K O N I N K L I J K N E D E R L A N D S
M E T E O R O L O G I S C H I N S T I T U U T

D e B i l t

WETENSCHAPPELIJK RAPPORT
W.R. 74-7

S.Kruizinga

Aanpassingstoetsen

De Bilt, 1974

Publikatienummer: K.N.M.I. W.R. 74-7 (S.B.)

U.D.C.: 519.2

Summary

For statistical purposes a study was performed on the subject of Tests of Fit. A number of tests were analyzed with the help of Monte Carlo methods. The general conclusion stated in this report is that especially the classical tests of fit are hardly useful for making the choice of a specified distribution in the case that a small sample is used.

Samenvatting

Ten behoeve van toekomstig statistisch onderzoek werd een literatuurstudie op het gebied van aanpassingstoetsen verricht. Van een aantal toetsen werden de eigenschappen nader onderzocht met behulp van Monte Carlo methoden. Als algemene konklusie komt naar voren dat in het bijzonder de klassieke toetsen bij kleine steekproeven nauwelijks geschikt zijn om een bepaalde keus voor de kansverdeling te rechtvaardigen.

1. Inleiding

Bij de meeste statistische bewerkingen is het noodzakelijk om veronderstellingen te maken ten aanzien van de verdeling waaruit de gegeven steekproef afkomstig is. Dit geldt vooral in die gevallen waarin gevraagd wordt naar een statistische extrapolatie, bijvoorbeeld bij vragen als "Welke grenswaarde wordt gemiddeld eens in de honderd jaar overschreden" terwijl men slechts over vijftig jaar waarnemingen beschikt. Dit laatste probleem is alleen op te lossen door een aanname te doen over de kansverdeling van de betreffende grootheid. Belangrijk hierbij is dan dat men de hypothese ten aanzien van de kansverdeling zo goed mogelijk tracht te toetsen op basis van het beschikbare materiaal. Hiervoor maakt men gebruik van een aanpassingstoets. In de statistische literatuur zijn vele aanpassingstoetsen te vinden doch de keus, welke toets men dient te gebruiken, wordt vaak bemoeilijkt doordat of de relevante gegevens niet beschikbaar of moeilijk te vinden zijn. In dit rapport wordt een overzicht gegeven van een aantal aanpassingstoetsen. Veel van de gegevens zijn afkomstig uit de literatuur; voor een aantal gevallen echter werd aanvullend onderzoek verricht. Veel nadruk werd gelegd op het onderscheidend vermogen van de toetsen bij steekproeven van kleine omvang.

Om het belang van een juiste aanpassing te illustreren geven we nog het volgende voorbeeld.[1] Uit een Gumbel verdeling, een veel gebruikte verdeling voor extreme waarden, hebben we een steekproef van twintig stuks getrokken. Hieraan hebben we zowel een normale als een Gumbel verdeling aangepast. In figuur 1 zijn de aangepaste verdelingen en de bronverdeling weergegeven. Op grond van de aangepaste normale verdeling berekenen we nu dat de waarde 3 slechts met 1% kans zal worden overschreden, in werkelijkheid is dit 5%. Met Gumbel aanpassing vinden we 3%. We zien hieruit dat een verkeerde keuze van de verdeling bij de aanpassing tot zeer slechte resultaten kan leiden.

2. Algemene opbouw van een aanpassingstoets

Een aanpassingstoets bestaat uit de volgende onderdelen:

2.1 Vooronderstellingen: meestal wordt niet meer geeist dan dat de elementen $x_1, \ldots x_N$ van de gegeven steekproef onderling onafhankelijk zijn getrokken.

2.2 Nulhypothese: hieronder wordt de stelling geformuleerd die men wenst te toetsen.
Bijv. a. De steekproef is afkomstig uit een normale verdeling met $\mu = 0$ en $\sigma^2 = 1$.
b. De steekproef is afkomstig uit een Gumbel verdeling.

De nulhypothese waarbij tevens alle parameters worden gespecificeerd noemen we enkelvoudig, alle andere samengesteld. In praktische situaties zullen we meestal te maken hebben met meervoudige hypothesen, waarbij de parameters nog geschat zullen worden uit de steekproef.

1) De in dit rapport genoemde verdelingen worden in appendix 1 nader beschreven.

2.3 Toetsingsgrootheid: dit is een hulpgrootheid die volgens een bepaald recept uit de steekproef wordt berekend en waarmee de toets wordt uitgevoerd.
Het is belangrijk op te merken dat T in feite een stochastische grootheid is.

2.4 De nulverdeling van T: indien we herhaaldelijk steekproeven ter grootte N trekken zullen we voor T steeds een ander antwoord vinden. Dus ook T heeft een zekere kansverdeling. De nulverdeling van T is de verdeling die ontstaat indien de steekproeven inderdaad afkomstig zijn uit de in de nulhypothese geformuleerde verdeling. Op grond van deze nulverdeling bepaalt men het kritieke gebied dat wil zeggen het gebied van uitkomsten van T die aanleiding geven tot niet accepteren van de nulhypothese.
Voorbeeld: stel we weten van een bepaalde T dat onder de nulhypothese de uitkomsten in de buurt van nul zullen liggen en dat bijvoorbeeld de kans op een uitkomst tussen -1 en +1 gelijk is aan 95 procent. Vinden we nu bij onze aktuele steekproef een uitkomst buiten (-1, +1), kritiek gebied dus $(-\infty, -1)$ en $(+1, +\infty)$, dan stellen we dat deze uitkomst zo zeldzaam is dat we eerder zullen geloven dat de nulhypothese niet waar is. We accepteren hierbij dus het risico van 5% dat we de nulhypothese ten onrechte verwerpen. Willen we dit risico verkleinen dan moeten we het kritieke gebied kleiner maken, dit heeft echter tot gevolg dat we minder kans zullen hebben op verwerpen van de nulhypothese als deze inderdaad onjuist is of m.a.w. meer kans op het ten onrechte accepteren daarvan . Algemeen wordt een grens van 95% geaccepteerd met dus een ingebouwd risico van 5% op een onjuiste beslissing. Dit is minder ernstig indien men in rekening brengt dat de steekproeven waarbij men ten onrechte de nulhypothese verwerpt vermoedelijk ook een slechte basis zullen zijn om verder op door te werken.
In bovenstaand voorbeeld was dus het kritieke gebied gekozen van $-\infty$, tot -1 en van +1 tot $+\infty$ overeenkomend met een kans van 5% op ten onrechte verwerpen van de nulhypothese. Meestal wordt de grootte van het kritieke gebied aangeduid met het percentage, we spreken dan dus van een kritiek gebied van 5%. In dit rapport zullen we de aanpassingstoetsen onderling, om redenen van rekentechnische aard, vergelijken voor een kritiek gebied van 10%.

2.5 Alternatieve hypothese: meestal wordt deze geformuleerd als "De nulhypothese is niet waar". Soms wordt hij ook konkreter geformuleerd bijvoorbeeld het alternatief bij 2.2 b. zou kunnen luiden "De steekproef is afkomstig uit een normale verdeling". De keuze van de toets kan hierdoor beinvloed worden.

2.6 Alternatieve verdeling van T: dit is de kansverdeling van T als het alternatief waar is. Dit is alleen een zinvol begrip indien het alternatief ook konkreet geformuleerd is. De alternatieve verdeling wordt gebruikt om het onderscheidend vermogen vast te stellen.

2.7 Het onderscheidend vermogen: hiermee duiden we aan de kans om te ontdekken dat de nulhypothese niet waar is onder omstandigheden dat de nulhypothese inderdaad niet waar is. Deze kans is dus gelijk aan de kans dat T in het kritieke gebied komt indien het alternatief waar is. Hieruit volgt dat we het onderscheidend vermogen alleen kunnen vaststellen indien het alternatief konkreet geformuleerd is. Uiteraard is het O.V. het belangrijkste argument bij de keuze van een toets. Dus in het geval dat het alternatief konkreet is omschreven, is de keus erg eenvoudig te maken. Indien het alternatief niet duidelijk is omschreven zou geen keus gemaakt kunnen worden. In dat geval kan men het beste de toets kiezen die een zo groot mogelijk onderscheidend vermogen heeft tegen allerlei mogelijke alternatieven.

In de volgende paragraaf worden een aantal toetsen beschreven waarbij de aandacht gericht zal zijn op de volgende punten:

a. Toegestane nulhypothesen
b. Toetsingsgrootheid
c. Nulverdeling
d. Onderscheidend vermogen

3. Eigenschappen van een aantal aanpassingstoetsen

In deze paragraaf zullen de volgende notaties gebruikt worden.

$x_1, \dots, x_N$: Steekproef ter grootte N

$x^*_1, \dots, x^*_N$: Geordende steekproef volgens $x^*_i \leq x^*_{i+1}$

$f(x, \Theta_1, \dots \Theta_I)$: Kansdichtheidsfunktie (d.f.) van x met parameters $\Theta_1, \dots \Theta_I$

$F(x, \Theta_1, \dots \Theta_I)$: Verdelingsfunktie (v.f.) x met parameters $\Theta_1, \dots, \Theta_I$

f_o en F_o : d.f. en v.f. behorend bij de nulhypothese

f_A en F_A : d,f. en v.f. behorend bij de alternatieve hypothese

$\hat{\Theta}_1, \dots, \hat{\Theta}_I$: Schatters van Θ_1 t/m Θ_I. Tenzij anders vermeld zijn de "maximum likelyhood" schatters bedoeld.

m : gemiddelde van de steekproefwaarden: $m = \frac{1}{N} \sum_{n=1}^{N} x_n$

s^2 : variantie van de steekproefwaarden: $s^2 = \frac{1}{n-1} \sum_{n=1}^{N} (x_n - m)^2$

$y_1, \dots, y_N$: De uit x_1 t/m x_N afgeleide steekproef volgens $y_n = F_o(x_n, \hat{\Theta}_1, \dots \hat{\Theta}_I)$ (terugtransformeren naar uniforme verdeling).

