

MATHEMATISCH CENTRUM

2e BOERHAAVESTRAAT 49

AMSTERDAM

STATISTISCHE AFDELING

Leiding: Prof. Dr D. van Dantzig

Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

Rapport S 113

Insuline-ijking met konijnen

door

Mevr. G.Klerk-Grobbe

en

H.J.Prins

1953.

Inhoud.

1. Inleiding.
2. Algemene beschouwingen en literatuurverwijzingen.
 - 2.1. Een maat voor de reactie.
 - 2.2. Invloeden op de reactie.
 - 2.3. Schema voor een ijking.
 - 2.4. Berekeningsmethoden.
 - 2.4.1. Methode van Marks e.a.
 - 2.4.2. Toepassing van variantie en regressie-analyse.
3. Analyse van het waarnemingsmateriaal.
 - 3.1. Overzicht van het materiaal.
 - 3.2. Normaliteit en homoscedasticiteit.
 - 3.3. Onderzoek in verband met een keuze van de reactiemaat.
 - 3.4. Gevoeligheid der konijnen.
 - 3.5. Het reactie-dosis verband.
 - 3.6. Conclusies.
4. Voorstellen over een gewijzigd ijkschema met berekeningsmethoden.
 - 4.1. Schema voor de routine-ijking.
 - 4.2. Schatting van de helling.
 - 4.3. Berekening van de sterkte van het praeparaat.
 - 4.4. Grafische controle van de helling van de ijklijn.
 - 4.5. Voorstel in verband met de gevoeligheid der konijnen.
 - 4.6. Opmerkingen.
5. Samenvatting.

Literatuurlijst.

Memoranda.

1. Inleiding.

Het doel van dit onderzoek was te zoeken naar een verbetering van de gebruikelijke ijkingmethoden van insuline-praeparaten met konijnen als proefdieren.

Ter oriëntering werden ons de resultaten van een aantal ijkingen en proeven gegeven; in paragraaf 3.1 zal dit materiaal uitvoeriger beschreven worden.

In dit rapport zullen in hoofdstuk 2 de verschillende problemen, welke bij een dergelijke verbeteringspoging naar voren komen, besproken worden. Cijfers tussen vierkante haken[] verwijzen naar de literatuurlijst aan het einde van dit rapport. Hoofdstuk 3 is gewijd aan de berekeningen en toetsen welke op het materiaal zijn toegepast, in verband met de problemen uit hoofdstuk 2. Hierbij zij nog opgemerkt dat een volledig onderzoek pas mogelijk zou zijn indien de resultaten van speciaal voor elk probleem opgezette proeven ter beschikking stonden. De suggesties over een eventueel verbeterd of vereenvoudigd proefschema voor de ijkingen zijn in hoofdstuk 4 samengevat.

2. Algemene beschouwingen en literatuurverwijzingen.

Ter wille van de overzichtelijkheid zijn de problemen in vier groepen ondergebracht

- 1) een maat voor de reactie,
- 2) invloeden op de reactie,
- 3) schema's voor een ijking en
- 4) berekeningsmethoden;

het zal echter blijken dat de problemen uit deze vier groepen nauw met elkaar samenhangen.

2.1. Een maat voor de reactie.

Wordt een konijn met een dosis insuline ingespoten dan zal zijn bloedsuikergehalte een tijdelijke daling vertonen, waarvan de duur en de diepte afhankelijk zijn van de soort en de hoeveelheid van het insuline-praeparaat. Van deze eigenschap van het konijn wordt nu gebruik gemaakt bij de ijking (dus sterktebepaling) van het insuline-praeparaat.

Het eerste probleem, dat nu aan de orde komt, is:

welke waarnemingen moeten er worden verricht en welke, uit de waarnemingen te berekenen, grootheid kan het doeltreffendst als maat voor de reactie van het konijn gebruikt worden. Nu zijn de mogelijke waarnemingen waar we van uit moeten gaan de volgende:

- 1) het suikergehalte van het konijnenbloed vlak voor de insulinspuiting met het insuline-praeparaat; dit gehalte zullen we $\bar{c}$

noemen (in de literatuur meestal I.B.S.);

2) de suikergehalten op verschillende tijdstippen na de insuline-injectie, bv. op 1, 2, 3, 4 en 5 uur of op 1, 2 en 3 uur na inspuiting. Het gemiddelde van deze gehalten (of bij één gehalte-bepaling na inspuiting, dit gehalte zelf) zullen we met e aangeven.

Uitgaande van een dergelijk waarnemingsmateriaal zijn nu verschillende grootheden als reactiemaat mogelijk. We noemen als voornaamste:

1) $\frac{100(b-e)}{b}$, de percentuele bloedsuikerdaling, de in de praktijk het meest gebruikte maat, waarbij e het gemiddelde van 3 à 5 gehalten na inspuiting is.

2) $b-e$, de absolute daling.

3) e , het absolute niveau dat een enkele gehalte-bepaling kan zijn of het gemiddelde van een aantal bepalingen. Een waarneming van b is hierbij overbodig geworden.

Als overwegingen welke bij een keus van de reactiemaat een rol spelen, kunnen we de volgende noemen:

1) physiologische overwegingen, overwegingen in verband met de klinische toepassing van het praeparaat en met wat reeds bekend is over het praeparaat (praeparaten met vertraagde werking!);

2) de tijd waarin de konijnen weer hersteld zijn en beschikbaar voor een volgende ijking;

3) de wens een van b onafhankelijke reactiemaat te vinden, zodat correcties voor b overbodig zijn;

4) de nauwkeurigheid van de ijking die met de verschillende reactiematen bereikt kan worden;

5) de bewerkelijkheid van het bepalen van de reactie en het berekenen van de ijkuitkomst.

De overwegingen onder 1) genoemd vallen natuurlijk buiten dit statistisch onderzoek. Ook punt 2) is geen statistische maar een biologische vraag. Wel is het duidelijk dat een reactiebepaling met kortere herstellingstijd te verkiezen is. K.L. SMITH e.a. ([B], [18], [17]) vermelden dat bij gebruik van het absolute niveau op één tijdstip na inspuiting (reactiemaat 3), met slechts één bloedafname per konijn!) de konijnen vlug hersteld zijn; zo zelfs dat ze op zes achtereenvolgende dagen voor proeven gebruikt kunnen worden ([B] p. 45).

In de literatuur wordt vermeld dat reactiemaat 1) en, in nog sterkere mate, reactiemaat 2) gecorreleerd zijn met b , maar dat verschillende groepen konijnen significant verschillende

waarden voor deze afhankelijkheden geven [5] (zie ook paragraaf 2.4.1). Deze correlatie zou voor het absolute niveau op bv. 1 uur of $1\frac{1}{2}$ uur na intraveneuze inspuiting veel minder zijn ([B] p. 62). (De gebruikelijke behandeling is een subcutane inspuiting.) K.L.SMITH behandelt een ijking met subcutane inspuiting en één bloedafname na $1\frac{1}{2}$ uur ([B] p. 44). Bestaat er een correlatie tussen δ en de reactie dan kan een correctie voor deze correlatie vaak de nauwkeurigheid van de uitkomst vergroten. De bewerking van de waarnemingen wordt hierdoor echter wel ingewikkelder en het zou dus aanbeveling verdienen te werken met een reactiemaat die niet met δ gecorreleerd is, mits natuurlijk de nauwkeurigheid van de uitkomst hieronder niet te veel lijdt.

In verband met het bovenstaande zullen we ons bij de analyse van het waarnemingsmateriaal alleen bezighouden met de reactiematen 1) en 3) en reactiemaat 2) verder buiten beschouwing laten.

Over de nauwkeurigheid in verband met de reactiemaat is in de literatuur niets gevonden. Wel worden een aantal ijkuitkomsten met hun nauwkeurigheid gegeven bij het gebruik van reactiemaat 1), en voor verschillende proefschema's (zie paragraaf 2.3). Zo geven BLISS en MARKS [10] als uitkomst voor de logaritmen van de sterkte-verhouding tussen praeparaat en standaard 0,12 met spreiding 0,03 voor een proefschema met 12 dieren welke twee maal 4 doses (2 praeparaat en 2 standaard) toegediend krijgen op in totaal 8 dagen. Voor hetzelfde schema geven YOUNG en ROMANS [17] als schatting voor de spreiding van de ijkuitkomst (gehalte aan insuline) 12% van de uitkomst (een betrouwbaarheidsinterval met 0,05 als onbetrouwbaarheid zou hier dus, ruw geschat een lengte van 48% van de uitkomst bezitten). FIELLER [12] bepaalt uit een serie van 7 gewone kruisproeven met elk 12 à 14 dieren een betrouwbaarheidsinterval met een lengte van 20% van de uitkomst.

Bij deze proeven werd de helling van de ijklijn uit hetzelfde materiaal als de sterkte van het praeparaat geschat.

We kunnen nog wijzen op een tabelletje van K.L.SMITH [B] waarin hij een schatting geeft van het aantal dieren dat voor de verschillende proefschema's nodig zal zijn om een betrouwbaarheidsinterval van bepaalde lengte te verkrijgen.

Voor reactiemaat 2) en 3) zijn geen gegevens over de nauwkeurigheid bekend.

Wat betreft de bewerkelijkheid van het bepalen van de reactie per konijn verdient het absolute niveau op één tijdstip na

inspuiting de voorkeur. Over het berekenen van de ijkuitkomst zal in paragraaf 2.4 nog een verhandeling volgen.

2.2. Invloeden op de reactie.

Deze invloeden kunnen we als volgt indelen:

- 1) De omstandigheden als ras, plaats, leeftijd, behandeling, voeding enz.
- 2) Tijdstip waarop de proef gedaan wordt: daginvloeden en seizoeninvloeden.
- 3) Individuele eigenschappen van het konijn.

Om van een groep konijnen vergelijkbare waarnemingen te krijgen is het van belang de omstandigheden in 1) genoemd zoveel mogelijk voor alle dieren gelijk te maken.

De dag- en seizoeninvloeden kunnen bij biologische ijk-ingen een grote rol spelen. Verschillende onderzoekers [2] wijzen op een grote daginvloed welke de waarnemingen van een hele groep konijnen kunnen verhogen of verlagen. In verband hiermee is het van belang, om bij vergelijking van reacties op verschillende praeparaten insuline, de proeven met beide praeparaten op dezelfde dagen te doen plaats vinden. In de volgende paragraaf, over een proefschema, wordt hier nog nader op in gegaan. Bij het onderzoek van het materiaal is de invloed van het seizoen nog onderzocht.

Ten slotte wordt de reactie van een konijn nog bepaald door zijn individuele eigenschappen. In hoeverre deze eigenschappen de plaats van een konijn in een groep vastleggen is in paragraaf 3.4 voor het beschikbare materiaal onderzocht. In hoofdstuk 4 wordt hierop nog een suggestie over de vervanging van konijnen bij routine-ijkingen gebaseerd.

2.3. Schema voor een ijking.

Daar het onmogelijk is alle invloeden genoemd in paragraaf 2.2 in rekening te brengen of voor iedere ijking gelijk te maken, is een absolute ijking van een insuline-paeparaat onmogelijk en zal een ijking steeds bestaan uit een vergelijking van de reacties op het onbekende praeparaat met de reacties op een praeparaat van bekende sterkte, waarbij beide praeparaten op dezelfde dagen en zoveel mogelijk onder gelijke omstandigheden worden ingespoten ([1] en [2]). Het algemeen gebruikelijke schema voor een dergelijke ijking is het kruisproefschema [2]. Hierbij worden twee groepen konijnen elk met het onbekende en het bekende praeparaat (in het vervolg aan te geven met praeparaat resp. standaard) behandeld. hiervoor zijn twee dagen nodig daar een konijn

niet op één dag zowel met standaard als praeparaat kan worden ingespoten. We kunnen het proefschema als in Schema I weergeven.

Schema I Kruisproef

konijnen groep	proefdag	
	1 ^e dag	2 ^e dag
G ₁	s	p
G ₂	p	s

In dit schema stelt s een dosis van de standaard en p een dosis van het praeparaat voor. Wanneer nu de som van alle reacties op de standaard-dosis vergeleken wordt met de som van de reacties op de praeparaat-dosis, dan zijn in beide sommen alle dieren en beide dagen vertegenwoordigd, zodat verwacht mag worden dat verschillen tengevolge van daginvloeden en individuele verschillen tussen de konijnen van beide groepen geen overwegende rol meer spelen bij het reactieverval tussen praeparaat en standaard.

Opgemerkt moet worden dat de ijking bij dit schema in wezen op een nulmethode berust. Dat wil zeggen, bij de statistische behandeling van een dergelijke proef is slechts één conclusie mogelijk, in de vorm van een betrouwbaarheidsinterval (met zekere onbetrouwbaarheid α) voor het verschil tussen de gemiddelden (verwachtingen) van de reacties op standaard en praeparaat. D.w.z. een uit de waarnemingen berekend interval, dat behoudens een onbetrouwbaarheid α de werkelijke (onbekende) waarde van dit verschil zal bevatten. Bij een kleinere waarde van α zal dus een langer interval gevonden worden. Bevat het, bij een bepaalde onbetrouwbaarheid α berekende, betrouwbaarheidsinterval de waarde 0 dan kunnen we bij deze onbetrouwbaarheid niet overgaan tot verwerping van de hypothese dat standaard en praeparaat even sterk zijn. Bevat het interval de waarde 0 niet, dan kan de hypothese wel verworpen worden (met een overschrijdingskans $\leq \alpha$, (zie het aan dit rapport toegevoegde memorandum S 47(M 61)), maar uit het betrouwbaarheidsinterval voor het verschil tussen de gemiddelden der reacties volgt nog niets over de grootte van het verschil (of de verhouding) tussen de sterkten van standaard en praeparaat.

Willen we komen tot een ijking die een schatting of betrouwbaarheidsinterval geeft voor het verschil in sterkte (of de verhouding van de sterkten) van praeparaat en standaard (waaruit dus de sterkte van het praeparaat volgt, daar die van de standaard bekend is) dan moet iets bekend zijn over het verband tussen reactie en toegediende dosis [1].

Door verschillende doses van een bekend praeparaat toe te dienen is dit verband verschillende malen onderzocht ([1], [4], [11] en [6]). Hierop zal in de volgende paragraaf nader worden ingegaan. Om een dergelijk verband ook bij elke ijking te toetsen zijn verschillende uitbreidingen van het kruisproefschema voorgesteld, waarbij meerdere doses van standaard en/of praeparaat gebruikt worden ([4], [12] en [B]). Hierbij zijn dan ook meer proefdagen en/of meer groepen konijnen nodig. Een bespreking van de voornaamste van deze schema's is te vinden bij EMMENS [B]. In hoofdstuk 3 is onderzocht in hoeverre het door MARKS [6] gevonden en veel toegepaste reactie-dosis-verband ook door ons materiaal bevestigd wordt. Van belang hierbij is onder andere de vraag of een gevonden relatie nog onderhevig is aan seizoen-invloeden.

