Een adaptieve modificatie van de SAS-routine NPARIWAY

TUE-RC 78322

Eindhoven University of Technology
Computing Centre Note 43

Een adaptieve modificatie
van de SAS-routine NPAR1WAY

Jan B. Dijkstra

Samengesteld voor de bijeenkomst van de Benelux Users Group SAS op
15 en 16 november 1988 in Scheveningen.

oktober 1988.

Een adaptieve modificatie van de SAS-routine NPAR1WAY

Jan B. Dijkstra

Samenvatting

De SAS-routine NPAR1WAY bevat een aantal verdelingsvrije toetsen voor de hypothese dat k steekproeven uit dezelfde continue verdeling komen. Middels een optie kan men een toets selecteren en de default-situatie is dat ze allemaal worden uitgevoerd.
Zonder voorkennis met betrekking tot de verdeling zal men in de praktijk geen verantwoorde keuze kunnen maken. En inspectie van de resultaten voor alle toetsen kan leiden tot tegenstrijdige conclusies. Waartoe moet men bijvoorbeeld besluiten als de Van der Waerden toets de hypothese verwerpt, terwijl de mediaantoets deze accepteert?
Er wordt een wijziging van NPAR1WAY middels een adaptieve procedure voorgesteld die resulteert in: (1) minder uitvoer, (2) geen inconsistenties en (3) asymptotisch een groter onderscheidend vermogen dan alle nu opgenomen toetsen voor een ruime klasse van symmetrische verdelingen. Een en ander wordt toegelicht aan de hand van een voorbeeld.

1. Inleiding

Deze notitie gaat over de nulhypothese H_0: $\mu_1 = \cdots = \mu_k$. Hierbij stelt iedere μ_i een populatiegemiddelde voor. Toetsen van de hypothese komt overeen met verificatie van het vermoeden dat k steekproeven afkomstig zijn uit populaties met dezelfde verwachting. De waarnemingen binnen de steekproeven worden aangeduid als x_{ij}. Hierbij geeft de index i de steekproef aan waar de waarneming toe behoort en j dient voor de nummering binnen de steekproeven. Er geldt $j = 1, \cdots, n_i$ dus er wordt van de steekproeven niet verwacht dat ze een even groot aantal waarnemingen bevatten. De concrete waarnemingen zijn realisaties van de stochastische variabelen X_{ij} waarvoor geldt $\mu_i = EX_{ij}$. De operator E duidt hierbij de verwachting aan.
Voor het toetsen van H_0 wordt veelal gebruik gemaakt van ANOVA of GLM. Deze routines eisen echter dat binnen het model $X_{ij} = \mu_i + E_{ij}$ de fouten E_{ij} onafhankelijk normaal verdeeld zijn met verwachting 0 en constante variantie σ^2. In de praktijk is hier niet altijd aan voldaan. Niet-normaliteit is soms op te lossen met tranformaties. Een van de machtigste hulpmiddelen hierbij is de Box-Cox (1964) transformatie die helaas niet in SAS is geimplementeerd. Maar zelfs dit hulpmiddel is niet altijd succesvol.
Achterin is een voorbeeld opgenomen van een SAS-sessie. Het betreft de behandeltijd van

Rekencentrum, Technische Universiteit, 5600 MB Eindhoven, telefoon 040 - 474535/474089

grotere software-problemen op het gebied van de statistiek bij het Rekencentrum van de TUE. De tijd is uitgedrukt in kwartieren; problemen die minder dan een half uur kostten zijn buiten beschouwing gelaten. Bovendien zijn de gegevens beperkt tot de tweedelijns consultatie. Problemen die aan de algemene inlichtingenbalie konden worden opgelost zijn niet opgenomen.
Er is begonnen met een klassieke verwerking van de gegevens. De resultaten zijn niet in het voorbeeld opgenomen, maar zullen nu worden samengevat. GLM leverde een overschrijdingskans op van 0.0351. Vervolgens werd middels de routine UNIVARIATE getoetst of de residuen redelijk door een normale verdeling beschreven konden worden. De overschrijdingskans bij deze toetsing bleek 0.0152 te zijn. Buiten SAS om werd vervolgens gezocht naar de gunstigste transformatie naar normaliteit binnen de Box-Cox familie. Deze werd gevonden met de TUE-routinebibliotheek PP-4 en op de oorspronkelijke data toegepast. Op de getransformeerde data werd GLM losgelaten en de residuen werden weer met UNIVARIATE onderzocht. De overschrijdingskans bij een toets op normaliteit bleek gestegen te zijn tot 0.0276 en dat is onvoldoende. Met een transformatie kan de nonnormalitiet dus niet worden gecompenseerd en dus moet de hypothese op een andere manier getoetst worden.