3.1 De χ^2-aanpassingstoets

Dit is de meest gebruikte aanpassingstoets. In elk statistisch studieboek wordt hieraan aandacht besteed.

a. Nulhypothese. Iedere willekeurige nulhypothese is toegestaan.

b. De toetsingsgrootheid, meestal aangeduid met X^2, wordt als volgt berekend:

1. Bereken, indien nodig, de schatters voor de parameters
2. Verdeel de uitkomstruimte voor x in K deelruimten en bereken voor ieder van de deelruimten V_k de kans p_k dat een waarneming in die deelruimte terecht komt. Gebruik hiervoor $F(x, \hat{\theta}_1, \ldots \hat{\theta}_I)$.

N.B. Onder uitkomstruimte wordt verstaan de verzameling van uitkomsten van de gegeven stochastische variabele, in de door ons beschouwde gevallen vaak de verzameling van de reële getallen.

3. Tel het aantal elementen n_k van de steekproef $x_1,\ldots,x_N$ dat in de deelruimte V_k ligt voor iedere k.

4. Bereken $X^2 = \sum\limits_{n=1}^{K} \dfrac{(n_k - N.p_k)^2}{N.p_k}$

In deze laatste formule wordt het verwachte aantal vergeleken met het aktuele aantal in de deelruimten. Indien de nulhypothese waar is, verwachten we X^2 klein. Het kritieke gebied ligt dan ook bij grote χ^2.
In het algemeen wordt aanbevolen alle p_k gelijk te maken en verder $K \approx 2(N-1)^{2/5}$ met als neven voorwaarde $N.p_k \geqslant 5$.

c. Nulverdeling. Voor grote N geldt dat χ^2 in zeer goede benadering verdeeld is volgens $\chi^2_{(\nu)}$ verdeling met $\nu=$(K-1 - aantal geschatte parameters) vrijheidsgraden. Deze benadering wordt ook gebruikt bij kleine N. In hoeverre dit is toegestaan is onbekend. Uit de χ^2 verdeling wordt de grenswaarde voor X^2 afgelezen.

d. Onderscheidend vermogen

Door Shapiro [2)] is het O.V. van het groot aantal toetsen, waaronder ook de χ^2 toets, bij verschillende alternatieven onderzocht. De nulhypothese was enkelvoudig en luidde
$F(x, \Theta_1, \Theta_2) = N(\mu, \sigma^2)$ = Normale verdeling met μ, σ^2 gelijk aan de theoretische waarden van de diverse alternatieven. Behalve voor de extreem afwijkende alternatieven waren de resultaten teleurstellend. Bijvoorbeeld voor een kritiek gebied van 10% en N = 20 werd, met als alternatief een uniforme verdeling, een O.V. van 18% bereikt (zie appendix 1 voor genoemde verdelingen). Bij de alternatieven Weibull (2) en χ^2 (4) was dit respektievelijk 12% en 20%. Bij N = 50 werd dit respektievelijk 22%, 17% en 69%.

Daar een enkelvoudige hypothese in de praktijk niet zoveel voorkomt hebben we deze toets ook nog onderzocht voor een samengestelde hypothese. Hierbij werd tevens onderzocht in hoeverre de nulverdeling nog door de χ^2 verdeling benaderd werd. We onderzochten de situatie $N = 20$, $K = 4$ en $p_k = \frac{1}{4}$. De nulhypothese luidde "De steekproef komt uit een normale verdeling".
We moeten twee parameters schatten dus de nulverdeling van de toetsingsgrootheid zou een $\chi^2(1)$ verdeling moeten zijn. Dit onderzoek geschiedde als volgt. Door de Wang tafelrekenmachine werden 1000 steekproeven van 20 stuks uit een normale verdeling getrokken. Van ieder van de steekproeven werd de X^2 berekend volgens het recept. De frekwentieverdeling van deze 1000 X^2 waarden werd beschouwd als een goede benadering van de nulverdeling van de toetsingsgrootheid. Om het O.V. voor een aantal alternatieven te schatten werden daarna uit ieder der alternatieven Uniform, Gumbel en Weibull (2), een groot aantal steekproeven getrokken. Voor ieder der steekproeven werd eveneens X^2 berekend. Hiervan werden weer frekwentieverdelingen opgesteld, de alternatieve verdelingen van T. Opgemerkt dient te worden dat in de gegeven situatie X^2 slechts $n \times 0.4$ kan zijn, met $n = 0, 1, \ldots 150$ (niet iedere n). Dus X^2 kan zeker geen $\chi^2(1)$ verdeling volgen, het is echter wel mogelijk dat $\chi^2(1)$ goede schatting geeft van de kans $P(X^2 > n \times 0.4)$. De resultaten zijn samengevat in tabel 1. Hierin is tevens de χ^2 schatting van de nulverdeling opgenomen. Vergelijking van kolom 2 en 3 resp. experimentele nulverdeling en X^2 levert een duidelijke diskrepantie op.We zien in de eerste kolom dat de grenswaarde voor X^2 voor een kritiek gebied van 10% gelijk is aan 10×0.4.

Tabel 1 Verdeling van X^2 onder nulhypothese en alternatieve hypothese

$n \times 0.4$	2 Normaal (1000)	3 $\chi^2(1)$	4 Uniform (1000)	5 Gumbel (1000)	6 Weibull (2) (750)
0	100.0	100.0	100.0	100.0	100.0
1	95.9	52.7	97.8	97.4	97.6
2	71.5	37.1	81.8	79.1	72.3
3	64.1	27.3	76.9	72.9	65.8
4	44.5	20.6	60.0	52.0	47.5
5	37.5	15.7	53.6	43.5	37.9
6	27.4	12.1	41.5	36.3	29.6
7	24.7	9.4	38.1	33.8	27.0
8	13.6	7.4	25.2	22.1	17.1
9	13.4	5.8	23.2	21.4	16.0
10	9.9	4.5	16.6	16.4	11.5
11	8.1	3.6	14.8	14.1	9.9
12	6.4	2.8	11.5	10.4	6.9
13	5.4	2.3	9.8	8.8	5.8
14	3.1	1.8	5.0	6.5	3.1
15	3.1	1.4	5.0	6.5	3.1

In de kolommen 4, 5 en 6 kunnen we nu aflezen dat het O.V. ongeveer 17%, 17% en 12% is voor respektievelijk de uniforme verdeling, de Gumbelverdeling en de Weibull (2) verdeling. Deze getallen wijken weinig af van die gegeven door Shapiro[2]. In appendix 2 wordt nog verder ingegaan op deze klassische toets.

3.2 Likelyhood variant van de X^2 toets

Deze toets is vrijwel identiek aan de voorgaande toets. Voor grote N zijn ze zelfs asymptotisch equivalent [7]. Echter wordt nu niet de grootheid X^2 berekend doch de logarithme van de zogenaamde likelyhoodratio L.L.R. welke als volgt bepaald wordt

$$\text{L.L.R.} = -2 \sum_{k=1}^{K} n_k \ \ln (p_k \cdot N/n_k)$$

in deze formule hebben n_k en p_k dezelfde betekenis als bij X^2. Voor de nulverdeling geldt hetzelfde als voor X^2 bij voorgaande toets. We hebben onderzocht of deze variant mogelijk betere eigenschappen heeft bij kleine N dan de voorgaande. Dit werd wederom gedaan met behulp van toevals trekkingen op de Wang tafelrekenmachine. De resultaten zijn samengevat in tabel 2. Ook hier is $\chi^2_{(1)}$ benadering niet bruikbaar. De grenswaarde voor een kritiek gebied ligt ergens tussen 3.5 en 4.0.

Tabel 2 Fraktie van L.L.R. groter of gelijk x in procenten.

x	Normaal 1000	$\chi^2_{(1)}$	Uniform (600)	Gumbel (500)	Weibull (2) (500)
0	100.0	100.0	100.0	100.0	100.0
0.5	70.2	47.9	84.0	80.2	72.2
1.0	61.9	31.7	79.2	74.0	62.4
1.5	39.6	22.1	61.0	53.4	37.0
2.0	32.2	15.7	56.0	47.2	31.6
2.5	22.3	11.4	42.0	35.4	23.8
3.0	17.9	8.3	38.9	27.0	19.4
3.5	12.2	6.1	24.7	21.8	12.0
4.0	7.3	4.5	17.8	15.8	8.4
4.5	5.5	3.4	17.8	15.8	8.4
5.0	5.5	2.5	17.8	14.0	6.2
5.5	4.9	1.9	15.2	11.6	5.2
6.0	3.5	1.4	9.0	9.0	4.4
6.5	2.2	1.1	9.0	6.0	4.4
7.0	1.9	0.8	3.5	5.6	1.8

Hieruit valt af te leiden dat het O.V. voor de drie reeds eerdergenoemde verdelingen nu ongeveer 20%, 18% en 10% is, wat voor de eerste twee alternatieven een lichte verbetering inhoudt vergeleken bij de vorige toets. Bedacht dient echter te worden dat dit steekproefresultaten zijn welke met een toevallige fout van zeker een paar procent behept zijn.