2.4. Berekeningsmethoden.

De in de literatuur vermelde berekeningsmethoden vallen uiteen in twee groepen. De oudere veel gebruikte berust op onderzoeken van MARKS, HEMMINGSEN e.a. De tweede groep steunt op de variantie- en regressieanalyse en is o.a. door FIELLER op insuline-ijkingen toegepast.

2.4.1. Methode van MARKS e.a.

Deze onderzoekers vonden uit hun materiaal de volgende betrekking tussen de verhouding van de sterkten van praeparaat en standaard (S_p resp. S_s) en de verhouding van de reacties op praeparaat en standaard ($\bar{y}$ resp. $\bar{x}$ ¹⁾), zie [6] p. 636:

$$(1) \quad \log \frac{S_p}{S_s} = \tau \left(\frac{\bar{y}}{\bar{x}} - 1 \right).$$

Hierin is $\bar{y}$ (resp. $\bar{x}$) het gemiddelde van de reacties van alle konijnen na inspuiting met de praeparaat-dosis (resp. standaard-dosis). Zijn de aantallen reacties op praeparaat en standaard gelijk, dan kan ook met de som der reacties in plaats van hun gemiddelde gewerkt worden. Als reactiemaat van een konijn werd door MARKS de percentuele bloedsuikerdaling gebruikt.

1) De reacties van de konijnen zijn stochastische grootheden, d.w.z. grootheden met een waarschijnlijkheidsverdeling; we geven dit aan door onderstreping van de symbolen. De verwachting, het theoretische gemiddelde, van een stochastische grootheid $\underline{z}$ geven we aan met $\bar{z}$.

Uit groot materiaal werd zo als schatting $\underline{t}$ van τ bepaald:
 $\underline{t} = 0,66$ (zie [6]).

Om deze berekeningsmethode te kunnen toepassen moet eerst worden nagegaan of ook onder gewijzigde omstandigheden deze relatie en deze schatting $\underline{t}$ voor τ voldoen (zie paragraaf 3.5). Hierbij is tevens een onderzoek naar seizoeninvloeden van belang.

Uit hun materiaal bepaalden MARKS en HEMMINGSEN [5] ook de afhankelijkheid tussen de reactie van een konijn en $\underline{b}$, zijn bloedsuikergehalte vlak voor de proef. Zij vonden een stijging van de reactie die 0,22 maal de stijging van $\underline{b}$ bedroeg en stelden een hierop gebaseerde correctie van de reactieverhouding voor. De waarde 0,22 is in paragraaf 3.3 met ons materiaal vergeleken.

Een ernstig nadeel van deze methode is dat ze geen mogelijkheid geeft de spreiding van, of betrouwbaarheidsgrenzen voor, $\log \frac{S_p}{S_s}$ te bepalen. Wel is er een methode om een betrouwbaarheidsinterval voor $\frac{\underline{t}}{\underline{t} \pm \underline{t}}$ te berekenen ([12] en [19]), indien we voor de reacties enkele onderstellingen (o.a. normaliteit) maken. Maar ook de schatting $\underline{t}$ voor τ is een stochastische grootheid (immers berekend uit reacties die stochastisch zijn) en deze wordt éénmaal uit een bepaald materiaal geschat en voor alle volgende berekeningen gebruikt. Hierdoor worden alle schattingen voor de sterkteverhoudingen op dezelfde, maar onbekende, wijze beïnvloed, zelfs indien voldaan is aan de onderstelling dat τ onafhankelijk is van plaats, tijd en ras. M.a.w. alle sterkteverhoudingen met behulp van deze schatting van τ zullen onder- of overschat worden, maar het is onbekend of het een onderschatting of overschatting is en hoe groot deze is. FIELLER e.a. [11] wijzen dan ook terecht op de wenselijkheid de helling uit eigen materiaal steeds opnieuw te schatten.

2.4.2. Toepassing van variantie- en regressieanalyse ²⁾.

FIELLER e.a. [12] gebruikten als berekeningsmethode een directe toepassing van de variantie- en regressieanalyse. De variantie- en regressieanalyse zelf zullen we hier niet behandelen, maar wel zullen we onderstellingen en resultaten bij deze methoden aangeven.

- 2) Deze paragraaf werd ingevoegd als theoretische basis voor de in hoofdstuk 4 voorgestelde gewijzigde schema's. Voor het verkrijgen van een globaal inzicht in het in de hoofdstukken 3 en 4 behandelde is lezing van deze paragraaf niet noodzakelijk.

Zoals steeds bij variantieanalyse het geval is, wordt uitgegaan van de volgende onderstellingen:

- 1) normaliteit der waarnemingen ³⁾,
- 2) homoscedasticiteit bij verschillende doses, etc. (gelijkheid der spreidingen),
- 3) onderlinge onafhankelijkheid van de waarnemingen.

Bovendien wordt nog ondersteld:

- 4) additiviteit van de bijdragen der verschillende invloeden op de reactie (daginvloed, grootte van de dosis, individuele eigenschappen van het konijn).

Bepalen we ons nu tot het gewone kruisproefschema (Schema I, paragraaf 2.3), dan kunnen we de, met dit schema te verkrijgen, waarnemingen als in Schema II samenvatten.

Schema II Overzicht der waarnemingen.

groep konijnen	waarnemingen		reactieverschillen
	1 ^e dag	2 ^e dag	
G ₁	<u>s</u> ₁₁	<u>p</u> ₁₁	<u>p</u> ₁₁ - <u>s</u> ₁₁ = <u>y</u> ₁₁
	.	.	.
	.	.	.
	<u>s</u> _{1n₁}	<u>p</u> _{1n₁}	<u>p</u> _{1n₁} - <u>s</u> _{1n₁} = <u>y</u> _{1n₁}
G ₂	<u>p</u> ₂₁	<u>s</u> ₂₁	<u>p</u> ₂₁ - <u>s</u> ₂₁ = <u>y</u> ₂₁
	.	.	.
	.	.	.
	<u>p</u> _{2n₂}	<u>s</u> _{2n₂}	<u>p</u> _{2n₂} - <u>s</u> _{2n₂} = <u>y</u> _{2n₂}

Uit dit schema zal tevens de gebruikte notatie duidelijk zijn. Geven we de bijdrage door de daginvloed aan met ϕ , de bijdrage van de individuele eigenschappen van het konijn met ψ en zijn δ_p en δ_s de doses insuline (uitgedrukt in internationale eenheden) van resp. praeparaat en standaard, dan specificeert FIELLER het verband tussen reactie en dosis (in overeenstemming met de onderstellingen 1)...4)) als:

3) Een grootheid x heet normaal verdeeld met gemiddelde μ en spreiding σ (aangegeven met $N(\mu, \sigma)$) indien de kans om een waarde voor x kleiner dan een gegeven x te vinden, $P[x \leq x]$, voldoet aan:

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

$$\underline{s}_{1i} = \delta + \phi_1 + \beta \log \gamma_s + \psi_{1i} + \underline{\varepsilon}_{s1i} \quad (i=1, \dots, n_1),$$

$$\underline{s}_{2i} = \delta + \phi_2 + \beta \log \gamma_s + \psi_{2i} + \underline{\varepsilon}_{s2i} \quad (i=1, \dots, n_2),$$

(1)

$$\underline{p}_{1i} = \delta + \phi_1 + \beta \log \gamma_p + \psi_{1i} + \underline{\varepsilon}_{p1i} \quad (i=1, \dots, n_1),$$

$$\underline{p}_{2i} = \delta + \phi_1 + \beta \log \gamma_p + \psi_{2i} + \underline{\varepsilon}_{p2i} \quad (i=1, \dots, n_2),$$

waarin de grootheden $\underline{\varepsilon}$ alle onderling onafhankelijk normaal verdeelde grootheden zijn met gemiddelde 0 en onbekende maar gelijke spreiding. Wil men bovendien de eventuele afhankelijkheid met de beginwaarde $\underline{b}$ van het bloedsuikergehalte onderzoeken dan kunnen we (1) uitbreiden tot

$$\underline{s}_{1i} - \kappa \underline{b}_{1i} = \delta + \phi_1 + \beta \log \gamma_s + \psi_{1i} + \underline{\varepsilon}_{s1i},$$

$$\underline{s}_{2i} - \kappa \underline{b}_{2i} = \delta + \phi_2 + \beta \log \gamma_s + \psi_{2i} + \underline{\varepsilon}_{s2i},$$

(2)

$$\underline{p}_{1i} - \kappa \underline{b}_{2i} = \delta + \phi_2 + \beta \log \gamma_p + \psi_{1i} + \underline{\varepsilon}_{p1i},$$

$$\underline{p}_{2i} - \kappa \underline{b}_{1i} = \delta + \phi_1 + \beta \log \gamma_p + \psi_{2i} + \underline{\varepsilon}_{p2i},$$

waarin de eerste index van $\underline{b}$ de dag aanwijst, de tweede index de groep en de derde index het nummer van het konijn in de groep. In deze vergelijkingen zijn κ , δ , ϕ_1 , ϕ_2 , β , ψ_{1i} en ψ_{2i} onbekende parameters. Uitgaande van de onderstellingen 1)...4) en van (2) is nu, bij voldoende waarnemingsmateriaal, bij verschillende doses van de standaard of het praeparaat, met behulp van de variantieanalyse te onderzoeken of verschillen bestaan tussen konijnen (ψ), tussen dagen (ϕ), of β en κ constant zijn over verschillende dagen (of seizoenen), of κ een van nul te onderscheiden waarde heeft en of de onderstelling van lineariteit van het verband tussen reactie en dagdosis en tussen reactie en beginwaarde bevestigd wordt.

Op deze toetsen gaan we hier niet verder in.

Wel zullen we een schatting voor β en voor de sterkte van het praeparaat geven. Hierbij gaan we uit van (1) en corrigeren dus niet voor eventuele afhankelijkheid van de beginwaarde. Dezelfde afleidingen zouden ook met (2) gegeven kunnen worden, waarbij dan tevens een schatting voor κ te vinden zou zijn.

Noemen we:

$$\underline{\varepsilon}_{p1i} - \underline{\varepsilon}_{s1i} = \underline{\varepsilon}_{1i} \quad ,$$

$$\underline{\varepsilon}_{p2i} - \underline{\varepsilon}_{s2i} = \underline{\varepsilon}_{2i}$$

en

$$\phi_2 - \phi_1 = \phi$$

dan volgt uit (1) en schema II:

$$(3) \quad \underline{y}_{1i} = \phi + \rho \log \frac{\delta_p}{\delta_s} + \underline{\varepsilon}_{1i} \quad (i = 1, \dots, n_1),$$

$$\underline{y}_{2i} = -\phi + \rho \log \frac{\delta_p}{\delta_s} + \underline{\varepsilon}_{2i} \quad (i = 1, \dots, n_2).$$

Hierin is dus δ_p de gezochte praeparaatsterkte.

Om een schatting voor ρ mogelijk te maken moeten een aantal kruisproeven met verschillende praeparaatdoses verricht worden. In overeenstemming met de voorstellen uit hoofdstuk 4 zullen we onderstellen dat steeds tegelijkertijd twee kruisproeven met verschillende praeparaatdoses worden uitgevoerd. Het schema van deze tweevoudige kruisproef is te vinden in hoofdstuk 4 (Schema III), terwijl de werkschema's 1,I; 2,I en 3,I, welke aan het einde van hoofdstuk 4 bijgevoegd zijn, een overzicht van de te verkrijgen waarnemingen geven en de notatie duidelijk zullen maken.

Deze tweevoudige kruisproef denken we ons een aantal malen in de loop van enige tijd uitgevoerd (met hetzelfde of met andere onbekende praeparaten). Hierbij onderstellen we dat de verhouding tussen de doses van het praeparaat bij een tweevoudige kruisproef steeds constant gehouden wordt. De grootheden bij de ν^e tweevoudige kruisproef zullen we voorzien van een index ν vóór het symbool. Analooq aan (3) kunnen we dan voor de reactieverschillen tussen praeparaat en standaard schrijven:

$$\nu \underline{y}_{1i} = \nu \phi + \rho \log \frac{\nu \delta_1}{\delta_s} + \nu \underline{\varepsilon}_{1i} \quad (i = 1, \dots, \nu n_1),$$

$$\nu \underline{y}_{2i} = \nu \phi + \rho \log \frac{\nu \delta_2}{\delta_s} + \nu \underline{\varepsilon}_{2i} \quad (i = 1, \dots, \nu n_2),$$

$$(4) \quad \nu \underline{y}_{3i} = -\nu \phi + \rho \log \frac{\nu \delta_1}{\delta_s} + \nu \underline{\varepsilon}_{3i} \quad (i = 1, \dots, \nu n_3),$$

$$\nu \underline{y}_{4i} = -\nu \phi + \rho \log \frac{\nu \delta_2}{\delta_s} + \nu \underline{\varepsilon}_{4i} \quad (i = 1, \dots, \nu n_4)$$

met

$$\nu = 1, \dots, k.$$

$\nu \delta_1$ en $\nu \delta_2$ zijn dus de onbekende hoeveelheden eenheden insuline in de 1^e en 2^e dosis van het praeparaat. Wel is hun ver-

houding bekend (uit de aantallen cc of mg die van het praeparaat zijn toegediend), zodat we dus kunnen schrijven:

$$\frac{\gamma_1}{\gamma_5} = \gamma \cdot p_1 \quad \text{en} \quad \frac{\gamma_2}{\gamma_5} = \gamma \cdot p_2,$$

waarin γ onbekend is maar p_1 en p_2 de bekende verhoudingsgetallen zijn. Definieren we nu:

$$\gamma \stackrel{\text{def}}{=} \beta \log \gamma \quad \text{en} \quad x_j \stackrel{\text{def}}{=} \log p_j \quad (j=1,2),$$

dan gaat (4) over in:

$$\begin{aligned} \gamma y_{1i} &= \gamma \phi + \gamma c + \beta x_1 + \gamma \varepsilon_{1i}, \\ \gamma y_{2i} &= \gamma \phi + \gamma c + \beta x_2 + \gamma \varepsilon_{2i}, \\ \gamma y_{3i} &= -\gamma \phi + \gamma c + \beta x_1 + \gamma \varepsilon_{3i}, \\ \gamma y_{4i} &= -\gamma \phi + \gamma c + \beta x_2 + \gamma \varepsilon_{4i}. \end{aligned} \quad (5)$$

