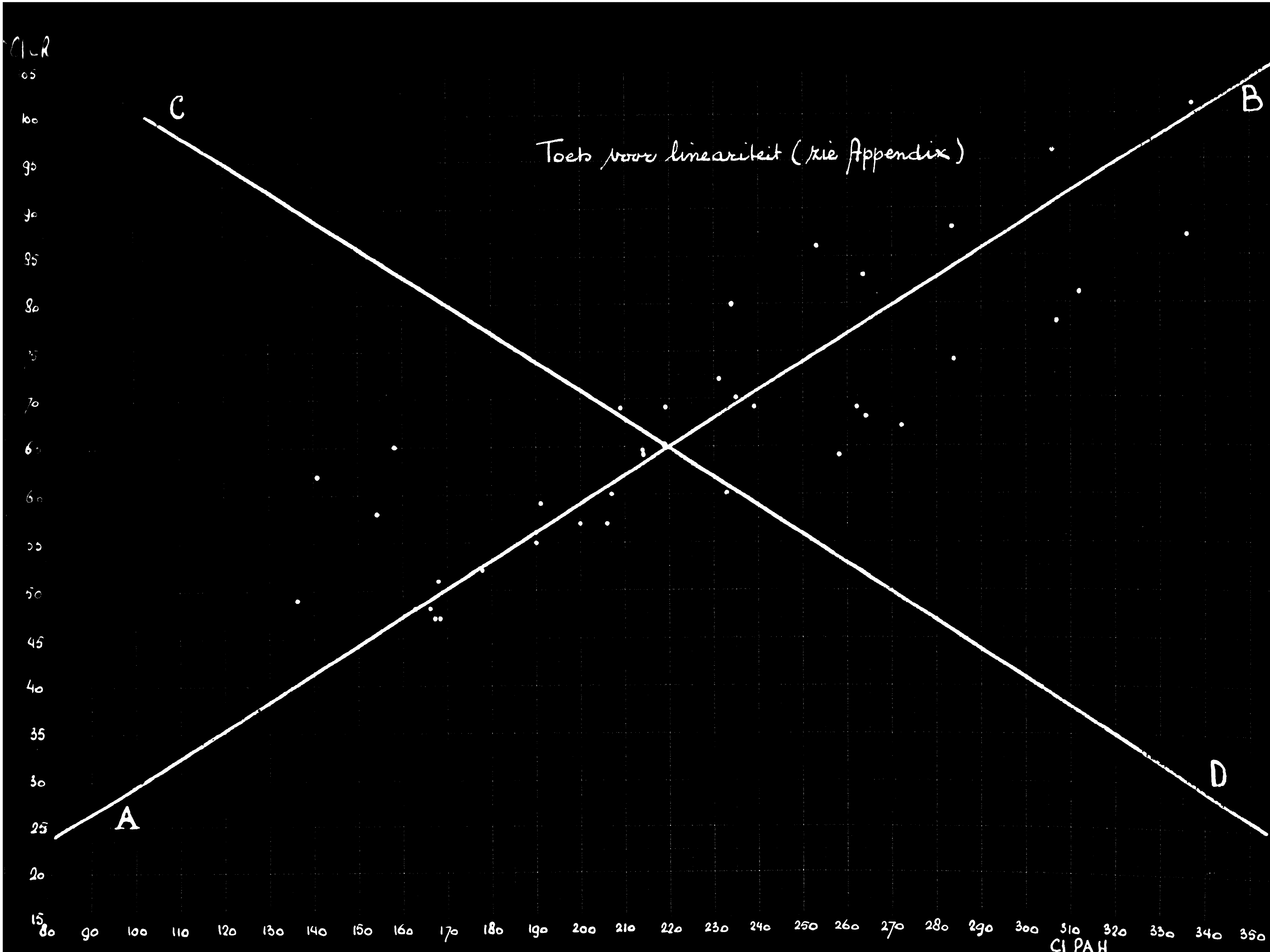
over de punten (asymptotisch, d.w.z. voor $n \to \infty$) onafhankelijk van elkaar zijn. Indien het aantal waarnemingen niet te klein is, b.v. minstens in de buurt van 40 ligt, kan men deze onafhankelijkheid toetsen door het toepassen van de χ^2-toets voor een 2 x 2-tabel. (Vgl. b.v. M.G.Kendall, The advanced theory of Statistics, I, p. 299 en 300).

Deze toets is in § 2.B van dit rapport gebruikt.

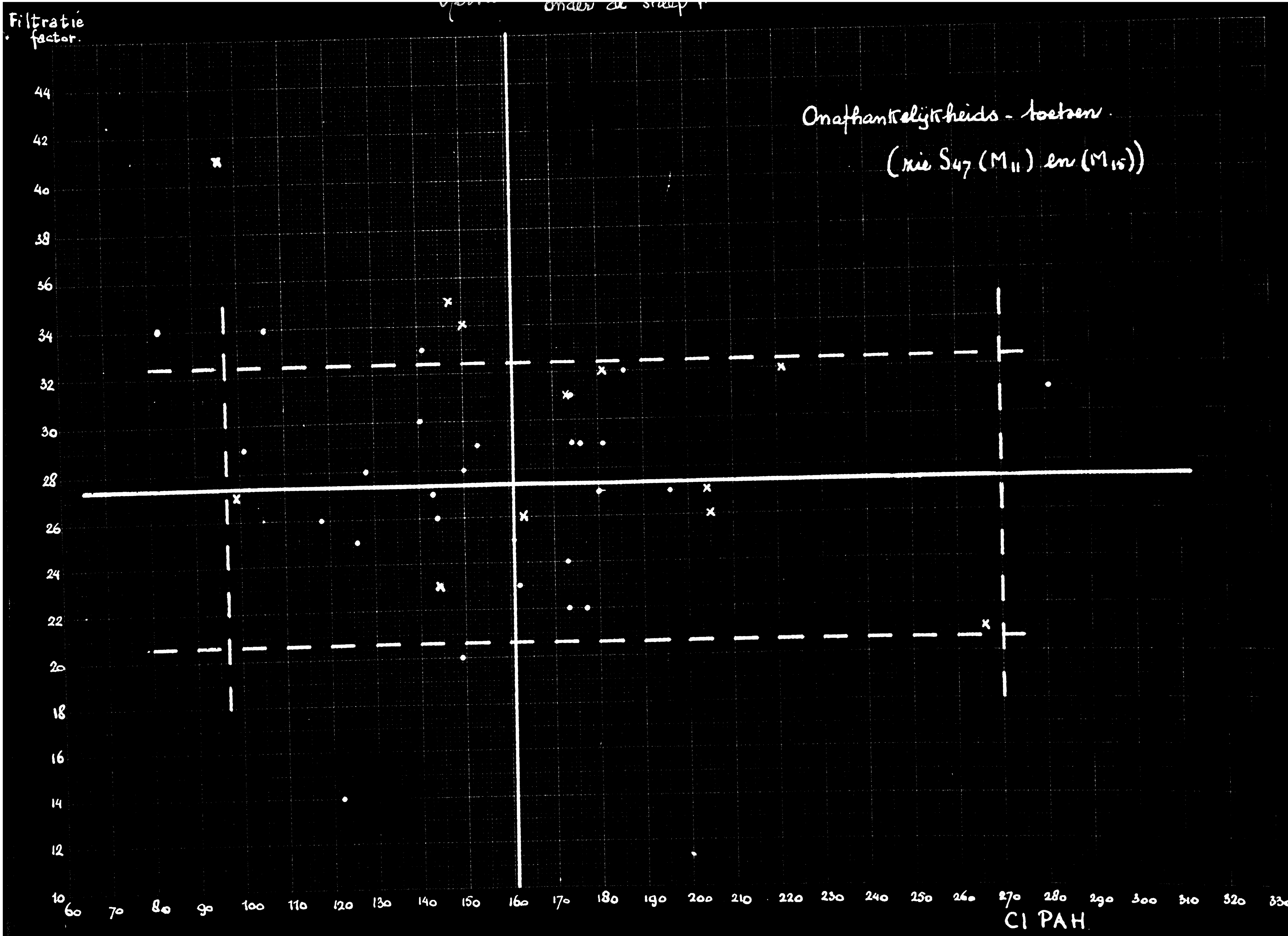
Symmetrietoets T_2, toegepast op de grootheden $m_i - x_i$ (zie $S_{47}(M_{10})$)

