

Diese Formelsammlung wurde für Studierende der Psychologie entwickelt, ist aber vermutlich auch für Studierende anderer Sozialwissenschaften hilfreich. Die Formelsammlung basiert auf einer alten Sammlung der Professur für Forschungsmethoden und Evaluation in der Psychologie an der TU Chemnitz, die weder Autoren noch Lizenz enthielt. Für eine sinnvolle Nutzung und Weiterentwicklung ist eine Lizenz notwendig, weshalb wir das vorliegende Dokument unter CC BY-SA 4.0 veröffentlichen. Die Formelsammlung orientiert sich am Lehrbuch von Sedlmeier und Renkewitz (2018). Die Autoren der Formelsammlung sind: Nils Heimhuber, Feline Baumgärtel und Johannes Titz. Weitere Mitwirkende: Vivien Lungwitz, Annika Sternkopf.

Einige Hinweise vorab: Lateinische Buchstaben (\bar{x} , s^2 , s) werden für die Stichprobe verwendet, griechische Buchstaben (μ , σ^2 , σ) für die Population. Ein Dach über einer Statistik ($\hat{\sigma}$) steht für eine Schätzung des Parameters. Abkürzungen:

MWU: Mittelwertsunterschied US: Unabhängige Stichproben AS: Abhängige Stichproben

Lagemaße

Modalwert

Der Modalwert ist der häufigste Wert aller Messwerte. Es kann mehrere Modal-Werte geben.

Mittelwert

Auch bezeichnet als arithmetisches Mittel, für den Mittelwert der Population wird der Buchstabe μ verwendet.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Median

Der Median ist der Wert in der Mitte aller in einer Rangreihe geordneten Messwerte.

$$\text{Tiefe}_{\text{Median}} = \frac{n+1}{2}$$

Quartile

Die Werte müssen in eine Rangreihe gebracht werden. Für das untere Quartil wird die Tiefe von unten gezählt, fürs obere Quartil von oben. Quartile werden manchmal auch als die Quantile Q_{25} , Q_{75} bezeichnet.

$$\text{Tiefe}_{\text{Quartil}} = \frac{\lfloor \text{Tiefe}_{\text{Median}} \rfloor + 1}{2}$$

Zäune bei Boxplots

Die Zäune werden benötigt um die Whiskers zu bestimmen. Die Whiskers sind durch tatsächlich vorkommende Werte repräsentiert. Von den Zäunen aus geht man in Richtung Box, bis man den ersten vorkommenden Wert findet.

$$\text{Zaun}_{\text{unten}} = Q_{25} - 1,5 \cdot \text{IQA}$$

$$\text{Zaun}_{\text{oben}} = Q_{75} + 1,5 \cdot \text{IQA}$$

Streuungsmaße

Spannweite (Range)

$$R = x_{\max} - x_{\min}$$

Varianz

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Standardabweichung

$$s = \sqrt{s^2}$$

Interquartilsabstand

$$\text{IQA} = Q_{75} - Q_{25}$$

Schätzung für Populationsvarianz

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{n}{n-1} \cdot s^2$$

Transformation

z-Standardisierung

$$z_i = \frac{x_i - \bar{x}}{s}$$

$$\bar{z} = 0$$

$$s_z = 1$$

Zentrierung

$$c_i = x_i - \bar{x}$$

IQ-Standardisierung

$$\text{IQ}_i = z_i \cdot 15 + 100$$

Standardfehler für Stichprobenverteilungen

Mittelwerte

$$\hat{\sigma}_{\bar{x}} = \sqrt{\hat{\sigma}_{\bar{x}}^2} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{s}{\sqrt{n-1}}$$

MWU für US

wenn $n_A = n_B$:

$$\hat{\sigma}_{\bar{x}_A - \bar{x}_B} = \sqrt{\hat{\sigma}_{\bar{x}_A}^2 + \hat{\sigma}_{\bar{x}_B}^2}$$

wenn $n_A \neq n_B$

$$\hat{\sigma}_{\bar{x}_A - \bar{x}_B} = \sqrt{\frac{(n_A - 1) \cdot \hat{\sigma}_A^2 + (n_B - 1) \cdot \hat{\sigma}_B^2}{(n_A - 1) + (n_B - 1)} \left(\frac{1}{n_A} + \frac{1}{n_B} \right)}$$

