STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 47 (M 7)

De toets van Wilcoxon

MATHEMATISCH CENTRUM,
2e Boerhaavestraat 49,
Amsterdam - O.
Statistische Afdeling
Rapport S 47 (M 7)

De toets van Wilcoxon [1)]

Deze methode dient tot het toetsen van de hypothese H_0, inhoudende, dat twee steekproeven $x_1, x_2, \dots, x_m$ en $y_1, y_2, \dots, y_n$ afkomstig zijn uit één collectie (ook wel populatie of universum genaamd).

De gegeven waarnemingsreeks kan dan worden samengevat als aangegeven in tabel I.

Tabel I

Voorkomende steekproef-waarden	Aantal malen dat z_i optreedt bij		totaal
	$\underline{y}$	$\underline{x}$	
z_1	$\underline{b}_1$	$\underline{a}_1$	$\underline{t}_1$
z_2	$\underline{b}_2$	$\underline{a}_2$	$\underline{t}_2$
.	.	.	.
.	.	.	.
.	.	.	.
z_k	$\underline{b}_k$	$\underline{a}_k$	$\underline{t}_k$
totaal	n	m	N

Hierin zijn $z_1, z_2, \dots, z_k$ de voorkomende steekproefwaarden, gerangschikt naar opklimmende grootte, n en m de steekproefgrootten, $t_1, t_2, \dots, t_k$ de grootten der groepen gelijke waarnemingen en $\underline{a}_i$ (resp. $\underline{b}_i$) het aantal malen dat de waarde z_i in de eerste (resp. tweede) steekproef optreedt; $\underline{a}_i$ kan dus de waarden $0, 1, \dots, t_i$ aannemen ($i = 1, 2, \dots, k$).

Voor het toetsen van de hypothese H_0 wordt gebruik gemaakt van een toetsingsgrootheid $\underline{W}$ [2)], die gelijk is aan twee maal

---

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

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

het aantal paren waarnemingen (x_i, y_j), waarvoor $x_i > y_j$ is, vermeerderd met het aantal paren (x_k, y_ℓ) waarvoor $x_k = y_\ell$ is.[3)]

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

Indien de hypothese H_0 juist is, dus als de twee steekproeven uit dezelfde verdeling afkomstig zijn, kan de verdeling van de toetsingsgrootheid $\underline{W}$ berekend worden. Deze verdeling heeft de volgende eigenschappen:

1. De verdeling van $\underline{W}$ is discreet, want $\underline{W}$ kan alleen gehele waarden aannemen.
2. De verdeling van $\underline{W}$ is in het algemeen niet symmetrisch, behalve,
   a) als $m = n$,
   b) als $t_1 = t_k$, $t_2 = t_{k-1}$ enz.
3. De verwachting (het theoretische gemiddelde) μ, van de verdeling van $\underline{W}$ is gelijk aan het product der steekproefgrootten

   (1) $$\mu = mn .$$
4. De variantie van de verdeling van $\underline{W}$ wordt gegeven door:

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

   waarin $D = t_1^3 + t_2^3 + \dots + t_k^3$

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

   $$N^3 - D = N^3 - N = N(N-1)(N+1)$$

   en in dit geval wordt de variantie

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

Wij zagen reeds dat de toetsingsgrootheid $\underline{W}$ een kleine waarde aan zal nemen, indien de x-waarden overwegend kleiner zijn dan de y-waarden een een grote waarde, indien het omgekeerde het geval is. Als tweezijdige kritieke zone nemen wij dus grote waarden van $|\underline{W} - \mu|$.

----

3) Gewoonlijk wordt voor de toets van WILCOXON de toetsingsgrootheid $\underline{U} = \frac{1}{2}\underline{W}$ gebruikt. Om gebroken getallen te vermijden, wordt hier $\underline{W}$ ingevoerd (zie [4], [6]).

Als m en n groot zijn en onderling niet te veel verschillen en als bovendien de groepen gelijke waarnemingen niet te veel in omvang verschillen, kan men gebruik maken van het feit, dat de grootheid $\underline{W}$, als de hypothese H_0 juist is, bij benadering normaal verdeeld is met gemiddelde en variantie volgens (1) en (2).

De tweezijdige overschrijdingskans k, behorende bij een waarde W van $\underline{W}$ is dus

$$k = \frac{2}{\sqrt{2\pi}} \int_{\frac{|W-\mu|-1}{\sigma}}^{\infty} e^{-\frac{1}{2}u^2}\, du \qquad 4)$$

en kan in een tabel van de normale verdeling worden gevonden.