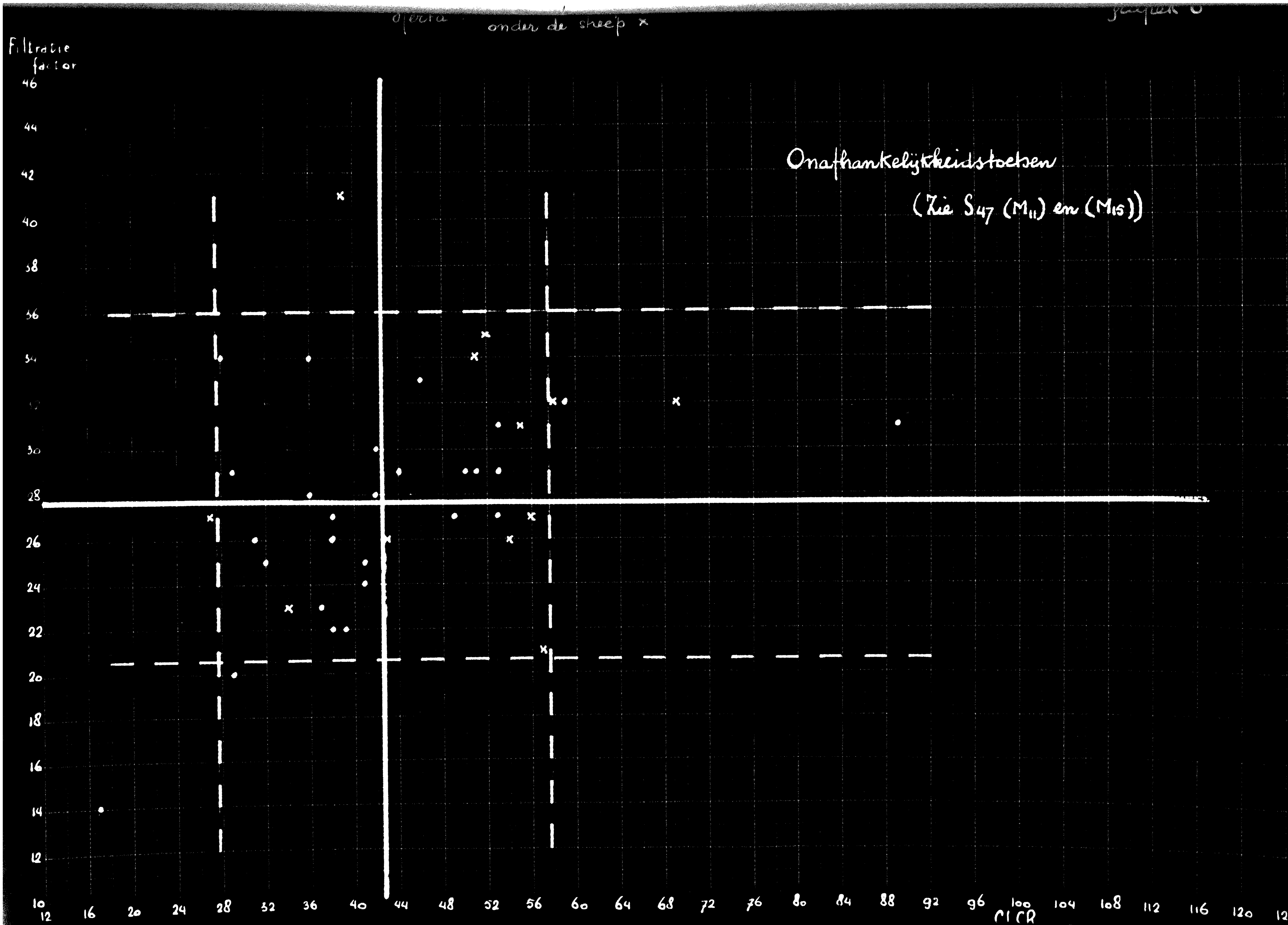

Herta •
Marja ×
CICR
Toets voor 2 steekproeven (zie S_{47} (M.19))
A
B
C
D
L_1
L_2
CJ PAH

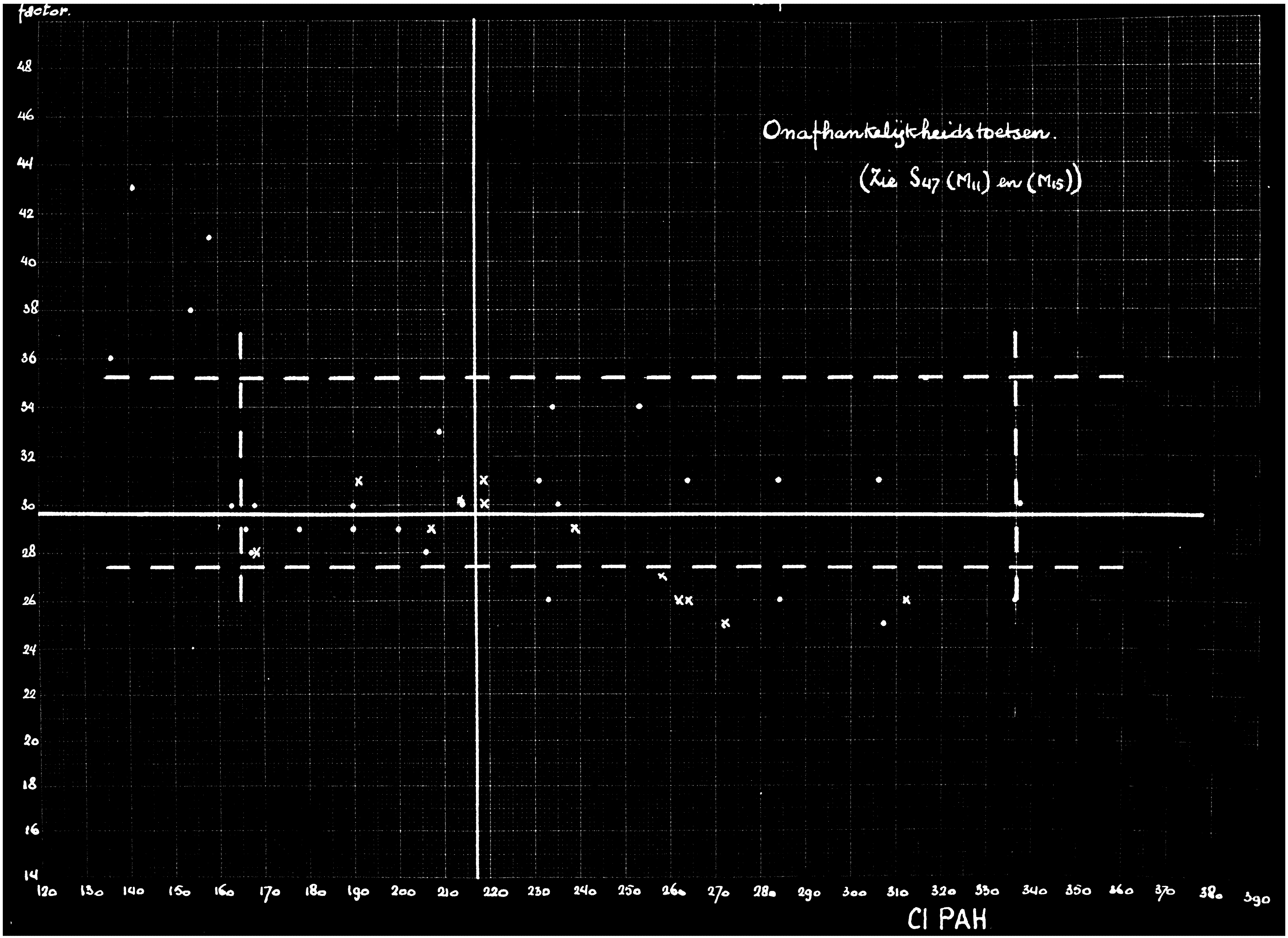

Cl CR.
Toets voor lineariteit (sie Appendix)
A
B
C
D
Cl PA H
105
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300

