

Statistische Afdeling

Rapport S 70,

door

Mej. G.Grobbe en

Dr J.Hemelrijk Jr.

Onderzoek ener ijkmethode.

Inhoud

	Pag.
1. Inleiding.	3
2. Methode van Irwin en exacte betrouwbaarheidsintervallen.	5
2.1 Notatie.	5
2.2 Het model van Irwin.	5
2.3 Schattingen voor β , α , γ en μ .	8
2.4 Spreiding der schattingen.	9
2.5 Exacte betrouwbaarheidsintervallen.	10
2.6 Numerieke voorbeelden.	12
3. Toetsingsmethoden voor de gemaakte onderstellingen.	16
3.1 Normaliteit.	16
3.2 Homoscedasticiteit.	17
3.3 Lineariteit.	18
3.4 Evenwijdigheid.	20
3.5 Toetsing van $\beta=0$.	22
4. Schatting van de spreiding van de logaritmische van de sterkteverhouding wanneer de homoscedasticiteit niet vervuld is.	24
5. Beantwoording der vragen.	29
5.1 Keuze der methode.	29
5.2 Ijking van meer dan één praeparaat tegelijk.	31
5.3 Aantal doses.	31
5.4 Duur van een ijking en combinatie van de waarnemingen van verschillende weken.	35
6. Opmerkingen en suggesties.	37
6.1 De ongelijkheid der spreidingen.	37
6.2 Oriënterende ijkingen.	37
6.3 Besparing rekenwerk.	38
7. Samenvatting.	39
8. Appendix.	40
8.1 Stochastische onafhankelijkheid van $\underline{Z}-\underline{Y}$ en $\underline{b}$.	40
8.2 Onzuiverheid van $\underline{M}$ als schatting van μ .	41
8.3 Voorwaarden voor de methode van Fieller.	41

8.4 Betrouwbaarheidsintervallen volgens Fieller.	43
8.5 Hulpstelling voor het niet homoscedastische geval.	45
Litteratuur.	47
Bijlagen: S 53 (M 21); S 53 (M 25); S 47 (M 6); S 53(M 24); S 47 (M 9).	

Onderzoek ener ijkmethode.

1. Inleiding.

Over de statistische berekeningen bij biologische ijkingen zoals deze door J.O.IRWIN: "Statistical Method Applied to Biological Assays", Suppl. to the Journal Roy Stat Soc 4 (1937) pp 1-12, beschreven wordt, werden ond de volgende vragen gesteld:

- 1) Is bovengenoemde methode juist?
- 2) Kan deze methode op meerdere praeparaten tegelijk worden toegepast (eventueel met 1 dosis van ieder praeparaat)?
- 3) Hoe is de ~~samenhang~~ tussen de aantallen doses van standaard en praeparaat en de nauwkeurigheid van de uitkomst?
- 4) Welk aantal weken geeft de nauwkeurigste uitkomsten?
- 5) Zijn de waarnemingen van verschillende weken gelijktijdig bruikbaar?

Ten behoeve van dit onderzoek werden ons de resultaten van vijf ijkingen ter beschikking gesteld. Elke ijking berust op de gewichtstoename gedurende enkele weken bij proefdieren, na geregelde toediening van verschillende doses van een standaardpraeparaat (verder kortweg standaard genoemd) en enkele te onderzoeken onbekende praeparaten (kortweg: praeparaten 1,2, enz.). Eén dosis van de standaard of van een praeparaat wordt aan een groep van 12 (een enkele maal 11) proefdieren gegeven. Van de standaard worden 3 tot 5 doses gebruikt, steeds met een verdunningsfactor 2 (de standaarddoses van ijking 5023 liggen onregelmatig), en van de praeparaten 1 of 2 doses.

Bovendien wordt van de standaard nog een dosis 0 gegeven, welke bij de methode van Irwin niet gebruikt wordt, maar na een uitgebreider onderzoek van het verband tussen de logaritmen van de doses en de gewichtstoename (meer doses, meer proefdieren per groep) eventueel gebruikt zou kunnen worden als controle op de bij de praeparaten verkregen uitkomsten. (vgl. § 6.2)

De gewichtstoename worden na een, twee, drie, vier en vijf weken opgenomen.

Doel van een ijking is het bepalen van de sterkteverhouding tussen standaard en praeparaat.

Bij het onderzoek van dit materiaal hebben we gebruik gemaakt van de in de litteratuur beschreven toetsen, welke vaak niet geheel geschikt zijn voor de beschouwde ijkingen, daar de voor deze methoden noodzakelijke voorwaarden niet steeds vervuld zijn. Algemener methoden, waarbij minder voorwaarden gesteld worden, zijn ons voor dit soort ijkingen echter niet bekend.

We zullen eerst in § 2 een overzicht geven van de methode van Irwin en van de afleiding van een methode voor de bepaling van een exact betrouwbaarheidsinterval voor de sterkteverhouding tussen standaard en praeparaat, die door E.C. FIELLER [2] ontwikkeld is. In § 3 worden methoden beschreven om te toetsen in hoeverre de onderstellingen waar de in § 2 behandelde afleidingen op berusten, vervuld zijn en worden deze methoden toegepast op het beschikbare materiaal. Daar bij dit onderzoek voornamelijk een afwijking van homoscedasticiteit (gelijkheid der spreidingen) tussen de waarnemingen uit verschillende groepen geconstateerd wordt, is in § 4 een benadering van de spreiding van de logaritme van de sterkteverhouding afgeleid en, voor de onderhavige gegevens, berekend, die met deze afwijkingen rekening houdt, teneinde een indruk te krijgen van de invloed van deze ongelijkheid der spreidingen op de nauwkeurigheid van de uitkomst. In § 5 vindt men de beantwoording van de op pag. 3 geformuleerde vragen, § 6 en § 7 geven nog enkele opmerkingen en een samenvatting.

Opmerking: de door Irwin gebruikte uitdrukking "log dose-response curve" is in het volgende met "werkingslijn" vertaald.

Ter oriëntering in het gebruikte statistische terminologie en methoden zijn aan het eind van dit rapport een aantal memoranda toegevoegd.

2. Methode van Irwin en exacte betrouwbaarheidsintervallen.

2.1. Notatie

Voor bekende of onbekende parameters of variabelen welke niet stochastisch zijn, gebruiken we Griekse letters. Stochastische grootheden (dat zijn grootheden, die een waarschijnlijkheidsverdeling bezitten) worden met Latijnse letters aangegeven, terwijl bovendien het stochastische karakter uitgedrukt wordt door onderstreping van deze letter. De onderstreping zullen we weglaten, wanneer er geen gevaar voor verwarring over de behandelde theorie bestaat en wanneer we ons in een formule de gevonden waarde voor de betreffende stochastische grootheid gesubstitueerd denken. De mathematische verwachting (het "theoretisch gemiddelde") van een stochastische grootheid wordt aangeduid met het symbool $\underline{E}$, dus b.v.:

$$\underline{E} \underline{x} = \underline{\xi}.$$

Gereduceerde grootheden, d.w.z. stochastische grootheden verminderd met hun mathematische verwachting geven we aan als $\underline{\tilde{x}}$, dus:

$$\underline{\tilde{x}} = \underline{x} - \underline{E} \underline{x} = \underline{x} - \underline{\xi}.$$

2.2. Het model van Irwin.

Bij ijkingen, als door Irwin in het eerste deel van zijn artikel [1] beschreven, wordt ondersteld dat de reactie van een proefdier (b.v. groei gedurende enkele weken) op een dosis van de standaard of van de praeparaten welke we met de standaard willen vergelijken, voorgesteld kan worden door

$$\begin{aligned} \underline{y} &= \alpha + \beta \underline{\xi} + \underline{u}, \text{ voor de standaard,} \\ \underline{z} &= \gamma + \beta \underline{\xi} + \underline{v}, \text{ voor een praeparaat,} \end{aligned}$$

waarin $\underline{\xi}$ de logarithme van de dosis is ¹⁾, α , β en γ onbekende parameters zijn en $\underline{u}$ en $\underline{v}$ onafhankelijk verdeelde grootheden, afkomstig uit een normale verdeling met gemiddelde 0 en onbekende spreiding σ (deze is dus onafhankelijk van $\underline{\xi}$ en van het gebruikte praeparaat).

¹⁾ Door Irwin wordt deze aangegeven met X .

Hieruit volgt voor de mathematische verwachtingen van de reacties:

$$\eta \stackrel{\text{df}}{=} \xi \quad \underline{y} = \xi (\alpha + \beta \xi + \underline{u}) = \alpha + \beta \xi,$$

$$\xi \stackrel{\text{df}}{=} \xi \quad \underline{z} = \xi (\gamma + \beta \xi + \underline{v}) = \gamma + \beta \xi.$$

(Voor de verschillende praeparaten zijn γ en dus ook ξ verschillend.)

De onderstellingen die we hierbij dus maken zijn:

- (2.2.1) $\left\{ \begin{array}{l} \text{I: normaliteit der waarnemingen;} \\ \text{II: gelijkheid van de spreidingen der waarnemingen;} \\ \text{III: lineariteit der werkingslijnen;} \\ \text{IV: evenwijdigheid der werkingslijnen van stan-} \\ \text{daard en praeparaat bij een ijking.} \end{array} \right.$

De sterkteverhouding van een praeparaat tot de standaard wordt nu gegeven door de verhouding van een dosis van de standaard en een dosis van het praeparaat welke dezelfde verwachting van de reactie hebben, dus waarvoor geldt:

$$\xi \underline{y} = \xi \underline{z}.$$

Zijn ξ_1 en ξ_2 de logarithmen van zulke doses van de standaard resp. het praeparaat, waarvoor dus geldt:

$$\alpha + \beta \xi_1 = \gamma + \beta \xi_2,$$

dan vinden we voor de logarithme van de sterkteverhouding:

$$(2.2.2) \quad \mu \stackrel{\text{df}}{=} \xi_1 - \xi_2 = \frac{\gamma - \alpha}{\beta}.$$

Opmerking: a) In een grafiek van reactie tegen log-dosis, dus η tegen ξ is μ de horizontale afstand van de werkingslijnen van standaard en praeparaat.

In werkelijkheid zijn α , γ en β onbekende parameters, welke uit de waarnemingen geschat moeten worden. Deze schattingen zijn stochastische grootheden, waarvan de verdelingen afhangen van de verdeling van $\underline{u}$; in het bijzonder hangen de spreidingen samen met σ , welke eveneens uit de waarnemingen moet worden geschat.

Bij een ijking worden meerdere doses van de standaard en van de praeparaten gebruikt en voor elke dosis een groep proefdieren. De gebruikte notatie blijkt uit onderstaand schema I.

2) $\stackrel{\text{df}}{=}$ betekent: is per definitie gelijk aan.

Schema I.

Notatie.

	standaard	praeparaat
dosis	$\xi_1, \dots, \xi_h$	$\pi_1, \dots, \pi_k$
reacties der proefdieren	$y_{11} \dots y_{h1}$ $\vdots$ $y_{1m_1} \dots y_{hm_h}$	$z_{11} \dots z_{k1}$ $\vdots$ $z_{1p_1} \dots z_{kp_k}$
gemiddelde reactie	$\bar{y}_1 \dots \bar{y}_h$	$\bar{z}_1 \dots \bar{z}_k$
verwachting van de reactie	$\zeta_1 \dots \zeta_h$	$\xi_1 \dots \xi_k$
totaal gemiddelde reactie	$\bar{y}$	$\bar{z}$
gemiddelde dosis	$\bar{\xi}$	$\bar{\pi}$

Moet in het geval van meer dan één praeparaat een voor elk praeparaat apart op dezelfde wijze berekende grootheid over alle praeparaten gesommeerd worden, dan geven we dit aan door het symbool $\sum^{(p)}$. In sommige formules komen ook grootheden voor, die op slechts een praeparaat betrekking hebben en waarvoor dus geen $\sum^{(p)}$ -teken staat. Alle grootheden van die aard in één formule behoren steeds bij hetzelfdepraeparaat en de formule geldt dan voorieder der praeparaten afzonderlijk. Bij sommatie over in dit schema vertikaal onder elkaar geplaatste grootheden zullen we steeds de index i gebruiken, bij sommatie in horizontale richting de index j .

In het schema is:

$$(2.2.3) \quad \bar{y} = \frac{\sum_{j=1}^h m_j \bar{y}_j}{N_s} \quad \text{en} \quad \bar{\xi} = \frac{\sum_{j=1}^h m_j \xi_j}{N_s}$$

$$\bar{z} = \frac{\sum_{j=1}^k p_j \bar{z}_j}{N_p} \quad \text{en} \quad \bar{\pi} = \frac{\sum_{j=1}^k p_j \pi_j}{N_p}$$

waarin

$$N_s = \sum_{j=1}^h m_j \quad \text{en} \quad N_p = \sum_{j=1}^k p_j$$

Het aantal praeparaten noemen we n .

We zullen verder nog:

$$(2.2.4) \quad \check{\xi}_j \stackrel{\text{def}}{=} \xi_j - \bar{\xi} \quad \text{en} \quad \check{\pi}_j \stackrel{\text{def}}{=} \pi_j - \bar{\pi}$$

gebruiken.

De spreiding van een grootheid $\underline{a}$ geven we aan met $\sigma_{\underline{a}}$ of $\sigma\{\underline{a}\}$ en een schatting voor deze spreiding met $\underline{s}_{\underline{a}}$ of $s\{\underline{a}\}$.

Opmerkingen: b) Als de spreiding van $\underline{Y}_{11}$ gelijk is aan $\sigma_{\underline{u}_1}$ dan is de spreiding van $\underline{Y}_1$:

$$(2.2.5) \quad \sigma\{\underline{Y}_1\} = \frac{\sigma_{\underline{u}_1}}{\sqrt{m_1}}, \quad \text{enz.}$$

c) Alle waarnemingen, dus alle $\underline{y}$'s en $\underline{z}$'s uit het schema zijn onderling volledig onafhankelijk verdeeld.

2.3 Schattingen voor $\underline{\beta}, \underline{\alpha}, \underline{\gamma}$ en $\underline{\mu}$. ³⁾

Uit de waarnemingen zijn met de methode der kleinste kwadraten de volgende schattingen voor $\underline{\beta}, \underline{\alpha}$ en $\underline{\gamma}$ af te leiden:

$$(2.3.1) \quad \begin{cases} \underline{b} = \frac{\sum_{j=1}^k m_j \cdot \underline{\bar{y}}_j \cdot \underline{\bar{y}}_j + \sum_{j=1}^k p_j \cdot \underline{\bar{z}}_j \cdot \underline{\bar{z}}_j}{\sum_{j=1}^k m_j \cdot \underline{\bar{y}}_j^2 + \sum_{j=1}^k p_j \cdot \underline{\bar{z}}_j^2}, \\ \underline{a} = \underline{\bar{y}} - \underline{b} \underline{\bar{x}}, \\ \underline{c} = \underline{\bar{z}} - \underline{b} \underline{\bar{\pi}}. \end{cases}$$

In het volgende zullen we de noemer van $\underline{b}$ met $\underline{Q}$ aangeven, dus

$$(2.3.2) \quad \underline{Q} = \sum_{j=1}^k m_j \cdot \underline{\bar{y}}_j^2 + \sum_{j=1}^k p_j \cdot \underline{\bar{z}}_j^2.$$

Opmerking: d) Noemen we een schatting $\underline{t}$ van een parameter $\underline{\theta}$ zuiver ⁴⁾, wanneer $\underline{\xi} \underline{t} = \underline{\theta}$ is, dan zijn onder veel algemenere onderstellingen dan (2.2.1) de schattingen in (2.3.1) nog zuivere schattingen, dus $\underline{\xi} \underline{b} = \underline{\beta}$, $\underline{\xi} \underline{a} = \underline{\alpha}$ en $\underline{\xi} \underline{c} = \underline{\gamma}$. Noodzakelijke en voldoende onderstellingen zijn:

$$(2.3.3) \quad \begin{cases} \text{I : } \underline{\xi} \underline{u} = \underline{\xi} \underline{v} = 0; \\ \text{II: lineariteit der werkingslijnen;} \\ \text{III: evenwijdigheid der werkingslijnen.} \end{cases}$$

Met behulp van de schattingen (2.3.1) is het nu mogelijk een schatting $\underline{M}$ voor $\underline{\mu}$ te geven:

$$(2.3.4) \quad \underline{M} = \frac{\underline{c} - \underline{a}}{\underline{b}} = \frac{\underline{\bar{z}} - \underline{b} \underline{\bar{\pi}} - \underline{\bar{y}} + \underline{b} \underline{\bar{x}}}{\underline{b}} = \frac{\underline{\bar{z}}}{\underline{b}} - \underline{\bar{\pi}} + \frac{\underline{\bar{y}} - \underline{\bar{x}}}{\underline{b}}.$$

³⁾ Vgl. voor de gebruikte termen memorandum S 53 (M 21) aan het eind van dit rapport.

⁴⁾ Engels: unbiased.

Opmerkingen: e) In de appendix, § 8.1 bewijzen we nog het volgende:

Voegen we aan de onderstellingen (2.3.3) nog toe:

$$(2.3.5) \quad \text{IV: } \bar{Z}^2 = \sigma_{\bar{Z}}^2 \quad ; \quad \bar{Y}^2 = \sigma_{\bar{Y}}^2 ,$$

d.w.z. de spreidingen zijn onafhankelijk van de dosis, maar kunnen nog afhankelijk zijn van de praeparaten, terwijl geen normaliteit der verdelingen geëist wordt, dan zijn $\bar{Z}$ - $\bar{Y}$ en $\bar{b}$ ongecorreleerd.

f) Onder de onderstellingen (2.3.3)+(2.3.5) is $\bar{M}$ geen zuivere schatting van μ (appendix, § 8.2). Zonder het bewijs hiervan te geven merken we nog op dat de schatting $\bar{M}$ wel bruikbaar ⁵⁾ is (zie memorandum S 53 (M 21)).

