stichting
mathematisch
centrum

AFDELING MATHEMATISCHE STATISTIEK SN 4/74 FEBRUARI

F.H. RUYMGAART

HANDLEIDING VOOR HET GEBRUIK VAN DE TOETS VAN m RANGSCHIKKINGEN EN EEN DAARAAN ONTLEENDE METHODE VOOR MEERVOUDIGE VERGELIJKING

Voorlopige uitgave

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS (MOS) onderwerpen classificatie schema (1970): 62G10, 62J15

Handleiding voor het gebruik van de toets van m rangschikkingen en een daaraan ontleende methode voor meervoudige vergelijking *)

F.H. Ruymgaart

SAMENVATTING

Men wenst de uitwerking van een p-tal behandelingen te onderzoeken. Daartoe beschikt men over m onderling onafhankelijke p-dimensionale stochastische vectoren. Van ieder der vectoren vertegenwoordigt de i-de component de uitwerking van de i-de behandeling. Een verdelingsvrije statistische procedure baseert men doorgaans op de methode van m rangschikkingen. In deze handleiding beschrijven we een tweetal op deze methode gebaseerde asymptotische procedures ($m \to \infty$), één voor het toetsen van de nulhypothese dat de behandelingen niet verschillen en één voor een meervoudige vergelijking van de behandelingen. Beide procedures blijven toepasbaar indien er vectoren met groepen onderling gelijke componenten (knopen) voorkomen. De nadruk ligt op toepassing in de praktijk; bewijzen worden niet gegeven.

---

*) Dit rapport is bedoeld voor publicatie in een tijdschrift.

1. HET PROBLEEM, NULHYPOTHESE EN ALTERNATIEF

Men wenst de uitwerking van een p-tal behandelingswijzen te vergelijken. Daartoe kiest men aselect uit m eventueel verschillende populaties een object en onderwerpt dit aan de eerste behandeling. Vervolgens kiest men uit ieder der m populaties aselect een tweede object en onderwerpt dit aan de tweede behandeling, enz. Tenslotte kiest men uit ieder der m populaties aselect een p-de object en onderwerpt dit aan de p-de behandeling. Na afloop van de behandeling wordt aan ieder object een grootheid gemeten die geacht wordt het effect van de behandeling te weerspiegelen. Omdat het j-de element uit de i-de populatie een element is van een aselecte steekproef van de omvang p uit deze populatie, is de variabele die aan dit element gemeten wordt een stochastische grootheid. Deze stochastische grootheid zullen wij met $\underline{x}_{ij}$ noteren*) $(i=1,\ldots,m;j=1,\ldots,p)$. Wij zeggen dat de grootheden $\underline{x}_{i1},\ldots,\underline{x}_{ip}$ de i-de *regel* vormen $(i=1,\ldots,m)$. Gezien de proefopzet bezitten de stochastische grootheden in iedere regel een simultane kansverdeling, zodat $(\underline{x}_{i1},\ldots,\underline{x}_{ip})$ voor iedere i een stochastische vector is. Bovendien zullen we aannemen dat de stochastische vectoren $(\underline{x}_{11},\ldots,\underline{x}_{1p}),\ldots,(\underline{x}_{m1},\ldots,\underline{x}_{mp})$ stochastisch onafhankelijk zijn. Dit betekent dat alle stochastische grootheden $\underline{x}_{11},\ldots,\underline{x}_{1p},\ldots,\underline{x}_{m1},\ldots,\underline{x}_{mp}$ een simultane kansverdeling bezitten.