2. Verdelingsvrije toetsen

De toetsen van NPAR1WAY zijn verdelingsvrij. Dat wil zeggen (in deze context) dat de nulhypothese stelt dat alle X_{ij} dezelfde continue verdeling F volgen zonder dat aan F bij voorbaat speciale eisen worden gesteld. In de praktijk zullen in dit soort situaties verdelingen F_i optreden die qua vorm en schaal sterk op elkaar lijken. Het enige merkbare verschil zit dan in de lokatieparameter. Toetsing van H_0: $F_1 = \cdots = F_k$ komt in deze situatie overeen met toetsing van H_0: $\mu_1 = \cdots = \mu_k$. Van dit type toetsen bevat NPAR1WAY er drie die hieronder besproken worden. De toets van Van der Waerden (1953) gebruikt de toetsingsgrootheid:

$$Q_{VdW} = \frac{N-1}{\sum_{i=1}^{N}[\Phi^{-1}(\frac{i}{N+1})]^2} \sum_{j=1}^{k} \frac{1}{n_j} [\sum_{i \in I_j} \Phi^{-1}(\frac{R_i}{N+1})^2]$$

Hierbij is $N = \sum_{i=1}^{k} n_i$ de totale steekproefgrootte. Φ stelt de standaardnormale verdelingsfunktie voor. I_j is de indexverzameling van groep j nadat alle waarnemingen x_{ij} vervangen zijn door hun rangnummers R_i binnen de gezamenlijke steekproef. Er zijn dus N rangnummers. Indien twee of meer waarnemingen gelijk zijn, worden hun rangnummers gemiddeld. De tweede verdelingsvrije toets is afkomstig van Kruskal en Wallis (1952):

$$Q_{KW} = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{1}{n_j} [\sum_{i \in I_j} R_i - n_j \frac{N+1}{2}]^2$$

De derde (en laatste) toets van dit type staat bekend als de mediaantoets en is ontleend aan Mood en Brown (1950). De toetsingsgrootheid wordt gegeven door:

$$Q_{Med} = 4\sum_{j=1}^{k}\frac{1}{n_j}(A_j-\tfrac{1}{2}n_j)^2$$

$$A_j = \sum_{i\in I_j}\tfrac{1}{2}(sign\,[R_i-\tfrac{1}{2}(N+1)]+1)$$

De kritieke waarden voor deze toetsingsgrootheden zijn getabelleerd voor kleine waarden van k en n_i. Voor grotere waarden moet met benaderingen gewerkt worden. Deze toetsingsgrootheden zijn onder de nulhypothese alle asymptotisch verdeeld als χ_ν^2 met $\nu = k-1$ vrijheidsgraden. Dit asymptotische resultaat geeft aanleiding tot goede benaderingen als $n_i \geqslant 5$ voor iedere steekproef. Zie hiervoor bijvoorbeeld Conover (1971).
Op grond van de frequentietabel uit de SAS-sessie moeten de programma's BMDP, SPSS, SPSS-PC en STATGRAPHICS dus uit het onderzoek verdwijnen omdat zij allen minder dan vijf keer voorkwamen. Voor de overige programma's is het gemiddelde en de standaarddeviatie voor de behandeltijd middels de routine MEANS uitgerekend. Vervolgens is de hypothese getoetst met NPAR1WAY in de default-setting.
NPAR1WAY begint in deze situatie met een klassieke ANOVA. De overschrijdingskans hierbij is 0.0366 (zie het voorbeeld), maar het weglaten van vier softwareprodukten heeft de non-normaliteit niet opgeheven. Onderzoek middels UNIVARIATE leverde namelijk een overschrijdingskans van 0.0124 op bij toetsing van de hypothese dat de residuen afkomstig waren uit een normale verdeling.
Na de klassieke toets wordt in het bijgevoegde voorbeeld overgegaan op verdelingsvrije methoden. Eerst wordt nu het minder belangrijke laatste deel van de uitvoer besproken. De toets van Savage (1962) geeft een overschrijdingskans van 0.0698. Deze toets is echter volgens Hajek (1969) voornamelijk gevoelig voor schaalverschillen en dus niet erg interessant binnen de context van dit onderzoek. De toets van Kolmogorov (1941) en Smirnov (1939) geeft geen overschrijdingskans, evenals die van Cramer (1946) en Von Mises (1931). Deze toetsen zijn gebaseerd op het vergelijken van de empirische verdelingsfunkties. Alleen als er slechts twee groepen worden vergeleken levert SAS hierbij de overschrijdingskans. Deze door SAS gekozen beperking is niet essentieel. Reeds in 1966 behandelde Miller de principes waarmee een simultane uitspraak over het vergelijken van k steekproeven kan worden gebaseerd op de resultaten voor de afzonderlijke paarsgewijze vergelijkingen.

3. Een schijnbare paradox

Nu komen de resultaten voor de relevante toetsen uit het voorbeeld aan de orde. Tabel 1 bevat de overschrijdingskansen, alsmede de uitspraak over de nulhypothese indien getoetst werd met een gekozen onbetrouwbaarheid van 5%.

