

TWEE F-TABELLEN VOOR RANGEN

Michiel J.W. Jansen*

SAMENVATTING

Drie rangtoetsgrootheden (van Wilcoxon-Mann-Whitney, Kruskal-Wallis en Friedman) worden omgerekend naar één standaard toetsgrootheid: de variantieverhouding, F , van de rangen. De hierdoor verkregen uniformiteit is niet alleen van praktisch nut voor berekeningen, maar kan ook bijdragen tot een beter begrip: de grote overeenkomst tussen de klassieke variantieanalyse en de besproken rangtoetsen wordt beter zichtbaar. Er worden tabellen gegeven voor proeven met equireplicate behandelingen, zonder blokken en in volledige blokken.

Sleutelwoorden: variantieanalyse, rangen, Wilcoxon-Mann-Whitney, Kruskal-Wallis, Friedman

* DLO-Groep Landbouwwiskunde, Postbus 100, 6700 AC, Wageningen

1. Inleiding

Met de tekentoets en de rangtoetsen van Wilcoxon-Mann-Whitney, Kruskal-Wallis en Friedman wordt dezelfde vraag onderzocht als met een klassieke variantieanalyse: zijn er aantoonbare verschillen in de onderzochte variabele ten gevolge van de toegediende behandelingen? De overeenstemming blijft niet beperkt tot de vraagstelling: de vier genoemde parameter vrije toetsen zijn namelijk equivalent aan variantieanalyse op rangen. Het is echter traditie dat bij de vier toetsen telkens andere toetsgrootheden worden gebruikt. Die veelvormigheid is verwarrend en beneemt het zicht op de onderliggende eenheid.

De gelijkwaardigheid van de vier parameter vrije toetsen met een variantieanalyse op rangen kan op twee manieren zichtbaar gemaakt worden:

- (a) men kan de traditionele parameter vrije toetsgrootheden definiëren en berekenen via variantieanalyse op rangen;
- (b) men kan die toetsgrootheden vervangen door de variantieverhouding, F , van de rangen.

Conover en Iman (1981) bepleiten bij rangtoetsen het gebruik van klassieke toetsgrootheden zoals Students t of de variantieverhouding F : zij zijn dus voor (b). Ze hebben echter niet de benodigde tabellen gemaakt.

In dit stuk worden, voor kleine datasets, tabellen gepresenteerd met kritieke waarden van rang- F voor proeven met k behandelingen in n -voud, zonder blokken en in volledige blokken. De tabellen zijn berekend uit bestaande tabellen met kritieke waarden van de traditionele toetsgrootheden (Lindley en Scott, 1984). Voor zover de kritieke rang- F -waarden niet in de tabel voorkomen kan men de gewone F -tabel gebruiken als benadering. Volgens Conover en Iman is deze benadering goed, soms beter dan de gebruikelijke χ^2 -benadering.

Het bestaan van allerlei rangtoetsgrootheden die gelijkwaardig zijn aan de variantieverhouding F kan voor een deel historisch worden verklaard. Dat er bij vergelijking van twee behandelingen gewerkt wordt met iets eenvoudigs als een rangsom is niet verwonderlijk. En dat men de tekentoets niet vervangt door een F -toets is vanzelfsprekend. Maar waarom werkten Friedman, Kruskal en Wallis niet gewoon met rang- F ? De auteurs wisten ongetwijfeld dat hun toetsen variantieanalyses op rangen waren. Maar de variantieverhouding F was in de jaren '30, toen Friedman zijn toets introduceerde, nog niet de standaard toetsgrootheid van de variantieanalyse. Toen Kruskal en Wallis in de jaren '50 hun toets introduceerden lag het voor de hand een toetsgrootheid te kiezen die zoveel mogelijk overeenstemde met die van Friedman.

2. Berekening van rang- F uit gebruikelijke rang-toetsgrootheden Terminologie

In de berekeningen wordt gebruik gemaakt van het engelse jargon van de variantieanalyse: 'df' staat voor 'degrees of freedom', 'ss' staat voor 'sum of squares', 'ms' staat voor 'mean square'. Verder worden afkortingen gebruikt zoals 'treat' voor treatments en 'res' voor residual. De rang- F wordt berekend met een bij de blokstructuur passende variantieanalyse op de rangen

$$F = ms_{\text{treat}} / ms_{\text{res}}.$$

(1)

