

Errata bij rapport S 176 (M 65)

pag.	regel	staat	moet staan
1	7 v.o.	waar	waar-
15	11 v.b.	(x)	(<u>x</u>)
15	14 v.b.	(y)	(<u>y</u>)
20	8 v.b.	als	als H ₀
22	16 v.b.	(x)	(<u>x</u>)
22	16 v.b.	(y)	(<u>y</u>)

Op pag. 9, schema 4

staat

(1)	(2)	(3)
	waarnemingen van	
	y	x
1	1	7
3	1	12
11	7	15
26	8	21
38	4	13
53	11	11
70	6	11
82	6	5
90	2	2
97	5	3
102	n=51	m=100

moet staan

(1)	(2)	(3)
	waarnemingen van	
	y	x
1	1	7
3	1	12
11	7	15
26	8	21
38	4	13
53	11	11
70	6	11
82	6	5
90	2	2
97	5	3
102	n=51	m=100

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 176(M 65)

Handleiding voor de toets van WILCOXON

door

Ir Doralien Wabeke

en

Constance van Eeden.

Juli

1955

The Mathematical Centre at Amsterdam, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

Voor deze uitgave werd een financiële bijdrage
verstrekkt door de KWALITEITSDIENST VOOR DE INDUSTRIE.

<u>Inhoud</u>	pag.
1. Inleiding	1
2. Het waarnemingsmateriaal en de getoetste hypothese	1
3. Wanneer leidt de toets tot verwerping van H_0 ?	2
4. Berekening van de toetsingsgrootte W	3
5. De verdeling van W onder de getoetste hypothese	10
6. Kritieke zones en overschrijdingskansen (algemene beschouwingen)	11
7. Tabellen voor kritieke zones en overschrijdingskansen	13
8. Benadering voor kritieke zones en overschrijdingskansen	15
9. Gevallen, waarin de exacte tabellen en de benaderingen niet te gebruiken zijn	19
10. Enige opmerkingen over de keuze tussen één- en tweezijdig toetsen	20
11. Het onderscheidingsvermogen van de toets van WILCOXON vergeleken met dat van de toets van STUDENT	23
12. Litteratuur	24

1. Inleiding

De toets van WILCOXON is de meest gebruikte "verdelingsvrije" toets voor het probleem van twee steekproeven. "Verdelingsvrij" wil zeggen, dat geen onderstellingen gemaakt worden over de vorm van de verdelingen, die aan de twee te vergelijken steekproeven ten grondslag liggen. Dit is het belangrijkste verschil tussen deze toets en de toets van STUDENT voor twee steekproeven, daar de laatstgenoemde toets alleen van toepassing is op steekproeven uit normale verdelingen met gelijke spreidingen, hetgeen de toepasbaarheid uiteraard aanzienlijk beperkt. Daar het wenselijk leek voor een zo belangrijke, maar relatief nog weinig bekende, toets als die van WILCOXON over een duidelijke, ook voor niet statistici bruikbare, handleiding te beschikken, werd in 1950 een door H.R. VAN DER VAART geschreven "Gebbruiksaanwijzing voor de toets van Wilcoxon" uitgegeven (zie [8]). Intussen is de toets op verschillende punten verder ontwikkeld, waardoor behoefte aan een nieuwe versie van de handleiding ontstond. Dit leidde tot de uitgave van de huidige. Deze bevat bovendien uitgebreidere tabellen dan tot nu toe beschikbaar waren. De berekeningen hiervoor werden uitgevoerd door J. BONNINGA, Mevr. E.C. BOS-LEVENBACH, J. KOKKES, RINA KORSWAGEN, J. ROOSEN en J.J.F.M. SCHLICHTING.

Een globale kennis van de gang van zaken bij het toetsen van hypothesen wordt voorondersteld.

2. Het waarnemingsmateriaal en de getoetste hypothese

Het waarnemingsmateriaal bestaat uit twee onafhankelijke steekproeven

$$x_1, x_2, \dots, x_m \text{ en } y_1, y_2, \dots, y_n,$$

terwijl de te toetsen hypothese H_0 inhoudt, dat deze steekproeven uit dezelfde populatie afkomstig zijn.

Men kan dit ook als volgt formuleren: gegeven is, dat

x_1, x_2, \dots, x_m onafhankelijke waarnemingen zijn van een stochastische
-----grootheid x ¹⁾,

1) Een stochastische grootheid is een grootheid, die een waarschijnlijkheidsverdeling bezit, hetgeen gewoonlijk blijkt uit het feit, dat bij herhaalde waarneming onder dezelfde omstandigheden verschillende uitkomsten verkregen worden. Stochastische grootheden worden door onderstreepte letters aangegeven; dezelfde letters, niet onderstreept, worden vaak gebruikt voor waarden, die zij aan kunnen nemen of aangenomen hebben.

y_1, y_2, \dots, y_m onafhankelijke waarnemingen zijn van een stochastische grootheid y ,

terwijl tevens de x -waarnemingen onafhankelijk van de y -waarnemingen zijn. De getoetste hypothese H_0 houdt nu in, dat x en y dezelfde waarschijnlijkheidsverdeling bezitten.

Voorbeeld 1

Men wil de treksterkte van twee partijen metaaldraden vergelijken en men neemt daartoe uit partij A een steekproef van $m = 8$ draden en uit B een van $n = 6$ draden. Men bepaalt van deze draden de treksterkte en men verkrijgt de volgende uitkomsten:

A: 62,9 68,1 64,6 67,9 65,3 67,4 66,9 67,2;

B: 68,2 66,3 65,2 65,4 68,0 67,5.

Deze twee reeksen waarnemingen (die in eenzelfde, niet ter zake doende, eenheid zijn uitgedrukt) komen nu overeen met de beide steekproeven x_1, x_2, \dots, x_m resp. y_1, y_2, \dots, y_m . De getoetste hypothese kan men nu als volgt interpreteren: de draden van de beide partijen zijn gelijkwaardig wat hun treksterkte betreft.

Opmerking 1: Men kan ook de hypothese toetsen, dat de verdeling van y wel dezelfde vorm heeft als die van x , maar dat de eerstgenoemde over een bepaalde afstand d verschoven is. Dan hebben $x+d$ en y dezelfde verdeling, hetgeen men toetst door de toets toe te passen op $x_1+d, x_2+d, \dots, x_m+d$ en y_1, y_2, \dots, y_m . De verdere beschrijving van de toets geven wij voor $d=0$; zij geldt geheel analoog voor $d \neq 0$.

3. Wanneer leidt de toets tot verwerping van H_0 ?

De toets leidt met grote waarschijnlijkheid tot verwerping van H_0 , indien de kans, dat x een grotere waarde aanneemt dan y ongelijk is aan de kans, dat x een kleinere waarde aanneemt dan y (zie [1] en [7]). In formule:

$$P[x > y] \neq P[x < y].$$

In praktische bewoordingen kan men dit uitdrukken door te zeggen, dat x systematisch grotere resp. systematisch kleinere waarden aanneemt dan y , hetgeen echter niet betekent, dat x altijd grotere resp. kleinere waarden dan y aanneemt.

De kans op verwerping van H_0 neemt in dit geval toe met toenemende m en n en nadert tot 1, indien men deze aantallen beide onbegrensd laat toenemen.

Opmerking 2: Als H_0 juist is dan geldt:

$$(1) \quad P[x > y] = P[x < y].$$

Als echter $P[x > y] = P[x < y]$ dan behoeft H_0 niet vervuld te zijn: aan (1) is nl. b.v. voldaan als x en y beide symmetrisch ten opzichte van 0 verdeeld zijn met verschillende spreidingen.

4. Berekening van de toetsingsgrootte W

De toetsingsgrootte W is gelijk aan twee maal het aantal paren waarnemingen (x_i, y_i) , waarvoor $x_i > y_i$ is, vermeerderd met het aantal paren (x_k, y_k) waarvoor $x_k = y_k$ is.²⁾

Het is duidelijk, dat W een kleine waarde aan zal nemen, indien de x -waarden overwegend kleiner zijn dan de y -waarden, een grote, indien het omgekeerde het geval is en een intermediaire waarde, indien geen van deze beide situaties zich voordoet.

De berekening van W (de door W in een bepaald geval aangenomen waarde) kan op velerlei wijzen uitgevoerd worden. Eén der in de praktijk zeer handig gebleken methoden demonstrenen wij aan de hand van voorbeeld 1. De hieronder beschreven werkzaamheden zijn voor dat voorbeeld in schema 1 uitgevoerd. De beschreven werkwijze ondergaat enige kleine, later te vermelden wijzigingen, indien er gelijke waarnemingen zijn, hetgeen bij dit voorbeeld niet het geval is.

1. Rangschik de waarden y_1, y_2, \dots, y_n naar opklimmende grootte in de tweede kolom van een schema, beginnende op de tweede regel en telkens één regel overslaande.

2. Nummer de regels in kolom (1) van 0 tot $2n$. Het nummer in iedere regel is dan gelijk aan twee maal het aantal y -waarden in vorige regels, vermeerderd met het aantal y -waarden in de betrokken regel.

3. Noteer vervolgens de waarden x_1, x_2, \dots, x_m op de regels tussen die y -waarden, waar zij tussenin liggen.

2) Gewoonlijk wordt voor de toets van WILCOXON de toetsingsgrootte $u = \frac{1}{2} W$ gebruikt. Om gebroken getallen te vermijden, wordt hier W ingevoerd.

4. De bijdrage tot W van iedere regel waarop een x voorkomt, wordt in kolom (4) genoteerd; deze wordt verkregen door het aantal x -waarden in kolom (3) te vermenigvuldigen met het in dezelfde regel staande getal in kolom (1).

