

Eine Korrelationsschätzung von nicht normalverteilten, alternativen Skalen

**Dipl.-Psych. Martin Junghöfer
Pionierstr. 2
53175 Bonn**

erarbeitet und verfaßt in Bad Neuenahr 1982/83
redaktionell ein wenig überarbeitet in Bonn 2005

Die Faktorenanalyse wird zunehmend häufiger angewandt sowohl bei der Konstruktion von Tests als auch bei der Erforschung interpretierbarer Zusammenhänge von Variablen, die mit Hilfe verschiedenster Methoden erhoben werden.

Im Rahmen dieser Testkonstruktion und Variablenenerhebung werden häufig dichotomierte Antworten bzw. Variablen benutzt (z.B. + -, stimmt – stimmt nicht, trifft zu – trifft nicht zu, richtig – falsch, usw.). Sind diese dichotomierten Variablen normalverteilt, wird der „tetrachorische Korrelationskoeffizient“ r_{tet} benutzt, um die Interkorrelationsmatrix zu erstellen. Kann die Normalverteilung nicht vorausgesetzt werden, wird allgemein der „Punkt-Vierfelder-Korrelationskoeffizient“, auch „Vierfelder-Koeffizient“ *pfrit* oder im weiteren Verlauf Φ (Phi) genannt, benutzt. 1)

Ausgang für die Berechnung von Φ ist ein Vierfelderschema:

		Y		
		-	+	
X	-	a	b	a + b
	+	c	d	c + d
		a + c	b + d	N = a + b + c + d

in dem a, b, c und d Häufigkeiten von aufgetretenen Koinzidenzen bedeuten.

Die Berechnung von Φ erfolgt nach der Formel:

$$\Phi = \frac{bc - ad}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}$$

Im folgenden wird nachgewiesen, daß diese Formel für Φ identisch ist mit der Formel für die Berechnung des „Produkt-Moment-Korrelationskoeffizienten“:

$$r = \frac{N * \Sigma XY - \Sigma X * \Sigma Y}{\sqrt{(N * \Sigma X^2 - (\Sigma X)^2) * (N \Sigma Y^2 - (\Sigma Y)^2)}}$$

Man setze beim Vierfelderschema statt + und – irgendwelche Zahlen (die Höhe der Zahlen ist ohne Effekt auf das Ergebnis), z.B. statt - : 1 und statt + : 2. Dann ergibt sich folgendes Vierfelderschema der Ausprägungen der Variablen X und Y:

Variable:	X	Y	Häufigkeit :
Ausprägung:	1	1	a
	2	1	b
	1	2	c
	2	2	d

Wenn man jetzt diese so umgeformten Daten in die Formel der „Produkt-Moment-Korrelation“ r einsetzt, erhalten wir:

$$\begin{aligned}
 N &= a + b + c + d \\
 \Sigma X &= a + c + 2(b + d) \\
 \Sigma Y &= a + b + 2(c + d) \\
 \Sigma XY &= a + 2(b + c) + 4d \\
 \Sigma X^2 &= a + c + 4(b + d) \\
 \Sigma Y^2 &= a + b + 4(c + d) \\
 (\Sigma X)^2 &= a^2 + c^2 + 4b^2 + 4d^2 + 2ac + 4(ab + ad + bc + cd) + 8bd \\
 (\Sigma Y)^2 &= a^2 + b^2 + 4c^2 + 4d^2 + 2ab + 4(ac + ad + bc + bd) + 8cd .
 \end{aligned}$$

Durch Kürzung erhalten wir die Formel zur Berechnung von Φ . Zu demselben Ergebnis kommt man genauso, wenn irgendwelche anderen Zahlen eingesetzt werden.

Daraus ergibt sich die Folgerung, daß Φ nach der Methode von r berechnet wird.

Die „Produkt-Moment-Korrelation“ r kann auch definiert werden als:

der Kosinus des Winkels, den die beiden jeweils auf X und Y minimalisierten, linearen Regressionsgraden miteinander bilden.

Ist die Übereinstimmung vollkommen, dann fallen die beiden Regressionsgraden aufeinander, dann ist der Winkel, den sie miteinander bilden $= 0^\circ$. $\cos 0^\circ = +1$. Ist überhaupt keine Übereinstimmung gegeben, dann bilden die beiden Regressionsgraden einen Winkel von 90° , dessen $\cos = 0$ ist. Ist die Übereinstimmung der Daten vollkommen, aber gegenläufig, fallen beide Regressionsgraden wieder aufeinander, bilden aber einen Winkel von 180° , dessen $\cos = -1$ ist.

Die Produkt-Moment-Korrelation ist also eine Sinus-Funktion der Daten X und Y.

Da Φ nach der Formel für r berechnet wird, ist also auch Φ eine Sinus-Funktion der Daten X und Y.

Diese Beziehung muß sowohl beim Vergleich verschiedener r -Werte, als auch verschiedener Φ -Werte berücksichtigt werden.

Der numerische Wert, den Φ erreichen kann, ist bestimmten Begrenzungen unterworfen. Er ist abhängig von der Ausprägung und dem Verhältnis der Randsummen sowie der Ausprägung und dem Verhältnis der Diagonal-Werte zueinander:

1. Φ kann die Extremwerte von $+1$ oder -1 nur dann erreichen, wenn zwei Diagonalfelder gleich sind, also:
 $a = d$ oder $b = c$.
 Dann ist das mögliche Maximum oder das mögliche Minimum von $\Phi = +1$ oder -1 .
2. Φ kann die Extremwerte von ± 1 nur dann erreichen, wenn die Diagonalfelder jeweils gleich sind, also:
 $a = d$ und $b = c$.

In allen anderen Fällen kann Φ den Wert von +1 oder -1 nicht erreichen UND die möglichen Extremwerte von Φ (d.h. Φ_{\max} und Φ_{\min}) bei randsummenäquivalenten Vierfelderschemata sind numerisch ungleich.