2.4 Spreiding der schattingen.

Bij het afleiden van betrouwbaarheidsintervallen zullen we nog schattingen voor de spreidingen van $\bar{b}$, $\bar{a}$ en $\bar{c}$ nodig hebben. Voor de werkelijke spreidingskwadraten van deze grootheden volgt uit (2.3.1) en (2.3.2):

$$(2.4.1) \quad \left\{ \begin{aligned} \sigma_{\bar{Z}}^2 &= \frac{\sum_{j=1}^k m_j \bar{Z}_j^2}{N} + \frac{\sum_{j=1}^k m_j \bar{Z}_j^2}{N} \\ \sigma_{\bar{a}}^2 &= \sigma_{\bar{Z}}^2 + \bar{F}^2 \sigma_{\bar{c}}^2 = \frac{\sum_{j=1}^k m_j \bar{Z}_j^2}{N} + \bar{F}^2 \sigma_{\bar{c}}^2 \\ \sigma_{\bar{c}}^2 &= \sigma_{\bar{Z}}^2 + \bar{F}^2 \sigma_{\bar{a}}^2 = \frac{\sum_{j=1}^k m_j \bar{Z}_j^2}{N} + \bar{F}^2 \sigma_{\bar{a}}^2 \end{aligned} \right.$$

Schattingen van deze spreidingskwadraten kunnen we vinden door in (2.4.1) schattingen voor $\sigma_{\bar{Z}_j}^2$ en $\sigma_{\bar{Z}_j}^2$ te substitueren. Werken we onder de onderstellingen (2.3.3), dus zonder gelijkheid der spreidingen voor de verschillende doses te onderstellen, dan moet ieder van deze spreidingen uit de bijbehorende groep waarnemingen geschat worden:

$$(2.4.2) \quad \hat{\sigma}_{\bar{Z}_j}^2 = \frac{\sum_{i=1}^{m_j} (Z_{ij} - \bar{Z}_j)^2}{m_j - 1} , \text{ enz.}$$

we vinden dan als schatting voor de spreidingskwadraten:

⁵⁾ Engels : consistent.

$$S_{\underline{e}}^2 = \frac{\sum_j m_j \cdot \frac{1}{m_j} \cdot S_{\underline{e}j}^2 + \sum_j^{(P)} \sum_i^{(P)} P_j \cdot \frac{1}{P_j} \cdot S_{\underline{e}j}^2}{L_3^2} ;$$

$$S_{\underline{a}}^2 = \frac{\sum_j m_j \cdot S_{\underline{e}j}^2}{N_S^2} + \frac{1}{L_3^2} S_{\underline{e}}^2 ;$$

$$S_{\underline{c}}^2 = \frac{\sum_j P_j \cdot S_{\underline{e}j}^2}{N_P^2} + \frac{1}{L_3^2} S_{\underline{e}}^2 .$$

Onder de onderstellingen (2.2.1) kunnen we voor een schatting uit het gehele waarnemingsmateriaal geven:

$$(2.4.3) \quad S_{\underline{u}}^2 = \frac{\sum_{d=1}^h \sum_{i=1}^{m_j} (y_{di} - \bar{y}_j)^2 + \sum_{d=1}^{(P)} \sum_{i=1}^{P_j} (z_{di} - \bar{z}_j)^2}{N_S - h + \sum^{(P)} (N_P - k)} =$$

$$= \frac{\sum_j (m_j - 1) S_{\underline{e}j}^2 + \sum_j^{(P)} \sum_i^{(P)} (P_j - 1) S_{\underline{e}j}^2}{N_S - h + \sum^{(P)} (N_P - k)} ;$$

waaruit volgt :

$$(2.4.4) \quad \left\{ \begin{array}{l} S_{\underline{e}}^2 = \frac{S_{\underline{u}}^2}{L_3^2} ; \\ S_{\underline{a}}^2 = \frac{S_{\underline{u}}^2}{N_S^2} + \frac{\frac{1}{L_3^2} S_{\underline{u}}^2}{L_3^2} ; \\ S_{\underline{c}}^2 = \frac{S_{\underline{u}}^2}{N_P^2} + \frac{\frac{1}{L_3^2} S_{\underline{u}}^2}{L_3^2} . \end{array} \right.$$

Opmerkingen: g) Zijn de aantallen waarnemingen per groep gelijk, dan komt (2.4.3) neer op het gemiddelde der schattingen (2.4.2) per groep.

$$(2.4.5) \quad h) \quad \frac{S_{\underline{u}}^2}{\sigma_{\underline{u}}^2} \times f \text{ met } f = N_S - h + \sum^{(P)} (N_P - k)$$

heeft een χ^2 -verdeling met f vrijheids graden.

2.5 Exakte betrouwbaarheidsintervallen.

Door FIELLER is in [4] en uitvoeriger in [2] aangegeven hoe exakte betrouwbaarheidsintervallen met onbetrouwbaarheid ε voor μ (zie (2.2.2)) berekend kunnen worden (zie ook IRWIN [5] en EMMENS [6] hoofdstuk 11.5). Bij deze berekening moeten we uitgaan van de strenge onderstellingen (2.2.1) (zie appendix, § 8.3). Deze betrouwbaarheidsgrenzen voor μ , behorende bij een waarnemingsreeks, die de waarden c, a, b enz. heeft gegeven, zijn de wortelsvoor μ van de vergelijking (zie appendix, § 8.4):

$$(8.4.3) \quad \mu^2 [b^2 - t_{\varepsilon}^2 S_{\underline{e}}^2] - 2\mu [b(c-a) - t_{\varepsilon}^2 (\bar{x} - \bar{\pi}) S_{\underline{e}}^2] +$$

$$+ (c-a)^2 - t_{\varepsilon}^2 [S_{\underline{a}}^2 (\frac{1}{N_S} + \frac{1}{N_P}) + (\bar{x} - \bar{\pi})^2 S_{\underline{e}}^2] = 0$$

waarin t_ε die waarde uit de Student verdeling met f vrijheidsgraden (vgl. (2.4.5)) is waarvoor geldt:

$$P[|t| \geq t_\varepsilon] = \varepsilon,$$

(ε is de onbetrouwbaarheid). Deze waarde kan bij gegeven ε in een tabel der Student-verdeling worden opgezocht.

Door uit vergelijking (8.4.3) μ op te lossen en hierin

$$(2.5.1) \quad C = \frac{b^2}{b^2 - t_\varepsilon^2 s_b^2}$$

te substitueren, vinden we voor de twee grenzen voor (zie appendix § 8.4):

$$(2.5.2) \quad (\bar{z} - \bar{y}) \pm \frac{C(\bar{z} - \bar{y})}{b} \pm t_\varepsilon \sqrt{\frac{s_b^2}{b^2} \left(\frac{1}{N_s} + \frac{1}{N_p} \right) + \frac{C(\bar{z} - \bar{y})^2 s_b^2}{b^4}}$$

Wanneer we hierin als benadering $C = 1$ invoeren en voor t_ε de bij de betrouwbaarheid ε behorende waarde t_ε^* uit de normale verdeling substitueren, d.w.z. waarvoor geldt:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t_\varepsilon^*} e^{-\frac{1}{2}u^2} du = \frac{1}{2} \varepsilon,$$

dan vinden we door Irwin gebruikte (oorspronkelijk op anderewijze afgeleide) benaderde grenzen:

$$(2.5.3) \quad (\bar{z} - \bar{y}) \pm \frac{\bar{z} - \bar{y}}{b} \pm t_\varepsilon^* \sqrt{\frac{s_b^2}{b^2} \left(\frac{1}{N_s} + \frac{1}{N_p} \right) + \frac{(\bar{z} - \bar{y})^2 s_b^2}{b^4}}.$$

De waarde t_ε^* kan men voor gegeven ε in een tabel der normale verdeling opzoeken.

In § 5 zal de fout die door deze benadering ontstaat besproken worden.

Daar steeds $C > 1$ is, zullen de grenzen van Irwin systematisch te nauw zijn en bovendien afhankelijk van het teken van $(\bar{z} - \bar{y})$ naar de lage waarden (bij (+)-teken) of naar de hoge waarden (bij (-)-teken) verschoven zijn t.o.v. de exacte betrouwbaarheidsgrenzen.

Bij bovenstaande afleiding is nog een zeer belangrijke opmerking te maken. Het betrouwbaarheidsinterval bestaat uit die waarden van μ waarvoor het linkerlid van (8.4.3) kleiner is dan 0 (zie appendix § 8.4), wanneer $b^2 - t_\varepsilon^2 s_b^2 > 0$ is, zijn dit de waarden tussen de beide gevonden grenzen (2.5.2), is echter $b^2 - t_\varepsilon^2 s_b^2 < 0$ dan zijn dit echter de waarden van μ buiten de gevonden grenzen. Dit laatste is natuurlijk bij ijkingen onbruikbaar en we moeten besluiten dat

voor de toepasbaarheid van deze intervalberekening van Fieller, voldaan moet zijn aan:

$$\left| \frac{b}{s_b} \right| > t_{\varepsilon}.$$

Dit betekent dat b significant van 0 moet verschillen volgens de toets van Student (met f vrijheidsgraden) met een betrouwbaarheidsdrempel ε . Dit geldt voor de methode van Irwin, zoals blijkt in § 8.5 bij de afleiding van vergelijking (8.5.1).

Was g de basis van de logaritmen die we gebruikten (voor log dosis, enz.), dan vinden we voor de sterkte-verhouding als betrouwbaarheidsinterval met onbetrouwbaarheid ε , het interval:

$$(2.4.5) \quad (g^{\mu_1}; g^{\mu_2}),$$

waarin μ_1 en μ_2 de beide grenzen van (2.5.2) voorstellen.

2.6 Numerieke voorbeelden.

Als eerste voorbeeld zullen we de berekeningen uitvoeren voor ijking 4926, 5 weken. Als basis voor de logaritmen van de dosis gebruiken we 2. De gegevens vinden we in tabel I.

Tabel I.

Gegevens van ijking 4926, 5 weken.

standaard					
dosis	m_j	$\bar{y}_j$	$\bar{s}_j$	$\bar{y}_j$	$\bar{s}_{\bar{y}_j}^2$
12,5	12	3,644	-2	64,9	130,08
25	12	4,644	-1	72,3	96,97
50	11	5,644	0	94,2	96,36
100	12	6,644	1	110,8	121,06
200	12	7,644	2	122,2	196,88
praeparaat 1.					
	p_j	π_j	$\bar{\pi}_j$	$\bar{z}_j$	$\bar{s}_{\bar{z}_j}^2$
30	12	4,907	$-\frac{1}{2}$	80,8	103,11
60	12	5,907	$+\frac{1}{2}$	101,8	216,20
praeparaat 2.					
7,9	12	2,982	0	74,7	63,33
praeparaat 3.					
31,9	12	4,995	0	62,0	135,64

In deze tabel is dus $\bar{y} = {}^2\log 12,5 = 3,644$, enz., en
 $S_{u_1}^2 = \frac{1}{m_1 - 1} \sum_i (y_{1j} - \bar{y}_j)^2 = 130,08$ enz. Verder zijn voor
 deze tabel nog berekend de grootheden:

$$\bar{y} = \frac{\sum_j m_j \bar{y}_j}{N_S} = 5,644 \quad \text{en} \quad N_S = \sum m_j = 59,$$

$$\bar{\pi}_1 = \frac{\sum_j p_j \pi_j}{N_P} = 5,407 \quad \text{en} \quad N_{P_1} = 24,$$

$$\bar{\pi}_2 = {}^2\log 7,9 = 2,982 \quad \text{en} \quad N_{P_2} = 12,$$

$$\bar{\pi}_3 = {}^2\log 31,9 = 4,995 \quad \text{en} \quad N_{P_3} = 12.$$

In deze paragraaf zullen we de grootheden behorende
 bij de verschillende praeparaten onderscheiden met de
 indices 1, 2 en 3 zoals hierboven al voor $\bar{\pi}$ is gedaan.
 Uit deze tabel berekenen we nu de grootheden:

$$\bar{y} = \frac{\sum m_j \bar{y}_j}{N_S} = 92,85,$$

$$\bar{z}_1 = \frac{\sum p_j \bar{z}_j}{N_{P_1}} = 91,3,$$

$$\bar{z}_2 = 74,7 \quad \text{en} \quad \bar{z}_3 = 62,0.$$

Volgens (2.3.1) is nu

$$b = \frac{\sum m_j \bar{y}_j^2 + \sum_{(P)} \sum_j p_j \bar{\pi}_j^2}{\sum m_j \bar{y}_j^2 + \sum_{(P)} \sum_j p_j \bar{\pi}_j^2} = 15,581,$$

$$Q = \sum m_j \bar{y}_j^2 + \sum_{(P)} \sum_j p_j \bar{\pi}_j^2 = 126,$$

$$a = \bar{y} - b \bar{y} = 4,91,$$

$$c_1 = \bar{z}_1 - b \bar{\pi}_1 = 7,1,$$

$$c_2 = \bar{z}_2 - b \bar{\pi}_2 = 28,3,$$

$$c_3 = \bar{z}_3 - b \bar{\pi}_3 = -15,8.$$

Daar van de praeparaten 2 en 3 slechts 1 dosis gebruikt
 is, dus $\bar{\pi} = 0$ is, is de helling b alleen afhankelijk
 van de waarnemingen voor standaard en praeparaat 1.

Voor de logaritmen van de sterkteverhoudingen vinden we volgens (2.3.4):

$$\begin{aligned} M_1 &= \frac{c_1 - a}{b} = 0,141, \\ (2.6.1) \quad M_2 &= \frac{c_2 - a}{b} = 1,51, \\ M_3 &= \frac{c_3 - a}{b} = -1,32. \end{aligned}$$

Volgens (2.4.3) vinden we:

$$S_u^2 = \frac{\sum (m_i - 1) s_{u_i}^2 + \sum_{j=1}^p \sum_{i=1}^2 (p_j - 1) s_{x_j}^2}{N_S - k + \sum_{j=1}^p (N_p - k)} = 129,179$$

en volgens (2.4.5):

$$f = N_S - h + \sum_{j=1}^p (N - k) = 98.$$

Uit de Student verdeling met 98 vrijheidsgraden en $\varepsilon = 0,05$ volgt: $t_{\varepsilon} = 1,98$. Daar (2.4.4):

$$S_b^2 = \frac{S_u^2}{Q} = \frac{129,179}{126} = 1,025,$$

is (2.5.1):

$$C = \frac{b^2}{b^2 - t_{\varepsilon}^2 S_b^2} = 1,0168.$$

Met behulp hiervan kunnen we de betrouwbaarheidsgrenzen volgens (2.5.2) (Fieller) berekenen:

$$(2.6.2) \quad \left\{ \begin{array}{l} \text{praeparaat 1: } (\mu_1; \mu_2) = (0,489; -0,217) \\ \text{praeparaat 2: } (\mu_1; \mu_2) = (1,965; 0,995) \\ \text{praeparaat 3: } (\mu_1; \mu_2) = (-0,832; -1,890). \end{array} \right.$$

Ditzelfde uitgerekend volgens de berekening van Irwin (2.5.3) geeft hier praktisch dezelfde uitkomsten:

$$\begin{aligned} \text{praeparaat 1: } (\mu_1'; \mu_2') &= (0,488; -0,212) \\ \text{praeparaat 2: } (\mu_1'; \mu_2') &= (1,981; -1,019) \\ \text{praeparaat 3: } (\mu_1'; \mu_2') &= (-0,806; -1,850). \end{aligned}$$

Betrouwbaarheidsgrenzen ($g^{\mu_1}; g^{\mu_2}$) voor de sterkteverhouding volgen nu uit (2.6.2):

$$\begin{aligned} \text{praeparaat 1: } (2^{\mu_1}; 2^{\mu_2}) &= (1,40; 0,860) \\ \text{praeparaat 2: } (2^{\mu_1}; 2^{\mu_2}) &= (3,90; 1,99) \\ \text{praeparaat 3: } (2^{\mu_1}; 2^{\mu_2}) &= (0,561; 0,270). \end{aligned}$$

Volgens de benadering van Irwin vinden we:

$$\text{praeparaat 1: } (2^{\mu_1'}; 2^{\mu_2'}) = (1,40 ; 0,863)$$

$$\text{praeparaat 2: } (2^{\mu_1'}; 2^{\mu_2'}) = (3,95 ; 2,03)$$

$$\text{praeparaat 3: } (2^{\mu_1'}; 2^{\mu_2'}) = (0,572 ; 0,277)$$

De lengtes van de intervallen zijn dus in dit geval voor beide methoden ongeveer gelijk. Voor de sterkteverhoudingen zelf vinden we uit (2.6.1) de volgende schattingen:

$$\text{praeparaat 1 : } 2^{\frac{M_1}{2}} = 2^{0,141} = 1,10 ,$$

$$\text{praeparaat 2 : } 2^{\frac{M_2}{2}} = 2^{1,51} = 2,85 ,$$

$$\text{praeparaat 3 : } 2^{\frac{M_3}{2}} = 2^{-1,32} = 0,400.$$

Om nog een voorbeeld te geven waarbij de overeenstemming tussen het betrouwbaarheidsinterval van Fieller en de benadering volgens Irwin minder goed is, zullen we dezelfde berekeningen uitvoeren voor ijking 4926, 2 weken.

Tabel II
Gegevens van ijking 4926, 2 weken.

standaard					
dosis	m_j	ξ_j	$\check{\xi}_j$	y_j	$s_{u_j}^2$
12,5	12	3,644	-2	36,7500	60,20
25	12	4,644	-1	35,7500	50,02
50	11	5,644	0	43,0000	53,20
100	12	6,644	1	46,5833	57,90
200	12	7,644	2	49,7500	35,84
praeparaat					
dosis	p_j	π_j	$\check{\pi}_j$	z_j	$s_{v_j}^2$
30	12	4,907	$-\frac{1}{2}$	40,5833	39,58
60	12	5,907	$+\frac{1}{2}$	44,4167	54,08

Uit tabel II berekenen we:

$$\bar{\xi} = 5,644 \quad ; \quad N_S = 59 ,$$

$$\bar{\pi} = 5,407 \quad ; \quad N_P = 24 ,$$

$$\bar{y} = \frac{\sum m_j y_j}{N_S} = 42,3559 \quad ; \quad \bar{z} = \frac{\sum p_j z_j}{N_P} = 42,5000 ,$$

$$Q = \sum m_j \check{\xi}_j^2 + \sum p_j \check{\pi}_j^2 = 126 ,$$

$$b = \frac{\sum m_j \bar{y}_j + \sum p_j \bar{z}_j}{Q} = 3,690$$

$$a = \bar{Y} - b \bar{x} = 21,529 ; c = \bar{Z} - b \bar{x} = 22,548,$$

$$M = \frac{c-a}{b} = 0,2761 ,$$

$$f = N_S + N_P - h - k = 76 ,$$

$$s_u^2 = \frac{\sum (m_j - 1) s_{y_j}^2 + \sum (p_j - 1) s_{z_j}^2}{f} = 50,085 ,$$

$$s_b^2 = \frac{s_u^2}{Q} = 0,3975 ; c = \frac{b^2}{b^2 - t_{\epsilon}^2 s_b^2} = 1,1322$$

$$(\mu_1 ; \mu_2) = (1,269 ; -0,707)$$

$$(2^{\mu_1} ; 2^{\mu_2}) = (2,41 ; 0,613) \quad (\text{Fieller})$$

$$(\mu_1' ; \mu_2') = (1,186 ; -0,634)$$

$$(2^{\mu_1'} ; 2^{\mu_2'}) = (2,27 ; 0,644) \quad (\text{Irwin})$$

Op de grootte van dit verschil komen wij in § 5.1 nog terug.