5. Door sommeren van de bedragen in kolom (4) krijgen wij de bij deze steekproeven behorende waarde van W .

Schema 1
Berekening van W bij voorbeeld 1

(1)	(2)	(3)	(4)
	waarnemingen van		bij- drage tot W
	y	x	
0		62,9 64,6	0
1	65,2		
2		65,3	2
3	65,4		
4			
5	66,3		
6		66,9 67,2 67,4	18
7	67,5		
8		67,9	8
9	68,0		
10		68,1	10
11	68,2		
12			
totaal	n=6	m=8	W=38

Voorbeeld 2

Door een wijziging in het productieproces kunnen weerstanden sneller gefabriceerd worden. Men heeft nu 2 partijen weerstanden (één volgens het oude (x) en één volgens het nieuwe procédé gefabriceerd (y)) en men wil onderzoeken of de elektrische weerstand voor de beide partijen verschillend is. Hiertoe neemt men van beide partijen een steekproef en men meet de weerstanden van de exemplaren daarvan. De resultaten zijn (in Ohm):

1e steekproef (x): 147, 150, 150, 153, 148, 144, 155, 151, 146, 151.
2e steekproef (y): 140, 145, 147, 147, 138, 144, 149, 150.

De te toetsen hypothese H_0 , kan nu als volgt geïnterpreteerd worden.

De twee partijen weerstanden zijn, wat hun weerstand betreft, gelijkwaardig.

De berekening van de toetsingsgrootte W , gaat, nu er slechts kleine groepjes van gelijke waarnemingen optreden, nagenoeg analoog aan voorbeeld 1 (zie schema 2).

1. Rangschik de waarden y_1, y_2, \dots, y_n naar opklimmende grootte in de tweede kolom van een schema, beginnende op de tweede regel en telkens een regel overslaand. Gelijke waarnemingen komen op dezelfde regel.

2. Noteer in kolom (3) de α -waarden, die niet gelijk zijn aan een y -waarde, op de regel tussen de y -waarden waar zij tussen in liggen, en noteer de α -waarden die wel gelijk zijn aan een y -waarde op dezelfde regel als die y -waarde.

3. De getallen in kolom (1) verkrijgt men nu door voor iedere regel twee maal het aantal y -waarden erboven te nemen, vermeerderd met het aantal y -waarden op dezelfde regel. Op de laatste regel in kolom (1) komt dan het getal $2n$ te staan (contrôle).

4. De bijdrage tot W van iedere regel waarop een α voorkomt, wordt in kolom (4) genoteerd. Deze wordt verkregen door het aantal α -waarnemingen in kolom (3) met het getal in kolom (1), dat op dezelfde regel staat, te vermenigvuldigen.

5. Door sommeren van de bedragen in kolom (4) krijgen wij de waarde van W .

6. In kolom (5) noteren wij de aantallen gelijke waarnemingen in iedere regel. Door sommatie moeten wij het totale aantal waarnemingen krijgen ($N = n + m$).

7. De derdemachten van de getallen in kolom (5) noteren wij in kolom (6); wij sommeren deze derdemachten en noemen hun som D .

De kolommen (5) en (6) hebben wij vooralsnog niet nodig doch in een der volgende paragrafen zal het nut van deze kolommen blijken.

Schema 2

Berekening van W bij voorbeeld 2

(1)	(2)	(3)	(4)	(5)	(6)
	waarnemingen van		bijdrage tot W	aantallen gelijken(t)	t ³
	y	x			
0					
1	138			1	1
2					
3	140			1	1
4					
5	144	144	5	2	8
6					
7	145			1	1
8		146	8	1	1
10	147 147	147	10	3	27
12		148	12	1	1
13	149			1	1
14					
15	150	150 150	30	3	27
16		151, 151, 153, 155	64	2, 1, 1	8, 1, 1
totaal	n=8	m=10	W=129	N=18	D=78

Voorbeeld 3

In dit voorbeeld stelt x de lengte van de mannen van een bepaalde bevolkingsgroep voor, en y de lengte van vrouwen van dezelfde groep (lengte in meters). Zowel van x als van y zijn een aantal waarnemingen verricht.

Waarnemingen

x : 1,72 1,68 1,77 1,73 1,68 1,59 1,73 1,81 1,74
 1,88 1,75 1,72 1,79 1,82 1,64

y : 1,56 1,65 1,62 1,69 1,74 1,74 1,65 1,65 1,74
 1,53 1,89 1,68 1,66 1,63 1,72 1,63 1,67 1,66
 1,83 1,61

Men wil nu onderzoeken of in deze bevolking de lengte van de mannen systematisch verschilt van de lengte van de vrouwen.

De nulhypothese H_0 luidt:

De mannen en de vrouwen van deze bevolkingsgroep verschillen niet systematisch in lengte.

De berekening van W gaat analoog aan voorbeeld 2, (zie schema 3).

Schema 3

Berekening van W bij voorbeeld 3

(1)	(2)	(3)	(4)	(5)	(6)
	waarnemingen van		bijdrage tot W	aantallen gelijken(t)	t^3
	y	x			
1	1,53			1	1
2					
3	1,56			1	1
4		1,59	4	1	1
5	1,61			1	1
6					
7	1,62			1	1
8					
8+2 = 10	1,63 1,63			2	8
12		1,64	12	1	1
12+3 = 15	1,65 1,65 1,65			3	27
18					
18+2 = 20	1,66 1,66			2	8
22					
23	1,67			1	1
24					
25	1,68	1,68 1,68	50	3	27
26					
27	1,69			1	1
28					
29	1,72	1,72 1,72	58	3	27
30		1,73 1,73	60	2	8
30+3 = 33	1,74 1,74 1,74	1,74	33	4	64
36		1,75 1,77 1,79 1,81 1,82	180	1, 1, 1, 1	1, 1, 1, 1, 1
37	1,83			1	1
38		1,88	38	1	1
39	1,89			1	1
40					
totaal	n=20	m=15	W=435	N=35	D=185

Voorbeeld 4

Gemeten is de tijd s in seconden nodig voor een bepaalde psychologische test. Het waarnemingsmateriaal bestaat uit: waarnemingen bij personen van 30-44 jaar (x_1, x_2, \dots, x_m) en waarnemingen bij personen van 45-60 jaar (y_1, y_2, \dots, y_n). Men wil onderzoeken of de tijdsduur afhangt van de leeftijd van de proefpersonen. De hypothese H_0 luidt nu: De tijdsduur is onafhankelijk van de leeftijd van de proefpersonen en verschilt dus niet systematisch voor de beide groepen. Het materiaal is geclassificeerd als aangegeven is in tabel 1.

Tabel 1

Tijdsduur (in seconden) van een psychologische test bij personen van 30-44 resp. 45-60 jaar.

klasse indeling voor s	Aantal waarnemingen	
	van x	van y
120 - 149	7	1
150 - 179	12	1
180 - 209	15	7
210 - 239	21	8
240 - 269	13	4
270 - 299	11	11
300 - 329	11	6
330 - 359	5	6
360 - 389	2	2
390 - 419	3	5

Uit deze tabel ziet men dus b.v. dat 7 personen tussen 30 en 44 jaar voor deze test een tijd nodig hadden die ligt tussen 120 en 149 seconden, etc.

Nu in de waarnemingen groepen van betrekkelijk veel gelijken voorkomen, gaat de bovenbeschreven methode ter berekening van W over in de volgende (zie schema 4):

1. Noteer in kolom (2) (zonder regels over te slaan) het aantal waarnemingen van y , dat in de verschillende klassen valt.
2. Noteer in kolom (3) het aantal waarnemingen van x in die klassen.

3. In kolom (1) noteert men nu op iedere regel twee maal het aantal y -waarnemingen, dat boven die regel staat, vermeerderd met éénmaal het aantal y -waarnemingen, dat op dezelfde regel staat.

Een snelle wijze van berekenen van deze aantallen is de volgende:

Op de eerste regel komt dus 1
 Op de tweede regel komt 1 + 1 + 1 = 3
 Op de derde regel komt 3 + 1 + 7 = 11
 Op de vierde regel komt 11 + 7 + 8 = 26
 enz.

In schema 4 geven de pijlen de optellingen aan. In de onderste regel moet twee maal het aantal y -waarnemingen komen te staan ($102 = 2 \times 51$).

4. Vermenigvuldig nu de getallen in kolom (1) en (3) met elkaar en noteer het product in kolom (4); tel de producten op. Dit geeft $W = 3482$.

Indien men over een rekenmachine beschikt, krijgt men W direct door accumulatief vermenigvuldigen, zonder dat dus voor iedere regel het resultaat van de vermenigvuldiging genoteerd behoeft te worden.

5. Noteer in kolom (5) en (6) de waarden van t en t^3 en bepaal de totalen $N = n + m$ en D .

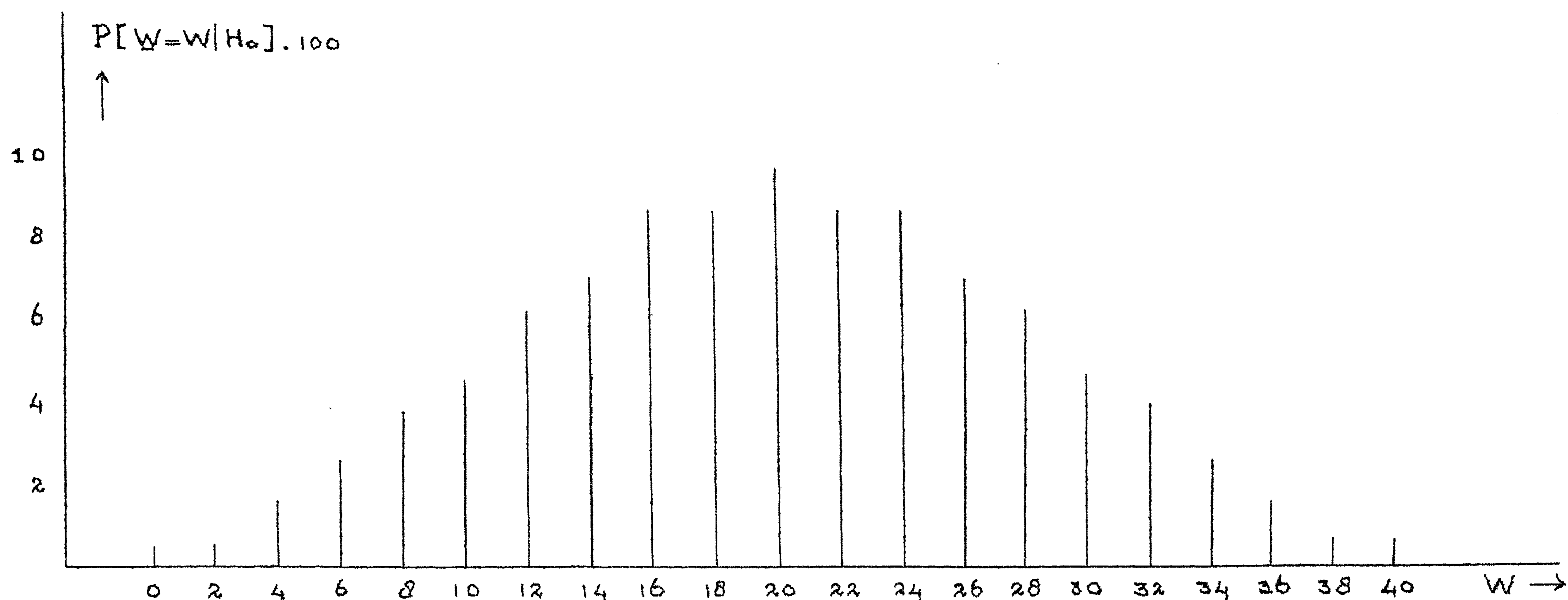
Schema 4

Berekening van W bij voorbeeld 4 (geclassificeerde waarnemingen).