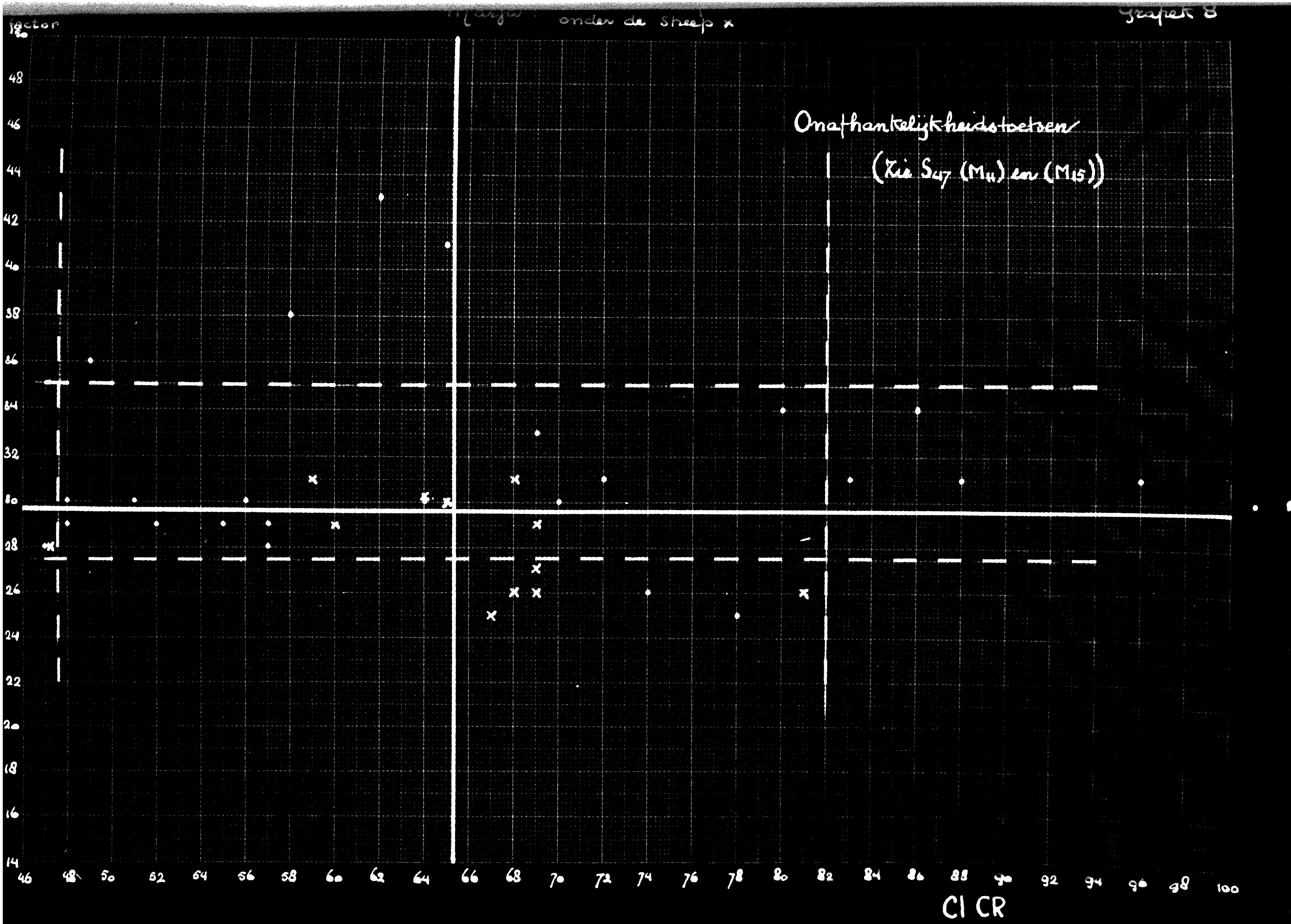


Toets voor lineariteit (zie Appendix)
Cl PAH
A
B
C
D

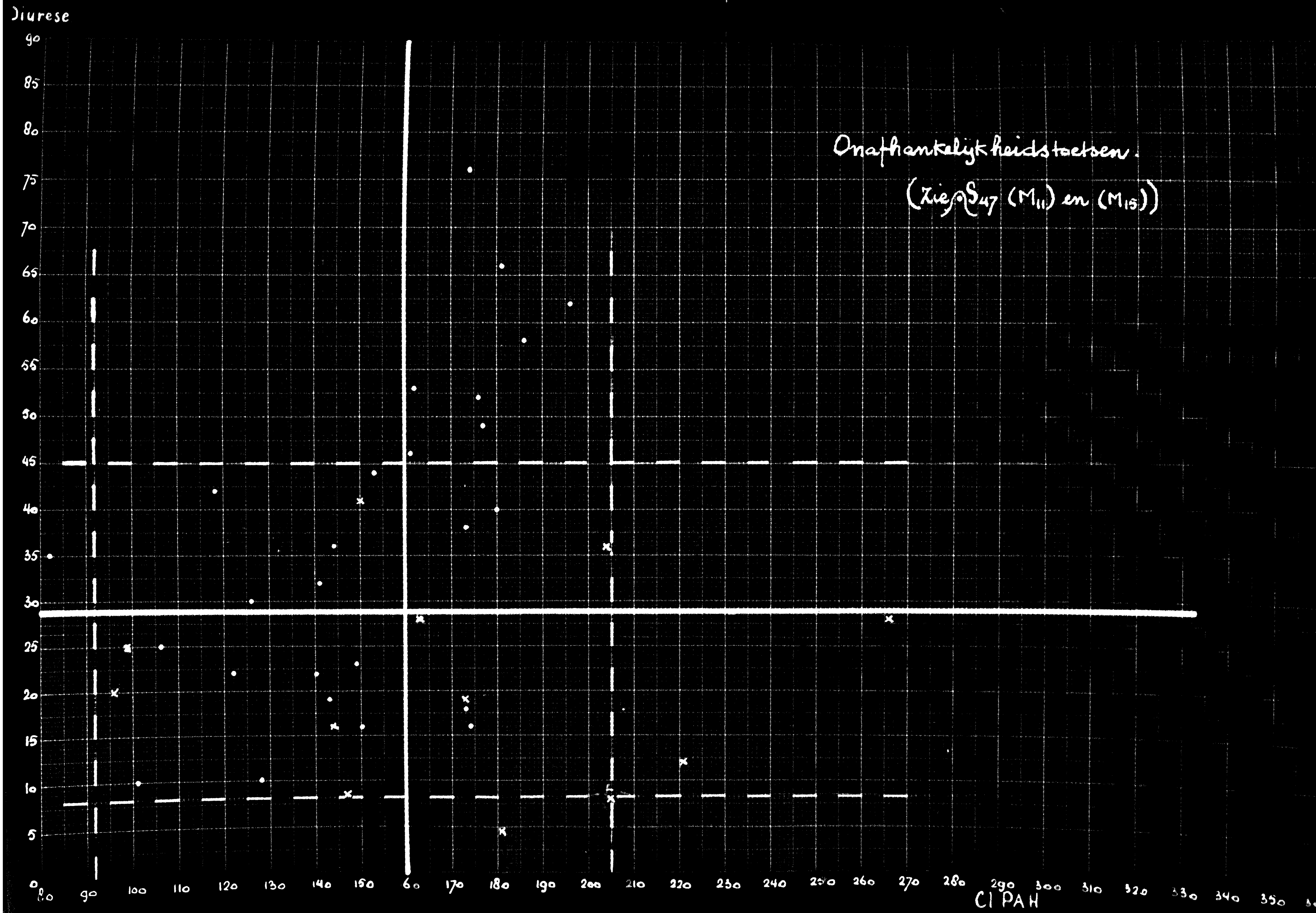

Filtratie factor.
44
42
40
38
36
34
32
30
28
26
24
22
20
18
16
14
12
10
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
Cl PAH
Onafhankelijkheids - toetsen.
(zie S_{47} (M_{11}) en (M_{15}))


onder de skeep ×
Filtrasie factor
Onafhankelijkheidstoetsen
(Zie S47 (M11) en (M15))
CICR
10 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88 92 96 100 104 108 112 116 120 124
12 14 16 18 20 22 24 26 28 30 34 36 38 40 42 44 46


factor.
Onafhankelijkheidstoetsen.
(Zie S47 (M11) en (M15))
48
46
44
42
40
38
36
34
32
30
28
26
24
22
20
18
16
14
120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390
Cl PAH


Grafiek 8
onder de sheep x
Onafhankelijkheidstoetsen
(Zie S_{47} (M_{11}) en (M_{15}))
factor
48
46
44
42
40
38
36
34
32
30
28
26
24
22
20
18
16
14
46
48
50
52
54
56
58
60
62
64
66
68
70
72
74
76
78
80
82
84
86
88
90
92
94
96
98
100
Cl CR

Grafiek 9


Diurese
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350
Cl PAH
Onafhankelijkheidstoetsen.
(Zie § 47 (M11) en (M15))

Marja : boven de sheep •
onder de sheep ×
Grafiek 10
Diurese
Onafhankelijkheidstoetsen
(Zie S47 (M11) en (M15))
Cl PAH

<u>Algemene gang van zaken bij het toetsen van een hypothese.</u>[1]

De toetsing van een hypothese $\mathcal{H}_0$ berust steeds op een aantal waarnemingen $x_1, x_2, \ldots, x_n$ van één of meer stochastische grootheden[2], of op enige groepen van waarnemingen (bv. twee steekproeven).

Bij een toets behoort een <u>toetsingsgrootheid</u> $\underline{u}$ (soms meer dan één), die een functie is van bovengenoemde stochastische grootheden en die, voor de waargenomen waarden $x_1, x_2, \ldots, x_n$ een waarde aanneemt, die berekend kan worden (bv.: het gemiddelde der waarnemingen, of de spreiding, of het verschil van de gemiddelden van twee waarnemingen).

De toetsingsgrootheid wordt steeds zo gekozen, dat men, op grond van de onderstelling, dat $\mathcal{H}_0$ juist is, de waarschijnlijkheidsverdeling van deze grootheid kan berekenen.

Vervolgens kiest men een verzameling Z van mogelijke uitkomsten van $\underline{u}$, en wel op zodanige wijze, dat de kans, dat $\underline{u}$ een in Z gelegen waarde aanneemt, onder de hypothese $\mathcal{H}_0$, gelijk is aan een gegeven getal α, zodat Z dus van α afhankelijk is . Z heet de <u>kritieke zône</u> van de toets, α de <u>onbetrouwbaarheidsdrempel</u> (Engels: level of significance). Voor α neemt men veelal de waarde 0,05 of 0,01.

Men verwerpt nu $\mathcal{H}_0$ op grond van de waarnemingen $x_1, x_2, \ldots, x_n$, indien de bij deze waarnemingen behorende waarde van $\underline{u}$ in Z ligt. Dit wordt vaak uitgedrukt door te zeggen, dat het resultaat van het experiment "<u>significant</u>" is. De waarde van α moet dan echter worden vermeld. De kans, dat dit zal gebeuren, is, indien $\mathcal{H}_0$ juist is, gelijk aan α. Derhalve is α de <u>kans op ten onrechte verwerping van de juiste hypothese</u>, ook de kans op een fout van de eerste soort genoemd. Indien men deze methode toepast, met $\alpha = 0{,}05$ resp. 0,01, zal men in gemiddeld ongeveer één op 20 resp. op 100 van de gevallen, <u>waarin de hypothese</u> die men toetst juist is, deze toch verwerpen.

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

2) Een stochastische grootheid is een grootheid, die een waarschijnlijkheidsverdeling bezit, of, anders gezegd, een grootheid, die voor de elementen van een <u>collectie</u>(<u>universum</u>,<u>populatie</u>) gedefinieerd is en daarop allerlei waarden aanneemt. Stochastische grootheden worden aangegeven door <u>onderstreepte</u> letters.

3) Soms kan men slechts bereiken, dat deze kans   is.

De toetsingstheorie biedt in het algemeen geen mogelijkheid om tot aanvaarding van een hypothese te komen. Indien een bepaalde hypothese $\mathcal{H}_0$ niet verworpen kan worden, is dit gewoonlijk met een hele verzameling van hypothesen tegelijk het geval. Niet-verwerpen staat dus niet gelijk met aanvaarden.

