

Statistische Afdeling  
Rapport S54, door:  
Mej. G.Grobbe.

Februari 1951

Onderzoek van de wekwerking van  
Theophylline en 8 Chl. Theophylline.

1. Inleiding.

Gevraagd werd te onderzoeken of de wekwerking van Theophylline sterker is dan die van 8 Chl Theophylline.

Hiertoe werden ons de resultaten van 12 proeven, genummerd I - XII, elk op 40 muizen, verschaft. Bij elke proef waren alle muizen eerst met het slaapmiddel Nembrital ( $65 \text{ mgr/kg}$ ,  $0,2 \text{ cc/20 gr}$ ) ingespoten; na 20 minuten werd de helft met een wekmiddel, de andere helft met eenzelfde hoeveelheid van een onwerkzame injectiestof (Na Cl 0,9%) of van het andere wekmiddel ingespoten en de duur van de slaap gemeten.

Het gewicht van elke muis was opgegeven. Bij elke proef was de verdeling van de gewichten der muizen over beide groepen ongeveer gelijk gemaakt.

Ter vergelijking werd een proef met een bekend wekmiddel, Picrotoxine, ingevoegd.

De proeven I en V werden terzijde gelegd, daar bij deze proeven 13 resp. 15 muizen stierven na behandeling met een wekmiddel, hetgeen deze proeven voor de statistische analyse ongeschikt maakt, daar de wekwerking en niet de eventuele lethale werking van het middel moet worden onderzocht.

Uit de histogrammen van bijlage 1, waar de aantallen muizen tegen de slaaptijden werden uitgezet voor de controle-groepen, ziet men dat de verdeling der slaaptijden weinig op een normale verdeling lijkt. In verband hiermee werden slechts toetsingsmethoden toegepast, waarbij de onderstelling der normaliteit der waarschijnlijkheidsverdelingen niet nodig is.

**2. Onderzoek naar de werking der middelen.**  
(door vergelijking met een controle-serie).

Ter bepaling van de werking werd voor de proeven VI; VII; VIII; III; IV; en II de toets van Wilcoxon (zie S47 (M7)) gebruikt voor de hypothese dat de werkmiddelen geen andere invloed op de slaaptijden hebben dan de controlestof NaCl <sup>1)</sup>).

De resultaten zijn in de volgende tabel samengevat.

Tabel I

Proef	gemidd. gewicht	onderzochte werkmiddel	dosis	invloed op slaapduur	overschrijdingskans
♂ VI	18.40 gr	Theophylline	600 <sup>mg</sup> /kgr	korter	0,011
♂ VII	19.65	Theophylline	300 <sup>mg</sup> /kgr	"	0,65
♂ VIII	19.65	Picrotoxine	-	"	0,005
♀ III	27.27	Theophylline	600 <sup>mg</sup> /kgr	"	0,022
♀ IV	18.67	Theophylline	600 <sup>mg</sup> /kgr	"	0,19
♀ II	26.25	S Chl Theophylline	300 <sup>mg</sup> /kgr	langer	0,020

In tabel I geeft de eerste kolom het nummer van de proef en het geslacht der muizen, de tweede kolom het gemiddelde gewicht van de muizen, de derde kolom het werkmiddel dat met de controlestof NaCl vergeleken wordt, de vierde kolom de dosis van dit werkmiddel, de vijfde kolom de invloed van het werkmiddel op de slaapduur en de zesde kolom de overschrijdingskansen.

Men ziet uit de tabel, dat slechts bij S Chl Theophylline een verlenging van de slaapduur optrad. Uit de grootte der overschrijdingskansen mogen wij concluderen dat Theophylline 600 mg/kg en Picrotoxine inderdaad een werking vertonen, die bij de picrotoxine duidelijker is dan bij de Theophylline. Bij Theophylline 300 mg/kg was de verkorting der slaapduur niet significant meer, hetgeen misschien bij een uitgebreider proefmateriaal wel het geval zou zijn geweest.

<sup>1)</sup> Zie ook S17 (M6) voor de algemene gang van zaken bij het toetsen van hypothesen.

Daarentegen heeft 8 Chl Theophylline een verlengende invloed op de slaapduur in vergelijking met de controle-injectie en moet dus, althans in de omstandigheden, waarin de proef werd verricht, eerder als een slaapmiddel dan als een wekmiddel worden beschouwd.