Het bepalen van de meest aannemelijke schattingen (Engels: maximum likelihood estimates) voor de onbekende parameters komt in gevallen als dit neer op het toepassen van de methode der kleinste kwadraten, dus het bepalen van die schattingen voor β , γc en $\gamma \phi$ (deze laatste interesseren ons minder) welke de uitdrukking:

$$\begin{aligned} \sum_{j=1}^k \left[\sum_{i=1}^{n_j} (\gamma y_{ji} - \gamma \phi - \gamma c - \beta x_j)^2 + \right. \\ \left. + \sum_{i=1}^{n_2} (\gamma y_{2i} - \gamma \phi - \gamma c - \beta x_2)^2 + \right. \\ \left. + \sum_{i=1}^{n_3} (\gamma y_{3i} + \gamma \phi - \gamma c - \beta x_1)^2 + \right. \\ \left. + \sum_{i=1}^{n_4} (\gamma y_{4i} + \gamma \phi - \gamma c - \beta x_2)^2 \right] \end{aligned}$$

minimaliseren. Voor de eenvoud der formules zullen we nu onderstellen dat $\gamma n_1 = n_2 = \gamma n_3 = \gamma n_4 = \gamma n$ is voor alle j ; in werkschema 3 zijn de formules voor ongelijke aantallen gebruikt. Bij uitwerking vinden we als kleinste kwadraten-schatting voor β :

$$(7) \quad \hat{\beta} = \frac{1}{x_1 - x_2} \cdot \frac{1}{2} \cdot \frac{\sum_j n_j (\bar{y}_1 + \bar{y}_3 - \bar{y}_2 - \bar{y}_4)}{\sum_j n_j},$$

waarin

$$\bar{y}_j \stackrel{\text{def}}{=} \frac{1}{n_j} \sum_{i=1}^{n_j} y_{ji} \quad (j = 1, 2, 3, 4).$$

In hoofdstuk 4 zal voorgesteld worden deze schatting voor β uit k tweevoudige kruisproeven te gebruiken bij de sterktebepaling van het praeparaat in de $(k+1)^e$ tweevoudige kruisproef. Dit kan dan als volgt geschieden: De kleinste kwadraten-schatting voor β uit één tweevoudige kruisproef is (we laten nu de index j weg):

$$\hat{\beta} = \frac{1}{4} (\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4) - \hat{\beta} \bar{x},$$

met

$$\bar{x} = \frac{1}{2} (x_1 + x_2).$$

Als schatting voor

$$\mu \stackrel{\text{def}}{=} \log y = \frac{c}{\beta}$$

volgt hieruit:

$$(8) \quad \hat{\mu} = \frac{\frac{1}{4} (\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4)}{\hat{\beta}} - \bar{x}$$

Hierin wordt dan de in (7) gegeven schatting voor β uit de k vooraafgaande ijkingen gebruikt.

Daar $y = \text{antilog } \mu$ is en $y_1 = y \cdot y_s \cdot p_1$ het aantal eenheden insuline in de eerste praeparaatdosis voorstelt, vinden we als schattingen:

$$\hat{y} = \text{antilog } \hat{\mu}$$

en

$$(9) \quad \hat{y}_1 = (\text{antilog } \hat{\mu}) \cdot y_s \cdot p_1.$$

Het is mogelijk voor μ en dus voor y , exacte betrouwbaarheidsgrenzen ⁴⁾ te bepalen en wel met de door FIELLER [12], [19]

4) We spreken van een betrouwbaarheidsinterval, met onbetrouwbaarheid α , voor een onbekende parameter indien de kans dat dit interval de werkelijke waarde van de parameter niet zal bevatten gelijk aan α is. Dit betekent dus dat, wanneer dergelijke intervallen vele malen bepaald worden, een gedeelte α van deze intervallen niet de werkelijke waarde van de bijbehorende onbekende parameter zal bevatten.

en [20] beschreven methode, met de aanvulling die in [21] wordt gegeven. We zullen deze methode niet behandelen; in hoofdstuk 4, werkschema 2, III worden de eindformules gegeven voor deze betrouwbaarheidsgrenzen. Wel vermelden we, zonder afleiding, nog een benadering voor de variantie van $\hat{\mu}$ die in de werkschema's 1, III en 3, III gebruikt wordt bij de bepaling van benaderde betrouwbaarheidsgrenzen. Deze benadering vereist minder rekenwerk en geeft een in vele gevallen goede benadering van de betrouwbaarheidsgrenzen (de benadering wordt beter naarmate $\frac{\hat{\beta}}{S_{\hat{\beta}}}$ groter is).

Noemen we:

$$(10) \quad \bar{A} = \frac{1}{4} (\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4)$$

dan wordt een benadering van de variantie van $\hat{\mu}$ gegeven door:

$$\sigma_{\hat{\mu}}^2 = \frac{\sigma_A^2}{\beta^2} + \frac{\mu^2 \sigma_{\hat{\beta}}^2}{\hat{\beta}^4},$$

... $\hat{\beta}$ onafhankelijk verdeeld zijn (dit wordt gewaarborgd doordat ze uit verschillend waarnemingsmateriaal bepaald worden). Een schatting $S_{\hat{\mu}}^2$ van $\sigma_{\hat{\mu}}^2$ verkrijgen we door μ en β door hun schattingen $\hat{\mu}$ en $\hat{\beta}$ en σ_A^2 en $\sigma_{\hat{\beta}}^2$ door hun schattingen S_A^2 en $S_{\hat{\beta}}^2$ te vervangen. Uit de uitdrukkingen voor $\bar{A}$ (10) en $\hat{\beta}$ (7) volgt:

$$S_A^2 = \frac{1}{16n(n-1)} \sum_{j=1}^4 \sum_{i=1}^n (\underline{y}_{ji} - \bar{y}_j)^2$$

en

$$S_{\hat{\beta}}^2 = \frac{\sum_{j=1}^k S_j^2}{\left[\sum_{j=1}^k n(x_1 - x_2) \right]^2},$$

waarin

$$S_j^2 = \frac{1}{4(n-1)} \sum_{j=1}^4 \sum_{i=1}^n (\underline{y}_{ji} - \bar{y}_j)^2.$$

De benadering voor een betrouwbaarheidsinterval voor μ is nu (betrouwbaarheid 0,05):

$$(11) \quad (\hat{\mu} - 1,96 S_{\hat{\mu}}; \hat{\mu} + 1,96 S_{\hat{\mu}}),$$

dus voor γ_i :

$$(12) \quad \left(\frac{\gamma_s \cdot p_i \cdot \text{antilog } \hat{\mu}}{\text{antilog}(1,96 S_{\hat{\mu}})}; \gamma_s \cdot p_i \cdot \text{antilog } \hat{\mu} \cdot \text{antilog}(1,96 S_{\hat{\mu}}) \right).$$

Dezelfde afleidingen kunnen gegeven worden indien de vier groepen in een tweevoudige kruisproef niet even groot zijn ($\gamma_1, \gamma_2, \gamma_3$ en γ_4 niet gelijk). De resultaten zijn te vinden in werkschema 3 (achter hoofdstuk 4). Een sterke vereenvoudiging van bovengegeven formules zou nog optreden indien gold $\gamma = n$ voor alle γ .

3. Analyse van het waarnemingsmateriaal.

3.1. Overzicht van het materiaal.

Ter demonstratie van de gebruikte ijkmethode werden de resultaten van een aantal kruisproeven verstrekt. De verderop genoemde "grafiekjes" zijn resultaten van proeven over het gedrag van konijnen bij verschillende doses insuline en over het verloop van het bloedsuikergehalte na insuline-inspuiting.

De gegevens bestonden uit:

- I: 9 kruisproeven uit de periode 20 December 1949 - 20 April 1950; geïnjecteerd $0,5 E$ van de standaards; per kruisproef ongeveer 22 konijnen.
- II: 8 kruisproeven uit de periode 28 Maart 1951 - 8 Juni 1951; geïnjecteerd $1 E$ van de standaard; per kruisproef ongeveer 22 konijnen.
- III: 56 grafiekjes van het verloop van het bloedsuikergehalte na injectie met insuline bij 56 konijnen, alleomstreeks April 1951. Elk konijn werd drie maal gebruikt, nl. met $1 E$, $1\frac{1}{2} E$ en $2 E$ insuline. Elke keer werd bepaald het bloedsuikergehalte vlak voor de injectie en op 1, 2, 3, 4 en 5 uur na de injectie.
- IV: 141 grafiekjes als boven, nu echter omstreeks September 1951, terwijl per konijn slechts 2 injecties werden uitgevoerd, nl. met $\frac{1}{2} E$ en $1 E$ insuline. Het bloedsuikergehalte werd vlak voor de injectie en op de tijdstippen 1, 2 en 3 uur na de injectie bepaald.

De proeven beschreven met de grafiekjes bleken meestal weer volgens een kruisproefschema verricht te zijn. Bij de 56 grafiekjes bleken de proeven met $1\frac{1}{2} E$ zich over een kleiner aantal dagen uit te strekken dan die met 1 en $2 E$ insuline.

3.2. Normaliteit en homoscedasticiteit.

Zoals in paragraaf 2.4.2 werd vermeld, zijn normaliteit en homoscedasticiteit van de waarnemingen noodzakelijke onderstellingen om variantieanalyse te kunnen toepassen.

Over de normaliteit is een klein oriënterend onderzoek uitgevoerd bij de waarnemingengroep IV. In figuur 1 zijn de fre-

quenties van de reacties tegen de reacties uitgezet voor $\frac{1}{2}$ E en 1 E insuline; als reactie werd hier genomen de percentuele daling.

Passen we de χ^2 -toets (zie memorandum S 47 (M 12)) voor aanpassing van een normale verdeling toe, dan vinden we bij $\frac{1}{2}$ E een overschrijdingskans (zie memorandum S 47 (M 6)) $k = 0,65$ en bij 1 E : $k = 0,60$. De hypothese van normaliteit kan dus zeker voor deze waarnemingsgroep niet verworpen worden.

Door de begrensdheid van het interval waarin de reacties kunnen liggen (bij percentuele bloedsuikerdaling in ieder geval tussen 0 en 100%, maar de grenzen zijn in de praktijk nauwer daar dieren met een erg hoge of lage reactie uitgesloten worden, verdacht van ziekte) zal de verdeling van de reacties scheef kunnen worden bij hele grote of hele kleine doses. Het zal daarom aanbeveling verdienen de dosis zoveel mogelijk zo te kiezen dat de reacties om het midden van het interval gegroepeerd liggen.

De homoscedasticiteit is niet speciaal onderzocht; wel zijn in de loop van het onderzoek een aantal varianties van de reacties uit de grafiekjes berekend, die we in tabel I zullen samenvatten. Hierbij moet nog bedacht worden dat de varianties uit alle grafiekjes berekend zijn en de waarnemingen zich dus over meer dagen uitstrekken dan gewoonlijk bij een ijking het geval is; hierdoor kan de homogeniteit der varianties ongunstig beïnvloed zijn. De varianties zijn berekend voor 1) de percentuele daling (bij het gemiddelde van 3 uurafnamen), 2) het absolute niveau (gemiddelde van 3 uurafnamen), 3) het absolute niveau van een bloedafname 1 uur na inspuiting.

Tabel I Varianties van reacties

periode	reactiemaat	dosis	variantie
April '51	percentuele daling	1 E	115,5
		$\frac{1}{2}$ E	137,5
		2 E	137,5
	abs. niveau 3 bloedafnamen	1 E	126,5
		$\frac{1}{2}$ E	208,6
		2 E	242,0
Sept. '51	abs. niveau 1 bloedafname	1 E	231,0
		$\frac{1}{2}$ E	192,5
		2 E	242,0
	percentuele daling	$\frac{1}{2}$ E	96,60
		1 E	106,40
	abs. niveau 3 bloedafn.	$\frac{1}{2}$ E	105,0
		1 E	101,8
	abs. niveau 1 bloedafn.	$\frac{1}{2}$ E	75,6
		1 E	74,2

We zien dat de verschillen in elke periode tussen de verschillende toegediende doses niet zeer groot zijn. Wel bestaan er grotere verschillen tussen de varianties gevonden in de twee perioden.

3.3. Onderzoek in verband met een keuze van de reactiemaat.

Een belangrijke vraag is, vooral in verband met de bewerkelijkheid der berekeningen bij een ijking, of de gekozen maat voor de reactie nog afhankelijk is van het begin-bloedsuikergehalte b . Bovendien is het van belang om, indien een dergelijke afhankelijkheid bestaat, te bepalen welke correctie hiervoor toegepast kan worden.

De afhankelijkheid tussen begin-bloedsuikergehalte en reactie is op verschillende manieren onderzocht.

1) Op de gevens bij de kruisproeven is een parameter vrije toetsingsmethode (hierbij wordt niets over de verdeling der waarnemingen ondersteld) toegepast, nl. de rangcorrelatiemethode van KENDALL (zie memorandum S 47 (M 13)). De te toetsen hypothese is hier: er bestaat geen correlatie tussen reactie en begin-bloedsuikergehalte. Voor elke dag en voor praeparaat en standaard apart werd de grootheid $\underline{S}$ van KENDALL bepaald. Is het aantal waarnemingen niet te klein, dan bezit deze grootheid bij benadering een normale verdeling met een bekende spreiding (alleen afhankelijk van het aantal waarnemingen en de aantallen gelijke waarnemingen). Het is dan dus mogelijk deze grootheden voor het hele materiaal of gedeelten ervan door optelling te combineren tot een grootheid welke weer normaal verdeeld is met bekende spreiding. In een tabel van de normale verdeling is dan de overschrijdingskans bij de gevonden waarde op te zoeken. (De overschrijdingskans is de kans dat deze of een nog verder van 0 verwijderde waarde gevonden zal worden indien de onafhankelijkheidshypothese juist is. Een kleine overschrijdingskans leidt dus tot verworping van deze hypothese.) Bij het kruisproefmateriaal werden de in tabel II vermelde resultaten gevonden. Als reactiemaat diende hier de percentuele bloedsuikerdaling.

Tabel II Correlatie tussen reactie en begingehalte

	overschrijdingskans
alle 1 ^e dagen	0.04
alle 2 ^e dagen	0.74
kruisproeven 1 - 9:	
1 ^e dagen	0.03 < k < 0.04
2 ^e dagen	
kruisproeven 1 ^A - 8 ^A :	
1 ^e dagen	0.77
2 ^e dagen	0.88

De eerste kolom geeft het gedeelte van het materiaal waarvoor de grootheden $\underline{S}$ gecombineerd zijn.

