MATHEMATISCH CENTRUM
AMSTERDAM
STATISTISCHE AFDELING

November 1950.

Rapport S 49
door T.J.Terpstra
en J.Hemelrijk

# Statistisch onderzoek omtrent het tellen van erythrocyten met telkamers.

## Inhoud:

## 1. Inleiding.

Voor het tellen van aantallen erythrocyten per volume-eenheid bloed worden tegenwoordig twee methoden gebruikt: de z.g. normale (vullings-) methode en de doorzuigmethode.

Deze 2 methoden verschillen hierin, dat bij eerstgenoemde methode de erythrocyten-suspensie door de capillaire krachten in de telkamer gezogen wordt, terwijl dit bij de tweede methode geschiedt met behulp van een filtreerpapiert

De normale methode was tot nu toe steeds gebruikelijk, maar de laatste tijd wordt in verschillende publicaties de doorzuigmethode aanbevolen.
(Zie o.a. E.Florijn en G.Smits, Het tellen van bloedcellen, N.T.V.G. 91.IV. 46, 1947).

Bij de doorzuigmethode zou volgens genoemde publicatie de verdeling van de aantallen cellen dermate regelmatig zijn, dat zonder bezwaar voor de betrouwbaarheid der uitkomsten op een willekeurige plaats in het netwerk geteld ka worden.

De normale methode (vullingsmethode) zou het nadeel bezitten, dat wanneer bij B, C (fig.1) de bloedsuspensie ondergebracht wordt, het front van de suspensie steeds de grootste concentratie bloedcellen bevat.

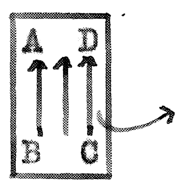



fig. 1

De bloedcellen zouden dus niet homogeen over het telvlak verdeeld zijn. Om dit effect zoveel mogelijk te beperken kiest men bij het tellen altijd veel hokjes, maar afgezien van het feit, dat het tellen in veel hokjes erg vermoeiend is, wordt de telfout bovendien nog nadelig beinvloed doordat vaker beslist moet worden of een op de rand van een telvakje gelegen cel al of niet meegeteld moet worden.

Aan de hand van waarnemingen met konijnenbloed verkregen, wordt in dit verslag statistisch onderzocht of de bei methoden bij telling in 4 hokjes werkelijk verschillen en zo ja, welke methode de nauwkeurigste tellingen geeft.

Verder wordt nagegaan of de tellingen in 2 verschille de telkamerdelen van eenzelfde telkamer systematisch uiteelopen, hoe het gedrag is van de spreiding der tellingen in 5 telvakjes (bepaling van de nauwkeurigheid der tellingen) en of er een dagperiodiciteit bestaat voor het aantal erythrocyten bij een konijn.

Voor een korte explicatie van de statistische toetsing theorie verwijzen wij naar S 47, M 6[1]

## 2. Statistisch onderzoek ter bepaling van de beste der 2 methoden.

2.1. Het beschikbare materiaal voor het vergelijken der 2 methoden bestaat uit volgens beide methoden verkregen tellingsresultaten met konijnenbloed, waarbij telkens 5 keer bloed afgenomen is van eenzelfde konijn en voor één bloedafname steeds 4 tellingen verricht worden op de plaatsen A, B, C en D (zie fig. 1), gelegen in één telkamerdeel.

Op deze wijze zijn volgens beide methoden 30 bloedmonsters geteld (zie bijlage 1).

2.2.1. Om na te gaan of met dit waarnemingsmateriaal het vermeende verschil in bloedcellen-concentraties (erythrocyten) bij B, C en A, D te constateren is zijn de verschillen v bepaald:

$$v = (n_B + n_C) - (n_A + n_D)$$

($n_A$ = aantal bloedcellen per volume-eenheid in telvlak A, etc.)

Wanneer namelijk geen systematisch verschil in concentratie bij (B,C) en (A,D) bestaat, dan is de stochastische grootheid v symmetrisch t.o.v. 0 verdeeld.

Om te toetsen of de berekende verschillen v beschouwd kunnen worden als een steekproef uit een symmetrische verdeling is de symmetrietoets $T_1$(zie S 47, M 10) toegepast. Deze toets geeft voor de normale (vullings-) methode een tweezijdige overschrijdingskans k = 0,18.

Daar de verschillen v voor de doorzuigmethode nog gunstiger gelegen zijn (zie grafiek 1), zal de symmetrietoets voor deze methode een overschrijdingskans geven, die zeker niet kleiner dan 0,18 is.

Dit betekent, dat wij, zonder het maken van verdere onderstellingen omtrent de waarschijnlijkheidsverdeling van het verschil $v = (n_B + n_C) - (n_A + n_D)$, geen afwijking van de symmetrie vinden, die niet aan het toeval geweten kan worden.

Dit geldt voor beide methoden.

---

1) Zie de appendix voor deze en volgende verwijzingen.

2.2.2. Gewoonlijk wordt in een geval als het onderhavige wel een extra onderstelling omtrent deze waarschijnlijkheidsverdeling gemaakt, daaruit bestaande, dat v een normale verdeling[1] bezit. Op deze onderstelling baseert men dan de toepassing van de toets van Student (zie S 47, M 8) voor de hypothese, dat het gemiddelde van de verdeling van v gelijk aan 0 is. Alvorens tot de toepassing van deze toets over te gaan, hebben wij de normaliteit van de verdeling van v getoetst met de normaliteitstoetsen van Geary en Pearson (zie S 47, M 1 met behulp van de waarnemingen van v. Daarbij werden, zowel voor de kurtosis als voor de symmetrie, overschrijdingskansen $> 0,20$ gevonden, zodat van een afwijking der normaliteit niets blijkt. Dit houdt niet in, dat de verdeling van v inderdaad normaal is, te meer daar het aantal waarnemingen van v slechts 30 bedraagt. Niettemin kan men er uit concluderen, dat een grove afwijking van normaliteit vermoedelijk niet aanwezig is.
Passen wij nu de bovengenoemde toets van Student toe, dan vinden wij voor de doorzuigmethode een tweezijdige overschrijdingskans van ongeveer 0,65 en voor de normale methode 0,048. Dit verschil is aanzienlijk en uit de tweede overschrijdingskans zou men geneigd zijn te concluderen tot verwerping van de getoetste hypothese bij de normale methode, dus tot een systematisch verschil tussen $n_B + n_C$ en $n_A + n_D$.
Er zijn echter enkele bezwaren tegen deze verwerping aan te voeren, die ons nopen dit resultaat in een voorzichtige vorm te geven, ondanks het feit, dat de genoemde toetsen voor normaliteit niet tot verwerping van deze onderstelling leiden. Deze bezwaren zijn het gemakkelijkst duidelijk te maken aan de hand van grafiek 1. Men ziet daar, dat de waarnemingen van v bij de normale methode alle op één na vrij regelmatig om het punt v=0 gegroepeerd liggen, terwijl er één zeer grote negatieve waarde optreedt (-102), die het resultaat in hoge mate beinvloedt. Het gemiddelde en de spreiding van de gehele steekproef bedragen -12,87 resp. 33,5 en indien deze waarneming weggelaten wordt -9,79 resp. 30,1.
Bij weglating van deze waarneming stijgt de overschrijdingskans voor de toets van Student tot 0,088, een waarde, die gewoonlijk te groot wordt geacht, om tot verwerping over te gaan.

---

1) Ook "verdeling van Gauss" genoemd.

Om enigszins na te gaan, in hoeverre al of niet weglating gerechtvaardigd geacht kan worden, kan men de volgende redenering toepassen: de normaliteit aanvaardende en aannemende, dat gemiddelde en spreiding ergen tussen de twee gevonden waarden -12,87 en -9,79 resp. 33,5 en 30,1 liggen, is de kans op het optreden van een minstens zo ver weg liggende waarneming als -102 onder 30 waarnemingen gelegen tussen 0,12 en 0,036. Ook hieruit valt dus geen duidelijke conclusie te trekken. De gevonden kans is van ongeveer dezelfde grootte als de bij de toets van Student gevonden overschrijdingkans, zodat de keuze tussen het weglaten van de kleinste waarneming, gepaard aan het niet verwerpen van de met de toets van Student getoetste hypothese en het niet weglaten daarvan, gepaard aan het verwerpen van deze hypothese, een moeilijke is.
Mede in verband met de in § 2.3.2. verkregen resultaten komen wij tot de conclusie dat er een aanwijzing is, dat de normale methode een minder regelmatig bloedbeeld geeft dan de doorzuigmethode, maar dat op grond van het verwerkte materiaal alleen geen definitieve uitspraak hierover kan worden gedaan.
Nadere bevestiging (of weerlegging) door verdere experimenten zou daarvoor nodig zijn.

2.3. Naast de verschillen v zijn ook de gemiddelden en spreidingen der viertallen waarnemingen in A, B, C en van belang voor het vergelijken der twee methoden. Hiervoor is weer het in bijlage 1 vermelde waarnemingsmateriaal gebruikt. Wij vergelijken eerst (in 2.3.1.) de gemiddelden der viertallen waarnemingen bij beide methoden. Indien hiertussen voor de twee methoden een systematisch verschil zou blijken te bestaan, zou de gemiddelde uitkomst van de tellingen dus van de methode afhankelijk zijn. Dit blijkt echter niet het geval te zijn. Vervolgens vergelijken wij (in 2.3.2.) de spreidingen der viertallen tellingen bij beide methoden. Indien één van beide methoden nauwkeuriger is dan de andere, zou men dit vermoedelijk kunnen constateren, doordat de spreidingen bij deze methode kleiner zijn dan bij de andere. Ook hiervan blijkt echter niets uit de verrichte waarnemingen.

2.3.1. Voor de viertallen tellingen zijn voor beide methoden de gemiddelden bepaald (zie bijlage 1). M.b.v. de Toets van Wilcoxon (zie S 47, M 7) is nu onderzocht of de beide reeksen gemiddelden, beschouwd kunnen worden als steekproeven uit eenzelfde verdeling. De uitkomst van deze toets is een tweezijdige overschrijdingskans 0,34 zodat niet besloten kan worden, dat de gemiddelden van de ene reeks systematisch afwijken van die van de ande M.b.v. deze gemiddelden is dus statistisch geen verschil tussen beide methoden te constateren.

2.3.2. Vervolgens hebben wij, eveneens met de toets van Wilcoxon, de hypothese getoetst, dat de spreidingen va de met beide methoden verkregen viertallen tellingen steekproeven uit eenzelfde verdeling zijn. Wij vonden daarbij een (tweezijdige) overschrijdingskans 0,65, hetgeen betekent, dat er geen verschil geconstateerd wordt, dat niet zeer wel door toeval kan zijn ontstaan Indien de regelmatigere verdeling der erythrocyten ove het telvlak, die bij de doorzuigmethode optrad (zie 2.2.2.), niet toevallig is, zou men verwachten, dat de spreidingen der viertallen tellingen bijdeze methode systematisch kleiner zouden zijn dan bij de normale methode.

Dit is echter met de toets van Wilcoxon, welke een vri scherpe toets is, niet aan te tonen. Er bestaat dus op grond van het beschikbare waarnemingsmateriaal slechts een zwakke aanwijzing, dat de doorzuigmethode de voorkeur verdient boven de normale methode.

3. <u>De tellingen in 2 verschillende telkamerdelen van eenzelfde telkamer.</u>

3.1. Het waarnemingsmateriaal, dat gebruikt wordt om 2 verschillende telkamerdelen van eenzelfde telkamer te ver gelijken, bestaat uit vijftallen tellingen op de plaat sen A, B, C, D en E voor de 2 telkamerdelen (zie fig.2 en bijlage 2)

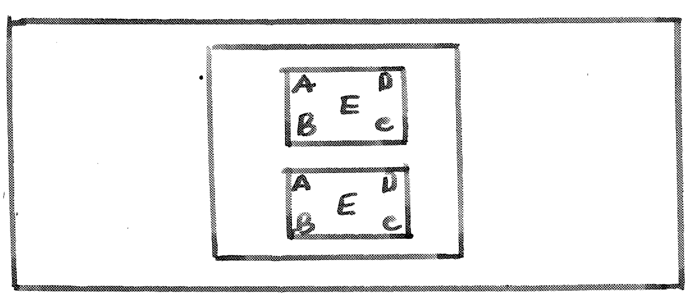



fig. 2

Telkens worden de twee vijftallen tellingen A, B, C, D en E voor de 2 telkamerdelen verricht met eenzelfde bloedmonster, zodat de resultaten der tellingen voor d 2 telkamerdelen uit paren van vergelijkbare vijftallen bestaan.