3.3 De X^2 aanpassingstoets met K = N

De voorschriften ten aanzien van het aantal deelruimten bij de X^2 toets zijn niet scherp omlijnd. Dit brengt een zekere vorm van willekeur in de toetsing, hierdoor is het soms mogelijk om door gunstige keuze van K en p_k de nulhypothese wel of niet te verwerpen. Door Kempthorne[4] is voorgesteld om hier wel een strakke regel voor te geven en wel een hele simpele, namelijk K = N en alle $p_k = 1/N$. Deze regel is in tegenspraak met de eis $N.p_k > 5$ want in dit geval is $N.p_k = 1$. Kempthorne toonde aan dat ook in dit geval de X^2 benadering voor de nulverdeling bruikbaar is hoewel voor X^2 slechts de even getallen en nul als uitkomsten kunnen voorkomen. Het O.V. is door Kempthorne alleen onderzocht voor enkele zeer scheve verdelingen. Daarom is ook deze toets nader onderzocht. De resultaten voor de nulverdeling en voor de alternatieven Uniform, Gumbel en Weibull (2) voor N=20 zijn gegeven in tabel 3. Het aantal vrijheidsgraden voor de benaderende X^2 verdeling is in dit geval (20 - 1 - 2 (aantal parameters) = 17.

Tabel 3 Fractie uitkomsten groter of gelijk n.

1 n	2 Normaal (1000)	3 $\chi^2_{(17)}$	4 Uniform (1000)	5 Gumbel (1000)	6 Weibull (2) (600)
10	96.2	91.5	99.1	97.9	97.5
12	89.7	80.0	96.9	94.4	92.1
14	78.4	66.0	91.7	86.4	83.9
16	64.2	52.0	83.3	72.7	69.7
18	49.4	38.5	71.7	58.1	54.0
20	35.4	27.5	59.5	42.2	40.0
22	25.1	18.0	44.9	28.6	28.6
24	15.8	11.7	34.2	19.7	19.0
26	10.0	7.2	24.4	12.7	12.8
28	5.3	4.5	17.2	9.3	9.2
30	3.3	2.7	11.1	5.7	6.0
32	2.1	1.5	8.2	3.8	4.3
34	0.9	0.8	5.0	2.2	2.5

Uit de tabel kolom 2 volgt dat we bij uitkomsten voor X^2 van 26 of groter tot verwerping moeten besluiten (kritiek gebied). De χ^2_{17} benadering (kol.3) leidt tot hetzelfde resultaat wat een bevredigende overeenstemming genoemd mag worden. De O.V. komen dan op 24%, 13% en 13% voor respektievelijk de alternatieven Uniform, Gumbel en Weibull (2).

3.4 De Neyman-Barton toets

Door Neyman en Barton[5] is voorgesteld om bij het toetsen van de aanpassing gebruik te maken van de transformatie

$$y_n = F_o\,(s_n,\ \theta_1, \ldots . \theta_r)$$

en van de zo verkregen steekproef $y_1, \ldots y_n$ te toetsen of deze steekproef

afkomstig kan zijn uit een uniforme verdeling. Het toetsen van de uniformiteit verschilt bij Neyman en Barton. We volgen nu verder Neyman.
Onafhankelijk van het feit of de nulhypothese wel of niet waar is, vormt de set y_1 t/m y_N een steekproef uit een verdeling op $[0, 1]$ met kansdichtheidsfunktie $f(y)$. Stel nu

$$f(y) = C \; e^{\sum_{k=o}^{\infty} a_k \cdot g_k(y)}.$$

waarin $g_k(y)$ een set orthonormale funkties op $[0, 1]$ (de Legendre polynomen) met $g_o(y) = 1$. Volgens de nulhypothese is $f(y)$ = konstant dus $a_1 = a_2 = \text{etc.} = 0$.
We toetsen nu deze nulhypothese door op basis van de steekproef a_1, enz. te schatten en daarna na te gaan of één van de a_k's signifikant afwijkt van nul. We kunnen dit niet van alle a_k's nagaan, we beperken ons daarom tot $k = 1,\ldots K$.
De keuze van K is vrij, dus ook hier geldt een zekere willekeur. Uit de theorie blijkt dat met

$$T = \sum_{k=1}^{K} \left\{ \frac{1}{N} \left(\sum_{n=1}^{N} g_k\,(y_n) \right)^2 \right\}$$

getoetst kan worden of één of meer van de a_k's afwijkt van nul.
Onder de nulhypothese blijft T vrij klein; bij afwijkingen wordt T groot. De toets heeft de volgende eigenschappen

a. <u>nulhypothese</u> Alleen enkelvoudige hypothesen zijn(in principe) toelaatbaar. Voor het terugtransformeren naar y_n dienen alle parameters bekend te zijn.

b. <u>toetsingsgrootheid</u>

$$T = \sum_{k=1}^{K} \left\{ \frac{1}{N} \left(\sum_{n=1}^{N} g_k(y_n) \right)^2 \right\}$$

c. <u>nulverdeling</u> De toetsingsgrootheid T volgt een χ^2 verdeling met K vrijheidsgraden.

d. <u>onderscheidend vermogen</u> Daar deze toets in deze vorm, waarin alleen enkelvoudige nul-hypothesen zijn toegelaten voor de praktijk niet interessant is, werd dit niet verder onderzocht. Omdat schrijver dezes echter de indruk heeft dat deze toets toch wel goede eigenschappen bezit werd onderzocht of de toets geschikt te maken was voor samengestelde hypothesen. Hierbij ontstond de volgende simpele toets:

a. <u>Nulhypothese</u>: De steekproef is afkomstig uit een normale verdeling.

b. Toetsingsgrootheid: Schat parameters μ en σ^2 uit de steekproef en transformeer naar y_1 t/m y_N volgens

$$y_n = F_o\,(x_n, \mu, \sigma^2)$$

en bereken

$$T = \frac{w_1}{N}\left(\sum_{n=1}^{N} g_1(y_n)\right)^2 + \frac{w_2}{N}\left(\sum_{n=1}^{N} g_2(y_n)\right)^2$$

hierin w_1 en w_2 nog konstanten afhankelijk van N.
Het verschil van Neyman en Barton zit dus in deze konstanten w_1 en w_2.

c. Nulverdeling: Bij benadering $\chi^2(2)$.

d. Onderscheidend vermogen: met een kritiek gebied van 10% en N = 20 werden de volgende O.V.'s gevonden tegen de alternatieven Uniform: 22%
Gumbel : 40%
Weibull (2) : 14%

Uit deze O.V.'s volgt dat deze toets ten opzichte van de klassieke toetsen een beter onderscheidend vermogen heeft. De toets van Kempthorne is ongeveer even goed voor de alternatieven Uniform en Weibull (2) doch beduidend slechter voor Gumbel.

N.B. 1: De konstanten in de toetsingsgrootheid werden voor N = 20 met behulp van toevalstrekkingen op de X8 bepaald. Deze konstanten zijn afhankelijk van N, het is niet bekend op welke manier.

N.B. 2: De O.V.'s hier gegeven berusten op een klein aantal steekproeven dus relatief grote afwijkingen zijn mogelijk.

3.5 De toets van Shapiro

Voor de theorie betreffende deze toets wordt verwezen naar Shapiro[1] en voor een uitvoerig onderzoek betreffende het onderscheidend vermogen naar Shapiro[2]. We geven hier alleen een korte beschrijving.

a. Nulhypothese: Alleen de samengestelde hypothese "De verdeling behorend bij x_1 t/m x_N is normaal" kan worden getoetst.

b. Toetsingsgrootheid:

$$W = \left(\sum_{n=1}^{N} a_n\, x_n^*\right)^2 \Big/ \left(\sum_{n=1}^{N} (x_n - m)^2\right)$$

waarin $x_1^*, \ldots, x_N^*$: geordende steekproef.

$$m = \frac{1}{N}\sum_{n=1}^{N} x_n$$

a_n : een aantal konstanten welke afhankelijk zijn van n en N. Ze zijn door Shapiro[1] theoretisch berekend en in tabelvorm bekend van N = 3 t/m N = 50.

c. Nulverdeling: In tabelvorm gegeven door Shapiro[1]. W is altijd kleiner dan 1. We verwerpen indien W te klein is.

d. Onderscheidend Vermogen: Bij een kritiek gebied van 10% en N = 20 vond Shapiro voor het alternatief Uniform dat het O.V. = 39% en voor het alternatief Weibull (2) een O.V. = 24% (Bij N = 50 resp. 96 en 59%). De Gumbelverdeling is niet als alternatief onderzocht door Shapiro, wij vonden hiervoor een O.V. van 40% (N = 20, 10%). De prestaties van deze toets steken ver boven die van de voorgaande toetsen uit. Een groot nadeel van deze toets is echter dat de konstanten a_n niet op een eenvoudige manier uit n en N zijn te berekenen. Alleen in tabelvorm zijn ze bekend voor N $\leqslant$ 50. Om deze bezwaren te omzeilen zijn een aantal varianten van deze toets ontworpen waarvan we er één in 3.6 zullen bespreken.
Een ander nadeel van de toets van Shapiro is dat slechts één nulhypothese is toegelaten terwijl aanpassing aan andere nulhypothesen een uiterst moeilijke zaak is. De voorgaande toetsen bieden in dit opzicht meer mogelijkheden.