De 1^e dagen van de kruisproeven 1 - 9 geven dus een zwakke aanwijzing voor afhankelijkheid tussen reactie en begin-bloedsuikergehalte.

2) Gaan we uit van de onderstelling dat $\underline{b}_i$ en y_i een ^{normale} simultane verdeling bezitten, dan is het mogelijk de correlatiecoëfficiënt ρ tussen $\underline{b}_i$ en y_i te schatten en te toetsen. De verdeling van $\hat{\rho}$ is door F.N.DAVID, Tables of the correlation coefficient, bepaald, met tabellen voor de toetsing van verschillende hypothesen van de vorm $\rho = \rho_0$. De schatting van ρ is

$$\hat{\rho} = \frac{\sum_i (y_i - \bar{y})(\underline{b}_i - \bar{\underline{b}})}{\sqrt{\sum_i (y_i - \bar{y})^2 \sum_i (\underline{b}_i - \bar{\underline{b}})^2}} ;$$

de door ons getoetste hypothese was $\rho = 0$. In tabel III zijn de resultaten vermeld voor twee reactiematen: de percentuele daling van het gemiddelde van drie uurafnamen na injectie en het absolute niveau uit één bloedafname op 1 uur na injectie. De berekening werd alleen uitgevoerd voor de grafiekjes $\frac{1}{2} E - 1 E$.

Tabel III Correlatiecoëfficiënt tussen reactie en begin-bloedsuikergehalte voor de periode September 1951.

reactie-maat	dosis	$\hat{\rho}$	overschrijdings-kans (éénzijdig)	aantal reacties
% daling	$\frac{1}{2} E$	0.155	0.04	133
	1 E	0.213	0.012	133
abs. niveau 1 bloedafn.	$\frac{1}{2} E$	0.1026	0.17	140
	1 E	0.2006	0.03	140

Deze onderzoeken geven dus wel enige aanleiding om voor beide reactiematen een afhankelijkheid van het begin-bloedsuikergehalte te onderstellen.

3) De afhankelijkheid tussen reactie en begingehalte is nog op een derde manier bekeken, welke tevens een mogelijkheid biedt een vergelijking te maken met de gebruikelijke correctiemethode (HEMMINGSSEN en MARKS [5]). We nemen nu het begingehalte als onafhankelijk veranderlijke en bepalen de regressiecoëfficiënt van de reactie t.o.v. het begingehalte. Deze regressiecoëfficiënt is per periode en per dosis bepaald, terwijl bij dit gedeelte van het onderzoek ook naar geslachten is onderscheiden. Noemen we voor het i^e konijn het begingehalte $\underline{b}_i$ en de reactie y_i dan is de regressiecoëfficiënt:

$$\underline{z} = \frac{\sum_i (y_i - \bar{y})(b_i - \bar{b})}{\sum_i (b_i - \bar{b})^2},$$

waarin

$$\bar{b} = \frac{1}{n} \sum_i b_i,$$

indien n het aantal dieren uit de beschouwde groep is. Als reactiemaat is hierbij de percentuele bloedsuikerdaling, uit alle bloedafnamen tezamen, en in een paar gevallen het absolute niveau na één uur, genomen. Is $\underline{s}_y^2$ de schatting voor de variantie van de reacties y_i , dan wordt de schatting voor de variantie van $\underline{z}$:

$$\underline{s}_z^2 = \frac{\underline{s}_y^2}{\sum_i (b_i - \bar{b})^2}.$$

Zijn de grootheden y_i normaal verdeeld met spreiding σ_y , dan is ook $\underline{z}$ normaal verdeeld en wel $N(0, \frac{\sigma_y^2}{\sum_i (b_i - \bar{b})^2})$, indien

te toetsen hypothese van onafhankelijkheid tussen y en b juist is. In een tabel van de STUDENT-verdeling met $(n-1)$ vrijheidsgraden kan dan de overschrijdingskans opgezocht worden van de waarde $\frac{\underline{z}}{\underline{s}_z}$ (zie memorandum S 47 (M 8)). In tabel IV zijn de uitkomsten voor beide groepen grafiekjes opgegeven. Voor het absolute niveau is de regressiecoëfficiënt alleen bepaald voor de markantste groepen bij de percentuele daling, in verband met het vele rekenwerk.

Tabel IV Regressiecoëfficiënt van reactie en begingehalte

reactie- maat	periode	ge- slacht	dosis	aan- tal	$\underline{z}$	$\underline{s}_z$	overschr. kans
percen- tuele bloed- suiker- daling	Sept. '51	♂	$\frac{1}{2} E$	62	0,36	0,43	>0,4
		♂	1 E	62	0,15	0,14	>0,3
		♀	$\frac{1}{2} E$	71	0,015	0,13	>0,9
		♀	1 E	71	0,20	0,17	>0,2
		♂ + ♀	$\frac{1}{2} E$	133	0,17	0,097	>0,05
		♂ + ♀	1 E	133	0,25	0,11	≈ 0,05 ←
		♂ + ♀	$\frac{1}{2}$ en 1 E	133	0,18	0,09	≈ 0,05 ←
	April '51	♂	1 E	31	0,25	0,27	>0,3
		♂	$1\frac{1}{2} E$	31	-1,46	0,27	<0,001 ←
		♂	2 E	31	-0,21	0,51	>0,7
		♀	1 E	25	0,45	0,20	≈ 0,05 ←
		♀	$1\frac{1}{2} E$	25	-0,07	0,20	≈ 0,7
		♀	2 E	25	0,23	0,19	>0,2
absolute niveau na 1 uur	Sept. '51	♂ + ♀	1 E	140	0,22	0,189	<0,03 ←
	April '51	♂	$1\frac{1}{2} E$	31	0,80	0,25	<0,005 ←

De resultaten geven dus geen aanleiding systematische verschillen tussen mannelijke en vrouwelijke dieren te veronderstellen. De zes kleine overschrijdingskansen (op een totaal van 12 onafhankelijke onderzoeken), wijzen wel op verband tussen begin-gehalte en percentuele daling; er is echter met deze aanwijzing door de uiteenlopende waarden (zelfs één negatieve onder de significant van 0 verschillende) van de regressiecoëfficiënten weinig te beginnen. En zeer zeker is in dit materiaal geen grond voor de onderstelling van een constant verband met regressiecoëfficiënt 0.22 zoals door HEMMINGSEN en MARKS wordt aangenomen [5].

3.4. Gevoeligheid der konijnen.

Tijdens het onderzoek kwam een eigenschap der konijnen naar voren, die wellicht van belang kan zijn voor de vergroting van de nauwkeurigheid van een ijking. Het bleek namelijk dat door de daginvloed wel alle reacties van een groep konijnen verhoogd of verlaagd werden, maar dat de volgorde van reactiehoogte binnen een groep praktisch niet door deze daginvloed beïnvloed wordt. Deze eigenschap is onderzocht door bij de kruisproeven de reacties van 2 dagen van dezelfde groep konijnen met elkaar te vergelijken wat hun volgorde van grootte betreft. De waarden, die de $\bar{S}$ van KENDALL (zie memorandum S 47 (M 13)), per groep berekend, aanneemt, worden voor het hele kruisproefmateriaal opgeteld en getoetst. Bij de grafiekjes van $\frac{1}{2} E - 1 E$ zijn alle reacties op $\frac{1}{2} E$ op deze wijze vergeleken met die van $1 E$, zodat hieruit één $\bar{S}$ -waarde berekend werd. Hierbij kan dus het feit dat niet alle dieren op dezelfde dagen behandeld werden, storend werken, hetgeen bij de kruisproeven niet het geval is. De resultaten zijn verzameld in tabel V.

Tabel V Gevoeligheid der konijnen

materiaal	$\sum S$	$\sigma_{\sum S}$	overschr.kans
kruisproeven	856	75.35	$\ll 0.01$
grafiekjes $\frac{1}{2} E - 1 E$	601	201.46	$0.02 > k > 0.01$

De kleine overschrijdingskansen geven hier dus een duidelijke aanwijzing voor de gecorreleerdheid der reacties binnen een groep, dus voor een systematisch verschil in gevoeligheid tussen de konijnen.

3.5. Het reactie-dosis verband.

Uit de betrekking tussen reactie en dosis welke door MARKS

[4] ondersteld werd (zie paragraaf 2.4.1):

$$\log \frac{S_p}{S_s} = \tau \left(\frac{t \bar{y}}{t \bar{x}} - 1 \right),$$

waarin $\bar{y}$ resp. $\bar{x}$ de gemiddelde reacties op praeparaat resp. standaard voorstellen, volgt:

$$\tau = \frac{t \bar{x}}{t(\bar{y} - \bar{x})} \log \frac{S_p}{S_s}.$$

Uit de grafiekjes waar we steeds twee bekende doses van een insuline-praeparaat vergelijken, dus $\log \frac{S_p}{S_s}$ bekend is, kunnen we nu een schatting t voor τ met een betrouwbaarheidsinterval, volgens de methode van FIELLER [19], bepalen. De resultaten zijn vermeld in tabel VI. Het betrouwbaarheidsinterval is met een betrouwbaarheid 0,95 bepaald. Dit betekent dus dat in 95% van de gevallen waarin we een dergelijk interval bepalen, dit interval de werkelijke parameterwaarde zal bevatten. Voor de grafiekjes 1 - 1½ - 2 E is de berekening voor 1 en 1½ E en voor 1 en 2 E uitgevoerd; deze bepalingen zijn dus niet onafhankelijk van elkaar.

Tabel VI Betrouwbaarheidsinterval voor de coëfficiënt van MARKS

periode	aantal dieren	geslacht	vergeleken doses	onder grens	t	boven grens
Sept. '51	71	♀	½ en 1 E	0.962	1.136	1.302
	62	♂	½ en 1 E	0.887	1.054	1.217
	133	♀ + ♂	½ en 1 E	0.981	1.099	1.215
April '51	25	♀	1 en 1½ E	-0.75	0.267	1.12
	25	♀	1 en 2 E	0.13	0,607	1.03
	30	♂	1 en 1½ E	-0.17	0.546	1.19
	30	♂	1 en 2 E	0.64	0.912	1.17
	55	♀ + ♂	1 en 1½ E	-0.14	0.418	0.93
	55	♀ + ♂	1 en 2 E	0.53	0.776	1.01

Bij deze berekeningen is nog weer naar eventuele verschillen tussen ♀ en ♂ gekeken; maar de resultaten wijzen ook hier niet op systematische verschillen.

Wel constateren we een duidelijk verschil tussen de 2 groepen grafiekjes onderling en met de door MARKS [6] gevonden schatting $t = 0,66$ en eveneens tussen de schatting voor τ bij vergelijking van 1 en 1½ E en van 1 en 2 E bij de tweede groep grafiekjes. Een verklaring voor de verschillen tussen de groepen grafiekjes is wellicht het verschillende seizoen waarin de waarnemingen verricht werden. Een mogelijke verklaring voor de ver-

schillen tussen $1-1\frac{1}{2} \bar{E}$ en $1-2 \bar{E}$ kan zijn dat het verband tussen reactie en dosis in dit vrij grote gebied niet lineair is maar als in figuur II geschetst.

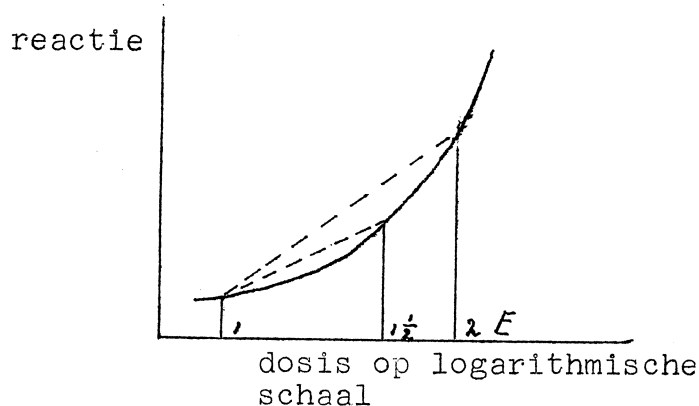


Fig. II Een mogelijk reactie-dosis verband

3.6. Conclusies.

1) Uit paragraaf 3.2 mogen we afleiden dat er geen ernstige bezwaren zijn tegen de onderstellingen van normaliteit en homoscedasticiteit van de reacties binnen een ijking. Wel waren er aanwijzingen voor een mogelijke tijdsafhankelijkheid van de spreidingen.

2) Uit paragraaf 3.3 volgt, dat er wel is waar aanwijzingen bestaan voor een correlatie tussen percentuele bloedsuikerdaling (en eveneens voor het absolute niveau na 1 uur) en begingehalte, maar dat dit verband sterk wisselend is tussen verschillende groepen waarnemingen, zodat het zeer twijfelachtig is, of een correctie hiervoor aangebracht kan worden. Voor de eenvoud der bepalingen en berekeningen verdient het daarom wellicht aanbeveling geen correctie voor de beginwaarde toe te passen.

3) Bij ons onderzoek naar het reactie-dosis verband, zoals dit door MARKS werd ondersteld en ook bij veel routine-ijkingen wordt gebruikt, vonden we schattingen voor helling τ na de ijklijn, welke sterk afweken van de door MARKS opgegeven waarde en ook onderling grote verschillen vertoonden; hierbij hebben vermoedelijk o.a. seizoeninvloeden een rol gespeeld (paragraaf 3.5). Op grond hiervan is het niet aan te raden de schatting 0,66 voor τ van MARKS zonder controle over te nemen. In hoofdstuk 4 zullen we een ijkschema en een berekeningsmethode voorstellen, waarbij een regressiecoëfficiënt uit eigen materiaal wordt geschat.

4) Bij de behandeling van het waarnemingsmateriaal kwam duidelijk naar voren dat er systematische gevoeligheidsverschillen tussen de konijnen bestaan, welke gedurende enige tijd behouden blijven, ondanks het feit dat door daginvloeden de gemid-

delde reactie van de hele groep aan sterke wisselingen onderhevig is (paragraaf 3.4). In hoofdstuk 4 zal een suggestie gedaan worden om met behulp van deze eigenschap een selectie op de konijnen toe te passen, waardoor wellicht de nauwkeurigheid van de ijkingen vergroot kan worden.