Tabel 1: Drie verdelingsvrije toetsen

Toets	Overschrijdingskans	Hypothese
Kruskal & Wallis	0.0423	Verworpen
Mediaan toets	0.0792	Geaccepteerd
Van der Waerden	0.0556	Geaccepteerd

De statisticus die aan de onderzoeker moet rapporteren of de hypothese al of niet verworpen is, zit nu met een dilemma. Twee toetsen accepteren de hypothese en een andere verwerpt hem. Er staan de statisticus nu verschillende wegen open:

1. Hij gooit een dobbelsteen op en spreekt met zichzelf af dat bij uitkomst 1 of 2 de toets van Kruskal & Wallis de uitslag zal bepalen. Bij uitkomst 3 of 4 wordt het de Mediaan toets en bij 5 of 6 de toets van Van der Waerden. Deze strategie is statistisch clean in die zin dat de gekozen onbetrouwbaarheid $\alpha = 0.05$ gelijk is aan de kans op verwerping van de nulhypothese indien deze waar is. Maar in deze aanpak wordt niets gebruikt van het voordeel dat SAS biedt in de vorm van drie verschillende toetsen.
2. Kies de kleinste overschrijdingskans. Deze aanpak geeft een aanzienlijke winst in het onderscheidend vermogen, maar de prijs die ervoor betaald wordt is hoog: de verwerpingskans van de geldige nulhypothese overschrijdt de gekozen onbetrouwbaarheid. Een strategie met deze eigenschap heet progressief en is statistisch onaanvaardbaar.
3. Probeer een indruk te krijgen van de verdeling waar de data uit afkomstig waren. Voor elk van de toetsen is er namelijk een verdeling waarvoor het onderscheidend vermogen asymptotisch optimaal is. Deze zijn gegeven in tabel 2.

Tabel 2: Asymptotische optimaliteit

Toets	Verdeling
Kruskal & Wallis	Logistisch
Mediaan	Dubbel exponentieel
Van der Waerden	Normaal

Als gevonden wordt met welke van de in tabel 2 genoemde verdelingen de empirische verdeling van de data het meest overeenkomt, dan kan daarop de keuze van de toets (en dus de te interpreteren overschrijdingskans) worden gebaseerd. In deze situatie is de gekozen onbetrouwbaarheid gelijk aan de verwerpingskans onder de nulhypothese indien het vergelijken van de verdelingsfunkties onafhankelijk gebeurde van de berekening van de toetsingsgrootheden. De volgende paragraaf gaat over een methode om dit te bewerkstelligen.

4. Selectie van de toets

In 1967 gaven Hajek en Sidak aan dat informatie in de gecombineerde steekproef onafhankelijk is van de rangnummers in de verschillende groepen. Voor de drie behandelde verdelingsvrije methoden worden de toetsingsgrootheden berekend op basis van de rangnummers in de verschillende groepen. Het is dus aantrekkelijk om te proberen een selectie van de optimale toets te baseren op uitsluitend informatie in de gecombineerde steekproef, en dus geen gebruik te maken van informatie die de groep van toebehoren voor iedere waarneming vastlegt.
Dit is de juiste plaats voor een kort theoretisch intermezzo. De toets van Kruskal & Wallis, de mediaan toets en de toets van Van der Waerden behoren tot een familie die in algemene gedaante als volgt kan worden genoteerd:

$$Q = \frac{N-1}{\sum_{i=1}^{N}(a_i-\bar{a})^2}\sum_{j=1}^{k}\frac{[S_j-E(S_j)]^2}{n_j}$$

De grootheden a_i stellen scores voor om een optimaal onderscheidend vermogen voor een zekere gekozen verdeling te krijgen. S_j is de som van de scores binnen de j-de steekproef. De scores a_i worden als volgt vastgelegd door de score-genererende functie ϕ:

$$a_i = \phi(\frac{i}{N+1},f)$$

Er wordt nu gestreefd naar asymptotische optimaliteit van verdeling F met kansdichtheid $f = \frac{d}{dx}F$. Dit doel wordt bereikt met de volgende keuze:

$$\phi(u,f) = -\frac{\frac{d}{du}f[F^{-1}(u)]}{f[F^{-1}(u)]}$$

Voor de toets van Kruskal & Wallis wordt asymptotisch een optimaal onderscheidend vermogen bereikt als de verdeling logistisch is. Hier geldt:

$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}$$

De toets van Mood & Brown ofwel de mediaantoets bereikt asymptotische optimaliteit voor de dubbelexponentiele verdeling. Deze onderscheidt zich voornamelijk van de logistische verdeling doordat de staarten aanzienlijk dikker zijn:

$$f(x) = \tfrac{1}{2}e^{-|x|}$$

De toets van Van der Waerden is asymptotsch optimaal voor de normale verdeling met als dichtheid:

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

De normale verdeling heeft juist aanzienlijk dunnere staarten dan de logistische. De gegeven dichtheden betreffen de genoemde verdelingen in de eenvoudigste vorm. Deze keuze is gemaakt omdat de berekening van Q niet wordt beinvloed door lokatie en schaal van de verdeling waarvoor asymptotische optimaliteit werd nagestreefd.
Op grond van bovenstaande moge duidelijk zijn dan een keuze tussen de drie verdelingsvrije methoden gebaseerd moet zijn op de staartdikte van de empirische verdeling van de samengevoegde steekproef.