3.6 Toets van d'Agostino

Deze toets is een variant van de toets van Shapiro waarbij de a_n wel eenvoudig zijn te berekenen (zie d'Agostino [8]).

a. Nulhypothese: Evenals bij Shapiro is alleen de samengestelde hypothese van een normale verdeling toegestaan.

b. Toetsingsgrootheid:

$$D = \left(\sum_{n=1}^{N} \left[(n - \tfrac{1}{2}(N+1)).x_n^*\right]\right) / (N^2.s)$$

waarin s: standaardafwijking steekproef.

c. Nulverdeling: In tabelvorm gegeven door d'Agostino zie literatuur [8][9].

d. Onderscheidend Vermogen: Door d'Agostino zelf onderzocht voor N = 50. In het algemeen iets slechter dan de toets van Shapiro doch nog beter dan de klassieke toetsen. Deze toets is door ons niet nader onderzocht.

3.7 Toets van Kolmogorov-Smirnov

Deze toets en de hierna volgende zijn gebaseerd op de steekproefverdelingsfunktie $S_N(x)$ welke als volgt is gedefinieerd.

$$S_N(x) = \begin{cases} 0, & x < x_1^* \\ \frac{n}{N}, & x_n^* \leq x < x_{n+1}^* \\ 1, & x_N^* \leq x \end{cases}$$

waarin x_1^*, ... x_N^* weer de geordende steekproef. De toets van Kolmogorov-Smirnov heeft de volgende eigenschappen,

a. Nulhypothese: Alle nulhypothesen zijn toegestaan.

b. Toetsingsgrootheid:

$$T = \max_x \left| F_o(x) - S_N(x) \right|$$

waarin F_o de nulverdeling is met eventueel geschatte parameters.

c. Nulverdeling Voor enkelvoudige hypothesen is de nulverdeling onafhankelijk van de nulhypothese en in tabelvorm bekend, zie bijv. Conover[6] tabel 14. Voor samengestelde hypothesen wordt de nulverdeling afhankelijk van de nulhypothese. Door Liliefors is de nulverdeling bepaald voor de nulhypothese "De verdeling is normaal", zie bijv. Conover[6] tabel 15.

d. Onderscheidend Vermogen: Door Shapiro[2] is het O.V. bepaald voor de enkelvoudige hypothese (Normale verdeling).
Bij N = 20 en kritiek gebied van 10% werd voor het alternatief Uniform :19% gevonden en voor Weibull (2) : 15%. Voor de samengestelde hypothese geeft Louter [3] het O.V. voor een aantal alternatieven bijv. Uniform: 21%. N = 20, 10%). Uit deze cijfers blijkt dat deze simpele toets het niet slechter doet dan de klassieke toetsen.

N.B. Door de eenvoud van de toetsingsgrootheid kunnen de tabellen voor de K.S.toets (enkelvoudig) ook nog op een andere manier gebruikt worden. Stel dat de tabel aangeeft voor zekere N dat met 95% kans $T < d_\alpha$ als nulhypothese waar is dan kunnen we omgekeerd stellen dat met 95% kans de ware verdeling in een band ligt ter breedte 2 d_α rond de steekproefverdelingsfunktie. Dit geeft ons een eenvoudige en altijd bruikbare betrouwbaarheidsband voor de steekproefverdelingsfunktie.(zie ook fig.6).

3.8 De Kuiper-toets

Deze toets is een eenvoudige variant van de Kolmogorov-Smirnov toets.

a. Nulhypothese: Alle nulhypothesen toegestaan.

b. Toetsingsgrootheid:

$$T = \max_x \left(F_o(x) - S_N(x)\right) - \min_x \left(F_o(x) - S_N(x)\right).$$

wederom in F_o eventueel schatters van parameters gebruiken.

c. Nulverdeling: Voor enkelvoudige hypothesen wederom onafhankelijk van de nulhypothese en in tabelvorm gegeven door Stephens [10]. Door Louter [3] wordt de nulverdeling gegeven voor de samengestelde hypothese "De verdeling is normaal".

d. Onderscheidend Vermogen: Voor enkelvoudige hypothesen onbekend. Voor samengestelde hypothese (normaal) geeft Louter[3] voor bijv. Uniform: 27% bij N=20 en kritiek gebied 10%. Uit eigen onderzoek blijkt voor alternatief Gumbel is het O.V. ongeveer 30% (N = 20, 10%) en voor Weibull (2) onder dezelfde omstandigheden ongeveer 20%. Uit het hiervoor geciteerde artikel blijkt tevens dat het O.V. redelijk snel toeneemt met de grootte van de steekproef, wat deze toets des te waardevoller maakt.

3.9 De toets van Cramèr-von Mises

Deze toets is wederom gebaseerd op de steekproefverdelingsfunktie. We geven hier slechts een korte beschrijving

a. Nulhypothese: Elke nulhypothese is toegestaan.

b. Toetsingsgrootheid:

$$W^2 = \int_{-\infty}^{+\infty} \left\{ S_n(x) - F_o(x) \right\}^2 d\, F_o(x)$$

of eenvoudiger

$$NW^2 = \frac{1}{12N} + \sum_{n=1}^{N} \left\{ F_o(x_n^*) - \frac{2n-1}{2N} \right\}^2$$

c. Nulverdeling: Alleen bekend voor enkelvoudige hypothesen. Anderson[11]

d. Onderscheidend Vermogen: Onderzocht door Shapiro[2] voor enkelvoudige hypothese. Er werden geen essentiële verschillen ten opzichte van Kolmogorov-Smirnov gevonden. Ditzelfde geldt voor de zogenaamde gewogen Cramèr-von Mises waarbij de toetsingsgrootheid gegeven wordt als

$$W^2 = \int_{-\infty}^{+\infty} (S_n(x) - F_o(x))^2 \quad \frac{dF_o(x)}{F_o(x)[1-F_o(x)]}$$

3.10 Samenvatting

In de volgende tabel zijn de resultaten met betrekking tot het onderscheidend vermogen van de diverse toetsen samengevat. De hier gegeven O.V.'s hebben betrekking op de situatie N = 20 en kritiek gebied van 10%. De toetsen zijn aangeduid met het nummer van de paragraaf. De O.V.'s zijn gegeven in procenten.

Alternatief	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Uniform	17	20	24	22	39	--	21	27	--
Gumbel	17	18	13	40	40	--	--	30	--
Weibull(2)	12	10	13	14	24	--	--	20	--

In dit overzicht zijn zeker niet alle toetsen ter sprake gekomen. Of een toets wel of niet opgenomen is, werd door een complex van (soms subjectieve)faktoren bepaald.

Benadrukt dient te worden dat de hier gegeven waarden voor het O.V. berusten op steekproeven en dus nog behept zijn met een toevallige fout die zeker een paar procent kan zijn.

4. Onderscheidend vermogen bij grotere steekproeven

In de voorgaande paragraaf zijn de diverse toetsen alleen maar vergeleken op basis van hun onderscheidend vermogen bij N=20. Algemeen geldt echter dat het onderscheidend vermogen van een toets toeneemt met toenemende grootte van de steekproef. Op grond hiervan kan men verwachten dat wanneer de grootte van de steekproef een bepaalde grens overschrijdt, de klassieke X^2-toets als voldoend onderscheidend beschouwd kan worden. In deze paragraaf zullen we een schatting maken van deze grensgrootte op basis van theoretische overwegingen.

We maken bij deze schatting gebruik van de volgende theoretische benaderingen:

A. Wanneer de alternatieve hypothese niet al te ver van de nulhypothese ligt, geldt voor de alternatieve verdeling van de toetsingsgrootheid X^2 van de X^2-toets, dat deze benaderd kan worden met een niet-centrale χ^2 verdeling met niet-centraliteitsparameter λ waarvoor geldt

$$\lambda = N \sum_{k=1}^{K} \frac{(p_{a,k} - p_{o,k})^2}{p_{o,k}}$$

waarin

$p_{a,k}$: de kans om een waarneming in de k^e deelruimte te vinden indien de alternatieve hypothese waar is.

$p_{o,k}$: deze zelfde kans indien de nulhypothese waar is.

(zie literatuur 5), deel II, 1e editie, pag.436)

B. De niet-centrale χ^2 verdeling kan als volgt benaderd worden:

$$P(X^2(\lambda, \nu) > x) \approx P(X^2(\nu^*) > \frac{x}{\theta})$$

waarbij geldt

$$\theta = 1 + \frac{\lambda}{\nu + \lambda}$$

en $\nu^* = \nu + \dfrac{\lambda^2}{\nu + 2\lambda}$

(zie literatuur 5), deel II, 1e editie, pag.229)

Met behulp van deze benaderingen hebben we, voor een aantal kombinaties van N en K, het onderscheidend vermogen bepaald van de χ^2 toets in de volgende situatie:

·Nulhypothese: normale verdeling, $\mu = 0$ en $\sigma^2 = 1$.
Alternatief : Gumbel verdeling, $u = -0.45$ en $x = 1{,}28$
of $\mu = 0$ en $\sigma^2 = 1$.