Zusammengefaßt heißt das: Φ kann bei randsummenäquivalenten Vierfelderschemata bestimmte Maxima und Minima erreichen, diese Extremwerte sind numerisch meist ungleich:

$$|\Phi_{\max}| \neq |\Phi_{\min}|,$$

erreichen ebenso meistens nicht den Wert $|1|$, mit Ausnahme der Fälle, in denen

1. $a = d$ und $b = c$ ist; dann sind beide möglichen Extremwerte = $|1|$,
2. $a = d$ oder $b = c$ ist; dann ist nur einer der beiden möglichen Extremwerte = $|1|$. 2)

Die Extremwerte kann man nach folgendem Verfahren errechnen, nachdem Φ berechnet worden ist:

$$\Phi_{\min} = \Phi - \frac{\text{Min}(b,c) * N}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

$$\Phi_{\max} = \Phi_{\min} + \frac{\text{Min}(\text{Randsummen}) * N}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

Da die möglichen Φ -Werte durch die Ausprägung der Randsummen in ihren möglichen Extremen begrenzt sind und die Größe dieser Extremwerte von den Randsummen abhängt, sind nur Φ -Werte von randsummenäquivalenten Vierfelderschemata miteinander vergleichbar!

Das bedeutet: numerisch gleiche Φ -Werte von Vierfelderschemata mit unterschiedlichen Randsummen sagen nicht aus, daß die Stärke des Zusammenhangs der jeweiligen Variablen gleich ist.

Das heißt auch, daß bei einer Faktorenanalyse über alternative Daten zur Erstellung der Interkorrelationsmatrix nicht die Produkt-Moment-Korrelation, die numerisch gleich dem Φ -Koeffizienten ist, benutzt werden darf, da nicht von durchgängig randsummenäquivalenten Vierfelderschemata ausgegangen werden kann. 3)

Um die Begrenzung der Vergleichbarkeit so weit als möglich aufzuheben, versucht man, Φ so zu korrigieren und relativieren, daß die möglichen Extreme die Werte ± 1 annehmen. In diesem Zusammenhang bekannt ist der „interspezifische Assoziationskoeffizient“ Φ_{korr} von Cole ⁴⁾. Diese Korrektur von Φ hat zum Ziel, aus Φ einen realistischen Schätzwert von r zu bilden.

Im folgenden will ich nachweisen, daß Φ_{korr} ein schlechter Schätzwert von r ist.

Φ_{korr} wird auf folgende Weise errechnet:

$$1. \quad \text{wenn } \Phi > 0, \text{ dann } \Phi_{\text{korr}} = \frac{\Phi}{\Phi_{\text{max}}}$$

$$2. \quad \text{wenn } \Phi < 0, \text{ dann } \Phi_{\text{korr}} = \frac{\Phi}{|\Phi_{\text{min}}|} .$$

Es wird also ein errechnetes Φ durch den Absolutwert des vorzeichenzugehörigen Extremwertes dividiert. Man relativiert also Φ mit dem vorzeichengleichen Extremwert.

Hier sind meine beiden Kritikpunkte verankert:

1. Es werden zwei Φ -Werte dividiert und
2. es wird nur jeweils ein Extremwert zur Relativierung herangezogen.

Ad 1: Da Φ eine Sinusfunktion von X und Y ist, dürfen zwei Φ -Werte nicht multipliziert, dividiert, addiert und subtrahiert werden. Diese Operationen dürfen nur mit den entsprechenden Winkeln (deren Kosinus = Φ ist) durchgeführt werden.

Insofern ist die Errechnung von Φ_{korr} eine irreguläre Operation.

Ad 2: **Da, wie oben festgestellt, die Extremwerte eines Vierfelderschemas nur im beschriebenen Ausnahmefall numerisch gleich sind, ist es unzulässig, für die Berechnung eines zu seinen Extremwerten relativierten, korrigierten Φ nur einen Extremwert zu benutzen, und den anderen, der numerisch meist ungleich ist, außer Acht zu lassen.**

Die Berechnung von Φ_{korr} führt also nicht zu Korrelations- oder Assoziationskoeffizienten, die einen Vergleich ermöglichen zwischen nicht-randsammenäquivalenten Vierfelderschemata.

Ja, wegen der oben angeführten Fehler ist m.E. Φ_{korr} ein aussageloser Wert.

Im folgenden stelle ich ein Verfahren vor, um aus Φ einen Schätzwert von r zu errechnen unter Auslassung der oben genannten Fehler.

Dieser Wert soll also den Forderungen genügen, daß

1. von einer Sinus-Beziehung der Φ -Werte ausgegangen wird und nicht von einer linearen Beziehung, und daß
2. beide Extremwerte berücksichtigt werden.

Ich nenne diesen Wert: r geschätzt von Φ , kurz: **$r_{\text{gesch } \Phi}$** .

Zur Erläuterung des Rationale von $r_{\text{gesch } \Phi}$ sei folgendes Beispiel aufgeführt:

a	b	19
c	d	31
14	36	50

Gegeben sei ein Vierfelderschema mit den beschriebenen Randsummen. Bei jedem Vierfelderschema gibt es: Minimum(Randsummen) + 1 unterschiedliche Möglichkeiten der Felderbesetzung, in diesem Beispiel: $14 + 1 = 15$, also auch 15 verschiedene Φ -Werte bei gleichen Randsummen incl. der beiden Extremwerte Φ_{max} und Φ_{min} . Diese 15 Φ -Werte sind in Tabelle 1 angegeben. Da Φ wie r der Kosinus des Winkels der beiden linearen Regressionen ist, ist in Tabelle 1 dieser Winkel φ° angegeben, zusätzlich noch – wegen einer Rechenvereinfachung – der Winkel $\alpha^\circ = 90^\circ - \varphi^\circ$.

