Berekeningsmethoden voor de (niet-centrale) t-verdeling

Eindhoven University of Technology
Computing Centre Note 42

Berekeningsmethoden voor de (niet-centrale) t-verdeling.

Jan B. Dijkstra

Bestemd voor de themadag "Rondom de Studentverdeling" op 14 oktober 1988 in Utrecht. Georganiseerd door de Landbouwkundige Sectie en Medisch-Biologische Sectie van de VVS.

september 1988.

Berekeningsmethoden voor de (niet-centrale) t-verdeling

Jan B. Dijkstra

Samenvatting

De (niet-centrale) t-verdeling wordt geplaatst in een historisch perspectief. Er wordt een vergelijking gemaakt met de normale verdeling die vroeger gebruikt werd in toepassingen waar de t-verdeling een juistere keuze was. Berekeningsmethoden worden gegeven voor mainframes, maar ook eenvoudige benaderingen voor pocket-calculators en personal computers. Voor het inverteren van de verdelingsfunktie worden een aantal methoden gegeven. Vroegere resultaten zijn gepubliceerd in tabellen of nomogrammen; hiernaar alsmede naar allerlei afleidingen zal uitgebreid worden verwezen.

1. Inleiding

Laat $x_1, \cdots, x_n$ normaal verdeelde grootheden zijn met dezelfde verwachting μ en variantie σ^2. Verder wordt ook onafhankelijkheid verondersteld. Het steekproefgemiddelde is $\bar{x}$ en de variantie hiervan is σ^2/n. De verwachting van het steekproefgemiddelde is μ en de standaarddeviatie is $\sigma/\sqrt{n}$ zodat:

$$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \simeq U$$

Hierbij stelt U een standaardnormaal verdeelde grootheid voor. Met deze uitdrukking kan men hypotheses betreffende μ toetsen of betrouwbaarheidsintervallen voor deze parameter construeren. Er is echter een probleem: σ zal bij veel toepassingen onbekend zijn. Vroeger maakte men zich daar niet erg druk over en men verving σ in dat soort gevallen door s, de steekproef-standaarddeviatie die wordt berekend als de wortel van de steekproef-variantie s^2:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$$

Reeds aan het begin van deze eeuw was echter duidelijk dat vervanging van σ door s resulteerde in een andere verdeling dan U zodat de hierboven genoemde toetsing en betrouwbaarheidsintervallen correcties behoefden. In 1908 publiceerde W.S. Gosset onder het pseudoniem Student de verdeling (met tabellen) van de volgende grootheid:

Rekencentrum Technische Universiteit, Postbus 513, 5600 MB Eindhoven, telefoon 040-474535/474089

$$Z = \sqrt{n-1}\frac{\bar{x}-\mu}{s/\sqrt{n}}$$

Z is verdeeld als U/χ_{n-1} en de ervoor staande factor $\sqrt{n-1}$ maakt dat deze uitdrukking toch niet helemaal de gezochte vorm heeft. In 1925 werd Z door Fisher gestandaardiseerd en ontstond wat nu bekend staat onder de naam Student's t-verdeling:

$$t_\nu \simeq \frac{U}{\sqrt{\chi_\nu^2/\nu}}$$

Dit is de verdeling van een normaal verdeelde grootheid met verwachting 0 die gedeeld wordt door een schatting voor zijn standaarddeviatie. De parameter ν stelt het aantal vrijheidsgraden voor waarmee de standaarddeviatie geschat werd. De toepasbaarheid hiervan is zeer ruim. Behalve de reeds genoemde toepassingen kan hierbij gedacht worden aan het vergelijken van twee steekproeven of aan toetsen met betrekking tot regressiecoefficienten.