Deze tellingen zijn voor telkamerdelen van 3 verschillende telkamers I, II en III verricht (in het oorspronkelijke waarnemingsmateriaal A, B en C genoemd; daar deze letters ook reeds voor telkamerdelen gebruik zijn, hebben wij ze vervangen).

3.2. Voor ieder vijftal tellingen zijn de gemiddelden bepaald en vervolgens de verschillen van gemiddelden va telkens 2 corresponderende vijftallen.
Wanneer de tellingen der telkamerdelen niet systematisch uiteenlopen, d.w.z. dat de tellingen van overeen komstige bloedmonsters bij beide telkamerdelen als onafhankelijke steekproeven uit dezelfde collectie kunner worden beschouwd, dan zijn de verschillen der gemiddel den van corresponderende vijftallen symmetrisch verdee t.o.v. 0. We kunnen dus op de berekende verschillen de symmetrietoets toepassen (zie S 47, M 10, daarbij is toets $T_1$ gebruikt), welke voor de 3 telkamers I, II en III de niet-significante tweezijdige overschrijdingskansen 0,41, 0,90 en 0,60 geeft, hetgeen inhoudt, dat geen verschil te constateren is in de gemiddelden van vijftallen tellingen. Dit betekent dus, dat tussen de 2 telkamerdelen van eenzelfde telkamer geen systematische verschillen in de tellingsresultaten worden gevonden.

4. <u>Afhankelijkheid van spreiding en gemiddelde der tellingen.</u>

4.1. Voor ieder vijftal tellingen uit het waarnemingsmateriaal beschreven in § 3.1. zijn gemiddelde en spreidir bepaald en door een punt in het x,y-vlak weergegeven. Op deze wijze ontstaat grafiek 2.
Om te toetsen of deze gemiddelden en spreidingen onafhankelijk van elkaar zijn, is de hoek-toets ("Corner test", zie S 47, M 11) toegepast, welke een duidelijke significantie geeft (overschrijdingskans 0,005; zie grafiek 2).
Voor de bepaling van de nauwkeurigheid der tellingen is het van belang deze afhankelijkheid op te heffen. Een vrij ruwe methode, die echter vaak voldoende is, bestaat daaruit, dat men het materiaal splitst in twee groepen, waarvan de ene de grote gemiddelden en de andere de kleine gemiddelden omvat. Voeren wij hier deze splitsing zo uit, dat beide groepen evenveel waarnemingen bevatten, dan verkrijgen wij een groep van 39 vijftallen met gemiddelden $<$ 280 en een groep van 39 vijftallen met gemiddelden $>$ 280. Passen wij dezelfde toe voor onafhankelijkheid van gemiddelde en spreiding op

deze twee groepen toe, dan vinden wij als overschrijdingskansen ongeveer 0,90 en 0,35, zodat we voor deze beide groepen <u>niet</u> tot afhankelijkheid van gemiddelde en spreiding mogen besluiten (zie grafiek 3).

4.2. Ter contrôle van deze resultaten is als tweede methode voor het toetsen van de onafhankelijkheid van gemiddelde en spreiding der vijftallen tellingen de methode der dubbele dichotomie toegepast (zie S 47, M 15), die voor alle waarnemingen tezamen een tweezijdige overschrijdingskans 0,003 (dus afhankelijkheid) geeft en voor beide in de vorige paragraaf genoemde groepen apart overschrijdingskansen 0,64 en 0,43 (dus geen aanwijzing voor afhankelijkheid) geeft.

Bij de verdere berekeningen voor de nauwkeurigheid der tellingen gaan wij nu van ieder dezer twee groepen apart uit. Wij verwaarlozen daarbij voor ieder dezer groepen de (eventueel nog aanwezige geringe) afhankelijkheid van spreiding en gemiddelde.

5. <u>Bepaling van de nauwkeurigheid der tellingen.</u>

In overeenstemming met het in de vorige paragraaf gevondene onderzoeken wij de nauwkeurigheid der tellingen voor de twee groepen waarnemings-vijftallen met gemiddelde $>$ resp. $<$ 280, apart. Daartoe onderzoeken wij eerst voor ieder dezer twee groepen de verdeling der gevonden spreidingen.

5.1. Indien wij het aantal erythrocyten per telvakje aangeven door $x$ en de gevonden waarden bij 5 tellingen in de vakjes A, B, C, D en E door $x_1, \ldots, x_5$, dan is de spreiding $s$ gedefinieerd als

$$s = \sqrt{\frac{1}{5}\sum_{1}^{5}{}^{i}\left(x_i - \bar{x}\right)^2}$$

waarin

$$\bar{x} = \frac{1}{5}\sum_{1}^{5}{}^{i} x_i$$

is. Indien nu $x$ een normale verdeling met spreiding $\sigma$ bezit, bezit $s$ een bekende verdeling (een z.g. $\Gamma$-verdeling), die het gemakkelijkst op de volgende wijze gegeven kan worden: de grootheid

$$t = \frac{5}{2}\,\frac{s^2}{\sigma^2}$$

bezit de waarschijnlijkheidsdichtheid

$$f(t) = e^{-t}.t$$

een z.g. $\Gamma(2)$-verdeling, waarvan tabellen bestaan. (K.Pearson, Tables of the incomplete $\Gamma$-function, Biometrika Office 1946).

In 2.2.2. bleek reeds, dat de hypothese van normaliteit voor het verschil van gemiddelden van tellingen niet verworpen kan worden. Het is daarom niet onredelijk, te vermoeden, dat de spreidingen der vijftallen tellingen uit beide groepen (met gemiddelden > resp. < 280) bovengenoemde verdeling zouden bezitten. Dit vermoeden is voor beide groepen op twee wijzen getoetst

1e. Met de $\chi^2$-methode voor aanpassing (zie S 47, M 1… waarbij als schatting voor $\sigma$ het gemiddelde der voor s gevonden waarden is genomen. Dit leverde voor de twee groepen als overschrijdingskansen de waarden 0,64 resp. 0,52, hetgeen op een goede overeenstemming van hypothese en waarnemingsmateriaal duidt.

2e. Met behulp van de methode der transformatie op een homogene verdeling van E.S.Pearson (zie S 47, M 1… waarbij dezelfde schatting voor $\sigma$ is gebruikt. Hier werden eveneens grote overschrijdingskansen gevonden (n.l. 0,96 resp. 0,85), hetgeen de vorige uitkomst bevestigt.

Dit wijst erop, dat de hypothese, inhoudende, dat de spreidingen s genoemde $\Gamma$-verdeling bezitten, een redelijke aanname is. Op deze aanname is nu de bepaling der nauwkeurigheid verder gebaseerd.

5.2. Wij bepalen nu een <u>bovengrens</u> voor de onbekende grootheid $\sigma$ in de vorm van een <u>alleen naar boven begrensd betrouwbaarheidsinterval</u> (zie S 47, M 18) en wel voor beide groepen (gemiddelde > resp. < 280) apart. De schatting voor $\sigma$ die in 5.1. voor beide toetsingen gebruikt is, is n.l. weinig geschikt als maat voor de nauwkeurigheid der tellingen, daar de werkelijke waarde zeer wel kleiner of groter kan zijn. De bovengrenzen, die wij bepalen bezitten echter een bekende betrouwbaarheid $1-\alpha = 0,95$, hetgeen dus betekent, dat wij bij herhaaldelijk bepalen van dergelijke bovengrenzen in ongeveer één van de 20 gevallen een getal als bovengrens zouden vinden, dat in werkelijkheid onder $\sigma$ ligt. De bepaling van de bovengrenzen berust nu op het volgende principe:

Wanneer de grootheid t een $\Gamma(2)$-verdeling bezit, dan bezit de som der 39 t-waarden $T = t_1+t_2+ ---- +t_{39}$ voor iedere groep een $\Gamma(2 \times 39) = \Gamma(78)$-verdeling, welke zeer goed benaderd wordt door een normale verdeling met gemiddelde 78 en spreidingskwadraat 78.

De grootheid $y = \dfrac{T-78}{\sqrt{78}}$ is dus bij benadering normaal verdeeld, met gemiddelde 0 en spreiding 1. (is N (0,1)). Voor de bepaling van een naar boven begrensd betrouwbaarheidsinterval is het dus noodzakelijk het alleen naar beneden begrensde betrouwbaarheidsinterval voor y te berekenen, waarvoor bij een betrouwbaarheid 0,95 het interval $-1,64 < y$ gevonden wordt, overeenkomende met $-1,64\sqrt{78} < T-78$.
Wij beschouwen nu de groepen vijftallen met gemiddelde groter en kleiner dan 280 apart. Voor ieder hiervan zi 39 waarden van $s^2$ berekend (zie bijlage 2). Noemen wij deze $s_1^2, \dots, s_{39}^2$, dan is

$$T = \frac{5}{2\sigma^2}\left(s_1^2 + \dots\dots + s_{39}^2\right)$$

zoals uit bovengegeven formule volgt. De voor $\frac{5}{2}(s_1^2+\dots+s$ gevonden waarden zijn voor beide genoemde groepen geli aan 9978,42 resp. 5988,76. Wij vinden dus uit de bove gegeven ongelijkheid voor T:

$$-1,64\sqrt{78} < \frac{9978,42}{\sigma^2} - 78,$$

resp. $$-1,64\sqrt{78} < \frac{5988,76}{\sigma^2} - 78.$$

Hieruit volgt voor de groep met gemiddelden groter da 280:

$$\sigma < 12,53$$

en voor de groep met gemiddelden kleiner dan 280:

$$\sigma < 9,71$$

Beide uitspraken gelden met betrouwbaarheid 0,95.
Met dezelfde betrouwbaarheid kunnen wij dus de uitspraak doen, dat het gemiddelde van vijf tellingen ee spreiding $< \frac{12,53}{\sqrt{5}} = 5,60$ bezit, indien dit gemiddelde $> 280$ is en een spreiding $< \frac{9,71}{\sqrt{5}} = 4,34$, indien het $< 280$ is. Dit wijst op een hoge graad van nauwkeurigheid. De variaties in aantallen erythrocyten bij verschillende konijnen en bij hetzelfde konijn op verschillende tijdstippen zijn aanzienlijk groter. Wij wijzen er echter op, dat deze nauwkeurigheid geldt vo de installatie en de waarnemer, waarmee het hier bewerkte waarnemingsmateriaal verkregen is en dat deze nauwkeurigheid voor andere telkamers en andere waarnemers zeer wel verschillend kan zijn. Op bovenstaand wijze kan echter uit de waarnemingen steeds een nauwkeurigheid worden bepaald.

## 6. Dag-periodiciteit van de aantallen erythrocyten.

6.1. Gedurende 3 achtereenvolgende dagen (5, 6 en 7 Augustus zijn met de telkamers I, II en III telkens voor 2 verschillende telkamerdelen tellingen verricht en wel op de tijdstippen 2 u., 8 u., 14 u., en 20 u. (zie bijlage 2).

De 3 uitkomsten der tellingen in één telkamerdeel voor elk der tijdstippen lagen steeds zo dicht bijeen, dat wij zonder gevaar voor verstoring van de mogelijke dagperiodiciteit het gemiddelde van deze 3 tellingsresultaten voor elk der 4 tijdstippen kunnen gebruiken. De uitkomsten, op deze wijze verkregen, zijn hieronder getabelleerd, waarbij $I_1$ de betekenis bezit van: eerste telkamerdeel van telkamer I, etc.

Tabel 1

| | 2 u. | 8 u. | 14 u. | 20 u. |
|---|---|---|---|---|
| $I_1$ | 321 | 297,6 | 272,6 | 298 |
| $II_1$ | 295 | 278 | 276,3 | 289,6 |
| $III_1$ | 265 | 234,6 | 247,3 | 257,6 |
| $I_2$ | 316,5 | 301,3 | 269 | 299,6 |
| $II_2$ | 294,5 | 272 | 274 | 290,6 |
| $III_2$ | 258 | 236,3 | 237,3 | 258 |

6.2. Om in deze gegevens een dagperiodiciteit op te sporen, is gebruik gemaakt van de z.g. "Methode der m rijen" van M.G.Kendall (zie S 47, M 14).