Wel zal men vaak in de loop van een statistische analyse bepaalde onderstellingen, die plausibel schijnen en voor de verdere analyse van nut zijn, toetsen, alvorens ze bij de verdere bewerking van het materiaal te gebruiken. Worden zij dan op grond van de toets niet verworpen, dan houdt dit in zo verre een rechtvaardiging van die onderstellingen in, dat een grote afwijking door de toets veelal wel zou zijn ontdekt. Indien men dan verder de onderstellingen gebruikt, verwaarloost men eventueel aanwezige afwijkingen van onbekende grootte, die echter niet zo groot zijn, dat zij door de toets zijn ontdekt.

Vele toetsen gelden zelf alleen onder bepaalde onderstellingen omtrent de waarschijnlijkheidsverdelingen der stochastische grootheden, waarvan waarnemingen zijn verricht. Deze nevenvoorwaarden dienen steeds uitdrukkelijk te worden vermeld en, zo mogelijk, zelf te worden getoetst.

In plaats van de onbetrouwbaarheidsdrempel α wordt vaak bij de uitslag van een toetsing de overschrijdingskans k opgegeven; dit is de kleinste waarde van α, waarbij in het betrokken geval, nog tot verwerping van $\mathcal{H}_0$ zou zijn overgegaan; anders gezegd: de kleinste α, waarvoor de gevonden waarde der toetsingsgrootheid nog juist in de (bij α behorende) kritieke zône Z ligt. Wordt dus de waarde k opgegeven en werkt men met onbetrouwbaarheidsdrempel α, dan wordt verworpen, indien $k \leq \alpha$ is.

Voor het onderscheid tussen één- en tweezijdige toetsing en de keuze tussen deze twee mogelijkheden vergelijke men bv. de tweede hieronder gegeven litteratuurplaats. Wij moeten hier volstaan met de opmerking, dat éénzijdige toetsing veelal eerder tot verwerping van $\mathcal{H}_0$ leidt, maar dat deze slechts onder bijzondere omstandigheden kan worden toegepast.

Litteratuur:

J.Neyman, First course in probability and statistics, New York, 1950, Chapter 5.

J.Hemelrijk en H.R. van der Vaart, Het gebruik van één- en tweezijdige overschrijdingskansen voor het toetsen van hypothesen, Statistica 4 (1950) p.54-66.

De toets van Wilcoxon.[1]

Deze methode dient tot het toetsen van de hypothese $\mathcal{H}_0$, inhoudend dat twee steekproeven $x_1, \ldots, x_n$ en $y_1, \ldots, y_m$ afkomstig zijn uit één collectie (ook populatie of universum genaamd). Zij is strict genomen, toepasbaar onder de voorwaarde, dat er geen enkel paar waarden (x_i, y_j) is met $x_i = y_j$. Verdere voorwaarden zijn voor de toepassing niet nodig, terwijl ook de zojuist genoemde, indien er niet teveel dergelijke paren zijn, de toepassing van de toets weinig hindert.

De toetsingsgrootheid $\mathcal{U}$ is het aantal paren (x_i, y_j) waarvoor $x_i > y_j$ is (het aantal "inversies"). Daar er $n.m$ dergelijke paren zijn, kan $\mathcal{U}$ alle gehele waarden van 0 tot en met $n.m$ aannemen. Is $\mathcal{U}$ groot, dan liggen er veel waarden x_i verder naar rechts dan waarden y_j, is $\mathcal{U}$ klein, dan juist weinig.

De kritieke zône $\mathcal{Z}$ neemt nu daarom de kleine en de grote waarden van $\mathcal{U}$ en wel van beide zoveel, dat de gekozen onbetrouwbaarheidsdrempel α niet overschreden wordt.

Voor éénzijdige toetsing, te onderscheiden in linker- en rechter-éénzijdige toetsing, gebruikt men kritieke zônes $\mathcal{Z}_1$, resp. $\mathcal{Z}_2$, die geheel bestaan uit kleine, resp. grote waarden van $\mathcal{U}$.

Verwerping van $\mathcal{H}_0$ ten gevolge van het vinden van een grote (resp.kleine) waarde van $\mathcal{U}$ wijst erop, dat $x_1, \ldots, x_n$ en $y_1, \ldots, y_m$ steekproeven uit verschillende collecties zijn, waarbij de op de x-collectie aangenomen waarden systematisch groter (resp.kleiner) dan de op de y-collectie aangenomen waarden zijn.

Litteratuur:

F. Wilcoxon, Individual comparisons by ranking methods, Biometrics 1 (1945), p.80-83.

H.B. Mann and D.R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Ann.Math.Stat.18 (1947),p.50-60. Bevat tabellen voor n en m ≦ 8.

H.R. van der Vaart, Some remarks on the power function of Wilcoxon's test for the problem of two samples, Proceedings van de Kon.Ned.Ak.v.Wet.,53 (1950), p.494-520.

H.R. van der Vaart, Gebruiksaanwijzing voor de toets van Wilcoxon, met tabellen voor n en m ≦ 10, Rapport S32 (M4)(1950).

D. van Dantzig, Kadercursus Mathematische Statistiek, Math. Centrum, Amsterdam(1947-50), hoofdstuk 6, 3.

1)Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

Toets van "Student" voor het gemiddelde van een normale verdeling.[1]

Gegeven: de steekproef $x_1, x_2, \ldots, x_n$ uit een normale collectie. Anders gezegd: $x_1, \ldots, x_n$ zijn onafhankelijke waarnemingen van de stochastische grootheid $\underline{x}$, die normaal verdeeld is (de zgn. waarschijnlijkheidsverdeling van Gauss bezit)[2].

$\mathcal{H}_0$ (te toetsen hypothese): Het gemiddelde van $\underline{x}$ bezit de waarde μ; μ is hierin een gegeven getal, b.v. 0.

Toetsingsgrootheid :

$$\underline{t} = (\underline{m} - \mu)/\underline{s}'$$

waarin de bij de steekproef behorende waarden van $\underline{m}$ en $\underline{s}'$ gegeven worden door $m = \frac{x_1 + x_2 + \cdots + x_n}{n}$ en $s' = \sqrt{\frac{(x_1-m)^2 + \ldots (x_n-m)^2}{n-1}}$ is.

Indien $\mathcal{H}_0$ juist is, zullen dicht bij 0 gelegen waarden vaker voorkomen dan ver van 0 gelegen waarden. Is echter het gemiddelde van $\underline{x}$ verschillend van μ, dan zullen verder van 0 af liggende waarden vaker voorkomen dan indien $\mathcal{H}_0$ juist is. Als kritieke zône Z kiest men daarom voor tweezijdige toetsing een gebied van de vorm

$$|\underline{t}| \geqq t_0$$

en voor éénzijdige toetsing

linker-toetsing	rechter-toetsing
$\underline{t} \leqq -t_1$	$\underline{t} \geqq t_2$

De waarden t_0, t_1 en t_2 zijn getabelleerd voor verschillende waarden van de onbetrouwbaarheidsdrempel α.

Litteratuur:

M.G. Kendall, The Advanced Theory of Statistics, London 1946, Vol.II, p.98-102; tabellen in deel I, p.440-41. Opmerking: Het bij de tabellen vermelde aantal vrijheidsgraden (aangegeven door ν) is gelijk aan $n-1$

A.M. Mood, Introduction to the theory of Statistics, London 1950, p.425.

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

2) d.w.z., dat de kans, dat $\underline{x} \leqq x$ is, gegeven wordt door:

$$P[\underline{x} \leq x] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\frac{(u-\mu)^2}{\sigma^2}} du$$

S47(M10)

Symmetrietoets.[1]

Hypothese $\mathcal{H}_0$: de waarnemingen $x_1, \ldots, x_{n'}$ zijn afkomstig van onafhankelijk verdeelde stochastische grootheden, die alle symmetrisch ten opzichte van 0 verdeeld zijn.[2] Van deze toets bestaan meerdere versies $\mathcal{T}_1, \ldots, \mathcal{T}_6$. We bespreken eerst $\mathcal{T}_1$ en $\mathcal{T}_2$.

Toetsingsgrootheden. Deze worden als volgt uit $x_1, \ldots, x_{n'}$ afgeleid:

1e de waarnemingen, die gelijk aan 0 zijn worden weggelaten. Stel er blijven over: $x_1, \ldots, x_n$

2e Hieruit worden de positieve waarnemingen gezocht. Stel dit zijn $x_1, \ldots, x_{n_1}$. dus n_1 in aantal.

3e De overblijvende negatieve waarnemingen worden van teken veranderd, zodat zij ook positief worden. Stel dit zijn dan $y_1, \ldots, y_{n_2}$

4e De grootheden $x_1, \ldots, x_{n_1}$ en $y_1, \ldots, y_{n_2}$ worden, door elkaar, in afdalende grootte-volgorde opgeschreven. Stel dit geeft: $w_1, \ldots, w_n$. (Komen er gelijken voor, dan worden deze in willekeurige volgorde geplaatst).

5e Uit de waarden $w_1, \ldots, w_n$ wordt een groep A van grote waarden afgezonderd op de volgende wijze: Men telt de waarden $w_1, \ldots, w_n$ op de rij af, tot men er $\frac{n}{2}$ (of als n oneven is $\frac{n+1}{2}$) heeft gepasseerd. Is de volgende waarde ongelijk aan de het laatst opgetelde, dan houdt men op; anders gaat men door tot men aan een eerste ongelijke toekomt, die dan niet bij de groep A genomen wordt. A bevat dus minstens de helft der waarden $w_1, \ldots, w_n$; noemen wij het restant B, dan is bovendien iedere in A bevatte waarde groter dan iedere in B bevatte waarde. Het aantal waarden in A noemen wij k

6e Het aantal waarden van $x_1, \ldots, x_n$, die in A voorkomen, noemen wij u.