2. De kansdichtheid en de verdelingsfunktie

De kansdichtheid van de t-verdeling met ν vrijheidsgraden is gegeven als volgt:

$$f_\nu(t) = \frac{1}{\sqrt{\nu}B(½,½\nu)}(1+\frac{t^2}{\nu})^{-½(\nu+1)}$$

B stelt hierbij de Beta-funktie voor. De grafiek hiervan lijkt veel op de klokvorm van de standaardnormale verdeling. De dichtheid is symmetrisch rond de y-as. In het middengebied is de dichtheid kleiner dan die van de standaardnormale verdeling, maar dit wordt gecompenseerd is de staarten. De buigpunten bevinden zich bij $\pm\sqrt{\nu/(\nu+2)}$. De standaardnormale verdeling heeft buigpunten bij ± 1 en dat is consistent met het feit dat de t-verdeling naar de standaardnormale convergeert als ν naar oneindig gaat. Een geschatte standaarddeviatie van een oneindig grote steekproef heeft immers de populatiewaarde. Vanwege de symmetrie geldt voor de verdelingsfunktie:

$$F_\nu(t) = \int_{-\infty}^{t} f_\nu(y)dy = \tfrac{1}{2}+\int_0^t f_\nu(y)dy$$

Voor $\nu = 1$ wordt dit een zeer eenvoudige uitdrukking: $F_1(t) = \frac{1}{2}+\frac{1}{\pi}\tan^{-1}t$. Laat nu $\theta = \tan^{-1}(t/\sqrt{\nu})$. Dan is voor ν oneven en groter dan 1:

$$F_\nu(t) = \tfrac{1}{2}+\frac{1}{\pi}[\theta+(\cos\theta+\frac{2}{3}\cos^3\theta+ \cdots +\frac{2*4* \cdots *(\nu-3)}{3*5* \cdots *(\nu-2)}\cos^{\nu-2}\theta)\sin\theta]$$

Deze uitdrukking is ontleend aan Zelen en Severo (1964) en is beschreven in het bekende Handbook of Mathematical Functions van Abramowitz en Stegun. Voor even ν krijgen we de uitdrukking:

$$F_\nu(t) = \tfrac{1}{2}+[1+\tfrac{1}{2}\cos^2\theta+\frac{1*3}{2*4}\cos^4\theta+ \cdots +\frac{1*3* \cdots *(\nu-3)}{2*4* \cdots *(\nu-2)}\cos^{\nu-2}\theta]\sin\theta$$

3. Het inverteren van de verdelingsfunktie

De grafiek van de verdelingsfunctie is monotoon stijgend met als domein de gehele verzameling der reele getallen. Asymptotisch worden de waarden 0 en 1 bereikt voor respectievelijk $t = -\infty$ en $t = \infty$. Voor de inversie wordt het punt t gezocht waarvoor F_ν een gegeven waarde tussen 0 en 1 heeft. Voor $\nu > 1$ kan dit alleen maar iteratief en is er een beginschatting nodig. De volgende keuze is ontleend aan Abramowitz en Stegun. Laat $F_\nu(t) = P(T \leqslant t)$ met T een t-verdeelde stochastische grootheid met ν vrijheidsgraden zijn. Deze kans zal verder worden aangeduid als P. Voor een zeer ruwe beginschatting kan de t-verdeling worden benaderd middels de standaardnormale verdeling en krijgen we $x = \Phi^{-1}(P)$ als oplossing. Hierbij stelt Φ de verdelingsfunktie van de standaardnormale verdeling voor. Ter verfijning van deze ruwe beginschatting worden nu de volgende funkties ingevoerd:

$$g_1(x) = \frac{1}{4}(x^3+x)$$

$$g_2(x) = \frac{1}{96}(5x^5+16x^3+3x)$$

$$g_3(x) = \frac{1}{384}(3x^7+19x^5+17x^3-15x)$$

$$g_4(x) = \frac{1}{92160}(79x^9+776x^7+1482x^5-1920x^3-945x)$$

Met deze funkties wordt nu de verbeterde beginschatting uitgerekend:

$$x_0 = x + \frac{g_1(x)}{\nu} + \frac{g_2(x)}{\nu^2} + \frac{g_3(x)}{\nu^3} + \frac{g_4(x)}{\nu^4}$$

Voor de na-iteratie ter verhoging van de nauwkeurigheid kan men kiezen uit verschillende methoden waaronder Bisectie, Regula Falsi en Newton-Raphson. Omdat de kansdichtheid van de t-verdeling middels een simpele uitdrukking beschreven kan worden die aanleiding geeft tot stabiele berekeningen ligt de keuze van Newton-Raphson hier voor de hand:

$$x_{i+1} = x_i - \frac{F_\nu(x_i) - P}{f_\nu(x_i)}$$

Deze iteratie eindigt zodra een vooraf gekozen nauwkeurigheid ϵ bereikt is. Voor $\nu = 1$ is deze aanpak van beginschatting en na-iteratie overbodig. Hier geldt namelijk $t = \tan[\pi(P - \frac{1}{2})]$.

