

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 19 (M 3)

Statistische bepaling van het lineair verband
tussen twee fysische grootheden

J. Hemelrijk

(Nederlands Tijdschrift voor Natuurkunde,
17(1951), p 147-158)



1951

Statistische bepaling van het lineair verband tussen twee fysische grootheden

door J. Hemelrijk

Summary:

A statistical method is developed for determining a joint confidence region for the coefficients of a linear equation, describing the connection between two quantities, both of which are subject to error. The conditions imposed on the errors are less stringent than usual, in particular no normality is assumed. The necessary number of observed points is small; confining ourselves to confidence levels $\leq 0,05$, 7 or more points enable us to find a confidence interval for the slope of the line, while 8 or more are sufficient to construct a joint confidence region for the slope and the intercept. The methods used are partly ranking methods.

1. Inleiding.

In natuur- en scheikunde treedt vaak de volgende situatie op: tussen twee grootheden ξ en η bestaat, naar men op bepaalde gronden aanneemt, een lineair verband, dat we zullen aangeven door

$$(1) \quad a\xi + \beta\eta + \gamma = 0$$

met onbekende a , β en γ . Ter bepaling van a , β en γ worden een aantal (noem dit n) waarnemingen van (niet noodzakelijk alle onderling verschillende) punten $Q_i \equiv (\xi_i, \eta_i)$, $i = 1, \dots, n$, van deze lijn, die we L zullen noemen, verricht. Deze waarnemingen zijn onderhevig aan een meetfout, zowel in de ξ - als in de η -richting, zodat we voor iedere i in plaats van het punt Q_i een punt $P_i \equiv (x_i, y_i)$ vinden met:

$$(2) \quad x_i = \xi_i + u_i \text{ en } y_i = \eta_i + v_i.$$

Twee vragen worden nu gewoonlijk gesteld:

- 1e. Hoe kan men naar aanleiding van de gevonden punten P_1, P_2, \dots, P_n een lijn trekken, waarvan men kan zeggen, dat hij goed bij het gevonden resultaat aansluit;
- 2e. Hoe kan men een gebied in het (ξ, η) -vlak vinden, waarvan men, met slechts een kleine kans dat deze uitspraak fout is, kan zeggen, dat L er in ligt.

Een antwoord op de eerste vraag is o.a. de bekende methode der kleinste kwadraten; (zie b.v. H. Cramér (1946) p. 548 en voor een overzicht van de methode J. Hemelrijk (1949b)). Het gebruiken van een op deze wijze gevonden lijn, alsof het de „ware”

is, heeft echter het nadeel, dat dit in werkelijkheid zelden of nooit precies het geval zal zijn, zodat men in feite practisch altijd met een verkeerde lijn werkt. De methode der kleinste kwadraten geeft, behalve in een zeer speciaal geval, geen oplossing van het tweede probleem, zodat men niet weet, welke lijnen, die van de gevondene verschillen, óók wel de „ware” lijn zouden kunnen zijn.

In dit artikel zal daarom een oplossing van het tweede vraagstuk worden gegeven met behulp van een methode, verwant met de in Engeland en Amerika ontwikkelde moderne methoden, waarbij uitsluitend op de volgorde naar grootte van de uitkomsten gelet wordt en niet op die grootte zelf (order statistics, ranking methods). Dit heeft het voordeel, dat over de meetfouten zeer weinig ondersteld behoeft te worden, i.h.b. niet, dat deze normaal (d.i. volgens de „fouten-kromme van Gauss”) verdeeld zijn. Deze onderstelling wordt bij de enige vroegere oplossing van het probleem, door A. Wald (1940), wel gemaakt en is zeer vaak niet vervuld; in het bijzonder zijn de afrondingsfouten bij het schatten van schaal-delen meestal verre van normaal verdeeld.

Een tweede methode, waarbij deze onderstelling niet wordt gemaakt, is onlangs ontwikkeld door H. Theil (1950).

2. Principes en begrippen van de gebruikte methode.

Ter oriëntering van lezers, die met de moderne methoden van de mathematische statistiek weinig vertrouwd zijn, volgt hier een korte uiteenzetting van de in dit artikel gebruikte beginselen:

Wij zullen een reële variabele x een *stochastische variabele* noemen, als x een waarschijnlijkheidsverdeling bezit, d.w.z. als voor ieder reëel getal x_0 de kans bepaald is, dat $x \leq x_0$ is, dus dat een waarneming van x kleiner of gelijk aan x_0 zal uitvallen. *) De kans dat $x \leq x_0$ is, geven we aan met $P[x \leq x_0]$. Deze kans is een functie van x_0 , die de verdelingsfunctie van x wordt genoemd, en behalve door de reeds vermelde notatie ook met $F(x_0)$ wordt aangeduid. Is $F(x_0)$ een continue functie van x_0 dan noemt men de verdeling van x continu (of ook: x is continu verdeeld). Dit houdt in, dat voor iedere x_0 de kans, $P[x = x_0]$, dat x de waarde x_0 aan zal nemen, gelijk aan 0 is, hetgeen echter niet betekent dat $x = x_0$ onmogelijk is. Een voorbeeld van een continue verdeling is bovengenoemde fouten-verdeling van Gauss.