Alvorens tot een beschrijving van het probleem en de procedures te kunnen overgaan, moeten wij eerst wat notaties invoeren. De simultane verdelingsfunctie $P(\underline{x}_{i1}\le x_1,\ldots,\underline{x}_{ip}\le x_p)$ zullen we met $F_i(x_1,\ldots,x_p)$ aangeven $(-\infty<x_1<\infty,\ldots,-\infty<x_p<\infty)$. De marginale verdelingsfuncties $P(\underline{x}_{ij}\le x)$ noteren we met $F_{ij}(x)$. Aan de F_i worden geen gladheidseisen opgelegd. In het bijzonder behoeven de F_i niet continu te zijn, waardoor het mogelijk is dat er onder de stochastische variabelen in de i-de regel met positieve kans gelijken voorkomen. Om de gedachten te bepalen zullen wij veronderstellen dat de in de i-de regel naar grootte geordende variabelen

$$\underline{x}_i^{(1)} \le \underline{x}_i^{(2)} \le \ldots \le \underline{x}_i^{(p)}$$

---

*) Stochastische grootheden worden onderstreept.

als volgt in groepen van gelijken, *knopen* geheten, uiteenvallen:

$$\underline{x}_i^{(1)} = \ldots = \underline{x}_i^{(\underline{\tau}_{i,1})} < \underline{x}_i^{(\underline{\tau}_{i,1}+1)} = \ldots = \underline{x}_i^{(\underline{\tau}_{i,1}+\underline{\tau}_{i,2})} < \ldots <$$
$$< \underline{x}_i^{(\underline{\tau}_{i,1}+\underline{\tau}_{i,2}+\ldots+\underline{\tau}_{i,\underline{\nu}_i-1}+1)} = \ldots = \underline{x}_i^{(p)}.$$

De i-de regel valt dus uiteen in $\underline{\nu}_i$ knopen, de ℓ-de knoop telt $\underline{\tau}_{i,\ell}$ waarnemingen ($\ell=1,\ldots,\underline{\nu}_i$) en $\sum_{\ell=1}^{\underline{\nu}_i} \underline{\tau}_{i,\ell} = p$. Zowel het aantal knopen als de omvangen der knopen zijn stochastische grootheden.

De in deze handleiding beschreven procedures zijn gebaseerd op het over iedere knoop afzonderlijk *gemiddelde rangnummer* $\underline{R}_{ij}$ van $\underline{x}_{ij}$ in de i-de regel ($i=1,\ldots,m$). Dit wordt als volgt gedefinieerd:

$$\begin{cases} \text{voor alle } \underline{x}_{ij} \text{ in de 1-ste knoop is } \underline{R}_{ij} = \dfrac{\sum_{k=1}^{\underline{\tau}_{i,1}} k}{\underline{\tau}_{i,1}}\,; \\[2ex] \text{voor alle } \underline{x}_{ij} \text{ in de 2-de knoop is } \underline{R}_{ij} = \dfrac{\sum_{k=\underline{\tau}_{i,1}+1}^{\underline{\tau}_{i,1}+\underline{\tau}_{i,2}} k}{\underline{\tau}_{i,2}}\,; \\ \ldots\ldots \\ \text{voor alle } \underline{x}_{ij} \text{ in de } \underline{\nu}_i\text{-de knoop is } \underline{R}_{ij} = \dfrac{\sum_{k=\underline{\tau}_{i,1}+\ldots+\underline{\tau}_{i,\underline{\nu}_i-1}+1}^{p} k}{\underline{\tau}_{i,\underline{\nu}_i}}\,. \end{cases}$$

Alle variabelen in één en dezelfde knoop krijgen dus een gelijk gemiddeld rangnummer, dat geen natuurlijk getal behoeft te zijn. Ongeacht de waarden der $\underline{\tau}_{i,\ell}$ geldt steeds

$$p^{-1} \sum_{j=1}^{p} \underline{R}_{ij} = p^{-1} \sum_{k=1}^{p} k = (p+1)/2.$$