MWU für AS

$$\hat{\sigma}_{\text{diff}} = \frac{\hat{\sigma}_{\text{diff}}}{\sqrt{n}}$$

Binomialverteilung

$$\sigma_{\text{Person}} = \sqrt{np(1-p)}$$

$$\sigma_{\text{Anteil}} = \sqrt{\frac{p(1-p)}{n}}$$

Zusammenhangsmaße

Kovarianz

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Korrelation

Pearson-Korrelationskoeffizient
(Produkt-Moment-Korrelation)

$$r = \frac{\text{cov}(x, y)}{s_x s_y} = \frac{1}{n} \sum_{i=1}^n z_{x_i} \cdot z_{y_i}$$

Partialkorrelation

$$r_{xy.z} = \frac{r_{xy} - r_{xz} \cdot r_{yz}}{\sqrt{1 - r_{xz}^2} \cdot \sqrt{1 - r_{yz}^2}}$$

Phi-Koeffizient

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

	nein	ja
nein	a	b
ja	c	d

Einfache lineare Regression

Regressionsgleichung

$$\hat{y}_i = a + b \cdot x_i$$

$$b = \frac{\text{cov}(x, y)}{s_x^2} = r \frac{s_y}{s_x}$$

$$a = \bar{y} - b\bar{x}$$

$$y_i = \hat{y}_i + e_i$$

Determinationskoeffizient

$$R^2 = \frac{s_y^2}{s_y^2}$$

Standardschätzfehler

$$s_e = s_y \cdot \sqrt{1 - r^2}$$

Multiple Regression

Regressionsgleichung

für Originalwerte:

$$\hat{y}_i = b_0 + b_1 \cdot x_{1i} + b_2 \cdot x_{2i}$$

wobei:

$$b_1 = \beta_1 \cdot \frac{s_y}{s_{x_1}}$$

$$b_2 = \beta_2 \cdot \frac{s_y}{s_{x_2}}$$

$$b_0 = \bar{y} - b_1 \cdot \bar{x}_1 - b_2 \cdot \bar{x}_2$$

für z-standardisierte Werte:

$$\hat{z}_{y_i} = \beta_1 \cdot z_{x_{1i}} + \beta_2 \cdot z_{x_{2i}}$$

wobei:

$$\beta_1 = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}$$

$$\beta_2 = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}$$

Gütemaße

Multipler Determinationskoeffizient

$$R^2 = \beta_1 \cdot r_{yx_1} + \beta_2 \cdot r_{yx_2}$$

Multipler Korrelationskoeffizient

$$R = \sqrt{R^2}$$

Standardschätzfehler
für Originalwerte

$$s_e = s_y \sqrt{1 - R^2}$$

für z-standardisierte Werte

$$s_e = \sqrt{1 - R^2}$$

Konfidenzintervalle

Binomialverteilung

nur nutzbar, wenn $\sigma_{\text{Person}}^2 = n \cdot p \cdot (1 - p) > 9$

für Anteile

$$p \pm \sigma_{\text{Anteil}} \cdot z_{\text{Konfidenz}}$$

für Personen/Objekte

$$p \cdot n \pm \sigma_{\text{Person}} \cdot z_{\text{Konfidenz}}$$

Mittelwert

$$\bar{x} \pm \hat{\sigma}_{\bar{x}} \cdot t_{\text{df, Konfidenz}}$$

MWU, US

$$(\bar{x}_A - \bar{x}_B) \pm \hat{\sigma}_{\bar{x}_A - \bar{x}_B} \cdot t_{\text{df, Konfidenz}}$$