Berekening van rangen

Als er geen blokken zijn worden de rangen berekend over alle eenheden. Bij blokken worden de rangen binnen blokken berekend. Als er knopen (groepen gelijke waarden) optreden wordt de middenrang toegekend aan de elementen van de groep: het gemiddelde van de rangen die men zou krijgen als men willekeurig zou doornummeren binnen de groep. De middenrang wordt kortweg als rangen aangeduid. Bij de berekeningen wordt gebruik gemaakt van de volgende formule voor een rang-kwadratsom

$$\sum_{i=1}^n (i - (n+1)/2)^2 = (n-1) * n * (n+1) / 12. \quad (2)$$

Wilcoxon-Mann-Whitney

De toets betreft de situatie van twee behandelingen zonder blokken. We beschouwen het geval dat de behandelingen even vaak, n maal, worden toegediend. De gebruikte toetsgrootte is de rangsom, R , de som van de rangen van een van de behandelingen. Men berekent rang- F als volgt uit R .

$$\begin{aligned} \text{mean} &= (2*n+1)/2 \\ SS_{\text{treat+res}} &= (2*n-1) * 2*n * (2*n+1) / 12 \\ SS_{\text{treat}} &= 2 * n * (R/n - \text{mean}) * (R/n - \text{mean}) \\ SS_{\text{res}} &= SS_{\text{treat+res}} - SS_{\text{treat}} \\ ms_{\text{treat}} &= SS_{\text{treat}} \\ ms_{\text{res}} &= SS_{\text{res}} / (n-2) \\ F &= ms_{\text{treat}} / ms_{\text{res}} \end{aligned}$$

Omgekeerd zou men desgewenst R uit F kunnen berekenen:

$$R = n \left(\text{mean} \pm \sqrt{\frac{SS_{\text{treat+res}}}{2n} \frac{F}{F+n-2}} \right)$$

Deze rang- F -toets is equivalent aan de tweezijdige toets van Wilcoxon-Mann-Whitney.

Tekentoets

De tekentoets wordt toegepast bij 2 behandelingen in volledige blokken. De toets is gelijkwaardig aan een rangtoets. Het aantal positieve tekens kan worden omgerekend in rang- F . Omgekeerd kan men uit F weer het aantal positieve tekens berekenen, mits men weet of dat aantal hoog of laag is.

Deze mogelijkheid wordt hier niet verder uitgewerkt omdat het geen vereenvoudiging is over te gaan van een toetsgrootte die exact binomiaal verdeeld is naar een die bij benadering F -verdeeld is.

Kruskal-Wallis

De toets betreft de situatie van k behandelingen zonder blokken. We beschouwen het geval dat de behandelingen even vaak, n maal, worden toegediend. De Kruskal Wallis grootheid K kan worden gedefinieerd als

$$K = ss_{treat} / ms_{treat+res} \quad (3)$$

Deze definitie geldt ook als er knopen zijn.

Als er geen knopen zijn kan (3) via (1) en (2) uitgewerkt worden als

$$K = \frac{12}{N(N+1)} \sum_{j=1}^k n \left(\frac{R_j}{n} - \frac{N+1}{2} \right)^2 = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n} - 3(N+1)$$

waarin $N = k \cdot n$, en waarin R_j de som is van de rangen van behandeling j. Als er knopen zijn wordt de uitwerking van (3) nog minder aantrekkelijk (zie bijvoorbeeld Lehmann, 1975).

De omrekening van K naar rang-F kan blijkens (1) en (3) als volgt gebeuren:

$$F = (df_{res} / df_{treat}) * (K / (df_{treat+res} - K)) ,$$

waarin $df_{treat+res} = n \cdot k - 1$, $df_{treat} = k - 1$, $df_{res} = n \cdot k - k$. Omgekeerd kan men desgewenst K uit rang-F berekenen:

$$K = df_{treat+res} * df_{treat} * F / (df_{res} + df_{treat} * F) .$$

Friedman

De toets betreft de situatie van k behandelingen in volledige blokken. We beschouwen het geval dat de behandelingen even vaak, n maal, worden toegediend. De Friedman grootheid Q wordt gedefinieerd als

$$Q = ss_{treat} / ms_{treat+res} \quad (4)$$

Deze definitie geldt ook als er knopen zijn.

Als er geen knopen zijn kan (4) via (1) en (2) uitgewerkt worden tot

$$Q = \frac{12}{k(k+1)} \sum_{j=1}^k n \left(\frac{R_j}{n} - \frac{k+1}{2} \right)^2 = \frac{12}{nk(k+1)} \sum_{j=1}^k R_j^2 - 3n(k+1)$$

waarin R_j de rangsom is van de j-de behandeling. Als er knopen zijn wordt de uitwerking van (4) weer weinig fraai (zie bijvoorbeeld Lehmann, 1975).