Een knoop noemen we *eigenlijk* als zijn omvang groter dan 1 is. De i-de regel noemen we *nutteloos* of *nuttig* al naar gelang $\underline{\nu}_i = 1$ (zodat de i-de regel slechts uit één enkele knoop bestaat) of $\underline{\nu}_i > 1$. Indien $\underline{\nu}_i = p$ en dus $\underline{\tau}_{i,1} = \ldots = \underline{\tau}_{i,p} = 1$ bevat de i-de regel in het geheel geen eigenlijke knopen. In het laatste geval komt het gemiddelde rangnummer van $\underline{x}_{ij}$ overeen met het *rangnummer* van $\underline{x}_{ij}$, hetgeen gedefinieerd is als het aantal variabelen in de i-de regel dat kleiner dan of gelijk aan $\underline{x}_{ij}$ is. De collectie rangnummers $\underline{R}_{i1},\ldots,\underline{R}_{ip}$ is dan een permutatie van de getallen $1,\ldots,p$. Hieronder volgt een rekenvoorbeeld voor $p = 7$ en een willekeurige regel i.

WAARNEMINGEN	
$x_{i1}=1.5$, $x_{i2}=2.3$, $x_{i3}=1.5$, $x_{i4}=2.1$, $x_{i5}=2.3$, $x_{i6}=1.5$, $x_{i7}=1.8$	
AANTAL KNOPEN	AANTAL EIGENLIJKE KNOPEN
$\nu_i=4$	2
OMVANG KNOPEN	
$\tau_{i,1}=3$, $\tau_{i,2}=1$, $\tau_{i,3}=1$, $\tau_{i,4}=2$	
GEMIDDELDE RANGNUMMERS	
$R_{i1}=2$, $R_{i2}=6.5$, $R_{i3}=2$, $R_{i4}=5$, $R_{i5}=6.5$, $R_{i6}=2$, $R_{i7}=4$	

In deze handleiding beschrijven we zowel een toetsings- als een meervoudige vergelijkingsprocedure voor het onderzoek naar de uitwerking der behandelingen. Het is duidelijk dat een nutteloze regel geen informatie geeft over het verschil in uitwerking der behandelingen. De statistische grootheden waarop de procedures gebaseerd zijn (zie (2.1) en (3.1)) leveren dan ook dezelfde waarden wanneer zij alleen over de nuttige regels berekend worden als wanneer zij over alle regels berekend worden. Wij zullen daarom aannemen dat reeds alle nutteloze regels weggelaten zijn, dus dat van meet af aan *alle* m *regels nuttig* zijn. Deze aanname maakt hier en daar een eenvoudiger formulering mogelijk zonder dat aan algemeenheid wordt ingeboet.

Het *toetsingsprobleem* is het toetsen van de nulhypothese H_0 dat er geen verschil is in de uitwerking van de behandelingen tegen het alternatief H_1 dat de uitwerking van tenminste één behandeling stochastisch groter is dan die van één der overige. Formeel:

H_0: De $\underline{x}_{i1},\ldots,\underline{x}_{ip}$ zijn verwisselbaar*), zodat o.a. $F_{i1}(x) = \ldots = F_{ip}(x)$ voor alle x en alle $i=1,\ldots,m$.

H_1: Voor tenminste één paar j,k (met $j\neq k$) geldt dat $P(\underline{x}_{ij}>\underline{x}_{ik}) > \frac{1}{2}$ voor alle $i=1,\ldots,m$ (of $< \frac{1}{2}$ voor alle $i=1,\ldots,m$).

Om de *meervoudige vergelijkingsprocedure* te kunnen beschrijven moeten wij het bovengenoemde alternatief inperken tot

H_1': Er bestaat een p-tal reële getallen $\theta_1,\ldots,\theta_p$, niet alle gelijk, zodat de $\underline{x}_{i1} + \theta_1,\ldots,\underline{x}_{ip} + \theta_p$ verwisselbaar zijn; dit betekent o.a. dat $F_{i1}(x-\theta_1) = \ldots = F_{ip}(x-\theta_p)$ voor alle x en alle $i=1,\ldots,m$.

De meervoudige vergelijkingsprocedure dient voor het opsporen van alle tweetallen θ_j, θ_k die als onderling verschillend kunnen worden aangewezen.