Bij deze methode worden de getallen in iedere rij van bovenstaande tabel volgens opklimmende grootte genummerd. Op deze wijze ontstaat tabel 2:

Tabel 2

| | 2 u. | 8 u. | 14 u. | 20 u. |
|---|---|---|---|---|
| $I_1$ | 4 | 2 | 1 | 3 |
| $II_1$ | 4 | 2 | 1 | 3 |
| $III_1$ | 4 | 1 | 2 | 3 |
| $I_2$ | 4 | 3 | 1 | 2 |
| $II_2$ | 4 | 1 | 2 | 3 |
| $III_2$ | 3,5 | 1 | 2 | 3,5 |

De methode der m rangschikkingen stelt ons in staat, te toetsen of er tussen deze rijen rangnummers een ove eenstemming bestaat, die niet op het toeval berust. De ze toetsingsmethode geeft een overschrijdingskans 0,0006, hetgeen betekent, dat de hypothese dat de over eenstemming een toevallige is met stelligheid verworpe kan worden. De gevonden duidelijke overeenstemming wij dus op een dagperiodiciteit.

6.3. Om enigszins een beeld te krijgen van de geconstateerd dagperiodiciteit in het aantal erythrocyten, is als volgt te werk gegaan:
Tabel 2 toont duidelijk aan, dat de aantallen erythrocyten 's nachts (2 u.) het grootst zijn. Overdag (8 u en 14 u.) echter zijn de aantallen het kleinst en bovendien ongeveer constant (hierop wijst het feit, dat in tabel 2 met één uitzondering de 2 kleinste rangnummers 1 en 2 afwisselend voorkomen in de kolommen onder 8 u. en 14 u.).
's Avonds neemt het aantal erythrocyten weer toe. Daar de aantallen overdag ongeveer constant zijn, ligt het voor de hand de verschillen te berekenen van de aantallen erythrocyten op de tijdstippen 20 u. en 2 u. met d gemiddelden van de twee overdag verrichte waarnemingen Deze berekeningen zijn uitgevoerd voor de telkamers I, IIen III en bij elk voor 2 verschillende telkamerdelen Daar reeds aangetoond is (zie § 3), dat de tellingsresultaten voor 2 telkamerdelen van eenzelfde telkamer nie significant verschillend zijn, hebben wij het gemiddel de van de uitkomsten van bovengenoemde berekeningen voor 2 telkamerdelen van eenzelfde telkamer genomen, waardoor de volgende tabel verkregen wordt, waarin ieder getal een verschil van het gemiddelde van een vijftal waarnemingen met het gemiddelde van een tienta overdag verrichte waarnemingen voorstelt:

Tabel 3

| Telkamer | datum | verschillen van de aantallen erythrocyten om 20 u. en 2 u. met de gemiddelden van overdag | |
|---|---|---|---|
| | | 20 u. | 2 u. |
| I | 5 Aug. | 15,25 | 24,24 |
| | 6 Aug. | 18,75 | 42,75 |
| | 7 Aug. | 6,75 | - |
| II | 5 Aug. | 14 | 25 |
| | 6 Aug. | 10,5 | 14,5 |
| | 7 Aug. | 21 | - |
| III | 5 Aug. | 16,5 | 29 |
| | 6 Aug. | 13 | 16 |
| | 7 Aug. | 27 | - |

Deze periodiciteit voor de 3 telkamers is in beeld gebracht in grafiek 4.
Tenslotte zijn spreidingen en gemiddelden berekend voor de in tabel 3 vermelde verschillen van de aantallen erythrocyten op de tijdstippen 20 u. en 2 u. met de daggemiddelden, gevonden voor de 3 telkamers.

Tabel 4

| | 20 u. | 2 u. |
|---|---|---|
| gemiddelde | 15,9 | 25,3 |
| spreiding | 5,6 | 7,3 |

Om 20 uur is dus het aantal erythrocyten per telvakje gemiddeld 15,9 (met schatting van de spreiding 5,6) hoger dan het gemiddelde aantal overdag en om 2 u. gemiddeld 25,3 (met spreiding 7,3).

7. Conclusies.

1. Er is een aanwijzing, dat de doorzuigmethode een regelmatigerverdeling der erythrocyten geeft dan de normale methode. Nadere bevestiging (of weerlegging) hiervan door een uitgebreider onderzoek is echter gewenst voor het formulere van een definitieve conclusie (zie 2.2.).

2. Een verschil in tellingsresultaten tussen de viertallen tellingen met beide methoden verkregen is met behulp van de toets van Wilcoxon niet te constateren, noch wat de gemidde den, noch wat de spreidingen dezer viertallen betreft (zie 2.3.). Dit laatste draagt ertoe bij, dat de in 1 genoemde aanwijzing slechts met voorbehoud kon worden vermeld.

3. Er werden geen systematische verschillen gevonden tussen de tellingen verricht in verschillende telkamerdelen (zie §

4. Tussen de gemiddelden en de spreidingen van vijftallen tellingen van bloedmonsters bleek een duidelijke afhankelijkheid te bestaan, die geëlimineerd kon worden door de bloedmonsters te splitsen in twee groepen, waarbij het gemiddelde van vijf tellingen > 280 resp. < 280 was. (zie § 4

5. Voor deze beide groepen werd een bovengrens der spreidin voor het gemiddelde van vijf tellingen bepaald, die dus een bovengrens voor de onnauwkeurigheid aangeeft. Deze bedroeg (behoudens onbetrouwbaarheid 0,05):

voor de groep met gemiddelden > 280 5,60
voor de groep met gemiddelden < 280 4,34

Dit betekent, dat een grote nauwkeurigheid der tellingen is bereikt. De verschillen tussen bloedmonsters op verschillen tijden en van verschillende dieren zijn aanzienlijk groter (n.l. van de orde van 3 tot 5 maal deze spreiding).

6. Een duidelijke dagperiodiciteit werd geconstateerd. Het aantal erythrocyten is overdag lager dan 's nachts. Van de tijdstippen, waarop metingen werden verricht, werd om 2 uur 's nachts het grootste aantal erythrocyten gevonden. Men vergelijke tabel 3 en 4 van § 6.2. en grafiek 4.

8. Appendix.

Deze bestaat uit de bijlagen
S 47: 6, 7, 8, 10, 11, 12, 14, 15, 16, 17, 18.

Bijlage 1.

DOORZUIGMETHODE

| konijn,telkamerdeel | | no. bloedafname | fototellingen (4 foto's) A | B | C | D | gem. m | spreiding s |
|---|---|---|---|---|---|---|---|---|
| 5 | a | 1 | 200 | 228 | 217 | 234 | 220 | 12,9 |
| | | 2 | 233 | 250 | 252 | 218 | 238 | 13,8 |
| | | 3 | 222 | 224 | 221 | 205 | 218 | 7,6 |
| | | 4 | 220 | 241 | 246 | 252 | 240 | 12,1 |
| | | 5 | 235 | 246 | 227 | 262 | 242,5 | 13,1 |
| 6 | b | 1 | 306 | 291 | 306 | 324 | 307 | 11,7 |
| | | 2 | 288 | 265 | 271 | 293 | 279 | 11,6 |
| | | 3 | 279 | 287 | 289 | 298 | 288 | 6,8 |
| | | 4 | 330 | 286 | 302 | 318 | 309 | 16,6 |
| | | 5 | 271 | 309 | 303 | 282 | 291 | 15,4 |
| 7 | c | 1 | 259 | 225 | 261 | 268 | 253 | 16,7 |
| | | 2 | 264 | 219 | 231 | 263 | 244 | 19,7 |
| | | 3 | 275 | 262 | 241 | 290 | 267 | 18,0 |
| | | 4 | 360 | 354 | 293 | 313 | 330 | 28,0 |
| | | 5 | 274 | 262 | 393 | 288 | 304 | 52,1 |
| 1 | d | 1 | 326 | 321 | 294 | 276 | 304 | 20,4 |
| | | 2 | 289 | 252 | 273 | 299 | 278 | 17,8 |
| | | 3 | 301 | 302 | 292 | 310 | 301 | 6,4 |
| | | 4 | 283 | 306 | 274 | 279 | 287 | 12,3 |
| | | 5 | 296 | 317 | 320 | 309 | 310,5 | 9,3 |
| 2 | e | 1 | 246 | 246 | 263 | 259 | 253,5 | 7,6 |
| | | 2 | 249 | 275 | 235 | 238 | 249 | 15,8 |
| | | 3 | 262 | 249 | 258 | 246 | 254 | 6,5 |
| | | 4 | 253 | 216 | 237 | 247 | 238 | 14,1 |
| | | 5 | 310 | 316 | 360 | 338 | 331 | 19,7 |
| 148 | f | 1 | 215 | 211 | 224 | 223 | 218 | 5,5 |
| | | 2 | 201 | 230 | 204 | 210 | 211 | 11,3 |
| | | 3 | 230 | 225 | 192 | 235 | 221 | 16,8 |
| | | 4 | 217 | - | - | 220 | 218,5 | 1,5 |
| | | 5 | 230 | 252 | 233 | 212 | 232 | 14,2 |

NORMALE METHODE

| konijn,telkamerdeel | | no. bloedafname | fototellingen (4 foto's) A | B | C | D | gem. m | spreiding s |
|---|---|---|---|---|---|---|---|---|
| 5 | $a^1$ | 1 | 264 | 212 | 266 | 247 | 247 | 21,7 |
| | | 2 | 231 | 228 | 239 | 249 | 237 | 8,1 |
| | | 3 | 238 | 233 | 245 | 212 | 232 | 12,3 |
| | | 4 | 240 | 270 | 263 | 269 | 261 | 12,1 |
| | | 5 | 265 | 255 | 246 | 232 | 249,5 | 12,1 |
| 6 | $b^1$ | 1 | 280 | 267 | 277 | 278 | 275,5 | 5,0 |
| | | 2 | 308 | 262 | 254 | 310 | 284 | 25,7 |
| | | 3 | 272 | 344 | 314 | 349 | 320 | 30,6 |
| | | 4 | 317 | 305 | 285 | 334 | 310 | 17,9 |
| | | 5 | 302 | 293 | 288 | 300 | 296 | 5,6 |
| 9 | $c^1$ | 1 | 278 | 285 | 273 | 260 | 274 | 9,1 |
| | | 2 | 299 | 296 | 315 | 274 | 296 | 14,6 |
| | | 3 | 260 | 286 | 273 | 286 | 276 | 10,8 |
| | | 4 | 253 | 236 | 223 | 225 | 234 | 11,9 |
| | | 5 | 307 | 288 | 289 | 267 | 288 | 14,2 |
| 10 | $d^1$ | 1 | 330 | 290 | 303 | 309 | 308 | 14,4 |
| | | 2 | 283 | 271 | 260 | 263 | 269 | 8,9 |
| | | 3 | 279 | 306 | 253 | 304 | 285,5 | 21,6 |
| | | 4 | 298 | 265 | 281 | 295 | 285 | 13,1 |
| | | 5 | 261 | 265 | 255 | 263 | 261 | 3,7 |
| 148 | $e^1$ | 1 | 231 | 224 | 219 | 253 | 232 | 13,0 |
| | | 2 | 220 | 257 | 281 | 259 | 254 | 21,9 |
| | | 3 | 235 | 215 | 215 | 225 | 222,5 | 8,3 |
| | | 4 | 232 | 220 | 239 | 242 | 233 | 8,5 |
| | | 5 | 248 | 215 | 254 | 249 | 241,5 | 15,5 |
| 223 | $f^1$ | 1 | 335 | 288 | 304 | 310 | 309 | 16,9 |
| | | 2 | 298 | 309 | 282 | 322 | 303 | 14,7 |
| | | 3 | 309 | 329 | 337 | 318 | 323 | 10,6 |
| | | 4 | 325 | 327 | 350 | 347 | 337 | 11,3 |
| | | 5 | 344 | 306 | 308 | 313 | 318 | 15,4 |