Twee stochastische variabelen x en y worden *stochastisch onafhankelijk* van elkaar genoemd, als voor iedere x_0 en y_0 geldt: $P[x \leq x_0 \text{ en } y \leq y_0] = P[x \leq x_0] \cdot P[y \leq y_0]$.

Uit de theorie van J. Neyman en E. S. Pearson voor het toetsen van statistische hypothesen (zie b.v. J. Neyman (1942)) zullen de volgende begrippen worden gebruikt:

Zij x een stochastische variabele, waarvan de verdeling bekend

*) Stochastische variabelen worden met schreefloze letter aangeduid.

is op een in de verdelingsfunctie voorkomende onbekende parameter θ na, waarvan men aanneemt, dat zij een ondubbelzinnig bepaalde, zij het onbekende, „ware” waarde bezit en zij x_1, \dots, x_n een n -tal onafhankelijke waarnemingen van deze variabele. Wordt nu gevraagd, aan de hand van deze serie waarnemingen de hypothese, dat θ een gegeven waarde θ_0 heeft, te toetsen, dan stelt men eerst de een of andere van de waarde θ_0 afhankelijke *kritieke situatie* voor de waarnemingen x_1, \dots, x_n vast, (b.v. hieruit bestaande dat het gemiddelde van x_1, \dots, x_n groter is dan een zekere van θ_0 en de verdelingsfunctie van x afhankende waarde) en indien deze kritieke situatie zich voordoet (en dus bijv. inderdaad het gemiddelde van x_1, \dots, x_n groter is dan deze uit de verdelingsfunctie verkregen waarde) dan verwerpt men de hypothese $\theta = \theta_0$.

Bij de keuze van een dergelijke kritieke situatie kan men zich door velerlei principes laten leiden, maar als eerste eis stelt men gewoonlijk dat men de waarschijnlijkheid *) kan berekenen, dat, als θ_0 de ware waarde van θ is, de genoemde kritieke situatie zich zal voordoen. Het is niet strikt nodig dat men deze berekening voor iedere waarde van θ kan uitvoeren, maar wel moet dit het geval zijn voor de te toetsen waarde θ_0 .

Deze waarschijnlijkheid, door Neyman en Pearson de „size” van de „critical region” genoemd, zullen we met p aanduiden en het *risico* of de *onbetrouwbaarheidsdrempel* van de toets noemen. Dit betekent, dat, als men volgens dezelfde methode en met dezelfde p een groot aantal experimenten uitvoert, in ongeveer een fractie p van deze gevallen de getoetste waarde van θ (die bij ieder van deze experimenten verschillend kan zijn), indien dit de ware waarde was, verworpen zou worden. Men noemt daarom p de *kans op ten onrechte verwerpen van de hypothese*. Men zal van p dus eisen, dat hij klein is; gewoonlijk neemt men genoegen met $p = 0,01$ of zelfs $0,05$. Hoe kleiner men p maakt (door verandering van de kritieke situatie) hoe kleiner echter ook de kans wordt, dat men de hypothese *terecht* zal verwerpen. Hierop zullen we hier niet nader ingaan. De opmerking dient slechts, om aan te geven, waarom men p gewoonlijk nog vrij groot neemt.

Indien men nu de kritieke situatie zo bepaalt, dat deze behalve van de waarden x_1, \dots, x_n slechts van θ_0 en van p afhankelijk is, noemt men de verzameling der waarden θ_0 , die, bij een bepaalde serie x_1, \dots, x_n en bij één zelfde waarde van p , niet voor verwerping in aanmerking komen (dus waarbij zich de kritieke situatie niet voordoet) een *betrouwbaarheidsgebied* voor θ . Dit betrouwbaarheidsgebied, dat voor één onbekende parameter gewoonlijk de vorm van een interval aanneemt, is afhankelijk van x_1, \dots, x_n en van p . Daar echter de getallen x_1, \dots, x_n op de collectie

*) Of een voldoende kleine bovengrens daarvan.

van alle mogelijke steekproeven stochastische variabelen zijn, alle met dezelfde verdeling, hangt ook het betrouwbaarheidsgebied van het toeval af. Uit de definitie blijkt, dat de kans, dat dit gebied zo zal uitvallen, dat het de ware parameterwaarde bevat gelijk is aan $1-p$. De bewering dat θ in het betrouwbaarheidsgebied ligt is dus, indien men deze methode herhaaldelijk toepast, in ongeveer een fractie p van deze gevallen fout. We noemen daarom p het *risico* verbonden aan het betrouwbaarheidsgebied of de *onbetrouwbaarheidsdrempel* daarvan.

Deze beschouwingen zijn niet gebonden aan één onbekende parameter, maar gaan onveranderd op voor meerdere.

In het volgende zal een betrouwbaarheidsgebied voor de gezochte lijn L geconstrueerd worden; dit is, analoog met het bovenstaande, een gebied afhankelijk van de uitgevoerde waarnemingen, dat de kans $1-p$ bezit om L geheel te bevatten. Daarbij doet zich nog een kleine complicatie voor, die daarin bestaat, dat bovendien bepaalde lijnen, die in het betrouwbaarheidsgebied liggen, toch voor verwerping in aanmerking komen. Dit zal echter ter plaatse nader besproken worden.