Voor $m \leq n \leq 10$, resp. voor $m+n \leq 40$, $n \geq 11$ en $m \leq n$ zijn tabellen van de eenzijdige overschrijdingskansen, resp. linker kritieke waarden beschikbaar (zie [6] tabel I en II). Deze tabellen gelden strikt genomen alleen als er geen gelijke waarnemingen zijn, maar zij geven voor het geval van weinig gelijke waarnemingen veelal een goede benadering.

Indien m en n klein zijn en bovendien onder de waarnemingen grote groepen gelijken voorkomen, kan men de overschrijdingskansen noch met behulp van de normale benadering van de verdeling van $\underline{W}$, noch met behulp van de genoemde tabellen bepalen. Hoe men in zulke gevallen dient te handelen, staat beschreven in [1].

<u>Opmerking</u>: Als er gelijke waarnemingen zijn, dan geldt

$$\frac{1}{3}\, m\, n\, (N+1) \;>\; \frac{m\, n\, (N^3 - D)}{3N(N-1)} .$$

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

----

4) De term -1 in de teller van $\frac{|W-\mu|-1}{\sigma}$ is de z.g. continuïteitscorrectie (zie [6], pag.16).

Literatuur

[1] van Eeden, Constance en Ir Doraline Wabeke, Handleiding voor de toets van WILCOXON (vervolg): Exacte behandeling als er gelijke waarnemingen zijn. Rapport S 176 (M 65A) van de Statistische Afdeling van het Mathematisch Centrum, Amsterdam, 1955.

[2] Hemelrijk, J., Note on WILCOXON's two sample test when ties are present, Ann.Math.Stat. 23 (1952), 133-135.

[3] Mann, H.B. and D.R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Ann.Math.Stat. 18 (1947), 50-60.

[4] van der Vaart, H.R., Gebruiksaanwijzing voor de toets van WILCOXON, Rapport S 32 (M 4) van de Statistische Afdeling van het Mathematisch Centrum, Amsterdam, 1950.

[5] Wilcoxon, F., Individual comparisons by ranking methods, Biometrics 1 (1945), 80-83.

[6] Wabeke, Ir Doraline en Constance van Eeden, Handleiding voor de toets van WILCOXON, Rapport S 176 (M 65) van de Statistische Afdeling van het Mathematisch Centrum, Amsterdam 1955.

Mathematisch Centrum,
2de Boerhaavestraat 49,
Amsterdam O.
Statistische Afdeling,
S47 (M7).

Maart, 1952.

De toets van Wilcoxon.[1]

Deze methode dient tot het toetsen van de hypothese H_o, inhoudende, dat twee steekproeven $x_1,\dots,x_n$ en $y_1,\dots,y_m$ afkomstig zijn uit één collectie (ook wel populatie of universum genaamd).

Voor het toetsen van de hypothese H_o wordt gebruik gemaakt van een toetsingsgrootheid $\underline{U}$ [2]), die als volgt uit de waarnemingen berekend wordt. Onderstellen we, dat de waarnemingen $x_1,\dots,x_n$ en $y_1,\dots,y_m$ naar opklimmende grootte gerangschikt zijn, dan bepalen we eerst het aantal waarnemingen uit de tweede steekproef, dat kleiner is dan de kleinste waarneming x_1 uit de eerste steekproef (bij gelijkheid tellen wij $\frac{1}{2}$ in plaats van 1). Noem dit aantal V_1. Vervolgens wordt het aantal waarnemingen uit de tweede steekproef bepaald, dat kleiner is dan de op één na kleinste waarneming x_2 uit de eerste steekproef (bij gelijkheid wordt weer $\frac{1}{2}$ in plaats van 1 geteld). Dit aantal noemen we V_2. Evenzo worden met betrekking tot $x_3,x_4,\dots,x_n$ de aantallen $V_3,V_4,\dots,V_n$ bepaald. De waarde U van de toetsingsgrootheid $\underline{U}$ wordt voor de twee steekproeven dan gegeven door

$$U=V_1+V_2+\dots+V_n.$$

Wanneer onder de waarnemingen niet te veel gelijken voorkomen, kan bewezen worden, dat de toetsingsgrootheid $\underline{U}$ onder de hypothese H_o voor grote waarden van n en m (beide $\geq$ 10) bij benadering een normale verdeling bezit. De waarnemingen $x_1,\dots,x_n$ en $y_1,\dots,y_m$ tezamen genomen vallen uiteen in een aantal groepen van gelijke waarnemingen. Noem het aantal van deze groepen k, dan is k minstens 1 (als alle waarnemingen gelijk zijn) en hoogstens m+n (als alle waarnemingen verschillend zijn).