Diese Winkel φ° bzw. α° sind in Abbildung 1 im Einheitskreis mit ihrer Nummerierung eingezeichnet.

	lfd. Nr.	Φ	φ°	α°	ρ°	$r_{\text{gesch } \Phi}$	Φ_{korr}
Φ_{max}	1	0,49	60,78	29,22	0,00	1,00	1,00
	2	0,40	66,65	23,35	12,89	0,97	0,80
	3	0,30	72,26	17,74	25,20	0,90	0,61
	4	0,21	77,71	12,29	37,16	0,80	0,42
	5	0,12	83,04	6,96	48,85	0,66	0,24
	6	0,03	88,32	1,68	60,44	0,49	0,06
	7	-0,06	93,58	-3,58	71,98	0,31	-0,07
	8	-0,15	98,87	-8,87	83,59	0,11	-0,17
	9	-0,25	104,24	-14,24	95,37	-0,09	-0,27
	10	-0,34	109,74	-19,74	107,44	-0,30	-0,37
	11	-0,43	115,44	-25,44	119,95	-0,50	-0,48
	12	-0,53	121,41	-31,41	133,05	-0,68	-0,59
	13	-0,61	127,81	-37,81	147,09	-0,84	-0,72
	14	-0,70	134,81	-44,81	162,45	-0,95	-0,85
Φ_{min}	15	-0,80	142,81	-52,81	180,00	-1,00	-1,00

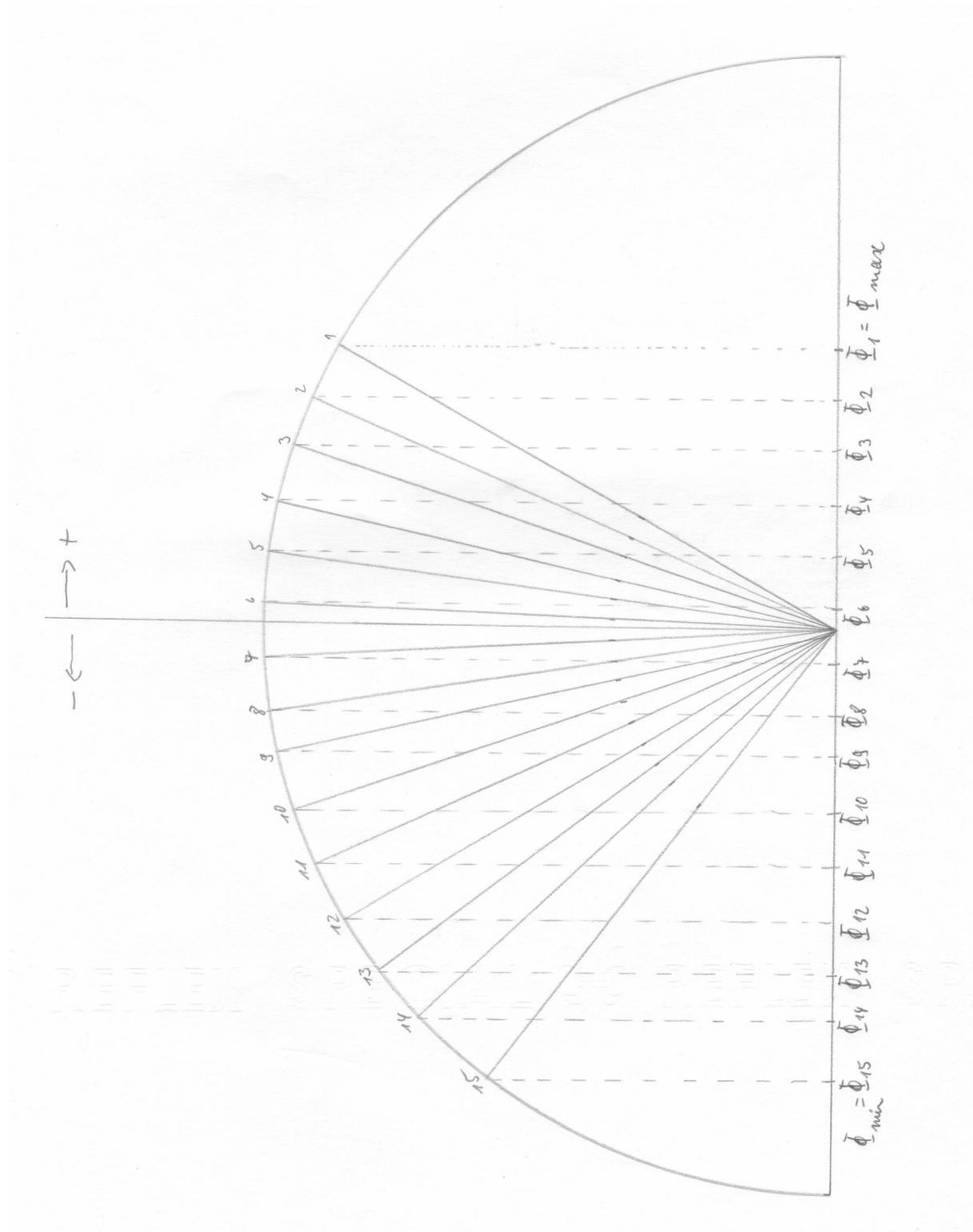
Tabelle 1: Auflistung aller möglichen Φ -Werte eines Vierfelderschemas mit gleichbleibenden Randsummen.

$$\Phi = \cos \varphi^\circ; \quad \alpha^\circ = 90^\circ - \varphi^\circ; \quad r_{\text{gesch } \Phi} = \cos \rho^\circ;$$

Φ_{korr} = der interspezifische Assoziationskoeffizient von Cole.

Abbildung 1: Alle 15 möglichen Φ des Vierfelderschemas:

a	b	19
c	d	31
14	36	50



Das Rationale, das der Berechnung von $r_{\text{gesch } \Phi}$ zugrunde liegt ist:

1. die Extremwerte werden ± 1 , d.h.