5. Asymptotische Relatieve Efficientie(ARE)

Het streven begint duidelijk te worden. We zoeken naar een techniek om op grond van alle waarnemingen (ongeacht de groep waartoe iedere waarneming behoort) de staart van de onderliggende verdeling te schatten. En vervolgens willen we nagaan waar deze het meest op lijkt: de logistische, de dubbelexponentiele of de normale verdeling. Daarna weten we welke toets de optimale keuze is, en alleen de daarbij behorende overschrijdingskans wordt

geinspecteerd.
Dit werk is alleen maar zinvol indien de onderscheidende vermogens van de toetsen voor de drie genoemde verdelingen aanzienlijke verschillen vertonen. Voor kleine steekproeven betreden we nu een theoretisch weinig ontgonnen terrein, maar asymptotische resultaten zijn er wel. Hier zal met de Asymptotische Relatieve Efficientie gewerkt worden. Deze staat ook wel bekend onder de naam Pitman Efficientie. Beschouw toetsen A en B met steekproefgroottes a en b. Laat α de gekozen onbetrouwbaarheid zijn waarmee de hypothese H_0 getoetst wordt tegen een klasse alternatieven H_c. Dan wordt $ARE_{A,B}$ gedefinieerd als de asymptotische waarde van $\frac{b}{a}$ wanneer a zodanig varieert dat de onderscheidende vermogens gelijk blijven terwijl $b \to \infty$ en $H_c \to H_0$. De waarde van de Asymptotische Relatieve Efficientie wordt in tabel 3 gegeven. Voor de afleiding van deze tabel zie Dijkstra (1984).

Tabel 3: Asymptotische Relatieve Efficientie

Verdeling	$ARE_{VdW,K\&W}$	$ARE_{VdW,Med}$	$ARE_{K\&W,Med}$
Normaal	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{3}{2}$
Logistisch	$\frac{3}{\pi}$	$\frac{4}{\pi}$	$\frac{4}{3}$
Dubbelexponentieel	$\frac{8}{3\pi}$	$\frac{2}{\pi}$	$\frac{3}{4}$

Sommige van de waarden in tabel 3 wijken substantieel af van 1. Daarom lijkt een adaptief schema (waarin de keuze van de toets wordt aangepast aan de geschatte staartdikte) zinvol. Er moet dus nu gezocht worden naar een geschikte maat voor de staartdikte. In Dijkstra (1987) worden een aantal maten vergeleken. De maat Q van Hogg, Fisher en Randles (1975) komt in dit werkje als beste naar voren:

$$Q = \frac{10(U_{.05}-L_{.05})}{U_{.5}-L_{.5}}$$

Hier staat $U_{.05}$ voor de som van de 5% grootste waarnemingen. Indien het aantal waarnemingen geen veelvoud is van 20, dan wordt een waarneming gedeeltelijk bij de berekening betrokken. De overige termen hebben een analoge betekenis, waarbij L de kleinste waarnemingen aanduidt. Deze maat is lokatie- en schaal-invariant. Tabel 4 geeft de waarden van Q voor oneindig grote steekproeven uit de relevante verdelingen, alsmede de middelpunten tussen qua staartdikte opeenvolgende verdelingen. Tabel 5 bevat een hieruit afgeleid selectiecriterium. In een appendix zullen de afleidingen behandeld worden. Het is van belang ons te realiseren dat dit selectiecriterium gebaseerd is op asymptotische resultaten. De winst voor (zeer) grote steekproeven is op grond van Asymptotische Relatieve Efficienties bekend. Maar het is niet a-priori zeker dat dezelfde resultaten ook voor kleine steekproeven worden bereikt.