3. Toetsingsmethoden voor de gemaakte onderstellingen.

In de vorige paragraaf zagen we dat om de methode van Irwin en de berekening van betrouwbaarheidsintervallen te kunnen toepassen de onderstellingen (2.2.1) vervuld moeten zijn. In het volgende zullen we voor elk van deze onderstellingen toetsingsmethoden bespreken.

3.1 Normaliteit.

Normaliteit van de waarnemingen is, behalve voor het bepalen van betrouwbaarheidsintervallen, bovendien noodzakelijk bij toetsing van homoscedasticiteit, lineariteit en evenwijdigheid. Bij het onderhavige onderzoek werd er vanaf gezien de normaliteit der gegevens te toetsen, daar wij slechts over groepjes van 12 waarnemingen beschikten. Het toetsen van normaliteit is dan nauwelijks op exacte wijze mogelijk, terwijl bovendien slechts zeer grove afwijkingen zullen kunnen worden ontdekt. Bij alle in het volgende besproken methoden, waarbij normaliteit ondersteld moet worden, dient men dan ook bij de interpretatie der resultaten rekening te houden met de mogelijkheid van afwijkingen. Het zou daarom nuttig zijn een apart experimenteel onderzoek met een grotere groep van dieren uit te voeren, om een beter inzicht in de aard van de waarschijnlijkheidsverdeling van de gewichtstoename te verkrijgen.

Wellicht kan hier gebruik gemaakt worden van de resultaten van vroegere proeven met dezelfde soort proefdieren.

3.2 Homoscedasticiteit.

De gelijkheid der spreidingen werd onderzocht met behulp van de toets van Bartlett (Vgl. de memoranda S 53 (M 25) en S 47 (M 6)). Dit gaf de in tabel III vermelde resultaten. In deze en volgende tabellen wordt steeds achter het nummer van de ijking (1^o kolom, tabel III) tussen haakjes het aantal doses van de standaard gegeven, terwijl achter de aanduiding van een praeparaat tussen haakjes het aantal doses van dit praeparaat volgt. In tabel III zijn de overschrijdingskansen volgens de toets van Bartlett gegeven, de onderstreepte kansen wijzen dus op significante afwijkingen van homoscedasticiteit. Deze afwijkingen bemoeilijken niet alleen de berekening van betrouwbaarheidsintervallen, maar ook het onderzoek naar de lineariteit en evenwijdigheid daar de in de literatuur beschreven toetsen, b.v. met behulp van variantieanalyse, steeds gelijke spreidingen in de groepen onderstellen.

Tabel III

Toets van Bartlett voor gelijkheid der spreidingen.

ijking nr.	combinatie van de groepen van	aantal weken				
		1	2	3	4	5
4926 (5)	standaard + praeparaat Un(2)	>0.05	>0.05	>0.05	>0.05	>0.05
	standaard	>0.05	>0.05	>0.05	>0.05	>0.05
503 (6)	standaard + praeparaten (1)	< <u>0.01</u>	< <u>0.01</u>	< <u>0.01</u>		
	standaard	>0.05	< <u>0.01</u>	< <u>0.01</u>		
5011 (4)	standaard + praeparaten (2)	>0.05	>0.05	>0.05	>0.05	< <u>0.05</u>
	standaard	>0.05	>0.05	>0.05	>0.05	>0.05
5017 (6)	standaard + praeparaten(2)	>0.05	< <u>0.05</u>	>0.05	< <u>0.05</u>	< <u>0.01</u>
	standaard	>0.05	< <u>0.05</u>	>0.05	>0.05	>0.05
5023 (4)	standaard + praeparaten (1)	< <u>0.05</u>	>0.05	>0.05	>0.05	>0.05
	standaard	>0.05	>0.05	>0.05	>0.05	>0.05

Conclusie: Ongelijkheid der spreidingen binnen de groepen zowel bij de standaard apart als bij standaard en praeparaten tezamen treedt veelvuldig op.

In verband met de moeilijkheden die deze afwijkingen van homoscedasticiteit bij het berekenen van betrouwbaarheidsintervallen veroorzaken, is in de volgende paragraaf een benaderde schatting van de spreiding van de logaritmische sterkteverhouding gegeven, waarbij de ongelijkheid der spreidingen in aanmerking genomen wordt. Hierbij zal blijken, dat bij een behoorlijke afwijking van homoscedasticiteit ook een vrij aanzienlijke verandering van de spreiding van de log sterkteverhouding gevonden wordt.

3.3 Lineariteit.

De lineariteitstoets, zoals deze beschreven wordt door EMMENS [6] p.40, is een toepassing van de variantie-analyse. Hierbij wordt dus normaliteit en gelijkheid der spreidingen geëist. Voor die ijkingen waarbij dit niet vervuld is, verliezen de uitkomsten van deze toets veel van hun betekenis.

Bij deze toets gebruiken we alleen gegevens van de standaard voor de schattingen van de helling $\underline{b}$ en de spreiding S_u binnen de groepen, daar, bij combinatie van de waarnemingen van standaard en praeparaten, door mogelijke afwijkingen van evenwijdigheid van de werkingslijnen of verschillen tussen de groepsspreidingen σ_u en σ_v van standaard resp. praeparaten de uitkomsten van de toets vertroebeld kunnen worden. Toetsing van lineariteit van de werkinglijn van een praeparaat was niet mogelijk daar nergens van een praeparaat meer dan twee doses gebruikt zijn.

Als toetsingsgrootte gebruiken we de $\underline{F}$ van Snedecor (zie memorandum S 53 (M 24)) om de som van de kwadraten van de residuen van de groepsgemiddelden der waarnemingen ten opzichte van de lineaire regressielijn (gedeeld door het aantal groepen verminderd met 2 dus $h-2$) te vergelijken met $\underline{S_u}^2$. In formule:

$$\underline{F} = \frac{\sum_{j=1}^h m_j (\bar{y}_j - \bar{a} - \bar{b}x_j)^2 / (h-2)}{\underline{S_u}^2}$$

Hierin hebben $\underline{b}_s$ en $\underline{S_u}^2$ analoge gedaanten als in (2.3.1) en (2.4.3), maar nu alleen toegepast voor de waarnemingen van de standaard.

Indien de hypothese van lineaire regressie juist is, zullen teller en noemer onafhankelijke schattingen zijn van het kwadraat van de spreiding der waarnemingen voor hun verwachting, hetgeen we hier niet bewijzen; zie b.v. A.M. MOOD [7] pg. 321 en 322), zodat hun quotient verdeeld zal zijn

volgens de verdeling van Snedecor met $h-2$ en N_S-h vrijheidsgraden. Bij afwijkingen van lineariteit zal de teller te groot worden, zodat bij de toets een rechter-éénzijdige critieke zone wordt gebruikt. Voor sommige van de beschikbare ijkingen is de waarde van F met de aantallen vrijheidsgraden en de overschrijdingskans opgenomen in tabel IV.

Tabel IV
Afwijkingen van lineariteit .

ijking no.	afwijkingen van lineariteit.		
	F	vrh gr.	ov.k.
4926 (5)			
1 week	0.86	3 en 55	>0.20
2 weken	1.02	3 en 55	>0.20
5 weken	1.40	3 en 55	>0.20
503 (5)			
1 week	1.15	3 en 55	>0.20
3 weken	3.2	3 en 55	<0.05
5011 (3)			
1 week	3.4	1 en 33	>0.05
3 weken	1.8	1 en 33	>0.20
5 weken	2.56	1 en 33	>0.05
5017 (4)			
1 week	0.10	2 en 44	>0.20
3 weken	0.27	2 en 44	>0.20
5 weken	0.9	2 en 44	>0.20
5023 (3)			
1 week	0.00	1 en 33	>0.20
3 weken	0.50	1 en 33	>0.20
5 weken	1.48	1 en 33	>0.20

In deze tabel zien we dat slechts bij ijking 503, 3 weken de lineariteit niet goed vervuld is (overschrijdingskans <0.05), zodat we tot de conclusie komen dat de voorwaarde van lineariteit van de standaardlijn bij deze proeven in behoorlijke mate vervuld is.

3.4 Evenwijdigheid.

Toetsing is uiteraard slechts mogelijk, indien van standaard en praeparaat elk 2 of meer doses gebruikt zijn, ijkningen 503 en 5023 blijven daarom buiten beschouwing. De hellingen van de werkingslijnen van standaard en praeparaat kunnen door middel van de toets van Student vergeleken worden (zie memorandum S 47 (M 9)). Hierbij wordt dus weer gebruik gemaakt van de normaliteit en homoscedastieiteit en voor de uitkomsten der toets geldt dezelfde opmerking als in § 3.3.

Bij deze toetsing hebben we onafhankelijke schattingen nodig voor de hellingen van de standaardwerkingslijn en van de praeparaatwerkingslijn en van hun spreidingskwadraten. Analooq aan (2.3.1) en (2.4.4) vinden we met de methode der kleinste kwadraten voor de standaard:

$$\underline{b}_s = \frac{\sum m_j \check{s}_j \cdot \check{y}_j}{\sum m_j \check{s}_j^2},$$

$$\underline{s}_{\underline{b}_s}^2 = \frac{\underline{s}_u^2}{\sum m_j \check{s}_j^2}$$

en voor het praeparaat:

$$\underline{b}_p = \frac{\sum p_j \check{\pi}_j \cdot \check{z}_j}{\sum p_j \check{\pi}_j^2},$$

$$\underline{s}_{\underline{b}_p}^2 = \frac{\underline{s}_u^2}{\sum p_j \check{\pi}_j^2}.$$

Voor de schatting van het spreidingskwadraat van $\underline{b}_s - \underline{b}_p$ vinden we:

$$\underline{s}_{\underline{b}_s - \underline{b}_p}^2 = \underline{s}_u^2 \left\{ \frac{1}{\sum m_j \check{s}_j^2} + \frac{1}{\sum p_j \check{\pi}_j^2} \right\}.$$

Onder de hypothese $\beta_s = \beta_p$ bezit de grootheid

$$\underline{t} = \frac{\underline{b}_s - \underline{b}_p}{\underline{s}_{\underline{b}_s - \underline{b}_p}}.$$

een Student verdeling met $f = N_s - h + \sum^{(p)} (N - k)$ vrijheidsgraden. Door de gevonden waarden van $\underline{t}$ te vergelijken met een tabel van de Student-verdeling vinden we de (tweezijdige) overschrijdingskansen welke in tabel V verzameld zijn.

Tabel V.
Evenwijdigheid der regressielijnen.

ijking no.	vergelijking tussen standaard en	t	vrijheids- graden	overschrij- dingskans
4926 (5) 5 weken	praep. Un (2)	1,201	99	>0,20
5011 (3) 5 weken	praep.SF (2)	1,28	77	0.20
	praep.HJ (2)	2,97	77	< <u>0,01</u>
4 weken	praep. HJ (2)	2,49	77	<u>0.02</u>
3 weken	praep.HJ. (2)	2,136	77	<u>0.05</u>
5017 (4) 5 weken	praep. (2)	0,75	66	≈0,5

De aantallen dieren per groep zijn voor de eenvoud der berekening weer steeds gelijk aan 12 ondersteld. De overige weken van deze ijkingsen vertonen geen significante afwijking meer van de evenwijdigheid.

Zijn de werkingslijnen van standaard en praeparaat niet evenwijdig, zoals we uit het bovenstaande voor het praeparaat HJ moeten besluiten, dan kan niet meer gesproken worden van "de" sterkteverhouding tussen praeparaat en standaard, daar de uitkomst nu afhankelijk geworden is van de gemiddelde toegediende doses van standaard en praeparaat. Het is dan nog mogelijk voor elke dosis van het praeparaat apart de sterkteverhouding ten opzichte van de standaard op te geven, maar deze sterkteverhouding geldt dan ook alleen voor de gebruikte praeparaat-dosis. Een volledig overzicht van de werking kan in een dergelijk geval alleen door de gehele werkingslijn van het praeparaat gegeven worden.

Het is mogelijk, dat in dit geval de evenwijdigheid zó duidelijk verstoord is, omdat de werkzame dosis van standaard en praeparaat sterk van elkaar verschillen. De reacties op de twee doses van het praeparaat zijn beide kleiner dan die op de standaard. De helling van de praeparaatwerkingslijn is groter dan die van de standaard. Indien de standaard-doses gedeeltelijk of geheel rechts van het lineariteitsgebied vallen, zou men dit effect juist kunnen verwachten. De doses van het praeparaat SF liggen

hoger dan van HJ. De werkingslijn van dit praeparat wijkt niet significant van die van de standaard af. Deze "verklaring" van het verschijnsel is echter geenszins de enig mogelijke:

Conclusie: De evenwijdigheid van standaard- en praeparatwerkingslijnen is bij 3 van de 4 gevallen bevredigend. In één geval is dit niet zo. Wellicht is dit een gevolg van een te groot verschil in werkzame doses van standaard en praeparat.

3.5 Toetsing van $\beta = 0$.

Toetsing van de hypothese $\beta = 0$ (d.w.z. onafhankelijkheid tussen reactie en dosis) is om de volgende redenen van belang. In de eerste plaats heeft bij onafhankelijkheid tussen reactie en dosis de ijking geen betekenis meer. Verder is in § 2.5 besproken, dat voor de toepasbaarheid van de methode van Fieller (of Irwin) ter bepaling van betrouwheidsintervallen voldaan moet zijn aan:

$$\left| \frac{b}{s_b} \right| > t_{\varepsilon} ,$$

waarin t_{ε} de waarde uit de Studentverdeling met $f = N_D - h + \sum (N_P - k)$ vrijheidsgraden is behorende bij de onbetrouwbaarheid ε . Tenslotte zal ook bij de afleiding van een benadering voor de spreiding van de logaritme van de sterkteverhouding als de homoscedasticiteit niet vervuld is gebruik gemaakt worden van de onderstelling $\sigma_b < \sigma_{\beta}$ dus $\beta \neq 0$ (§ 4 en § 8.5).

Toetsing van de hypothese $\beta = 0$ kan gebeuren door de waarde van

$$\underline{t} = \frac{\underline{b}}{\underline{s}_b} = \frac{\underline{b}\sqrt{Q}}{\underline{s}_u} ,$$

welke bij een ijking gevonden wordt te vergelijken met een tabel voor de Studentverdeling met $f = N_D - h + \sum (N_P - k)$ vrijheidsgraden. De voor de verschillende ijkingen op deze wijze gevonden overschrijdingskansen zijn verzameld in tabel VI.

EMMENS [6] gebruikt als toetsingsgrootheid

$$\underline{F} = \frac{\underline{b}^2 Q}{\underline{s}_u} = \underline{t}^2 ,$$

welke onder de hypothese $\beta = 0$ verdeeld is als de $\underline{F}$ van Snedecor (zie memorandum S 53 (M 24)) met 1 en

$N_S - h + \sum^{(P)} (N_P - k)$ vrijheidsgraden. In het algemeen geldt dat, wanneer een grootte verdeeld is volgens een Studentverdeling met ν vrijheidsgraden, het kwadraat van deze grootte en F verdeling met 1 en ν vrijheidsgraden bezit. Tweezijdige toetsing van t komt overeen met éénzijdige toetsing van $F = t^2$.

Tabel VI
Toetsing van de hypothese $\beta = 0$.

ijking no.	aantal weken	$t = \frac{b}{s_b}$	f	overschr. kans
4926 (5)	1	0,73	99	$> 0,40$
praep. 1 (2)	2	5,66	99	< 0.001
praep. 2 (1)	3	11,09	99	< 0.001
praep. 3 (1)	4	13,97	99	< 0.001
	5	15,40	99	< 0.001
503 (5)	1	9,26	99	< 0.001
praeparaten	2	14,89	99	< 0.001
1 t/m 4 (1)	3	16,37	99	< 0.001
5011 (3)	1	6,06	77	< 0.001
praep.SP (2)	2	10,90	77	< 0.001
praep.HJ (2)	3	10,06	77	< 0.001
	4	9,97	77	< 0.001
	5	10,12	77	< 0.001
5017 (4)	1	1,91	66	> 0.05
praeparaat (2)	2	5,17	66	< 0.001
	3	6,27	66	< 0.001
	4	7,45	66	< 0.001
	5	8,69	66	< 0.001
5023 (3)	1	7,68	77	< 0.001
praeparaten	2	12,91	77	< 0.001
1 t/m 4 (1)	3	14,44	77	< 0.001
	4	13,67	77	< 0.001
	5	14,07	77	< 0.001

In tabel VI zien wedat slechts bij de ijkingen 4926, 1 week en 5017, 1 week de hypothese $\beta = 0$ niet verworpen kan worden (overschrijdingskansen resp. > 0.40 en > 0.05). We komen dus tot de conclusie dat bij deze ijkingen met een duur van 2 of meer weken, de hypothese van onafhankelijkheid tussen reactie en dosis verworpen kan worden en $\frac{b}{s_b}$ voldoende groot is om betrouwbaar-

heidsintervallen volgens § 2.5 of § 4 te kunnen bepalen.

4. Schatting van de spreiding van de logaritmische van de sterkteverhouding wanneer de homoscedasticiteit niet vervuld is.

In de appendix, § 8.5 is een benadering afgeleid voor het spreidingskwadraat van het quotient van twee normaal verdeelde grootheden $\underline{r}$ en $\underline{t}$ onder de voorwaarde, dat de mathematische verwachting τ van $\underline{t}$ groot is ten opzichte van de spreiding van $\underline{t}$. Deze benadering luidt:

$$(8.5.1) \quad \sigma^2 \left\{ \frac{\underline{r}}{\underline{t}} \right\} \approx \frac{\sigma_{\underline{r}}^2}{\tau^2} + \frac{\rho^2 \sigma_{\underline{t}}^2}{\tau^4} - 2 \frac{\rho \tau (\tilde{\underline{r}} \cdot \tilde{\underline{t}})}{\tau^3},$$

waarin ρ de verwachting van $\underline{r}$ voorstelt.