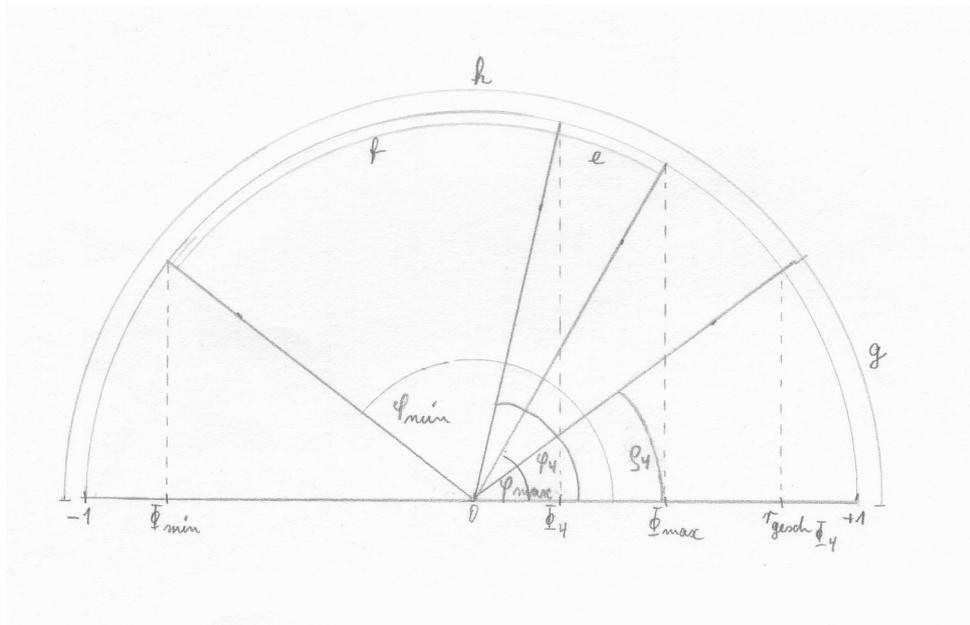
$$\Phi_{\text{max}} \rightarrow r_{\text{gesch } \Phi_{\text{max}}} = +1$$

$$\Phi_{\text{min}} \rightarrow r_{\text{gesch } \Phi_{\text{min}}} = -1.$$

2. Das Verhältnis der Kreisbogenabschnitte von Φ_{max} und Φ_i zu Φ_i zu Φ_{min} ist gleich dem Verhältnis der Kreisbogenabschnitte von $r_{\text{gesch } \Phi_{\text{max}}}$ und $r_{\text{gesch } \Phi_i}$ zu $r_{\text{gesch } \Phi_{\text{min}}}$.

Dies ist beispielhaft in Abbildung 2 an Φ_4 aus dem o.g. Beispiel aufgetragen.

Abbildung 2: Beispiel (Φ_4 aus Tab. 1) der gleichen Verhältnisse von Kreisbogenabschnitten bei Φ und $r_{\text{gesch } \Phi}$. Erläuterung im Text.



In dieser Abbildung bedeuten:

- die Strecke **e** ist der Kreisbogenabschnitt, den der Winkel $(\varphi_4 - \varphi_{\text{max}})$ einschließt;
- die Strecke **f** ist der Kreisbogenabschnitt, den der Winkel $(\varphi_{\text{min}} - \varphi_4)$ einschließt;
- Φ_{max} , Φ_4 und Φ_{min} sind als Kosinus auf der Abszisse aufgetragen;
- die Strecke **g** ist der Kreisbogenabschnitt, den $(\rho_4 - \rho_{\text{max}})$ einschließt, wobei $r_{\text{gesch } \Phi_{\text{max}}} = \cos \rho_{\text{max}} = 1$ ist und ρ_{max} entsprechend 0° ist;
- die Strecke **h** ist der Kreisbogenabschnitt, den der Winkel $(\rho_{\text{min}} - \rho_4)$ einschließt, wobei $r_{\text{gesch } \Phi_{\text{min}}} = \cos \rho_{\text{min}} = -1$ und $\rho_{\text{min}} = 180^\circ$ ist.

(Die Zusätze max und min zu ρ beziehen sich hier auf den zugehörigen Kosinus, so daß das Paradoxon entsteht, daß $\rho_{\max} = 0^\circ$ und $\rho_{\min} = 180^\circ$ ist.)

Die Forderung nach der Gleichheit der Verhältnisse der Bogenabschnitte läßt sich also so ausdrücken:

$$\frac{\mathbf{e}}{\mathbf{f}} = \frac{\mathbf{g}}{\mathbf{h}} .$$

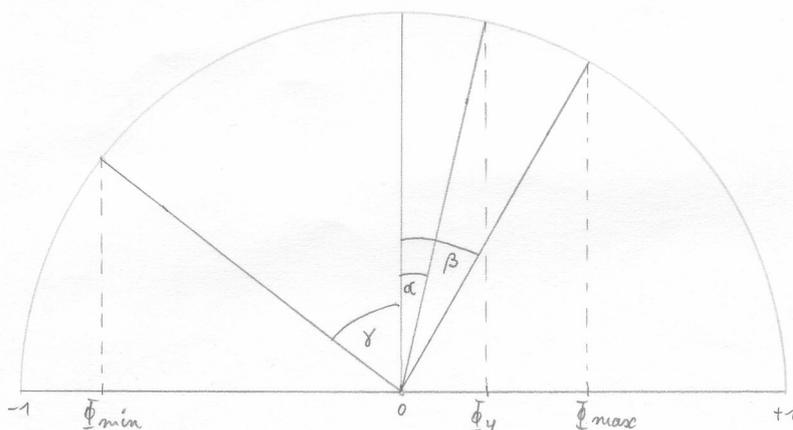
Der Kosinus des so errechneten Winkels ρ ist $r_{\text{gesch } \Phi}$. Dieser aus Φ errechnete Schätzwert von r ist in Verhältnis gesetzt zu beiden Extremwerten von Φ , und er hat er hat eigene Extremwerte von ± 1 , erfüllt also beide o.g. Forderungen.