Tabel 4: Criterium Q

Verdeling	Q	Criterium
Normaal	2.58	
		2.72
Logistisch	2.86	
		3.08
Dubbelexponentieel	3.30	

Tabel 5: Selectie met Q

$Q < 2.72$	Van der Waerden
$2.72 \leqslant Q < 3.08$	Kruskal & Wallis
$3.08 \leqslant Q$	Mediaantoets

6. Evaluatie middels simulatie

Er wordt nu onderzocht of dit adaptieve schema aanleiding geeft tot een groter onderscheidend vermogen dan de aparte toetsen voor een ruime klasse van symmetrische verdelingen. Daartoe lijkt het niet verstandig om bij voorbaat het onderzoek te beperken tot de reeds genoemde verdelingen. Er worden nog twee extra opgenomen, namelijk de uniforme verdeling met $f(x) = 1$ voor $-\frac{1}{2} \leqslant x < \frac{1}{2}$ en $f(x) = 0$ daarbuiten. Deze heeft uiteraard dunnere staarten dan de normale verdeling (voor zover men uberhaupt van staarten kan spreken). Ook met de mogelijkheid van dikkere staarten wordt rekening gehouden middels de Cauchy verdeling:

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$$

Eerst wordt nu onderzocht hoe goed het selectiecriterium werkt. Er wordt gesimuleerd met vier steekproeven ($k = 4$). Voor de steekproefgroottes worden de waarden 15, 40 en 65 geprobeerd. Uit elk van de vijf genoemde symmetrische verdelingen worden steeds 600 datasets gegenereerd waarbij de lokatieparameters enigszins verschillend werden genomen. Hierdoor kon zowel de juistheid van de selectie als het uiteindelijke onderscheidend vermogen worden geverifieerd. De resultaten voor de selectie zijn weergegeven in tabel 6 en die voor het onderscheidend vermogen in tabel 7.

Tabel 6: Resultaten selectie

Steekproefgrootte	Correct	Naburig	Tegenover
15	2210	675	115
40	2449	547	4
65	2610	390	0

Bij de beoordeling van de selectie wordt onderscheid gemaakt tussen twee situaties. De Van der Waerden toets en de Mediaan toets zijn elkaars uiterste. De een voor de ander

Tabel 7: Onderscheidend vermogen				
Steekproefgrootte	K&W	Med	VdW	Adaptief
15	41.90	27.63	42.93	43.53
40	80.03	70.47	76.70	82.00
65	92.47	88.27	89.63	94.30

selecteren is dus de grootst mogelijke fout die gemaakt kan worden. Ten onrechte de toets van Kruskal & Wallis kiezen is volgens de asymptotische resultaten veel minder ernstig. We zien dat voor kleine steekproeven (waar de staart minder herkenbaar is) soms de grootste fout wordt gemaakt (met een geschatte kans van 3.8%). Maar voor de grootste steekproeven komt dit niet meer voor.
Het onderscheidend vermogen van de adaptieve toets is superieur, maar de winst ten opzichte van de middelste (Kruskal & Wallis) is gering.

7. Conclusie

De keuze van SAS om bij wijze van default de Van der Waerden toets, de mediaan toets en de toets van Kruskal & Wallis alle drie met overschrijdingskans te produceren is niet optimaal. Het zou eenvoudiger en consistenter geweest zijn om alleen maar Kruskal & Wallis bij wijze van standaard af te drukken. De andere twee zouden middels een optie opvraagbaar moeten blijven, maar de gebruiker moet deze optie alleen gebruiken indie hij weet dat de staarten van de onderliggende verdeling uitzonderlijk dik of dun zijn.
Fraaier zou het zijn indien NPAR1WAY een adaptief schema zou bevatten zoals dat hier behandeld is. De consistentie en de bescheiden omvang van de uitvoer blijven gewaarborgd, maar nu komt er ook nog een winst in het asymptotisch onderscheidend vermogen bij.
Er is nog een kleine winst in het ondersheidend vermogen mogelijk met een ingewikkelder selectieschema dat niet alleen Q, maar ook de gemiddelde steekproefgrootte $\bar{n} = N/k$ gebruikt. Het aantal steekproeven en hun omvang zijn doorgaans reeds voor het experiment bekend. Daarom kan deze informatie in het selectieschema worden betrokken zonder het gevaar dat de feitelijke verwerpingskans onder de nulhypothese de gekozen onbetrouwbaarheid overschrijdt. Een dergelijk schema is ontwikkeld door Van den Heuvel (1987) en Kentstra (1988). De winst in onderscheidend vermogen is echter slechts ongeveer 1.7% en daarom zal dit schema hier niet behandeld worden.
Reeds bij een selectie op asymptotische resultaten bleek dat NPAR1WAY op eenvoudige wijze consistent gemaakt kan worden. Hierbij wordt dan tevens papier uitgespaard en onderscheidend vermogen gewonnen. Hopelijk zal SAS deze weg eens inslaan.