De omrekening van Q naar rang-F kan blijkens (1) en (4) als volgt gebeuren:

$$F = (df_{res} / df_{treat}) * (Q / (df_{treat+res} - Q)) ,$$

waarin $df_{treat+res} = n \cdot k - n$, $df_{treat} = k - 1$, $df_{res} = (n-1) \cdot (k-1)$. Omgekeerd kan men desgewenst Q uit rang-F berekenen:

$$Q = df_{treat+res} * df_{treat} * F / (df_{res} + df_{treat} * F) .$$

3. Voorbeelden

Wilcoxon-Mann-Whitney

Een volledig gelote proef met twee behandelingen in 8-voud leverde de volgende rangen

behandeling	rangen							
a	1	2	3	4	7	9	10	15
b	5	6	8	11	12	13	14	16

Variantieanalyse (zonder blokken) op de rangen levert $F = 3.78$. De rang-F tabel levert $P \leq 0.10$. Wilcoxon's rangsom van behandeling a is 51 (de ook wel gehanteerde Mann-Whitney grootte is $51 - 8 * 4.5 = 15$) een en ander natuurlijk weer met bijbehorende tweezijdige $P \leq 0.10$.

Kruskal-Wallis

Een volledig gelote proef met 4 behandelingen in 4-voud leverde de volgende rangen

controle	13	14	15	16
a	5.5	5.5	7	8
b	9	10	11	12
c	1	2	3.5	3.5

De Kruskal-Wallis grootte is volgens de lastige formule met knopen-correctie uit Lehmann gelijk aan $K = 14.16$, met bijbehorende $P \leq 0.001$.

Variantieanalyse (zonder blokken) op de rangen levert $F = 67.37$, hetgeen een $P \leq 0.01$ levert volgens onze tabel (de kleinste P daarin is 0.01.) Uit rang-F berekenen we vlot K

$$K = df_{\text{treat}+\text{res}} * df_{\text{treat}} * F / (df_{\text{res}} + df_{\text{treat}} * F) = 14.16.$$

Friedman

In Hollander en Wolfe (1973) wordt de volgende dataset met knopen besproken als voorbeeld bij de toets van Friedman.

blok	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
a	1	3	1	3	3	1	2.5	2	3	3	3	3	3	3	2.5	1	2.5	1	3	3	3	2.5
b	2	1	3	2	2	2	2.5	3	2	2	2	2	1	2	2.5	3	2.5	2	2	2	2	2.5
c	3	2	2	1	1	3	1	1	1	1	1	2	1	1	2	1	3	1	1	1	1	1

De Friedman toetsgrootte met correctie voor knopen is volgens de auteurs $Q = 11.1$. Blijkens de Friedman-tabel is de bijbehorende $P \leq 0.01$.

Variantieanalyse (met blokken) op de rangen levert $F = 7.12$. Ook hieruit kan men de Friedman grootte berekenen:

$$Q = df_{\text{treat+res}} * df_{\text{treat}} * F / (df_{\text{res}} + df_{\text{treat}} * F) = 11.1.$$

Het is makkelijker rang-F direct te vergelijken met de kritieke waarden uit de rang-F tabel voor volledige blokken. Het resultaat is natuurlijk weer $P \leq 0.01$.

4. Opmerking bij de tabellen

De kritieke F-waarden worden in het algemeen kleiner als het aantal herhalingen stijgt, doch niet altijd. De uitzonderingen worden veroorzaakt doordat de exacte p-waarden niet steeds gelijk zijn aan de onbetrouwbaarheidsdrempels.

Literatuur

Conover, W.J. & Iman, R.L., *Rank transformations as a bridge between parametric and nonparametric statistics*, The American Statistician, aug. 1981, p124-133

Hollander, M. & Wolfe, D.A., *Nonparametric statistical methods*, Wiley, 1973, p140-141

Lehmann, E.L., *Nonparametrics: statistical methods based on ranks*, McGraw-Hill, 1975

Lindley, D.V. & Scott, W.F., *New Cambridge elementary statistical tables*, Cambridge University Press, 1984

ontvangen	5 - 5-1991
geaccepteerd	3 - 12-1992

Twee tabellen met kritieke waarden van rang-F

De tabellen betreffen proeven waarin de behandelingen even vaak zijn toegepast, zonder blokken of in volledige blokken.

De kritieke waarden zijn getablleerd voor onbetrouwbaarheidsdrempels: 10%, 5% en 1%.

Als er blokken zijn worden de rangen berekend binnen blokken, anders over de hele dataset. Rang-F wordt berekend met een bij de blokstructuur passende variantieanalyse.