Kritiek gebied: 10%

N	K	O.V.
20	4	15%
50	10	35%
100	10	60%
200	10	98%
100	20	50%
200	20	85%

De resultaten zijn samengevat in de tabel hiernaast, aan de uitkomsten mag zeker geen al te grote nauwkeurigheid worden toegekend, ze geven echter duidelijk aan dat de klassieke X^2-toets bij $N > 200$ tot hoge prestaties komt. Uit deze tabel blijkt tevens dat ook keuze van K invloed heeft op het onderscheidend vermogen. Een algemene konklusie is op basis hiervan niet mogelijk, wel kan gesteld worden dat hoge K-waarden nadelig werken.

5. Konklusie

In voorgaande studie is vrij veel aandacht besteed aan de klassieke X^2 toets en een aantal varianten daarvan. De reden hiervoor was dat deze toetsen verdelingsvrij en parametervrij zijn, wat ze bijzonder geschikt maakt voor algemene toepassing. Op grond van deze studie moet echter gesteld worden dat het onderscheidend vermogen bij kleine steekproeven (N=20) in vele gevallen zo gering is dat ze niet als kriterium over wel of niet aanpassen gebruikt kunnen worden. Uit Shapiro[2)] blijkt tevens dat lang niet altijd zal gelden dat het onderscheidend vermogen snel beter wordt bij toenemende N. Wil men dus bij kleine steekproefgrootte (20-200) (zie ook 4) een nulhypothese kritisch toetsen dan zal men gebruik dienen te maken van andere toetsen. In het geval dat de nulhypothese luidt "De verdeling is normaal" is dan de keuze eenvoudig. Grote voorkeur genieten dan de toetsen van Shapiro en d'Agostino. Ook voor enkelvoudige hypothesen zijn makkelijk andere toetsen te vinden. Zowel de toets van Neyman en Barton als de toets van Kuiper zijn in deze situatie verdelingsvrij. Vermoedelijk zal de toets van Neyman en Barton dan de beste zijn, dat is echter hier niet onderzocht.

In het geval dat men een algemene samengestelde hypothese wenst te toetsen zijn geen alternatieve mogelijkheden beschikbaar, tenzij men een toets voor de gegeven situatie kan aanpassen. Wederom valt dan de keus op of Neyman en Barton of Kuiper. De eerste hiervan vergt dan aanmerkelijk meer studie, de tweede kan niet zonder meer worden toegepast omdat de nulverdelingen niet bekend zijn, het zou daarom ten zeerste aanbeveling verdienen om deze nulverdelingen te gaan vervaardigen.

Vanaf N=200 is de keuze niet meer zo kritisch en kan in het algemeen de klassieke X^2 toets worden toegepast. Ook de andere toetsen blijven natuurlijk toepasbaar; in veel gevallen zal dit echter gepaard gaan met problemen van zowel rekentechnische als theoretische aard.

Appendix 1

In deze appendix zal nader worden ingegaan op de eigenschappen van de in het rapport genoemde kansverdelingen en op het genereren van steekproeven hieruit. Een kansverdeling kunnen we karakteriseren door de kansdichtheidsfunktie (d.f.) $f(x, \Theta_1, \ldots \Theta_I)$ verdelingsfunktie (v.f.) $F(x, \Theta_1, \ldots \Theta_I)$. Hierin is x de variabele en $\Theta_1, \ldots \Theta_I$ noemen we de parameters van de kansverdeling. De samenhang tussen deze twee funkties wordt gegeven door

$$F(x, \Theta_1, \ldots \Theta_I) = \int_{-\infty}^{x} f(x', \Theta_1 \ldots \Theta_I)\, dx'$$

Verder geldt nog

$$F(-\infty, \Theta_1 \ldots \Theta_I) = 0 \text{ en } F(+\infty, \Theta_1 \ldots \Theta_I) = 1.$$

De betekenis van de funkties volgt uit de volgende twee relaties:

$$P(\underline{x} < x) = F(x, \Theta_1 \ldots \Theta_I)$$

en $P(x < \underline{x} < x + dx) = f(x, \Theta_1 \ldots \Theta_I)\, dx$

waarin dan de uitdrukking $P(\underline{x} < x)$ gelezen moet worden als "De kans P dat we bij een trekking uit de gegeven verdeling een waarde vinden voor $\underline{x}$ die kleiner is dan x" en de tweede uitdrukking $P(x < \underline{x} < x+dx)$ op overeenkomstige wijze.

Belangrijk is nog de definitie van de volgende grootheden:

1. De verwachtingswaarde van $\underline{x} = E(\underline{x}) = \mu = \int_{-\infty}^{+\infty} f(x, \Theta_1 \ldots \Theta_I).\, x.dx$
   (wordt ook wel gemiddelde van $\underline{x}$ genoemd).

2. De variantie van $\underline{x} = \sigma_x^2 = \int_{-\infty}^{+\infty} f(x, \Theta_1 \ldots \Theta_I)(x - \mu)^2\, dx.$

Deze twee grootheden bestaan vrijwel altijd doch zijn niet altijd even geschikt om er de verdeling mee vast te leggen.

Hierna zullen we een overzicht geven van de in het rapport genoemde verdelingen. Van de niet-normale verdelingen hebben we bovendien ter vergelijking de verdelingsfunktie uitgezet op normaal waarschijnlijkheidspapier tezamen met v.f. van de normale verdeling met dezelfde μ en σ^2.

1. De normale verdeling

   Dit is de meest bekende en meest gebruikte verdeling, die wordt gespecificeerd door de parameters μ en σ^2.

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

$F(x, \mu, \sigma^2)$ = niet in gesloten vorm bekend, wel in tabelvorm.

Indien we de v.f. van de normale verdeling uitzetten op normaal waarschijnlijkheidspapier vinden we een rechte lijn.

2. Gumbel verdeling

Deze verdeling die vaak gebruikt wordt als de verdeling van extreme waarden, wordt gespecificeerd door de parameters x en u. De d.f. en v.f. zien er als volgt uit:

$$f(x,\ x,\ u) = \alpha\ e^{-\alpha(x-u)}\ e^{-e^{-\alpha(x-u)}}$$

$$F(x,\ x,\ u) = e^{-e^{-\alpha(x-u)}}$$

verder geldt $\mu = u + \frac{0.57721566}{\alpha}$

$$\sigma^2 = \frac{\pi^2}{6\alpha^2}$$

In figuur 2 hebben we een Gumbel verdeling met $\mu = 0$ en $\sigma^2 = 1$ of $u = -0{,}45$ en $\alpha = 1{,}28$ uitgezet tezamen met een normale verdeling met $\mu = 0$ en $\sigma^2 = 1$. Belangrijk is om op te merken dat deze twee verdelingen vooral in de staarten sterk verschillen. Ook in figuur 1 vinden we de Gumbel verdeling.

3. Weibull verdeling

Dit is een tamelijk algemeen toepasbare verdeling, welke eveneens wordt gespecificeerd door twee parameters. De eerste, aangeduid met λ, is een simpele schaalparameter. De tweede, aangeuid met k, een vorm-parameter. De aanduiding in het rapport Weibull (2) geeft aan dat we een Weibull verdeling met vorm parameter k = 2 gebruiken. De d.f. en v.f. zien er als volgt uit:

$$\left.\begin{aligned} f(x, \lambda, k) &= \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^{k}} \\ F(x, \lambda, k) &= 1 - \exp^{-\left(\frac{x}{\lambda}\right)^{k}} \end{aligned}\right\}$$

Verder geldt $\mu = \lambda\ \Gamma\left(1 + \frac{1}{k}\right)$

$$\sigma^2 = \lambda^2 \left\{ \Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{v}{k}\right) \right\}$$

(Voor toepassing van Weibull verdeling zie bijv. P.J.Rijkoort 12)).

In figuur 3 is de Weibull verdelingsfunktie uitgezet voor $\lambda = 1$ en $k = 2$ op normaal waarschijnlijkheidspapier. Eveneens is hier de rechte lijn gegeven behorend bij een normale verdeling met $\mu = 0.8862$ en $\sigma^2 = 0.2146$.

4. Uniforme verdeling

Dit is de meest eenvoudige verdeling met echter weinig praktische toepassingen. Deze verdeling wordt gespecificeerd met twee parameters respektievelijk ondergrens en bovengrens van de verdeling. De beschrijvende funkties zien er als volgt uit:

$$f(x, a, b) = \begin{cases} 0 & x < a \\ 1/(b-a) & a \leqslant x \leqslant b \\ 0 & b < x \end{cases}$$

$$F(x, a, b) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \leqslant x \leqslant b \\ 1 & b < x \end{cases}$$

verder geldt $\mu = (b-a)/2$, $\sigma^2 = (b-a)^2/12$.
De verdeling is weergegeven in fig.4 met $a = -1$ en $b = +1$ tezamen met de normale verdeling met $\mu = 0$ en $\sigma^2 = 1/3$.