8. Appendix: berekening van Q

In deze paragraaf wordt de waarde van Q berekend voor oneindig grote steekproeven uit de uniforme, de normale, de logistische, de dubbelexponentiele en de Cauchy verdeling. De grootheid Q wordt gedefinieerd als:

$$Q = \frac{10(U_{.05}-L_{.05})}{U_{.5}-L_{.5}}$$

$U_{.05}$ staat voor de som van de grootste 5% waarnemingen. Als het aantal waarnemingen naar oneindig gaat, hoeven we ons geen zorgen te maken over de vraag of dit aantal wel door 20 deelbaar is. Voor oneindige steekproeven uit symmetrische verdelingen kan Q berekend worden middels:

$$Q = \frac{10\int_c^\infty xf(x)dx}{\int_0^\infty xf(x)dx}$$

Hierbij staat c voor het bovenste 5% punt van de verdeling F met kansdichtheid f. Voor de uniforme verdeling beschouwen we het interval van $-\frac{1}{2}$ tot $\frac{1}{2}$ met dichtheid $f(x) = 1$. Dit resulteert in:

$$Q_U = \frac{10\int_{.45}^{.5} xdx}{\int_0^{.5} xdx} = 1.9$$

Voor de standaardnormale verdeling geldt $c = 1.645$. Daarom kan Q hier als volgt worden berekend:

$$Q_N = \frac{10\frac{1}{\sqrt{2\pi}}\int_{1.645}^\infty xe^{-\frac{1}{2}x^2}dx}{\frac{1}{\sqrt{2\pi}}\int_0^\infty xe^{-\frac{1}{2}x^2}dx} = 10e^{-\frac{1}{2}1.645^2} = 2.58$$

Voor de logistische verdeling nemen we de eenvoudigste vorm waarin $F(x) = (1+e^{-x})^{-1}$. Eenvoudig is in te zien dat $c = \log_e 19$. Dit resulteert in:

$$Q_L = \frac{10\int_c^\infty \frac{xe^{-x}}{(1+e^{-x})^2}dx}{\int_0^\infty \frac{xe^{-x}}{(1+e^{-x})^2}dx} = \frac{10\int_c^\infty x\sum_{j=1}^\infty (-1)^{j-1}je^{-jx}dx}{\sum_{j=1}^\infty (-1)^{j-1}j^{(-1)}} =$$

$$\frac{10\sum_{j=1}^\infty (-1)^{j-1}(c+\frac{1}{j})e^{-jc}}{\log_e 2} = \frac{10[\frac{ce^{-c}}{1+e^{-c}}+\log_e(1+e^{-c})]}{\log_e 2} =$$

$$10\frac{\frac{\log_e 19}{20}+\log_e(\frac{20}{19})}{\log_e 2} = 2.86$$

In de stap waarin het integraal-teken uit de teller wordt verwijderd, is gebruik gemaakt van de volgende relatie:

$$\int_c^{\infty} x j e^{-jx} dx = (c+\frac{1}{j})e^{-jc}$$

Voor de dubbelexponentiele verdeling beschouwen we de standaarvorm. Ter vereenvoudiging bekijken we de dichtheid van de absolute waarden, zodat de linkerstaart in de as van symmetrie wordt gespiegeld. We weten reeds dat:

$$\int_0^{\infty} x e^{-x} = 1$$

In de teller gebruiken we $c = \log_e 10$ omdat dan $1-e^{-c} = 0.90$. Dit resulteert in de volgende waarde voor Q:

$$Q_D = 10 \int_{\log_e 10}^{\infty} x e^{-x} dx = 10(\log_e 10+1)e^{-\log_e 10} = 3.30$$

Voor de Cauchy verdeling wordt de waarde van Q als volgt gegeven:

$$Q_C = \frac{10\int_c^{\infty}\frac{x}{1+x^2}dx}{\int_0^{\infty}\frac{x}{1+x^2}dx} = \frac{10(\log_e \infty - \log_e c)}{\log_e \infty} = 10$$

In deze uitdrukking is ∞ gemakshalve gebruikt voor een hulpvariabele die naar oneindig gaat. Het is aardig om te zien dat het resultaat onafhankelijk is van c. Dit betekent dat de staarten van de Cauchy verdeling uiteindelijk altijd het middenstuk overheersen.

9. Litteratuur

SAS/STAT Guide for Personal Computers (1987) Version 6 Edition
SAS Institute Inc. (Cary, NC, USA)

Box, G.E.P. and D.R. Cox (1964) An analysis of transformations
Journal of the Royal Statistical Society (B 26) 211-243

SAS Procedures Guide for Personal Computers (1985) Version 6 Edition
SAS Institute Inc. (Cary, NC, USA)

RC-Informatie PP-4.3 (1987) Meervoudige Regressie en Correlatie
TUE-RC 69889 (Eindhoven University of Technology)

Waerden, B.L. van der (1952) Order tests for the two-sample problem and their power
Indagationes Math. (14) 453-458

Kruskal, W.H. and W.A. Wallis (1952) Use of ranks in one-criterioon variance analysis
Journal of the American Statistical Association (47) 583-621

Brown, G.W. and A.M. Mood (1950) On median tests for linear hypotheses
Proceedings of the second Berkeley Symposium. 159-166

Savage, I.R. (1962) Bibliography of Nonparametric Statistics
Harvard University. Cambridge, Mass.