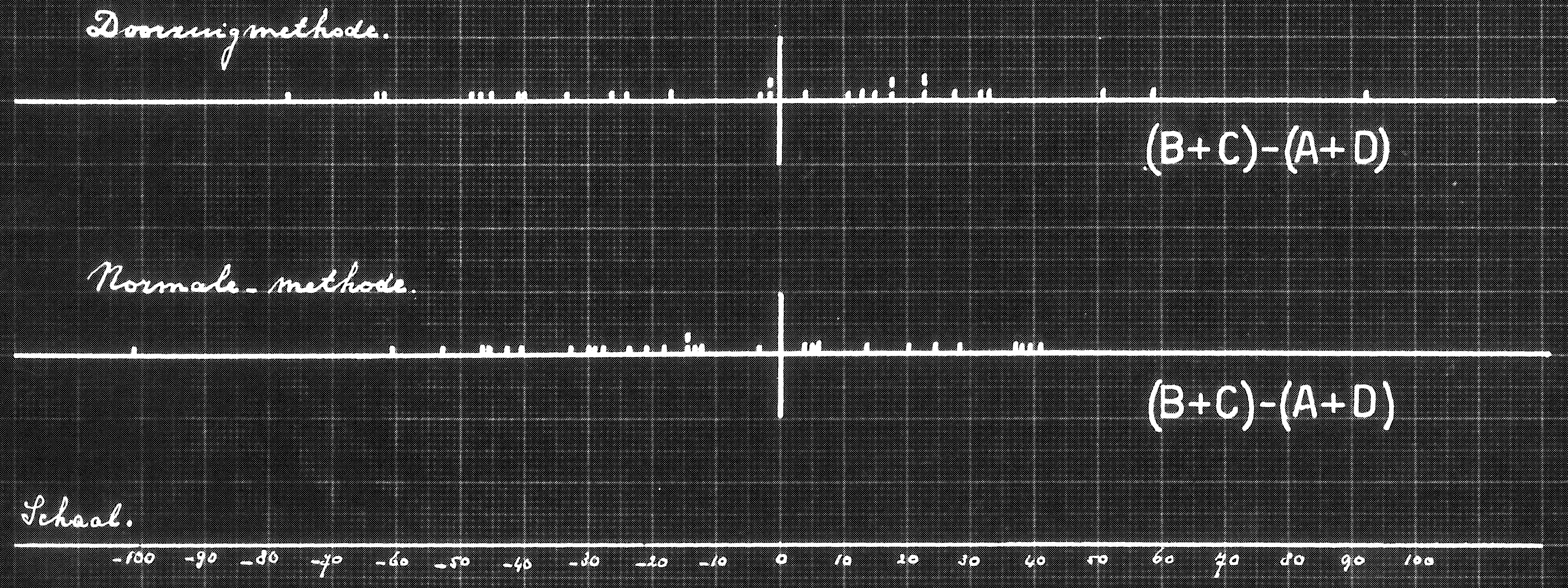
MATHEMATISCH CENTRUM
AMSTERDAM

Bijlage 2.

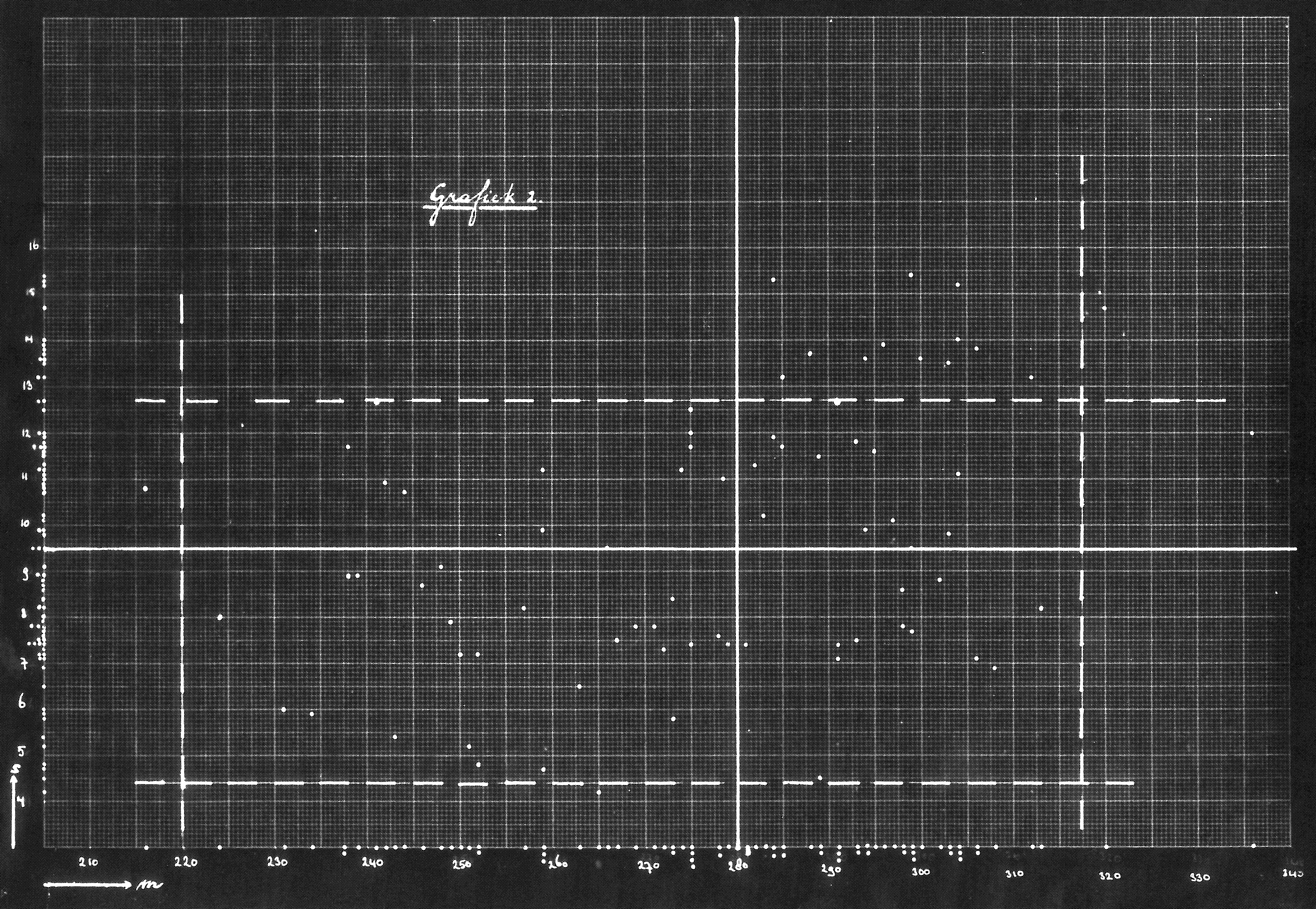
TELKAMER I.

| konijn | datum | uur | no. bloedafname | 1ste telkamerdeel 5 foto's A | B | C | D | E | gem. m | spreiding s | 2de telkamerdeel 5 foto's F | G | H | I | K | gem. m | spreiding s |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2-7 | - | 1 | 323 | 297 | 306 | 315 | 279 | 304 | 15,2 | 286 | 302 | 274 | 293 | 314 | 294 | 13,6 |
| | 3-7 | - | 2 | 282 | 293 | 318 | 312 | 315 | 304 | 14,0 | 298 | 287 | 320 | 305 | 310 | 304 | 11,1 |
| | 5-7 | 8 | 3 | 300 | 287 | 311 | 297 | 299 | 299 | 7,7 | 316 | 288 | 288 | 283 | 319 | 299 | 15,4 |
| | 5-7 | 14 | 4 | 287 | 289 | 295 | 262 | 291 | 285 | 11,7 | 288 | 266 | 300 | 266 | 300 | 284 | 15,3 |
| | 5-7 | 20 | 5 | 289 | 288 | 288 | 293 | 318 | 295 | 11,6 | 314 | 312 | 299 | 307 | 296 | 306 | 7,1 |
| | 6-7 | 2 | 6 | 298 | 302 | 289 | 324 | 317 | 306 | 13,8 | 318 | 304 | 310 | 326 | 306 | 313 | 8,2 |
| 1 | 6-7 | 8 | 7 | 283 | 296 | 306 | 324 | 306 | 303 | 13,5 | 330 | 294 | 301 | 313 | 322 | 312 | 13,2 |
| | 6-7 | 14 | 8 | 239 | 261 | 251 | 247 | 252 | 250 | 7,2 | 228 | 231 | 242 | 253 | 236 | 238 | 8,9 |
| | 6-7 | 20 | 9 | 320 | 310 | 301 | 306 | 302 | 308 | 6,9 | 321 | 281 | 299 | 308 | 292 | 300 | 13,6 |
| | 7-7 | 2 | 10 | 323 | 338 | 347 | 351 | 322 | 336 | 12,0 | 324 | 296 | 318 | 342 | 321 | 320 | 14,7 |
| | 7-7 | 8 | 11 | 297 | 28[illegible] | 313 | 280 | 282 | 291 | 12,7 | 286 | 275 | 295 | 307 | 304 | 293 | 11,8 |
| | 7-7 | 14 | 12 | 302 | 281 | 273 | 276 | 282 | 283 | 10,15 | 296 | 268 | 301 | 271 | 288 | 285 | 13,2 |
| | 7-7 | 20 | 13 | 304 | 287 | 288 | 283 | 295 | 291 | 7,4 | 284 | 285 | 294 | 297 | 304 | 293 | 7,5 |
| | | | | | | | TELKAMER II. | | | | | | | | | | |
| | 2-7 | - | 1 | 317 | 300 | 312 | 291 | 296 | 303 | 9,8 | 311 | 294 | 286 | 306 | 287 | 297 | 10,1 |
| | 2-7 | - | 2 | 290 | 274 | 281 | 291 | 308 | 289 | 11,5 | 292 | 281 | 288 | 289 | 294 | 289 | 4,45 |
| | 5-7 | 8 | 3 | 270 | 258 | 294 | 278 | 275 | 275 | 11,7 | 258 | 243 | 253 | 267 | 262 | 257 | 8,2 |
| | 5-7 | 14 | 4 | 281 | 266 | 290 | 290 | 265 | 278 | 11,0 | 289 | 290 | 262 | 261 | 275 | 275 | 12,5 |
| | 5-7 | 20 | 5 | 280 | 276 | 278 | 291 | 268 | 279 | 7,4 | 282 | 304 | 302 | 310 | 296 | 299 | 9,5 |
| | 6-7 | 2 | 6 | 301 | 305 | 286 | 308 | 311 | 302 | 8,8 | 296 | 297 | 288 | 314 | 295 | 298 | 8,6 |
| | 6-7 | 8 | 7 | 284 | 269 | 277 | 290 | 286 | 281 | 7,4 | 271 | 273 | 286 | 304 | 287 | 284 | 11,9 |
| 2 | 6-7 | 14 | 8 | 276 | 265 | 281 | 262 | 262 | 269 | 7,8 | 282 | 279 | 252 | 281 | 276 | 274 | 11,2 |
| | 6-7 | 20 | 9 | 311 | 296 | 312 | 279 | 282 | 296 | 13,9 | 293 | 268 | 285 | 264 | 264 | 275 | 12,0 |
| | 7-7 | 2 | 10 | 301 | 294 | 271 | 303 | 273 | 288 | 13,7 | 280 | 286 | 296 | 300 | 291 | 291 | 7,1 |
| | 7-7 | 8 | 11 | 293 | 277 | 277 | 271 | 274 | 278 | 7,6 | 276 | 274 | 267 | 288 | 269 | 275 | 7,4 |
| | 7-7 | 14 | 12 | 264 | 277 | 280 | 294 | 294 | 282 | 11,3 | 287 | 267 | 274 | 274 | 262 | 273 | 8,4 |
| | 7-7 | 20 | 13 | 288 | 285 | 300 | 311 | 287 | 294 | 9,9 | 309 | 287 | 295 | 305 | 296 | 298 | 7,8 |
| | | | | | | | TELKAMER III. | | | | | | | | | | |
| | 2-7 | - | 1 | 268 | 260 | 266 | 280 | 259 | 267 | 7,5 | 281 | 268 | 280 | 263 | 267 | 272 | 7,3 |
| | 3-7 | - | 2 | 254 | 265 | 258 | 277 | 277 | 266 | 9,5 | 249 | 279 | 248 | 260 | 257 | 259 | 11,2 |
| | 5-7 | 8 | 3 | 229 | 240 | 244 | 263 | 228 | 241 | 12,7 | 254 | 237 | 231 | 239 | 259 | 244 | 10,7 |
| | 5-7 | 14 | 4 | 236 | 248 | 261 | 234 | 232 | 242 | 10,7 | 243 | 253 | 240 | 242 | 237 | 243 | 5,4 |
| | 5-7 | 20 | 5 | 248 | 259 | 254 | 247 | 232 | 248 | 9,1 | 255 | 256 | 268 | 271 | 266 | 263 | 6,5 |
| | 6-7 | 2 | 6 | 269 | 269 | 263 | 268 | 286 | 271 | 7,8 | 265 | 257 | 266 | 269 | 268 | 265 | 4,2 |
| | 6-7 | 8 | 7 | 247 | 231 | 228 | 251 | 240 | 239 | 8,9 | 237 | 253 | 250 | 246 | 261 | 249 | 7,9 |
| 3 | 6-7 | 14 | 8 | 231 | 228 | 233 | 232 | 245 | 254 | 5,9 | 228 | 224 | 235 | 241 | 229 | 231 | 6,0 |
| | 6-7 | 20 | 9 | 246 | 254 | 258 | 256 | 247 | 252 | 4,8 | 246 | 247 | 262 | 259 | 245 | 252 | 7,2 |
| | 7-7 | 2 | 10 | 245 | 259 | 252 | 274 | 264 | 259 | 9,9 | 249 | 243 | 252 | 252 | 259 | 251 | 5,2 |
| | 7-7 | 8 | 11 | 230 | 232 | 217 | 229 | 212 | 224 | 8,0 | 232 | 199 | 212 | 215 | 221 | 216 | 10,8 |
| | 7-7 | 14 | 12 | 242 | 250 | 259 | 233 | 244 | 246 | 8,7 | 234 | 248 | 227 | 227 | 256 | 238 | 11,7 |
| | 7-7 | 20 | 13 | 281 | 269 | 266 | 279 | 271 | 273 | 5,8 | 259 | 262 | 266 | 253 | 255 | 259 | 4,7 |