---

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

2) Stochastische grootheden worden door onderstreping aangeduid.

Zijn $t_1,\dots,t_k$ de aantallen waarnemingen in deze groepen van gelijken, dan worden het gemiddelde μ en de variantie σ^2 van de toetsingsgrootheid $\underline{U}$ gegeven door

$$\mu(\underline{U})=\tfrac{1}{2}nm,$$

en

$$\sigma^2=\mathrm{Var}\ (\underline{U})=\frac{1}{12}\frac{nm}{(n+m)(n+m-1)}\left\{(n+m)^3+ - (t_1^3+t_2^3+\dots+t_k^3)\right\} \qquad {}^{1})$$

De grootheid $\mu(\underline{U})$ is dus onafhankelijk van de waarden vast. Indien de hypothese H_o niet vervuld is, zal de grootheid $\underline{U}$ grote of kleine waarden bezitten, al naar gelang $\underline{y}$ systematisch kleiner of groter is dan $\underline{x}$.
De (tweezijdige) toets bestaat nu daarin, dat men H_o verwerpt indien de gevonden waarde U van $\underline{U}$ te sterk van μ afwijkt, d.w.z. als

$$\frac{|U-\mu|}{\sigma} > \xi_\alpha\ , \qquad {}^{2})$$

waarin α de onbetrouwbaarheidsdrempel is en ξ_α volgt uit

$$\frac{1}{\sqrt{2\pi}}\int_{\xi_\alpha}^{\infty} e^{-\frac{1}{2}x^2}dx=\tfrac{1}{2}\alpha\ ,$$

en in een tabel van de normale verdeling kan worden opgezocht.
De (tweezijdige) overschrijdingskans k , behorende bij T, is gedefiniëerd als

$$k=\frac{2}{\sqrt{2\pi}}\int_{\left|\frac{U-\mu}{\sigma}\right|}^{\infty} e^{-\frac{1}{2}x^2}dx \qquad {}^{2})$$

en kan ook in een tabel van de normale verdeling worden gevonden.
Bij eenzijdige toetsing wordt α door 2α vervangen, resp. k gehalveerd.
Een bijzonder geval van het bovenstaande is, dat onder de waarnemingen voor $\underline{x}$ en $\underline{y}$ in 't geheel geen gelijken voorkomen. In dat geval kan de uitdrukking voor de variantie herleid worden tot

$$\sigma^2=\frac{1}{12}\,nm(n+m+1).$$

---

1) Deze formule is een door T.J.Terpstra gegeven vereenvoudiging van de door J.Hemelrijk([5] en [7]) afgeleide formule. De afleiding van deze vereenvoudigde formule zal nog gepubliceerd worden.

2) Deze formules berusten op de normale benadering van de verdeling van $\underline{U}$.

Indien n en m kleiner zijn dan 10, zijn tabellen beschikbaar voor het berekenen van de overschrijdingskans $\underline{k}$ voor de uit de steekproef bepaalde waarde U van $\underline{U}$ (zie [2] en [4]). Dergelijke tabellen bestaan echter niet voor het geval, dat gelijke waarnemingen optreden.

Opmerking. Men kan gemakkelijk bewijzen, dat de variantie van $\underline{U}$ door het optreden van gelijke waarnemingen vermindert. Het verschil, dat door deze gelijken optreedt, is echter in het algemeen gering. Men kan daarom in eerste instantie deze correctie op σ^2 verwaarlozen. De overschrijdingskansen, die men dan vindt, zijn iets te groot.

Litteratuur:

1.	F.Wilcoxon,	Individual comparisons by ranking methods, Biometrics 1 (1945), p.80-83.
2	H.B.Mann and D.R.Whitney	On a test of whether one of two random variables is stochastically larger than the other, Amer.Math.Stat. 18 (1947),p. 50-60.
3	H.R.van der Vaart	Some remarks on the power function of Wilcoxon's test for the problem of two samples, Proceedings van de Kon. Ned.Ak.v.Wet., 53 (1950),p. 494-520.
4	H.R.van der Vaart	Gebruiksaanwijzing voor de toets van Wilcoxon, met tabellen voor n en m$\leq$10, Rapport S32 (M4) (1950).
5	H.R.van der Vaart	De toets van Wilcoxon voor het probleem van twee steekproeven. (Cursus "Parametervrije Methoden", 1951-'52).
6	D.van Dantzig	Kadercursus Mathematische Statistiek, Math. Centrum, Amsterdam (1947-'50), hoofdst. 6, § 3.
7	J.Hemelrijk	Note on Wilcoxon's two sample test, when ties are present, Ann.Math.Stat. 23 (1952) no. 2.