MWU, AS

$$\overline{\text{diff}} \pm \hat{\sigma}_{\overline{\text{diff}}} \cdot t_{\text{df, Konfidenz}}$$

t-Test

Mittelwert gegen Konstante

wobei c (constant) vorgegeben

$$t = \frac{\bar{x} - c}{\hat{\sigma}_{\bar{x}}}$$

$$\text{df} = n - 1$$

MWU, US

$$t = \frac{\bar{x}_A - \bar{x}_B}{\hat{\sigma}_{\bar{x}_A - \bar{x}_B}}$$

$$\text{df} = (n_A - 1) + (n_B - 1)$$

MWU, AS

$$t_{\text{AS}} = \frac{\overline{\text{diff}}}{\hat{\sigma}_{\overline{\text{diff}}}}$$

$$\text{df}_{\text{AS}} = n - 1$$

F-Wert

$$F(df_{zw}, df_{inn}) = \frac{\hat{\sigma}_{zw}^2}{\hat{\sigma}_{inn}^2}$$

Populationsvarianzen

$$\hat{\sigma}_{zw}^2 = \frac{QS_{zw}}{df_{zw}}$$

$$\hat{\sigma}_{inn}^2 = \frac{QS_{inn}}{df_{inn}}$$

Quadratsummenzerlegung

$$QS_{ges} = QS_{zw} + QS_{inn}$$

Quadratsummen

$$QS_{zw} = \sum_j^k n_j \cdot (\bar{x}_j - \bar{\bar{x}})^2$$

$$QS_{inn} = \sum_j^k \sum_i^{n_j} (x_{ij} - \bar{x}_j)^2$$

Freiheitsgrade

$$df_{zw} = k - 1$$

$$df_{inn} = \sum_j^k (n_j - 1) = N - k$$

Notation

k : Anzahl der Bedingungen

n_j : Anzahl der Messwerte pro Bedingung j

N : Gesamtzahl der Messwerte

F-Wert

$$F = \frac{\hat{\sigma}_{UV}^2}{\hat{\sigma}_{res}^2}$$

Populationsvarianzen

$$\hat{\sigma}_{UV}^2 = \frac{QS_{UV}}{df_{UV}}$$

$$\hat{\sigma}_{res}^2 = \frac{QS_{res}}{df_{res}}$$

Quadratsummen

$$QS_{ges} = \sum_j^k \sum_i^n (x_{ji} - \bar{\bar{x}})^2$$

$$QS_{UV} = \sum_j^k n \cdot (\bar{x}_j - \bar{\bar{x}})^2$$

$$QS_{Person} = \sum_i^n k \cdot (\bar{x}_i - \bar{\bar{x}})^2$$

$$QS_{res} = QS_{ges} - QS_{UV} - QS_{Person}$$

Quadratsummenzerlegung

$$QS_{ges} = QS_{UV} + QS_{Person} + QS_{res}$$

Freiheitsgrade

$$df_{gesamt} = N - 1$$

$$df_{UV} = k - 1$$

$$df_{Person} = n - 1$$

$$df_{res} = (k - 1)(n - 1)$$

Notation

k : Anzahl der Bedingungen

n : Anzahl Messwerte pro Bedingung

N : Gesamtzahl der Messwerte

$\bar{\bar{x}}$: Gesamtmittelwert

Zweifaktorielle Varianzanalyse

Hinweis: gilt nur bei balancierten Designs (gleiche Stichprobengröße)

F-Werte

$$F_A = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_{\text{inn}}^2}$$

$$F_B = \frac{\hat{\sigma}_B^2}{\hat{\sigma}_{\text{inn}}^2}$$

$$F_{A \times B} = \frac{\hat{\sigma}_{A \times B}^2}{\hat{\sigma}_{\text{inn}}^2}$$