4. Tabellen en nomogrammen

Reeds in 1925 produceerde Student tabellen voor $F_\nu(t)$ met een nauwkeurigheid van vier cijfers. Hij deed dit voor $\nu = 1(1)20$ en $t = 0(0.1)6$. Hierbij betekent (h) "met stappen ter grootte h". Veel uitgebreider zijn de tabellen van Pearson en Hartley uit Biometrika Tables for Statistician (1958). Hier wordt $F_\nu(t)$ in vijf cijfers gegeven voor $\nu = 1(1)24, 30, 40, 60, 120, \infty$ en $t = 0(0.1)4(0.2)8$ voor $\nu \leqslant 20$ en $t = 0(0.05)2(0.1)4, 5$ voor $\nu \geqslant 20$.

Pearson en Hartley geven ook de inverse t in drie cijfers voor $\nu = 1(1)30, 40, 60, 100, \infty$ en $P = 0.6, 0.75, 0.9, 0.95, 0.975, 0.99, 0.995, 0.9975, 0.999, 0.9995$. In de praktijk zal men bij de berekening van t voornamelijk geinteresseerd zijn in waarden van P die maar weinig onder 1 liggen. Dan is de verdelingsfunctie nogal vlak en is de inversie numeriek problematisch. De kansdichtheid ligt dan dicht in de buurt van 0 en deze funktie wordt in de noemer van het Newton-Raphson proces gebruikt. Kortom: er wordt hier gevraagd om moeilijkheden. Hierdoor is Federighi (1959) reeds behoorlijk uitgedaagd en middels enige hoogstandjes uit de numerieke trukendoos heeft hij t in drie cijfers getabelleerd voor extreme waarden van P tot $(1-10^{-7})$.

Johnson en Kotz (1970) geven een zeer uitgebreid historisch overzicht met verwijzingen naar allerlei tabellen. Hier zullen deze slechts genoemd worden: Fisher en Yates (1966), Vesela (1964), Owen (1962), Hald (1952) en Cotterman en Knoop (1968).

Een alternatief voor een tabel wordt gevormd door een zogenaamd nomogram. James-Levy heeft een dergelijke figuur voor de t-verdeling uitgewerkt. Het bestaat uit drie krommen op het platte vlak. Achtereenvolgens zijn hier schaalverdelingen op aangebracht voor P, ν en t. Als twee van deze grootheden gegeven zijn kan de derde hierbij behorende waarde gevonden worden door een lijn te trekken die de twee gegeven punten verbindt. Het snijpunt met de derde kromme geeft de gevraagde waarde. Soortgelijke figuren zijn ook ontwikkeld door Stammberger (1967) en Babanin (1952).

5. Benaderingen voor kleine rekenmachines

De kansdichtheidsfunktie van de t-verdeling laat zich op een aantal manieren uitschrijven als een reeksontwikkeling. Deze reeksen zijn van het type waarin de termen in absolute waarde snel dalen en iedere term de staart domineert. Het is dus mogelijk om de reeks op een verantwoorde plaats af te breken. Vervolgens kan men middels eenvoudige integratie een uitdrukking krijgen die de verdelingsfunktie benadert. En met enige algebra kan deze uitdrukking dan weer geinverteerd worden om t als funktie van P te krijgen. Velen hebben zich met dit soort zaken beziggehouden. Johnson en Kotz noemen: Fisher en Cornisch (1960), Dickey (1967), Hendricks (1936) en Hotelling en Frankel (1938).

Anderen zien de t-verdeling als een normale verdeling met correcties. Laat $U = \Phi^{-1}(p)$ waarin Φ de cumulatieve standaardnormale verdeling voorstelt. Dan komt Peiser (1943) tot de volgende benadering:

$$t = U + \frac{U^3+U}{4\nu}$$

De resultaten hiervan zijn in drie cijfers nauwkeurig voor ν groter of gelijk aan 30. Goldberg en Levine (1946) hebben er nog een term aan toegevoegd:

$$t = U + \frac{U^3+U}{4\nu} + \frac{5U^5+16U^3+3U}{96\nu^2}$$

De gewenste nauwkeurigheid van drie cijfers wordt hier reeds bereikt voor $\nu \geqslant 20$. En als P niet groter is dan 0.975 volstaat het zelfs om ν groter of gelijk aan 10 te nemen. In Abramowtz en Stegun worden nog twee extra termen gegeven. Dan ontstaat de eerder genoemde uitdrukking die als beginschatter was voorgesteld voor grote computers.