5. De $\chi^2_{(v)}$-verdelingen

De χ^2 verdelingen vormen een klasse van bekende en belangrijke theoretische verdelingen. Het zijn van de normale verdeling afgeleide verdelingen en als volgt gedefinieerd. Stel we voeren een stochastische variabele in volgens

$$\underline{X}^2 = \sum_{n=1}^{N} x_n^2 ,$$

waarin x_n trekkingen zijn uit de normale verdeling met $\mu = 0$ en $\sigma^2 = 1$, dan is X^2 per definitie χ^2 verdeeld met $\nu = N$ vrijheidsgraden.
De d.f. van deze verdeling ziet er als volgt uit

$$f(X^2, \nu) = \frac{1}{2^{\frac{\nu}{2}}\,\Gamma(\frac{\nu}{2})} (X^2)^{\frac{\nu}{2}-1} e^{-\frac{1}{2}X^2} \qquad X^2 \geqslant 0$$

$F(X^2, \nu)$ is alleen in tabelvorm bekend.

Verder geldt heel eenvoudig

$$\mu = \nu \text{ en } \sigma^2 = 2\nu$$

In fig.5 is een verdelingsfunktie gegeven voor $\nu = 4$ tezamen met een normale verdeling met $\mu = 4$ en $\sigma^2 = 8$.

Voor het genereren van trekkingen uit verdelingen is het volgende van belang. Indien we bij een trekking x_i uit een gegeven verdeling met v.f. $F(x)$ een nieuwe grootheid definiëren

$$z_i = F(x_i) \text{ of in statistische notatie}$$

$$\underline{z} = F(\underline{x})$$

dan is $\underline{z}$ een variabele die, indien de verdeling van x kontinu was, automatisch uniform verdeeld is met $a = 0$ en $b = 1$ (merk op $\underline{z} = P(\underline{x} < x)$. Deze relatie mogen we ook omgekeerd toepassen. Indien we z_i trekken en de vergelijking $z_i = F(x_i)$ oplossen, vinden we een x_i die gezien mag worden als een trekking uit de verdeling met v.f. $F(x)$.

De Wang tafelrekenmachine beschikt over een procedure die onafhankelijke trekkingen levert uit de uniforme verdeling met $a = -1$ en $b = 1$. Door het nemen van absolute waarde wordt dit een uniforme verdeling met $a = 0$ en $b = 1$. De trekkingen x voor de in dit rapport gebruikte verdelingen werden daarna met de volgende formules uit de trekking z berekend.

a. <u>Normale verdeling</u>:

Voor $z \geq 0.5$ dan $t = \sqrt{-2\ln(1-z)}$

en

$$x = t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}$$

Voor z kleiner dan 0.5 wordt de x berekend door spiegelen ten opzichte van 0.5.

b. <u>Gumbel verdeling</u>:

$x = -\ln(-\ln z)$.

c. <u>Weibull verdeling</u>: <u>k=2</u>

$$x = \sqrt{-\ln(1-z)}$$

Literatuur

1. Shapiro, S.S. en Wilk, M.B.(1965) An analysis of variance test for normality (complete samples). Biometrika 52, 591-611.

2. Shapiro, S.S., Wilk, M.B. en Chen, Mrs.H.J. (1968) A comparative study of various test for normality. Jasa 1963, 1343-1372.

3. Louter, A.S.en Koerts, J. On the Kuiper test for normality with mean and variance unknown. Statistica Neerlandica 24(1970) nr.2.

4. Kempthorne, O. The classical problem of inference - Goodness of Fit. Fifth Berkeley Symposion.

5. Kendall, M.G. en Stuart, A. The advanced Theory of Statistics, Vol.2, (First ed.) 444.

6. Conover, W.J. Practical Non-parametric Statistics.

7. Leppink, G.J. Standaarddictaat Statistiek R.U. te Utrecht 1970.

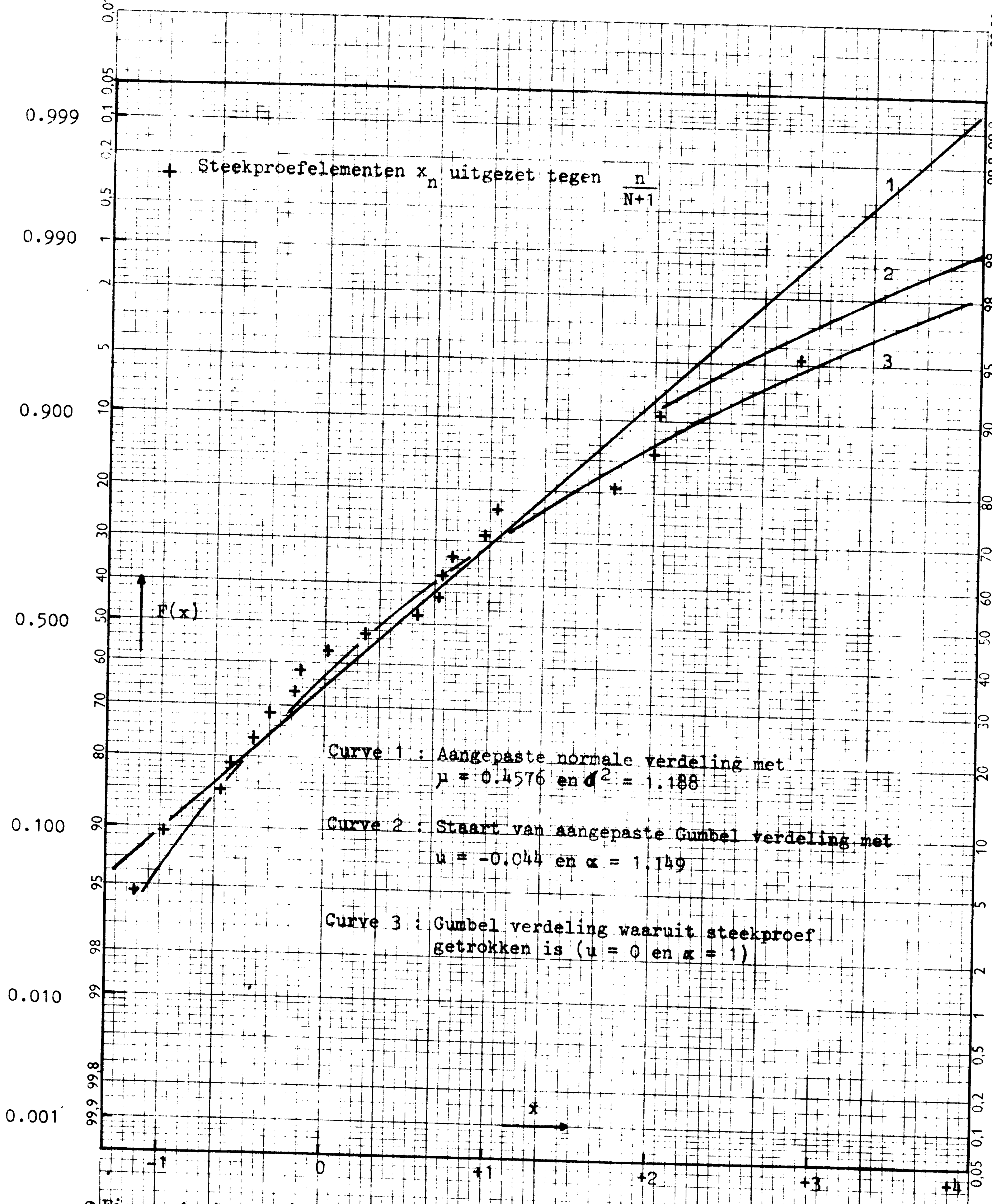
8. d'Agostino, R.B. (1971) An omnibus test of normality for moderate and large size samples. Biometrika 58, 341-348.

9. d'Agostino, R.B. (1972) Small sample probability points for the D test of normality. Biometrika 59, 219-221.