### 3. Vergelijking van Theophylline en 8 Chl Theophylline

Om de werking van Theophylline en 8 Chl Theophylline onderling te vergelijken werd op de proeven X; XII; XI en IX eveneens de toets van Wilcoxon toegepast.

De resultaten van deze toetsing zijn in tabel II te vinden.

Tabel II

Proef	gemidd. gewicht	dosis van beide wekmiddelen	kortste slaaptijd trad op bij	overschrijdingkans
X ♂	20.27 gr	300 <sup>mg</sup> /kg	Theophylline	0,049 ←
XII ♂	17.15	300 <sup>mg</sup> /kg	Theophylline	0,25
XI ♀	18.32	300 <sup>mg</sup> /kg	Theophylline	← 0,0026 ←
IX ♀	18.20	200 <sup>mg</sup> /kg	Theophylline	0,34

Conclusie: Theophylline geeft in dosis 300<sup>mg</sup>/kg bij mannelijke en vrouwelijke dieren een kortere slaapduur dan 8 Chl Theophylline. Ook bij de proeven XII en IX, waar geen significantverschil gevonden werd, wees het resultaat in de richting van een sterkere werking van Theophylline.

Dit resultaat stemt overeen met het in de vorige paragraaf beschrevene.

### 4. Invloed van het gewicht van de muizen.

Daar er bij vergelijking van proef X met XII (tabel II) en van proef III met IV (tabel I) een aanwijzing leek te bestaan voor een invloed van het gewicht der muizen op de werking der wekmiddelen, werd getracht deze invloed op te sporen.

Daar bij de proeven te weinig waarnemingen voorkwamen met dezelfde dosis van één wekmiddel toegepast op muizen, waarvan de gewichten een behoorlijke spreiding vertoonden, (alleen bij de proeven III + IV kon de werking van Theophylline/ 600<sup>mg</sup>/kg tussen ongeveer

20 zware en 20 lichte muizen vergeleken worden (tabel III laatste regel); bij de groepen X en XII was het gewichtsverschil te gering om effect te verwachten), werd onderzocht of deze invloed bij de controle groepen merkbaar zou zijn. Hiertoe werd weer de toets van Wilcoxon toegepast.

Tabel III geeft de resultaten van dit onderzoek.

Tabel III

Proef	gemidd. gewicht	idem zware m.	injec- tiestof m.	kortste slaap- tijd trad op bij	overschrij- dingskans
I+V+VI ♂	16.23	19.08	Contrôle	lichte muizen	0,42
VI ♂	16.56	19.09	Contrôle	zware "	0,44
I + V ♂	15.61	17.91	Contrôle	lichte "	0,19
VII+VIII ♂	18.40	20.90	Contrôle	zware "	0,70
II+III+IV ♀	19.57	28.00	Contrôle	lichte "	0,35
III + IV ♀	18.70	27.70	Theophylline	lichte ..	~0,90

Er zijn blijkens de grootte der overschrijdingskan- sen in dit materiaal geen aanwijzingen voor invloed van het gewicht op de wekwerkingen.

5. Vergelijking der proeven op verschillende dagen.

Tenslotte werd nog, om te onderzoeken of er mis- schien verschil zou bestaan tussen de muizen die in ver- schillende proeven, dus op verschillende dagen werd ge- bruikt, met de toets van Wilcoxon de reacties der muizen op de onwerkzame stof NaCO bij verschillende proeven vergeleken.

Tabel IV

Geslacht	proef tegen	proef	kortste slaaptijd werd gevonden bij	overadhridd. kans
♀	II	III	III	0,52
♀	II	IV	IV	0,22
♀	III	IV	IV	0,86
♂	VI	I + V	VI	0,072

Volgens de resultaten in tabel IV is voor 't aanne- men van een verschil geen reden.