In § 3 zullen eerst een aantal onderstellingen besproken worden over de meetfouten, waarop de daaropvolgende beschouwingen gebaseerd zijn. In § 4 zullen dan de definities worden gegeven van wat wij als kritieke situaties beschouwen. In § 5 zal het hiermee samenhangende risico p worden bepaald, dat men ten onrechte de hypothese verwerpt dat de gekozen lijn de ware lijn is. Tenslotte zal in § 6 het betrouwbaarheidsgebied voor L worden bepaald.

Opmerking: Het is in de theorie voor het toetsen van hypothesen in het algemeen niet mogelijk een hypothese van de vorm $\theta = \theta_0$ te *bevestigen* (met risico p). In het algemeen zullen nl., als de waarde θ_0 niet voor verwerping in aanmerking komt, ook naburige waarden van θ niet verworpen kunnen worden. Het meest positieve resultaat, dat men met deze theorie kan bereiken is het *verwerpen* van een dergelijke hypothese (met risico p). Daarom stelt men, indien dit mogelijk is, een hypothese ter toetsing op, waarvan de verwerping een waardevolle conclusie is.

Als gevolg van deze moeilijkheid zal in het algemeen een betrouwbaarheidsgebied, dat, in overeenstemming met bovenstaande opmerking, verkregen wordt door die waarden van θ , die wel verworpen kunnen worden, weg te laten, niet uit één enkele waarde van θ bestaan, maar uit een verzameling van dergelijke waarden, die alle als gelijkelijk „bevestigd” dienen te worden beschouwd. Het is daarom beter de term „bevestiging” te vermijden en slechts te spreken van hypothesen, „die niet voor verwerping in aanmerking komen.”

§ 3. *Onderstellingen omtrent de meetfouten.*

Onderstelling I: a) De waarschijnlijkheidsverdeling van de meetfouten u_i en v_i is voor iedere i dezelfde en de meetfouten u_i en v_i zijn stochastisch onafhankelijk van alle overige meetfouten u_j en v_j (met $j \neq i$). *)

b) De kans, dat een waargenomen punt P_i boven L ligt is gelijk aan de kans, dat het eronder ligt en de kans, dat het erop ligt is gelijk aan 0.

c) Voor iedere vaste lijn evenwijdig aan L is de kans, dat één der waargenomen punten erop ligt, gelijk aan nul.

Gevolgtrekkingen: geven wij de hoogte van P_i boven de (onbekende) lijn L, gemeten in de η -richting, aan met z_i , waarbij z_i negatief wordt gerekend, als P_i onder L ligt, dan volgt uit onderstelling Ia, dat iedere z_i beschouwd kan worden als een waarneming van een stochastische variabele z_i , waarbij deze variabelen z_i ($i = 1, \dots, n$) onderling onafhankelijk verdeeld zijn volgens dezelfde waarschijnlijkheidsverdeling.

De n bij de waargenomen punten behorende waarden z_i kunnen dus beschouwd worden als een steekproef van n onafhankelijke waarnemingen van één stochastische variabele z . (Dit geldt ook als we de afstand van P_i tot L niet in de η -richting meten, maar in een willekeurige andere vaste richting).

Uit onderstelling Ic volgt bovendien, dat de kans, dat twee of meer der z_i gelijk zullen zijn, gelijk aan 0 is, terwijl uit Ib volgt dat de kans, dat z_i positief is evenals de kans, dat z_i negatief is, gelijk is aan $\frac{1}{2}$.

Opmerkingen: a) De onderstellingen Ia, b en c zijn b.v. vervuld, als de u_i voor iedere i dezelfde symmetrische en continue verdeling, met gemiddelde nul, bezitten en analoog voor de v_i , terwijl bovendien alle meetfouten stochastisch onafhankelijk van elkaar zijn. Dit betekent, dat er geen systematische meetfouten worden gemaakt.

b) Is aan onderstelling Ib niet voldaan, dan kan men wel op de in het volgende beschreven wijze een betrouwbaarheidsinterval voor de richtingscoëfficiënt van de lijn L bepalen, maar de bepaling voor een betrouwbaarheidsgebied voor L is dan, zonder dat enige verdere onderstelling gemaakt wordt, niet meer mogelijk.

Onderstelling II: De „ware” punten Q_i ($i = 1, \dots, n$) liggen op L en daar uiteraard ondersteld wordt, dat zij alle eindige coördinaten bezitten, liggen zij binnen een eindig interval van L, zodat zij, indien men dit interval in één van de twee mogelijke richtingen doorloopt, een bepaalde volgorde bezitten. In § 1 is gezegd, dat de punten Q_i niet noodzakelijk alle verschillend zijn,

*) Voor eenzelfde i behoeven u_i en v_i niet onafhankelijk te zijn, als ook de afhankelijkheid voor iedere i maar dezelfde is.

m.a.w. dat er bij één „waar” punt meerdere punten P_i (waarnemingen van dat punt) kunnen behoren.

Onderstelling II bestaat nu uit de volgende *twee delen*:

1^o. We onderstellen, dat van ieder van de twee uiterste „ware” punten, die we Q_1 en Q_n zullen noemen, slechts één waarneming gedaan is, of anders gezegd, dat er slechts één eerste en één laatste „waar” punt is, die uiteraard verschillend moeten zijn.