Voor meer informatie kan worden verwezen naar Johnson en Kotz. Daarin worden nog de volgende benaderingen genoemd: Simaika (1942), Anscombe (1950), Chu (1956), Walace (1959), Peizer en Pratt (1968), Cornish (1969), Hill (1969), Cuconi (1962), Gardiner en Bombay (1965), Moran (1966), Kramer (1966), Zelen en Severo (1964) en Gentlemen en Jenkins (1968).

6. Het verband tussen de t-verdeling en de F-verdeling

In allerlei toetsen die gebaseerd zijn op het quotient van geschatte varianties wordt gebruik gemaakt van de F-verdeling. Deze heeft twee parameters: het aantal vrijheidsgraden p voor de noemer en q voor de teller. De kansdichtheid ziet er als volgt uit:

$$f_q^p(t) = \frac{p^{\frac{1}{2}p} q^{\frac{1}{2}q}}{B(\frac{1}{2}p, \frac{1}{2}q)} t^{\frac{1}{2}p-1}(q+pt)^{-\frac{1}{2}(p+q)}$$

De verdelingsfunktie is:

$$F_q^p(t) = \int_0^t f_q^p(y)dy$$

Paulson (1942) is een der velen die een benadering ontwikkelde voor de F-verdeling. Alle resultaten voor de F-verdeling zijn toepasbaar voor de t-verdeling door gebruik te maken van de relatie $t_\nu^2 \simeq F_\nu^1$. Veronderstel dat we een numerieke procedure hebben die de verdelingsfunktie van de F-verdeling met parameters p en q en argument x berekent. En we willen de waarde van de verdelingsfunktie van de t-verdeling met ν vrijheidsgraden en argument t weten. De eerste stap betreft dan de volgende substituties: $p = 1$, $q = \nu$ en $x = t^2$. Door deze aanpak is het teken van t verdonkermaand. Dit kan worden hersteld door de volgende transformatie: $P = (1+F)/2$ als $t \geqslant 0$ en $P = (1-F)/2$ als $t < 0$. Hierbij stelt F het resultaat voor de F-verdeling voor. Via Paulson kan men zo het volgende resultaat voor de t-verdeling verkrijgen:

$$P(|T| \leqslant t) = P(U \leqslant \frac{1}{3\sqrt{2}}[(9-\frac{2}{\nu})t^{2/3}-7][\frac{t^{4/3}}{\nu}+1]^{-\frac{1}{2}})$$

Er is dan nog een procedure nodig voor de verdelingsfunktie van de standaardnormale verdeling. Maar dat zal in het algemeen geen probleem zijn. De nauwkeurigheid van de methode van Paulson is niet geweldig (2 of 3 cijfers buiten de uiterste staarten), maar het gebruik van een routine voor de F-verdeling biedt ook andere mogelijkheden. In 1984 publiceerde Lackritz een werkelijk schitterend (stabiel en efficient) algorithme voor de F-verdeling. Dat wordt nu in de routinebibliotheek van het Rekencentrum van de TUE als standaard gebruikt voor de t-verdeling.

7. De niet-centrale t-verdeling

De niet-centrale t-verdeling heeft een parameter meer dan de gewone (centrale) t-verdeling, namelijk de niet-centraliteitsparameter δ. De definitie is als volgt:

$$t_{\nu;\delta} \simeq \frac{U+\delta}{\sqrt{\chi_\nu^2/\nu}}$$

Hierbij volgt U de standaardnormale verdeling. De niet-centrale t-verdeling is onder andere nuttig voor het bepalen van het onderscheidend vermogen van de t-toets. Denk bijvoorbeeld dat we bij een normale verdeling op grond van een steekproef het hypothese H_0: $\mu = \mu^*$ willen toetsen met onbetrouwbaarheid α. Daartoe berekenen we de volgende toetsingsgrootheid:

$$t = \frac{\bar{x}-\mu^*}{s/\sqrt{n}}$$

De beslissingsregel is dat we H_0 verwerpen als $|t| > t_{n-1}(0.025)$. De onbetrouwbaarheid $\alpha = 0.05$ is per definitie de kans om de nulhypothese te verwerpen als hij waar is. Vanwege de tweezijdigheid van de toetsing wordt hier de staartkans gelijk genomen aan de helft van de gekozen onbetrouwbaarheid. Onder de nulhypothese volgt t een t-verdeling met ν vrijheidsgraden. Daarom geldt:

$$\alpha = \int_{-\infty}^{-t_\nu} f_\nu(t)dt + \int_{t_\nu}^{\infty} f_\nu(t)dt = 2\int_{t_\nu}^{\infty} f_\nu(t)dt$$

Hierbij geldt $\nu = n-1$ en de grootheid t_ν stelt de kritieke waarde $t_{n-1}(0.025)$ voor. Met de niet-centrale t-verdeling kunnen we de kans op verwerping uitrekenen voor een gekozen verschuivingsalternatief $\mu = \mu^* + \Delta$ en een bekende populatievariantie σ^2. De niet-centraliteitsparameter δ is dan gelijk aan $\Delta/(\sigma/\sqrt{n})$. En de kans op verwerping wordt:

$$\beta = \int_{-\infty}^{-t_\nu} f_{\nu;\delta}(t)dt + \int_{t_\nu}^{\infty} f_{\nu;\delta}(t)dt$$

Op een soortgelijke manier kan het onderscheidend vermogen worden uitgerekend van een 2-steekproeven t-toets. Ook wordt de niet-centrale t-verdeling wel gebruikt om de steekproefgrootte n te berekenen terwijl β gekozen is.

8. De variatiecoefficient

Met behulp van de niet-centrale t-verdeling is het zeer eenvoudig om betrouwbaarheidsintervallen te construeren voor het quotient μ/σ, het omgekeerde van de variatiecoefficient. Stel dat we een steekproef hebben uit een normale verdeling met verwachting μ en variantie σ^2. Het steekproefgemiddelde $\bar{x}$ heeft dan de variantie σ^2/n bij steekproefgrootte n. Voor niet-centraliteitsparameter $\delta = \mu/(\sigma/\sqrt{n})$ krijgen we nu het volgende:

$$\frac{\bar{x}}{s/\sqrt{n}} \simeq t_{n-1;\delta}$$

Een 95% betrouwbaarheidsinterval voor μ/σ kan nu geconstrueerd worden uit de volgende twee vergelijkingen:

$$t_{n-1;\sqrt{n}\,\theta_2}(0.025) = \frac{\sqrt{n}\,\bar{x}}{s}$$

$$t_{n-1;\sqrt{n}\,\theta_1}(0.975) = \frac{\sqrt{n}\,\bar{x}}{s}$$

De waarden van θ_1 en θ_2 vormen dan de grenzen van het gezochte betrouwbaarheidsinterval. Door de reciproce waarden te nemen kan men een betroubaarheidsinterval voor σ/μ krijgen. Met deze problematiek hebben McKay (1932) en Iglewicz, Myers en Howe (1968) zich beziggehouden.

9. De kansdichtheid en de verdelingsfunktie

Laat T een stochastische variabele zijn die een niet-centrale t-verdeling volgt met ν vrijheidsgraden en niet-centraliteitsparameter δ. Dan geldt het volgende:

$$P(T \leqslant t) = P(U+\delta \leqslant \frac{t\chi_\nu}{\sqrt{\nu}})$$

Dit leidt tot de verdelingsfunktie F van de niet-centrale t-verdeling:

$$F_{\nu;\delta}(t) = P(T \leqslant t) = \frac{1}{2^{\frac{1}{2}\nu-1}\Gamma(\frac{1}{2}\nu)}\int_0^\infty x^{\nu-1}e^{\frac{1}{2}x^2}\frac{1}{\sqrt{2\pi}}\int_0^{tx/\sqrt{\nu}} e^{-\frac{1}{2}(u-\delta)^2}du\,dx$$