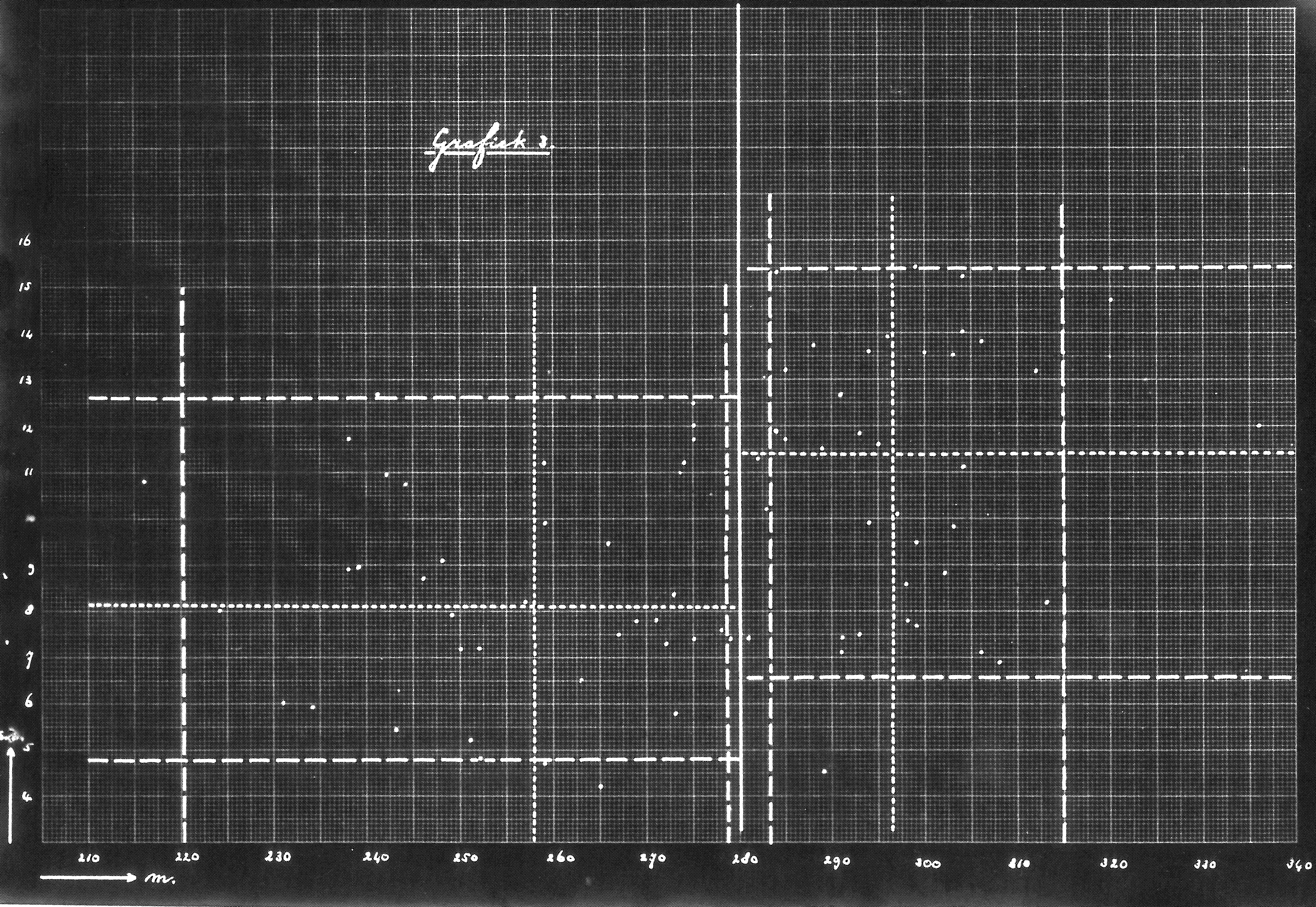
## Grafiek 1.

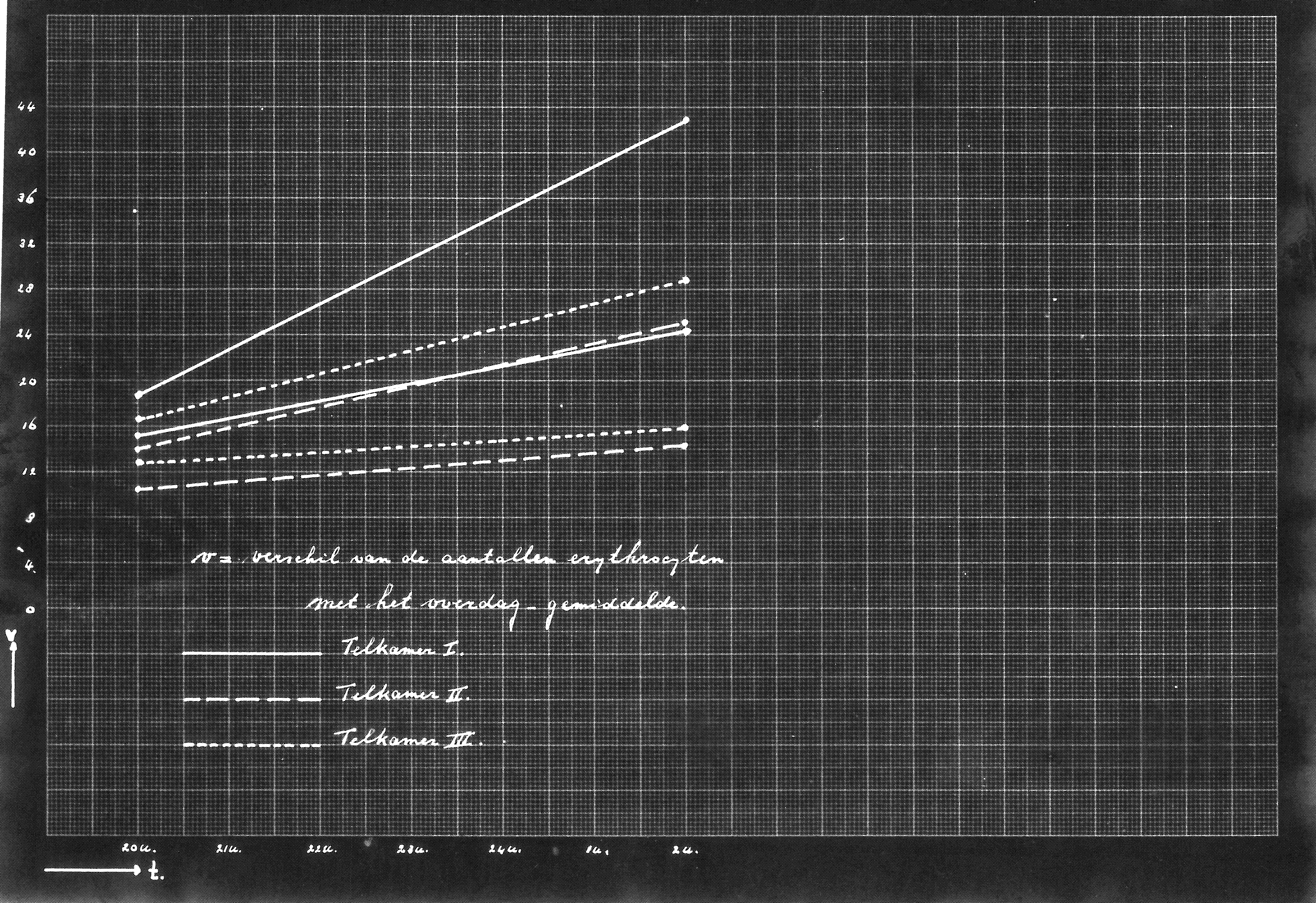


Doorzenigmethode.
(B+C)-(A+D)
Normale-methode.
(B+C)-(A+D)
Schaal.
-100 -90 -80 -70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70 80 90 100



Grafiek 2.
16
13
12
11
10
9
8
7
6
5
4
210
220
230
240
250
260
270
280
290
300
310
320
330
340
m


Grafiek 3.
16
15
14
13
12
11
9
8
7
6
5
4
210
220
230
240
250
260
270
280
290
300
310
320
330
340
m.


44
40
36
32
28
24
20
16
12
8
4
0
v
v = verschil van de aantallen erythrocyten met het overdag-gemiddelde.
Telkamer I.
Telkamer II.
Telkamer III.
20u.
21u.
22u.
23u.
24u.
1u.
2u.
t.

MATHEMATISCH CENTRUM
Amsterdam
Statistische Afdeling.
S47(M6)

# Algemene gang van zaken bij het toetsen van een hypothese.[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 toetsingsgrootheid $\underline{u}$ (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 $\alpha$, zodat Z dus van $\alpha$ afhankelijk is[3)]. Z heet de kritieke zône van de toets, $\alpha$ de onbetrouwbaarheidsdrempel (Engels: level of significance). Voor $\alpha$ 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 "significant" is. De waarde van $\alpha$ moet dan echter worden vermeld. De kans, dat dit zal gebeuren, is, indien $\mathcal{H}_0$ juist is, gelijk aan $\alpha$. Derhalve is $\alpha$ de kans op ten onrechte verwerping van de juiste hypothese, 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, waarin de hypothese 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 collectie (universum, populatie) gedefinieerd is en daarop allerlei waarden aanneemt. Stochastische grootheden worden aangegeven door onderstreepte letters.

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

De toetsingstheorie biedt in het algemeen <u>geen</u> mogelijkheid om tot <u>aanvaarding</u> van een hypothese te komen. Indien een bepaalde hypothese $H_0$ 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 <u>nevenvoorwaarden</u> dienen steeds uitdrukkelijk te worden vermeld en, zo mogelijk, zelf te worden getoetst.

In plaats van de onbetrouwbaarheidsdrempel $\alpha$ wordt vaak bij de uitslag van een toetsing de <u>overschrijdingskans</u> $k$ opgegeven; dit is de kleinste waarde van $\alpha$, waarbij in het betrokken geval, nog tot verwerping van $H_0$ zou zijn overgegaan; anders gezegd: de kleinste $\alpha$, waarvoor de gevonden waarde der toetsingsgrootheid nog juist in de (bij $\alpha$ behorende) kritieke zône Z ligt. <u>Wordt dus de waarde $k$ opgegeven en werkt men met onbetrouwbaarheidsdrempel $\alpha$, dan wordt verworpen, indien $k \leq \alpha$ is.</u>

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 $H_0$ leidt, maar dat deze slechts onder bijzondere omstandigheden kan worden toegepast.

<u>Litteratuur:</u>

J.Neyman, <u>First</u> 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 <u>4</u> (1950) p.54-66.

MATHEMATISCH CENTRUM
Amsterdam
Statistische Afdeling.

S47 (M7)

## De toets van Wilcoxon.[1)]

Deze methode dient tot het toetsen van de hypothese $H_0$, 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 $U$ 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 $U$ alle gehele waarden van 0 tot en met $n.m$ aannemen. Is $U$ groot, dan liggen er veel waarden $x_i$ verder naar rechts dan waarden $y_j$, is $U$ klein, dan juist weinig.

De kritieke zône $K$ neemt nu daarom de kleine en de grote waarden van $U$ en wel van beide zoveel, dat de gekozen onbetrouwbaarheidsdrempel $\alpha$ niet overschreden wordt.

Voor éénzijdige toetsing, te onderscheiden in linker- en rechter-éénzijdige toetsing, gebruikt men kritieke zônes $K_1$, resp. $K_2$, die geheel bestaan uit kleine, resp. grote waarden van $U$.

Verwerping van $H_0$ ten gevolge van het vinden van een grote (resp. kleine) waarde van $U$ 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 ($M_4$)(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.