2^o. We onderstellen, dat wij onder de waargenomen punten P_i met zekerheid een puntenpaar, dat we met P_1P_n zullen aangeven, kunnen aanwijzen, dat bij het puntenpaar van „ware” punten Q_1Q_n behoort (daarbij is het niet van belang of P_1 een waarneming van Q_1 is en P_n van Q_n , of P_1 van Q_n en P_n van Q_1); bovendien moet het zeker zijn, dat bij herhaling van de serie waarnemingen van dezelfde punten Q_i weer aan het voorgaande voldaan zal zijn.

Opmerking: Deze vrij ingewikkelde onderstelling is nodig om de stochastische onafhankelijkheid van de z_i bij de in het volgende beschreven bewerking te behouden. Indien men niet op andere gronden dan de waargenomen waarden van de coördinaten weet, welke punten waarnemingen van de twee uiterste punten zijn, betekent het, dat de meetfouten klein moeten zijn in vergelijking met de afstand van de punten P_1 en P_n tot de overige punten. Populair uitgedrukt: de meetfouten moeten zo klein zijn, dat het eerste en het tweede punt (en evenzo het op één na laatste en het laatste) niet ten gevolge van de meetfouten van volgorde kunnen verwisselen.

§ 4. Toetsing van de hypothese, dat L een gegeven lijn L' is.

De in § 6 te bespreken constructie van een betrouwbaarheidsgebied voor L berust op zekere afspraken over het al of niet aanvaardbaar achten van een voorgestelde lijn L' .

Daarom stellen wij eerst op een „recept” voor het verwerpen van de hypothese, dat een gegeven lijn L' het werkelijke lineaire verband tussen ξ en η aangeeft:

Trek door P_1 en P_n de lijnen L_1' en L_n' evenwijdig aan L' (zie fig. 1).

We verwerpen nu de hypothese dat L' de onbekende lijn, dus $L' = L$, is, als zich minstens één van beide volgende situaties voordoet, welke wij „kritieke situaties” zullen noemen.

Eerste kritieke situatie: van de n punten P_1, \dots, P_n liggen slechts een nader te bepalen aantal m ($m < n - 2$), of minder, buiten de door L_1' en L_n' begrensde strook (de lijnen L_1' en L_n' worden bij de strook gerekend).

Tweede kritieke situatie: van de n punten P_1, \dots, P_n liggen er een nader te bepalen aantal ν ($\nu < \frac{1}{2}(n - 3)$), of minder, aan één kant van L' .

Opmerkingen: a) In figuur 1 verkeert L' in de eerste kritieke situatie als wij $m = 1$ stellen en wordt dus op grond daarvan ver-