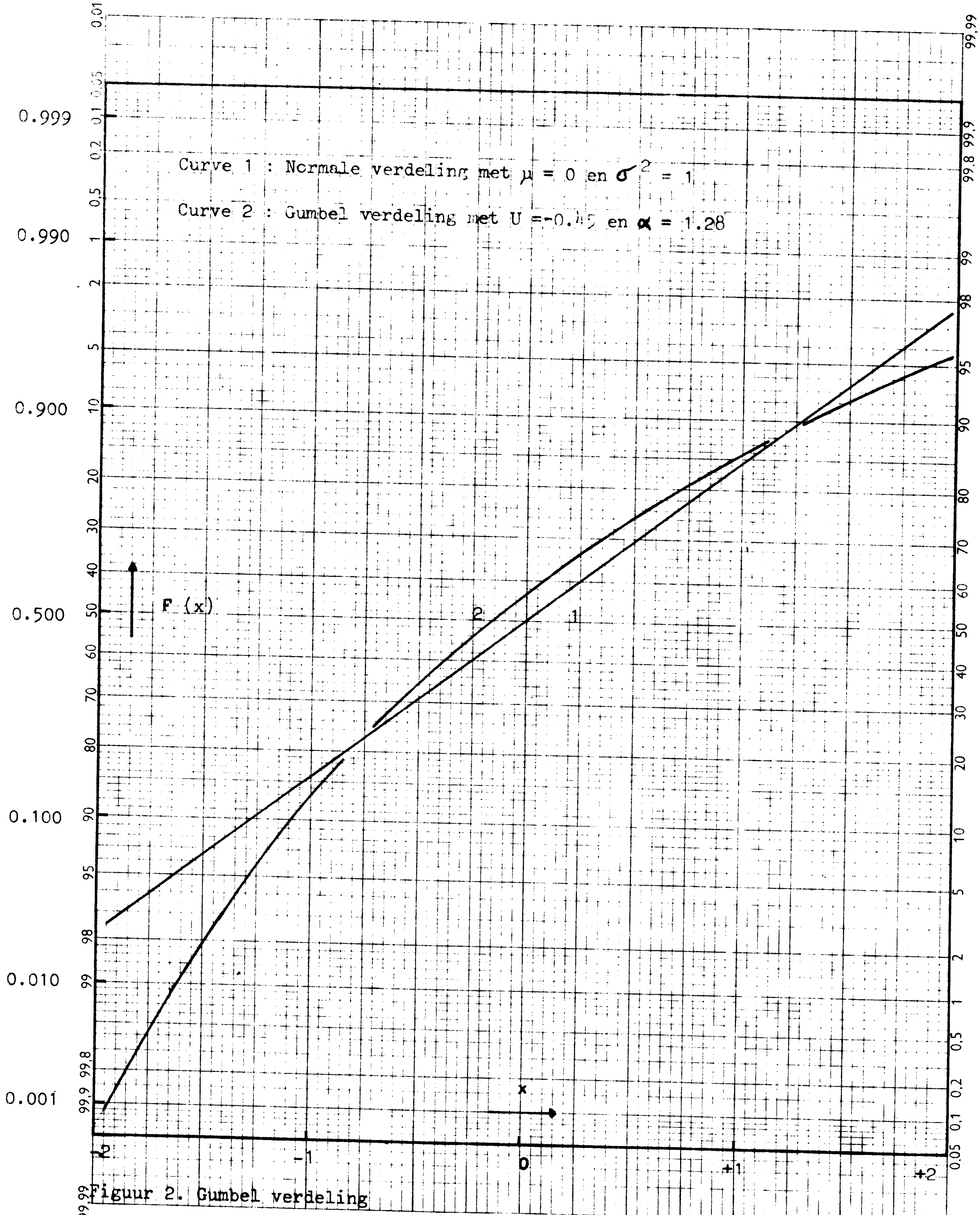
10. Stephens, M.A. (1965) The goodness of fit Statistic V_N: distribution and significance points. Biometrika 52, 309-321

11. Anderson, T.W. and Darling, D.A. (1952) Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. The Annals of Mathematical Statistics (23), 193-212.

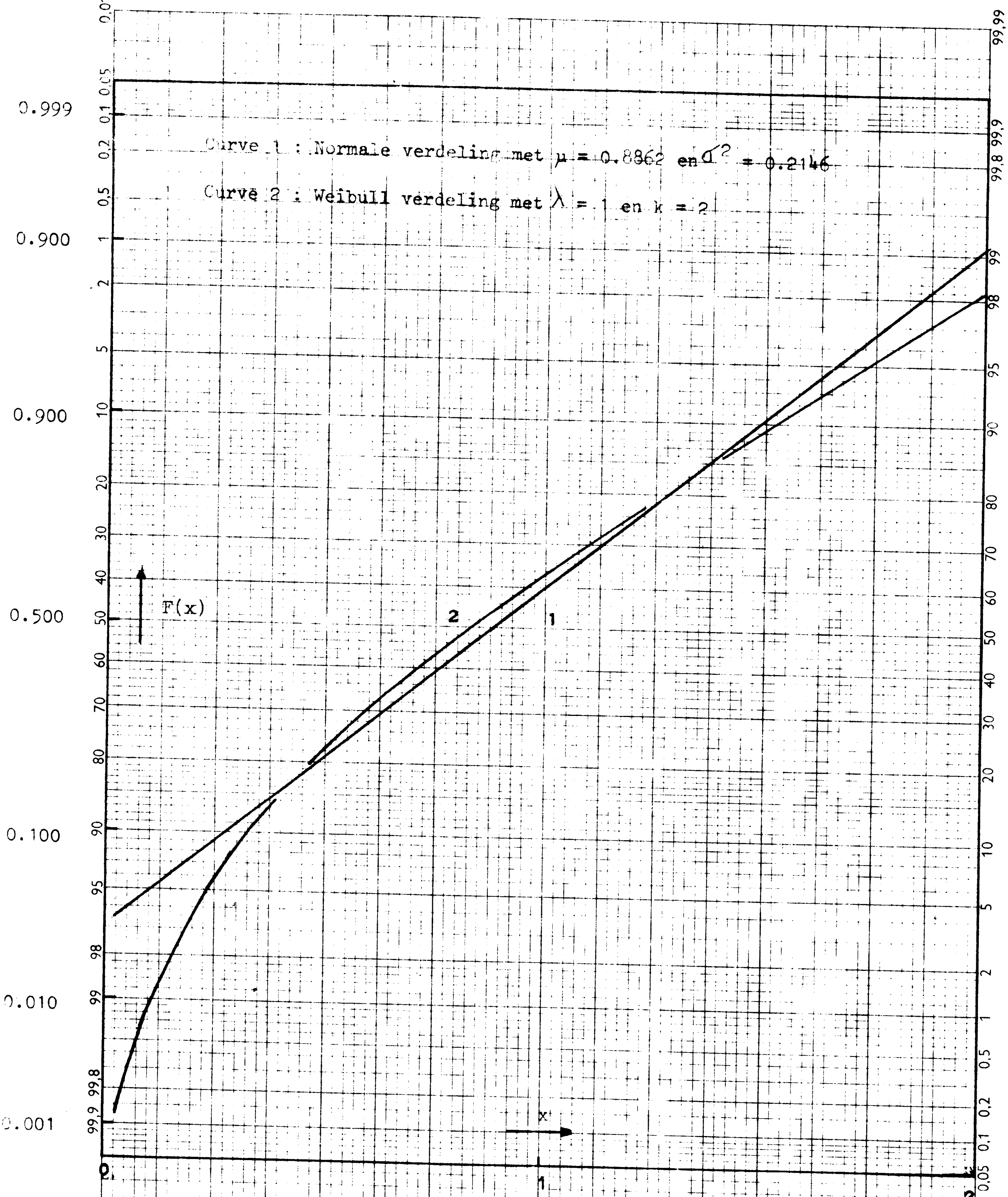
12. Rijkoort, P.J. De variatie van de windsnelheidsverdeling volgens waarnemingen op 10, 40 en 80 m hoogte aan de meteorologische meetmast te Vlaardingen. KNMI W.R. 72-4.



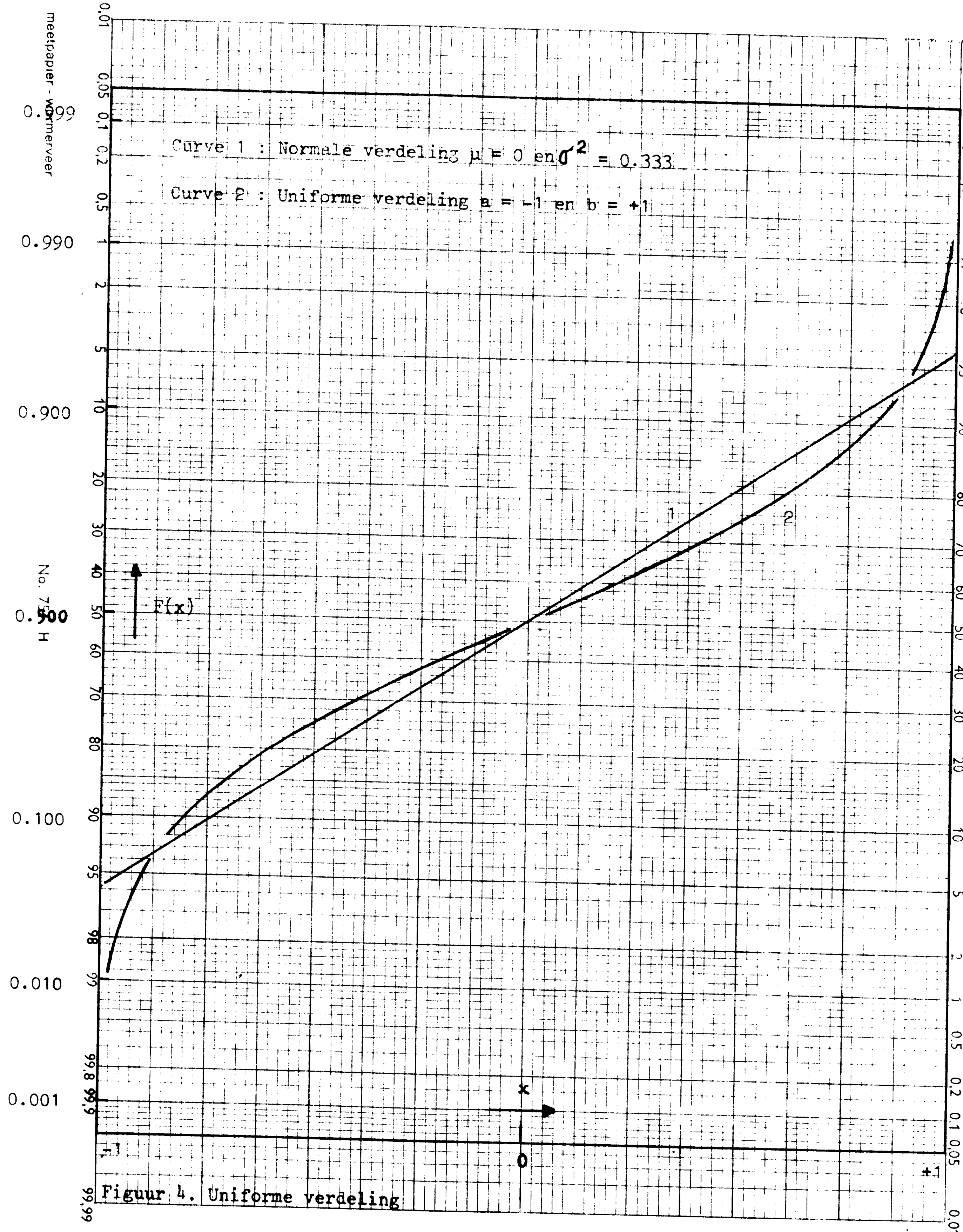
Figuur 1. Aanpassingen op steekproef (N = 20) uit standaard Gumbel verdeling