Dit passen we toe op de logaritmische van de sterkteverhouding. Uit (2.3.4) volgt:

$$\sigma_{\underline{r}}^2 = \sigma^2 \left\{ \frac{\tilde{\underline{z}} - \tilde{\underline{y}}}{\underline{t}} \right\}.$$

In (8.5.1) substitueren we de volgende schattingen:

$$\text{voor } \rho : \underline{\tilde{\underline{z}}} - \underline{\tilde{\underline{y}}},$$

$$\text{voor } \tau : \underline{b},$$

$$\text{voor } \sigma_{\underline{z}}^2 : \underline{S}_{\underline{z}}^2 = \underline{S}^2 \left\{ \underline{\tilde{\underline{z}}} - \underline{\tilde{\underline{y}}} \right\} = \underline{S}_{\underline{\tilde{\underline{z}}}}^2 + \underline{S}_{\underline{\tilde{\underline{y}}}}^2 = \frac{\sum p_j \cdot \tilde{z}_j^2}{N_p} + \frac{\sum m_j \cdot \tilde{z}_j^2}{N_s}$$

$$\text{voor } \sigma_{\underline{t}}^2 : \underline{S}_{\underline{t}}^2 = \frac{\sum m_j \cdot \tilde{t}_j^2 \cdot \underline{S}_{\underline{z}_j}^2 + \sum p_j \cdot \tilde{t}_j^2 \cdot \underline{S}_{\underline{y}_j}^2}{Q^2}.$$

We denken ons hier, voor de eenvoudigheid van de formules, de helling alleen uit waarnemingen van de standaard en één praeparaat geschat (Bij méér dan één praeparaat wordt in de laatste formule - niet in de voorafgaanden - de tweede term in de teller nog van een $\sum^{(r)}$ -teken voorzien en ondergaat ook Q een dergelijke verandering.) Verder gaat $\underline{z}(\underline{r}, \underline{t})$ over in:

$$\frac{1}{b} \left(\frac{\sum p_j \cdot \tilde{t}_j \cdot \underline{S}_{\underline{z}_j}^2}{N_p} - \frac{\sum m_j \cdot \tilde{t}_j \cdot \underline{S}_{\underline{y}_j}^2}{N_s} \right)$$

(zie (8.1.1)).

Er ontstaat nu:

$$\underline{S}_{\underline{M}}^2 = \frac{1}{b^2} \left\{ \frac{\sum p_j \cdot \underline{S}_{\underline{z}_j}^2}{N_p^2} + \frac{\sum m_j \cdot \underline{S}_{\underline{y}_j}^2}{N_s^2} + \frac{(\tilde{\underline{z}} - \tilde{\underline{y}})^2 (\sum m_j \cdot \tilde{t}_j^2 \cdot \underline{S}_{\underline{z}_j}^2 + \sum p_j \cdot \tilde{t}_j^2 \cdot \underline{S}_{\underline{y}_j}^2)}{b^2 Q^2} \right.$$

(4.9.1)

$$\left. - \frac{2(\tilde{\underline{z}} - \tilde{\underline{y}})}{bQ} \left(\frac{\sum p_j \cdot \tilde{t}_j \cdot \underline{S}_{\underline{z}_j}^2}{N_p} - \frac{\sum m_j \cdot \tilde{t}_j \cdot \underline{S}_{\underline{y}_j}^2}{N_s} \right) \right\}.$$

Om deze nieuwe formule voor het spreidingskwadraat van $\underline{M}$ eenvoudiger numeriek te kunnen vergelijken met de formule die Irwin gebruikt onder het wortelteken in (2.5.3) zullen we onderstellen dat in alle groepen het aantal proefdieren gelijk is en wel gelijk aan n . Hierdoor gaat (4.0.1) over in:

$$(4.0.2) \quad s_{\underline{M}}^2 = \frac{1}{b^2} \left\{ \frac{\sum s_{u_j}^2}{nk^2} + \frac{\sum s_{v_j}^2}{nh^2} + \frac{(\bar{Z} - \bar{Y})^2}{b^2} \frac{(\sum \check{s}_{j^2}^2 s_{u_j}^2 + \sum \check{\pi}_{j^2}^2 s_{v_j}^2)}{b^2} + \right. \\ \left. - \frac{2(\bar{Z} - \bar{Y})}{b^2} \left(\frac{\sum \check{\pi}_{j^2}^2 s_{v_j}^2}{k} - \frac{\sum \check{s}_{j^2}^2 s_{u_j}^2}{h} \right) \right\}.$$

De benadering van Irwin, gelijkheid der spreidingen onderstellende, is uit (4.0.2) af te leiden, wanneer we bedenken, dat gelijkheid van de spreidingen ten gevolge heeft, dat de covariantieterm $\check{Z}(\underline{r}, \underline{t}) = 0$ is (§ 8.1) en alle $s_{v_j}^2 = s_{u_j}^2 = s_{\underline{u}}^2$ zijn. We vinden dan:

$$(4.0.3) \quad s_{\underline{M}}^2 = \frac{1}{b^2} \left\{ \frac{s_{\underline{u}}^2}{nk} + \frac{s_{\underline{u}}^2}{nh} + \frac{(\bar{Z} - \bar{Y})^2}{b^2} \frac{s_{\underline{u}}^2}{Q} \right\}.$$

Bedenken we dat in het eenvoudige geval met gelijke aantallen per groep en alleen waarnemingen van de standaard en één praeparaat:

$$s_{\underline{u}}^2 = \frac{\sum s_{u_j}^2 + \sum s_{v_j}^2}{h + k},$$

(zie (2.4.3)) dan vinden we voor het verschil $s_{\underline{M}}^2 - s_{\underline{M}}'^2$, dus voor de "correctie term" op het spreidingskwadraat volgens Irwin:

$$s_{\underline{M}}^2 - s_{\underline{M}}'^2 = \frac{1}{nb^2} \left\{ \frac{\sum s_{v_j}^2}{k^2} + \frac{\sum s_{u_j}^2}{h^2} - \frac{s_{\underline{u}}^2}{k} - \frac{s_{\underline{u}}^2}{h} \right\} + \\ + \frac{(\bar{Z} - \bar{Y})^2}{b^4 Q^2} \left\{ \sum \check{s}_{j^2}^2 (s_{u_j}^2 - s_{\underline{u}}^2) + \sum \check{\pi}_{j^2}^2 (s_{v_j}^2 - s_{\underline{u}}^2) \right\} + \\ - \frac{2(\bar{Z} - \bar{Y})}{b^3 Q} \left(\frac{\sum \check{\pi}_{j^2}^2 s_{v_j}^2}{k} - \frac{\sum \check{s}_{j^2}^2 s_{u_j}^2}{h} \right) = \\ = \frac{1}{nb^2} \left(\frac{1}{h} - \frac{1}{k} \right) \left(\frac{\sum s_{u_j}^2}{h} - \frac{\sum s_{v_j}^2}{k} \right) + \\ + \frac{n(\bar{Z} - \bar{Y})}{b^4 Q^2} \left(\sum \check{s}_{j^2}^2 (s_{u_j}^2 - s_{\underline{u}}^2) + \sum \check{\pi}_{j^2}^2 (s_{v_j}^2 - s_{\underline{u}}^2) \right) +$$

$$+ \frac{2(\bar{Z}-\bar{Y})}{b^3 Q} \left(\frac{\sum_j \check{f}_j \cdot \underline{s}_{u,j}^2}{h} - \frac{\sum_j \check{f}_j \cdot \underline{s}_{v,j}^2}{k} \right).$$

De "correctie" op het spreidingskwadraat volgens Irwin bestaat dus uit drie bestanddelen:

$$A = \frac{1}{nb^2} \left(\frac{1}{h} - \frac{1}{k} \right) \left(\frac{\sum \underline{s}_{u,j}^2}{h} - \frac{\sum \underline{s}_{v,j}^2}{k} \right), \text{ welke ontstaan}$$

is als correctie op $s^2 \{ \bar{Z}-\bar{Y} \}$;

$$(4.0.4) \quad B = \frac{n(\bar{Z}-\bar{Y})^2}{b^4 Q^2} \left(\sum_j \check{f}_j^2 (\underline{s}_{u,j}^2 - \underline{s}_u^2) + \sum_j \check{f}_j^2 (\underline{s}_{v,j}^2 - \underline{s}_v^2) \right),$$

ontstaan als correctie op s_b^2 ;

$$C = \frac{2(\bar{Z}-\bar{Y})}{b^3 Q} \left(\frac{\sum_j \check{f}_j \cdot \underline{s}_{u,j}^2}{h} - \frac{\sum_j \check{f}_j \cdot \underline{s}_{v,j}^2}{k} \right), \text{ een extra term,}$$

dooradrt bij ongelijkheid der spreidingen de covariantie-term tussen $\bar{Z} - \bar{Y}$ en b niet meer wegvalt.

Bij ijkings met meer dan één praeparaat is de correctie-term A niet meer in deze overzichtelijke vorm te schrijven, de term B krijgt een $\sum^{(p)}$ -teken voor de term

$$\sum_j \check{f}_j^2 (\underline{s}_{v,j}^2 - \underline{s}_v^2), \text{ de term C blijft onveranderd, terwijl}$$

overal de schattingen b en $\underline{s}_u^2$ en ook Q uit de standaard en alle praeparaten als in (2.3.1), (2.4.3) en (2.3.2) berekend moet worden.

In tabel VII zijn deze correctietermen voor de verschillende uitgevoerde ijkings opgenomen, met $S'_M{}^2$ en de verhouding tussen de totale correctie en $S'_M{}^2$. Hierbij is steeds $n = 12$ gesteld, ook in die weinig voorkomende gevallen, waar eigenlijk $n=11$ is.

Bij vergelijking van tabel VII met tabel III blijkt dat de ijkings 503 en 5017 de belangrijkste correctie op $S'_M{}^2$ en de grootste afwijkingen van homoscedasticiteit bezitten, terwijl de overige ijkings, die beter aan de eis van homoscedasticiteit voldeden, slechts kleine correcties geven.

De invloed van deze correcties op het betrouwbaarheidsinterval zullen we laten zien aan de hand van enkele getallen voorbeelden, opgenomen in tabel VIII. In deze tabel zijn de grenzen van het interval uitgedrukt in % van de sterkteverhouding en, analoog aan (2.5.4), als volgt berekend:

Tabel VII
Correctietermen bij S_M^2 in het niet-homoscedastische geval.

ijkning no.	aan- tal velden	S_M^2	correctietermen			totale correctie	tot. corr. S_M^2
			A	B	C		
4926 (5)	1	15.965	1.035	-0.0523	0.814	1.797	0.118
praeparaat	2	0.1557	-0.0085	-0.00000	-0.00014	-0.0086	-0.055
1 (2)	3	0.05308	+0.0026	0.00002	0.0015	0.0041	0.077
	4	0.0389	0.0057	0.00001	0.00031	0.0060	0.154
	5	0.0330	0.0032	0.00001	-0.000022	0.0032	0.097
4926 (5)	1	89.309	-9.684	-2.925	-2.55	-15159	-0.170
praeparaat	2	0.4036	0.092	-0.00012	-0.0021	0.090	0.223
2 (1)	3	0.1035	0.0081	0.00099	0.0069	0.016	0.155
	4	0.0804	-0.0129	0.00082	-0.00034	-0.012	-0.149
	5	0.0536	-0.0178	0.00107	-0.00241	-0.019	-0.354
4926 (5)	1	25.134	3.991	-0.0142	-0.178	3.809	0.152
praeparaat	2	0.411	0.046	-0.0017	0.0078	0.052	0.127
3 (1)	3	0.1302	0.027	0.0038	0.0135	0.044	0.338
	4	0.0848	0.017	0.0029	-0.00065	0.020	0.236
	5	0.0699	0.0020	0.0031	-0.0041	0.001	0.014
503 (5)	1	0.1730	0.0795	-0.0138	0.0209	0.0866	0.501
praep.1 (1)	2	0.0601	0.0401	-0.0058	0.0097	0.0450	0.749
	3	0.0417	0.0252	-0.0044	0.0083	0.0291	0.698
503 (5)	1	0.1549	0.0617	-0.0099	0.0177	0.0693	0.447
praep.2 (1)	2	0.0544	0.0516	-0.0017	0.0057	0.0556	1.022
	3	0.0459	0.0626	-0.0016	0.0050	0.0660	1.438
503 (5)	1	0.1175	0.0266	-0.0008	0.0050	0.0308	0.262
praep. 3 (1)	2	0.0472	0.0288	-0.0014	0.0052	0.0326	0.691
	3	0.0364	0.0282	-0.0016	0.0050	0.0316	0.868
503 (5)	1	0.1887	0.0167	-0.0267	-0.0291	-0.0391	-0.207
	2	0.0570	-0.0088	-0.0110	-0.0146	-0.0344	-0.604
	3	0.0407	-0.0018	-0.0088	-0.0118	-0.0224	-0.550
5011 (3)	1	0.1842	-0.0053	0.00005	0.0019	-0.0034	-0.018
praep.3F (2)	2	0.0398	-0.0018	0.00000	-0.0001	-0.0019	-0.048
	3	0.0465	0.0056	0.00002	-0.0005	0.0051	0.110
	4	0.0452	0.0053	0.00002	-0.0001	0.0052	0.115
	5	0.0431	0.0032	-0.00003	0.0012	0.0044	0.102
5011 (3)	1	0.2277	-0.0007	0.0011	0.0308	0.0312	0.137
praep.HJ (2)	2	0.0659	0.0012	0.0013	0.0185	0.0210	0.319
	3	0.0828	0.0047	0.0009	0.0306	0.0362	0.437
	4	0.0806	0.0062	0.0011	0.0164	0.0237	0.294
	5	0.0966	0.0063	-0.0008	0.0284	0.0339	0.351
5017 (4)	1	6.892	1.344	-0.011	0.991	2.324	0.337
praep. (2)	2	1.177	0.049	-0.055	0.160	0.154	0.131
	3	0.791	0.130	-0.013	0.231	0.348	0.440
	4	0.546	0.106	-0.009	0.190	0.287	0.526
	5	0.420	0.075	-0.002	0.201	0.274	0.652
5023 (3)	1	0.02128	-0.00157	0.00001	-0.00041	-0.00117	-0.055
praep.1 (1)	2	0.00847	-0.00105	0.00092	-0.00042	-0.00055	-0.065
	3	0.00802	-0.00117	0.00108	-0.00036	-0.00045	-0.056
	4	0.00907	-0.00186	0.00136	0.00032	-0.00018	-0.020
	5	0.00835	-0.00177	0.00110	-0.00017	-0.00084	-0.101
5023 (3)	1	0.01432	-0.00250	0.00016	0.00019	-0.00215	-0.150
praep.2 (1)	2	0.00672	-0.00104	0.00037	0.00027	-0.00040	-0.060
	3	0.00520	-0.00141	0.00024	0.00017	-0.00100	-0.192
	4	0.00643	-0.00148	0.00037	-0.00017	-0.00094	-0.146
	5	0.00651	-0.00140	0.00046	0.00011	-0.00083	-0.127
5023 (3)	1	0.01864	-0.00162	0.00055	-0.00034	-0.00141	-0.076
praep.3 (1)	2	0.00570	-0.00210	0.00023	-0.00021	-0.00208	-0.365
	3	0.00505	-0.00105	0.00012	-0.00012	-0.00105	-0.206
5023 (3)	1	0.02119	0.00556	0.00035	-0.00027	0.00564	0.266
praep.4 (1)	2	0.00664	0.00020	0.00013	-0.00016	0.00017	0.026
	3	0.00542	-0.00059	0.00016	-0.00014	-0.00057	-0.105
	4	0.00597	-0.00150	0.00022	0.00013	-0.00115	-0.193
	5	0.00550	-0.00135	0.00014	-0.00006	-0.00127	-0.231

De bovengrens:

$$\frac{100 \cdot g^{M+t_{\varepsilon}^* \cdot S'_M}}{g^M} = 100 \cdot g^{t_{\varepsilon}^* \cdot S'_M} \% \text{ van } M;$$

de benedengrens:

$$\frac{100 \cdot g^{M-t_{\varepsilon}^* \cdot S'_M}}{g^M} = 100 \cdot g^{-t_{\varepsilon}^* \cdot S'_M} \% \text{ van } M,$$

en net zo voor S_M in plaats van S'_M . Hierin is g de basis van de logaritmen (in "log dosis" enz.) en t de bij de onbetrouwbaarheid ε behorende waarde uit de normale verdeling. In tabel VIII hebben we $\varepsilon = 0.05$, dus $t_{\varepsilon}^* = 1.96$ genomen.

Tabel VIII

Invloed van de correctie op S'_M op het betrouwbaarheidsinterval.

ijking no.	aan- tal weken	S'_M	S_M	grenzen in % van de sterkteverhouding	
				met S'_M	met S_M
4926 (5) praep. 3 (1)	3	0.1302	0.1742	162 - 63	175 - 57
503 (5) praep. 1 (1)	2	0.0601	0.1051	139 - 72	153 - 65
praep. 2 (1)	2	0.0544	0.1100	140 - 71	161 - 62
	3	0.0459	0.1119	132 - 76	156 - 64
5017 (4) praep. (2)	5	0.420	0.694	160 - 62	174 - 57

Uit tabel VIII blijkt dat bij afwijking van homoscedasticiteit met een behoorlijke verbreding van het betrouwbaarheidsinterval rekening moet worden gehouden. Het is derhalve van belang na te gaan, hoe de correcties (4.0.4) zo klein mogelijk gehouden kunnen worden. Dit is, voor het geval van één praeparaat, uit deze formules gemakkelijk te zien: Bij een ijking beheerst men de spreiding niet, zodat men niet voor homoscedasticiteit kan zorgen. Wel kan men trachten aan de volgende drie voorwaarden zo goed mogelijk te voldoen:

- 1° $h=k$,
- 2° $\bar{y} = \bar{z}$,
- 3° Q groot, dus de doses ver uit elkaar.

Door 1° wordt correctie A gelijk aan nul (voor meer dan één praeparaat is er niet zulk een eenvoudige weg om deze

correctie te doen verdwijnen); 2° en 3° maken B en C klein (dit blijft voor meer dan één praeparaat mutatis mutandis wel geldig). De betekenis van 2° is: de gemiddelde reacties op standaard en praeparaat stemmen zo goed mogelijk overeen.