Voor beide problemen bestaat een procedure gebaseerd op de toetsingsgrootheid van FRIEDMAN, de ontwerper van de besproken methode. De voor de procedures gebruikte statistische grootheden zijn slechts functies van de gemiddelde rangnummers. Onder de nulhypothese H_0 zijn de voorwaardelijke verdelingen der statistische grootheden, gegeven de omvangen der knopen, onafhankelijk van de onderliggende verdelingsfuncties $F_1,\ldots,F_m$. Onder de nulhypothese H_0 zijn de onvoorwaardelijke verdelingen der statistische

---

*) Een p-tal stochastische variabelen $\underline{y}_1,\ldots,\underline{y}_p$ heet *verwisselbaar* indien voor iedere permutatie $i_1,\ldots,i_p$ van $1,\ldots,p$ de simultane kansverdeling van de vector $(\underline{y}_{i_1},\ldots,\underline{y}_{i_p})$ gelijk is aan die van de vector $(\underline{y}_1,\ldots,\underline{y}_p)$. In het bijzonder zijn deze stochastische variabelen verwisselbaar indien zij stochastisch onafhankelijk zijn en dezelfde verdeling bezitten, dus een steekproef uit de een of andere kansverdeling vormen. Ook voor steekproeven zonder teruglegging uit eindige populaties geldt verwisselbaarheid.

grootheden slechts dan onafhankelijk van $F_1,\ldots,F_m$ als er in geen van de regels eigenlijke knopen optreden. (In dit geval neemt immers de vector van rangnummers $\underline{R}_{i1},\ldots,\underline{R}_{ip}$ iedere permutatie der getallen $1,\ldots,p$ met gelijke kans $1/p!$ aan.) *Wel geldt echter dat de onvoorwaardelijke kansverdeling der statistische grootheden asymptotisch, voor* $m \to \infty$, *onder de nulhypothese* H_0 *onafhankelijk is van de onderliggende kansverdelingen* $F_1,\ldots,F_m$, *ongeacht of deze continu zijn of niet.* Om deze reden en ook omdat voorwaardelijke exacte verdelingen niet getabelleerd zijn zullen we ons tot *asymptotische procedures* beperken.

Laten we van de grote hoeveelheid literatuur over dit onderwerp naast een artikel van BENARD & Van ELTEREN (1953) slechts enkele van de meest recente boeken vermelden. Zowel in het boek van HÁJEK & ŠIDÁK (1967) als in dat van HÁJEK (1969) wordt de toets besproken. In het eerste boek ligt de nadruk op de theorie. Het tweede is een inleiding in de theorie van de rangtoetsen die met het oog op toepassing in de praktijk geschreven is; er wordt dan ook veel aandacht aan de behandeling van knopen geschonken. Laten we tenslotte de boeken van MILLER (1966) en PURI & SEN (1971) vermelden, waarin niet alleen de toets maar ook de meervoudige vergelijking wordt behandeld.

2. DE TOETSINGSPROCEDURE

De toets is gebaseerd op de toetsingsgrootheid

$$(2.1) \qquad \underline{Q}_m = \frac{12\, m \sum_{j=1}^{p} \left[m^{-1} \sum_{i=1}^{m} \underline{R}_{ij} - (p+1)/2\right]^2}{p(p+1) - \dfrac{\sum_{i=1}^{m} \sum_{\ell=1}^{\underline{\nu}_i} \underline{\tau}_{i,\ell}(\underline{\tau}_{i,\ell}^2 - 1)}{m(p-1)}} \; ;$$

de door $\underline{Q}_m$ aangenomen waarde noteren we met Q.

Men kan bewijzen dat onder de nulhypothese geldt:

$$(2.2) \qquad \lim_{m\to\infty} P(\underline{Q}_m \le Q \mid H_0) = P(\underline{\chi}_{p-1}^2 \le Q),$$

voor alle Q. Hierin stelt $\underset{\sim}{\chi}^2_{p-1}$ een stochastische grootheid voor, die een chi-kwadraat verdeling met p-1 vrijheidsgraden bezit. Hierop kan men een toetsingsprocedure baseren die asymptotisch, voor $m \to \infty$, de onbetrouwbaarheidsdrempel α $(0<\alpha<1)$ bezit, door H_0 ten gunste van H_1 (of H_1') te verwerpen indien de aangenomen waarde Q voldoet aan

$$(2.3) \qquad P(\underset{\sim}{\chi}^2_{p-1} \geq Q) \leq \alpha.$$

Voor een tabel van de kansverdeling van $\underset{\sim}{\chi}^2_{p-1}$ zie bijv. MOOD & GRAYBILL (1963; tabel III), OWEN (1962; paragraaf 3.1) en PEARSON & HARTLEY (1958; tabel 7).