Hieruit kan door differentiatie de kansdichtheidsfunktie worden verkregen:

$$f_{\nu;\delta}(t) = \frac{d}{dt}F_{\nu;\delta}(t) = \frac{e^{\frac{1}{2}\delta^2}}{2^{\frac{1}{2}(\nu-1)}\sqrt{\pi\nu}\,\Gamma(\frac{1}{2}\nu)}\int_0^\infty x^\nu \exp[-\tfrac{1}{2}([1+\frac{t^2}{\nu}]x^2-2t\sqrt{\nu}x)]dx$$

F en f zijn ook op andere manieren voor de niet-centrale t-verdeling gegeven. Maar altijd met een numeriek weinig aantrekkelijke representatie. Johnson en Kotz noemen bijvoorbeeld: Fisher (1931), Airey (1931), Amos (1964) en Hodges en Lehmann (1965).

10. Methoden voor de berekening

Omdat de niet-centrale t-verdeling met ν vrijheidsgraden en niet-centraliteitsparameter δ verdeeld is als het quotient van $U+\delta$ en $\chi_\nu/\sqrt{\nu}$ kunnen we de volgende relatie opschrijven:

$$P(T \leqslant t) = P(U - t\frac{\chi_\nu}{\sqrt{\nu}} \leqslant -\delta)$$

T volgt hierbij bovengemoende niet-centrale t-verdeling en U de standaardnormale verdeling. Benaderingen zijn eenvoudig te construeren door te veronderstellen dat de verdeling van $U - t\chi_\nu/\sqrt{\nu}$ niet al te veel van de normale verdeling afwijkt. Hiermee zijn benaderingen gevonden door: Jennett en Welch (1939), Pearson en Hartley (1954), van Eeden (1958) en Johnson en Welch (1940). De nauwkeurigheid van deze methoden is redelijk, behalve als ν zeer klein wordt.

De veronderstelde normaliteit van $U - t\chi_\nu/\sqrt{\nu}$ is niet essentieel. Door gewone algebraische manipulaties met de verdelingsfunktie is het ook mogelijk benaderingen te construeren. Johnson en Kotz noemen hierbij: Merrington en Pearson (1958), Pearson (1961), Azorin (1953), Laubscher (1960), Harley (1957), Hogben, Pinkham en Wilk (1964) en Halperin

(1963).

De nu volgende methode is ontleend aan Owen (1962). Voor oneven ν geldt:

$$P(T \leqslant t) = P^*\left(\frac{-\delta\sqrt{\nu}}{\sqrt{\nu+t^2}}\right)+2T\left(\frac{\delta\sqrt{\nu}}{\sqrt{\nu+t^2}},\frac{t}{\sqrt{\nu}}\right)+$$

$$2[M_1+\frac{1}{3}M_3+\frac{1}{3*5}M_5+\cdots+\frac{1}{3*5*7*\cdots*(\nu-2)}M_{\nu-2}]$$

Hierbij stelt T een stochastische variabele voor die een niet-centrale t-verdeling volgt met niet-centraliteitsparameter δ en ν vrijheidsgraden. In bovenstaande uitdrukking zijn de volgende hulpgrootheden gebruikt:

$$P^*(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(\frac{-t^2}{2}\right)dt$$

$$T(a,b) = \frac{1}{2\pi}\int_0^b\frac{e^{-\frac{1}{2}a^2(1+t^2)}}{1+t^2}dt$$

$$M_{-1} = 0$$

$$M_0 = \frac{t}{\sqrt{\nu+t^2}}Z\left(\frac{\sqrt{\nu}\delta}{\sqrt{\nu+t^2}}\right)P^*\left(\frac{\delta t}{\sqrt{\nu+t^2}}\right)$$

$$Z(x) = \frac{1}{\sqrt{2\pi}}\exp-\frac{x^2}{2}$$

$$M_1 = \frac{\delta t\sqrt{\nu}}{\nu+t^2}M_0+\frac{t\sqrt{\nu}}{\nu+t^2}\frac{1}{\sqrt{2\pi}}Z(\delta)$$

$$M_k = \frac{\delta t\sqrt{\nu}}{\nu+t^2}M_{k-1}+\frac{(k-1)\nu}{\nu+t^2}M_{k-2}$$

Voor even waarden van ν krijgen we:

$$P(T \leqslant t) = P^*(-\delta)+\sqrt{2\pi}[M_0+\frac{1}{1!*2}M_2+\frac{1}{2!2^2}M_4+\cdots+\frac{1}{((\nu-2)/2)!2^{(\nu-2)/2}}M_{\nu-2}]$$

Voor de inverse van de niet-centrale t-verdeling is de Newton-Raphson methode niet erg geschikt omdat de afgeleide (de kansdichtheidsfunktie) alleen middels een onaantrekkelijke uitdrukking berekend kan worden. Bisectie en Regula Falsi zijn hier de aangewezen alternatieven.