De toetsingsgrootheden zijn n en u, k is een hulpgrootheid.

V.B. 22 waarden x_i: 7,4/6,3/3,6/3,5/3,4/2,9/2,5/1,1/0/0/-1,3/-2,5/-3,2/ -4,6/-4,6/-4,6/-6,?/-7,0/-7,9/-8,0/-8,7.

$\therefore \quad k = 11, \quad n_1 = 8$

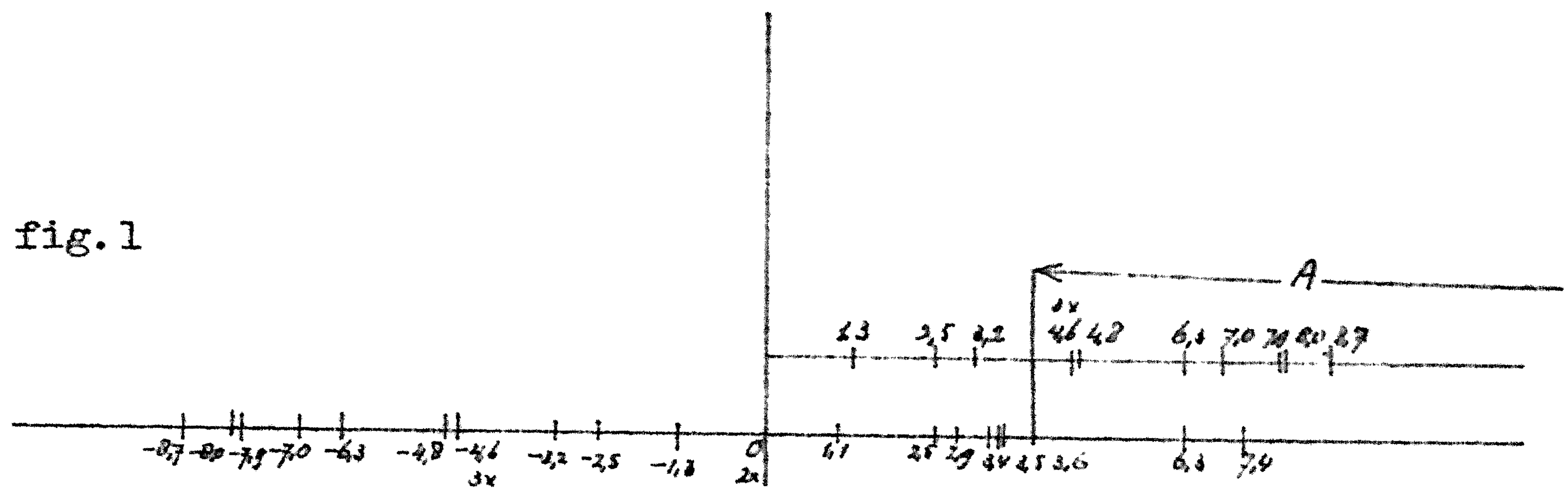
$u = 2$

---

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of exactheid.

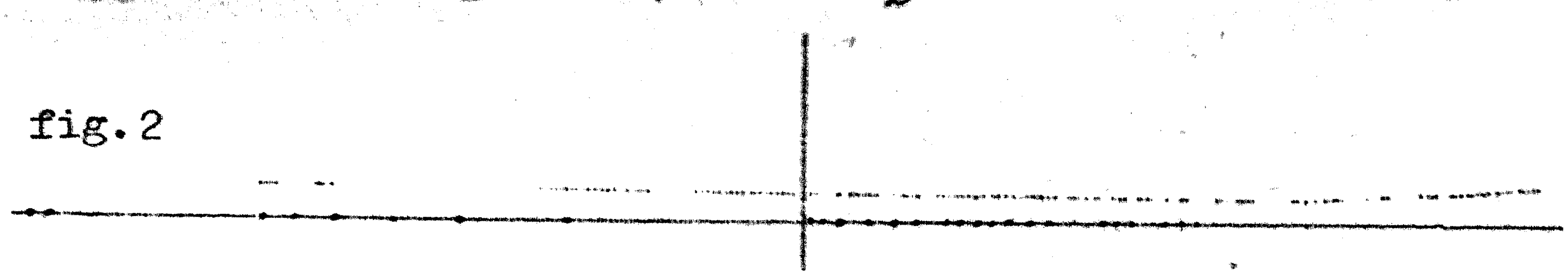
2) Zetten wij hier in plaats van 0, dan geldt voor

fig. 1



<u>Kritieke zônes.</u> Waarden van n_1, die direct bij 0 of dicht bij n liggen zullen, als H_0 juist is weinig, maar als H_0 onjuist is vaker, voorkomen. Grote resp. kleine waarden van u zullen eveneens, als H_0 juist is weinig voorkomen. Hierop berust de keuze van de bij T_1 en T_2 behorende kritieke zônes Z_1 resp. Z_2
Z_1 bevat grote en kleine waarden van n_1, en grote en kleine waarden van u, terwijl Z_2 bij grote waarden van n_1 in hoofdzaak grote waarden van u en bij kleine waarden van n_1 in hoofdzaak kleine waarden van u bevat. T_1 leidt bij voldoende grote n vrijwel steeds tot verwerping als de hypothese niet is vervuld. T_2 leidt echter alleen tot verwerping van H_0 als er veel <u>positieve</u> (resp. negatieve) waarden zijn, die <u>verder</u> van 0 verwijderd liggen dan de <u>negatieve</u> (resp. positieve). In sommige gevallen is het juist van belang om deze laatste afwijkingen van H_0 te ontdekken. In dat geval gebruikt men T_2 liever dan T_1. In fig. 2 is een schematisch voorbeeld gegeven van een serie waarnemingen, waarbij het aantal positieve groter is dan het aantal negatieve, terwijl deze positieve dichter bij 0 liggen dan de negatieve, zodat T_2 niet tot verwerping leidt.

fig. 2



Van T_1 en T_2 bestaan ook éénzijdige versies, waarvan de beschrijving te ver zou voeren.

<u>T_3 en T_4</u>.

<u>Toetsingsgrootheden</u>.

1e, 2e en 3e als boven. Eerste toetsingsgrootheid: n_1
4e: op $x_1, \ldots, x_{m_1}$ en $y_1, \ldots, y_{m_2}$ wordt de toets van Wilcoxon toegepast (vgl. S47(M8)). De toetsingsgrootheden zijn n_1 en de U van deze toets van Wilcoxon.

<u>Kritieke zônes</u>. Overwegingen analoog aan die voor T_1 en T_2 (met U in plaats van u) leiden tot analoge kritieke zônes Z_3 en Z_4, behorend bij T_3 en T_4.

Opmerkingen: $\mathcal{T}_1$ en $\mathcal{T}_2$ zijn bijzonder geschikt voor een niet te groot aantal waarnemingen. Zij gelden ook voor niet continue verdelingen. $\mathcal{T}_3$ en $\mathcal{T}_4$ zijn alleen geschikt, als er geen of weinig paren (x_i, y_j) met $x_i = y_j$ zijn. Voor grote aantallen zijn $\mathcal{T}_3$ en $\mathcal{T}_4$ geschikter dan $\mathcal{T}_1$ en $\mathcal{T}_2$. Er is ook een versie voor grote aantallen ($\mathcal{T}_5$ en $\mathcal{T}_6$), die geheel analoog is met $\mathcal{T}_3$ en $\mathcal{T}_4$ met dien verstande, dat $\underline{u}$ in plaats van $\underline{\mathcal{U}}$ wordt gebruikt.

Litteratuur:

J. Hemelrijk, A family of parameterfree tests for symmetry with respect to a given point, I,II. Proceedings van de Kon.Ned.Ak.v.Wet. 53,(1950),p.945-955. Indagationes Mathematicae 12 (1950),p.340-350.

---, Symmetrietoetsen, Diss., Den Haag 1950, Excelsio[r]

Hoek - toets voor onafhankelijkheid [1])

Deze methode dient om de onafhankelijkheid te toetsen van 2 continuë stochastische grootheden. De toets is geldig zonder verdere onderstellingen over de vorm van de waarschijnlijkheidsverdelingen. Zij berust op n paren waarnemingen $(x_1, y_1), (x_2, y_2), \ldots\ldots (x_n, y_n)$. Deze getallenparen worden als punten in een vlak getekend.

We trekken nu 2 lijnen AB (horizontaal) en CD (verticaal), zo, dat onder en boven AB evenveel punten liggen, en evenzo evenveel punten links als rechts van CD. Van de vier kwadranten, die ze ontstaan, worden de kwadranten rechts boven en links onder van een + teken en de beide andere van een - teken voorzien.

(zie fig. 1.)