4. Voorstellen over een gewijzigd ijkschema met berekeningsmethoden.

Deze voorstellen zullen omvatten:

1) een nieuw ijkschema, waarbij het mogelijk zal zijn uit iedere ijking gegevens te putten ter schatting van de helling van de ijklijn (reactie-log dosis verband),

2) berekeningsmethoden voor de sterkte van het insulinepraeparaat, waarbij gebruik wordt gemaakt van een uit een aantal onmiddellijk voorafgaande ijkingen geschatte ijklijn,

3) een selectie van de konijnen in verband met hun gevoeligheid waardoor wellicht de nauwkeurigheid vergroot kan worden en

4) de inrichting van een grafische controle, waardoor de veranderlijkheid van de helling van de ijklijn onderzocht kan worden en enige controle op iedere ijking mogelijk is.

In verband met de onduidelijke uitkomsten voor het verband tussen reactie en begingehalte (zie conclusie 2)) is correctie hiervoor achterwege gelaten.

De benodigde waarnemingen en berekeningen zijn in de aan het eind van dit hoofdstuk bijgevoegde werkschema's samengevat. Deze schema's worden in de volgende paragrafen uitvoeriger besproken. Het zijn:

Werkschema 1. Dit behandelt het geval waarbij per ijking de groepen konijnen, die de verschillende doses toegediend krijgen, even groot zijn en geeft een benadering voor de betrouwbaarheids grenzen voor de sterkte van het praeparaat.

Werkschema 2. Dit behandelt hetzelfde geval maar geeft exacte betrouwbaarheids grenzen.

Werkschema 3. Dit daarentegen geeft benaderde betrouwbaarheids grenzen voor het geval dat ook binnen een ijking de verschillende konijnengroepen in grootte verschillen.

De vierde mogelijkheid: exacte betrouwbaarheids grenzen bij ongelijke groepen is niet uitgewerkt, daar de optredende formules zeer ingewikkeld en dus minder geschikt voor routine-berekeningen zullen worden.

De formules van werkschema 1 volgen uit die van werkschema

3 door $n_1 = n_2 = n_3 = n_4 = n$ te stellen. Heeft in een serie ijkingen een gedeelte wel gelijke groepen dan kunnen voor deze ijkingen dus in werkschema 3 de formules uit werkschema 1 gebruikt worden. Mocht in een serie ijkingen met overigens gelijke groepen per ijking een enkele maal één dier uit één groep wegvallen, dan zullen de resultaten niet veel beïnvloed worden wanneer toch de voor gelijke groepen opgestelde formules gebruikt worden.

Het aantal dieren per groep mag, ook bij gebruik van werkschema 1 of 2, wel van ijking tot ijking variëren.

Als reactiemaat kan iedere maat gebruikt worden die in voldoende mate aan de onderstellingen van paragraaf 2.4.2 voldoet en dit zal vermoedelijk zowel voor percentuele daling als voor absolute daling van één bloedafname gelden, zolang de gemiddelden van de reacties per dosis niet te dicht bij de grenzen van het toegelaten interval komen. Het komt ons aanbevelingswaardig voor de bruikbaarheid van het absolute niveau uit één bloedafname (bv. 1 of $1\frac{1}{2}$ uur na injectie) te onderzoeken (zie overwegingen uit paragraaf 2.1), vooral in verband met de vereenvoudiging van de reactiebepaling die dit geeft en het eventuele vluggere herstel der konijnen.

4.1. Schema voor de routine-ijking (I in werkschema's)

Op grond van de bij het onderzoek van het waarnemingsmateriaal verkregen resultaten (conclusie 2) en 3) uit paragraaf 3.6) leek het ons wenselijk te komen tot een ijking waarbij de helling van de ijklijn uit eigen materiaal geschat en gecontroleerd kan worden, terwijl bovendien een schatting voor de nauwkeurigheid van de ijking mogelijk is. In verband met dit laatste zijn we uitgegaan van de berekeningsmethode van FIELLER (zie paragraaf 2.4.2). Om een bijdrage tot een schatting van de helling te verkrijgen worden per ijking twee verschillende doses van het praeparaat toegediend volgens het tweevoudige kruisproefschema (zie schema III)

Schema III Tweevoudig kruisproefschema.

groepen konijnen	eerste dag	tweede dag
G_1	p_1	s
G_2	p_2	s
G_3	s	p_1
G_4	s	p_2

In schema III is s de standaarddosis en zijn p_1 en p_2 twee verschillende praeparaat doses, die bij een serie routine-ijkin-

gen steeds dezelfde verhouding dienen te hebben (doet men dit niet, dan worden de berekeningen ingewikkelder dan nodig is). Bij de keuze van de doses p_1 en p_2 spelen twee tegengestelde overwegingen een rol, nl. enerzijds bevordert een groot verschil tussen p_1 en p_2 de nauwkeurigheid van de hellingschatting, anderzijds zullen dan de gemiddelde reacties op p_1 en p_2 niet meer beide in de buurt van het midden van het reactie-interval kunnen liggen, zoals voor de normaliteit gewenst was (zie paragraaf 3.2). Wellicht is de verhouding 4:5 of 9:10 een geschikt compromis.

Werkschema 1,I en 2,I zijn opgesteld voor even grote groepen konijnen G_1 , G_2 , G_3 en G_4 ; werkschema 3,I moet gebruikt worden indien deze groepen niet dezelfde aantallen bevatten. In de drie werkschema's 1,I, 2,I en 3,I vinden we de waarnemingen en de eerste berekeningen hieruit, welke bij de verdere berekeningen gebruikt worden. De gebruikte notaties zullen uit de schema's zelf duidelijk zijn.

Opmerkingen:

- 1) De groepen G_1 en G_3 worden met dezelfde praeparatodosi behandeld en vormen dus één kruisproef evenals de groepen G_2 en G_4 .
- 2) Bij de verdeling van de konijnen in de vier groepen is het belangrijk dit werkelijk aselekt (Engels: random) te doen (bv. met behulp van "random numbers"), vooral in verband met de weglating van een correctie voor het begingehalte.
- 3) Bij de berekening van S_1^2 kan men in elk der drie werkschema's met vrucht gebruik maken van de identiteit:

$$\sum_{i=1}^n (y_{ji} - \bar{y}_j)^2 = \sum_{i=1}^n y_{ji}^2 - n \bar{y}_j^2$$

waardoor de berekeningen eenvoudiger kunnen worden uitgevoerd.

4.2. Schatting van de helling (II in de werkschema's).

Daar een schatting van de helling per ijking een te grote onnauwkeurigheid zou bezitten en het toch van belang is seizoeninvloeden in rekening te brengen (zie conclusie 3) paragraaf 3.6) wordt voorgesteld deze helling uit een aantal onmiddellijk voorafgaande ijkings te schatten⁵⁾. Op deze manier wordt een soort

-
- 5) De waarnemingen uit de in behandeling zijnde ijking kunnen beter niet voor de bepaling van β gebruikt worden om de onafhankelijkheid tussen $\bar{y}$ en β te verkrijgen welke nodig is bij de benadering van $\sigma_{\hat{\beta}}^2$, die we zullen gebruiken (zie paragraaf 4.3).

voortschrijdende ijklijn verkregen, die beter dan een vaste ijklijn aangepast is aan de gemiddelde omstandigheden in de proefperiode. Het is duidelijk dat aan de ene kant om de aanpassing aan de omstandigheden gevoelig te maken, de waarnemingen voor de helling-bepaling zich over een niet te lange periode moeten uitstrekken, terwijl aan de andere kant een vergroting van het aantal waarnemingen een vergroting van de nauwkeurigheid van de schatting van β tengevolge zal hebben. Hiertussen zal een middenweg gezocht moeten worden, waarbij de grafische controle, zoals deze in paragraaf 4.4 besproken zal worden, wellicht van dienst kan zijn.

Geven we de berekende grootheden bij de ν^e ijking, uit de serie van m ijkingen waaruit we de helling willen schatten, aan met de index ν vóór het symbool en geven we de sommatie van een grootheid over bovengenoemde serie aan met $\sum_{\nu}$ dan is de meest aannemelijke (Engels: maximum likelihood) schatting voor β :

$$\hat{\beta} = \frac{\sum_{\nu} n (\bar{y}_1 - \bar{y}_2 + \bar{y}_3 - \bar{y}_4)}{2(x_1 - x_2) \sum_{\nu} n},$$

waarin $x_1 - x_2 = \log \frac{p_1}{p_2}$ is (hierin is $\frac{p_1}{p_2}$ de voor alle ijkingen gelijke verhouding van de twee praeparatdoses) en n het aantal dieren per groep is in de ν^e ijking. Als schatting voor de variantie van β wordt bepaald:

$$s_{\hat{\beta}}^2 = \frac{\sum_{\nu} n s_1^2}{(x_1 - x_2)^2 (\sum_{\nu} n)^2}$$

De werkschema's 1,II en 3,II geven een overzicht van de berekening van $\hat{\beta}$ en $s_{\hat{\beta}}^2$ uit een aantal ijkingen; iedere regel in de eerste vier kolommen van deze schema's heeft betrekking op één waarde van ν . De in de tabel genoteerde waarden zijn dus afkomstig van een aantal aan deze ijking voorafgaande ijkingen. Het werkschema 2,II wijkt van de bovengenoemde schema's alleen af door de toevoeging van een kolom voor het aantal dieren per groep in een ijking en dit aantal verminderd met 1; de totalen van deze kolommen zullen bij verdere berekeningen gebruikt worden. Bovendien ontbreekt de berekening van $s_{\hat{\beta}}^2$ welke door berekeningen in 2,III vervangen wordt.

De afleidingen van de schattingen $\hat{\beta}$ en $s_{\hat{\beta}}^2$ zullen we niet uitvoerig behandelen. $\hat{\beta}$ is de meest aannemelijke schatting voor β (in gevallen als deze komt dit overeen met de kleinste kwa-

dratenschatting), uitgaande van onderstellingen over de reacties, geformuleers in paragraaf 2.4.2 en vastgelegd in de formules (5) van die paragraaf.

4.3. Berekening van de sterkte van het praeparaat (III in de werkschema's).

In paragraaf 2.4.2 zijn de schattingen $\hat{\mu}$ en $\hat{\gamma}_{P_1}$ (het aantal eenheden insuline in de eerste praeparaat dosis) afgeleid ((8) en (9)). De werkschema's 1,III en 2,III geven deze onveranderd, werkschema 3,III geeft de schattingen bij ongelijke groepen.

De werkschema's 1,III en 3,III bepalen benaderde betrouwbaarheidsgrenzen voor de sterkte van het praeparaat met behulp van een schatting, $S_{\hat{\mu}}^2$, voor de variantie van $\hat{\mu}$ (zie paragraaf 2.4.2).

Werkschema 2,III geeft de berekening van exacte betrouwbaarheidsgrenzen voor γ_{P_1} . Hierbij moet $t_{f\alpha}$ opgezocht worden in een tabel van de verdeling van STUDENT met f vrijheidsgraden; $t_{f\alpha}$ is nl. die waarde waarvoor geldt:

$$P[|t| \geq t_{f\alpha}] = \alpha$$

wanneer t een STUDENT-verdeling met f vrijheidsgraden bezit. Tabellen van deze verdeling zijn o.a. te vinden in [A].

De betrouwbaarheidsgrenzen voor μ zijn afgeleid uit die voor $\frac{\hat{\mu}}{\hat{\sigma}_{\hat{\mu}}}$ (door er $\frac{1}{2}(\chi_1 + \chi_2)$ bij op te tellen), welke zijn:

$$\frac{A \hat{\beta}}{\hat{\beta}^2 - t_{\epsilon}^2 a_{22}} \pm \frac{\sqrt{A^2 \hat{\beta}^2 - (A^2 - t_{\epsilon}^2 a_{11})(\hat{\beta}^2 - t_{\epsilon}^2 a_{22})}}{\hat{\beta}^2 - t_{\epsilon}^2 a_{22}}$$

Voor de afleiding hiervan verwijzen we naar [20] en [21].

Opmerking. In de uitdrukking voor $S_{\hat{\mu}}^2$ in werkschema's 1,III en 3,III stelt $\frac{1}{4} s_2^2$ een schatting voor de variantie van A voor. Deze schatting kan zonder bezwaar vervangen worden door een simultane schatting uit het gehele materiaal (dus met gebruikmaking van de variantie-schattingen uit de voorafgaande ijkingen). Dit zal de nauwkeurigheid van de schatting $S_{\hat{\mu}}^2$ vergroten. In verband met de complicaties die hieruit voor het werkschema optreden is dit achterwege gelaten. Indien het in de praktijk wenselijk zou blijken deze verfijning aan te brengen kan dit alsnog geschieden.

4.4. Grafische controle van de helling van de ijklijn (IV van de werkschema's).

Uit iedere ijking apart is een schatting voor de helling

van de ijklijn te verkrijgen, nl.:

$$\hat{\beta}' = \frac{B}{X}$$

en voor de spreiding van $\hat{\beta}'$:

$$S_{\hat{\beta}'} = \sqrt{\frac{S_1^2}{X^2}}$$

Door deze waarden in een grafiekje te zetten tegen de datum van de ijking, krijgt men een indruk van de fluctuaties van helling en spreiding in de loop van een serie ijkingsen en van de grootte van de seizoeninvloeden.

Het zo verkregen inzicht kan belangrijk zijn bij het bepalen van het aantal ijkingsen dat voor de schatting van de helling gebruikt zal worden.

Behalve voor het bepalen van dit aantal heeft deze grafische controle nog een ander, directer, nut. Plotselinge afwijkende waarden van $\hat{\beta}'$ en/of $S_{\hat{\beta}'}$ kunnen nl. een aanwijzing zijn voor fouten bij de bepaling (dit kunnen zijn: afwijkend gedrag van de konijnen, fouten bij de uitvoering van de ijking, fouten bij de berekening, enz.). Grenzen voor $\hat{\beta}'$ en $S_{\hat{\beta}'}$, waarbuiten een nader onderzoek van de ijking gewenst is, zullen voornamelijk experimenteel bepaald moeten worden. Leidraad hierbij kan het volgende zijn (bij de werkschema's 1, IV en 2, IV):

$\frac{\hat{\beta}' - \beta}{S_{\hat{\beta}'}}$ heeft een STUDENT-verdeling met $4(n-1)$ vrijheidsgraden (zie [A] of [B]); hierin is β onbekend, als schatting hiervoor zou de helling $\hat{\beta}$ gebruikt kunnen worden, die bij de berekening van de praeparaatsterkte gebruikt wordt (dit verstoort de STUDENT-verdeling natuurlijk enigszins). De grenzen voor $\hat{\beta}'$, waarbinnen 95% van de gevonden waarden dient te liggen, worden dan benaderd door:

$$\hat{\beta} - t_{0.05} S_{\hat{\beta}} \quad \text{en} \quad \hat{\beta} + t_{0.05} S_{\hat{\beta}}$$

waarin $t_{0.05}$ bepaald is door:

$$P[|t| \geq t_{0.05}] = 0,05$$

als t een STUDENT-verdeling met $4(n-1)$ vrijheidsgraden bezit. Als benadering kan gelden $t_{0.05} = 2,02$ (dit geldt bij $n = 11$).

