STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 47 (M 14)

Methode der m rangschikkingen

MATHEMATISCH CENTRUM
2de Boerhaavestr. 49
A M S T E R D A M O.
Statistische Afdeling
S47 (M14)

Methode der m rangschikkingen [1])

Een duidelijk voorstelling van deze toetsingsmethode verkrijgt men door n elementen te beschouwen, die een bepaald kenmerk, eventueel in verschillende mate, bezitten. dit kenmerk wordt door m waarnemers beoordeeld en ieder van deze waarnemers rangschikt deze n elementen volgens zijn beoordeling naar opklimmende waardering. Op deze wijze ontstaan m rijen van rangschikkingen. We willen nu een maat aangeven voor de overeenstemming tussen deze rangschikkingen, m.a.w. een maat voor de overeenstemming tussen de m beoordelingen. De hypothese $\mathcal{H}_0$, die met deze methode getoetst kan worden, houdt in dat er geen overeenstemming tussen de waarnemers bestaat; precieser gezegd, dat alle rangschikkingen onafhankelijk van elkaar op toevallige wijze zijn ontstaan. Dit is b.v. het geval, als het betrokken kenmerk in werkelijkheid voor alle elementen dezelfde waarde bezit.

We kunnen de afleiding voor de maat van overeenstemming het eenvoudigst geven aan de hand van een voorbeeld.

elementen		A	B	C	D	E	F
rangnummers toegekend door waarnemer	a	5	4	1	6	3	2
	b	2	3	1	5	6	4
	c	4	1	6	3	2	5
	d	4	3	2	5	1	6
		15	11	10	19	12	17

De som van alle rangnummers is $\frac{1}{2} n m(n+1)$. Onder de hypothese $\mathcal{H}_0$ is het theoretische gemiddelde van iedere kolom: $\frac{1}{2} m(n+1)$

We beschouwen nu de afwijkingen van dit gemiddelde. In ons voorbeeld is het theoretisch kolomgemiddelde gelijk aan 14. De afwijkingen daarvan zijn

$$1 \quad -3 \quad -4 \quad 5 \quad -2 \quad 3$$

---

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

De som der kwadraten van deze afwijkingen noemen wij S.
In ons voorbeeld is S = 64.
Als alle m rangschikkingen gelijk zijn wordt het maximum van S bereikt.
Dit maximum is $\frac{1}{12} m^2(n^3-n)$.
We definiëren nu als coëfficiënt van overeenstemming

$$W = \frac{12\ S}{m^2(n^3-n)}$$

In ons voorbeeld is $W = \frac{12 \times 64}{16 \times 210} = 0{,}229$.

W varieert dus tussen 0 en 1.

De verdeling van S onder de hypothese H_o is exact berekend voor een aantal waarden van n en m [1] , terwijl voor grote m en n benaderingen bekend zijn.

De meeste gebruikelijke benaderingen zijn de volgende.

1°. De χ^2-benadering:

$\chi_r^2 = m(n-1)\underline{W} = \frac{12\ \underline{S}}{mn(n+1)}$ heeft voor $m \to \infty$ een χ^2-verdeling met n-1 vrijheidsgraden ([1] pg. 84 [2] pg. 36-37).

2°. De z-benadering:

$\underline{V} = (m-1)\frac{\underline{W}}{1-\underline{W}}$ is bij benadering verdeeld als $\underline{F} = e^{2\underline{z}}$
(F is de F van Snedecor, z de z van Fisher) met
$\nu_1 = n-1-\frac{2}{m}$
$\nu_2 = (m-1)\nu_1$ vrijheidsgraden ([1] pg. 84 [2] pg.33-36)

Met behulp van de verdelingen van S of W onder de hypothese H_o, kan deze hypothese getoetst worden, waarbij H_o verworpen wordt als W waarden dichtbij 1 (resp. S dichtbij $\frac{1}{12} m^2(n^3-n)$) aanneemt, de kritieke zône is dus van de vorm $W \geqq W_o$ (resp. $S \geqq S_o$).

Het kan voorkomen dat de waarnemers geen onderscheid ontdekken in de mate waarin verschillende elementen het kenmerk bezitten. Ze geven deze elementen dan gelijke rangnummers.

Veronderstel, dat door een waarnemer geen onderscheid wordt gemaakt tussen de elementen, die de rangnummers 3 t/m 6 moeten dragen. Dan wordt als rangnummer van ieder van deze elementen het gemiddelde van de rangnummers $\frac{1}{4}(3 + 4 + 5 + 6) = 4\frac{1}{2}$ gebruikt.

Daar het maximum van S nu verandert, moeten wij een correctie op de formule voor W toepassen. Deze vindt men in [1] (pg.82) en [2] (pg. 28-30). Eveneens veranderen dan de formules voor de χ^2-benadering ([1] pg. 86, [2] pg. 37) en voor de z-benadering ([1] pg. 86 [2] pg. 34), doch deze correcties zijn van weinig betekenis, tenzij het aantal gelijken groot is.

Literatuur: [1] M.G.Kendall, Rank correlation methods, London 1948, Hoofdstuk 6, pag. 80.
Tabel van de verdelingsfunctie van S voor:
n = 3 m = 2 t/m 10
n = 4 m = 2 t/m 6
n = 5 m = 3
op pag. 146-149.
Tabel van de waarden van S, waarvan de overschrijdingskansen onder de hypothese H_0 gelijk zijn aan 0,05 of 0,01, berekend met behulp van de z-benadering voor:
n = 3 m = 8,9,10,12,14,15,16,18,20
n = 4 m = 4,5,6,8,10,15,20
n = 5 t/m 7 m = 3,4,5,6,8,10,15,20
op pag. 150.

[2] Ph.van Elteren, Methode der m rangschikkingen, Cursus "Parametervrije Methoden", Hoofdstuk II, Rapport S 59, Mathematisch Centrum (1951), Blz. 18-45.