	(1)	(2)	(3)	(4)	(5)	(6)
klasse- indeling van s		waarnemingen van		bijdrage tot W	aantal gelijkt	t^3
		y	x			
120-149	1	1	7	$1 \times 7 = 7$	8	512
150-179	3	1	12	$3 \times 12 = 36$	13	2197
180-209	11	7	15	$11 \times 15 = 165$	22	10648
210-239	26	8	21	$26 \times 21 = 546$	29	24389
240-269	38	4	13	$38 \times 13 = 494$	17	4913
270-299	53	11	11	$53 \times 11 = 583$	22	10648
300-329	70	6	11	$70 \times 11 = 770$	17	4913
330-359	82	6	5	$82 \times 5 = 410$	11	1331
360-389	90	2	2	$90 \times 2 = 180$	4	64
390-419	97	5	3	$97 \times 3 = 291$	8	512
	102	$n=51$	$m=100$	$W = 3482$	$N=151$	$D=60127$

5. De verdeling van \underline{W} onder de getoetste hypothese H_0 .

Indien de hypothese H_0 juist is, dus als de twee steekproeven uit dezelfde verdeling afkomstig zijn, kan de verdeling van de toetsingsgrootte \underline{W} berekend worden. Voor $m=4$ en $n=5$ b.v. en voor het geval er onder de waarnemingen geen gelijken voorkomen, ziet de verdeling van \underline{W} er uit als in de onderstaande figuur.



figuur 1

De verdeling van \underline{W} onder de hypothese H_0 voor $m=4$ en $n=5$ (geen gelijke waarnemingen).

Op de X-as zijn de waarden, die \underline{W} aan kan nemen, uitgezet en op de Y-as, de kans dat \underline{W} de betreffende waarde aanneemt als H_0 juist is. Bijvoorbeeld:

$$P[\underline{W} = 12 | H_0] = 0,06.$$

Dit betekent: de kans dat \underline{W} de waarde 12 aanneemt, indien H_0 juist is, is gelijk aan 0,06.

$$\begin{aligned} P[\underline{W} \geq 36 | H_0] &= P[\underline{W} = 36 | H_0] + P[\underline{W} = 38 | H_0] + P[\underline{W} = 40 | H_0] = \\ &= 0,0159 + 0,0079 + 0,0079 = 0,0317. \end{aligned}$$

Dit betekent: de kans dat \underline{W} een waarde 36 of groter aanneemt, als H_0 juist is, is gelijk aan 0,0317.

Eigenschappen van de verdeling van W indien H_0 juist is.

1. De verdeling van W is discreet, want W kan alleen gehele waarden aannemen.
2. De verdeling van W is in het algemeen niet symmetrisch, behalve,
 - a) als $m = n$,
 - b) als de omvang ³⁾ van de 1e groep gelijke waarnemingen gelijk is aan die van de laatste, de omvang van de tweede groep gelijk is aan die van de op één na laatste, enz. Stellen wij het aantal waarnemingen dat de kleinste waargenomen waarde bezit voor door t_1 , het aantal dat de op één na kleinste waarde bezit door t_2 , etc., en worden er in het totaal k verschillende waarden aangenomen dan is de verdeling dus o.a. symmetrisch als

$$t_1 = t_k, t_2 = t_{k-1} \quad \text{enz.}$$

3. De verwachting (het theoretisch gemiddelde) μ , van de verdeling van W is gelijk aan het product der steekproefgrootten $\mu = mn$.
4. De variantie van de verdeling van W wordt gegeven door:

$$\sigma^2 = \frac{mn(N^2 - D)}{3N(N-1)}.$$

Bevatten de beide steekproeven geen gelijke waarnemingen dan wordt $D = N$ en dan is dus

$$N^2 - D = N^2 - N = N(N-1)(N+1)$$

en in dit geval wordt de variantie

$$\sigma^2 = \frac{1}{3} mn(N+1) \quad (\text{geen gelijke waarnemingen}).$$

6. Kritieke zones en overschrijdingskansen (algemene beschouwingen).

Wij zagen reeds in par. 4 dat de toetsingsgrootheid W een kleine waarde aan zal nemen, indien de x -waarden overwegend kleiner zijn dan de y -waarden en een grote waarde indien het omgekeerde het geval is.

Wenst men nu de hypothese H_0 te toetsen tegen de alternatieve hypothese, inhoudende dat \underline{x} systematisch groter of kleiner is dan y , d.w.z. tegen de alternatieve hypothese

$$H: P[\underline{x} > y] \neq P[\underline{x} < y],$$

3) het aantal elementen.

dan kiest men een tweezijdige kritieke zone Z voor W , bestaande uit grote en kleine waarden van W . Wil men de hypothese H_0 alleen toetsen tegen de alternatieve hypothese dat \underline{x} systematisch groter is dan y , dus tegen de alternatieve hypothese

$$H_1: P[\underline{x} > y] > P[\underline{x} < y],$$

dan kiest men een rechtséénzijdige kritieke zone Z_r bestaande uit grote waarden van W . Een linkséénzijdige kritieke zone Z_l gebruikt men als men H_0 wil toetsen tegen de alternatieve hypothese

$$H_2: P[\underline{x} > y] < P[\underline{x} < y].$$

Wordt een waarde W gevonden, die in de gebruikte kritieke zone ligt, dan wordt H_0 verworpen. Ook indien H_0 juist is, kan dit gebeuren, hetgeen dan een ten onrechte verwerpen van H_0 inhoudt. De kans op dit ten onrechte verwerpen wordt de onbetrouwbaarheid van de kritieke zone (of ook: van de toets) genoemd en de kritieke zone wordt zo gekozen, dat deze onbetrouwbaarheid niet groter is dan een van de te voren gegeven grens α , de onbetrouwbaarheidsdrempel.

Wordt H_0 verworpen omdat W in een rechter-kritieke zone valt (dus "te groot" is), dan trekt men de conclusie, dat:

$$P[\underline{x} > y] > P[\underline{x} < y],$$

of, in praktische bewoordingen: dat \underline{x} systematisch hogere waarden aanneemt dan y . Bij een waarde in een linker-kritieke zone is dit juist andersom.

Het resultaat van een toets kan nog iets vollediger uitgedrukt worden door de overschrijdingskans op te geven.

Een algemene definitie van een overschrijdingskans luidt: de onbetrouwbaarheid van de kleinste kritieke zone, die de gevonden waarde van \underline{W} nog juist bevat. Daarbij worden dan ook kritieke zones met een onbetrouwbaarheid $> \alpha$ in de beschouwing betrokken.

Voor de rechts- (resp. links-) éénzijdige toets wordt de overschrijdingskans met k_r (resp. k_l) aangeduid en de definitie komt dan neer op: de kans op het optreden, onder de hypothese H_0 van een minstens even grote (resp. hoogstens even grote) waarde als de werkelijk gevondene.

In het tweezijdige geval wordt het symbool k gebruikt.

In vele praktijkgevallen kan men volstaan met voor k twee maal de kleinste éénzijdige overschrijdingskans van de gevonden waarde van \underline{W} te nemen. (vergelijk echter par. 9). Het vinden van een

overschrijdingskans α is equivalent met het vinden van een waarde van W in de kritieke zone.

Opmerking 3. De keuze tussen één- en tweezijdige toetsen wordt in par. 10 behandeld. Vooreerst bespreken wij hoe de kritieke zones en overschrijdingskansen, zowel één- als tweezijdig, bepaald kunnen worden.

7. Tabellen voor kritieke zones en overschrijdingskansen.

De tabellen voor overschrijdingskansen en kritieke zones, die aan dit rapport zijn toegevoegd (Tabel I en II, zie pag. 26 tot 37) gelden voor het geval er geen gelijke waarnemingen zijn. De verdeling van W onder de hypothese H_0 is dan symmetrisch (zie par. 5), zodat slechts de helft opgenomen behoeft te worden. Hiervoor is de linkerhelft gekozen.

Tabel I bevat voor $m \leq n \leq 10$ de exacte linkséénzijdige overschrijdingskansen van W . Is in een bepaald geval $m > n$, dan kunnen m en n in de tabel verwisseld gedacht worden. De tabel blijft dan gelden. Voor een waarde van W die $> mn$ is, wordt de rechteroverschrijdingskans gevonden door in de tabel de linkeroverschrijdingskans van $2mn - W$ te zoeken. De tweezijdige overschrijdingskans is exact gelijk aan twee maal de kleinste éénzijdige overschrijdingskans.