MATHEMATISCH CENTRUM
Amsterdam.
Statistische Afdeling.
S47(M3)

## 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].

$H_0$ (te toetsen hypothese): Het gemiddelde van $\underline{x}$ bezit de waarde $\mu$; $\mu$ 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 $m = \dfrac{x_1 + x_2 + \ldots\ldots + x_n}{n}$ en $s' = \sqrt{\dfrac{(x_1-m)^2 + \ldots (x_n-m)^2}{n-1}}$ is.

Indien $H_0$ 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 $\mu$, dan zullen verder van 0 af liggende waarden vaker voorkomen dan indien $H_0$ 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 $\alpha$.

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 $\nu$) 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:

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

MATHEMATISCH CENTRUM
Amsterdam
Statistische Afdeling.

S47(M10)

## Symmetrietoets.[1]

Hypothese $H_0$: de waarnemingen $z_1, \ldots, z_{n'}$ zijn afkomstig van onafhankelijk verdeelde stochastische grootheden, die alle symmetrisch ten opzichte van 0 verdeeld zijn.[2] Van deze toets bestaan meerdere versies $T_1, \ldots, T_6$. We bespreken eerst $T_1$ en $T_2$.

Toetsingsgrootheden. Deze worden als volgt uit $z_1, \ldots, z_{n'}$ afgeleid:

1e de waarnemingen, die gelijk aan 0 zijn worden weggelaten. Stel er blijven over: $z_1, \ldots, z_n$

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 $h$

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

De toetsingsgrootheden zijn $n_1$ en $u$, $h$ is een hulpgrootheid.

V.B. 22 waarden $z_i$: 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/-6,3/-7,0/-7,9/-8,0/-8,7.

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

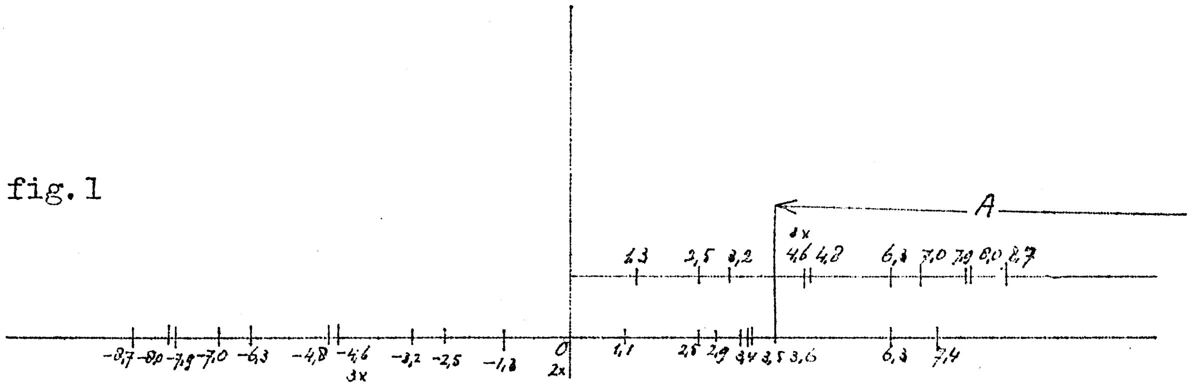
$u = 2$

---

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

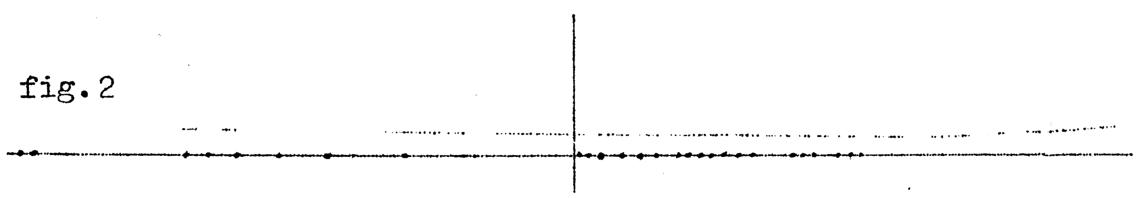
2) Zetten wij hier in plaats van 0, dan geldt voor

fig. 1




Kritieke zônes. 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 positieve (resp. negatieve) waarden zijn, die verder van 0 verwijderd liggen dan de negatieve (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.

$T_3$ en $T_4$.

Toetsingsgrootheden.

1e, 2e en 3e als boven. Eerste toetsingsgrootheid: $n_1$

4e: op $x_1, \ldots, x_{n_1}$ en $y_1, \ldots, y_{n_2}$ wordt de toets van Wilcoxon toegepast (vgl. S47(M8)). De toetsingsgrootheden zijn $n_1$ en de $U$ van deze toets van Wilcoxon.

Kritieke zônes. 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 $\underline{U}$ 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, Excelsior.

MATHEMATISCH CENTRUM
2de Boerhaavestr. 49
A M S T E R D A M O.
Statistische Afdeling
S47 (M11)

## 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 dekwadranten rechts boven en links onder van een + teken en de beide andere van een - teken voorzien.

(zie fig. 1.)

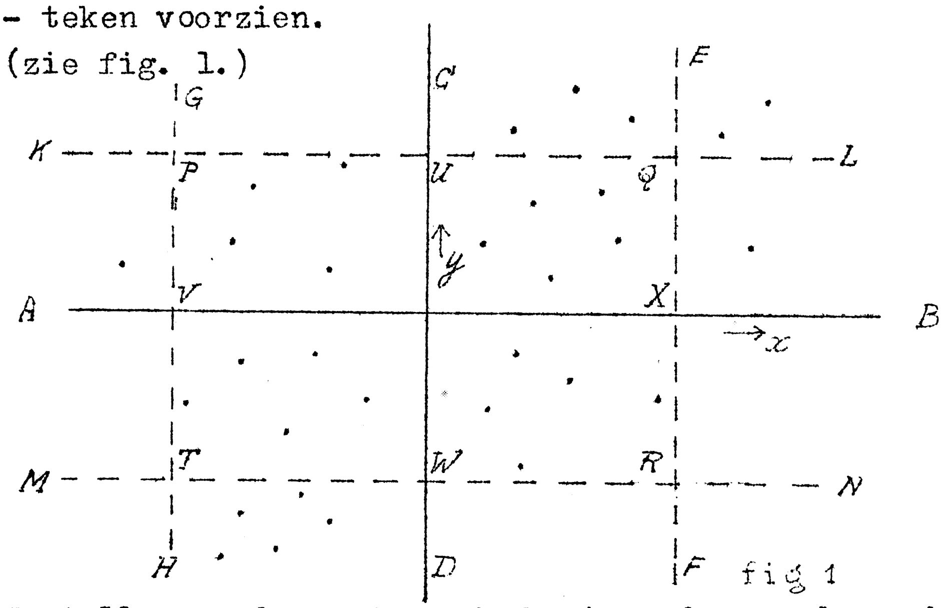



fig 1

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 op klimmende $x$ 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 $S$ van deze aantallen geïsoleerde punten. Onder deze kwadrantensom 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.

Vb. in fig. 1. wordt :

| | | |
|---|---|---|
| Hoek | $CUL$ | $+5$ |
| ,, | $EXB$ | $+3$ |
| ,, | $GVA$ | $-1$ |
| ,, | $MWD$ | $+5$ |
| | | $S = +13$ |

De verdelingsfunctie van $S$ is bekend en vrijwel onafhankelijk van het aantal waarnemingen. Is aan de onafhankelijkheid voldaan, dan kan men betrekkelijk kleine waarden van $|S|$ verwachten. Is er afhankelijkheid, dan is er meer kans op grote waarden van $|S|$. Daarom neemt men kritieke zônes van de vorm $|S| \geq |S_0|$. Voor een onbetrouwbaarheidsdrempel $\alpha = 0{,}05$ is $S_0 = 11$, voor $\alpha = 0{,}01$ is $S_0 = 19$.
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.

Litteratuur

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.

MATHEMATISCH CENTRUM
2de Boerhaavestraat 49
A M S T E R D A M O.
Statistische Afdeling
S47 (M12)

## De $\chi^2$-toets voor aanpassing 1)

De $\chi^2$-toets voor aanpassing dient voor het toetsen van de hypothese $H_0$, dat een stelsel waarden $x_1, x_2, ---, x_n$ een steekproef is uit een continue waarschijnlijkheidsverdeling met verdelingsdichtheid $f(x)$. Hierbij wordt een grootheid $\chi^2$ waarvan de verdelingsfunctie bekend is, als toetsingsgrootheid gebruikt.
De verdelingsfunctie van $\chi^2$ wordt nog nader gekarakteriseerd door een grootheid $\nu$ welke het aantal vrijheidsgraden van $\chi^2$ aangeeft. De grootte van $\chi^2$ en het getal $\nu$ worden op de volgende wijze bepaald:

Wanneer $f(x)$ de onderstelde verdelingsdichtheid is (fig.1)

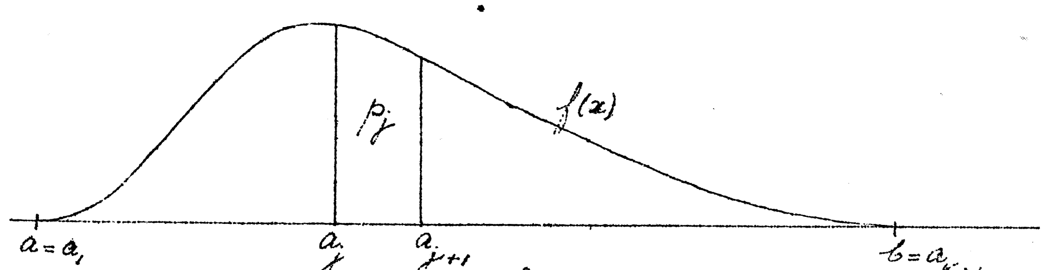


fig.1

wordt het interval $[a, b]$ in $K$ stukken verdeeld, zodanig dat boven ieder interval $[a_j, a_{j+1}]$ een oppervlak $p_j$ ligt $(j=1, \ldots, K)$ Daar het gehele oppervlak onder de kromme de totale waarschijnlijkheidsmassa voorstelt, dus gelijk aan 1 is, geldt de betrekking $p_1 + p_2 + \cdots\cdots + p_K = 1$
Bestaat nu de steekproef uit $n$ waarnemingen $x_1, x_2, \ldots\ldots, x_n$ dan zal men onder de hypothese $H_0$ in het algemeen verwachten, dat ongeveer $np_j$ van de $n$ steekproef-waarnemingen in het interval $[a_j, a_{j+1}]$ komen te liggen. Is het werkelijke aantal waarnemingen uit de steekproef, dat in het interval $[a_j, a_{j+1}]$ valt $n_j$ (dus $n_1$ waarden in $[a_1, a_2]$ etc.) dan wordt $\chi^2$ als volgt gedefinieerd:

$$\chi^2 = \frac{(n_1 - np_1)^2}{np_1} + \frac{(n_2 - np_2)^2}{np_2} + \cdots\cdots\cdots + \frac{(n_K - np_K)^2}{np_K}$$

---

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

Deze $\chi^2$ is een maat voor de afwijking van de werkelijke verdelingsdichtheid van $x$ van de ondersteldc verdelingsdichtheid $f(x)$. Is de werkelijke verdelingsdichtheid van $x$ inderdaad gelijk aan $f(x)$ zoals volgens $H_0$ ondersteld wordt, dan zal $\chi^2$ in het algemeen klein zijn; is de verdelingsdichtheid echter verschillend van $f(x)$, dan worden grote waarden van $\chi^2$ waarschijnlijker dan het geval is, als $H_0$ juist is. De kritieke zône voor deze toets geeft men daarom de gedaante $\chi^2 > \chi_0^2$, waarbij $\chi_0^2$ correspondeert met een onbetrouwbaarheidsdrempel $\alpha$ en uit tabellen voor de $\chi^2$-verdeling opgezocht kan worden, waarbij rekening gehouden moet worden met het in de tabellen aangegeven.

<u>Aantal vrijheidsgraden</u> $\nu$. Dit getal wordt als volgt bepaald: zijn er $k$ parameters van $f(x)$ (b.v. het gemiddelde de spreiding e.d.) uit de steekproef bepaald, dan is $\nu = n-k-1$, waarin $n$ het aantal waarnemingen is. Is $f(x)$ van tevoren geheel gespecifieerd, dan is dus $\nu = n-1$

<u>Voorbeelden</u>: bij toetsing op bovenstaande wijze van de hypothese, dat de steekproef $x_1, \ldots, x_n$ uit een normale verdeling ("verdeling van Gauss") afkomstig is, worden gemiddelde en spreiding van de steekproef gebruikt om $f(x)$ te bepalen. Door gemiddelde en spreiding is de normale verdeling geheel bepaald. Dan is dus $\nu = n-3$. Wenst men echter de hypothese te toetsen, dat de steekproef afkomstig is uit een volledig gegeven normale verdeling, waarbij dus gemiddelde en spreiding ook gegeven zijn, dan behoeven deze parameters niet uit de steekproef geschat te worden en is .