Hajek, J. (1969) A Course in Nonparametric Statistics
Holden-Day, San Francisco

Kolmogorov, A.N. (1941) Confidence limits for an unknown distribution function
Annals of Mathematical Statistics (12) 461-463

Smirnov, N.V. (1939) Estimate of deviation between empirical distribution functions in two independent samples
Bull. Moscow Univ. (2) 3-16

Cramer, H. (1946) Mathematical Methods of Statistics
Princeton University Press

Mises, R. von (1931) Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und Theoretischen Physik
F. Deuticke, Leipzig und Vienna

Miller, R.G. (1966) Simultaneous statistical inference
McGraw-Hill Book Company, New York

Hajek, J. and Z. Sidak (1967) Theory of Rank Tests
Academia, Prague

Dijkstra, J.B. (1984) Nonparametric comparison of several mean values with mild adaptation to some sample characteristics
COMPSTAT, Prague (Computing Centre Note 20, Eindhoven University of Technology)

Dijkstra, J.B. (1987) Analysis of means in some non-standard situations
Proefschrift (Eindhoven University of Technology)

Hogg, R.V., D.M. Fisher and R.H. Randles (1975) A two-sample adaptive distribution-free test
Journal of the American Statistical Association (70) 656-661

Heuvel, M. van den (1987) Toetsen voor het vergelijken van gemiddelde waarden met aanpassing aan de staartdiktes
Computing Centre Note 37 (Eindhoven University of Technology)

Kentstra, W.P.A. (1988) Toetsen voor het vergelijken van gemiddelde waarden met aanpassing aan de staartdiktes 2
Computing Centre Note 40 (Eindhoven University of Technology)

SAS 12:28 Friday, July 29, 1988

------------------------- SOFTWARE=bmdp -------------------------

OBS	TIJD
1	5

------------------------- SOFTWARE=glim -------------------------

OBS	TIJD
2	4
3	14
4	12
5	8
6	6
7	3
8	3
9	3
10	3
11	14
12	10
13	6
14	17
15	5

------------------------- SOFTWARE=pp4 -------------------------

OBS	TIJD
16	3
17	3
18	5
19	7
20	12
21	3
22	15
23	6
24	11
25	15
26	7
27	7
28	12
29	12
30	16
31	11
32	9
33	13
34	3
35	4

SAS 12:28 Friday, July 29, 1988

-------------------------- SOFTWARE=ppS --------------------------

OBS	TIJD
36	3
37	5
38	6
39	7
40	4
41	3

-------------------------- SOFTWARE=sas --------------------------

OBS	TIJD
42	12
43	4
44	8
45	15
46	12
47	7
48	12
49	10
50	7
51	4
52	5
53	6
54	3
55	3
56	5
57	5
58	5
59	8
60	5

------------------------- SOFTWARE=saspc -------------------------

OBS	TIJD
61	7
62	4
63	5
64	10
65	10
66	3
67	5
68	8
69	5
70	5
71	3
72	6

SAS 12:28 Friday, July 29, 1988

------------------------- SOFTWARE=spss --------------------------

OBS	TIJD
73	3
74	4
75	3

------------------------- SOFTWARE=spsspc --------------------------

OBS	TIJD
76	4

------------------------- SOFTWARE=spssx --------------------------

OBS	TIJD
77	4
78	3
79	3
80	4
81	3

------------------------- SOFTWARE=statgrap --------------------------

OBS	TIJD
82	4
83	3
84	4

SAS 12:28 Friday, July 29, 1988

SOFTWARE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
bmdp	1	1.2	1	1.2
glim	14	16.7	15	17.9
pp4	20	23.8	35	41.7
pp5	6	7.1	41	48.8
sas	19	22.6	60	71.4
saspc	12	14.3	72	85.7
spss	3	3.6	75	89.3
spsspc	1	1.2	76	90.5
spssx	5	6.0	81	96.4
statgrap	3	3.6	84	100.0

SAS 12:28 Friday, July 29, 1988

Analysis Variable : TIJD

SOFTWARE=glim

N Obs	N	Minimum	Maximum	Mean	Std Dev
14	14	3.0000000	17.0000000	7.7142857	4.8425767

SOFTWARE=pp4

N Obs	N	Minimum	Maximum	Mean	Std Dev
20	20	3.0000000	16.0000000	8.7000000	4.4497191

SOFTWARE=pp5

N Obs	N	Minimum	Maximum	Mean	Std Dev
6	6	3.0000000	7.0000000	4.6666667	1.6329932

SOFTWARE=sas

N Obs	N	Minimum	Maximum	Mean	Std Dev
19	19	3.0000000	15.0000000	7.1578947	3.5002088

SOFTWARE=saspc

N Obs	N	Minimum	Maximum	Mean	Std Dev
12	12	3.0000000	10.0000000	5.9166667	2.3915888