$\frac{\sum_{i=1}^m (n_i - 1) S_{\hat{\beta}_i}^2}{(n-1) S_{\hat{\beta}}^2}$ heeft bij benadering een F-verdeling met $4(n-1)$ en $4 \sum_{i=1}^m (n_i - 1)$ vrijheidsgraden (zie [A] of [B]), waarin n de aantallen dieren per groep zijn voor de ijkingsen waaruit

$\hat{\beta}$ bepaald is en n het aantal dieren per groep van de ijking waarvan we $\hat{\beta}'$ bekijken. Als grenzen voor $S_{\hat{\beta}'}^2$ kunnen genomen worden:

$$\frac{n-1}{\sum (n_i-1)} \cdot S_{\hat{\beta}}^2 F_{0,975} \quad \text{en} \quad \frac{n-1}{\sum (n_i-1)} \cdot S_{\hat{\beta}}^2 F_{0,025}.$$

Bij het werkschema 3, IV gelden dezelfde beschouwingen als bij werkschema 1, IV en 2, IV; met gewijzigde aantallen vrijheidsgraden voor bovengenoemde verdelingen: de STUDENT-verdeling bezit nu $(n_1 + n_2 + n_3 + n_4 - 4)$ vrijheidsgraden en de F -verdeling $(n_1 + n_2 + n_3 + n_4 - 4)$ en $\sum (n_i - 1)$ vrijheidsgraden.

4.5. Voorstel in verband met de gevoeligheid der konijnen.

Het systematische verschil in gevoeligheid van de konijnen gedurende enige tijd (zie paragraaf 3.4) kan wellicht als volgt gebruikt worden voor het homogeniseren van de stal en het vergroten van de nauwkeurigheid van de ijkingen. Heeft een konijn bij een ijking zowel op de standaardsdosis als op de praeparatosis de grootste of kleinste reactie van zijn groep gegeven, dan zetten we dit konijn voor volgende ijkingen op non-actief en vullen de groep aan met een nieuw konijn willekeurig gekozen uit een reservegroep. Komt een dergelijke situatie te weinig voor, dan kan het voorschrift ook zo gemaakt worden dat een konijn vervangen wordt zodra zijn reacties op standaard en praeparat tot de hoogste twee of tot de laagste twee behoren. Op deze manier kan een geregelde verversing van de ijkingsgroepen verkregen worden, terwijl konijnen die herhaalde malen wegens hoge of lage reacties vervangen moesten worden, verder totaal kunnen worden uitgesloten. De toepasbaarheid van dit voorstel hangt in sterke mate af van de tijd gedurende welke een konijn voor ijkingen te gebruiken is en van de frequentie van dit gebruik en kan dan ook slecht door ons beoordeeld worden. De uitwerking van deze maatregel valt niet met zekerheid te voorspellen, maar zal in de praktijk moeten blijken.

Werkschema 1. Routine-ijkingen.

Datum: Standaarddos $\gamma_S = \dots$.Einsuline.

Praeparaatdoses $p_1 = \dots$ cc en $p_2 = \dots$ cc (vaste verhouding $\frac{p_1}{p_2}$)

I	groep konij- nen	reacties		reactie verschil- len	groeps- gemid- delden	
		praep.	stand.			
	Q_1	$p_{11} -$ $\vdots$ $p_{1n} -$	$s_{11} =$ $\vdots$ $s_{1n} =$	y_{11} $\vdots$ y_{1n}	$\bar{y}_1$	$x_1 = \log p_1 = \dots$ $x_2 = \log p_2 = \dots$
	Q_2	$p_{21} -$ $\vdots$ $p_{2n} -$	$s_{21} =$ $\vdots$ $s_{2n} =$	y_{21} $\vdots$ y_{2n}	$\bar{y}_2$	$A = \frac{1}{4}(\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4) = \dots$ $B = \frac{1}{2}n(\bar{y}_1 - \bar{y}_2 + \bar{y}_3 - \bar{y}_4) = \dots$
	Q_3	$p_{31} -$ $\vdots$ $p_{3n} -$	$s_{31} =$ $\vdots$ $s_{3n} =$	y_{31} $\vdots$ y_{3n}	$\bar{y}_3$	$S_1^2 = \frac{n}{4(n-1)} \sum_{j=1}^4 \sum_{i=1}^n (y_{ji} - \bar{y}_j)^2 = \dots$
	Q_4	$p_{41} -$ $\vdots$ $p_{4n} -$	$s_{41} =$ $\vdots$ $s_{4n} =$	y_{41} $\vdots$ y_{4n}	$\bar{y}_4$	$S_2^2 = \frac{S_1^2}{n^2} = \dots$ $X = n(x_1 - x_2) = \dots$

II	Berekening $\hat{\beta}$ en $S_{\hat{\beta}}^2$ uit <u>voorafgaande</u> ijkingen			
	datum ijking	B	X	S_1^2
	$\dots$	$\dots$	$\dots$	$\dots$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$\dots$	$\dots$	$\dots$	$\dots$
	Totalen	T = ...	C = ...	V = ...

$\hat{\beta} = \frac{T}{C} = \dots$
 $S_{\hat{\beta}}^2 = \frac{V}{C^2} = \dots$

III	Berekening sterkte praeparaat	
	$\hat{\mu} = \frac{A}{\hat{\beta}} - \frac{1}{2}(x_1 + x_2) = \dots$ $S_{\hat{\mu}}^2 = \sqrt{\frac{S_1^2}{4\hat{\beta}^2} + \frac{A^2}{\hat{\beta}^4} S_{\hat{\beta}}^2} = \dots$ Aantal eenheden insuline in praeparaatdos p_1 : $\hat{\gamma}_{p_1} = \gamma_S \cdot p_1 \cdot \text{antilog } \hat{\mu} = \dots$ Grenzen van betrouwbaarheidsinterval (Onbetrouw- baarheid ongeveer 0.05): $\frac{\hat{\gamma}_{p_1}}{\text{antilog}(1,96 S_{\hat{\mu}})} = \dots$ en $\hat{\gamma}_{p_1} \cdot \text{antilog}(1,96 S_{\hat{\mu}}) = \dots$	

In een grafiek uitzetten tegen de datum van de ijking, met vermelding van de waarde van $4(n-1)$:

$\hat{\beta}' = \frac{B}{X} = \dots$

en $S_{\hat{\beta}'}^2 = \sqrt{\frac{S_1^2}{X}} = \dots$

Werkschema 2. Routine-ijkingen met berekening van exacte betrouwbaarheidsgrenzen voor de sterkte van het praeparaat.

Datum: Standaarddosis $\gamma_s = \dots$ E insuline

Praeparaatdoses $p_1 = \dots$ cc en $p_2 = \dots$ cc (vaste verhouding $\frac{p_1}{p_2}$)

I	groep konijnen	reacties		reactie verschillen	groeps-gemiddelden	
		praep.	stand.			
	G_1	$p_{11} -$. . . $p_{1n} -$	$S_{11} =$. . . $S_{1n} =$	y_{11} . . . y_{1n}	$\bar{y}_1$	$x_1 = \log p_1 = \dots; x_2 = \log p_2 = \dots$ $A = \frac{1}{4}(\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4) = \dots$
	G_2	$p_{21} -$. . . $p_{2n} -$	$S_{21} =$. . . $S_{2n} =$	y_{21} . . . y_{2n}	$\bar{y}_2$	$B = \frac{1}{2}n(\bar{y}_1 - \bar{y}_2 + \bar{y}_3 - \bar{y}_4) = \dots$ $S_1^2 = \frac{n}{4(n-1)} \sum_{j=1}^4 \sum_{i=1}^n (y_{ji} - \bar{y}_j)^2 = \dots$
	G_3	$p_{31} -$. . . $p_{3n} -$	$S_{31} =$. . . $S_{3n} =$	y_{31} . . . y_{3n}	$\bar{y}_3$	$S_3^2 = \sum_{j=1}^4 \sum_{i=1}^n (y_{ji} - \bar{y}_j)^2 = \dots$
	G_4	$p_{41} -$. . . $p_{4n} -$	$S_{41} =$. . . $S_{4n} =$	y_{41} . . . y_{4n}	$\bar{y}_4$	$X = n(x_1 - x_2) = \dots$

II Berekening $\hat{\beta}$ uit voorafgaande ijkingen

datum ijking	n	$n-1$	B	X	S_3^2	$\hat{\beta} = \frac{T}{C} = \dots$
...	...	...	...	...	...	
...	...	...	...	...	...	
...	...	...	...	...	...	
Totalen	$N = \dots$	$M = \dots$	$T = \dots$	$C = \dots$	$S_4^2 = \dots$	

III Berekening sterkte praeparaat

$$\hat{\mu} = \frac{A}{\hat{\beta}} - \frac{1}{2}(x_1 + x_2) = \dots \quad \hat{\beta} = 4(n-1+M) = \dots \quad t_{\hat{\beta}, \alpha} = \dots$$

$$S_5^2 = S_4^2 + S_3^2 = \dots$$

$$a_{11} = \frac{1}{\hat{\beta}} \cdot \frac{1}{4n} \cdot S_5^2 = \dots \quad F = A^2 - t_{\hat{\beta}, \alpha}^2 \cdot a_{11} = \dots$$

$$a_{22} = \frac{1}{\hat{\beta}} \cdot \frac{N}{C^2} \cdot S_5^2 = \dots \quad G = \hat{\beta}^2 - t_{\hat{\beta}, \alpha}^2 \cdot a_{22} = \dots \quad H = \frac{A\hat{\beta}}{G} = \dots$$

grenzen voor μ (onbetrouwbaarheid α)

$$\mu_1 = -\frac{1}{2}(x_1 + x_2) + H - \sqrt{H^2 - \frac{F}{G}} = \dots$$

$$\mu_2 = -\frac{1}{2}(x_1 + x_2) + H + \sqrt{H^2 - \frac{F}{G}} = \dots$$

Aantal eenheden insuline in praeparaatdosis p_1 :

$$\hat{\gamma}_{p_1} = \gamma_s \cdot p_1 \cdot \text{antilog } \hat{\mu} = \dots$$

grenzen van betrouwbaarheidsinterval (onbetrouwbaarheid α) voor het aantal eenheden insuline in p_1 :

$$\gamma_s \cdot p_1 \cdot \text{antilog } \mu_1 = \dots$$

$$\gamma_s \cdot p_1 \cdot \text{antilog } \mu_2 = \dots$$

IV In een grafiek uitzetten tegen de datum van de ijking, met vermelding van het aantal vrijheidsgraden $4(n-1)$ (eventueel ook $4M$):

$$\hat{\beta}' = \frac{B}{X} = \dots \quad \text{en} \quad S_{\hat{\beta}'}^2 = \sqrt{\frac{S_1^2}{X^2}} = \dots$$

Werkschema 3. Routine-ijkingen wanneer de groepen niet even groot zijn.

Datum: ... Standaarddosis $\gamma_s = \dots$ E insuline

Praeparaatdoses $p_1 = \dots$ cc en $p_2 = \dots$ cc (vaste verhouding $\frac{p_1}{p_2}$)

I	groep konijnen	reacties		reactie verschillen	aantal- len	groeps- gemidd.	groeps- varianties	
		praep.	stand.					
	G_1	p_{11} : : p_{1n_1}	$-s_{11}$: : $-s_{1n_1}$	$= y_{11}$: : $= y_{1n_1}$	n_1	$\bar{y}_1$	s_1^2	$x_1 = \log p_1 = \dots ; x_2 = \log p_2 = \dots$ $A = \frac{1}{2} \left(\frac{n_1 \bar{y}_1 + n_2 \bar{y}_2}{n_1 + n_2} + \frac{n_3 \bar{y}_3 + n_4 \bar{y}_4}{n_3 + n_4} \right) = \dots$
	G_{12}	p_{21} : : p_{2n_2}	$-s_{21}$: : $-s_{2n_2}$	$= y_{21}$: : $= y_{2n_2}$	n_2	$\bar{y}_2$	s_2^2	$B = \frac{n_1 n_2}{n_1 + n_2} (\bar{y}_1 - \bar{y}_2) + \frac{n_3 n_4}{n_3 + n_4} (\bar{y}_3 - \bar{y}_4) = \dots$
	G_{13}	p_{31} : : p_{3n_3}	$-s_{31}$: : $-s_{3n_3}$	$= y_{31}$: : $= y_{3n_3}$	n_3	$\bar{y}_3$	s_3^2	$S_1^2 = \left(\frac{n_1 n_2}{n_1 + n_2} \right)^2 \left(\frac{1}{n_1} s_1^2 + \frac{1}{n_2} s_2^2 \right) + \left(\frac{n_3 n_4}{n_3 + n_4} \right)^2 \left(\frac{s_3^2}{n_3} + \frac{s_4^2}{n_4} \right) = \dots$
	G_{14}	p_{41} : : p_{4n_4}	$-s_{41}$: : $-s_{4n_4}$	$= y_{41}$: : $= y_{4n_4}$	n_4	$\bar{y}_4$	s_4^2	$S_2^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{(n_1 + n_2)^2} + \frac{n_3 s_3^2 + n_4 s_4^2}{(n_3 + n_4)^2} = \dots$ $X = (x_1 - x_2) \left(\frac{n_1 n_2}{n_1 + n_2} + \frac{n_3 n_4}{n_3 + n_4} \right) = \dots$

II Berekening $\hat{\beta}$ en $s_{\hat{\beta}}^2$ uit voorafgaande ijkingen

datum ijking	B	X	S_1^2	$\hat{\beta} = \frac{T}{C} = \dots$ $s_{\hat{\beta}}^2 = \frac{V}{C^2} = \dots$
...	...	...	...	
...	...	...	...	
Totalen	$T = \dots$	$C = \dots$	$V = \dots$	

III Berekening sterkte praeparaat

$$\hat{\mu} = \frac{A}{\hat{\beta}} - \frac{1}{2} \left(\frac{n_1 x_1 + n_2 x_2}{n_1 + n_2} + \frac{n_3 x_1 + n_4 x_2}{n_3 + n_4} \right) = \dots$$

$$s_{\hat{\mu}} = \sqrt{\frac{S_1^2}{4\hat{\beta}^2} + \frac{A^2}{\hat{\beta}^4} s_{\hat{\beta}}^2} = \dots$$

Aantal eenheden insuline in praeparaatdosis p :

$$\hat{\gamma}_p = \gamma_s \cdot p_1 \cdot \text{antilog } \hat{\mu} = \dots$$

Grenzen van betrouwbaarheidsinterval (onbetrouwbaarheid ongeveer 0.05):

$$\frac{\hat{\gamma}_p}{\text{antilog}(1.96 s_{\hat{\mu}})} = \dots \quad \text{en} \quad \hat{\gamma}_p \cdot \text{antilog}(1.96 s_{\hat{\mu}}) = \dots$$

IV In een grafiek uitzetten tegen de datum met vermelding van de waarde van $(n_1 + n_2 + n_3 + n_4)$:

$$\hat{\beta}' = \frac{B}{X} = \dots \quad \text{en} \quad s_{\hat{\beta}'} = \sqrt{\frac{S_1^2}{X^2}} = \dots$$

5. Samenvatting.

Dit rapport over de gebruikelijke insulineijking en een eventuele verbetering hiervan bestaat uit drie gedeelten:

- 1) een literatuuroverzicht,
- 2) een onderzoek van het waarnemingsmateriaal dat ons ter beschikking was gesteld en
- 3) een uitwerking van voorstellen over een nieuw ijkingsschema.