<u>Opm.</u> In de practijk worden de getallen $p_1, p_2, .. p_k$ meestal even groot gekozen, dus alle gelijk $\frac{1}{k}$. Het aantal delen $k$ laat men bovendien nog afhangen van het aantal waarnemingen, waaruit de steekproef bestaat. De benadering, die bij het toepassen van de $\chi^2$-toets gebruikt wordt, is alleen dan behoorlijk indien het aantal $k$ zodanig gekozen wordt, dat het verwachte aantal waarden in de intervallen $[a_j, a_{j+1}]$ d.i. $np_j$ niet al te klein is, bv. $\geqq 10$

Litteratuur:

M.G.Kendall, The advanced Theory of Statstics.
London 1946, Deel II hfdst. 12

MATHEMATISCH CENTRUM
2de Boerhaavestr. 49
A M S T E R D A M   O.
Statistische Afdeling
S47 (M14)

## Methode der $m$ rangschikkingen [1]

Een duidelijk voorstelling van deze toetsingsmethode verkrijgt men door $n$ elementen te beschouwen, die een bepaald kenmerk, eventueel in verschillende mate, bezitten. dit kenmerk wordt door $m$ waarnemers beoordeeld en ieder van deze waarnemers rangschikt deze $n$ elementen volgens zijn beoordeling naar opklimmende waardering. Op deze wijze ontstaan $m$ rijen van rangschikkingen. We willen nu een maat aangeven voor de overeenstemming tussen deze rangschikkingen, m.a.w. een maat voor de overeenstemming tussen de $m$ beoordelingen. De hypothese $\mathcal{H}_0$ , die met deze methode getoetst kan worden, houdt in dat er geen overeenstemming tussen de waarnemers bestaat; precieser gezegd, dat alle rangschikkingen onafhankelijk van elkaar op toevallige wijze zijn ontstaan. Dit is b.v. het geval, als het betrokken kenmerk in werkelijkheid voor alle elementen dezelfde waarde bezit.

We kunnen de afleiding voor de maat van overeenstemming het eenvoudigst geven aan de hand van een voorbeeld.

| elementen | | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|
| rangnummers toegekend door waarnemer | a | 5 | 4 | 1 | 6 | 3 | 2 |
| | b | 2 | 3 | 1 | 5 | 6 | 4 |
| | c | 4 | 1 | 6 | 3 | 2 | 5 |
| | d | 4 | 3 | 2 | 5 | 1 | 6 |
| | | 15 | 11 | 10 | 19 | 12 | 17 |

De som van alle rangnummers is $\frac{1}{2} n\, m(n+1)$ . Onder de hypothese $\mathcal{H}_0$ is het theoretische gemiddelde van iedere kolom: $\frac{1}{2} m(n+1)$

We beschouwen nu de afwijkingen van dit gemiddelde. In ons voorbeeld is het theoretisch kolomgemiddelde gelijk aan 14. De afwijkingen daarvan zijn

$$1 \quad -3 \quad -4 \quad 5 \quad -2 \quad 3$$

---

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

De som der kwadraten van deze afwijkingen noemen we $S$
In ons voorbeeld is $S = 64$
Als alle $m$ rangschikkingen gelijk zijn wordt het maximum van $S$ bereikt.
Dit maximum is $\frac{1}{12} m^2(n^3-n)$
We definiëren nu als <u>coëfficiënt van overeenstemming</u>

$$W = \frac{12\, S}{m^2(n^3-n)}$$

In ons voorbeeld is $W = \frac{12 \times 64}{16 \times 210} = 0{,}229$
$W$ varieert dus tussen $0$ en $1$
De verdeling van $W$ onder de hypothese $H_0$ is exact berekend voor een aantal waarden van $n$ en $m$ [1] [illegible] terwijl voor grote $n$ en $m$ benaderingen bekend zijn. [1] Hiermee kan $H_0$ dus getoetst worden, waarbij $H_0$ verworpen wordt, als $W$ waarden dichtbij $1$ aanneemt. De kritieke zône is dus van de vorm $W \geqq W_0$
Het kan voorkomen dat de waarnemers geen onderscheid ontdekken in de mate waarin verschillende elementen het kenmerk bezitten. Ze geven deze elementen dan gelijke rangnummers.
Veronderstel, dat door een waarnemer geen onderscheid wordt gemaakt tussen de elementen, die de rangnummers $3$ t/m $6$ moeten dragen. Dan wordt als rangnummer van iedere van deze elementen het gemiddelde van de rangnummers $\frac{1}{4}(3+4+5+6) = 4\frac{1}{2}$
gebruikt.
Daar het maximum van $S$ nu verandert, moeten we een correctie op de formule van $W$ toepassen; deze vindt men in het boek van Kendall behandeld.

Litteratuur:

[1] M.G.Kendall, Rank correlation methods, London 1948
hoofdstuk 6 p.80.

tabel verdelingsfunctie van $W$ voor

$n = 3$ $m = 2$ t/m $10$
$n = 4$ $m = 2$ t/m $6$
$n = 5$ $m = 3$
op p. 146-149

MATHEMATISCH CENTRUM,
2de Boerhaavestraat 49,
A m s t e r d a m - O.
Statistische Afdeling.



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.

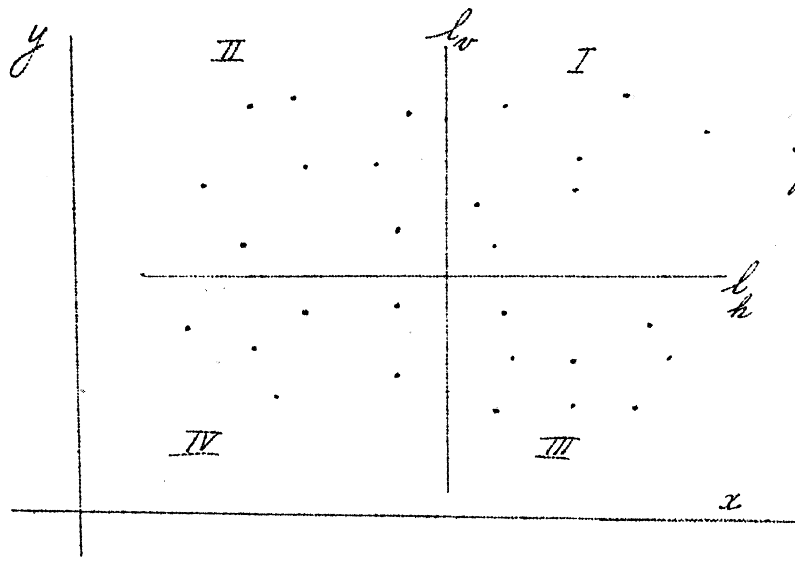



fig 1

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 $\bar{A}$ '') (rechts resp. links van verticale streep) en B of $\bar{B}$ (boven resp. onder horizontale streep). Er ontstaan dus vier groepen elementen met de kenmerken

| $AB$ | $A\bar{B}$ | $\bar{A}B$ | $\bar{A}\bar{B}$ |
|---|---|---|---|
| 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.

'') $\bar{A}$ 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 $\sigma_{\underline{a}}^2 = \frac{n\,m\,r\,s}{N^2(N-1)}$
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.

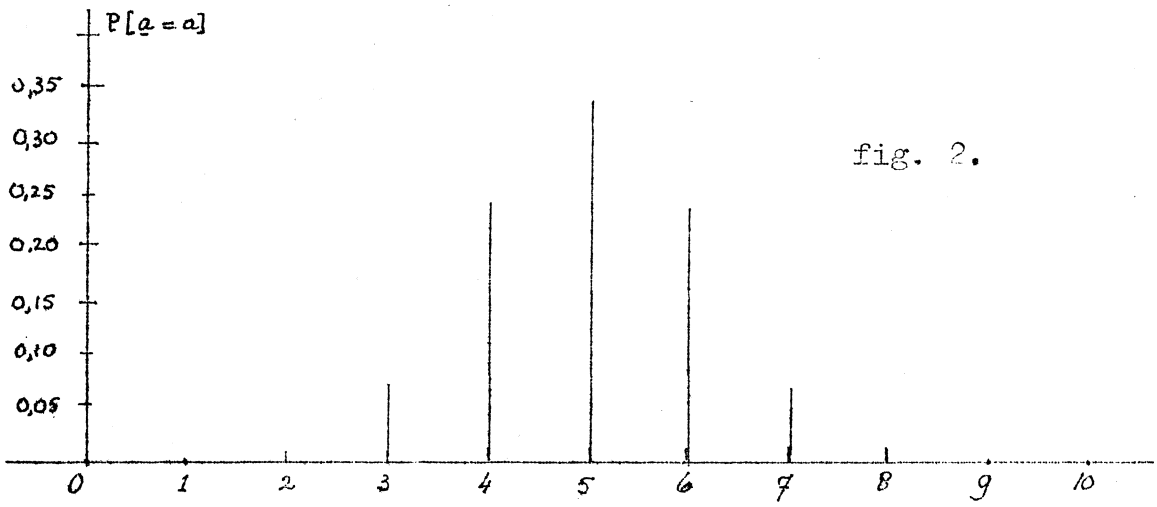



fig. 2.

mogelijkheden:

| 0 \| 10 | 1 \| 9 | 2 \| 8 | 3 \| 7 | 4 \| 6 | 5 \| 5 | 6 \| 4 | 7 \| 3 | 8 \| 2 | 9 \| 1 | 10 \| 0 |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 \| 0 | 9 \| 1 | 8 \| 2 | 7 \| 3 | 6 \| 4 | 5 \| 5 | 4 \| 6 | 3 \| 7 | 2 \| 8 | 1 \| 9 | 0 \| 10 |

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) [i])

[i]) 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 $< \alpha$ ($\alpha$: 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 $a, b, n-a, m-b$ 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.

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
A m s t e r d a m - O.

Statistische Afdeling
S 47 (M 16)

## Normaliteitstoetsen van Geary en Pearson[1].

Om de hypothese $\mathcal{H}_0$ te toetsen, dat een reeks waarnemingen $x_1, x_2, \ldots . x_n$ beschouwd kan worden als een steekproef uit een normale verdeling, d.i. een verdeling met verdelingsdichtheid

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

zijn door Geary en Pearson de verdelingen, onder hypothese $\mathcal{H}_0$, berekend van enkele statistische grootheden, welke betrekking hebben op de vorm van de kromme $f(x)$.

De eerste grootheid is de volgende:

$$(1) \quad \underline{a} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{\sqrt{n \sum_{i=1}^{n}(x_i - \bar{x})^2}} \qquad \text{waarin } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ is.}$$

Deze grootheid $a$ is een maat voor de "kurtosis" ("platheid" of "slankheid" der kromme $f(x)$). In figuur 1 zijn 3 verdelingsdichtheden $f(x)$, $f_1(x)$ en $f_2(x)$ getekend waarvan $f(x)$ een normale is, terwijl $f_1(x)$ te slank is (en te dik in de staarten), en $f_2(x)$ te plat (en te dun in de staarten). Pearson noemt $f_1(x)$ "leptocurtic" en $f_2(x)$ "platycurtic".

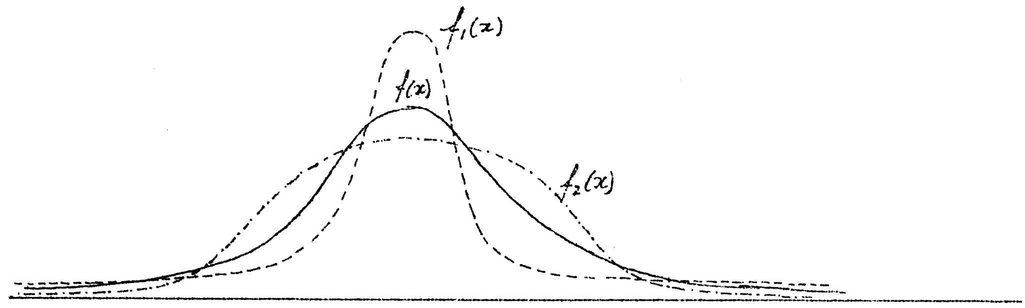



fig. 1

De kritieke zône voor $\underline{a}$ bestaat uit grote en kleine waarden van $a$.