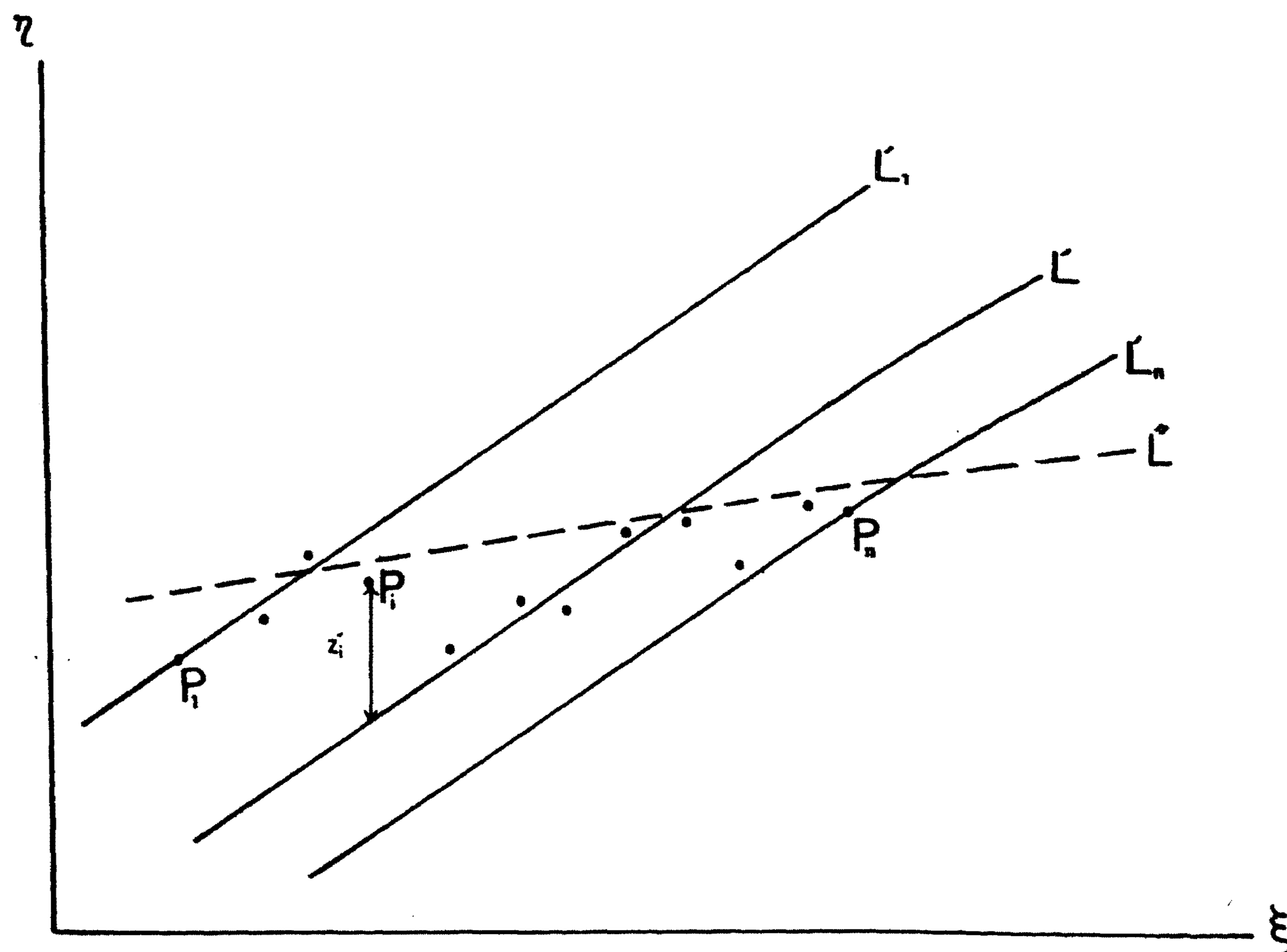


Fig. 1. Toetsing van $L' = L$ of $L'' = L$. De lijnen L_1' en L_n' zijn door de punten P_1 en P_n evenwijdig L' getrokken. n is gelijk aan 12.

worpen; L' verkeert echter niet in de tweede kritieke situatie. De lijn L'' echter wordt op grond van de tweede kritieke situatie verworpen als wij ook $\nu = 1$ stellen, maar verkeert, zoals direct blijkt, indien men de lijnen L_1'' en L_n'' evenwijdig aan L'' door P_1 en P_n trekt, niet in de eerste kritieke situatie.

b) De getallen m en ν worden bepaald aan de hand van het risico, dat men wenst te aanvaarden, (vgl. § 5). De keuze van kritieke situaties ter verwerping van een hypothese is steeds in zekere mate willekeurig. In dit geval hangt de keuze van de eerste kritieke situatie nauw samen met de overweging, dat de verbindingslijn $P_1 P_n$ van alle verbindingslijnen $P_i P_j$, de beste voorlopige schatting van de helling van L is, daar Q_1 en Q_n van alle paren (Q_i, Q_j) de grootste onderlinge afstand bezitten. De keuze van de tweede kritieke situatie ligt in verband met onderstelling Ib, zeer voor de hand.

§ 5. Berekening van het risico uit n , ν en m .

Het risico van de in § 4 beschreven toetsing-methode wordt als volgt berekend:

Indien L' de gezochte (onbekende) lijn L , dus $L = L'$ is, dan is de in figuur 1 met z'_i aangegeven, in de η -richting gemeten afstand van P_i tot L' juist de in § 3 ingevoerde z_i , zodat de z'_i ($i = 1,$

..., n) beschouwd kunnen worden als n onafhankelijke waarnemingen van een continu verdeelde stochastische variabele z , die gelijke kans bezit om positief of negatief te zijn.

De eerste kritieke situatie zal zich dan en slechts dan voordoen, als z_1' en z_n' op hoogstens m van de z_i' na de twee uiterste waarden zijn. Daar echter alle z_i' gelijkwaardig zijn (het zijn immers onafhankelijke waarnemingen van dezelfde stochastische variabele z) is de kans, dat z_1' precies de j^e waarde van boven af heeft, gelijk aan $1/n$. Is dit het geval (met $j \leq m+1$) dan treedt de eerste kritieke situatie op, als bovendien z_n' één van de $m+2-j$ laagste waarden van z_1', \dots, z_n' bezit, daar er dan hoogstens m van de waarden z_i' buiten het interval met z_1' en z_n' als eindpunten liggen. De kans daarop is, daar er, als z_1' bepaald is, nog $n-1$ punten over zijn, gelijk aan $(m+2-j)/(n-1)$ en daar in de vorige zinnen ook „van boven af” door „van onderen af”, benevens „laagste” door „hoogste” vervangen kan worden, geldt voor de kans (die we p_1 noemen) op het optreden van de eerste kritieke situatie, als $L = L'$ is:

$$(3) \quad p_1 = 2 \sum_{j=1}^{m+1} \frac{1}{n} \cdot \frac{m+2-j}{n-1} = \frac{(m+1)(m+2)}{n(n-1)}.$$

Bij het berekenen van deze kans is geen gebruik gemaakt van onderstelling Ib. (Vgl. opmerking b van § 3). Tevens mogen we de in de η -richting gemeten z_i' vervangen door de in een willekeurige andere richting gemeten analoge grootheden (vgl. eveneens § 3); dit is b.v. noodzakelijk, als de lijn L' verticaal zou lopen.