Populationsvarianzen

$$\hat{\sigma}_A^2 = \frac{QS_A}{df_A}$$

$$\hat{\sigma}_B^2 = \frac{QS_B}{df_B}$$

$$\hat{\sigma}_{A \times B}^2 = \frac{QS_{A \times B}}{df_{A \times B}}$$

$$\hat{\sigma}_{\text{inn}}^2 = \frac{QS_{\text{inn}}}{df_{\text{inn}}}$$

Freiheitsgrade

$$df_{\text{ges}} = N - 1$$

$$df_{\text{inn}} = k \cdot m(n - 1)$$

$$df_{\text{inn}} = \sum_j^k \sum_l^m (n_{jl} - 1) = N - k$$

$$df_A = k - 1$$

$$df_B = m - 1$$

$$df_{A \times B} = (k - 1)(m - 1)$$

Quadratsummen

$$QS_{\text{ges}} = QS_A + QS_B + QS_{A \times B} + QS_{\text{inn}}$$

$$QS_{\text{ges}} = \sum_j^k \sum_l^m \sum_i^n (x_{ijl} - \bar{\bar{x}})^2$$

$$QS_{\text{inn}} = \sum_j^k \sum_l^m \sum_i^n (x_{ijl} - \bar{x}_{jl})^2$$

$$QS_A = n \cdot m \sum_j^k (\bar{x}_{j.} - \bar{\bar{x}})^2$$

$$QS_B = n \cdot k \sum_l^m (\bar{x}_{.l} - \bar{\bar{x}})^2$$

$$QS_{A \times B} = QS_{\text{ges}} - QS_A - QS_B - QS_{\text{inn}}$$

Notation

N : Gesamtzahl der Messwerte der Untersuchung
 n : Anzahl der Messwerte pro Bedingung
 k : Anzahl der Stufen des Faktors A
 m : Anzahl der Stufen des Faktors B
 $\bar{\bar{x}}$: Gesamtmittelwert
 x_j : Mittelwert für Faktor A
 x_i : Mittelwert für Faktor B

Kontrastanalyse US

Quadratsummen

$$F_{\text{Kontrast}} = \frac{\hat{\sigma}_{\text{Kontrast}}^2}{\hat{\sigma}_{\text{inn}}^2}$$

oder

$$F = t^2$$

$$|t_{\text{Kontrast}}| = \sqrt{F_{\text{Kontrast}}}$$

$$QS_{\text{Kontrast}} = \frac{\left(\sum_{i=1}^k \lambda_i \bar{x}_i \right)^2}{\sum_{i=1}^k \frac{\lambda_i^2}{n_i}}$$

Populationsvarianz

$$\hat{\sigma}_{\text{Kontrast}}^2 = \frac{QS_{\text{Kontrast}}}{df_{\text{Kontrast}}}$$

$$\hat{\sigma}_{\text{inn}}^2 = \frac{QS_{\text{inn}}}{df_{\text{inn}}}$$

oder

$$\hat{\sigma}_{\text{inn}}^2 = \frac{\sum_{j=1}^k \hat{\sigma}_j^2}{k}$$

Freiheitsgrade

$$df_{\text{Kontrast}} = 1$$

$$df_{\text{inn}} = N - k$$

Notation

k : Gruppe
 n_i : Anzahl der Messwerte pro Bedingung i
 n_j : Anzahl der Messwerte pro Bedingung j

(ohne Subgruppen)

$$t_{\text{Kontrast}} = \frac{\bar{L}}{\sqrt{\frac{\hat{\sigma}_L^2}{n}}}$$

$$L_i = \sum_{j=1}^k (x_{ij} \cdot \lambda_j)$$

$$\hat{\sigma}_L^2 = \frac{1}{n-1} \sum_{i=1}^n (L_i - \bar{L})^2$$

$$\text{df}_{\text{Kontrast}} = n - 1$$

Notation \bar{L} : Mittelwert L -Werte $\hat{\sigma}_i^2$: geschätzte Populationsvarianz der Gruppe i n : Anzahl der Objekte/Personen,
für die mehrere Messungen durchgeführt

Nonparametrische Verfahren

 χ^2 -Anpassungstests

für eine Variable

$$\chi^2 = \sum_i \frac{(f_{b,i} - f_{e,i})^2}{f_{e,i}}$$

$$f_{e,i} = N \cdot P_i$$

$$\text{df} = k - 1$$

für zwei Variablen

$$\chi^2 = \sum_i \sum_j \frac{(f_{b,ij} - f_{e,ij})^2}{f_{e,ij}}$$

$$f_{e,ij} = N \cdot P_{ij}$$

$$\text{df} = k \cdot m - 1$$

 χ^2 -Unabhängigkeitstest

$$\chi^2 = \sum_i \sum_j \frac{(f_{b,ij} - f_{e,ij})^2}{f_{e,ij}}$$

$$f_{e,ij} = \frac{Z_i \cdot S_j}{N}$$

$$\text{df} = (k - 1) \cdot (m - 1)$$

Bei 2 dichotomen Variablen: $\chi^2 = \phi^2 \cdot N$ **U-Test nach Mann und Whitney (Wilcoxon Rangsummen-Test)**