---

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

Als tweede toetsingsgrootheid wordt gebruikt:

$$(2) \qquad \sqrt{\underline{b}_1} = \sqrt{n}\, \frac{\sum_{i=1}^{n} (\underline{x}_i - \underline{\bar{x}})^3}{\left\{\sum_{i=1}^{n} (\underline{x}_i - \underline{\bar{x}})^2\right\}^{\frac{3}{2}}}$$

Deze grootheid is een maat voor de "scheefheid" (Engels: "skewness") van de verdeling. In figuur 2 zijn 3 verdelingsdichtheden getekend, waarvan $f(x)$ symmetrisch is, terwijl $f_1(x)$ "positief scheef" (omdat de "dikke staart" rechts ligt) en $f_2(x)$ "negatief scheef" is.

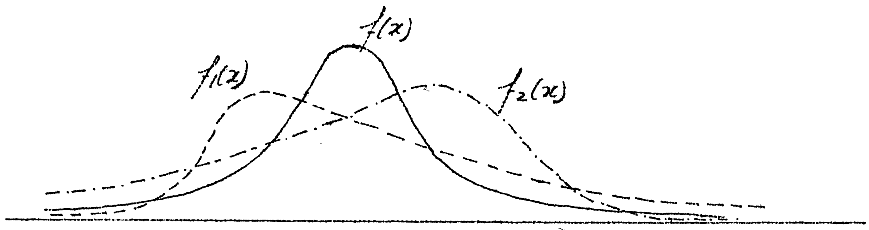



fig. 2

De kritieke zône voor de grootheid $\sqrt{b_1}$, bestaat uit grote en kleine waarden van $\sqrt{b_1}$, (daarbij wordt gewoonlijk ook voor negatieve waarden de notatie $\sqrt{b_1}$ gebruikt).

De derde toetsingsgrootheid

$$(3) \qquad \underline{b}_2 = \frac{n \sum_{i=1}^{n} (\underline{x}_i - \underline{\bar{x}})^4}{\left\{\sum_{i=1}^{n} (\underline{x}_i - \underline{\bar{x}})^2\right\}^2}$$

is weer een maat voor de "kurtosis" der kromme en wordt alleen voor zeer grote steekproeven berekend. Voor kleinere steekproeven is de maat $a$ voldoende.

De kritieke zône bestaat uit grote en kleine waarden van $b_2$.

<u>Nomogrammen der 3 grootheden:</u>

1e. Voor de grootheid $a$ zijn onder de hypothese $\mathcal{H}_0$ nomogrammen en tabellen berekend voor de onbetrouwbaarheidsdrempels 0,01; 0,05 en 0,10, waarbij $10 \leqq n \leqq 1000$.

2e. Voor de grootheid $\sqrt{b_1}$ zijn nomogrammen berekend voor de onbetrouwbaarheidsdrempels 0,01 en 0,05, $25 \leqq n \leqq 1000$.

3e. Voor de grootheid $\sqrt{b_2}$ zijn nomogrammen berekend voor de onbetrouwbaarheidsdrempels 0,01 en 0,05. Hierbij is $100 \leqq n \leqq 1000$.

Literatuur:

R.C.Geary, Moments of the ratio of the mean deviation to the standard deviation for normal samples, Biometrika 28 (1936) p.295.

R.C.Geary and E.S.Pearson, Tests of Normality, Biometrika Office, London 1938.

Bovenstaande nomogrammen en tabellen komen in deze beide publicaties voor.
Alle daarin genoteerde overschrijdingskansen zijn éénzijdig.

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
Amsterdam-O.
Statistische Afdeling
S 47 (M 17)

# Transformatie op een homogene verdeling.[1]

Ter toetsing van verschillende hypothesen kan men gebruik maken van een transformatie, waardoor een stochastische grootheid $\underline{x}$ met gegeven waarschijnlijkheids-verdeling, overgaat in een homogeen verdeelde grootheid $\underline{x}'$. Dit is een grootheid, waarvoor geldt:

$$(1)\quad P[\underline{x}' \leqq x'] = \begin{cases} 0 & \text{voor } x' < 0 \\ x' & \text{voor } 0 \leqq x' \leqq 1 \\ 1 & \text{voor } x' > 1 \end{cases}$$

dus met verdelingsdichtheid $h(x')$ (frequentiekromme) als in figuur 1

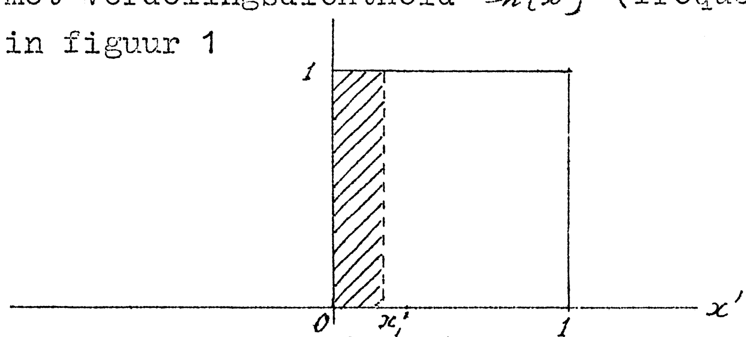


fig. 1

Indien nu de verdelingsdichtheid $f(x)$ van $\underline{x}$ gegeven wordt door de kromme van figuur 2:

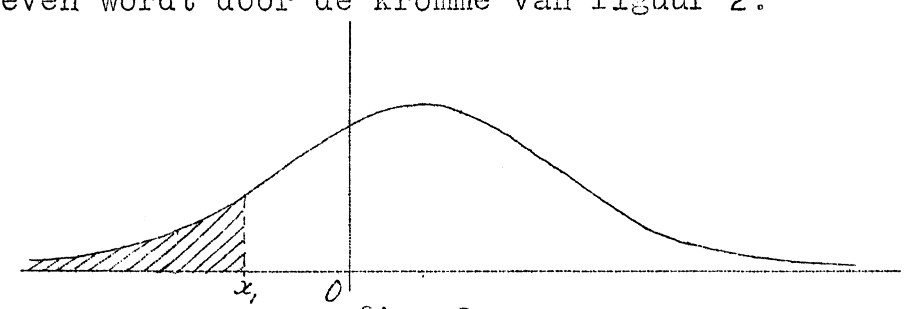


fig. 2

en we noemen de oppervlakte links van een waarde $x_1$ onder de kromme gelegen $F(x_1)$, dan bestaat deze transformatie op een homogene verdeling daaruit, dat $x_1$ vervangen wordt door dat punt $x_1'$ (zie fig. 1), waarvoor geldt

$$(2)\qquad x_1' = F(x_1)$$

zodat links van $x_1'$ dezelfde oppervlakte onder de kromme $h(x')$ ligt, als links van $x_1$ onder $f(x)$.

Men kan deze transformatie ook zeer duidelijk voorstellen door figuur 3, waarin $F$ de verdelingsfunctie van $\underline{x}'$ en $G$ die van $\underline{x}$ voorstelt.

---

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

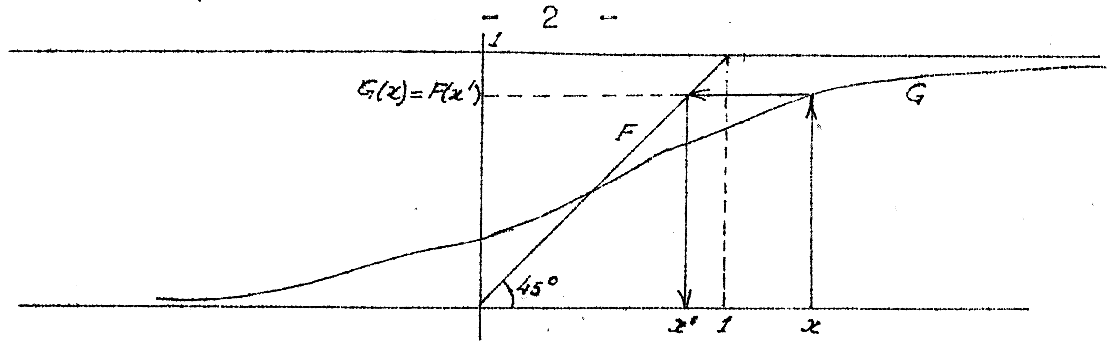



fig. 3

Voor een punt $x$ is dus $G(x) = P[\underline{x} \leq x]$ . Zoeken wij nu het punt $\underline{x}'$, waarvoor $x' = F(x') = G(x)$ is, dan is dit het bij $x$ behorende getransformeerde punt. De transformatie is in de figuur door een drietal pijlen aangegeven.

Willen wij nu de hypothese $\mathcal{H}_0$ toetsen, dat een aantal waarden $x_1, \ldots\ldots x_n$ een steekproef vormen uit een gegeven verdeling met verdelingsdichtheid $f(x)$ , dan transformeren wij deze steekproef door de transformatie:

(3) $x_1' = F(x_1)$ , . . . . . . . . . , $x_n' = F(x_n)$

en toetsen in plaats van $\mathcal{H}_0$ , de hypothese $\mathcal{H}_0^*$, dat $x_1', \ldots . x_n'$ een steekproef uit een homogene verdeling vormen.

Deze toets verloopt als volgt:
Vorm het product der waarden $x_1', \ldots . x_n'$ en zoek hiervan de natuurlijke logarithme op. Het getal

(4) $Q = -2\{\ln x_1' + \ln x_2' + \ldots\ldots\ldots \ln x_n'\}$ is

de bij de steekproef behorende waarde van een statistische grootheid $Q$ , die, zoals E.S.Pearson bewezen heeft, een $\chi^2$-verdeling bezit met $2n$ vrijheidsgraden.

Is dus $Q \geqq Q_0$ , waarin $Q_0$ een van $n$ en de onbetrouwbaarheidsdrempel $\alpha$ afhankelijke grootheid is, die getabelleerd is, dan wordt $\mathcal{H}_0$ verworpen.

Literatuur:

E.S.Pearson, The Probability-Integral Transformation for testing goodness of fit and combining independent tests of significance. .
Biometrika 30 (1938 ) p.134-148.

MATHEMATISCH CENTRUM
Amsterdam.
Statistische Afdeling.

S47(M18)

## Betrouwbaarheidsintervallen (algemeen).[1]

Zij $\underline{x}$ een stochastische grootheid, die een verdelingsfunctie bezit die, op een onbekende parameter $\theta$ na, geheel bekend is ( $\theta$ kan bv. het gemiddelde van $\underline{x}$ zijn, of de spreiding of iets dergelijks), dan kan men de vraag stellen uit een aantal waarnemingen van $\underline{x}$ een schatting voor $\theta$ af te leiden.

Een betrouwbaarheidsinterval $\mathcal{I}$ voor $\theta$ is een interval, waarvan de grenzen afhankelijk zijn van de waarnemingen $x_1, \ldots, x_n$ van $\underline{x}$ , en dat de eigenschap bezit, behoudens een zekere gegeven onbetrouwbaarheid $\alpha$, de juiste waarde van $\theta$ te bevatten. Dit betekent, dat bij een serie bepalingen van betrouwbaarheidsintervallen slechts in ongeveer een fractie $\alpha$ van deze gevallen het interval $\mathcal{I}$ zo zal uitvallen, dat het $\theta$ niet bevat. Hierbij is dus $\theta$ constant en het interval $\mathcal{I}$ veranderlijk (en wel stochastisch). Hierin ligt het grote verschil met een zgn. voorspellingsinterval, d.i. een gegeven vast interval, waar een stochastisch punt met een zekere waarschijnlijkheid in valt.

Het algemene principe ter bepaling van een betrouwbaarheidsinterval is het volgende: zij $\mathcal{T}$ een toets voor de hypothese $\theta = \theta_0$ (vgl. S47(M6)), dan is $\mathcal{I}$ de verzameling van die waarden $\theta_0$ die bij toepassing van $\mathcal{T}$ op grond van de gevonden waarnemingen $x_1, \ldots, x_n$ niet voor verwerping in aanmerking komen. Is $\mathcal{T}$ toegepast met een onbetrouwbaarheidsdrempel $\alpha$, dan is dit ook de onbetrouwbaarheidsdrempel van het betrouwbaarheidsinterval.

Litteratuur:

M.G. Kendall, The Advanced Theory of Statistics, London 1946, deel II, p.62-84.

A.M. Mood, Introduction to the theory of Statistics, London 1950, p.220.

J. Neyman, First course in probability and statistics, N.Y.1950.

---

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