Indien de tweede kritieke situatie zich voordoet, zijn er van de n waarden z_1', \dots, z_n' hoogstens ν positief en de rest negatief, of andersom. Volgens onderstelling Ib is de kans, dat z positief is, evenals de kans, dat z negatief is, gelijk aan $\frac{1}{2}$, dus geldt voor de kans (die we p_2 noemen) op het optreden van de tweede kritieke situatie, als $L = L'$ is:

$$(4) \quad p_2 = 2 \cdot 2^{-n} \sum_{i=0}^{\nu} \binom{n}{i} = 2^{-n+1} \sum_{i=1}^{\nu} \binom{n}{i}.$$

Immers men heeft in dit geval te maken met de binomiale verdeling met kansen $\frac{1}{2}$ (zoals bij „kruis en munt”) en daarvan moeten de beide uiteinden in rekening gebracht worden.

Voor het risico p , dit is de kans, dat één van beide kritieke situaties op zal treden, als $L = L'$ is, geldt nu:

$$(5) \quad p = p_1 + p_2 - p_1 p_2.$$

Het bewijs van deze betrekking, dat berust op de onafhankelijkheid van de twee kritieke situaties, zullen wij achterwege laten. Men ziet gemakkelijk in, dat p zeker niet groter kan zijn dan $p_1 + p_2$ en de term $p_1 p_2$ is gewoonlijk zo klein, dat hij verwaarloosd kan

worden. Wij zullen deze term verder dan ook buiten beschouwing laten.

Men zal ν en m zo bepalen, dat p een waarde krijgt, die men niet te groot acht. Veelal neemt men p hoogstens 0,05, soms kleiner. Verkleining van p betekent echter vermindering van het aantal lijnen, waarvan men (met risico p) kan beweren, dat zij van L verschillende, dus vergroting van het nog te construeren betrouwbaarheidsgebied.

In de keuze van ν en m zal, bij een vast maximum voor p , nog vrijheid bestaan als n niet te klein is. Hierdoor heeft men dan nog een zekere vrijheid van keuze voor het betrouwbaarheidsgebied.

§ 6. *Betrouwbaarheidsgebied voor L.*

Een betrouwbaarheidsgebied G voor L vinden we nu als de verzameling van al die lijnen, waarvan, volgens de in § 4 besproken