Im Verlauf der Ableitung und Vereinfachung des Rechengangs für $r_{\text{gesch } \Phi}$ hat es sich als günstiger erwiesen, die jeweiligen Winkel nicht von der Abszisse aus zu betrachten, sondern von der Halbierungsgraden, bei der $r = 0$ ist, also:

$$\begin{aligned} \Phi &= \cos \varphi = \sin \alpha, & \alpha &= 90 - \varphi \\ \Phi_{\max} &= \cos \varphi_{\max} = \sin \beta, & \beta &= 90 - \varphi_{\max} \\ \Phi_{\min} &= \cos \varphi_{\min} = \sin \gamma, & \gamma &= 90 - \varphi_{\min} . \end{aligned}$$

Dies ist in Abbildung 3 veranschaulicht.

Abbildung 3: Beispiel (Φ_4 aus Tabelle 1) für die Lage der Winkel α , β und γ .



Der Winkel α hat das selbe Vorzeichen wie Φ , ebenso der Winkel β . Da der Winkel γ im folgenden nur als Summand ($\beta + |\gamma|$) auftaucht, dessen Winkelsumme den Kreisbogenabschnitt zwischen Φ_{\max} und Φ_{\min} einschließt, wird γ nur als $|\gamma|$ benutzt.

Wir berechnen jetzt die Strecken e und f:

$$\begin{aligned} e &= \text{arc}(\beta - \alpha) \\ f &= \text{arc}(\gamma + \alpha) \\ e + f &= \text{arc}(\beta + \gamma). \end{aligned}$$

Die Strecke $g + h$ ist gleich $\text{arc } 180^\circ$, also gleich π :

$$\begin{aligned} g + h &= \pi \\ h &= \pi - g. \end{aligned}$$

Setzt man die so gefundene Gleichung in die Forderung:

$$\frac{e}{f} = \frac{g}{h} \quad \text{ein, so ergibt sich:}$$

$$\frac{e}{f} = \frac{g}{\pi - g}$$

$$e(\pi - g) = fg$$

$$e\pi - eg = fg$$

$$e\pi = g(e + f)$$

$$g = \frac{e\pi}{e + f}$$

$$g = \frac{\text{arc}(\beta - \alpha)\pi}{\text{arc}(\beta + \gamma)}$$

$$g = \frac{(\beta - \alpha) * \text{arc } 1^\circ * \pi}{(\beta + \gamma) * \text{arc } 1^\circ}$$

$$g = \frac{(\beta - \alpha)\pi}{(\beta + \gamma)}$$

Die Strecke g ist der Kreisbogenabschnitt des Winkels ρ , dessen Kosinus $r_{\text{gesch } \Phi}$ ist. Also:

$$g = \text{arc } \rho$$

$$\rho = \frac{g}{\text{arc } 1^\circ}$$

$$\rho = \frac{(\beta - \alpha) * \pi}{(\beta + \gamma) * \text{arc } 1^\circ}$$

$$\frac{\pi}{\text{arc } 1^\circ} = 180$$

$$\rho = \frac{(\beta - \alpha) * 180}{(\beta + \gamma)}$$

Also ist:

$$r_{\text{gesch } \Phi} = \cos \rho$$

$$r_{\text{gesch } \Phi} = \cos \left[\frac{\beta - \alpha}{\beta + |\gamma|} * 180 \right]$$

Abbildung 4: Alle 15 möglichen, randsummenäquivalenten $r_{\text{gesch } \Phi}$ des o.g. Beispiel-Vierfelderschemas:

a	b	19
c	d	31
14	36	50

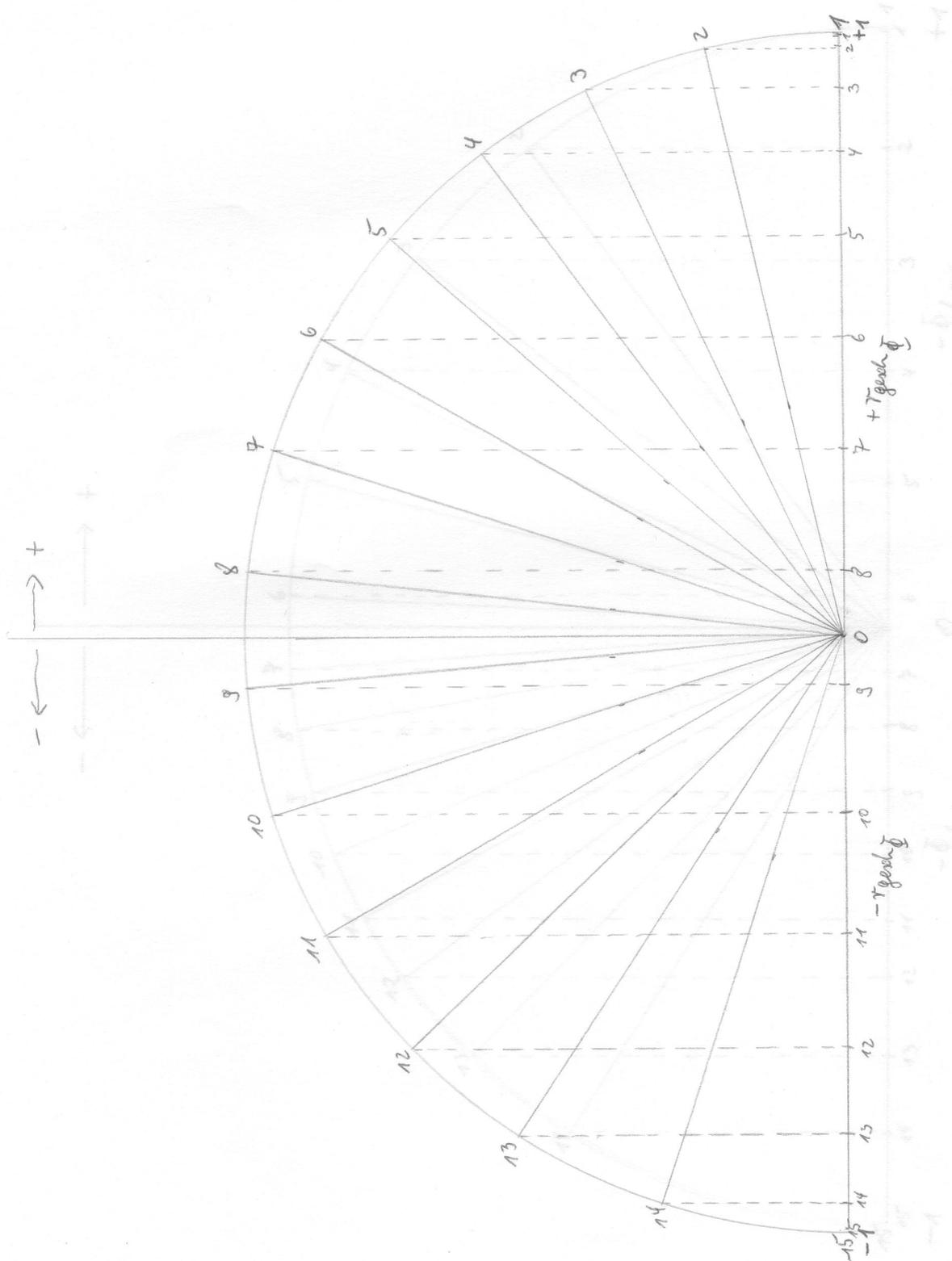
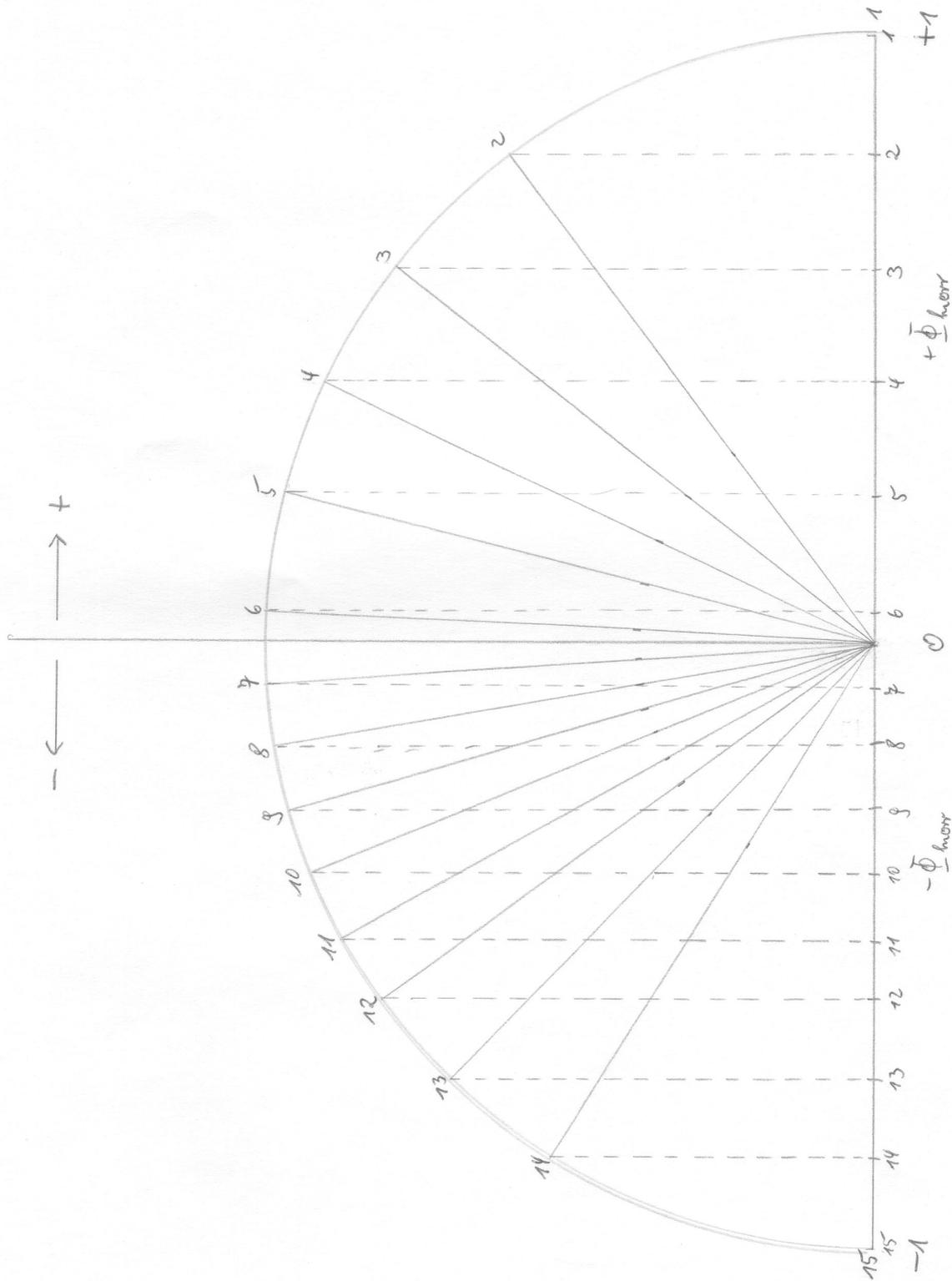


Abbildung 5: Alle 15 möglichen, randsummenäquivalenten Φ_{korrr} des o.g. Beispiel-Vierfelderschemas:

a	b	19
c	d	31
14	36	50



In Abbildung 4 sind die von Φ nach $r_{\text{gesch } \Phi}$ transformierten Werte aus dem Beispiel in Tabelle 1 eingezeichnet. Zum Vergleich dazu die entsprechenden Φ_{korr} -Werte nach Cole in Abbildung 5.

Im Vorausgehenden habe ich versucht nachzuweisen, daß Φ und Φ_{korr} schlechte Schätzwerte von r sind unter dem Gesichtspunkt der Vergleichbarkeit, die ja notwendige Voraussetzung z.B. bei der Faktorenanalyse ist.