SOFTWARE=spssx

N Obs	N	Minimum	Maximum	Mean	Std Dev
5	5	3.0000000	4.0000000	3.4000000	0.5477226

SAS 12:28 Friday, July 29, 1988

N P A R 1 W A Y P R O C E D U R E

Analysis of Variance for Variable TIJD
Classified by Variable SOFTWARE

SOFTWARE	N	Mean
glim	14	7.71428571
pp4	20	8.70000000
pp5	6	4.66666667
sas	19	7.15789474
saspc	12	5.91666667
spssx	5	3.40000000

Among MS	Within MS
35.3827820	13.9861923
F Value	**Prob > F**
2.530	0.0366

Average Scores were used for Ties

SAS 12:28 Friday, July 29, 1988

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable TIJD
Classified by Variable SOFTWARE

SOFTWARE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
glim	14	562.000000	539.000000	73.9336996	40.1428571
pp4	20	931.500000	770.000000	83.9830392	46.5750000
pp5	6	158.500000	231.000000	51.4288983	26.4166667
sas	19	783.500000	731.500000	82.5841692	41.2368421
saspc	12	420.500000	462.000000	69.5445635	35.0416667
spssx	5	70.000000	192.500000	47.2820995	14.0000000

Average Scores were used for Ties

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ= 11.502 DF= 5 Prob > CHISQ= 0.0423

SAS 12:28 Friday, July 29, 1988

N P A R 1 W A Y   P R O C E D U R E

Median Scores (Number of Points above Median)
for Variable TIJD
Classified by Variable SOFTWARE

SOFTWARE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
glim	14	6.0	6.07894737	1.68619127	0.428571429
pp4	20	13.0	8.68421053	1.91538457	0.650000000
pp5	6	1.0	2.60526316	1.17292872	0.166666667
sas	19	9.0	8.25000000	1.88348082	0.473684211
saspc	12	4.0	5.21052632	1.58608911	0.333333333
spssx	5	0.0	2.17105263	1.07835350	0.000000000

Average Scores were used for Ties
Median 1-Way Analysis (Chi-Square Approximation)
CHISQ= 9.8641 DF= 5 Prob > CHISQ= 0.0792

SAS 12:28 Friday, July 29, 1988

N P A R 1 W A Y   P R O C E D U R E

Van der Waerden Scores (Normal) for Variable TIJD
Classified by Variable SOFTWARE

SOFTWARE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
glim	14	1.32002609	0.0	3.15938741	0.09428758
pp4	20	6.63615938	0.0	3.58882294	0.33180797
pp5	6	-3.03692904	0.0	2.19769625	-0.50615484
sas	19	2.03014362	0.0	3.52904543	0.10684966
saspc	12	-1.92807170	0.0	2.97182773	-0.16067264
spssx	5	-5.02132836	0.0	2.02049229	-1.00426567

Average Scores were used for Ties
Van der Waerden 1-Way (Chi-Square Approximation)
CHISQ= 10.793 DF= 5 Prob > CHISQ= 0.0556

SAS 12:28 Friday, July 29, 1988

N P A R 1 W A Y   P R O C E D U R E

Savage Scores (Exponential) for Variable TIJD
Classified by Variable SOFTWARE

SOFTWARE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
glim	14	3.36045860	0.0	3.27501661	0.240032757
pp4	20	7.57974048	0.0	3.72016888	0.378987024
pp5	6	-3.16995265	0.0	2.27812888	-.528325442
sas	19	-0.19528782	0.0	3.65820360	-.010278306
saspc	12	-3.62002017	0.0	3.08059250	-.301668348
spssx	5	-3.95493843	0.0	2.09443950	-.790987686

Average Scores were used for Ties

Savage 1-Way (Chi-Square Approximation)
CHISQ= 10.197 DF= 5 Prob > CHISQ= 0.0698

SAS 12:28 Friday, July 29, 1988

N P A R 1 W A Y P R O C E D U R E

Kolmogorov-Smirnov Test for Variable TIJD
Classified by Variable SOFTWARE

SOFTWARE	N	EDF at maximum	Deviation from Mean at maximum
glim	14	0.4	0.10549786
pp4	20	0.2	-0.35306336
pp5	6	0.5	0.41899167
sas	19	0.2	-0.51618540
saspc	12	0.2	-0.27348171
spssx	5	1.0	1.50051930
--------	----	-----------	
	76	0.3	

Maximum Deviation occurred at Observation 39
Value of TIJD at maximum 4.00000000

Kolmogorov-Smirnov Statistic (Asymptotic)
KS = 0.195480 KSa = 1.70415

SAS 12:28 Friday, July 29, 1988

N P A R 1 W A Y P R O C E D U R E

Cramer-von Mises Test for Variable TIJD
Classified by Variable SOFTWARE

SOFTWARE	N	Summed Deviation from Mean
glim	14	0.060529180
pp4	20	0.387725980
pp5	6	0.225928828
sas	19	0.098597645
saspc	12	0.123891359
spssx	5	0.783349431

Cramer-von Mises Statistic (Asymptotic)
CM = 0.022106 CMa = 1.68002