Tabel II geldt voor $m+n \leq 40$, $n \geq 11$ en $m \leq n$. Hierin staan echter niet de overschrijdingskansen, doch alleen de linker kritieke waarden W_ℓ van W , behorende bij drie éénzijdige onbetrouwbaarheidsdrempels nl.: $\alpha = 0,01$; $0,025$ en $0,05$.⁴⁾ Dit betekent, dat de linker kritieke zone Z_ℓ met de gekozen onbetrouwbaarheidsdrempel bestaat uit alle waarden van $W \leq W_\ell$, als W_ℓ het in de tabel opgezochte getal is. De rechter kritieke waarde, behorende bij dezelfde α , is gelijk aan

$$W_\alpha = 2mn - W_\ell$$

en de rechter kritieke zone Z_α bestaat dan uit alle waarden van $W \geq W_\alpha$. De tweezijdige kritieke zone met onbetrouwbaarheidsdrempel 2α wordt verkregen door Z_ℓ en Z_α tezamen te nemen. Toetsen wij b.v. tweezijdig met onbetrouwbaarheidsdrempel $0,05$ dan wordt H_0 verworpen als de gevonden waarde van $W \geq W_\alpha$ en als $W \leq W_\ell$ is, waarbij W_α (resp. W_ℓ) de rechter (resp. linker) kritieke waarden, behorende bij $\alpha = 0,025$, is. Toetsen wij linkséénzijdig met onbetrouwbaarheidsdrempel $0,05$, dan wordt H_0 verworpen als $W \leq W_\ell$ (de linkerkritieke waarde, behorende bij ⁴⁾ Deze kritieke waarden zijn gedeeltelijk berekend met behulp van de exacte verdeling van W , gedeeltelijk met behulp van de benadering met de normale verdeling (zie par. 8). De benaderde waarden zijn in de tabel onderstreept.

$\alpha = 0,05$) is. Analooq voor rechtséénzijdig toetsen.

Opmerking 4. De tabellen I en II gelden strikt genomen alleen als er geen gelijke waarnemingen zijn, maar zij geven voor het geval van weinig gelijke waarnemingen veelal een goede benadering.

Opmerking 5. In tabel I staan meer decimalen opgenomen dan voor toepassing van de toets noodzakelijk is. Gewoonlijk kan men daarbij zonder bezwaar met 2 of 3 decimalen volstaan. Het opnemen van meer decimalen is uitsluitend uitgevoerd, om in voorkomende gevallen de tabellen ook voor andere doeleinden, waarbij een grotere precisie nodig is, geschikt te maken.

Wij kunnen nu de voorbeelden 1 en 2 nader uitwerken.

Voorbeeld 1

Bij dit voorbeeld was: $m=8$, $n=6$ en $W=38$.

Wij toetsen tweezijdig en kiezen als onbetrouwbaarheidsdrempel $\alpha = 0,05$. In tabel I zoeken wij in de tabel met $n=8$ en $m=6$ (N.B. m en n zijn in de tabellen verwisselbaar, zie boven) de éézijdige overschrijdingskans bij de regel $W=38$. Deze bedraagt: 0,29. De tweezijdige overschrijdingskans van $W=38$ is dus gelijk aan 0,58.

De tweezijdige overschrijdingskans is hier dus groter dan de onbetrouwbaarheidsdrempel α , en wij kunnen H_0 , inhoudende dat de draden van beide partijen, wat hun treksterkte betreft, equivalent zijn, dus niet verworpen.

Voorbeeld 2

Voor de onbetrouwbaarheid van de toets nemen wij weer $\alpha = 0,05$.

Hoewel bij dit voorbeeld enkele gelijke waarnemingen voorkomen maken wij toch gebruik van tabel I, daar de aantallen gelijken slechts klein zijn. Het getal $W=129$ is groter dan $m \cdot n = 80$, dus berekenen wij $W' = 2 \cdot 80 - 129 = 31$. Deze waarde komt in de tabel niet voor, echter wel de waarden 30 en 32. Wij nemen nu als benadering van de éézijdige overschrijdingskans van $W'=31$ het gemiddelde van de éézijdige overschrijdingskansen van $W=30$ en $W=32$. De tweezijdige overschrijdingskans van $W'=31$ is dus gelijk aan de som van de éézijdige overschrijdingskansen van $W=30$ en $W=32$:

$$k = 0,013 + 0,017 = 0,03.$$

Deze overschrijdingskans is kleiner dan $\alpha = 0,05$ en wij kunnen dus H_0 verwerpen. Wij concluderen dus dat de weerstanden, gefabriceerd volgens de twee procédés, systematisch verschillen en daar $W > m n$ liggen de waarden van de volgens het oude procédé gefabriceerde weerstanden systematisch hoger dan de anderen.

Voorbeeld 5

Ter vergelijking van het gehalte aan stof S in het bloed van gezonde en aan ziekte A lijdende mannen heeft men de volgende gegevens verzameld (de waarnemingen zijn reeds in volgorde van opklimmende grootte gerangschikt).

Gehalte S bij gezonde mannen (x):

9,4; 9,9; 10,0; 10,2; 10,5; 10,6; 11,2; 11,7; 11,8; 12,4; 12,6; 13,6; 13,8; 15,1.

Gehalte S bij mannen die aan A lijden (y):

10,1; 11,4; 12,6; 12,9; 14,7; 15,5; 15,8; 16,2.

De getoetste hypothese houdt in, dat het gehalte aan S in het bloed bij lijders aan A niet systematisch verschilt van dit gehalte bij gezonden, hetgeen dus betekent, dat de ziekte A dit gehalte niet beïnvloedt. Bij berekening vinden wij: $W = 51$ (de lezer controleer deze uitkomst).

Wij toetsen tweezijdig met onbetrouwbaarheidsdrempel 0,05. In tabel II zoeken wij de linker kritieke waarde van W op voor $m=8$ en $n=14$ en $\alpha=0,025$ (éénzijdig); deze bedraagt: $W_{\alpha} = 52$.

De gevonden waarde van W is dus kleiner dan W_{α} .

Conclusie: Lijders aan ziekte A hebben een systematisch hoger gehalte aan S in het bloed dan gezonden.

Opmerking 6. Bij $\alpha=0,01$ (éénzijdig) en $m=8$ en $n=14$ vinden wij $W_{\alpha} = 44$. Aangezien de gevonden waarde van W groter is dan 44, is de tweezijdige overschrijdingskans $k > 0,02$.

Dus:

$$0,02 < k < 0,05.$$

8. Benadering voor kritieke zones en overschrijdingskansen

Als m en n groot zijn en onderling niet te veel verschillen en als bovendien de groepen gelijke waarnemingen niet te veel in omvang verschillen kan men gebruik maken van het feit dat de grootheid W , als de hypothese H_0 juist is, bij benadering normaal verdeeld is met gemiddelde

$$\mu = m n$$

en variantie

$$\sigma^2 = \frac{m n (N^2 - D)}{3 N (N - 1)} \quad (\text{zie par. 5}).$$

De linkséénzijdige overschrijdingskans is dan bij benadering het oppervlak der normale verdeling met gemiddelde 0 en spreiding 1 links van

$$\frac{W - \mu + 1}{\sigma}$$

en kan dus in een tabel der normale verdeling worden opgezocht.

Opmerking 7: Daar hier een discrete verdeling benaderd wordt met een continue, wordt de z.g. continuïteitscorrectie toegepast. Hiervoor neemt men de halve afstand tussen de opeenvolgende waarden der toetsingsgrootheid. Deze afstand is hier, als er geen gelijke waarnemingen zijn, gelijk aan 2; in dit geval wordt de correctie dus gelijk aan 1. Dit is de 1 in de teller van bovenstaande formule.

Treden er wel gelijke waarnemingen op dan is de afstand tussen de opeenvolgende waarden van W in het algemeen niet constant, zodat een algemene regel voor de correctie dan moeilijk aan te geven is. Het verwaarlozen der correctie geeft echter soms een kleinere overschrijdingskans dan gerechtvaardigd is, zodat men beter doet ook hier 1 als correctie te nemen.

Als echter

- a. de groepen gelijke waarnemingen alle dezelfde omvang t hebben of als
- b. er slechts twee groepen gelijke waarnemingen zijn, dus als $k = 2$,⁵⁾

dan is de afstand tussen de opeenvolgende waarden van W wel constant en gelijk aan $2t$ in geval a en gelijk aan N in geval b; de continuïteitscorrectie is dan dus t resp. $\frac{1}{2}N$.

Analoog is de rechtséénzijdige overschrijdingskans bij benadering het oppervlak der normale verdeling met gemiddelde 0 en spreiding 1 rechts van

$$\frac{W - \mu - 1}{\sigma}$$

en de tweezijdige overschrijdingskans bij benadering twee maal het oppervlak rechts van

$$\frac{|W - \mu| - 1}{\sigma}.$$

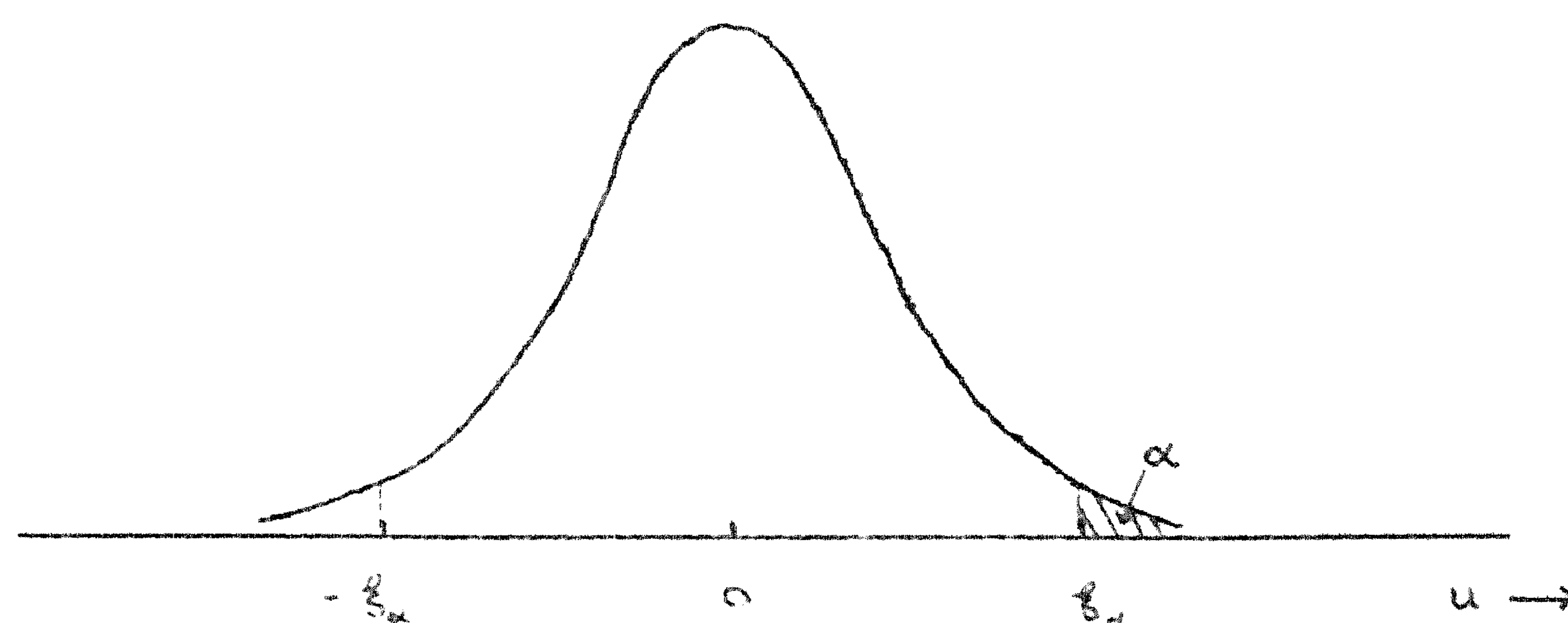
5) In dat geval gaat het schema der waarnemingen over in een 2×2 -tabel.