Hoewel over het teken van de totale correcties weinig te voorspellen valt, blijkt toch uit tabel VII, dat de belangrijkste correcties voor het merendeel positief zijn. Berekening van het betrouwbaarheidsinterval met S'_M geeft in deze gevallen een te klein betrouwbaarheidsinterval. Tenslotte blijkt nog uit tabel VII dat de correctieterm B, bij de onderhavige ijkingen geen belangrijke bijdrage voor de totale correctie levert.

Conclusie: De invloed van de afwijkingen van homoscedasticiteit op de lengte van de betrouwbaarheidsintervallen is zo groot dat het de moeite loont de correctie te berekenen en te trachten hen zo klein mogelijk te houden. Dit laatste kan gebeuren door voor standaard en praeparaat evenveel doses te nemen (bij een ijking met een standaard en slechts één praeparaat), te zorgen dat de gemiddelde reacties van standaard en praeparaat zo goed mogelijk overeen stemmen en door de doses zo ver als het lineariteitsgebied toelaat uit elkaar te kiezen (de laatste twee punten gelden ook bij meer dan één praeparaat).

5. Beantwoording der vragen.

We zullen nu overgaan tot een bespreking van de in § 1 gestelde vragen.

5.1 Keuze der methode.

Is een ijking verricht dan hangt de keuze van de methode om de sterkteverhouding met zijn nauwkeurigheid te berekenen af van het al dan niet vervuld zijn der onderstellingen (2.2.1). Is niet voldaan aan de lineariteit of evenwijdigheid der werkingslijnen dan is een bepaling van de sterkteverhouding in het geheel niet mogelijk (§ 3.4). Is aan deze voorwaarden wel voldaan, dan wordt getoetst, of de gezamenlijke helling van de standaard en praeparaat van nul verschilt. Deze toets dient een significant resultaat te geven, daar ijking bij helling 0 geen zin heeft. Tenslotte kunnen we dan nog twee gevallen onderscheiden:

- a) er zijn geen afwijkingen tussen de spreidingen binnen de groepen, dus voorwaarden (2.2.1) zijn vervuld,
- b) er zijn wel afwijkingen tussen de spreidingen binnen de groepen.

In geval a) kan de methode van IRWIN toegepast worden met een berekening van exacte betrouwbaarheidsintervallen volgens FIELLER (§ 2). Is de spreiding van $\underline{b}$ klein genoeg t.o.v. de gevonden waarde $\underline{b}$ dan kunnen deze betrouwbaarheidsintervallen volgens de benadering van Irwin (2.5.3) bepaald worden. Deze benadering ontstaat, doordat in de formule van de exacte betrouwbaarheidsintervallen (2.5.2)

$$C = \frac{b^2}{b^2 - t_{\epsilon}^* \cdot 2 S_b^2} = 1$$

gesteld wordt. Door Irwin is de fout die hierbij gemaakt wordt in [5] besproken en numeriek aan een aantal voorbeelden onderzocht. Hierbij bleek dat wanneer $\frac{b}{S_b}$ kleiner is

dan ongeveer 7 de berekende betrouwbaarheidsgrenzen merkbaar van de exacte beginnen af te wijken. In § 2 zagen we dat de benadering van Irwin een vernauwing en een systematische verschuiving van het interval tegevolge heeft. De grootte van deze vernauwing en van de verschuiving hangt, behalve van de grootte van C , ook af van $(Z-\bar{Y})$, de richting van de verschuiving, hangt af van het teken van $(Z-\bar{Y})$ (§ 2,5). Voor de beschikbare ijkingen zijn, uit alle waarnemingen, $\frac{b}{S_b}$ berekend en in tabel VI verzameld. Uit die

tabel blijkt dat bij deze ijkingen bij 3 of meer weken $\frac{b}{S_b}$ voldoende groot is om de benadering van Irwin te rechtvaardigen. Is bij een ijking niets over de verhouding $\frac{b}{S_b}$

bekend dan verdient het aanbeveling eerst te controleren of deze voldoende groot is. In dit geval, nl. als toch al $\underline{b}$ en S_b berekend zijn, is het echter niet veel extra werk meer om de exacte betrouwbaarheidsintervallen volgens Fieller te berekenen.

In geval b), wanneer dus niet aan de eis van homoscedasticiteit voldaan is, is geen exacte methode der bepaling van betrouwbaarheidsintervallen beschikbaar (§ 2.5). Het is nu beter gebruik te maken van de schatting S_M^2 (4.0.2) en voor de grenzen van μ te nemen:

$$(5.1.1) \quad M \pm t_{\epsilon}^* \cdot S_M$$

waarbij dan de onbetrouwbaarheid niet gelijk aan ϵ is, maar toch van ϵ niet sterk zal afwijken. Met behulp van

deze schatting verkrijgt men een betere indruk van de nauwkeurigheid dan met de schatting S'_M van Irwin. Volgens tabel VII leidt de benadering van Irwin bij de meeste van de onderzochte ijkingen, waar afwijkingen van homoscedasticiteit werden geconstateerd, tot te nauwe intervallen, in vergelijking met de door (5.1.1) berekende.

Volgens dit recept moeten uit de ijkingen 503 en 5017 de grenzen voor de sterkteverhouding volgens (5.1.1) berekend worden en uit de ijkingen 4926, 1 en 2 weken, en 5011, 1 week volgens de exacte methode van Fieller. Bij de overige ijkingen kan met de benadering van Irwin worden volstaan.

Samenvatting: Bij een ijking wordt eerst lineariteit en evenwijdigheid getoetst; indien deze voorwaarden vervuld zijn, wordt de helling getoetst, die $\neq 0$ behoort te zijn. Vervolgens toetst men de homoscedasticiteit. De uitslag van deze toets bepaalt de keus der methode: is de homoscedasticiteit vervuld en $\frac{b}{s_b}$ groot (> 7) dan volgen de grenzen voor μ uit (2.5.3); is $\frac{b}{s_b}$ klein dan uit (Fieller)/(2.5.2); is de homoscedasticiteit niet vervuld dan volgen de grenzen voor μ uit (5.1.1).

5.2 Ijking van meer dan één praeparaat tegelijk.

In § 2 is de methode van Irwin reeds uitgebreid voor meer dan één praeparaat tegelijk. De schattingen $\underline{b}$ en $\underline{s}_u^2$ worden uit het gehele waarnemingsmateriaal berekend, waarna voor ieder praeparaat een schatting van de sterkteverhouding ten opzichte van de standaard met bijbehorend betrouwbaarheidsinterval volgens één der 3 beschouwde methoden berekend worden. Ijking met slechts één dosis van het praeparaat is zeker mogelijk, een bezwaar is dan echter de onmogelijkheid lineariteit en evenwijdigheid te toetsen en de grotere onnauwkeurigheid van de uitkomst. Dit laatste komt ter sprake bij de beantwoording van de volgende vraag.

5.3 Aantal doses.

Om de vraag naar het geschiktste aantal doses van standaard en praeparaat, in verband met de nauwkeurigheid van de uitkomst, te beantwoorden, denken wij ons de bepaling van de sterkteverhouding $\underline{M}$ uitgevoerd door een ijking, waarbij aan de voorwaarden (2.2.1) voldaan is en we de

werkelijke waarde van de helling, de spreiding in de groepen en de sterkteverhouding al kennen. Dit maakt het nl. mogelijk, de invloed der aantallen doses op eenvoudige wijze te berekenen. We weten dus:

$$\begin{aligned} \sum \underline{b} &= \beta, \\ \sum \underline{M} &= \mu, \end{aligned}$$

en de spreiding binnen de groepen: $\sigma_{\underline{u}}^2$.

$\sigma_{\underline{M}}^2$ kan nu met behulp van de volgende formules berekend worden:

$$(2.3.4) \quad \underline{M} = \frac{c-a}{\underline{b}},$$

$$(8.5.1) \quad \sigma^2 \left\{ \frac{c-a}{\underline{b}} \right\} = \frac{\sigma^2 \{c-a\}}{\beta^2} + \frac{(\gamma - \alpha)^2 \sigma_{\underline{b}}^2}{\beta^4} + \frac{2(\gamma - \alpha) \sum \{(\tilde{c} - \tilde{a}) \tilde{b}\}}{\beta^3},$$

$$\sigma^2 \{c-a\} = \sigma^2 \{ \underline{Z} - \underline{Y} + \underline{b}(\bar{\xi} - \bar{\pi}) \} =$$

(volgens § 8.1)

$$= \sigma^2 \{ \underline{Z} - \underline{Y} \} + (\bar{\xi} - \bar{\pi})^2 \sigma_{\underline{b}}^2 =$$

$$= \sigma_{\underline{u}}^2 \left(\frac{1}{N_S} + \frac{1}{N_P} \right) + (\bar{\xi} - \bar{\pi})^2 \sigma_{\underline{b}}^2,$$

$$\sum \{(\tilde{c} - \tilde{a}) \tilde{b}\} = \sum \{(\tilde{Z} - \tilde{Y} + \tilde{b}(\bar{\xi} - \bar{\pi})) \tilde{b}\} = \sum \{(\tilde{Z} - \tilde{Y}) \tilde{b}\} + (\bar{\xi} - \bar{\pi}) \sum \tilde{b}^2 =$$

(volgens § 8.1)

$$= 0 + (\bar{\xi} - \bar{\pi}) \sigma_{\underline{b}}^2,$$

uit (2.4.1)

$$\sigma_{\underline{b}}^2 = \frac{\sigma_{\underline{u}}^2}{\sum_j m_j \cdot \tilde{\xi}_j^2 + \sum_j p_j \cdot \tilde{\pi}_j^2}.$$

We vinden:

$$(5.3.1) \quad \sigma_{\underline{M}}^2 = \frac{\sigma_{\underline{u}}^2}{\beta^2} \left\{ \frac{1}{N_S} + \frac{1}{N_P} + \frac{(\bar{\pi} - \bar{\xi} + \mu)^2}{\sum_j m_j \cdot \tilde{\xi}_j^2 + \sum_j p_j \cdot \tilde{\pi}_j^2} \right\}.$$

We bereiken een maximale nauwkeurigheid bij de bepaling van $\underline{M}$, wanneer $\sigma_{\underline{M}}^2$ minimaal is. De eerste twee termen zijn onafhankelijk van de verdeling der dieren over de verschillende doses. $\frac{1}{N_S} + \frac{1}{N_P}$ is minimaal wanneer we van alle beschikbare dieren de helft voor de standaard en eb

de helft voor het praeparaat gebruiken. Willen we echter meer dan één praeparaat gelijktijdig ijken en alle praeparaten op gelijke wijze behandelen, dan moeten we voor elk praeparaat evenveel dieren gebruiken. We vinden dan de beste waarden voor N_S en N_P door $\frac{1}{N_S} + \frac{1}{N_P}$ te minimaliseren onder de voorwaarde $N_S + nN_P = N$, als n het aantal praeparaten en N het totaal aantal beschikbare dieren voorstelt. We vinden dan:

$$(5.3.2) \quad N_S = \frac{N}{1 + \sqrt{n}} \quad \text{en} \quad N_P = \frac{N}{n + \sqrt{n}}.$$

De laatste term in (5.3.1), wordt minimaal, wanneer

$$\sum m_j \check{\xi}_j^2 \quad \text{en} \quad \sum p_j \check{\pi}_j^2 \quad \text{maximaal worden. In de eerste}$$

plaats kunnen we opmerken dat het steeds beter is de dieren voor één praeparaat over meer dan één dosis te verdelen, daar bij één praeparaat-dosis $\sum p_j \check{\pi}_j^2 = 0$ wordt.

Gaan we er nu van uit dat we de log doses op gelijke afstanden kiezen en voor elke dosis evenveel dieren en (en wel m) gebruiken bij de standaard (dezelfde argumenten kunnen we voor elk der praeparaten aanvoeren), dus d

dat:

$$\bar{\xi} = \frac{\sum \xi_j}{h}$$

In het midden van het interval tussen de laagste en de hoogste dosis ligt, dan volgt, dat

$$\sum \check{\xi}_j^2 = \sum (\xi_j - \bar{\xi})^2$$

maximaal wordt, wanneer we alle dieren over de laagste en de hoogste doses verdelen. Wat dit betreft zou men dus het beste 2 doses kunnen nemen. Dit valt echter slechts dan zonder voorbehoud aan te bevelen, indien de lineariteit (dus ook het lineariteitsgebied) reeds voldoende bekend is, daar de lineariteit bij 2 doses niet getoetst kan worden.

Tenslotte wordt de grootte van σ_M^2 nog beïnvloed door $(\bar{\pi} - \bar{\xi} + \mu)^2$. Kiezen we de afstand tussen de gemiddelde doses van standaard en praeparaat zo goed mogelijk gelijk aan μ , dan zal $(\bar{\pi} - \bar{\xi} + \mu)^2$ klein blijven en weinig bijdargen tot σ_M^2 . Dit komt overeen met ongeveer gelijke gemiddelde reacties bij standaard en praeparaat.

Uit het bovenstaande volgt dat we een zo nauwkeurig mogelijke ijking krijgen, wanneer we (b.v. uit vroegere ijkingen) reeds weten waar het lineariteitsgebied van standaard en praeparaat ongeveer ligt, hoe groot ongeveer de sterkteverhoudingen zullen zijn, voor de standaard $N_S \approx \frac{N}{1+\sqrt{n}}$ en voor de praeparaten $N_P \approx \frac{N}{n+\sqrt{n}}$ dieren gebruiken en zowel bij standaard als praeparaten de dieren over twee doses verdelen die we zo ver als het lineariteitsgebied toelaat uit elkaar kiezen.

Opmerkingen:1.: Over het verband tussen σ_M en het aantal praeparaten n kunnen we nog het volgende afleiden:

Gebruiken we voor de standaard $N_S = \frac{N}{1+\sqrt{n}}$ dieren

en voor elk van de praeparaten $N_P = \frac{N}{n+\sqrt{n}}$, verdelen

we deze dieren steeds over twee doses, waarbij steeds de log doses op gelijke afstanden liggen, noemen we deze afstand tussen de log doses 1, dan is steeds

$$\check{s}_j^2 = \check{\pi}_j^2 = \frac{1}{4}$$

en volgt uit (5.3.1):

$$\begin{aligned} \sigma_M^2 &= \frac{\sigma_u^2}{s^2} \left\{ \frac{1}{N/(1+\sqrt{n})} + \frac{1}{N/(n+\sqrt{n})} + \frac{4(\bar{x} - \bar{f} + \mu)^2}{N} \right\} \\ (5.3.3) \quad &= \frac{\sigma_u^2}{s^2 N} \left\{ (1+\sqrt{n})^2 + 4(\bar{x} - \bar{f} + \mu)^2 \right\} \end{aligned}$$

Kiezen we $\bar{f} - \bar{x}$ zo goed mogelijk gelijk aan μ dan is dus σ_M^2 ongeveer evenredig met $(1+\sqrt{n})^2$. Bij overgang van 1 op 2 praeparaten zal dus σ_M voor ieder praeparaat ongeveer 20 à 25% groter worden. De invloed van n is dus vrij aanzienlijk.

2. Bij een ijking, alsboven beschreven, zal men in het algemeen de afstand tussen de log doses van de standaard groter kunnen nemen dan die van het praeparaat, omdat het lineariteitsgebied bij de standaard in de regel beter bekend is.

5.4 Duur van een ijking en combinatie van de waarnemingen van verschillende weken.

Ter beantwoording van de vraag naar het aantal weken of de combinatie van weken, welke de nauwkeurigste uitkomst voor de sterkteverhouding zal opleveren is S_M^2 berekend met behulp van (4.0.3), (2.4.3) en (2.3.1) uit de waarnemingen na 1, 2, 3, 4 en 5 weken (dit waren de gegeven waarnemingen); verder uit het materiaal dat we verkrijgen door van de waarnemingen uit de verschillende weken die uit de eerste week af te trekken (in tabel IX onder het hoofd "aantal weken" met 2-1, 3-1, 4-1 en 5-1 aangegeven), en uit de sommen van de waarnemingen na 1, 2, 3, 4 en 5 resp. na 3, 4 en 5 weken (aangegeven met $1+2+3+4+5$). Er is vanaf gezien bij de ijkingen die afwijkingen van homoscedasticiteit vertonen S_M^2 (4.0.2) in plaats van S_M^2 te berekenen, ter besparing van extra rekenwerk, daar het te verwachten is dat S_M^2 door de verschillende weken heen hetzelfde gedrag zal vertonen als S_M^2 .

F_{uap}. 3+4+5

Uit tabel IX blijkt, dat een keus betreffende het aantal weken of een combinatie van weken gedaan zal moeten worden uit de weken 3, 4, 5, 1+2+3+4+5 en 3+4+5 terwijl ook 3-1 geen gek figuur slaat; 3 weken geeft bij de ijkingen 4926 en 5017 nog een vrij grote onnauwkeurigheid in vergelijking met 4 en 5 weken, maar bij de ijkingen 5011 en 5023 geeft verlenging van de duur geen nauwkeuriger uitkomst meer. In verband met de tijdsbesparing welke het oplevert valt het te overwegen, de ijkingen slechts 3 weken te laten duren.