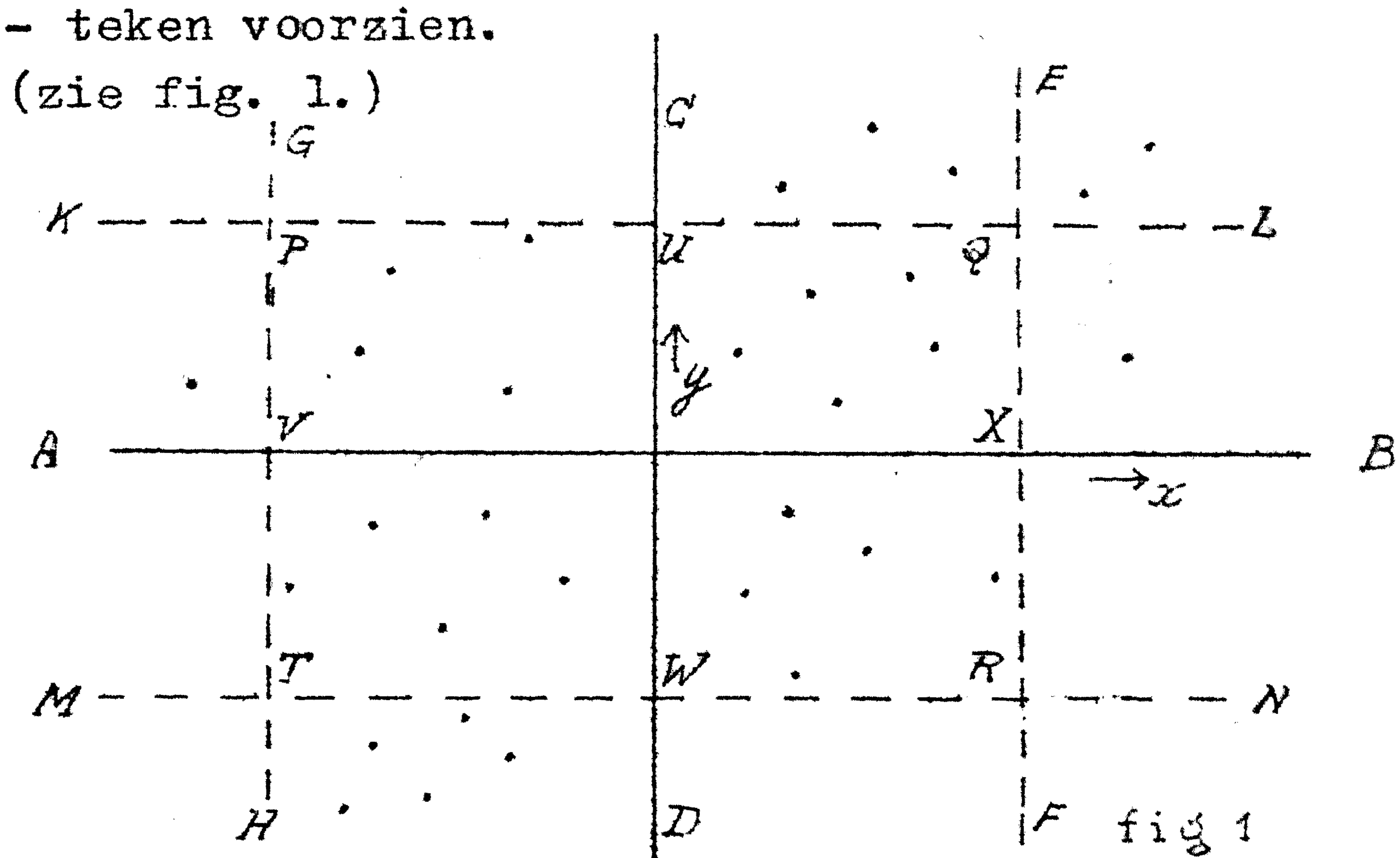


fig 1

We tellen nu de punten af, beginnende aan de rechterkant van het diagram en naar links gaande, volgens afnemende x, tot we, om een volgend punt te vinden, de horizontale lijn AB moeten passeren. Tussen dit volgende punt en het laatst getelde punt trekken we de verticale lijn EF. Van links af tellend, volgens opklimmende x vinden we op analoge wijze de lijn GH.

Van bovenaf beginnend, tellen we de punten volgens afnemende y tot we de lijn CD moeten passeren. Daar trekken we de rechte KL. Aan de onderzijde bepalen we op analoge wijze de rechte MN.

---

[1]) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

Er zijn nu een aantal punten geïsoleerd buiten de op deze wijze verkregen vierhoek $PQRT$ (zie fig. 1)
We bepalen nu de zgn. kwadranten som $\underset{\sim}{S}$ van deze aantallen geïsoleerde punten. Onder deze <u>kwadrantensom</u> verstaan we de som van het aantal punten boven KL , onder MN, links van GH en rechts van EF in het eerste en derde kwadrant verminderd met het aantal punten buiten deze lijnen in het tweede en vierde kwadrant. De punten in de hoeken EQL, MTH FRN en GPK worden dus dubbel geteld.

<u>Vb.</u> in fig. 1. wordt :

Hoek	CUL	$+5$
,,	EXB	$+3$
,,	GVA	-1
,,	MWD	$+5$
		$\underset{\sim}{S} = +13$

De verdelingsfunctie van $\underset{\sim}{S}$ is bekend en vrijwel onafhankelijk van het aantal waarnemingen. Is aan de onafhankelijkheid voldaan, dan kan men betrekkelijk kleine waarden van $/\underset{\sim}{S}/$ verwachten. Is er afhankelijkheid, dan is er meer kans op grote waarden van $/\underset{\sim}{S}/$. Daarom neemt men kritieke zônes van de vorm $/\underset{\sim}{S}/ \geqq /S_o/$. Voor een onbetrouwbaarheidsdrempel $\alpha = 0,05$ is $S_o = 11$, voor $\alpha = 0,01$ is $S_o = 19$.
Voor het geval dat één of meer lijnen door één of meer waargenomen punten gaan wordt een kleine verandering in de werkwijze aangebracht. Men zie hiervoor de litteratuur.

<u>Litteratuur</u>

P.S. Olmstead, John W. Tukey
A Corner Test for Association
Ann. of Math. Stat. 1947,
Vol. 18, p. 495-513

A.M. Mood,
Introduction to the Theory of Statistics, 1950
p. 410.

S 47 (M 15).

Toets voor onafhankelijkheid van 2 grootheden met behulp van de methode der Dubbele Dichotomie ')

Deze methode dient o.a. om de onafhankelijkheid te toetsen van twee continue grootheden; (x_1, y_1), (x_2, y_2) --- (x_n, y_n) zijn waarnemingsparen van de stochastische grootheden $\underline{x}$ en $\underline{y}$. Deze getallenparen worden als punten in een vlak getekend.

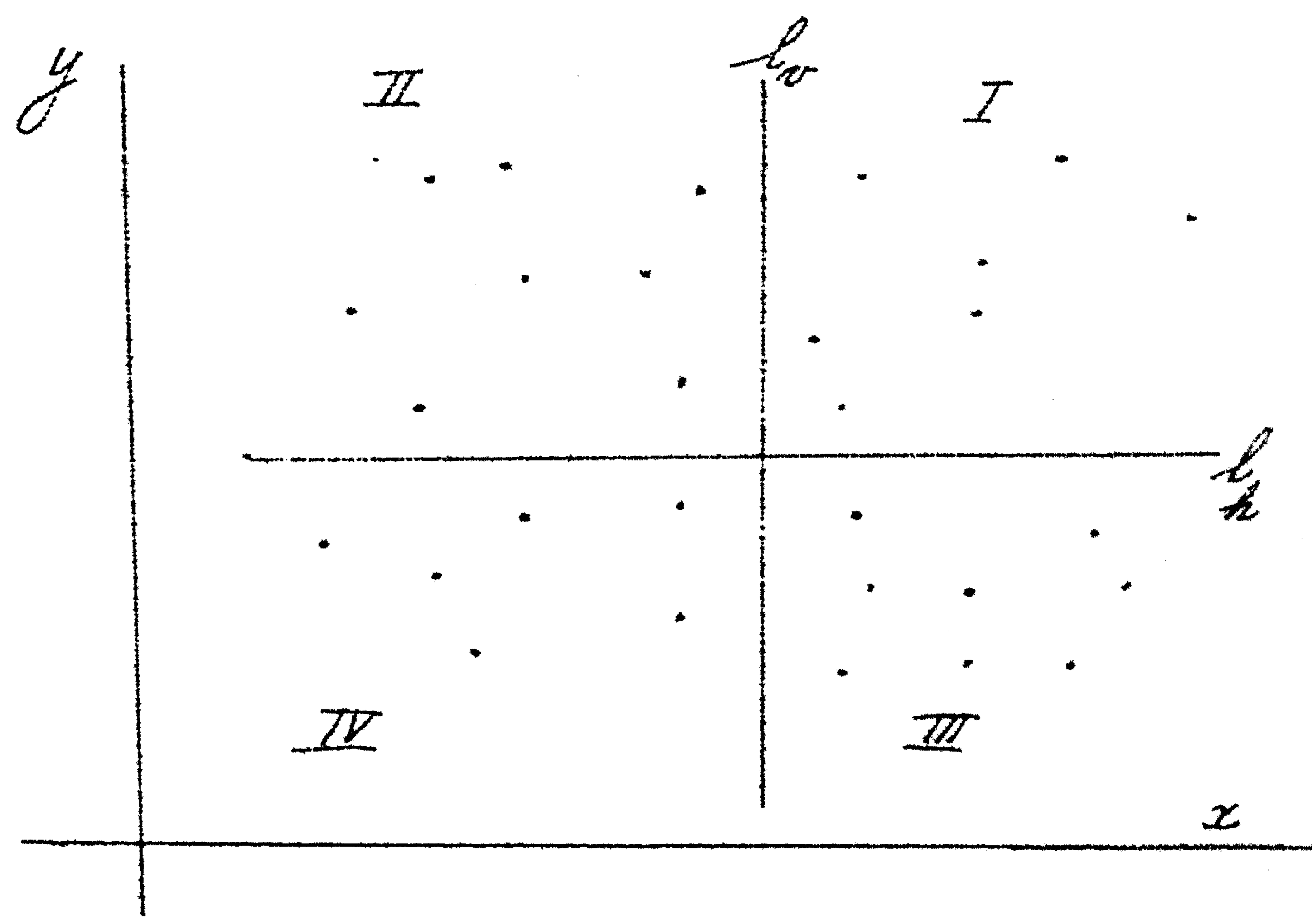


fig 1

We verdelen de puntenwolk door een verticale en een horizontale rechte (l_v resp. l_h) zo, dat links en rechts van l_v, en ook boven en onder l_h evenveel punten liggen. Het vlak is nu verdeeld in vier gebieden. Is er geen afhankelijkheid dan verwachten we dat in alle vier gebieden ongeveer evenveel punten zullen liggen. Een exacte afleiding, (geschikt voor kleine aantallen waarnemingen) verkrijgen we als volgt:

We beschouwen de puntne in ons vlak als waarnemingen van elementen die ieder tegelijk twee kenmerken bezitten; A of $\overline{A}$ '') (rechts resp. links van verticale streep) en B of $\overline{B}$ (boven resp. onder horizontale streep). Er ontstaan dus vier groepen elementen met de kenmerken

AB	$A\overline{B}$	$\overline{A}B$	$\overline{A}\,\overline{B}$
vak I	vak III	vak II	vak IV

We tekenen dit als volgt in een 2 x 2 tabel: (Zie blz 2)

') Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

'') A betekent: "non-A"

	A	$\bar{A}$	
B	a	$n-a$	n
$\bar{B}$	b	$m-b$	m
	r	s	N

Het totale aantal elementen is N. Het aantal elementen dat kenmerk B bezit, is n, het aantal dat kenmerk $\bar{B}$ bezit, is m. Hetzelfde geldt voor de kenmerken A en $\bar{A}$ met de aantallen r resp. s. In het bovenstaande voorbeeld werd de puntenwolk zo verdeeld, dat $n = m = r = s = \frac{1}{2} N$

Dit principe wordt vaak toegepast, maar is niet noodzakelijk. De toets kan ook worden gebruikt met andere waarden van m, n, r, s mits deze steeds op grond van de proefopstelling, niet naar aanleiding van de uitkomst van het experiment worden gekozen.

Onafhankelijkheid van de kenmerken wil zeggen, dat de kans dat een element het kenmerk B bezit even groot is, indien het tevens het kenmerk A bezit, als het tevens het kenmerk Ā bezit. Wij toetsen nu deze hypothese van onafhankelijkheid.

Daar de rand-totalen gegeven zijn, is er nog één vrijheidsgraad, d.w.z. indien één van de waarden in één der vier binnenvakjes bepaald is, volgen daaruit de andere waarden.
Wij beschouwen het getal a in het linker-boven vakje. Dit kan verschillende waarden aannemen en is dus een stochastische grootheid $\underline{a}$, waarvan de waarschijnlijkheidsverdeling, onder aanname van de te toetsen hypothese, een z.g. hypergeometrische verdeling is, die gegeven wordt door

$$P[\underline{a} = a] = \frac{\binom{n}{a}\binom{m}{r-a}}{\binom{m+n}{r}} = \frac{n!\,m!\,r!\,s!}{a!\,b!\,(n-a)!\,(m-b)!\,N!}$$

Het gemiddelde van deze verdeling is $\mathcal{E}\underline{a} = \frac{nr}{N}$ en het spreidings kwadraat is $\sigma^2_{\underline{a}} = \frac{n\,m\,r\,s}{N^2(N-1)}$

Als voorbeeld nemen wij $N=20$, $r=s=n=m=10$

De waarschijnlijkheidsverdeling van $\underline{a}$ is nu in fig. 2 en tabel I weergegeven.

getallen v.b.

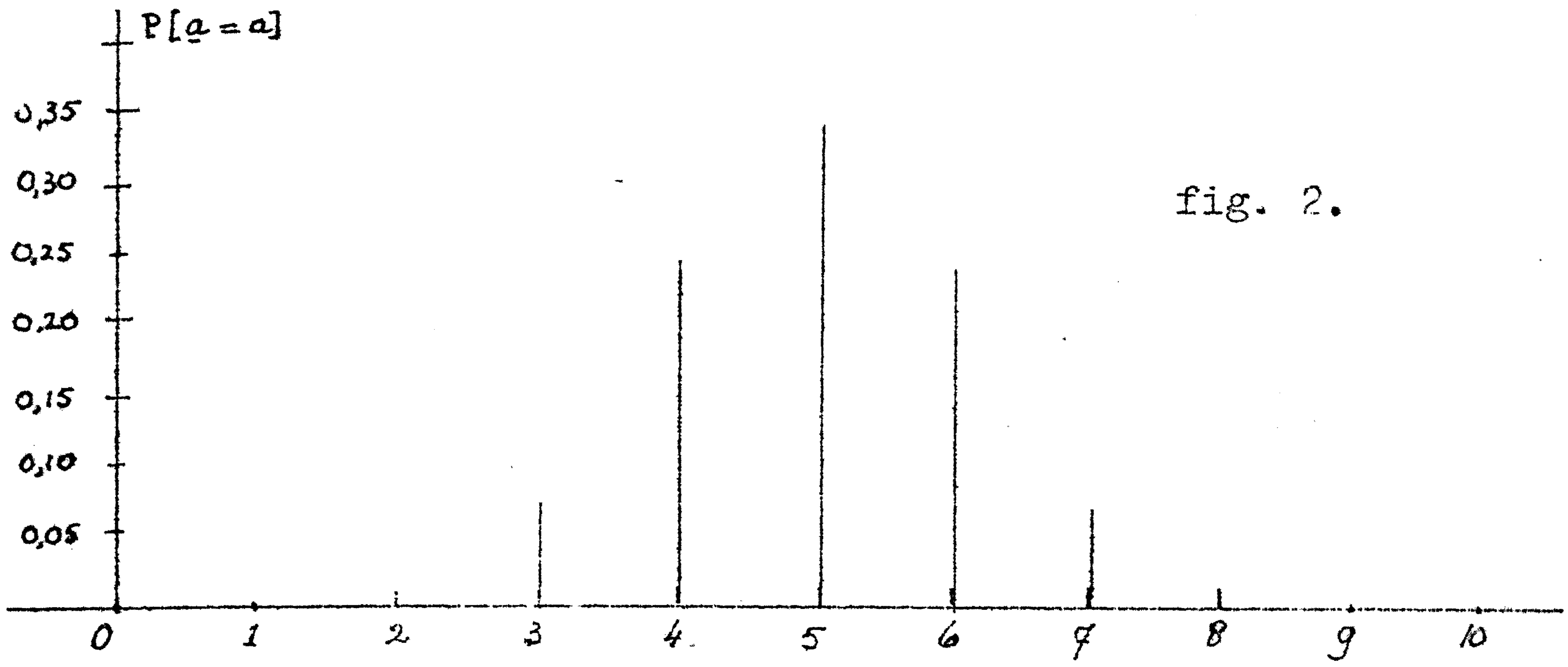


fig. 2.

mogelijkheden:

$\begin{array}{c|c} 0 & 10 \\ \hline 10 & 0 \end{array}$ $\begin{array}{c|c} 1 & 9 \\ \hline 9 & 1 \end{array}$ $\begin{array}{c|c} 2 & 8 \\ \hline 8 & 2 \end{array}$ $\begin{array}{c|c} 3 & 7 \\ \hline 7 & 3 \end{array}$ $\begin{array}{c|c} 4 & 6 \\ \hline 6 & 4 \end{array}$ $\begin{array}{c|c} 5 & 5 \\ \hline 5 & 5 \end{array}$ $\begin{array}{c|c} 6 & 4 \\ \hline 4 & 6 \end{array}$ $\begin{array}{c|c} 7 & 3 \\ \hline 3 & 7 \end{array}$ $\begin{array}{c|c} 8 & 2 \\ \hline 2 & 8 \end{array}$ $\begin{array}{c|c} 9 & 1 \\ \hline 1 & 9 \end{array}$ $\begin{array}{c|c} 10 & 0 \\ \hline 0 & 10 \end{array}$

Tabel I

a	$P[\underline{a} = a]$
0	0,00001
1	0,00054
2	0,01096
3	0,07794
4	0,24356
5	0,34372
6	0,24356
7	0,07794
8	0,01096
9	0,00054
10	0,00001

Stel nu dat de volgende waarden gevonden zijn:

	A	$\bar{A}$	
B	2	8	10
$\bar{B}$	8	2	10
	10	10	20

We berekenen

$$P[\underline{a} = 2] = \frac{(10!)^4}{20!(8!)^2(2!)^2} = 0{,}01096$$

en zoeken verder alle waarden van a bijeen, waarvoor $P[\underline{a} = a] \leq P[\underline{a} = 2]$ is

Vervolgens tellen we de daarbij behorende waarschijnlijkheden $P[\underline{a} = a]$ op. (In ons voorbeeld dus voor $\underline{a}$ = 0,1,2,8,9,10). Deze som is per definitie de overschrijdingskans, behorende bij het gevonden resultaat, (in ons voorbeeld 0,02302) [1])

[1]) De bij deze definitie van de overschrijdingskans behorende kritieke zône bestaat uit alle waarden voor a met overschrijdingskans $\leq \alpha$,in ons voorbeeld 0,1,2,8,9 en 10.

Is deze $< \alpha$ (α is de onbetrouwbaarheidsdrempel), dan wordt de hypothese van onafhankelijkheid verworpen (In ons voorbeeld treedt dus, als men $\alpha = 0{,}05$ neemt, verwerping op).
Voor grote aantallen waarnemingen maken we gebruik van het feit dat $\frac{a - \mathcal{E}a}{\sigma_a}$ bij benadering normaal verdeeld is, met gemiddelde 0 en spreiding 1.