Opmerking 8: De linker kritieke waarde van W met éézijdige onbetrouwbaarheidsdrempel α is dus bij benadering de grootste waarde van W waarvoor de benaderde linkséézijdige overschrijdingskans $\leq \alpha$ is, dus de grootste waarde van W , waarvoor

$$\frac{W - \mu + 1}{\sigma} \leq -\xi_{\alpha}$$

is, waarbij ξ_{α} gevonden wordt met behulp van een tabel der normale verdeling (vgl. fig. 2):

$$\frac{1}{\sqrt{2\pi}} \int_{\xi_{\alpha}}^{\infty} e^{-\frac{1}{2}u^2} du = \alpha.$$



figuur 2

Definitie van ξ_{α} bij een normale verdeling.

De berekening van de overschrijdingskans bij voorbeeld 3 gaat nu als volgt:

$$W = 435,$$

$$N = 35,$$

$$D = 185,$$

$$\mu = mn = 300,$$

$$\sigma^2 = \frac{mn(N^3 - D)}{3N(N-1)} = 3587,39,$$

$$\sigma = 59,9.$$

De tweezijdige overschrijdingskans is dus bij benadering tweemaal het oppervlak der normale verdeling met gemiddelde 0 en spreiding 1 rechts van

$$\frac{|W - \mu| - 1}{\sigma} = 2,24.$$

Hiervoor vinden we in een tabel der normale verdeling

$$k = 2 \cdot 0,0125 = 0,025.$$

Indien wij, in de formule voor de variantie van W , de gelijke waarnemingen niet in rekening brengen, d.w.z. als wij de formule

$$\sigma^2 = \frac{1}{3} nm(N+1) \quad (\text{zie par. 5})$$

gebruiken, dan vinden wij $\sigma^2 = 3600$, dus $\sigma = 60$ en

$$\frac{|W - \mu| - 1}{\sigma} = 2,23.$$

Voor de overschrijdingskans vinden we dan

$$k = 2.0.0129 = 0,0258.$$

De invloed der gelijke waarnemingen is hier dus klein.

Gebruiken wij een onbetrouwbaarheidsdrempel 0,05, dan wordt, daar de overschrijdingskans kleiner is dan 0,05, de hypothese H_0 verworpen. H_0 houdt in, dat mannen en vrouwen niet systematisch in lengte verschillen, en deze hypothese wordt dus verworpen ten gunste van de hypothese dat mannen systematisch langer zijn dan vrouwen (de gevonden waarde van W is n.l. groter dan mn).

Opmerking 9: Als er gelijke waarnemingen zijn dan geldt

$$\frac{1}{3} mn(N+1) > \frac{mn(N^3 - D)}{3N(N-1)}.$$

Dus als wij, in de formule voor de variantie van W , de gelijke waarnemingen niet in rekening brengen, vinden we een te grote waarde voor deze variantie en dus een te grote waarde voor de overschrijdingskans. Is deze te grote waarde voor de overschrijdingskans $\leq \alpha$ dan wordt H_0 dus verworpen en hoeft men de berekening der derde machten niet uit te voeren.

Uitwerking van voorbeeld 4.

De toets wordt weer tweezijdig uitgevoerd met $\alpha = 0,01$. De berekening van de overschrijdingskans gaat als volgt:

$$W = 3482,$$

$$N = 151,$$

$$D = 60127,$$

$$\mu = mn = 5100,$$

$$\sigma^2 = \frac{mn(N^3 - D)}{3N(N-1)} = 253898,49,$$

$$\sigma = 503,88.$$

Dus

$$\frac{|W - \mu| - 1}{\sigma} = 3,2..$$

In een tabel der normale verdeling vinden we voor de tweezijdige overschrijdingskans

$$k = 2.0.0007 = 0,0014.$$

Brengen we in de formule voor de variantie de gelijke waar-

nemingen niet in rekening dan vinden we

$$\sigma^2 = \frac{1}{3} m m (N+1) = 258400,$$

$$\sigma = 508,3,$$

$$\frac{|W - \mu| - 1}{\sigma} = 3,18,$$

$$k = 2.0,0007 = 0,0014.$$

De gelijke waarnemingen hebben hier dus geen invloed.

Daar de overschrijdingskans kleiner is dan de onbetrouwbaarheidsdrempel verwerpen wij de hypothese H_0 ten gunste van de hypothese dat jongere mensen systematisch minder tijd voor de test nodig hebben dan oudere mensen.

9. Gevallen, waarin de exacte tabellen en de benadering niet te gebruiken zijn

Wanneer m en n klein zijn en bovendien de groepen gelijke waarnemingen veel in omvang verschillen, kan men noch van de in par.7 genoemde tabellen, noch van de in par.8 beschreven benadering gebruik maken. Dit geval doet zich voor bij het volgende voorbeeld:

Voorbeeld 6

Men wil twee geneesmiddelen A en B vergelijken wat betreft hun uitwerking bij de behandeling van een bepaalde ziekte; 5 patiënten lijdende aan deze ziekte krijgen een injectie met geneesmiddel A en 10 patiënten een injectie met B. Twee weken na deze injectie wordt voor iedere patiënt de graad van verbetering genoteerd; deze verbeteringsgraad wordt aangegeven met:

1. geen of geringe verbetering,
2. matige verbetering,
3. goede verbetering,
4. geheel genezen.

Deze waarnemingen kunnen we als volgt samenvatten.

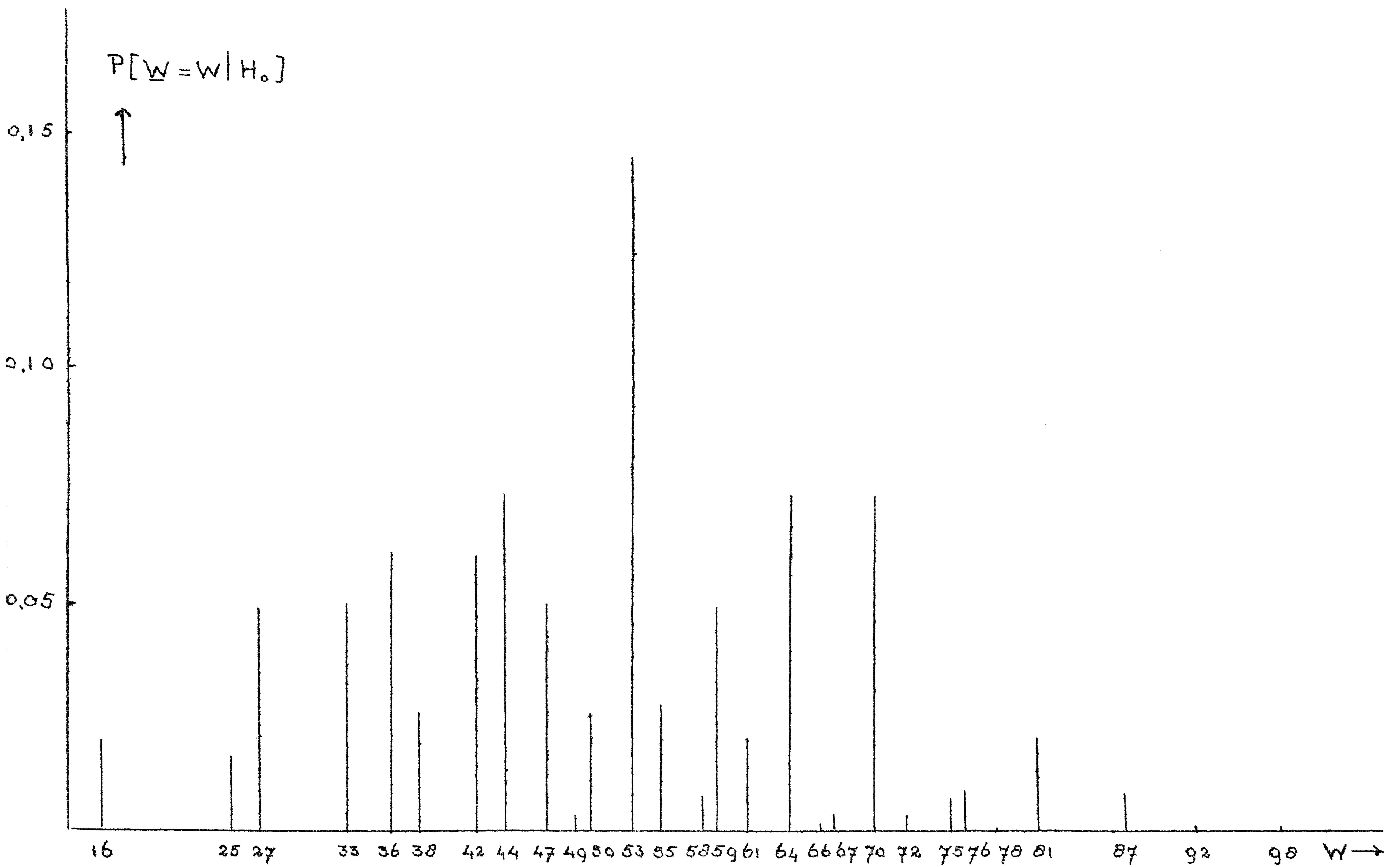
Tabel 2

Verbeteringsgraad van patiënten, behandeld met geneesmiddelen A en B:

verbeterings- graad	aantal patiënten met de betref- fende verbeteringsgraad bij geneesmiddel		totaal
	A	B	
1	1	0	1
2	3	5	8
3	1	2	3
4	0	3	3
totaal	5	10	15

In gevallen als die van voorbeeld 6 zal men, om de overschrijdingskansen te vinden, de exacte verdeling van \underline{W} onder de hypothese H_0 moeten berekenen. Hoe men deze exacte verdeling berekent en hoe men in deze gevallen de tweezijdige overschrijdingskansen definiëert zullen we hier niet uiteenzetten; dit valt buiten het kader van deze handleiding en wij verwijzen hiervoor naar [2].