### 6. Samenvatting.

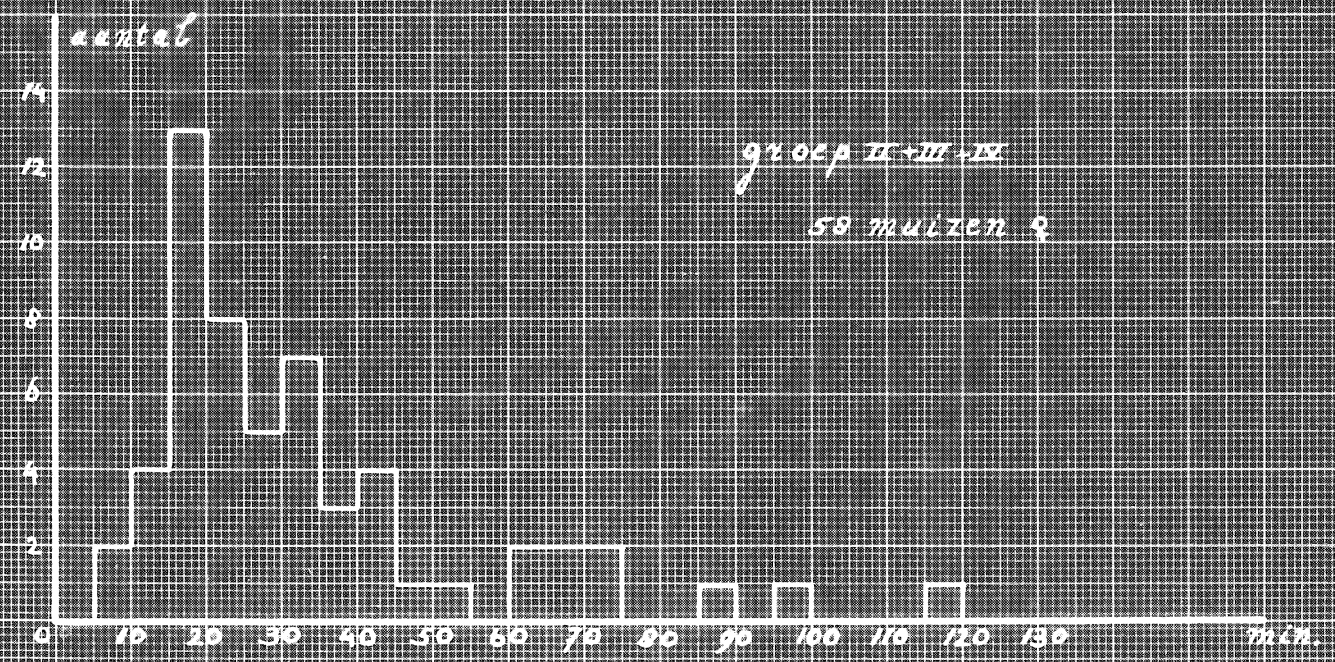
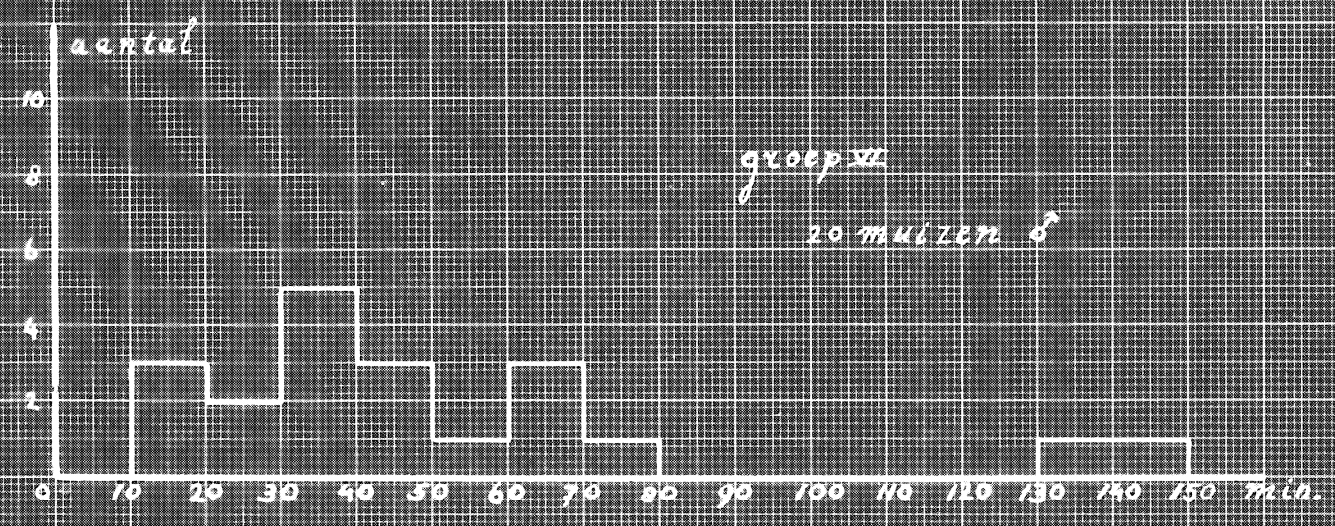
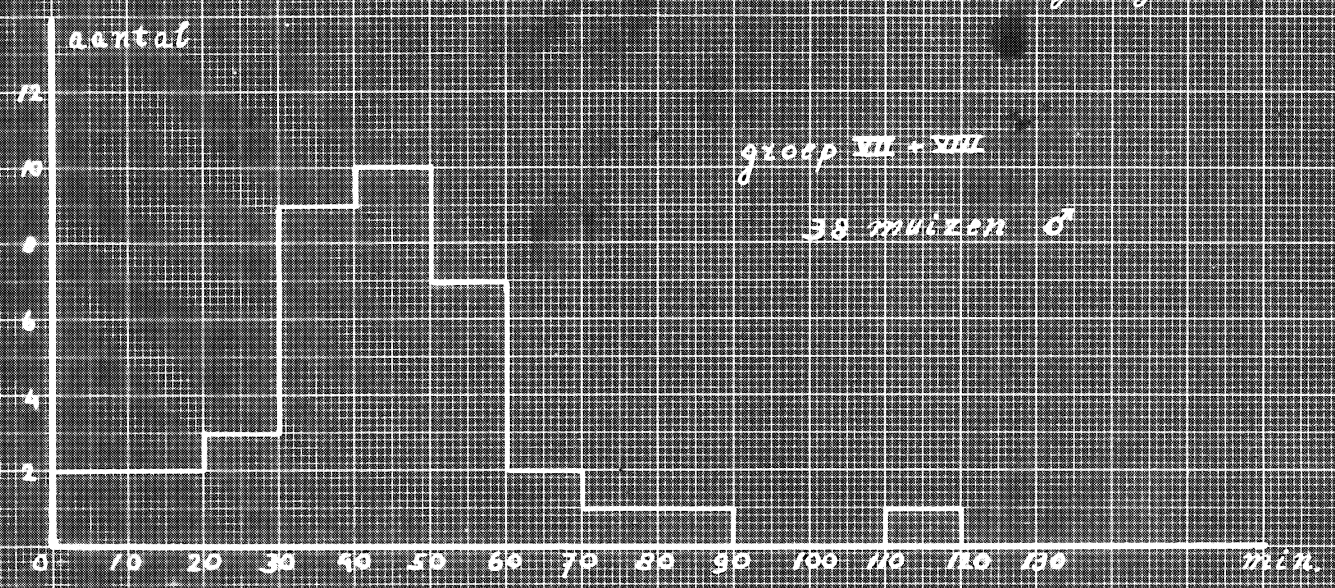
- 1) Theophylline met dosis 600 mg/kg vertoont een duidelijke wekwerking.
- 2) Van een dosis van 300 mg/kg Theophylline viel de werking niet meer vast te stellen. Dit is wellicht het gevolg van de kleinere dosis, waarvan de eventuele zwakkere werking met het geringe aantal verrichte waarnemingen niet kon worden geconstateerd.
- 3) Picrotoxine heeft een zeer duidelijke wekwerking.
- 4) In vergelijking met de controlestof NaCl werkte 8 Chl Theophylline (dosis 300 mg) verlengend op de slaapduur.
- 5) Theophylline in dosis 300 mg heeft een sterkere wekwerking dan 8 Chl Theophylline in dezelfde dosis.
- 6) Tussen 200 mg/kg Theophylline | kon geen verschil in wekwerking worden vastgesteld. Hiervoor geldt dezelfde opmerking als onder 2).
- 7) Afhankelijkheid tussen slaapduur, (wekwerking) en gewicht van de muizen is niet te constateren.
- 8) Voor verschillen tussen proeven van verschillende dagen is geen aanwijzing te vinden. Dit laatste pleit voor de betrouwbaarheid van de proefopzet.