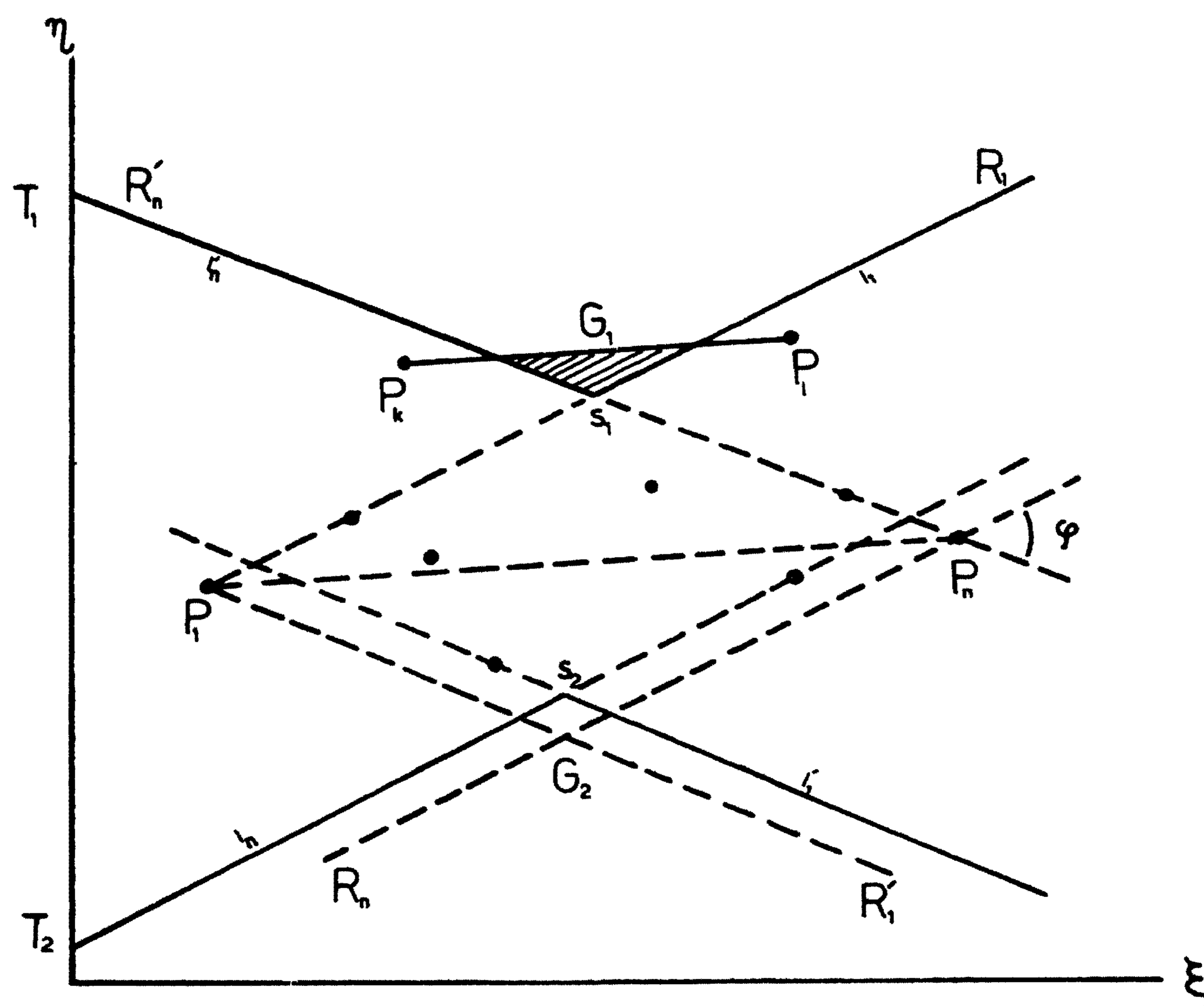


Fig. 2. ($n = 10$, $\nu = 1$, $m = 1$). Het betrouwbaarheidsgebied voor L is het gebied van het platte vlak, binnen de zwaar getrokken lijnen, waarin de 10 meetpunten liggen.

methode, de hypothese, dat zij gelijk aan de lijn L zijn, niet voor verwerping in aanmerking komt. Men kan G op de volgende wijze construeren (vgl. fig. 2, waar $n = 10$, $\nu = 1$ en $m = 1$ genomen is; in deze figuur zijn de punten P_1, \dots, P_{10} met het oog op de duidelijkheid ongunstig, d.w.z. sterk van lineair verband afwijkend,

getekend; dit kan men door schaalverandering steeds bereiken).

Wij trekken door P_1 en P_n twee evenwijdige lijnen in willekeurige richting en gaan na, of de eerste kritieke situatie optreedt. Is dit het geval, dan gaan wij de lijnen, die steeds evenwijdig gehouden worden, om de punten P_1 en P_n met de wijzers van de klok meewentelen, tot zij gekomen zijn in een dergelijke stand, dat bij verdere draaiing deze kritieke situatie niet meer optreedt. De zo verkregen lijnen noemen wij P_1R_1 en P_nR_n . Treedt in de beginstand de eerste kritieke situatie niet op, dan draaien wij juist in de tegenovergestelde richting, tot verdere draaiing de lijnen in de eerste kritieke situatie zou brengen en verkrijgen op die wijze dezelfde lijnen P_1R_1 en P_nR_n . Vervolgens draaien wij de lijnen verder (resp. terug) met de wijzers van de klok mee, zodat zij een hoek van richtingen doorlopen, waarbij de eerste kritieke situatie niet optreedt, en wel wentelen wij ze zover, dat verdere wenteling hen weer in de eerste kritieke situatie zou brengen. In deze stand noemen we de lijnen P_1R_1' en P_nR_n' . Men ziet nu gemakkelijk in, dat iedere lijn L' , waarvan de richting binnen de zo verkregen, in fig. 2 met φ aangegeven, hoek van richtingen valt, niet in de eerste kritieke situatie verkeert. Door de twee evenwijdige lijnen door P_1 en P_n alle mogelijke standen te laten doorlopen, vindt men op deze wijze een eindig aantal dergelijke hoeken van „toegelaten” richtingen; een voldoende, maar niet noodzakelijke voorwaarde, opdat er slechts één dergelijke hoek is, is dat er een stand van de twee evenwijdige lijnen door P_1 en P_n is, waarbij alle punten P_1, \dots, P_n tussen deze lijnen in liggen. We zullen verder onderstellen, dat er slechts één dergelijke hoek φ is en dat deze niet alle richtingen bevat. Dit is weliswaar niet onmogelijk, maar kan toch slechts in zeer bijzondere gevallen voorkomen.

Vervolgens schuift men de lijnen P_1R_1 en P_nR_n met behoud van hun richting zover naar elkaar toe, dat er aan beide zijden precies ν punten buiten de door hen begrensde strook liggen, terwijl ze niet dicht bij elkaar gebracht kunnen worden zonder dat dit aantal punten wordt vergroot *).

Is deze stand bereikt, dan geven we de verschoven lijn P_1R_1 aan met l_1 (in fig. 2 is dit P_1R_1 zelf) en de verschoven lijn P_nR_n met l_n . Hetzelfde doet men met P_1R_1' en P_nR_n' , die in de nieuwe stand met l_1' en l_n' aangegeven worden (in fig. 2 is P_nR_n' gelijk aan l_n').

Noem nu het snijpunt van l_1 en l_n' : S_1 en dat van l_1' en l_n : S_2 . De lijnen l_1, l_1', l_n en l_n' sluiten een parallellogram in, waarvan S_1 en S_2 overstaande hoekpunten zijn. Het gebied gelegen binnen de overstaande hoek van hoek S_1 van dit parallellogram noemen we G_1^* en dat gelegen binnen de overstaande hoek van hoek S_2 van het parallellogram G_2^* (de gebieden G_1^* en G_2^* zijn in figuur 2 niet aangegeven).

*) Indien $\nu < m$ is, kan het ook noodzakelijk zijn de lijnen van elkaar af te schuiven, of beide in dezelfde richting te verplaatsen.