Bij voorbeeld 4 heeft de exacte verdeling van \underline{W} , als juist is, de in fig.3 getekende vorm. Uit deze figuur ziet men dat de verdeling in dit geval een zeer onregelmatig karakter heeft.



figuur 3

Verdeling van \underline{W} onder de hypothese H_0 voor $n=5, m=10, t_1=1, t_2=8, t_3=t_4=3$.

10. Enige opmerkingen over de keuze tussen één- en tweezijdig toetsen

In par.6 hebben wij reeds opgemerkt, dat men de hypothese H_0 kan toetsen tegen de alternatieve hypothesen

$$H: P[x > y] \neq P[x < y]$$

of tegen

$$H_1: P[\underline{x} > y] > P[\underline{x} < y]$$

of tegen

$$H_2: P[\underline{x} > y] < P[\underline{x} < y].$$

Men gebruikt dan een tweezijdige, een rechtséénzijdige, resp. linkséénzijdige kritieke zone (overschrijdingskans). Bij rechtséénzijdig (en mutatis mutandis bij linkséénzijdig) toetsen kan men nimmer besluiten tot: \underline{x} is systematisch kleiner dan y , want een kleine waarde van W (d.w.z. kleiner dan m_n) kan bij rechtséénzijdige toetsing nooit in de kritieke zone liggen.

Hiermede zal men rekening moeten houden, indien men overweegt eenzijdig te toetsen. Anderzijds zal, bij rechtséénzijdige toetsing sneller ontdekt worden, dat \underline{x} systematisch groter is dan y , indien dit het geval is; immers de rechter kritieke zone is dan groter.

De rechtséénzijdige toets kan men dus toepassen als men er uitsluitend in geïnteresseerd is H_0 te verwerpen ten gunste van de alternatieve hypothese, dat \underline{x} systematisch groter is dan y , dus tegen $H_1: P[\underline{x} > y] > P[\underline{x} < y]$.

Dit zal b.v. het geval zijn, indien men, onafhankelijk van het gegeven waarnemingsmateriaal, weet dat \underline{x} niet systematisch kleiner is dan y . Ook praktische overwegingen kunnen echter tot deze keuze leiden (zie voorbeeld 7 verderop).

Men gebruikt dan een rechtséénzijdige kritieke zone en men besluit tot: \underline{x} is systematisch groter dan y , als de berekende waarde van W in de kritieke zone ligt.

Opmerking 10. Het is echter onjuist éénzijdig te toetsen naar aanleiding van de gevonden waarde van W . Zou men n.l., omdat b.v. $W > \mu$ gevonden is, een rechtséénzijdige kritieke zone nemen "omdat W toch niet in een linker kritieke zone ligt", dan dient men te overwegen, dat men, als $W < \mu$ gevonden was, hetzelfde met een linkséénzijdige zone zou hebben gedaan. In feite werkt men dan dus toch tweezijdig, maar zou men nu de rechtséénzijdige (resp. linkséénzijdige) kritieke zone de onbetrouwbaarheid α geven, dan werkt men dus eigenlijk met een tweezijdige kritieke zone met onbetrouwbaarheid 2α . Het feit, dat $W > \mu$ (resp. $< \mu$) gevonden is, doet zich gelden bij de conclusie, indien H_0 verworpen wordt. Immers men besluit dan niet alleen, dat H_0 onjuist is, maar ook, dat \underline{x} systematisch groter (resp. kleiner) dan y is.

Indien men éénzijdig kan toetsen verdient het aanbeveling dit inderdaad te doen, daar de éénzijdige toets, zoals boven reeds vermeld, bij dezelfde onbetrouwbaarheidsdrempel een groter onderscheidingsvermogen ⁶⁾ bezit dan de tweezijdige (zie par.11). De tot nu toe gegeven voorbeelden zijn allen tweezijdig getoetst omdat nergens redenen aanwezig waren, die het besluit van éénzijdig toetsen zouden rechtvaardigen.

Bij het volgende voorbeeld kunnen wij echter wel éénzijdig toetsen.

Voorbeeld 7.

Een textielfabrikant wil door toepassing van een prepareermethode A zijn garens sterker maken.

Trekproeven, verricht op monsters onbehandelde en met A behandelde garens, gaven de volgende resultaten (uitgedrukt in een of andere sterkte-eenheid).

Met A behandelde garens (x)	Onbehandelde garens (y)
11,79	10,34
11,21	11,40
13,20	10,19
12,66	12,10
13,37	11,46

De hypothese die wij willen toetsen luidt: de behandelde en de onbehandelde garens zijn even sterk.

Aangezien de fabrikant de prepareermethode voor deze garenssoort echter alleen wil invoeren indien deze methode de garens sterker maakt, is hij er dus alleen in geïnteresseerd H_0 te verwerpen ten gunste van de alternatieve hypothese dat de met A behandelde garens systematisch sterker zijn dan de niet behandelde. De twee mogelijkheden: a) de prepareermethode helpt niet en b) de prepareermethode schaadt, leiden beide tot dezelfde beslissing: deze niet in te voeren, en zijn dus bij dit probleem, praktisch gezien, equivalent, zodat zij niet van elkaar onderscheiden behoeven te worden. Het linker deel van een tweezijdige kritieke zone zou dus doelloos zijn.

We toetsen dus rechtséénzijdig en nemen als onbetrouwbaarheidsdrempel 0,05.

Voor de toetsingsgrootte vinden we hier $W=42$ en in tabel 1 vinden we bij $m=n=5$ voor de rechtséénzijdige overschrijdings-

6) Het onderscheidingsvermogen van een toets is de kans op het verwerpen van H_0 , indien H_0 onjuist is.

kans $k_{\alpha} = 0,048$.

Daar de overschrijdingskans kleiner dan de onbetrouwbaarheidsdrempel is, verwerpen wij de hypothese H_0 ten gunste van de hypothese dat de prepareermethode de garens sterker maakt.

Indien wij hier tweezijdig getoetst hadden, hadden wij voor de (tweezijdige) overschrijdingskans $k = 2 \cdot 0,048 = 0,096$ gevonden; de hypothese H_0 hadden wij dan, bij dezelfde onbetrouwbaarheidsdrempel, niet kunnen verwerpen.

De onbetrouwbaarheidsdrempel α van de éézijdige toets heeft in dit geval de volgende praktische interpretatie: de kans op het ten onrechte invoeren van A (als dit niet helpt of zelfs schaadt) is hoogstens gelijk aan α .⁷⁾

11. Het onderscheidingsvermogen van de toets van WILCOXON, vergeleken met dat van de toets van STUDENT

Het onderscheidingsvermogen van de toets van WILCOXON is voor speciale gevallen onderzocht. Men zou verwachten, dat het bij toepassing op twee steekproeven uit normale verdelingen met gelijke spreidingen aanzienlijk geringer zou zijn dan dat van de toets van STUDENT, daar bij eerstgenoemde toets van het gegeven der normaliteit geen gebruik wordt gemaakt.

VAN DER VAART [9] heeft het onderscheidingsvermogen van beide toetsen vergeleken voor zeer kleine steekproeven (n.l. $m+n \leq 6$) en daarbij bleek, dat het onderscheidingsvermogen van de toets van WILCOXON slechts weinig kleiner is dan dat van de toets van STUDENT.

Voor grote steekproeven ($m \rightarrow \infty$ en $n \rightarrow \infty$) bleek uit (niet gepubliceerde) onderzoekingen van G.E. NOETHER en E.J.G. PITMAN, dat de verhouding van de aantallen waarnemingen, die nodig zijn om de toets van WILCOXON en die van STUDENT in de buurt van H_0 hetzelfde onderscheidingsvermogen te geven, tot $\frac{3}{\pi} \approx 0,955$ nadert. Tevens bewezen zij, dat het onderscheidingsvermogen van de toets van WILCOXON voor scheve verdelingen - waarvoor dus strikt genomen de toets van STUDENT niet toepasbaar is - aanzienlijk groter kan zijn dan dat van de toets van STUDENT. De toets van WILCOXON heeft verder een belangrijk voordeel boven de toets van STUDENT, n.l. dat ver weg liggende waarnemingen, die wellicht (of wellicht niet) door een vergissing van enige aard veroorzaakt zijn, weinig invloed hebben op de uitkomst van de toets. Dit behoedt de onderzoeker voor het vaak

7) En wel, indien A niet helpt, doch schaadt, zelfs veel kleiner dan α .

zeer onaangename dilemma: een waarneming, die er onaannemelijk uitziet, doch waarvan men niet weet, of hij "fout" is, al of niet weg te laten.

Litteratuur over verdere onderzoekingen vindt men in par.12

12. Litteratuur

- [1] van Dantzig, D., On the consistency and the power of WILCOXON's two sample test, Proc.Kon.Ned.Akad.v.Wet. A 54 (1951), 1-8; Indagationes Mathematicae 13 (1951), 1-8.
- [2] van Eeden, Constance en Ir Doraline Wabeke, Handleiding voor de toets van WILCOXON (vervolg): Exacte behandeling als er gelijke waarnemingen zijn. Rapport S176(M65A) van de Statistische Afdeling van het Mathematisch Centrum, Amsterdam, 1955.
- [3] Hemelrijk, J. Note on WILCOXON's two sample test when ties are present, Ann.Math.Stat. 23 (1952), 133-135.
- [4] Lehmann, E.L. Consistency and unbiasedness of certain non-parametric tests, Ann.Math.Stat. 22 (1951), 165-179.
- [5] Lehmann, E.L. The power of rank tests, Ann.Math.Stat. 24 (1953), 23-43.
- [6] Mann, H.B. 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), 50-60.
- [7] Stoker, D.J. Oor 'n klas van toetsingsgrootthede vir die probleem van twee steekproewe, Dissertatie Amsterdam, 1955.
- [8] van der Vaart, H.R. Gebruiksaanwijzing voor de toets van WILCOXON, Rapport S32 (M4) van de Statistische Afdeling van het Mathematisch Centrum, Amsterdam, 1950.
- [9] van der Vaart, H.R., Some remarks on the power function of WILCOXON's test for the problem of two samples, Proc. Kon.Ned.Akad.v.Wet. A 53 (1950), 494-506 en 507-520; Indagationes Mathematicae 12 (1950), 146-158 en 159-172.