Dagegen habe ich versucht, mit $r_{\text{gesch } \Phi}$ einen aus Φ berechneten Schätzwert von r zu entwickeln, der durch Ausschaltung von zwei formalen Fehlern in der Berechnung von Φ_{korr} , die Vergleichbarkeit bei nicht randsummenäquivalenten Vierfelderschemata in höherem Maße herstellen soll.

Um ein Beispiel anzubringen, wie sich die Unterschiede von Φ , Φ_{korr} und $r_{\text{gesch } \Phi}$ formal, ohne Berücksichtigung der Interpretation, in einer Faktorenanalyse auswirken, habe ich aus einem Test, bei dessen Entwicklung dieses Problem auftauchte, über die ersten vier Fragen in einer Faktorenanalyse nach der Hauptkomponentenmethode die ersten zwei Faktoren herausgezogen und zwar jeweils für Φ , Φ_{korr} und $r_{\text{gesch } \Phi}$. Die Korrelationsmatrizen, die Ladungen auf den beiden Faktoren und die Werte a^2 / h^2 dieser drei Faktorenanalysen sind in Tabelle 2 zu ersehen.

Methode:	Korrelationsmatrix:				Faktorenmatrix:		a^2 / h^2 :			
Φ	1	2	3	4	A1	A2	A1	A2		
1	(1,0)	-0,25	-0,23	-0,01	1	0,67	-0,15	1	0,96	0,05
2	-0,25	(1,0)	-0,24	-0,24	2	-0,63	-0,68	2	0,46	0,54
3	-0,23	-0,24	(1,0)	-0,20	3	-0,36	0,87	3	0,15	0,88
4	-0,01	-0,24	-0,20	(1,0)	4	0,61	-0,03	4	1,01	0,00
Φ_{korr}	1	2	3	4	A1	A2	A1	A2		
1	(1,0)	-0,31	-0,27	-0,03	1	0,53	-0,01	1	1,00	0,00
2	-0,31	(1,0)	-0,26	-0,44	2	-0,64	-0,74	2	0,43	0,57
3	-0,27	-0,26	(1,0)	-0,41	3	-0,51	0,84	3	0,27	0,73
4	-0,03	-0,44	-0,41	(1,0)	4	0,80	-0,05	4	1,00	0,00
$r_{\text{gesch } \Phi}$	1	2	3	4	A1	A2	A1	A2		
1	(1,0)	-0,09	-0,06	0,33	1	0,55	0,73	1	0,36	0,63
2	-0,09	(1,0)	0,21	-0,25	2	-0,60	0,42	2	0,67	0,33
3	-0,06	0,21	(1,0)	-0,27	3	-0,60	0,50	3	0,59	0,41
4	0,33	-0,25	-0,27	(1,0)	4	0,78	0,19	4	0,95	0,06

Tabelle 2: Beispiel für unterschiedliche Ergebnisse bei einer Faktorenanalyse nach der Hauptkomponentenmethode, je nach Erstellung der Korrelationsmatrix nach den Methoden für Φ , Φ_{korr} und $r_{\text{gesch } \Phi}$.

Auffallend ist, daß Φ und Φ_{KORR} nahezu identische Faktoren liefern und daß der Faktor 1 bei $r_{\text{gesch } \Phi}$ noch einige Ähnlichkeiten aufweist mit den Faktoren 1 von Φ und Φ_{KORR} , daß aber beim zweiten Faktor sowohl in der Höhe der relevanten Ladungen als auch in den Vorzeichen so gut wie keine Ähnlichkeit mehr aufzuweisen ist.

Das bedeutet, daß spätestens bei der Extraktion des zweiten Faktors mit gänzlich anderen Ergebnissen zu rechnen ist, wenn die Korrelationsmatrix statt nach den Methoden von Φ und Φ_{KORR} mit der Methode von $r_{\text{gesch } \Phi}$ erstellt wird.

Der Frage, ob dieses Verfahren zu inhaltlich-psychologisch besseren Interpretationen führt, konnte im Rahmen dieser Untersuchung nicht nachgegangen werden. Rein formal gesehen scheint das aber zu erwarten zu sein, da ja – wie oben nachgewiesen - Φ und Φ_{KORR} bei nichttrandsummenäquivalenten Vierfelderschemata nicht miteinander vergleichbare Korrelations- bzw. Assoziationskoeffizienten sind.

Literatur und Bemerkungen:

- 1) LIENERT, G.A.: Testaufbau und Testanalyse, Weinheim 1969,
hier S. 136
- 2) CLAUSS, G. & EBNER, H.: Grundlagen der Statistik, Ffm o.J.
hier Seite 254: hier irren Claus & Ebner in der
Behauptung, daß Φ max und Φ min die Werte ± 1 nur
dann erreichen können, wenn $a = d = 0$ oder $b = c = 0$
sind.
- 3) FAHRENBERG, J., SELG, H. & HAMPEL, R.: Das Freiburger
Persönlichkeitsinventar FPI, Handanweisung,
Göttingen 1973
hier Seite 17
- 4) zitiert nach CLAUSS & EBNER, s. 2)
hier S. 256

zusätzlich:

- SIEGEL, S.: Nonparametric Statistics for the behavioral Sciences,
New York 1956
- LEWIS, V.E.: Statistical Analysis, Princeton 1963
- FAHRENBERG, J. et al.: Psychophysiologische Aktivierungsforschung, München 1979

Anhang 1: Berechnung von $r_{\text{gesch } \Phi}$ mit einem Taschenrechner:

Mit einem Taschenrechner mit mindestens einen Speicher und Sinus-, Kosinus- und deren Umkehrfunktion ist die Berechnung von $r_{\text{gesch } \Phi}$ recht einfach und sei hier beschrieben.