11. Tabellen

Johnson en Kotz verwijzen naar de volgende tabellen: Fisher (1931), Neyman, Iwaskiewicz en Kolodziejczyk 1935), Neyman en Tokarska (1936), Johnson en Welch (1940), Resnikoff en Lieberman (1957), Owen (1963), Locks, Alexander en Byars (1963) en Pearson en Hartley (1954). De eerste toepassingen van de niet-centrale t-verdeling betroffen het onderscheidend vermogen van de t-toets. Daarom hebben de oudere tabellen niet δ, maar β als entry. Een voorbeeld hiervan is de tabel van Neyman en Tokarska. Hierin geldt $\nu = 1(1)30(5)100(10)200, \infty$ en $\alpha = 0.005, 0.01, 0.025, 0.05$. De tabellen van Johnson en Welch zitten anders in elkaar. Het doel is hierbij een hulpgrootheid λ te vinden die als volgt gebruikt wordt:

$$t = \frac{\delta+\lambda[1+\frac{1}{2}(\delta^2-\lambda^2)\nu^{-1}]^{\frac{1}{2}}}{1-\frac{1}{2}\lambda^2\nu^{-1}}$$

Hierbij is λ gegeven voor $\nu = 4(1)9, 16, 36, 144, \infty$ en $\alpha = 0.005, 0.01, 0.025, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.975, 0.99, 0.995$. De tabellen van Owen vormen een aanzienlijke uitbreiding hiervan. Hoewel in de praktijk minder gevraagd wordt naar de verdelingsfunktie F dan naar de inverse t is ook voor F een tabel beschikbaar. Deze is gegeven door Locks, Alexander en Byars.

12. Litteratuur

Lackritz, J.R. (1984) Exact p values for F and t tests
The American Statistician (38) 312-314

Student (1908) On the probable error of the mean
Biometrika (6) 1-25

Fisher, R.A. (1925) Applications of Student's distribution
Metron (5) 90-104

Zelen, M. en N.C. Severo (1964) Probability Functions
In: Handbook of Mathematical Functions. Editors: M. Abramowitz en I.A. Stegun. National Bureau of Standards. Applied Mathematics Series (55) Washington D.C.

Student (1925) New tables for testing the significance of observations
Metron (5) 105-108, 114-120

Pearson, E.S. en H.O. Hartley (1958) Biometrika Tables for Statisticians (deel 1)
London. Cambridge University Press

Federighi, E.T. (1959) Extended tables of the percentage points of Student's t-distribution
Journal of the American Statistical Association (54) 683-688

Johnson, N.L. en S. Kotz (1970) Continuous Univariate Distributions (deel 2)
Wiley-Interscience Publication

James-Levy, G.E. (1956) A nomogram for the integral law of Student's distribution
Teoriya Veroyatnostei i ee Primeneniya (1) 271-274, 246-248

Peiser, A.M. (1943) Asymptotic formulas for the significance levels of certain distributions
Annals of Mathematical Statistics (14) 56-62

Goldberg, H. en H. Levine (1946) Approximate formulas for the percentage points and normalization of t and χ^2
Annals of Mathematical Statistics (17) 216-225

Paulson, E. (1942) An approximate normalization of the analysis of variance distribution
Annals of Mathematical Statistics (13) 233-235

Owen, D.B. (1962) Handbook of Statistical Tables
Addison-Wesley Publishing Company. Reading, Massachusetts

Neyman, J. en B. Tokarska (1936) Errors of the second kind in testing Student's hypothesis

Journal of the American Statistical Association (31) 318-326

Locks, M.O., M.J. Alexander en B.J. Byars (1963) New tables of the noncentral t-distribution
Report ARL63-19. Wright-Patterson Air Force Base