1) Als voornaamste resultaten van het literatuuronderzoek (hoofdstuk 2) vonden we:

Zowel de percentuele bloedsuikerdaling als de absolute daling als maat voor de reactie van een konijn op een insuline-inspuiting bleken gecorreleerd met het beginbloedsuikergehalte; de in de literatuur opgegeven waarden voor deze correlatie lopen echter sterk uiteen. Bij een correctie voor deze correlatie is bij routine-ijkingen een regressiecoëfficiënt 0,22 gebruikelijk. Latere onderzoekers melden, dat het absolute niveau op één tijdstip na inspuiting als reactiemaat minder afhankelijk van het begingehalte zou zijn (paragraaf 2.1).

Ook de helling van de ijklijn (het verband tussen reactie en dosis, deze moet gebruikt worden om uit de reacties af te leiden hoe groot de dosis van het onbekende praeparaat was) is door verschillende onderzoekers bepaald. MARKS e.a. vinden als helling voor hun ijklijn (verband tussen reactieverhouding en logaritmie van dosesverhouding) $t = 0,66$; dit is een vaak bij routine-ijkingen gebruikte waarde van de helling (paragraaf 2.4.1). Door FIELLER e.a. werd het verband tussen reactieverschillen en de logaritmie van de dosesverhouding onderzocht (paragraaf 2.4.2). De tweede methode heeft op de eerste voor dat ze een betere nauwkeurigheidsbepaling toelaat.

FIELLER waarschuwt tegen het gebruik van een vaste éénmaal geschatte ijklijn, daar hierdoor systematische afwijkingen in de ijkuitkomsten geïntroduceerd worden. Beter is het de ijklijn steeds opnieuw uit eigen materiaal te schatten.

Het gebruikelijke proefschema is de kruisproef of uitbreidingen hiervan (meer doses, meer dagen), dat een vergelijking tussen praeparaat en standaard onder zoveel mogelijk voor beide gelijke omstandigheden mogelijk maakt (paragraaf 2.3).

2) Bij de analyse van het waarnemingsmateriaal (hoofdstuk 3) bleek, dat ook hier afhankelijkheid tussen reactie (reactiematen: percentuele bloedsuikerdaling en absoluut niveau) en beginbloedsuikergehalte aangenomen moet worden; de resultaten geven echter geen aanleiding een constant verband te onderstellen (vermoedelijk grote seizoen- en daginvloeden). Correctie voor deze

afhankelijkheid met een constante regressiecoëfficiënt 0,22 is dan ook niet aan te bevelen.

Voor de helling van de ijklijn volgens MARKS werden sterk verschillende waarden gevonden, vermoedelijk o.a. als gevolg van seizoeninvloeden. Hieruit blijkt dus het belang van een ijklijn bepaling uit eigen materiaal. Een schatting van de helling van de ijklijn van FIELLER kon uit dit waarnemingsmateriaal niet gegeven worden.

Bij de berekeningen bleek nog dat er systematische gevoeligheidsverschillen tussen de konijnen bestaan, welke zich ondanks grote daginvloeden zeker enige tijd handhaven.

3) De voorstellen over een nieuw ijkschema (hoofdstuk 4) gaan uit van een tweevoudige kruisproef (twee praeparatodoses in plaats van één); hierdoor wrdt het mogelijk de helling van de ijklijn steeds opnieuw uit eigen materiaal te schatten. Om de nauwkeurigheid te vergroten zonder hiervoor het aantal konijnen per ijking te moeten uitbreiden, worden hiervoor een aantal onmiddellijk aan de pas uitgevoerde ijking voorafgaande ijkingen gebruikt. Als ijklijn wordt, als **bij FIELLER** e.a., het verband tussen reactieverschillen en logaritmen van dosesverhoudingen gebruikt, waardoor een goede nauwkeurigheidsbepaling mogelijk wordt.

In verband met de keus van een reactiemaat vestigen we nog eens de aandacht op het belang van verdere experimenten hierover, vooral wat betreft het absolute niveau op één tijdstip na inspuiting. De bepaling hiervan zal de ijking sterk bekorten en zou ook verdere voordelen bezitten (betere onafhankelijkheid van het beginbloedsuikergehalte, sneller herstel van de konijnen, zie hoofdstuk 2).

Verder wordt voorgesteld de systematische gevoeligheidsverschillen tussen de konijnen te gebruiken bij de selectie der konijnen, waardoor wellicht een homogener stal en daardoor nauwkeuriger ijkingen verkregen worden.

4) We willen er ten slotte nog op wijzen dan een verder experimenteel onderzoek naar de besproken punten zeer gewenst zou zijn.

Literatuur.

- [1] The biological standardisation of insulin,
Publ. League of Nations C.H. 398 (1926).
- [2] H.P.Marks, The biological assay of insulin praeparations
in comparison with a stable standard,
Brit. Med. J., 2 (1925), pp. 1102-1104.
- [3] K.Culhane, H.P.Marks, D.A. Scott, J. Trevan, Assay of
crystallin insulin, The Biometrical J., 23 (1929),
pp. 397-409.
- [4] H.P.Marks, A comparison of the rabbit and mouse methods
in the assay of highly active insulin praeparations,
Quart. J. of Pharmacy and Pharmacology, 5 (1932), pp.
255-273.
- [5] A.M.Hemmingsen, H.P.Marks, The correlation between the
bloodsugarfall and the initial bloodsugar in rabbits
injected with insulin,
Quart. J. of Pharmacy and Pharmacology, 5 (1932), pp.
245-254.
- [6] H.P.Marks, C.Pak, Evaluation of the new international
standard insulin by the rabbit and mouse methods of
assay,
Quart. Bull., Volkenb.uitg., 5 (1936), pp. 631-651.
- [7] R.W.Spanhoff, Enkele opmerkingen over de sterkte van het
international insuline-standaard-paeparaat; benevens
een vergelijking tussen de subcutane en intraveneuze
methode van inspuiten bij konijnen,
Pharm. Weekblad, 34 (1938), pp. 933-944.
- [8] S.E.de Jongh, E.Laqueur, Eichung von Insulin,
Sonderdruck aus Handbuch der biologischen Arbeits-
methoden, abt. V., Teil 3 B (1938), pp. 1475-1492.
- [9] J.H.Gaddum, The accuracy of a cross over test,
Quart. J. Pharm. Pharmacol., 12 (1939), pp. 60-65.
- [10] C.I.Bliss, H.P.Marks, The biological assay of insulin I
and II,
Quart. J. Pharm. Pharmacol., 12 (1939), pp. 82-110 en
182-205.
- [11] E.C.Fieller, J.O.Irwin, H.P.Marks, E.A.G.Shrimpton, The
response-dosage relation in the cross-over rabbit
test for insulin I and II,
Quart. J. Pharm. Pharmacol., 12 (1939), pp. 206-211
en 724-742.
- [12] E.C.Fieller, The biological standardisation of insulin,
Suppl. J. Roy. Stat. Soc., 7 (1940), pp. 1-53.

- [13] A.H.Lacey, The rabbit method of insulin assay,
Endocrinology, 29 (1941), pp. 866-876.
- [14] R.H.Thorp, The effect of exercise upon the response of
rabbits to insulin,
Quart. J. Pharm. Pharmacol., 17 (1944), pp. 75-88.
- [15] K.W.Smith, H.P.Marks, E.C.Fieller, W.A.Brown, An extended
cross-over design and its use in the insulin assay,
Quart. J. Pharm. Pharmacol., 17 (1944), pp. 108-117.
- [16] A.H.Lacey, Further observations on the rabbit method of
insulin assay,
Endocrinology, 39 (1946), pp. 334-
- [17] D.M.Young, R.G.Romans, Assays of insulin with one blood
sample per rabbit per test day,
Biometric, 4 (1948), pp. 122-
- [18] L.J.Pugsley, Sybil Rampton, Variables affecting the assay
of insulin,
Endocrinology, 42 (1948), pp. 31-37.
- [19] E.C.Fieller, A fundamental formula in the statistics of
biological assay and some applications,
Quart. J. of Pharm. Pharmacol, 17 (1944), pp. 117-
123.
- [20] G.Klerk-Grobbe, Een betrouwbaarheidsgebied voor het quo-
tient van de verwachtingen van twee stochastische
grootheden, die een simultane normale verdeling be-
zitten,
Rapport S 90 (M 36) van de Statistische Afdeling van
het Mathematisch Centrum te Amsterdam, 1952.
- [21] H.J.Prins, Een betrouwbaarheidsinterval voor het quotient
van twee verwachtingen (aanvullend rapport),
Rapport S 90 (M 36a) van de Statistische Afdeling van
het Mathematisch Centrum te Amsterdam, 1953.

Boeken.

- [A] C.W.Emmens, Principles of Biological Assay,
London, Chapman & Hall Ltd. 1948.
- [B] C.W.Emmens, Hormone Assay,
Academic Press Inc., Publishers, New York, 1950
(Hoofdstuk over insuline van K.L.Smith).

Frequentiegrafieken van de waarnemingsgroep IV
met aangepaste normale verdelingen

— reacties op $\frac{1}{2}$ E insuline
- - - reacties op 1 E insuline

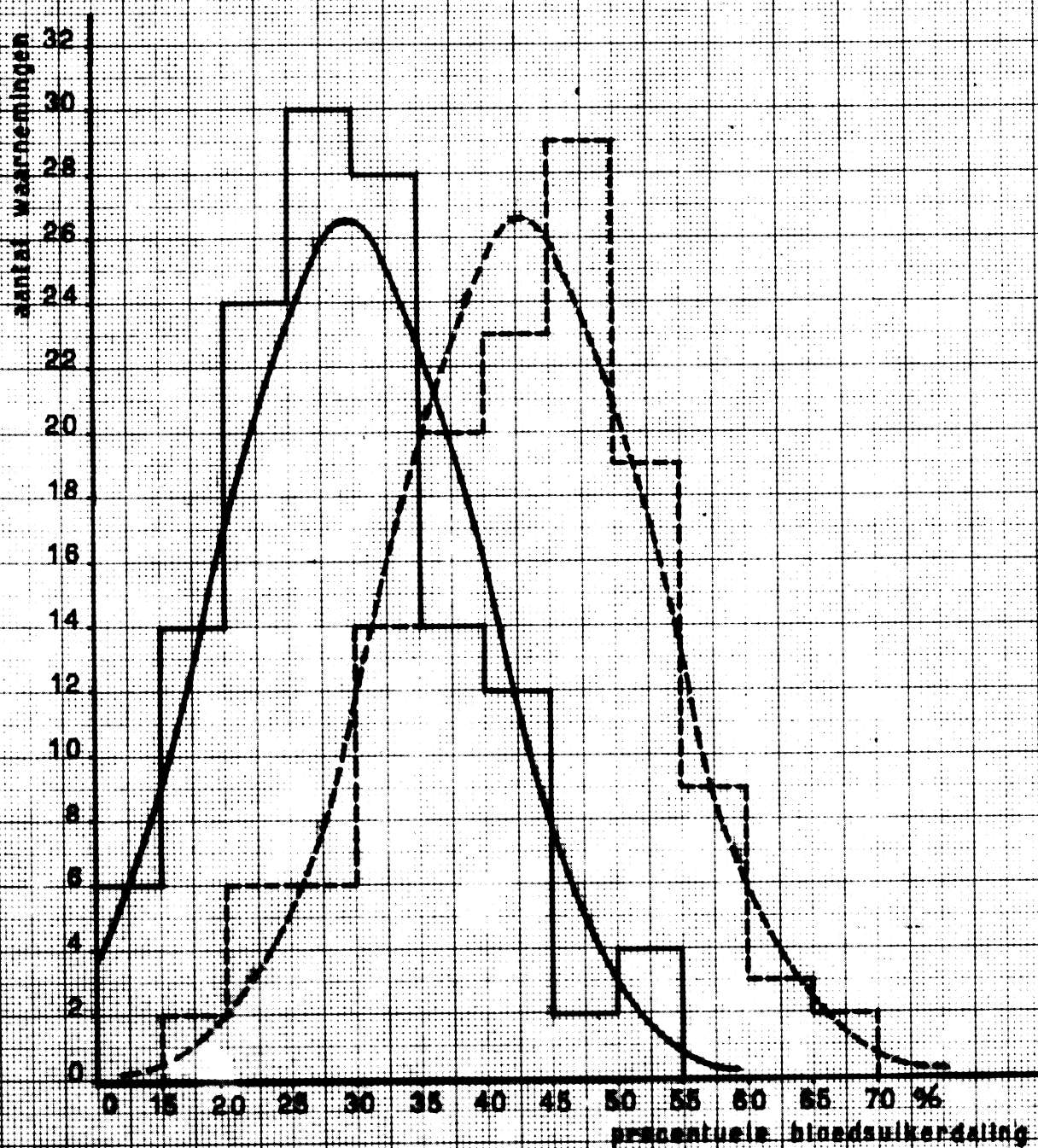


Fig. 1

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 grootheden²⁾, 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. Z heet de kritieke zone van de toets, α de onbetrouwbaarheidsdrempel (Engels: level of significance). Voor α neemt men veelal de waarde 0,05 of 0,01.

Men verworpt 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.