Een sterretje duidt aan dat er geen kritieke waarde bestaat met de gewenste eigenschap. Kritieke waarde oneindig duidt aan dat het residu gelijk aan 0 is.

Het geval van een volledige blokkenproef met 2 behandelingen is niet getabelleerd: voor dit geval, de tekentoets, wordt verwezen naar een tabel van de binomiale verdeling.

Voor zover een proefopzet niet voorkomt in de tabel kan men een klassieke F-tabel gebruiken als benadering.

Kritieke waarden van rang-F
k behandelingen in n volledige blokken

k	n	10%	5%	1%
3	2	*	*	*
3	3	∞	∞	*
3	4	9.00	13.00	∞
3	5	4.33	7.11	21.00
3	6	4.00	7.00	15.00
3	7	3.80	6.25	10.33
3	8	3.42	4.49	9.00
3	9	3.57	4.23	9.05
3	10	3.00	4.04	8.31
3	11	3.01	4.23	7.54
3	12	3.02	4.09	7.21
3	13	2.70	4.09	6.78
3	14	2.93	3.65	6.30
3	15	2.76	3.80	5.94
3	16	2.70	3.82	6.22
3	17	2.80	3.51	6.02
3	18	2.60	3.63	5.67
3	19	2.76	3.66	6.07
3	20	2.65	3.55	5.76
3	21	2.67	3.40	5.64
3	22	2.53	3.37	5.47
3	23	2.66	3.52	5.64
3	24	2.72	3.44	5.49
3	25	2.60	3.32	5.24
3	26	2.57	3.31	5.45
3	27	2.50	3.25	5.48
3	28	2.40	3.55	5.32
3	29	2.66	3.40	5.26
3	30	2.56	3.34	5.30
3	31	2.54	3.21	5.29
3	32	2.49	3.24	5.24
3	33	2.50	3.24	5.15
3	34	2.49	3.23	5.15

Kritieke waarden van rang-F
k behandelingen in n volledige blokken
(vervolg)

k	n	10%	5%	1%
4	3	5.50	9.25	∞
4	4	3.32	5.57	12.00
4	5	2.94	4.33	7.90
4	6	2.76	3.65	6.54
4	7	2.65	3.55	6.05
4	8	2.49	3.28	5.44
4	9	2.38	3.17	5.28
4	10	2.42	3.10	4.98
4	11	2.35	3.04	4.83
4	12	2.33	2.99	4.71
4	13	2.24	3.00	4.63
4	14	2.31	2.92	4.55
4	15	2.27	2.90	4.49
4	16	2.27	2.91	4.43
4	17	2.26	2.89	4.43
4	18	2.26	2.84	4.31
4	19	2.26	2.88	4.31
4	20	2.21	2.84	4.31
5	3	3.29	4.92	10.83
5	4	2.71	3.67	7.00
5	5	2.49	3.25	5.62
5	6	2.38	3.04	4.89
5	7	2.30	2.91	4.57
5	8	2.22	2.82	4.37
5	9	2.19	2.76	4.22
6	3	2.77	3.83	7.26
6	4	2.45	3.18	5.23
6	5	2.25	2.89	4.50
6	6	2.16	2.72	4.16

Kritieke waarden van rang-F
k behandelingen in n-voud
(geen blokken)

k	n	10%	5%	1%
2	3	13.50	*	*
2	4	8.40	19.20	*
2	5	4.31	9.19	25.00
2	6	3.93	6.50	14.80
2	7	3.56	6.23	13.41
2	8	3.78	5.06	11.90
2	9	3.38	5.43	10.63
2	10	3.41	5.05	9.59
2	11	3.37	4.72	9.43
2	12	3.30	4.75	9.14
2	13	3.21	4.71	8.80
2	14	3.12	4.39	8.45
2	15	3.02	4.54	8.10
2	16	3.07	4.43	8.07
2	17	2.96	4.32	7.98
2	18	2.97	4.35	7.85
2	19	2.97	4.22	7.71
2	20	2.95	4.22	7.75
3	2	15.98	*	*
3	3	4.10	7.00	27.00
3	4	3.30	4.83	10.29
3	5	2.90	4.22	8.00
3	6	2.82	3.88	7.02
3	7	2.68	3.69	6.49
3	8	2.62	3.54	6.12
3	9	2.57	3.48	5.89
4	2	5.67	9.87	26.69
4	3	3.23	4.67	9.25
4	4	2.73	3.73	6.50
5	2	4.32	5.86	14.62
5	3	2.75	3.68	6.71
6	2	3.44	4.93	10.24