Tabel IX

(zie blz. 36)

Tabel IX

S_M^2 na de verschillende weken of combinaties van weken.

ijking no.	aantal we- ken	S_M^2	ijking no.	aantal we- ken	S_M^2
4926 (5) praep.1(2)	1	15.188	5017(4)	1	6.892
	2	0.229	praep.(2)	2	1.177
	3	0.0599		3	0.791
	4	0.0385		4	0.546
	5	0.0310		5	0.420
	2-1	0.236		2-1	0.815
	3-1	0.0467		3-1	0.581
	4-1	0.0313		4-1	0.383
	5-1	0.0521		5-1	0.315
	1+2+3+4+5	0.0516		1+2+3+4+5	0.672
	3+4+5	0.0378		3+4+5	0.434
4926 (5) praep.2 (1)	1	99.32	503 (5)	1	0.1876
	2	0.396	praep.1(1)	2	0.0712
	3	0.110		3	0.0582
	4	0.0717		2-1	0.0737
	5	0.0588		3-1	0.0573
	2-1	0.408		1+2+3	0.0675
	3-1	0.0888	503 (5)	1	0.174
	4-1	0.0598	praep.2(1)	2	0.0607
	5-1	0.122		3	0.0497
	1+2+3+4+5	0.145		2-1	0.0611
	3+4+5	0.0713		3-1	0.0480
5011 (3) praep.1(2) (SF)	1	0.1353		3-1	0.0475
	2	0.0426		1+2+3	0.0847
	3	0.0501	5023 (3)	1	0.0225
	4	0.0511	praep.1(1)	2	0.00753
	5	0.0506		3	0.00694
	2-1	0.0538		4	0.00786
	3-1	0.0692		5	0.00727
	4-1	0.0734		2-1	0.00864
	5-1	0.0721		3-1	0.00769
	1+2+3+4+5	0.0415		4-1	0.00928
	3+4+5	0.0465		5-1	0.00842
5011 (3) praep.2(2) (HJ)	1	0.1565		1+2+3+4+5	0.00723
	2	0.0612		3+4+5	0.00758
	3	0.0929	5023 (3)	1	0.01763
	4	0.0877	praep.4(1)	2	0.00531
	5	0.0985		3	0.00441
	2-1	0.0969		4	0.00500
	3-1	0.1610		5	0.00459
	4-1	0.1473		2-1	0.00601
	5-1	0.1654		3-1	0.00456
	1+2+3+4+5	0.0710		4-1	0.00560
	3+4+5	0.0835		5-1	0.00509
5023 (3) praep.3(1)	1	0.01973		1+2+3+4+5	0.00471
	2	0.00560		3+4+5	0.00499
	3	0.00431			
	2-1	0.00611			
	3-1	0.00438			

6. Opmerkingen en suggesties.

6.1 De ongelijkheid der spreidingen.

Daar de ongelijkheid der spreidingen, die herhaalde malen optreedt, erg onbevredigend is, is gezocht naar een afhankelijkheid tussen spreiding en gemiddelde reactie van een groep, daar een duidelijk verband wellicht in de statistische analyse betrekken zou kunnen worden. De spreidingen vertonen echter een zeer onregelmatig gedrag, waarin geen lijn te vinden is, zodat deze poging geen resultaat gaf. In verband met de belangrijkheid der homoscedasticiteit voor de berekening der betrouwbaarheidsintervallen en het toetsen van lineariteit en evenwijdigheid is een uitgebreider onderzoek naar deze kwestie sterk aan te bevelen. Wellicht kunnen hiervoor vroegere resultaten gebruikt worden. Mogelijke oorzaken voor afwijkingen van homoscedasticiteit kunnen schuilen in de verdeling der dieren over de verschillende groepen. Zie hierover JERNE en WOOD [8] vooral de onderstellingen A1 en B6, of EMMENS [6] hoofdstuk 6.

6.2 Oriënterende proeven.

In §5.3 is besproken dat de ijking nauwkeuriger wordt, wanneer we voor standaard en praeparaten elk de dieren over 2 doses verdelen (in plaats van hetzelfde aantal dieren voor meer dan twee doses te gebruiken), en deze doses zover mogelijk uit elkaar te kiezen. Bij een dergelijke ijking is de mogelijkheid de lineariteit te toetsen verloren gegaan, terwijl juist de kennis van het lineariteitsgebied in verband met de afstand der log doses van veel belang is. Het is daarom wenselijk, indien men deze tactiek volgt, van tijd tot tijd controleijkingen uit te voeren met van standaard en praeparaat meer dan twee doses, teneinde de lineariteit te toetsen, de grenzen van lineariteit te bepalen. Dit schema van ijkingen is b.v. zeer geschikt voor regelmatige controle van een praeparaat, dat geregeld wordt gefabriceerd.

Indien volledig onbekende praeparaten moeten worden geijkt is toetsing van lineariteit, evenwijdigheid en helling steeds van belang. Daartoe heeft men zowel van standaard als van praeparaat 3 doses nodig. Vooral, indien de gemiddelde werking van het praeparaat strek verschilt van die van de standaard, moet men op afwijkingen van lineariteit in één van beide of beide en op niet-evenwijdigheid verdacht zijn. Voor dit laatste punt levert ijking 5011,

praeparaat HJ een argument: bij deze ijking gaat immers een afwijking van evenwijdigheid gepaard met een groot verschil tussen de gemiddelde reacties op standaard en praeparaat.

In plaats van bovengenoemde controle-ijkingen zou ook een uitgebreid onderzoek van een veel gebruikte standaard zijn nut kunnen hebben. Bij biologische ijkingen wordt ondersteld dat het praeparaat een zuivere verdunning is van de standaard, wat hierin tot uitdrukking komt dat de werkingslijnen van beide slechts horizontaal t.o.v. elkaar verschoven zullen zijn. Hieruit volgt dat wanneer we weten welke reacties op de standaard nog binnen het lineariteitsgebied vallen, we reacties van het praeparaat welke buiten dit gebied vallen moeten afkeuren. Nu zal de absolute hoogte van de reacties binnen het lineariteitsgebied vermoedelijk van ijking tot ijking verschillen. Misschien is het mogelijk hier gebruik te maken van de "O-dosis" van de standaard, b.v. doordat het lineariteitsgebied relatief t.o.v. de reacties op deze O-dosis ongeveer constant blijft. Het bovenstaande is slechts een suggestie waarvan de bruikbaarheid door een uitgebreid onderzoek bewezen zal moeten worden.

6.3 Besparing rekenwerk.

Bij de bepaling van de sterkteverhouding met betrouwbaarheidsinterval en ook bij het toetsen van de verschillende onderstellingen, wordt veel rekenwerk bespaard door voor beide doses van een praeparaat of van de standaard evenveel dieren te gebruiken. Ook zou het werk verminderd worden wanneer de verhouding tussen beide doses steeds dezelfde is (afstand tussen log doses gelijk) maar dit is uit een ander oogpunt minder wenselijk, zoals uit opmerking 2 van § 5.3 blijkt.

7. Samenvatting.

1. De methode van IRWIN is toepasbaar op de ijkingen indien aan de onderstellingen: normaliteit en homoscedasticiteit van de waarnemingen en lineariteit en evenwijdigheid van de werkingslijnen en de helling $\neq 0$ voldaan is.

2. Van deze eisen levert vooral de homoscedasticiteit moeilijkheden op, slechts bij ijking 4926 is deze voor alle weken vervuld (§ 3.2). Een onderzoek hiernaar is sterk aan te bevelen.

3. Bij afwijkingen van homoscedasticiteit is het beter de in § 4 afgeleide schatting (4.0.1) te gebruiken voor het spreidingskwadraat van de log sterkteverhouding daar de schatting van Irwin (2.5.3) in vele gevallen te nauwe grenzen geeft (tabel VIII).

4. De formule (2.5.3) van Irwin voor een betrouwbaarheidsinterval is een benadering van de formule (2.5.2) van Fieller voor exacte betrouwbaarheidsintervallen, welke benadering echter pas voor $\frac{b}{s_b} < 7$ duidelijk

begint af te wijken. Bij de gegeven ijkingen was steeds $\frac{b}{s_b} > 7$ (op enkele eerste weken na) (§ 5.1).

5. De lineariteit was goed vervuld bij de onderzochte ijkingen. De evenwijdigheid was in één geval, van de weinige die gecontroleerd konden worden, niet vervuld n.l. bij ijking 5011 voor praeparaat HJ (oorzaak hiervan was wellicht dat de werkzame doses van praeparaat en standaard te ver uiteen lagen) (§ 3.3 en § 3.4). Behalve enkele malen in de eerste week, was steeds de helling significant verschillend van 0. De waarde van de hierbij gebruikte toetsen, wordt verminderd wanneer aan de eis van homoscedasticiteit niet voldaan is.

6. De methode van Irwin is in § 2 beschreven voor meer dan één praeparaat tegelijk.

7. Bij een gegeven aantal praeparaten verkrijgen we een zo nauwkeurig mogelijke ijking, door van N proefdieren er $N_S \approx \frac{N}{1 + \sqrt{n}}$ voor de standaard en $N_P \approx \frac{N}{n + \sqrt{n}}$ voor elk der n praeparaten te gebruiken en zowel voor standaard als voor praeparaten de dieren gelijk over twee doses te verdelen die zover als het lineariteitsgebied toelaat uit elkaar liggen, waarbij de afstand van de twee log doses voor praeparaten en standaard niet dezelfde hoeft te zijn (§ 5.3). Hierbij können proef- of controle-ijkingen goede diensten bewijzen.

8. Richtten we een ijking in op de in punt 7. beschreven wijze, dan vinden we dat, wanneer n het aantal praeparaten is, de spreiding van de logaritme van de sterkteverhouding ongeveer evenredig is met $(1 + \sqrt{n})$, wanneer de afstanden tussen de gemiddelde log doses van standaard

en praeparaten in goede overeenstemming met de logarithme der sterkteverhouding gekomen zijn (§ 5.3 opmerking 1).

9. Uit § 5.4 blijkt dat in het algemeen de nauwkeurigheid van de uitkomst bij 3, 4 en 5 weken en de combinaties (1+2+3+4+5) weken en (3+4+5) weken niet veel uiteenloopt, zodat met 3 weken zou kunnen worden volstaan, hetgeen een aanzienlijke tijdsbesparing inhoudt.

10. In verband met een nader onderzoek naar de homoscedasticiteit, normaliteit en vooral naar het lineariteitsgebied is het zeer aan te bevelen oriënterende proeven en proefijkingen te verrichten.

8. Appendix.

8.1 Stochastische onafhankelijkheid van $\bar{Z}-\bar{Y}$ en $\underline{b}$.

Dat $\bar{Z}-\bar{Y}$ en $\underline{b}$ onder de onderstellingen in opmerking e) (§ 2.3) ongecorreleerd zijn, zien we als volgt in:

$$\begin{aligned} & E \{ (\bar{Z}-\bar{Y}) \underline{b} \} = \\ & = E \left\{ \left(\frac{\sum_j p_j \bar{z}_j}{N_P} - \frac{\sum_j m_j \bar{y}_j}{N_S} \right) \times \frac{\sum_j m_j \bar{x}_j \bar{y}_j}{Q} + \frac{\sum_j r_j \sum_j p_j \bar{\pi}_j \bar{z}_j}{Q} \right\} \end{aligned}$$

Daar alle $\underline{y}$'s en $\underline{z}$'s ongecorreleerd zijn gaat dit over in:

$$E \left\{ \frac{\sum_j p_j^2 \bar{\pi}_j^2 \bar{z}_j^2}{N_P} - \frac{\sum_j m_j^2 \bar{x}_j^2 \bar{y}_j^2}{N_S} \right\} \frac{1}{Q} =$$

(vgl. (2.2.5))

$$= \frac{1}{Q} \left(\frac{\sigma_{\pi}^2 \sum_j p_j^2 \bar{\pi}_j^2}{N_P} - \frac{\sigma_{\bar{x}}^2 \sum_j m_j^2 \bar{x}_j^2}{N_S} \right) = 0,$$

daar $\sum_j p_j \bar{\pi}_j = 0$ en $\sum_j m_j \bar{x}_j = 0$, krachtens de definitiën van $\bar{\pi}$, $\bar{x}$, $\bar{z}$ en $\bar{y}$ ((2.2.3) en (2.2.4)).

Was de onderstelling van homoscedasticiteit niet vervuld, dan vinden we:

$$\begin{aligned} & E \left\{ \frac{\sum_j p_j^2 \bar{\pi}_j^2 \bar{z}_j^2}{N_P} - \frac{\sum_j m_j^2 \bar{x}_j^2 \bar{y}_j^2}{N_S} \right\} \frac{1}{Q} = \\ (8.1.1) \quad & = \frac{1}{Q} \left(\frac{\sum_j p_j^2 \bar{\pi}_j^2 \sigma_{\pi_j}^2}{N_P} - \frac{\sum_j m_j^2 \bar{x}_j^2 \sigma_{\bar{x}_j}^2}{N_S} \right). \end{aligned}$$

Door hierin schattingen voor $\sigma_{\pi_j}^2$ en $\sigma_{\bar{x}_j}^2$ in te substitueren vinden we een schatting van $E \{ (\bar{Z}-\bar{Y}) \underline{b} \}$.

8.2 Onzuiverheid van $\underline{M}$ als schatting van μ

Om te laten zien dat $\underline{M}$ onder geen der onderstellingen (2.2.1) of (2.3.3)+(2.3.5) een zuivere schatting van μ is, maken we gebruik van het volgende. Bezitten $\underline{x}$ en $\underline{y}$ een simultane verdeling en geven we de voorwaardelijke verwachting van $\underline{x}$, onder de voorwaarde dat $\underline{y}=\underline{y}$ is, aan met $\bar{x}(\underline{x} | \underline{y})$, dan vinden we de onvoorwaardelijke verwachting van $\underline{x}$ door van $\bar{x}(\underline{x} | \underline{y})$ de verwachting over $\underline{y}$ te bepalen, wanneer we $\underline{y}$ weer als stochastische grootheid opvatten. In formule:

$$\bar{x} = \bar{y} \int \bar{x}(\underline{x} | \underline{y}) \{ \underline{y} \} .$$

Op deze manier kunnen we $\bar{M}$ bepalen. We weten:

$$\bar{a} = \alpha ,$$

$$\bar{c} = \gamma ,$$

$$\bar{b} = \beta .$$

Wegens opmerking e) is nu:

$$\bar{a}(\underline{a} | \underline{b}) = \bar{y} \int \bar{y} - b \bar{x} | \underline{b} \{ \underline{y} \} = \bar{y} - b \bar{x} ,$$

$$\bar{c}(\underline{c} | \underline{b}) = \bar{y} \int \bar{y} - b \bar{\pi} | \underline{b} \{ \underline{y} \} = \bar{y} - b \bar{\pi} .$$

Dus:

$$\bar{a}(\underline{c-a} | \underline{b}) = \frac{\bar{y} - \bar{x} - b(\bar{x} - \bar{\pi})}{b} = \frac{\bar{y} - \bar{x}}{b} - (\bar{x} - \bar{\pi}) ,$$

$$\bar{a}(\underline{c-a} | \underline{b}) = \bar{y} \int \frac{\bar{y} - \bar{x}}{b} - (\bar{x} - \bar{\pi}) \{ \underline{y} \} = (\bar{y} - \bar{x}) \bar{a}(\underline{1} | \underline{b}) - (\bar{x} - \bar{\pi})$$

Wanneer we aannemen dat b essentieel positief is, geldt volgens de ongelijkheid van Jensen (D.V.DANTZIG [3] pg.94):

$$\bar{a}(\underline{1} | \underline{b}) - \frac{1}{\bar{b}} > 0 ,$$

waaruit volgt

$$\bar{a}(\underline{c-a} | \underline{b}) > \frac{\bar{y} - \bar{x}}{\bar{b}} - \frac{\beta(\bar{x} - \bar{\pi})}{\bar{b}} = \frac{\gamma - \alpha}{\beta} .$$

Delogarithme van de sterkteverhouding wordt dus systematisch overschat.

8.3 Voorwaarden voor de methode van Fieller.

Bij de bepaling van een betrouwbaarheidsinterval volgens FIELLER [2] voor het quotient van twee grootheden

($\mu = \frac{\gamma - \alpha}{\beta}$), waarbij voor de teller en voor de noemer, stochastische schattingen bekend zijn (hier $\underline{c-a}$ en $\underline{b}$),

wordt er gebruik van gemaakt dat de grootheid $\frac{\underline{Z}}{\underline{S}_Z}$,
waarin $\underline{Z} = \underline{c} - \underline{a} - \mu \underline{b}$, en $\underline{S}_Z$ een schatting is voor de spreiding van $\underline{Z}$, verdeeld is volgens een Studentverdeling. Dit is vervuld, wanneer:

1) $\underline{Z}$ verdeeld is volgens een normale verdeling met gemiddelde 0, $f \underline{S}_Z^2$

2) $\frac{f \underline{S}_Z^2}{\underline{Z}}$ een χ^2 -verdeling met f vrijheidsgraden bezit (over f hieronder),

3) $\underline{Z}$ en $\underline{S}_Z$ onderling onafhankelijk verdeeld zijn.

Onder destrengere onderstellingen (2.2.1) is aan deze eisen voldaan:

1) Daar $\underline{Y}_j$ en $\underline{Z}_j$ normaal verdeeld zijn is dit ook het geval met $\underline{c}$, $\underline{a}$ en $\underline{b}$ (zie (2.3.1)) en dus met $\underline{Z} = \underline{c} - \underline{a} - \mu \underline{b}$. Verder geldt dat:

$$\mathcal{L}(\underline{c} - \underline{a}) = \gamma - \alpha$$

$$\mathcal{L}\underline{b} = \mu \beta,$$

dus:

$$\mathcal{L}\underline{Z} = \mathcal{L}(\underline{c} - \underline{a} - \mu \underline{b}) = \gamma - \alpha - \mu \beta = 0.$$

2) Daar $\underline{c}$, $\underline{a}$ en $\underline{b}$ elk homogene lineaire combinaties zijn respectievelijk van $\underline{Z}_j$, van $\underline{Y}_j$ en van $\underline{Z}_j$ en $\underline{Y}_j$ samen, is $\underline{Z}$ te schrijven als:

$$\underline{Z} = \sum_j c_j \underline{Y}_j + \sum_j d_j \underline{Z}_j,$$

waarin c_j en d_j constanten zijn. Het spreidingskwadraat van $\underline{Z}$ wordt:

$$\sigma_{\underline{Z}}^2 = \sum_j \frac{c_j^2}{m_j} \sigma_{\underline{Y}_j}^2 + \sum_j \frac{d_j^2}{p_j} \sigma_{\underline{Z}_j}^2.$$

Een schatting hiervoor verkrijgen we door schattingen voor $\sigma_{\underline{Y}_j}$ en $\sigma_{\underline{Z}_j}$ en te vullen. Volgens de strenge onderstellingen zijn alle spreidingen gelijk:

$$\sigma_{\underline{Y}_j}^2 = \sigma_{\underline{Z}_j}^2 = \sigma_{\underline{u}}^2.$$

Schatten we $\sigma_{\underline{u}}^2$ volgens (2.4.3), dan wordt:

$$\underline{S}_Z^2 = \underline{S}_u^2 \left\{ \sum_j \frac{c_j^2}{m_j} + \sum_j \frac{d_j^2}{p_j} \right\},$$

en:

$$\frac{\underline{S}_Z^2}{\sigma_{\underline{Z}}^2} = \frac{\underline{S}_u^2}{\sigma_{\underline{u}}^2}$$

en dus is volgens (2.4.5) $\frac{f \cdot \underline{S}_Z^2}{\sigma_Z^2}$ met:

$f = N_S - h + \sum^{(p)} (N_P - k)$, verdeeld volgens een χ^2 -verdeling met f vrijheidsgraden.