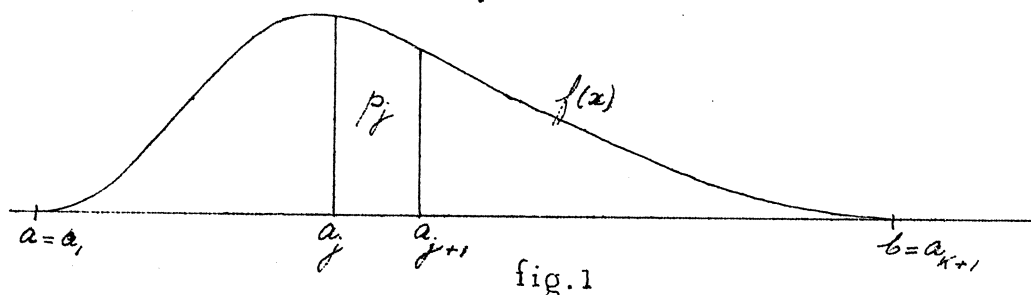
J.Hemelrijk en H.R. van der Vaart, Het gebruik van één- en tweezijdige overschrijdingskansen voor het toetsen van hypothesen, Statistica 4 (1950) p.54-66.

De χ^2 -toets voor aanpassing 1).

De χ^2 -toets voor aanpassing dient voor het toetsen van de hypothese $\mathcal{H}_0$, dat een stelsel waarden $x_1, x_2, \dots, x_n$ een steekproef is uit een continue waarschijnlijkheidsverdeling met verdelingsdichtheid $f(x)$. Hierbij wordt een grootte χ^2 waarvan de verdelingsfunctie bekend is, als toetsingsgrootte gebruikt.

De verdelingsfunctie van χ^2 wordt nog nader gekarakteriseerd door een grootte ν welke het aantal vrijheidsgraden van χ^2 aangeeft. De grootte van χ^2 en het getal ν worden op de volgende wijze bepaald:

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



wordt het interval $[a, b]$ in K stukken verdeeld, zodanig dat boven ieder interval $[a_j, a_{j+1}]$ een oppervlak p_j ligt ($j=1, \dots, K$). Daar het gehele oppervlak onder de kromme de totale waarschijnlijkheidsmassa voorstelt, dus gelijk aan 1 is, geldt de betrekking $p_1 + p_2 + \dots + p_K = 1$

Bestaat nu de steekproef uit n waarnemingen $x_1, x_2, \dots, x_n$ dan zal men onder de hypothese $\mathcal{H}_0$ in het algemeen verwachten, dat ongeveer np_j van de n steekproef-waarnemingen in het interval $[a_j, a_{j+1}]$ komen te liggen. Is het werkelijke aantal waarnemingen uit de steekproef, dat in het interval $[a_j, a_{j+1}]$ valt n_j (dus n_1 waarden in $[a_1, a_2]$ etc.) dan wordt χ^2 als volgt gedefinieerd:

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

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

Deze χ^2 is een maat voor de afwijking van de werkelijke verdelingsdichtheid van $\underline{x}$ van de onderstelde verdelingsdichtheid $f(x)$. Is de werkelijke verdelingsdichtheid van $\underline{x}$ inderdaad gelijk aan $f(x)$ zoals volgens $\mathcal{H}_0$ ondersteld wordt, dan zal χ^2 in het algemeen klein zijn; is de verdelingsdichtheid echter verschillend van $f(x)$, dan worden grote waarden van χ^2 waarschijnlijk dan het geval is, als $\mathcal{H}_0$ juist is. De kritieke zône voor deze toets geeft men daarom de gedaante $\chi^2 > \chi_0^2$, waarbij χ_0^2 correspondeert met een onbetrouwbaarheidsdrempel α en uit tabellen voor de χ^2 -verdeling opgezocht kan worden, waarbij rekening gehouden moet worden met het in de tabellen aangegeven:

Aantal vrijheidsgraden ν . Dit getal wordt als volgt bepaald: zijn er l parameters van $f(x)$ (b.v. het gemiddelde, de spreiding e.d.) uit de steekproef bepaald, dan is $\nu = k - l - 1$, waarin k het aantal intervallen $[a_j, a_{j+1}]$ (zie boven) is. Is $f(x)$ van tevoren geheel gespecificeerd, dan is dus $\nu = k - 1$.

Voorbeelden: bij toetsing op bovenstaande wijze van de hypothese, dat de steekproef $x_1, \dots, x_n$ uit een normale verdeling ("verdeling van Gauss") afkomstig is, worden gemiddelde en spreiding van de steekproef gebruikt om $f(x)$ te bepalen. Door gemiddelde en spreiding is de normale verdeling geheel bepaald. Dan is dus $\nu = k - 3$. Wenst men echter de hypothese te toetsen, dat de steekproef afkomstig is uit een volledig gegeven normale verdeling, waarbij dus gemiddelde en spreiding ook gegeven zijn, dan behoeven deze parameters niet uit de steekproef geschat te worden en is $\nu = k - 1$.

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

Litteratuur:

M.G.Kendall, The advanced Theory of Statistics
London 1947, Deel 1 hfdst. 12.

Rangcorrelatie¹⁾

1. Beschrijving van de methode.

De door M.G. Kendall ontwikkelde methode der rangcorrelatie is toepasbaar op de volgende situatie:

De stochastische grootheden x en y bezitten een simultane verdeling. Over deze verdeling zelf behoeft niets ondersteld te worden.

(x_i, y_i) ($i = 1, \dots, n$), zijn onafhankelijke waarnemingsparen van deze stochastische grootheden

Voorbeeld:

$i =$	1	2	3	4	5	6
x_i	0,11	0,12	0,10	0,11	0,15	0,13
y_i	3,4	3,0	3,2	3,5	3,5	3,5

Wij zeggen dat de waarnemingsparen (x_i, y_i) en (x_j, y_j) positief gecorreleerd zijn, als de volgorde van x_i en x_j hetzelfde is als die van y_i en y_j (bv. $x_i < x_j$ en $y_i < y_j$); zij zijn negatief gecorreleerd als de volgorde van x_i en x_j tegengesteld is aan de volgorde van y_i en y_j (bv. $x_i > x_j$ en $y_i < y_j$) en zij zijn niet gecorreleerd als $x_i = x_j$ of $y_i = y_j$.

In tabel 1 hebben wij van alle tweetallen (x_i, y_i) en (x_j, y_j) uit ons voorbeeld nagegaan of zij positief, negatief dan wel niet gecorreleerd zijn. Een positieve correlatie is aangeduid met +1, een negatieve met -1, terwijl het ontbreken van correlatie wordt aangegeven door een 0.

De toetsingsgrootheid van de methode van rangcorrelatie is nu het aantal positief gecorreleerde tweetallen verminderd met het aantal negatief gecorreleerde, of wel de som van de getallen, die in tabel 1 in de kolom "correlatie" voorkomen.

De verdeling van S voor het geval dat x en y onafhankelijk zijn is bekend (zie § 2). De hypothese dat x en y onafhankelijk

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

Tabel 1

Berekening van S
voor het voorbeeld

i	j	Correlatie
1	2	-1
1	3	+1
1	4	0
1	5	+1
1	6	+1
2	3	-1
2	4	-1
2	5	+1
2	6	+1
3	4	+1
3	5	+1
3	6	+1
4	5	0
4	6	0
5	6	0

$$S = +5$$

zijn, kan dus getoetst worden.

Is de hypothese van onafhankelijkheid niet vervuld, dan is de waarschijnlijkheid van grote positieve of grote negatieve waarden van S groter, dan wanneer dit wel het geval is. De kritieke zône is daarom van de vorm $|S| \geq S_0$, en bij ééNZijddige toetsing van de vorm $S \geq S'_0$ (rechtszijdige toetsing) of $S \leq S''_0$ (linkszijdige toetsing).

2. Verdeling van S als x en y onafhankelijk zijn.

Als er noch bij de x_i noch bij de y_j gelijke waarden voorkomen kunnen wij gebruik te maken van exacte tabellen, die voorkomen in [1] pg 141 ($n = 4$ t/m, 10) en in [2] (tables I and II, $n = 4$ t/m 40). Bovendien vindt men in [2] table III de kleinste waarden van $\underline{S}$, waarvan de overschrijdingskansen onder de hypothese van onafhankelijkheid hoogstens gelijk zijn aan α voor $\alpha = 0,005; 0,01; 0,025; 0,05$ en $0,10$ en $n = 4, 5, 6, \dots, 40$.

Als er bij de x_i óf bij de y_i , doch niet bij beide tweetallen of drietallen gelijken voorkomen, kan men voor $n \leq 10$ gebruik maken van de tabel van Sillitto [4].

Voor grote waarden van n is de verdeling van $\frac{S}{\sigma_S}$ (waarin

σ_S de spreiding van $\underline{S}$ is, die uit een hieronder op te geven formule berekend kan worden) bij benadering normaal met gemiddelde 0 en spreiding 1. Hiervan kunnen we gebruik maken om de hypothese van onafhankelijkheid te toetsen in de gevallen waar de exacte verdeling niet getabelleerd is. Dit geschiedt dan, door in een tabel van de normale verdeling de

overschrijdingskans op te zoeken, die behoort bij de gevonden waarde van $\frac{\sigma_{\underline{S}}}{\sigma_{\underline{S}}}$.

Om $\sigma_{\underline{S}}$ te berekenen, nemen wij in de rij der waarnemingen x_i de gelijke waarnemingen in groepen bij elkaar. De aantallen waarnemingen in die groepen duiden wij aan met t_h , waarin $h = 1, 2, \dots, k_1$. Evenzo doet men in de rij der waarnemingen y_j , waar we de overeenkomstige aantallen aanduiden met u_l , waarin $l = 1, 2, \dots, k_2$. $\sigma_{\underline{S}}$ kan dan gevonden worden uit de volgende formule:

$$\begin{aligned} (1) \quad \sigma_{\underline{S}}^2 &= \frac{1}{18} \left\{ n(n-1)(2n+5) - \sum_{h=1}^{k_1} t_h(t_h-1)(2t_h+5) - \right. \\ &- \sum_{l=1}^{k_2} u_l(u_l-1)(2u_l+5) \left. \right\} + \\ &+ \frac{1}{9n(n-1)(n-2)} \sum_{h=1}^{k_1} t_h(t_h-1)(t_h-2) \sum_{l=1}^{k_2} u_l(u_l-1)(u_l-2) + \\ &+ \frac{1}{2n(n-1)} \sum_{h=1}^{k_1} t_h(t_h-1) \sum_{l=1}^{k_2} u_l(u_l-1). \end{aligned}$$

In ons voorbeeld van § 1 komt in de rij x_i één tweetal gelijken (dus $k_1=1$ en $t_1=2$) en in de rij y_j één drietal gelijken ($k_2=1$, $u_1=3$) voor. Dus geldt:

$$\begin{aligned} t_1(t_1-1)(2t_1+5) &= 2.1.9 = 18 \\ u_1(u_1-1)(2u_1+5) &= 3.2.11 = 66 \\ t_1(t_1-1)(t_1-2) &= 0, (t_1-1)(t_1-2) = 0 \\ t_1(t_1-1) &= 2.1 = 2 \\ u_1(u_1-1) &= 3.2 = 6 \\ n(n-1)(2n+5) &= 6.5.17 = 510 \\ n(n-1) &= 6.5 = 30 \end{aligned}$$

zodat:

$$\sigma_{\underline{S}}^2 = \frac{1}{18} \{ 510 - 18 - 66 \} + \frac{1}{60} \times 2 \times 6 = 23,87$$

en $\sigma_{\underline{S}} = 4,89$ is.

Als alle t_h en alle u_l gelijk zijn aan 1 en er dus in geen van beide rijen gelijken voorkomen, gaat formule (2) over in:

$$(2) \quad \sigma_{\underline{S}} = \sqrt{\frac{1}{18} n(n-1)(2n+5)}$$

Een tabel van deze functie voor $n = 40, 41, \dots, 100$ vindt men in [2] (table IV).

3. Rangcorrelatiecoëfficiënt

Als maat voor de correlatie in de rij van waarnemingsparen $(x_1, y_1), \dots, (x_n, y_n)$ heeft Kendall de coëfficiënt τ gedefinieerd, die +1 is als de volgorden der waarnemingen in beide rijen $x_1, \dots, x_n$ en $y_1, \dots, y_n$ volledig overeenstemmen en -1 is, als deze volgorden volkomen tegengesteld zijn. De definitie van τ is:

$$(3) \quad \tau = \frac{2S}{\left\{ n(n-1) - \sum_{h=1}^{k_1} t_h(t_h-1) \right\}^{\frac{1}{2}} \left\{ n(n-1) - \sum_{j=1}^{k_2} u_j(u_j-1) \right\}^{\frac{1}{2}}}$$

Als er in geen van beide rijen gelijke waarnemingen voorkomen wordt deze formule:

$$(4) \quad \tau = \frac{2S}{n(n-1)}.$$

Literatuur:

- [1] M.G. Kendall. Rank correlation Methods London 1948, Hoofdstuk 1.
- [2] L. Kaarsemaker en A. van Wijngaarden. Tables for use in rank correlation. (1952)
Report R 73 of the Computation Department of the Mathematical Centre.
- [3] J. Hemelrijk. Kendall's rangcorrelatie-coëfficiënt. Hoofdstuk I der cursus "Parameter vrije Methoden" Rapport S 59 (1951) Mathematisch Centrum, blz. 1-17.
- [4] G.P. Sillitto. "The Distribution of Kendall's coefficient of rankcorrelation in rankings containing ties. Biometrika 34 (1947) p. 36-40.

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

Gegeven: de steekproef $x_1, x_2, \dots, x_n$ uit een normale collectie. Anders gezegd: $x_1, \dots, x_n$ zijn onafhankelijke waarnemingen van de stochastische grootte x , die normaal verdeeld is (de zgn. waarschijnlijkheidsverdeling van Gauss bezit)²⁾.

H_0 (te toetsen hypothese): Het gemiddelde van x bezit de waarde μ ; μ is hierin een gegeven getal, b.v.0.

Toetsingsgrootte :

$$t = \sqrt{n}(\bar{m} - \mu) / s'$$

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

Indien H_0 juist is, zullen dicht bij 0 gelegen waarden vaker voorkomen dan ver van 0 gelegen waarden. Is echter het gemiddelde van x verschillend van μ , dan zullen verder van 0 af liggende waarden vaker voorkomen dan indien H_0 juist is. Als kritieke zone Z kiest men daarom voor tweezijdige toetsing een gebied van de vorm

$$|t| \geq t_0$$

en voor éénzijdige toetsing

linker-toetsing

$$t \leq -t_1$$

rechter-toetsing

$$t \geq t_2$$

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

Litteratuur:

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

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

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

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

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