Zur Erläuterung nehme ich wieder das oben angeführte Beispiel Φ_4 :

Gegeben sei folgendes Vierfelderschema:

3	16	19
11	20	31
14	36	50

Jetzt berechnet man zuerst eine Konstante (k):

$$k = \frac{1}{\sqrt{\text{Randsummenprodukt}}}$$

Im Beispiel:

$$k = \frac{1}{\sqrt{19 * 31 * 36 * 14}} = \frac{1}{\sqrt{296.856}} = 0,0018353846$$

und setze dieses Ergebnis als Konstante (k) in den Speicher.

Jetzt wird Φ berechnet:

$$\Phi = (bc - ad) * k$$

Im Beispiel:

$$\Phi = (16 * 11 - 3 * 20) * k = 116 * k = 0,2129$$

Jetzt errechne man den Winkel α :

$$\Phi = \sin \alpha$$

Im Beispiel:

$$\begin{aligned} \Phi &= \sin 12,2926^\circ \\ \alpha &= 12,2926^\circ \end{aligned}$$

Durch Drücken der Sinus-Taste bekommt man wieder Φ und kann nun Φ_{\min} berechnen:

$$\Phi_{\min} = \Phi - \text{Min} (b, c) * N * k$$

Im Beispiel:

$$\Phi_{\min} = 0,2129 - 11 * 50 * k = -0,7966.$$

Durch Wechsel des Vorzeichens und Betätigen der inversen Sinusfunktion erhalten wir den Winkel $|\gamma|$:

$$|\Phi_{\min}| = \sin |\gamma|$$

Im Beispiel:

$$0,7966 = \sin 52,8026^\circ$$

$$|\gamma| = 52,8026^\circ.$$

Man holt Φ_{\min} zurück und errechnet Φ_{\max} :

$$\Phi_{\max} = \Phi_{\min} + \text{Min} (\text{Randsummen}) * N * k$$

Im Beispiel:

$$\Phi_{\max} = -0,7966 + 14 * 50 * k = 0,4882$$

Nun wird der Winkel β aus Φ_{\max} errechnet:

$$\Phi_{\max} = \sin \beta$$

Im Beispiel:

$$0,4882 = \sin 29,2231^\circ$$

$$\beta = 29,2231^\circ$$

Jetzt berechnet man den Winkel ρ :

$$\rho = \frac{\beta - \alpha}{\beta + |\gamma|} * 180$$

Im Beispiel:

$$\rho = \frac{29,2231 - 12,2926}{29,2231 + 52,8026} * 180 = 37,1529^\circ$$

Durch die Kosinusfunktion erhalten wir $r_{\text{gesch } \Phi}$:

$$r_{\text{gesch } \Phi} = \cos \rho$$

Im Beispiel:

$$r_{\text{gesch } \Phi} = \cos 37,1529^\circ = 0,7970 .$$

Anhang 2: Berechnung von $r_{\text{gesch } \Phi}$ mit Excel:

Gegeben sei das Vierfelderschema:

a	b	a+b
c	d	c + d
a+c	b+d	a+b+c+d

$r_{\text{gesch } \Phi} =$

COS(BOGENMASS((ARCSIN((b*c-a*d)*1/WURZEL((a+b)*(c+d)*(a+d)*(b+d))-
MIN(b;c)*(a+b+c+d)*1/WURZEL((a+b)*(c+d)*(a+c)*(b+d))+MIN((a+b);(c+d);(a+c);(b+d))
*(a+b+c+d)*1/WURZEL((a+b)*(a+c)*(c+d)*(b+d)))-ARCSIN((b*c-a*d)*1/WURZEL
((a+b)*(c+d)*(a+c)*(b+d))))*180/(ARCSIN((b*c-a*d)*1/WURZEL((a+b)*(c+d)*(a+c)*
(b+d))-MIN(b;c)*(a+b+c+d)*1/WURZEL((a+b)*(c+d)*(a+c)*(b+d)))+
MIN((a+b);(c+d);(a+c);(b+d))*(a+b+c+d)*1/WURZEL((a+b)*(c+d)*(a+c)*(b+d)))+ARCSIN
(ABS(((b*c-a*d)*1/WURZEL((a+b)*(c+d)*(a+c)*(b+d)))-
MIN(b;c)*(a+b+c+d)*1/WURZEL((a+b)*(c+d)*(a+c)*(b+d))))))

Auf meinem Rechner setze ich:

a in die Zelle B5,

b in C5,

c in B6 und

d in die Zelle C6, daraus ergibt sich folgende Formel:

COS(BOGENMASS((ARCSIN((C5*B6-
B5*C6)*1/WURZEL((B5+C5)*(B6+C6)*(B5+B6)*(C5+C6))-
MIN(C5;B6)*(B5+C5+B6+C6)*1/WURZEL((B5+C5)*(B6+C6)*(B5+B6)*(C5+C6))+MIN
((B5+C5);(B6+C6);(B5+B6);(C5+C6))*(B5+C5+B6+C6)*1/WURZEL((B5+C5)*(B6+C6
)*(B5+B6)*(C5+C6)))-ARCSIN((C5*B6-
B5*C6)*1/WURZEL((B5+C5)*(B6+C6)*(B5+B6)*(C5+C6))))*180/(ARCSIN((C5*B6-
B5*C6)*1/WURZEL((B5+C5)*(B6+C6)*(B5+B6)*(C5+C6))-
MIN(C5;B6)*(B5+C5+B6+C6)*1/WURZEL((B5+C5)*(B6+C6)*(B5+B6)*(C5+C6))+MIN
((B5+C5);(B6+C6);(B5+B6);(C5+C6))*(B5+C5+B6+C6)*1/WURZEL((B5+C5)*(B6+C6
)*(B5+B6)*(C5+C6)))+ARCSIN(ABS(((C5*B6-
B5*C6)*1/WURZEL((B5+C5)*(B6+C6)*(B5+B6)*(C5+C6))-
MIN(C5;B6)*(B5+C5+B6+C6)*1/WURZEL((B5+C5)*(B6+C6)*(B5+B6)*(C5+C6))))))