Buiten hoek S_1 van het parallelogram liggen hoogstens 2ν punten (in figuur 2 de punten P_k en P_i ; het kunnen er minder zijn dan 2ν , daar ook punten in G_1^* kunnen vallen; in figuur 2 is dit niet het geval). Deze punten verbinden wij 2 aan 2; van de zo verkregen verbindingslijnen zijn alleen die van belang, waarvan de richting binnen φ valt, die G_1^* snijden en waarbij bovendien aan beide zijden meer dan $\nu - 2$ van de punten P_1, \dots, P_n liggen. Dergelijke lijnen behoeven er niet onder de verbindingslijnen te zijn; zijn er echter wel, dan verminderen wij het gebied G_1^* met de driehoeken, die deze lijnen er van afsnijden (in fig. 2 het gearceerde driehoekje) en noemen het aldus verkregen gebied G_1 . Analoge operaties worden op G_2^* uitgevoerd ter verkrijging van G_2 (in fig. 2 snijdt P_1P_n het gebied G_2^* niet, zodat $G_2 = G_2^*$ is). De grenzen van G_1 en G_2 kunnen naar willekeur gerekend worden al of niet tot G_1 en G_2 te behoren. Indien nu L' een lijn is, die een punt van G_1 of van G_2 bevat en de richting van L' binnen de door φ aangegeven hoek van toegelaten richtingen valt (zodat de eerste kritieke situatie zich niet voordoet), dan verkeert L' in de tweede kritieke situatie, omdat er geen of slechts één punt ($\nu = 1$) aan éénzijde van L' ligt!

Wij noemen nu dat gedeelte van het vlak, dat we verkrijgen door G_1 en G_2 eruit weg te laten G . De bewering:

„L heeft geen punt met G_1 of G_2 gemeen en verloopt binnen G bovendien zo, dat aan beide zijden minstens $\nu + 1$ van de punten P_1, \dots, P_n liggen”,

bezit nu het risico p ; immers deze bewering is equivalent met: „de lijn L verkeert in geen van beide kritieke situaties”; en de kans dat L wel in één van beide kritieke situaties verkeert is, zoals uit de vorige paragraaf blijkt, gelijk aan p . Preciezer gezegd: de kans, dat de punten P_1, \dots, P_n zo zullen uitvallen, dat de met behulp van deze punten op boven beschreven wijze geconstrueerde gebieden G_1 en G_2 één van beide of beide een punt met de ware lijn L gemeen zullen hebben of dat aan één van beide kanten van L hoogstens ν punten zullen vallen, is gelijk aan p . We noemen G het betrouwbaarheidsgebied van L en de in de hoek φ gelegen richtingen het betrouwbaarheidsinterval voor de richting van L . Het risico van de bij dit laatste betrouwbaarheidsinterval behorende uitspraak:

„de richting van L is één van de binnen φ vallende richtingen”, is gelijk aan p_1 .

Verder bezit de uitspraak:

„het snijpunt van L met de η -as ligt tussen de punten T_1 en T_2 ”, (d.w.z. binnen de doorsnijding van G en de η -as; zie fig. 2) het risico p .

Opmerking: Uit dit betrouwbaarheidsinterval volgt direct een betrouwbaarheidsinterval voor α/β uit (1); voor ieder van de toegelaten waarden α/β kan men nu met behulp van de eerste kritieke situatie een betrouwbaarheidsinterval voor γ/β vinden. Op die wijze

vindt men dan een betrouwbaarheidsgebied in het $(\alpha/\beta, \gamma/\beta)$ -vlak met risico ϕ .

Volgens principes analoog met de hier beschrevene kan men de hypothese toetsen, dat L door een bepaald punt gaat; dit wordt hier niet nader beschreven, daar de constructie van het betrouwbaarheidsgebied G, zij het op enigszins omslachtige wijze, een oplossing voor dit probleem inhoudt.

Een tabel van ϕ_1 en ϕ_2 is opgenomen in J. H e m e l r i j k (1949a). Het minimum aantal punten, dat nodig is om de methode met een niet te groot risico toe te kunnen passen, is 7 (dan is nl. $\phi_1 = 0,048$, $\phi_2 = 0,016$ als $m = \nu = 0$ genomen wordt).

Statistische Afdeling
van het Mathematisch Centrum
Amsterdam.

LITTERATUUR:

- A. W a l d (1940), The fitting of straight lines if both variables are subject to error, *Annals Math. Statistics* **11**, p. 284–300.
- J. N e y m a n (1942), Basic ideas and some results of the theory of testing statistical hypotheses, *Jrn. Royal Statistical Soc.* **105**, p. 292–327.
- H. C r a m é r, (1946) *Mathematical Methods of Statistics*, Princeton 1946.
- J. H e m e l r i j k (1949a), Construction of a confidence region for a line, *Proc. Kon. Ned. Ak.* **52** p. 995–1005, *Indag. Math.* **11**, p. 374–384.
- J. H e m e l r i j k (1949b), Over de bepaling van betrouwbaarheidsintervallen en schattingen van de coëfficiënten van een rechte lijn uit een aantal onnauwkeurig waargenomen punten, Rapport S.O. 1 van het Mathematisch Centrum, Amsterdam.
- H. T h e i l (1950), A rank-invariant method of linear and polynomial regression analysis I, *Proc. Kon. Ned. Ak.* **5g**, p. 386–392.