- [10] van der Waerden, B.L., Order tests for the two sample problem and their power, Proc.Kon.Ned.Akad. v.Wet. A 55 (1952), 453-458; Indagationes Mathematicae 14 (1952), 453-458. Corrigenda *ibid.* (1953), p.80.
- [11] Wilcoxon, F., Individual comparison's by ranking methods, *Biometrics* 1 (1945), 80-82.

TABEL I : linkséénzijdige overschrijdingskansen voor de toets van Wilcoxon.

n = 3

W \ m	1	2	3
0	0,250000	0,100000	0,050000
2	0,500000	0,200000	0,100000
4	0,750000	0,400000	0,200000
6		0,600000	0,350000
8			0,500000
10			0,650000

n = 4

W \ m	1	2	3	4
0	0,200000	0,066667	0,028571	0,014286
2	0,400000	0,133333	0,057143	0,028571
4	0,600000	0,266667	0,114286	0,057143
6		0,400000	0,200000	0,100000
8		0,600000	0,314286	0,171429
10			0,428571	0,242857
12			0,571429	0,342857
14				0,442857
16				0,557143

n = 5

W \ m	1	2	3	4	5
0	0,166667	0,047619	0,017857	0,007936	0,003968
2	0,333333	0,095238	0,035714	0,015873	0,007936
4	0,500000	0,190476	0,071429	0,031746	0,015873
6	0,666667	0,285714	0,125000	0,055556	0,027778
8		0,428571	0,196429	0,095238	0,047619
10		0,571429	0,285714	0,142857	0,075397
12			0,392857	0,206349	0,111111
14			0,500000	0,277778	0,154762
16			0,607143	0,365079	0,210318
18				0,452381	0,273810
20				0,547619	0,345238
22					0,420635
24					0,500000
26					0,579365

TABEL I : linkseénzijdige overschrijdingskansen (vervolg).

n = 6

W \ m	1	2	3	4	5
0	0,142857	0,035714	0,011905	0,004762	0,002164
2	0,285714	0,071429	0,023810	0,009524	0,004329
4	0,428571	0,142857	0,047619	0,019048	0,008658
6	0,571428	0,214286	0,083333	0,033333	0,015152
8		0,321429	0,130952	0,057143	0,025974
10		0,428571	0,190476	0,085714	0,041126
12		0,571429	0,273810	0,128571	0,062770
14			0,357143	0,176190	0,088744
16			0,452381	0,238095	0,123376
18			0,547619	0,304762	0,164502
20				0,380952	0,214286
22				0,457142	0,268398
24				0,542857	0,331169
26					0,396104
28					0,465368
30					0,534632

n = 7

W \ m	1	2	3	4	5
0	0,125000	0,027778	0,008333	0,003030	0,001263
2	0,250000	0,055556	0,016667	0,006061	0,002525
4	0,375000	0,111111	0,033333	0,012121	0,005051
6	0,500000	0,166667	0,058333	0,021212	0,008838
8	0,625000	0,250000	0,091667	0,036364	0,015152
10		0,333333	0,133333	0,054545	0,023990
12		0,444444	0,191667	0,081818	0,036616
14		0,555556	0,258333	0,115152	0,053030
16			0,333333	0,157576	0,074495
18			0,416667	0,206061	0,101010
20			0,500000	0,263636	0,133838
22			0,583333	0,324242	0,171717
24				0,393939	0,215909
26				0,463636	0,265152
28				0,536364	0,319444
30					0,377525
32					0,438131
34					0,500000
36					0,561869

TABEL I : linkséénzijdige overschrijdingskansen (vervolg)

n = 6

W \ m	6
0	0,001082
2	0,002164
4	0,004329
6	0,007576
8	0,012987
10	0,020563
12	0,032468
14	0,046537
16	0,066017
18	0,089827
20	0,120130
22	0,154762
24	0,196970
26	0,242424
28	0,294372
30	0,349567
32	0,409091
34	0,468614
36	0,531385

n = 7

W \ m	6	7
0	0,000583	0,000291
2	0,001165	0,000583
4	0,002331	0,001166
6	0,004079	0,002040
8	0,006993	0,003497
10	0,011072	0,005536
12	0,017482	0,008741
14	0,025641	0,013112
16	0,036713	0,018939
18	0,050699	0,026515
20	0,068764	0,036422
22	0,090326	0,048660
24	0,117133	0,064103
26	0,147436	0,082459
28	0,182983	0,104312
30	0,222610	0,129662
32	0,266899	0,158800
34	0,314102	0,191434
36	0,365384	0,227856
38	0,417832	0,267483
40	0,472610	0,310023
42	0,527389	0,355187
44		0,402389
46		0,450758
48		0,500000
50		0,549242

TABEL I : linkséénzijdige overschrijdingskansen (vervolg)

n = 8

W \ m	1	2	3	4	5
0	0,111111	0,022222	0,006061	0,002020	0,000777
2	0,222222	0,044444	0,012121	0,004040	0,001554
4	0,333333	0,088889	0,024242	0,008081	0,003108
6	0,444444	0,133333	0,042424	0,014141	0,005439
8	0,555556	0,200000	0,066667	0,024242	0,009324
10		0,266667	0,096970	0,036364	0,014763
12		0,355556	0,139394	0,054545	0,022533
14		0,444444	0,187879	0,076768	0,032634
16		0,555556	0,248485	0,107071	0,046620
18			0,315152	0,141414	0,063714
20			0,387879	0,183838	0,085470
22			0,460606	0,230303	0,111111
24			0,539394	0,284848	0,142191
26				0,341414	0,177156
28				0,404040	0,217560
30				0,466667	0,261849
32				0,533333	0,310800
34					0,362082
36					0,416472
38					0,471639
40					0,528361

TABEL I : linkséénzijdige overschrijdingskansen (vervolg)

n = 8

W \ m	6	7	8
0	0,000333	0,000155	0,000078
2	0,000666	0,000311	0,000155
4	0,001332	0,000622	0,000311
6	0,002331	0,001088	0,000544
8	0,003996	0,001865	0,000932
10	0,006327	0,002953	0,001476
12	0,009990	0,004662	0,002331
14	0,014652	0,006993	0,003496
16	0,021312	0,010256	0,005206
18	0,029637	0,014452	0,007382
20	0,040626	0,020047	0,010334
22	0,053946	0,027040	0,014064
24	0,070929	0,036053	0,018959
26	0,090576	0,046931	0,024942
28	0,114219	0,060295	0,032479
30	0,141192	0,075991	0,041492
32	0,172494	0,094639	0,052448
34	0,206793	0,115928	0,065190
36	0,245421	0,140482	0,080264
38	0,286380	0,167832	0,097436
40	0,331002	0,198446	0,117249
42	0,377289	0,231702	0,139316
44	0,425907	0,267910	0,164102
46	0,474858	0,306294	0,191142
48	0,525141	0,347164	0,220901
50		0,389433	0,252680
52		0,433255	0,286869
54		0,477544	0,322688
56		0,522455	0,360450
58			0,399223
60			0,439238
62			0,479565
64			0,520435

TABEL I : linkséénzijdige overschrijdingskansen (vervolg)

n = 9

W \ m	1	2	3	4	5
0	0,100000	0,018182	0,004545	0,001399	0,000500
2	0,200000	0,036364	0,009091	0,002797	0,000999
4	0,300000	0,072727	0,018182	0,005594	0,001998
6	0,400000	0,109091	0,031818	0,009790	0,003496
8	0,500000	0,163636	0,050000	0,016783	0,005994
10	0,600000	0,218182	0,072727	0,025175	0,009490
12		0,290909	0,104545	0,037762	0,014486
14		0,363636	0,140909	0,053147	0,020979
16		0,454545	0,186364	0,074126	0,029970
18		0,545455	0,240909	0,099301	0,041458
20			0,300000	0,130070	0,055944
22			0,363636	0,165035	0,073426
24			0,431818	0,206993	0,094905
26			0,500000	0,251748	0,119880
28			0,568182	0,302098	0,148851
30				0,355244	0,181818
32				0,412587	0,218781
34				0,469930	0,259240
36				0,530070	0,303196
38					0,349650
40					0,398601
42					0,449051
44					0,500000
46					0,550949

TABEL I : linkséénzijdige overschrijdingskansen (vervolg)

n = 9

W \ m	6	7	8	9
0	0,000200	0,000087	0,000041	0,000021
2	0,000400	0,000175	0,000082	0,000041
4	0,000799	0,000350	0,000164	0,000082
6	0,001399	0,000612	0,000288	0,000144
8	0,002398	0,001049	0,000494	0,000247
10	0,003796	0,001651	0,000782	0,000391
12	0,005994	0,002622	0,001234	0,000617
14	0,008791	0,003934	0,001851	0,000926
16	0,012787	0,005769	0,002756	0,001378
18	0,017982	0,008217	0,003949	0,001995
20	0,024775	0,011451	0,005553	0,002818
22	0,033167	0,015560	0,007610	0,003887
24	0,043956	0,020892	0,010325	0,005306
26	0,056743	0,027448	0,013698	0,007096
28	0,072328	0,035577	0,017976	0,009379
30	0,090509	0,045367	0,023200	0,012217
32	0,111888	0,057080	0,029618	0,015734
34	0,136064	0,070804	0,037228	0,019992
36	0,163836	0,086888	0,046360	0,025154
38	0,194206	0,105245	0,056972	0,031263
40	0,227972	0,126136	0,069395	0,038503
42	0,264335	0,149563	0,083587	0,046956
44	0,303496	0,175524	0,099794	0,056746
46	0,344455	0,203934	0,117935	0,067956
48	0,387812	0,234878	0,138297	0,080749
50	0,431968	0,268007	0,160633	0,095125
52	0,477322	0,303234	0,185191	0,111209
54	0,522677	0,340297	0,211724	0,129042
56		0,378846	0,240354	0,148663
58		0,418532	0,270712	0,170054
60		0,459091	0,302920	0,193254
62		0,500000	0,336487	0,218141
64		0,540909	0,371493	0,244714
66			0,407104	0,272851
68			0,444179	0,302406
70			0,481283	0,333237
72			0,518716	0,365220
74				0,398087
76				0,431654
78				0,465714
80				0,500000
82				0,534286