Is aan de homoscedasticiteit niet voldaan, dus moeten we de spreiding van iedere groep apart schatten, dan vinden we:

$$\underline{S}_Z^2 = \sum_j \frac{c_j^2}{m_j} \underline{S}_{u_j}^2 + \sum_j \frac{d_j^2}{p_j} \underline{S}_{v_j}^2.$$

We weten dat voor iedere waarde van j de grootheden:

$$\frac{(m_j-1)\underline{S}_{u_j}^2}{\sigma_{u_j}^2} \text{ en } \frac{(p_j-1)\underline{S}_{v_j}^2}{\sigma_{v_j}^2} \text{ verdeeld zijn volgens}$$

χ^2 -verdelingen met (m_j-1) resp. (p_j-1) vrijheidsgraden.

Nu is:

$$\begin{aligned} \frac{f \underline{S}_Z^2}{\sigma_Z^2} &= \sum_j \frac{c_j^2}{m_j} \cdot \frac{f \underline{S}_{u_j}^2}{\sigma_Z^2} + \sum_j \frac{d_j^2}{p_j} \cdot \frac{f \underline{S}_{v_j}^2}{\sigma_Z^2} = \\ &= \sum_j \frac{c_j^2 f \sigma_{u_j}^2}{m_j(m_j-1) \sigma_Z^2} \cdot \frac{(m_j-1)\underline{S}_{u_j}^2}{\sigma_{u_j}^2} + \frac{d_j^2 f \sigma_{v_j}^2}{p_j(p_j-1) \sigma_Z^2} \cdot \frac{(p_j-1)\underline{S}_{v_j}^2}{\sigma_{v_j}^2} \end{aligned}$$

en dit heeft in het algemeen niet weer een χ^2 -verdeling, daar het een lineaire combinatie van grootheden, verdeeld volgens χ^2 -verdelingen, is met in het algemeen ongelijke coëfficiënten. Over verdelingen van dergelijke lineaire combinaties is zeer weinig bekend.

3) Aan de eis dat $\underline{Z}$ en $\underline{S}_Z$ onafhankelijk van elkaar verdeeld zijn is voldaan., Indien de waarnemingen binnen iedere groep een normale verdeling bezitten. Dit volgt direct uit de stochastische onafhankelijkheid van het gemiddelde en de spreiding van een steekproef uit een normale verdeling (zie b.v. A.M.Mood [7] p.201).

8.4 Betrouwbaarheidsintervallen volgens Fieller.

Daar $\frac{\underline{Z}}{\underline{S}_Z}$ een onder de voorwaarde (2.2.1) Studentver-

deling bezit met $f = N_S - h + \sum^{(p)} (N_P - k)$ vrijheidsgraden (zie § 8.3) is het mogelijk een betrouwbaarheidsinterval voor μ af te leiden. Is t_ϵ die waarde uit de Student verdeling met f vrijheidsgraden, waarvoor geldt:

$$P[|t_1| > t_\varepsilon] < \varepsilon,$$

dan is

$$(8.4.1) \quad |z| \leq t_\varepsilon \cdot \underline{s}_z,$$

behoudens een onbetrouwbaarheid ε . Nu is:

$$\underline{z} = \underline{c} - \underline{a} - \mu \underline{b}$$

en dus:

$$\sigma_{\underline{z}}^2 = \sigma^2 \{ \underline{c} - \underline{a} - \mu \underline{b} \}.$$

Volgens (2.3.1) gaat dit over in:

$$\begin{aligned} \sigma_{\underline{z}}^2 &= \sigma^2 \{ \underline{Z} - \underline{b} \bar{\pi} - \underline{Y} + \underline{b} \bar{\xi} - \mu \underline{b} \} = \\ &= \sigma^2 \{ \underline{Z} - \underline{Y} + \underline{b} (\bar{\xi} - \bar{\pi} - \mu) \}. \end{aligned}$$

In verband met § 8.1 wordt dit:

$$\begin{aligned} \sigma_{\underline{z}}^2 &= \sigma^2 \{ \underline{Z} - \underline{Y} \} + (\bar{\xi} - \bar{\pi} - \mu)^2 \sigma_{\underline{b}}^2 = \\ &= \sigma_{\underline{u}}^2 \left(\frac{1}{N_P} + \frac{1}{N_S} \right) + (\bar{\xi} - \bar{\pi} - \mu)^2 \sigma_{\underline{b}}^2. \end{aligned}$$

Een schatting voor $\sigma_{\underline{z}}^2$ is:

$$\underline{s}_z^2 = \underline{s}_u^2 \left(\frac{1}{N_P} + \frac{1}{N_S} \right) + (\bar{\xi} - \bar{\pi} - \mu)^2 \underline{s}_b^2.$$

Substitueren we deze uitdrukkingen voor $\underline{z}$ en $\underline{s}_z$ in (8.4.1) dan vinden we dat de werkelijke waarde μ , behoudens een onbetrouwbaarheid ε zal voldoen aan:

$$(8.4.2) \quad |c - a - \mu b| \leq t_\varepsilon \cdot \sqrt{\underline{s}_u^2 \left(\frac{1}{N_S} + \frac{1}{N_P} \right) + (\bar{\xi} - \bar{\pi} - \mu)^2 \underline{s}_b^2}.$$

Hierin is μ onbekend. Alle waarden, die, voor μ ingevuld, aan deze ongelijkheid voldoen, vormen tezamen een betrouwbaarheidsinterval voor de gezochte waarde van met onbetrouwbaarheid ε . Om de grenzen van dit interval te vinden kunnen we beide leden van (8.4.2) kwadrateren en het rechterlid naar links brengen:

$$(c - a - \mu b)^2 - t_\varepsilon^2 \left[\underline{s}_u^2 \left(\frac{1}{N_S} + \frac{1}{N_P} \right) + (\bar{\xi} - \bar{\pi} - \mu)^2 \underline{s}_b^2 \right] \leq 0$$

of

$$\begin{aligned} (8.4.3) \quad &\mu^2 [b^2 - t_\varepsilon^2 \underline{s}_b^2] - 2\mu [b(c - a) - t_\varepsilon^2 (\bar{\xi} - \bar{\pi}) \underline{s}_b^2] + \\ &+ (c - a)^2 - t_\varepsilon^2 \left[\underline{s}_u^2 \left(\frac{1}{N_S} + \frac{1}{N_P} \right) + (\bar{\xi} - \bar{\pi})^2 \underline{s}_b^2 \right] \leq 0. \end{aligned}$$

De grenzen voor het interval voor μ zijn de wortels van deze vergelijking met het $\pm$ -teken.

We vinden:

$$\frac{b(c-a) - t_{\varepsilon}^2 (\bar{f} - \bar{\pi}) s_b^2}{b^2 - t_{\varepsilon}^2 s_b^2} + \frac{\sqrt{[b(c-a) - t_{\varepsilon}^2 (\bar{f} - \bar{\pi}) s_b^2]^2 - [b^2 - t_{\varepsilon}^2 s_b^2] [(c-a)^2 - t_{\varepsilon}^2 s_u^2 (\frac{1}{N_S} + \frac{1}{N_P}) - t_{\varepsilon}^2 (\bar{f} - \bar{\pi})^2 s_b^2]}}{b^2 - t_{\varepsilon}^2 s_b^2}$$

Volgens (2.3.1) is

$$c-a = \bar{Z} - \bar{Y} + b(\bar{f} - \bar{\pi}).$$

Werken we bovenstaande grenzen met behulp hiervan uit, dan ontstaat:

$$(\bar{f} - \bar{\pi}) + \frac{b(\bar{Z} - \bar{Y})}{b^2 - t_{\varepsilon}^2 s_b^2} \pm \frac{\sqrt{t_{\varepsilon}^2 s_u^2 (\frac{1}{N_S} + \frac{1}{N_P}) + \frac{t_{\varepsilon}^2 (\bar{Z} - \bar{Y})^2 s_b^2}{b^2 - t_{\varepsilon}^2 s_b^2}}}{\sqrt{b^2 - t_{\varepsilon}^2 s_b^2}}.$$

Substitueren we hierin:

$$c = \frac{b^2}{b^2 + t_{\varepsilon}^2 s_b^2},$$

dan ontstaat:

$$(8.4.4) \quad (\bar{f} - \bar{\pi}) + \frac{(\bar{Z} - \bar{Y})c}{b} \pm t_{\varepsilon} \sqrt{c} \sqrt{\frac{s_u^2}{b^2} (\frac{1}{N_S} + \frac{1}{N_P}) + \frac{c(\bar{Z} - \bar{Y})^2 s_b^2}{b^4}}.$$

Is $b^2 - t_{\varepsilon}^2 s_b^2 > 0$ dan bestaat het interval voor μ uit de tussen de grenzen (8.4.4) gelegen waarden, is echter $b^2 - t_{\varepsilon}^2 s_b^2 < 0$, dan bestaat het interval voor μ , dat aan (8.4.3) voldoet, juist uit buiten deze grenzen gelegen waarden.

8.5 Hulpstelling voor het homoscedastische geval.

Om een benadering af te leiden voor het spreidingskwadraat van het quotient van twee normaal verdeelde grootheden $\underline{x}$ en $\underline{t}$ zullen we onderstellen dat de mathematische verwachting τ van $\underline{t}$ groot is ten opzichte van de spreiding van $\underline{t}$. De grootheden kunnen als volgt geschreven worden:

$$\underline{x} = \rho + \tilde{\underline{x}} \quad \text{en} \quad \underline{t} = \tau + \tilde{\underline{t}}.$$

Het quotient van beide is:

$$\frac{\underline{x}}{\underline{t}} = \frac{\rho + \tilde{\underline{x}}}{\tau + \tilde{\underline{t}}} = \frac{\rho + \tilde{\underline{x}}}{\tau} \left(1 - \frac{\tilde{\underline{t}}}{\tau} + \frac{\tilde{\underline{t}}^2}{\tau^2} - \frac{\tilde{\underline{t}}^3}{\tau^3} + \dots \right).$$

Uit de onderstelling $\sigma_{\underline{t}} \ll \underline{t}$ volgt, dat ook de hogere momenten van de verdeling van $\underline{t}$ klein zullen zijn t.o.v. τ , en verwaarloosd mogen worden. Voor de verwachting van $\frac{\underline{r}}{\underline{t}}$ volgt nu:

$$\mathbb{E} \frac{\underline{r}}{\underline{t}} \approx \frac{\rho}{\tau} - \frac{\mathbb{E}(\underline{\tilde{t}} \cdot \underline{\tilde{t}})}{\tau^2} + \frac{\rho \sigma_{\underline{\tilde{t}}}^2}{\tau^3},$$

daar $\mathbb{E} \underline{\tilde{t}} = \mathbb{E} \underline{\tilde{r}} = 0$. Voor het spreidingskwadraat van $\frac{\underline{r}}{\underline{t}}$ vinden we:

$$\begin{aligned} \sigma^2 \left\{ \frac{\underline{r}}{\underline{t}} \right\} &= \mathbb{E} \left\{ \left(\frac{\underline{r}}{\underline{t}} - \mathbb{E} \left(\frac{\underline{r}}{\underline{t}} \right) \right)^2 \right\} = \mathbb{E} \left(\frac{\underline{r}}{\underline{t}} \right)^2 - \left(\mathbb{E} \frac{\underline{r}}{\underline{t}} \right)^2 = \\ &\approx \mathbb{E} \left\{ \frac{(\rho + \underline{\tilde{t}})^2}{\tau^2} \left(1 - \frac{\underline{\tilde{t}}}{\tau} + \frac{\underline{\tilde{t}}^2}{\tau^2} \right)^2 \right\} - \left(\frac{\rho}{\tau} - \frac{\mathbb{E}(\underline{\tilde{t}} \cdot \underline{\tilde{t}})}{\tau^2} + \frac{\rho \sigma_{\underline{\tilde{t}}}^2}{\tau^3} \right)^2 = \\ &\approx \mathbb{E} \left\{ \frac{\rho^2 + 2\rho \underline{\tilde{t}} + \underline{\tilde{t}}^2}{\tau^2} \left(1 - 2\frac{\underline{\tilde{t}}}{\tau} + 3\frac{\underline{\tilde{t}}^2}{\tau^2} \right) \right\} - \left(\frac{\rho^2}{\tau^2} - 2\frac{\rho \mathbb{E}(\underline{\tilde{t}} \cdot \underline{\tilde{t}})}{\tau^3} + 2\frac{\rho^2 \sigma_{\underline{\tilde{t}}}^2}{\tau^4} \right) \\ &\approx \frac{\rho^2}{\tau^2} + \frac{\sigma_{\underline{\tilde{t}}}^2}{\tau^2} - \frac{4\rho \mathbb{E}(\underline{\tilde{t}} \cdot \underline{\tilde{t}})}{\tau^3} + 3\frac{\rho^2 \sigma_{\underline{\tilde{t}}}^2}{\tau^4} - \frac{\rho^2}{\tau^2} + 2\frac{\rho \mathbb{E}(\underline{\tilde{t}} \cdot \underline{\tilde{t}})}{\tau^3} + \\ (8.5.1) \quad &- \frac{2\rho^2 \sigma_{\underline{\tilde{t}}}^2}{\tau^4} = \frac{\sigma_{\underline{\tilde{t}}}^2}{\tau^2} + \frac{\rho^2 \sigma_{\underline{\tilde{t}}}^2}{\tau^4} - 2\frac{\rho \mathbb{E}(\underline{\tilde{t}} \cdot \underline{\tilde{t}})}{\tau^3}. \end{aligned}$$

Door hierin voor de onbekende parameters schattingen in te vullen ontstaat een schatting van het spreidingskwadraat. De voorwaarde $\sigma_{\underline{t}} \ll \underline{t}$ houdt voor de in § 4 vermelde toepassing vandeze stelling in, dat $\mathbb{E} \underline{\tilde{t}} = \beta$ aanzienlijk groter moet zijn dan $\sigma_{\underline{\tilde{t}}}$, dus dat $\beta = 0$ niet toestaan is.

Een deel der berekeningen van dit rapport is uitgevoerd door de Rekenafdeling van het Mathematisch Centrum.

Litteratuur

- [1] J.O.Irwin Statistical Method Applied to Biological Assays. Suppl.to the Journal Roy Stat Soc 4 (1937) pp 1-12.
- [2] E.C.Fieller A fundamental formula in the statistics of biological assay and some applications. Quart.J.of Pharmacy and Pharmacology.Vol XVII (1944) pp 117-123.
- [3] Prof.Dr D.van Dantzig College Capita Selecta. Waarschijnlijkheidsrekening § 3.
- [4] E.C.Fieller The biological standardisation of Insulin: Appendix (e). Suppl.to the J.Roy Stat Soc 7 (1940-41) pp 50-53.
- [5] J.O.Irwin On the calculation of the error of biological assays. The Journal of Hygiene, Vol XLIII (1943-44), pp 121 e.v.
- [6] C.W.Emmens Principles of Biological Assay London, Chapman & Hall Ltd, 1948.
- [7] A.M.Mood Introduction to the Theory of Statistics. New York, Toronto, London, McGraw-Hill Book Company, Inc., 1950, First Edition.
- [8] N.K.Jerne and E.C.Wood, The validity and meaning of the results of biological assays. Biometrics 5 (1949) pp 273-299.

Enige definities uit de schattingstheorie.¹⁾

1. Inleiding.

Een veel voorkomend probleem in de statistiek is het schatten van één of meer onbekende parameters (b.v. het gemiddelde of de spreiding) van een overigens bekende waarschijnlijkheidsverdeling W , indien een aantal waarnemingen gegeven zijn. Men schat de onbekende parameter dan door een functie van de waarnemingen. Zijn $x_1, \dots, x_n$ de waarnemingen en zijn deze onderling onafhankelijk dan kan men b.v. het gemiddelde van de waarschijnlijkheidsverdeling W schatten door de grootte

$$(1) \quad \bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i,$$

dus door het gemiddelde van de waarnemingen. Een dergelijke schatting geeft in het algemeen niet de werkelijke waarde van de gezochte parameter, maar een daarbij in de buurt liggende waarde. Men heeft steeds de keus tussen een aantal verschillende functies van $x_1, \dots, x_n$ als schattingen voor dezelfde onbekende parameter. Zo kan men b.v. als schatting voor het gemiddelde van de waarschijnlijkheidsverdeling ook nemen de functie

$$(2) \quad \bar{x}^* = \frac{2}{n(n+1)} (x_1 + 2x_2 + \dots + nx_n) = \frac{2}{n(n+1)} \sum_{i=1}^n i \cdot x_i,$$

Men kiest nu uit deze mogelijkheden naar aanleiding van de eigenschappen, die de verschillende schattingen bezitten. Daartoe denkt men zich het experiment vele malen herhaald, d.w.z. men onderstelt, dat vele reeksen van onderling onafhankelijke waarnemingen $x_1, \dots, x_n$ zijn verricht en berekent voor ieder van deze reeksen de verschillende schattingen. Van ieder dezer schattingen verkrijgt men zodoende een groot aantal waarden, waaruit men eigenschappen omtrent de nauwkeurigheid der verschillende schattingen kan afleiden. In mathematische formulering komt dit daarop neer, dat men ieder

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

der waarnemingen $x_1, \dots, x_n$ beschouwt als een stochastische grootheid¹⁾, die verdeeld is volgens de waarschijnlijkheidsverdeling W , waarvan de onbekende parameter geschat moet worden. Deze stochastische grootheden $x_1, \dots, x_n$ zijn, indien de waarnemingen onafhankelijk zijn, bovendien onafhankelijk verdeeld, alle volgens deze zelfde waarschijnlijkheidsverdeling W . Wij beperken ons tot de schatting van één onbekende parameter van W , die wij met θ angeven en laten het geval van de gezamenlijke schatting van verschillende parameters buiten beschouwing.

De schatting van θ uit n waarnemingen $x_1, \dots, x_n$ geven we aan met

$$(3) \quad \hat{u}_n = u_n(x_1, \dots, x_n)$$

daar u_n een functie van $x_1, \dots, x_n$ is en dus zelf ook een waarschijnlijkheidsverdeling bezit.