Continuïteitscorrectie

Deze behoeft alleen bij kleine aantallen toegepast te worden en bestaat daarin, dat men alle getallen $a, s, n-a, m-s$ met $\frac{1}{2}$ vermeerdert of vermindert, zodanig dat de randtotalen dezelfde blijven $|a-\bar{a}|$ kleiner wordt.

Literatuur:

M.G.Kendall, The advanced theory of Statistics, Vol.I, (London 1947), p.303;

E.S.Pearson, The choice of statistical tests illustrated on the interpretation of data classed in a 2 x 2 table, Biometrika 34 (1949) p.139-167.

Statistische Afdeling
S47 (M19)
door J.Hemelrijk

Parametervrije toets voor twee meerdimensionale steekproeven. 1)

Wij beschrijven de toets voor twee 2-dimensionale steekproeven (x_1', y_1'), ...1, (x_n', y_n') en (x_1'', y_1''), ..., (x_m'', y_m''). De te toetsen hypothese H_0 luidt, dat deze beide steekproeven uit dezelfde tweedimensionale collectie ("populatie") afkomstig zijn.

Ieder der waargenomen getallenparen kan men uitzetten als een punt in een vlak met coördinaten x en y. De voorwaarde waaronder de toets geldt is, dat voor iedere lijn L van dit vlak de kans, dat een punt op L zal worden waargenomen, gelijk aan nul is. Dit is dus een continuïteitsvoorwaarde van een type, dat vaak bij parametervrije toetsingsmethoden noodzakelijk is. Voor de practijk houdt deze eis b.v. in, dat de waarnemingen niet gegroepeerd moeten zijn en dat de verdelingen niet discreet zijn. Kleine afwijkingen van deze voorwaarden oefenen weinig invloed uit.

Zetten wij de punten (x_i', y_i') en (x_j'', y_j'') uit in een vlak, dan verkrijgen wij een "puntenwolk" (zie fig. 1). De punten (x_i', y_i') geven wij daarbij aan met kruisjes en de punten (x_j'', y_j'') met kringetjes.

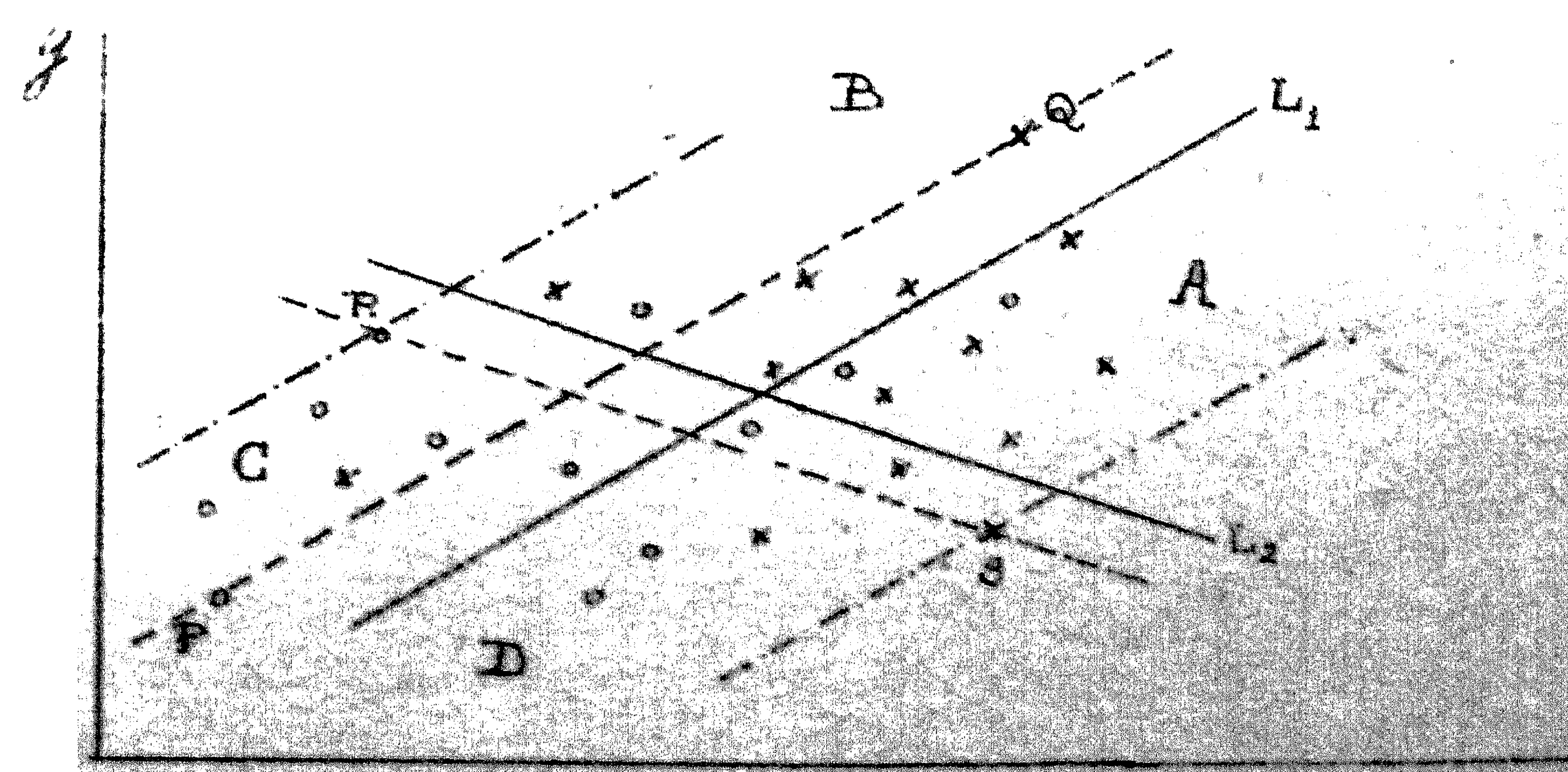


fig. 1 ($n = 14$; $m = 12$)

1) Dit memorandum dient slechts ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

Wij verdelen nu de puntenwolk in een aantal delen, zonder daarbij onderscheid te maken tussen de punten en kringetjes. Dit kan op verschillende wijzen geschieden.

Men kan b.v. de puntenwolk in 4 delen verdelen door het aanbrengen van een horizontale en een verticale lijn, die beide zodanig zijn aangebracht, dat ter weerszijden van ieder evenveel punten liggen. Indien x en y vergelijkbare en in de zelfde eenheden gemeten grootheden zijn kan men ook de volgende methode toepassen 1):
Wij zoeken eerst het puntenpaar (P, Q) met de grootste onderlinge afstand. De lijn PQ verschuiven wij, de richting constant houdende, tot er ter weerszijden evenveel punten liggen. De zo verkregen lijn noemen wij L_1. Vervolgens verschuiven wij hem verder tot alle punten aan één zijde van de lijn liggen. Het laatst gepasseerde punt noemen wij S. Verschuiving naar de andere kant wijst een overeenkomstig punt R aan de andere zijde van de puntenwolk aan. De lijn RS verschuiven wij nu, de richting weer constant houdende, tot ter weerszijden evenveel punten liggen. Deze lijn is in fig. 1 aangegeven met L_2.

De puntenwolk wordt nu door de lijnen L_1 en L_2 in 4 delen verdeeld, die we aangeven met A, B, C en D. Aan ieder punt kennen wij nu twee kenmerken toe

1° diegene der letters A, ... , D, die het vak aangeeft, waar het punt in ligt;

2° het kenmerk o of x naar gelang van de steekproef, waar het punt toe behoort.

Is nu de te toetsen hypothese H_0 juist, d.w.z. stellen alle punten waarnemingen uit éénzelfde tweedimensionale collectie voor, dan is de verdeling van de kenmerken o en x over de punten van de puntenwolk stochastisch onafhankelijk van die van de kenmerken A, ... , D. Deze onafhankelijkheid kan voor niet te kleine n getoetst worden met behulp van de χ^2-toets voor een 2 x 4- tabel (vgl. b.v. M.G.Kendall, Advanced theory of statistics, I, p. 299 en 300). De bovengenoemde voorwaarde, dat de verdeling van de puntenwolk in

---

1) Deze methode is totaal willekeurig, indien x en y niet in dezelfde eenheden gemeten zijn daar de constructie niet invariant is tegen schaalverandering van x en y. Is echter aan de voor x en y gestelde voorwaarde wel voldaan, dan bezit de hier beschreven methode voor bepaalde alternatieve hypothesen, een groter onderscheidingsvermogen dan de methode met verdeling door een horizontale en een verticale lijn.

een aantal delen moet plaats vinden zonder de kruisjes van de kringetjes te onderscheiden, is nodig om de stochastische onafhankelijkheid van de kenmerken A, ... , D t.o.v. de kenmerken o en x niet te verstoren.

Daar de χ^2-toets een asymptotische toets is, d.w.z. slechts voor grote n met een voldoende kleine benaderings-fout behept is, heeft men vrij veel punten nodig. Als minimum mag men wel $n+m=40$ stellen (of daaromtrent). Indien men nog veel meer punten heeft kan men de puntenwolk op, aan de bovenstaande wijze verwante manieren, in meer dan 4 delen verdelen en de zelfde χ^2-toets toepassen. Verdeelt men de puntenwolk in ν delen, dan wordt dit dus een χ^2-toets voor een $2 \times \nu$-tabel

Litteratuur. Een publicatie over deze toets is in voorbereiding.