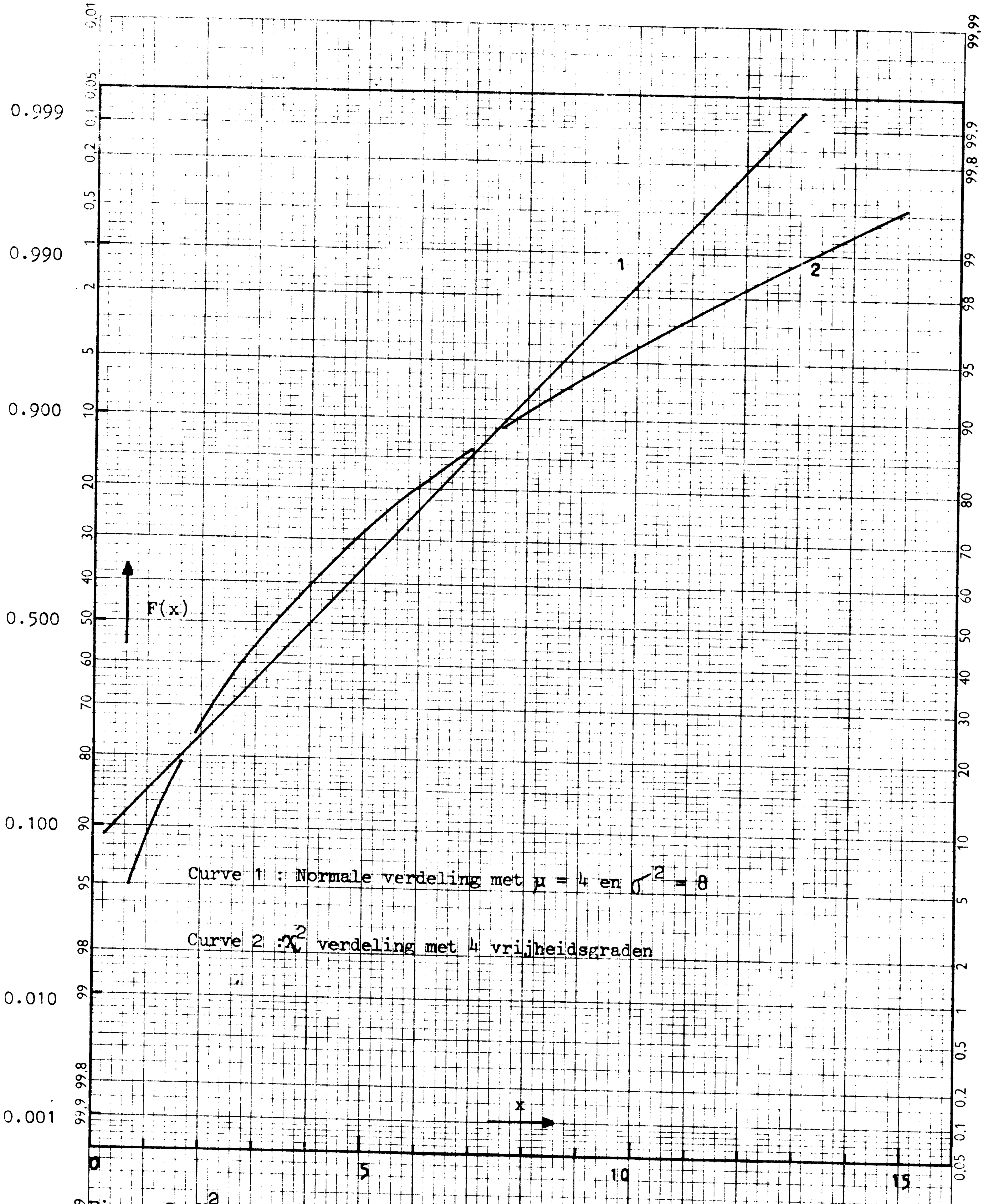
Figuur 2. Gumbel verdeling



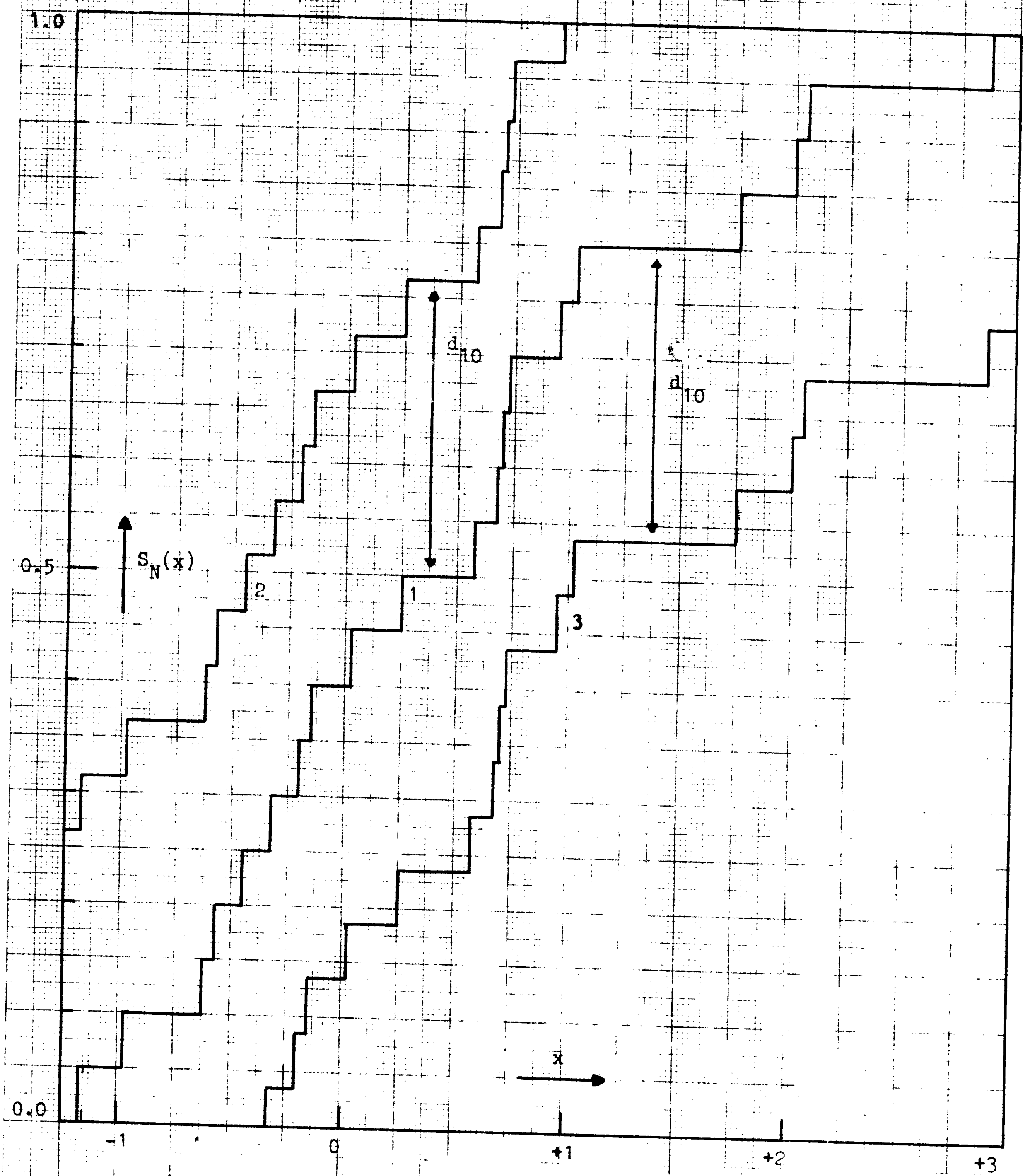
Figuur 3. Weibull verdeling.



Figuur 4. Uniforme verdeling



Figuur 5. χ_4^2 verdeling



Stapjescurve 1: steekproefverdelingsfunktie van de steekproef uit fig.1

Stapjescurve 2: Bovenste grens } volgens Kolmogorov

Stapjescurve 3: Onderste grens

Er geldt dus "Met 90% kans ligt de werkelijke F(x) in zijn geheel binnen de band gevormd door 2 en 3"

Figuur 6 Kolmogorov betrouwbaarheidsband