*In dezelfde dosis  
8 Chl Theophylline*

### Opmerking

Het verdient misschien aanbeveling, teneinde te onderzoeken in hoeverre NaCl zelf enige werking heeft, de volgende proeven te verrichten:

- a) inspuiten met NaCl  $\rightarrow$  alleen een prik geven.
- b) Inspuiten met NaCl  $\leftrightarrow$  inspuiten met eigen bloed.
- c) Inspuiten met eigen bloed  $\rightarrow$  alleen een prik geven.



Histogrammen der slaaptijden.  
(controlgroepen).

Algemene gang van zaken bij het toetsen van een <sup>1)</sup>  
hypothese.

De toetsing van een hypothese  $H_0$  berust steeds op een aantal waarnemingen  $x_1, x_2, \dots, x_n$  van één of meer stochastische grootheden <sup>2)</sup>, 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  $\alpha$ , zodat  $Z$  dus van  $\alpha$  afhankelijk is. <sup>3)</sup>  $Z$  heet de kritieke zône van de toets,  $\alpha$  de onbetrouwbaarheidsdrempel (Engels: level of significance). Voor  $\alpha$  neemt men veelal de waarde 0,05 of 0,01.

Men verwierpt 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  $\alpha$  moet dan echter worden vermeld. De kans, dat dit zal gebeuren, is, indien  $H_0$  juist is, gelijk aan  $\alpha$ . Derhalve is  $\alpha$  de kans op ten onrechte verwerping van de juiste hypothese, ook de kans op een fout van de eerste soort genoemd. Indien men deze methode toepast, met  $\alpha = 0,05$  resp. 0,01, zal men in gemiddeld ongeveer één op 20 resp. op 100 van de gevallen, waar in 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  $\alpha$  wordt vaak bij de uitslag van een toetsing de overschrijdingskans  $k$  opgegeven; dit is de kleinste waarde van  $\alpha$ , waarbij in het betrokken geval, nog tot verwerping van  $H_0$  zou zijn overgegaan, anders gezegd: de kleinste  $\alpha$ , waarvoor de gevonden waarde der toetsingsgrootte nog juist in de (bij  $\alpha$  behorende) kritieke zone  $Z$  ligt. Wordt dus de waarde  $k$  opgegeven en werkt men met onbetrouwbaarheidsdrempel  $\alpha$ , dan wordt verworpen, indien  $k \leq \alpha$  is.

Voor het onderscheid tussen één- en tweezijdige toetsing en de keuze tussen deze twee mogelijkheden vergelijkte 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 toets van Wilcoxon.<sup>1)</sup>

Deze methode dient tot het toetsen van de hypothese  $H_0$ , inhoudend dat twee steekproeven  $x_1, \dots, x_n$  en  $y_1, \dots, y_m$  afkomstig zijn uit één collectie (ook populatie of universum genaamd). Zij is strict genomen, toepasbaar onder de voorwaarde, dat er geen enkel paar waarden  $(x_i, y_j)$  is met  $x_i = y_j$ . Verdere voorwaarden zijn voor de toepassing niet nodig, terwijl <sup>ook</sup> de zojuist genoemde, indien er niet teveel dergelijke paren zijn, de toepassing van de toets weinig hindert.

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

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

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

Verwerping van  $H_0$  ten gevolge van het vinden van een grote (resp. kleine) waarde van  $U$  wijst erop, dat  $x_1, \dots, x_n$  en  $y_1, \dots, y_m$  steekproeven uit verschillende collecties zijn, waarbij de op de  $x$ -collectie aangenomen waarden systematisch groter (resp. kleiner) dan de op de  $y$ -collectie aangenomen waarden zijn.

Litteratuur:

- F. Wilcoxon, Individual comparisons by ranking methods, Biometrics 1 (1945), p.80-83.
- H.B. Mann and D.R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Ann.Math.Stat. 18 (1947), p.50-60. Bevat tabellen voor  $n$  en  $m \leq 8$ .
- H.R. van der Vaart, Some remarks on the power function of Wilcoxon's test for the problem of two samples, Proceedings van de Kon.Ned.Ak.v.Wet., 53 (1950), p.494-520.
- H.R. van der Vaart, Gebruiksaanwijzing voor de toets van Wilcoxon, met tabellen voor  $n$  en  $m \leq 10$ , Rapport S32 (M<sub>4</sub>)(1950).
- D. van Dantzig, Kadercursus Mathematische Statistiek, Math. Centrum, Amsterdam(1947-50), hoofdstuk 6, 3.

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