Indien er in geen van de regels eigenlijke knopen optreden is de noemer van de toetsingsgrootheid $\underset{\sim}{Q}_m$ in (2.1) gelijk aan $p(p+1)$, zodat

$$(2.4) \qquad \underset{\sim}{Q}_m = \frac{12 \sum_{j=1}^{p} \left(\sum_{i=1}^{m} \underset{\sim}{R}_{ij} \right)^2}{mp(p+1)} - 3m(p+1).$$

Onder de aanname dat H_0 vervuld is, heeft men in dit geval de *exacte* rechter overschrijdingskansen berekend voor kleine waarden van m en p. Zie bijv. OWEN (1962; paragraaf 14.1) voor $p = 3$, $m = 2(1)15$ en $p = 4$, $m = 2(1)8$. Indien men de door $\underset{\sim}{Q}_m$ aangenomen waarde Q vervangt door

$$(2.5) \qquad Q' = \frac{mp(p+1)Q}{12},$$

kan men deze rechter overschrijdingskansen ook vinden in PEARSON & HARTLEY (1958; tabel 46) voor $p = 3$, $m = 3(1)10$ en $p = 4$, $m = 3(1)6$ en tenslotte $p = 5$, $m = 3$. Voor grote waarden van m kan men weer de boven beschreven asymptotische procedure toepassen, waarbij voor niet al te grote waarden van m toepassing van een *continuïteitscorrectie* zin heeft. Men dient dan in (2.3) de door $\underset{\sim}{Q}_m$ aangenomen waarde Q te vervangen door

$$(2.6) \qquad Q'' = \frac{m^2(p^3-p)Q - 12m(p-1)}{m^2(p^3-p)+24},$$

zie bijv. KENDALL (1962).

3. DE MEERVOUDIGE VERGELIJKINGSPROCEDURE

Beschouw voor alle $1 \leq j < k \leq p$ de stochastische grootheden

$$\underline{W}_{m,j,k} = \frac{\sqrt{12m}\,\left|\sum_{i=1}^{m} m^{-1}(\underline{R}_{ij}-\underline{R}_{ik})\right|}{\sqrt{p(p+1)-\dfrac{\sum_{i=1}^{m}\sum_{\ell=1}^{\underline{v}_i}\underline{\tau}_{i,\ell}(\underline{\tau}_{i,\ell}^2-1)}{m(p-1)}}}, \tag{3.1}$$

en zij $W_{j,k}$ de door $\underline{W}_{m,j,k}$ aangenomen waarde.

Men kan bewijzen dat onder de nulhypothese geldt:

$$\lim_{m\to\infty} P\left(\max_{1\leq j<k\leq p}\{\underline{W}_{m,j,k}\} \leq W \,|\, H_0\right) = P(\underline{\rho}_p \leq W), \tag{3.2}$$

voor alle W. Hierin is de stochastische grootheid $\underline{\rho}_p$ als volgt gedefinieerd. Zij $\underline{u}_1,\ldots,\underline{u}_p$ een steekproef van de omvang p uit een standaardnormale verdeling en laat $\underline{u}^{(1)}\leq\underline{u}^{(2)}\leq\ldots\leq\underline{u}^{(p)}$ de naar grootte geordende variabelen voorstellen. Dan is $\underline{\rho}_p = \underline{u}^{(p)} - \underline{u}^{(1)}$, de zgn. *steekproefwijdte*. Wij zullen onze aandacht nu in het bijzonder op het alternatief H_1' richten en wijzen die tweetallen θ_j, θ_k $(j\neq k)$ als verschillend aan, waarvoor