TABEL I : linkséénzijdige overschrijdingskansen (vervolg)

n = 10

W \ m	1	2	3	4	5
0	0,090909	0,015152	0,003496	0,000999	0,000333
2	0,181818	0,030303	0,006993	0,001998	0,000666
4	0,272727	0,060606	0,013986	0,003996	0,001332
6	0,363636	0,090909	0,024476	0,006993	0,002331
8	0,454545	0,136364	0,038462	0,011988	0,003996
10	0,545455	0,181818	0,055944	0,017982	0,006327
12		0,242424	0,080420	0,026973	0,009657
14		0,303030	0,108392	0,037962	0,013986
16		0,378788	0,143357	0,052947	0,019980
18		0,454545	0,185315	0,070929	0,027639
20		0,545455	0,234266	0,093906	0,037629
22			0,286713	0,119880	0,049617
24			0,346154	0,151848	0,064602
26			0,405594	0,186813	0,082251
28			0,468532	0,226773	0,103230
30			0,531469	0,269730	0,127206
32				0,317682	0,154845
34				0,366633	0,185481
36				0,419580	0,219780
38				0,472527	0,256743
40				0,527472	0,297036
42					0,339327
44					0,383949
46					0,429570
48					0,476523
50					0,523476

TABEL I : linkséénzijdige overschrijdingskansen (vervolg)

n = 10

W \ m	6	7	8	9	10
0	0,000125	0,000051	0,000023	0,000011	0,000005
2	0,000250	0,000103	0,000046	0,000022	0,000011
4	0,000500	0,000206	0,000091	0,000043	0,000022
6	0,000874	0,000360	0,000160	0,000076	0,000038
8	0,001499	0,000617	0,000274	0,000130	0,000065
10	0,002373	0,000977	0,000434	0,000206	0,000103
12	0,003746	0,001542	0,000686	0,000325	0,000162
14	0,005494	0,002314	0,001028	0,000487	0,000244
16	0,007992	0,003393	0,001531	0,000725	0,000362
18	0,011239	0,004833	0,002194	0,001050	0,000525
20	0,015609	0,006787	0,003108	0,001494	0,000752
22	0,020979	0,009255	0,004273	0,002067	0,001044
24	0,027972	0,012494	0,005827	0,002836	0,001440
26	0,036339	0,016505	0,007770	0,003810	0,001943
28	0,046703	0,021544	0,010261	0,005066	0,002558
30	0,058941	0,027663	0,013323	0,006636	0,003420
32	0,073551	0,035119	0,017140	0,008606	0,004465
34	0,090285	0,043912	0,021710	0,011009	0,005748
36	0,109890	0,054401	0,027264	0,013964	0,007344
38	0,131743	0,066536	0,033800	0,017493	0,009271
40	0,156593	0,080625	0,041570	0,021737	0,011615
42	0,183816	0,096616	0,050551	0,026738	0,014402
44	0,213911	0,114767	0,060994	0,032627	0,017731
46	0,246129	0,134924	0,072855	0,039446	0,021628
48	0,281093	0,157394	0,086407	0,047360	0,026212
50	0,317682	0,181921	0,101536	0,056377	0,031506
52	0,356393	0,208659	0,118492	0,066650	0,037627
54	0,396228	0,237350	0,137140	0,078200	0,044604
56	0,437437	0,268099	0,157708	0,091157	0,052561
58	0,478896	0,300442	0,179967	0,105512	0,061502
60	0,521104	0,334533	0,204122	0,121403	0,071570
62		0,369806	0,229878	0,138756	0,082746
64		0,406262	0,257393	0,157689	0,095158
66		0,443387	0,286302	0,178116	0,108781
68		0,481129	0,316719	0,200090	0,123725
70		0,518870	0,348233	0,223484	0,139930
72			0,380913	0,248349	0,157499
74			0,414278	0,274480	0,176340
76			0,448375	0,301890	0,196524
78			0,482700	0,330360	0,217936
80			0,517300	0,359847	0,240625
82				0,390092	0,264424
84				0,421052	0,289370
86				0,452412	0,315264
88				0,484119	0,342105
90				0,515880	0,369682
92					0,397968
94					0,426714
96					0,455898
98					0,485256
100					0,514743

TABEL II : linkerkritieke waarden van W voor de toets van Wilcoxon. 8)
 $\alpha = 0,01$ (eenzijdig)

n \ m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	-	-	2	8	14	18	24	30	36	44	50									
12	-	-	4	10	16	22	28	34	42	48	56	62								
13	-	0	4	10	18	24	32	40	46	54	62	70	78							
14	-	0	4	12	20	26	34	44	52	60	68	76	86	94						
15	-	0	6	14	22	30	38	48	56	66	74	84	94	102	112					
16	-	0	6	14	24	32	42	52	62	72	82	92	102	112	120	130				
17	-	0	8	16	26	36	46	56	66	76	88	98	110	120	130	140	152			
18	-	0	8	18	28	38	48	60	72	82	94	106	118	128	140	150	162	174		
19	-	2	8	18	30	40	52	64	76	88	100	112	126	136	148	162	174	186	200	
20	-	2	10	20	32	44	56	68	80	94	106	120	132	144	158	172	186	198	212	226
21	-	2	10	22	34	46	60	72	86	100	114	128	140	154	168	182	196	210	224	
22	-	2	12	22	36	48	62	76	90	106	120	134	148	162	178	192	208	222		
23	-	2	12	24	38	52	66	80	96	110	126	140	156	172	186	202	218			
24	-	2	12	26	40	54	70	84	100	116	132	148	164	180	196	214				
25	-	2	14	26	42	58	72	90	106	122	138	154	172	188	206					
26	-	2	14	28	44	60	76	94	110	128	144	162	180	198						
27	-	4	14	30	46	62	80	98	116	132	150	168	188							
28	-	4	16	32	48	66	84	102	120	138	156	176								
29	-	4	16	32	50	68	86	106	126	144	164									
30	-	4	18	34	52	70	90	110	130	148										
31	-	4	18	36	54	74	94	114	134											
32	-	4	18	36	56	76	98	118												
33	-	4	20	38	58	80	100													
34	-	6	20	40	60	82														
35	-	6	22	40	62															
36	-	6	22	42																
37	-	6	22																	
38	-	6																		
39	-																			

8) "-" betekent dat voor de betreffende waarden van n en m voor geen enkele waarde van W de linkséénzijdige overschrijdingskans $\leq \alpha$ is. De onderstreepte waarden zijn gevonden met behulp van de normale benadering.

TABEL II : linkerkritieke waarden van W voor de toets van Wilcoxon (vervolg)
 $\alpha = 0,05$ (eenzijdig)

n \ m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	-	2	10	16	24	32	38	46	54	62	68									
12	-	4	10	18	26	34	42	52	60	68	76	84								
13	-	4	12	20	30	38	48	56	66	74	84	94	102							
14	-	6	14	22	32	42	52	62	72	82	92	102	112	122						
15	-	6	14	24	36	46	56	66	78	88	100	110	122	<u>132</u>	<u>144</u>					
16	-	6	16	28	38	50	60	72	84	96	108	120	130	142	154	166				
17	-	6	18	30	40	52	66	78	90	102	114	128	140	<u>152</u>	<u>166</u>	<u>178</u>	192			
18	-	8	18	32	44	56	70	82	96	110	122	136	<u>150</u>	<u>164</u>	<u>176</u>	<u>190</u>	<u>204</u>	218		
19	0	8	20	34	46	60	74	88	102	116	130	<u>144</u>	<u>160</u>	<u>174</u>	<u>188</u>	<u>202</u>	<u>216</u>	<u>232</u>	246	
20	0	8	22	36	50	64	78	94	108	124	<u>138</u>	<u>154</u>	<u>168</u>	<u>184</u>	<u>200</u>	<u>214</u>	<u>230</u>	<u>246</u>	<u>260</u>	<u>276</u>
21	0	10	22	38	52	68	82	98	114	130	146	162	178	194	210	226	242	258	276	
22	0	10	24	40	56	72	88	104	120	136	<u>154</u>	<u>170</u>	<u>188</u>	<u>204</u>	<u>222</u>	<u>238</u>	<u>256</u>	<u>272</u>		
23	0	10	26	42	58	74	92	108	126	<u>144</u>	<u>162</u>	<u>180</u>	<u>196</u>	<u>214</u>	<u>232</u>	<u>250</u>	<u>268</u>			
24	0	12	26	44	60	78	96	114	132	<u>150</u>	<u>170</u>	<u>188</u>	<u>206</u>	<u>224</u>	<u>244</u>	<u>262</u>				
25	0	12	28	46	64	82	100	120	<u>138</u>	<u>158</u>	<u>176</u>	<u>196</u>	<u>216</u>	<u>236</u>	<u>254</u>					
26	0	12	30	48	66	86	106	124	144	164	184	204	226	<u>246</u>						
27	0	14	30	50	70	90	110	130	<u>150</u>	<u>172</u>	<u>192</u>	<u>214</u>	<u>234</u>							
28	0	14	32	52	72	92	114	136	<u>156</u>	<u>178</u>	<u>200</u>	<u>222</u>								
29	0	14	34	54	76	96	118	<u>140</u>	<u>162</u>	<u>186</u>	<u>208</u>									
30	0	14	34	56	78	100	122	<u>146</u>	<u>170</u>	<u>192</u>										
31	0	16	36	58	80	104	128	<u>152</u>	<u>176</u>											
32	0	16	38	60	84	108	132	<u>156</u>												
33	0	16	38	62	86	112	<u>136</u>													
34	0	18	40	64	90	114														
35	0	18	42	66	92															
36	0	18	42	68																
37	0	20	44																	
38	0	20																		
39	2																			