De definities van enkele eigenschappen uit de schattingstheorie geven wij eerst in populaire vorm en vervolgens (in § 3) in streng mathematische vorm.

2. Definities (populair).

A. Een schatting²⁾ $\hat{u}_n$ van een parameter θ heet zuiver²⁾, indien het gemiddelde van de waarschijnlijkheidsverdeling van $\hat{u}_n$ gelijk aan θ is.

Voorbeelden: De door (1) en (2) gegeven schattingen van het gemiddelde van de waarschijnlijkheidsverdeling W zijn beide zuivere schattingen. De schatting

$$(4) \quad \hat{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

waarin $\bar{x}$ het door (1) gegeven gemiddelde der waarnemingen voorstelt, is een zuivere schatting van de variantie (het spreidingskwadraat) van W .

B. Een schatting $\hat{u}_n$ van θ heet asymptotisch zuiver³⁾, indien het gemiddelde van de waarschijnlijkheidsverdeling van $\hat{u}_n$ voor $n \rightarrow \infty$ tot θ nadert.

- 1) Een stochastische grootheid is een grootheid, die een waarschijnlijkheidsverdeling bezit; wij geven een dergelijke grootheid aan door onderstreepte letter, terwijl niet onderstreepte letters gewone algebraïsche grootheden of door stochastische grootheden aangenomen waarden voorstellen.
- 2) Engels: "estimate" resp. "unbiased". De vertaling der termen is van Prof. Dr D. van Dantzig; zie de literatuurverwijzing aan het einde van dit memorandum.
- 3) Engels: "asymptotically unbiased".

Voorbeeld:

$$(5) \quad \underline{s}^2 = \frac{1}{n} \sum_{i=1}^n (\alpha_i - \underline{\alpha})^2$$

is een voor eindige n onzuivere, maar asymptotisch zuivere, schatting van de variantie van W .

C. Een schatting u_n van θ heet bruikbaar¹⁾, indien men, door het verrichten van een voldoende groot aantal waarnemingen kan zorgen dat, op een willekeurig kleine gegeven waarschijnlijkheid na, het verschil tussen de gevonden waarde van u_n en de gezochte waarde θ kleiner is dan een willekeurig klein gegeven getal. Anders uitgedrukt: indien men, op een willekeurig kleine waarschijnlijkheid na, iedere gewenste nauwkeurigheid kan bereiken door n voldoende groot te kiezen.

Voorbeeld: Vrijwel alle gebruikelijke schattingen zijn bruikbaar. De door (1) gegeven schatting van het gemiddelde van W gaat b.v. in een niet bruikbare schatting over, indien wij de factor $\frac{1}{n}$ vervangen door $\frac{1}{2n}$. Bij een toenemend aantal waarnemingen zal dan de schatting vrijwel zeker vlak bij de helft van het gemiddelde terecht komen.

D. Een schatting u_n van θ heet de doeltreffendste²⁾ schatting van θ indien zij zuiver is en bovendien van alle zuivere schattingen de kleinste spreiding bezit.

Voorbeeld: De door (1) gegeven schatting van het gemiddelde van W is, als W een normale waarschijnlijkheidsverdeling is, de doeltreffendste schatting van het gemiddelde.

E. Een schatting u_n van θ heet een asymptotisch meest doeltreffende³⁾, indien zij bruikbaar is en onder alle bruikbare schattingen voor $n \rightarrow \infty$ de kleinste spreiding bezit.

Voorbeeld: de schatting (5) is evenals (4) een asymptotisch meest doeltreffende schatting van de variantie van W , indien W een normale waarschijnlijkheidsverdeling is. Daarentegen voldoet (4) wel, maar (5) niet aan de definitie D. .

1) Engels: "consistent".

2) Engels: "most efficient estimate". Deze naam wordt ook vaak gegeven aan de onder E genoemde schattingen.

3) Engels: "asymptotically most efficient estimate". Vgl. ook voetnoot 2) van deze pagina.

F. Een schatting u_n van θ heet de aannemelijkste schatting¹⁾ van θ , indien de waarschijnlijkheid, of bij continue $\mathbb{W}$ de waarschijnlijkheidsdichtheid van het gevonden resultaat (d.i. van de gevonden waarnemingen $x_1, \dots, x_n$) maximaal wordt bij substitutie van u_n voor θ .

Voorbeeld: Indien onder n worpen met een munt a maal kruis voorkomt, is a/n de aannemelijkste schatting van de kans op kruis.

Opmerking: In de volgende paragraaf worden deze definities gepreciseerd en nog een enkele verdere definitie toegevoegd. Verder worden enkele eigenschappen vermeld.

3. Precisering van de definities.

Zij $F(x/\theta)$ de verdelingsfunctie en $f(x/\theta)$, indien deze bestaat, de waarschijnlijkheidsdichtheid van de waarschijnlijkheidsverdeling $\mathbb{W}$ met de onbekende parameter θ .

Zij verder

$$(6) \quad \ell_{\theta} \varphi(x) = \int_0^1 \varphi(x) dF(x|\theta) = \int_{-\infty}^{\infty} \varphi(x) f(x|\theta) dx$$

Dan is:

A. u_n een zuivere schatting van θ , indien geldt:

$$(7) \quad \ell_{\theta} u_n = \theta$$

B. u_n een asymptotisch zuivere schatting van θ , indien geldt:

$$(8) \quad \lim_{n \rightarrow \infty} \ell_{\theta} u_n = \theta$$

C. u_n een bruikbare schatting van θ , indien voor iedere $\varepsilon > 0$ geldt:

$$(9) \quad \lim_{n \rightarrow \infty} P[|u_n - \theta| > \varepsilon] = 0$$

D. u_n de doeltreffendste schatting van θ , indien u_n aan (7) voldoet en

$$(10) \quad \sigma_{\theta}(u_n) \leq \sigma_{\theta}(v_n) \quad 2)$$

is voor iedere v_n , waarvoor (7) geldt.

E. u_n een asymptotisch doeltreffendste schatting van θ , indien u_n aan (9) voldoet, terwijl voor iedere v_n , die aan (9) voldoet geldt:

1) Engels: "maximum likelihood estimate". Dit begrip is afkomstig uit de theorie der aannemelijkste schattingen van R.A.Fisher.

$$2) \quad \sigma_{\theta} \{ \psi(x) \} = \left[\ell_{\theta} \{ \psi(x) \} - \ell_{\theta} \psi(x) \right]^2 \Big|^{1/2}$$

$$(11) \quad \lim_{n \rightarrow \infty} \frac{\sigma_{\theta}(\underline{u}_n)}{\sigma_{\theta}(\underline{v}_n)} \leq 1.$$

D', E'. Men noemt $\underline{u}_n$ doeltreffender resp. asymptotisch doeltreffender dan $\underline{v}_n$ als schatting van θ , indien (10) resp. (11) geldt met het kleiner-teken.

Men noemt de grootheid:

$$(12) \quad \lim_{n \rightarrow \infty} \frac{\sigma_{\theta}^2(\underline{u}_n)}{\sigma_{\theta}^2(\underline{v}_n)}$$

de asymptotische doeltreffendheid¹⁾ van $\underline{v}_n$, indien $\underline{u}_n$ een asymptotisch meest doeltreffende schatting van θ is.

F. De aannemelijkheid²⁾ van het stelsel waarnemingen $x_1, \dots, x_n$ wordt gedefinieerd door

$$(13) \quad \ell = \sum_{i=1}^n \ln f(x_i | \theta)$$

of, indien W een discontinue waarschijnlijkheidsverdeling is, door:

$$(14) \quad \ell^* = \sum_{i=1}^n \ln P[x_i = x_i | \theta]$$

ℓ en ℓ^* worden in de theorie der aannemelijkste schattingen beschouwd als functies van θ en de aannemelijkste schatting $\underline{u}_n$ van θ , bij gegeven $x_1, \dots, x_n$, wordt gedefinieerd als die waarde van θ , waarvoor ℓ resp. ℓ^* een absoluut maximum bezitten.

Enkele eigenschappen.

Een bruikbare schatting is altijd asymptotisch zuiver.

Een zuivere schatting is bruikbaar, indien geldt:

$$(15) \quad \lim_{n \rightarrow \infty} \sigma_{\theta}(\underline{u}_n) = 0 \quad 3)$$

Onder vrij algemene voorwaarden geldt, dat de aannemelijkste schatting bruikbaar en asymptotisch meest doeltreffend is. Eén van de voorwaarden, waaronder deze stelling bewezen is, is, dat W een continue waarschijnlijkheidsverdeling is.

Litteratuur: D. van Dantzig, Kadercursus Mathematische Statistiek, Mathematisch Centrum, Amsterdam 1946-1950, Hoofdstuk 4.

1) Engels: "efficiency"; ook in het Nederlands laat men de toevoeging "asymptotisch" meestal weg.

2) Engels: "likelihood function" of "likelihood".

3) Deze voorwaarde is voldoende, maar niet noodzakelijk.

MATHEMATISCH CENTRUM
 2de Boerhaavestraat 49,
A m s t e r d a m -0.
 Statistische Afdeling.
 S 53 (M 25).

Toets van Bartlett¹⁾

Stel we hebben N waarnemingen:

$$\begin{array}{cccc} x_{11} & x_{21} & \dots & x_{n_1 1} \\ x_{21} & x_{22} & \dots & x_{n_2 2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ x_{1p} & x_{2p} & & x_{n_p p} \end{array}$$

die verdeeld zijn in p subgroepen, terwijl het aantal waarnemingen in de j^e subgroep n_j is.

Veronderstel verder dat ieder der subgroepen een steekproef is uit een normale verdeling. De te toetsen hypothese H_0 is dan dat de spreidingen der subgroepen gelijk zijn.

Is $\bar{x}_j$ het gemiddelde van de waarnemingen in de j^e subgroep, $\nu_j = n_j - 1$, $N' = \sum_{j=1}^p \nu_j = N - p$

$$s_j^2 = \frac{1}{\nu_j} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

$$s^2 = \sum_{j=1}^p \frac{\nu_j}{N'} s_j^2$$

dan definiëren wij:

$$L_1^* = \prod_{j=1}^p \left(\frac{s_j^2}{s^2} \right)^{\frac{\nu_j}{N'}}$$

L_1^* is dus de verhouding van het geometrisch en het rekenkundig gemiddelde van de varianties der subgroepen en is dus ≤ 1 .

Als toetsingsgrootheid wordt nu genomen:

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

$$M = -N' \lg L_1^* = N' \lg \sum_{j=1}^p \frac{\nu_j}{N'} s_j^2 - \sum_{j=1}^p \nu_j \lg s_j^2$$

en M is dus ≥ 0 .

Als H_0 juist is zal M in het algemeen klein zijn en de kritieke zone bestaat dus uit grote waarden van M .

Deze toets van Bartlett is analoog aan een enkele jaren eerder door J. Neymann en E.S. Pearson afgeleide toets, waarbij in de toetsingsgrootte n_j en N in plaats van ν_j en N' gebruikt worden. De met L_1^* overeenkomende grootte noemden zij L_1 . De toets wordt, in deze oudere vorm, L_1 -toets genoemd.

Litteratuur: H.O. Hartley: Testing the homogeneity of a set of estimated variances. Biometrika 31 (1940) p. 249.

Tabellen: H.O. Hartley en E.S. Pearson: Tables for testing the homogeneity of a set of estimated variances, Biometrika 33 (1946) p. 296.

Algemene gang van zaken bij het toetsen van een ¹⁾
hypothese.

De toetsing van een hypothese H_0 berust steeds op een aantal waarnemingen $x_1, x_2, \dots, x_n$ van één of meer stochastische grootheden ²⁾, of op enige groepen van waarnemingen (bv. twee steekproeven).

Bij een toets behoort een toetsingsgrootheid u (soms meer dan één), die een functie is van bovengenoemde stochastische grootheden en die, voor de waargenomen waarden $x_1, x_2, \dots, x_n$ 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 H_0 juist is, de waarschijnlijkheidsverdeling van deze grootheid kan berekenen.

Vervolgens kiest men een verzameling Z van mogelijke uitkomsten van u , en wel op zodanige wijze, dat de kans, dat u een in Z gelegen waarde aanneemt, onder de hypothese H_0 , gelijk is aan een gegeven getal α , zodat Z dus van α afhankelijk is ³⁾. Z heet de kritieke zône van de toets, α de onbetrouwbaarheidsdrempel (Engels: level of significance). Voor α neemt men veelal de waarde 0,05 of 0,01.

Men verwerpt nu H_0 op grond van de waarnemingen $x_1, x_2, \dots, x_n$, indien de bij deze waarnemingen behorende waarde van u in Z ligt. Dit wordt vaak uitgedrukt door te zeggen, dat het resultaat van het experiment "significant" is. De waarde van α moet dan echter worden vermeld. De kans, dat dit zal gebeuren, is, indien H_0 juist is, gelijk aan α . Derhalve is α 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 $\leq \alpha$ is.

De toetsingstheorie biedt in het algemeen geen mogelijkheid om tot aanvaarding 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 nevenvoorwaarden dienen steeds uitdrukkelijk te worden vermeld en, zo mogelijk, zelf te worden getoetst.

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

Voor het onderscheid tussen één- en tweezijdige toetsing en de keuze tussen deze twee mogelijkheden vergelijke men bv. de tweede hieronder gegeven literatuurplaats. 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.

Litteratuur:

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

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

z-toets van Fisher¹⁾

Stel we hebben twee onafhankelijke steekproeven

$$\begin{array}{ll} x_1 \dots x_{n_1} & \text{met gemiddelde } \bar{x} \\ y_1 \dots y_{n_2} & \text{met gemiddelde } \bar{y}, \end{array}$$

beide uit een normale verdeling.

De te toetsen hypothese is dan, dat de spreidingen van deze normale verdelingen gelijk zijn.

Is

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2$$

en

$$s_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2$$

dan is de toetsingsgrootheid:

$$z = \frac{1}{2} \lg \frac{s_1^2}{s_2^2}$$

waarbij van s_1^2 en s_2^2 de grootste in de teller van de breuk genomen wordt.

Als H_0 juist is zal z in het algemeen klein zijn en de kritieke z -one bestaat dus uit grote waarden van z . Men kan als toetsingsgrootheid ook nemen

$$F = e^{2z} = \frac{s_1^2}{s_2^2}.$$

De kritieke z -one bestaat dan uit grote waarden van e^{2z} . De toets wordt dan gewoonlijk de F-toets van Snedecor genoemd.

In de verdelingen van z en van e^{2z} komen twee parameters voor: $\nu_1 = n_1 - 1$ en $\nu_2 = n_2 - 1$, die het aantal vrijheidsgraden in de teller resp. in de noemer genoemd worden.

Litteratuur: M.G.Kendall: The advanced theory of statistics: deel II pag.115.

P.G.Hoel: Introduction to mathematical statistics; pag. 152-154.

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

Vervolg litteratuuropg.:

Tabellen: M.G.Kendall: The advanced theory of statistics; deel I pag. 442-443.

P.G.Hoel: Introduction to mathematical statistics; pag. 250-253.

χ^2_1 : 1 tot 500; χ^2_2 : 1 tot 1000; onbetrouwbaarheidsdrempels 0,05 en 0,01.

R.A.Fisher en F.Yates: Statistical tables for biological, agricultural and medical research; pag. 34-43.

χ^2_1 : 1 tot 24; χ^2_2 : 1 tot 120; onbetrouwbaarheidsdrempels:

0,20; 0,10; 0,05; 0,01; 0,001.

De toets van Student voor 2 steekproeven¹⁾.

Deze toets van Student wordt gebruikt voor het toetsen van de hypothese $\mathcal{H}_0$, dat 2 steekproeven $x_1, x_2, \dots, x_n$ en $y_1, y_2, \dots, y_m$ beschouwd kunnen worden als steekproeven uit eenzelfde normale verdeling.

De grootte $\underline{t}$, welke hierbij als toetsingsgrootte gebruikt wordt en waarvan Student de verdelingsfunctie berekend heeft, is de volgende:

$$\underline{t} = \frac{\bar{m}_x - \bar{m}_y}{S} \sqrt{\frac{nm}{n+m}}$$

Hierin zijn $\bar{m}_x$ en $\bar{m}_y$ de gemiddelden van de 2 steekproeven terwijl S een schatting is van de spreiding σ van de normale verdeling, waaruit volgens $\mathcal{H}_0$ de steekproeven getrokken zijn. De bij de gevonden steekproeven behorende waarden van $\bar{m}_x$, $\bar{m}_y$ en S zijn op volgende wijze uit de waarnemingen te berekenen:

$$\bar{m}_x = \frac{1}{n} (x_1 + x_2 + \dots + x_n)$$

$$\bar{m}_y = \frac{1}{m} (y_1 + y_2 + \dots + y_m)$$

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{m}_x)^2 + \sum_{j=1}^m (y_j - \bar{m}_y)^2}{n+m-2}$$

Bovengenoemde grootte $\underline{t}$ bezit als verdeling de verdeling van "Student" met $\nu = n+m-2$ vrijheidsgraden.

Deze verdeling is getabelleerd voor $\nu = 1$ t/m 20 [1] voor $\nu = 20$ t/m 100 zijn waarden van $\underline{t}$ berekend, welke behoren bij de onbetrouwbaarheidsdrempels $\alpha = 0,01$ en $\alpha = 0,05$. [2]

Is $\mathcal{H}_0$ juist, dan zullen dicht bij nul gelegen waarden vaker voorkomen, dan ver van nul gelegen waarden. De kritieke zone kiezen we dus zodanig, dat ze gevormd worden door grote en kleine waarden van $\underline{t}$. De kritieke zone bestaat dus, bij gegeven waarde van α , uit twee stukken: $\underline{t} < -t_0$ en $\underline{t} > t_0$, waarin t_0 de bij α en ν behorende kritieke waarde is, die in [1] en [2] getabelleerd is. Voor ééNZijdige toetsing is de kritieke zone van de vorm $\underline{t} \leq -t_1$ of $\underline{t} \geq t_1$.

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

Literatuur:

- [1]. M.G.Kendall, The Advanced Theory of Statistics,
London 1946, Vol. II p.109, tabellen in
deel I p. 440-41;
Opmerking: het bij de tabellen vermelde
aantal vrijheidsgraden (aangegeven door
 ν) is gelijk aan $n + m - 2$.
- [2]. Elizabeth M.Baldin, Table of Percentage Points of the
t-distribution, Biometrika 33 (1943),
p. 362. (Deze tabel worden tweezijdige
overschrijdingskansen gegeven).