$$W_{j,k} \geq \rho_{p,\alpha}. \tag{3.3}$$

Hierin is $0 < \alpha < 1$ en $\rho_{p,\alpha}$ het getal waarvoor

$$P(\underline{\rho}_p \geq \rho_{p,\alpha}) = \alpha. \tag{3.4}$$

Daarbij wordt geconcludeerd tot $\theta_j > \theta_k$ indien

$$\sum_{i=1}^{m} R_{ij} > \sum_{i=1}^{m} R_{ik}, \tag{3.5}$$

en omgekeerd.

Uit (3.2) volgt dat onder de nulhypothese H_0 de kans om ten onrechte tenminste één paar θ_j, θ_k als verschillend aan te wijzen asymptotisch, voor $m \to \infty$, gelijk is aan α. Is H_0 niet vervuld, zodat niet alle $\theta_1,\ldots,\theta_p$ onderling gelijk zijn, dan is de kans om tenminste één tweetal θ_j, θ_k ten onrechte als verschillend aan te wijzen asymptotisch, voor $m \to \infty$, niet groter dan onder H_0, dus kleiner dan of gelijk aan α.

Een tabel van de exacte kansverdeling van $\underline{\rho}_p$ kan men vinden in OWEN (1962; paragraaf 6.1) voor p = 2(1)20, p = 22(2)40 en p = 50(10)100 en in PEARSON & HARTLEY (1958; tabel 23, zie ook tabel 22) voor p = 2(1)20. De kansverdeling van de steekproefwijdte (Engels: *range*) van een steekproef uit de standaardnormale verdeling kan men opvatten als de kansverdeling van een bijzonder geval van de zgn. *studentized range*. Daarom kan men de kansverdeling van $\underline{\rho}_p$ doorgaans ook vinden in tabellen van de kansverdeling van de studentized range. Een dergelijke tabel kunnen we voor ons doel toepassen door de parameter die het aantal vrijheidsgraden aangeeft van de chi-kwadraat grootheid, welke in de definitie van de studentized range optreedt, ∞ te kiezen en door de tweede parameter gelijk aan p te stellen. (Zie bijv. MILLER (1966; tabel I) en MOOD & GRAYBILL (1963; tabel VI-VIII).)

Indien er in geen van de regels eigenlijke knopen optreden kunnen wij de $\underline{W}_{m,j,k}$ schrijven als

$$(3.6) \qquad \underline{W}_{m,j,k} = \frac{\sqrt{12}\,\left|\sum_{i=1}^{m} (\underline{R}_{ij} - \underline{R}_{ik})\right|}{\sqrt{mp(p+1)}} ,$$

aangezien in dit geval de noemer in (3.1) gelijk is aan $\sqrt{p(p+1)}$.

LITERATUUR

[1] BENARD, A. & ELTEREN, Ph. Van (1953). A generalization of the method of m rankings. *Indagationes Math.* 15, 358-369.

[2] HÁJEK, J. (1969). *A Course in Nonparametric Statistics.* Holden-Day, San Francisco.

[3] HÁJEK, J. & ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[4] KENDALL, M.G. (1962). *Rank Correlation Methods.* Derde druk, Griffin, London.

[5] MILLER, R.G. (1966). *Simultaneous Statistical Inference.* McGraw-Hill, New York.

[6] MOOD, A.M. & GRAYBILL, F.A. (1963). *Introduction to the Theory of Statistics.* Tweede druk, McGraw-Hill, New York.

[7] OWEN, D.B. (1962). *Handbook of Statistical Tables.* Addison-Wesley, Reading, Mass.

[8] PEARSON, E.S. & HARTLEY, H.O. (1958). *Biometrika Tables for Statisticians*, vol. 1. Tweede druk, Cambridge.

[9] PURI, M.L. & SEN, P.K. (1971). *Nonparametric Methods in Multivariate Analysis.* Wiley, New York.