1. Messwerte insgesamt in Rangreihe bringen und jedem Messwert einen Rangplatz zuweisen
2. T_1 : Summe der Rangplätze der Gruppe 1 T_2 : Summe der Rangplätze der Gruppe 2
3. $U = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 - 1)}{2} - T_1$
 $U' = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - T_2 = n_1 \cdot n_2 - U$
4. Prüfgröße für die bei uns verwendete Tabelle ist der kleinere der beiden U-Werte

Wilcoxon-Test für AS (Vorzeichenrangtest)

1. Differenzen der Messwertpaare bilden
2. den Beträgen der Differenzen Rangplatz zuweisen (beim kleinsten Betrag mit Rangplatz 1 beginnen)
3. T_- : Summe der Rangplätze der negativen Differenzen T_+ : Summe der Rangplätze der positiven Differenzen
4. Prüfgröße für die Tabelle ist der kleinere der beiden T-Werte

Notation $f_{b,i}$: beobachtete Häufigkeit $f_{e,i}$: erwartete Häufigkeit P_i : in Merkmalsausprägung erwarteter Anteil N : Anzahl der Untersuchungsteilnehmer $k ; m$: Anzahl der Merkmalsausprägungen beider Merkmale Z_i : Zeilenhäufigkeit S_j : Spaltenhäufigkeit ϕ : Phi-Koeffizient

Faktorenanalyse

1. Faktorladung: $a_{mk} = r(m, k)$

2. Eigenwert: $\lambda_k = \sum_{m=1}^M a_{mk}^2$

3. Kommunalität: $h_m^2 = \sum_{k=1}^f a_{mk}^2$

4. durch Faktor aufgeklärter Varianzanteil: $\frac{\lambda_k}{M}$

Notation

m : Variable,

k : Anzahl der Faktoren,

i : Person/Objekt,

M : Anzahl der Variablen = Gesamtvarianz

r : Korrelation

f : Anzahl der ausgewählten Faktoren

Clusteranalyse

Nominalskalierte Variablen

	Fall x	
	+	-
Fall y	+	-
	a	c
	-	b
	d	

Tanimoto-Koeffizient

$$T = \frac{a}{a + b + c}$$

M-Koeffizient

$$M = \frac{a + d}{a + b + c + d}$$

Intervallskalierte Variablen

a, b : Fälle

J : Anzahl der Variablen (Dimensionen)

r : Minkowski-Konstante

Minkowski-Metrik

$$d_{a,b} = \left[\sum_{j=1}^J |x_{aj} - x_{bj}|^r \right]^{\frac{1}{r}}$$

Euklidische Distanz

$$d_{a,b} = \sqrt{\sum_{j=1}^J |X_{aj} - X_{bj}|^2}$$

Manhattan-Distanz / City-Block-Metrik

$$d_{a,b} = \sum_{j=1}^J |X_{aj} - X_{bj}|$$

Konventionen nach Cohen

"These values are necessarily somewhat arbitrary, but were chosen so as to seem reasonable. The reader can render his own judgment as to their reasonableness." (Cohen, 1962)

	d/g	r/w/ ϕ	η^2
klein	± 0.2	± 0.1	0.01
mittel	± 0.5	± 0.3	0.06
groß	± 0.8	± 0.5	0.14

t-Test: Einstichprobenfall

$$g = \frac{t}{\sqrt{n}}$$

$$d = \frac{t}{\sqrt{df}}$$

t-Test: US

$$g = t_{US} \cdot \sqrt{\frac{n_A + n_B}{n_A \cdot n_B}}$$

$$d = t_{US} \cdot \frac{n_A + n_B}{\sqrt{df} \cdot \sqrt{n_A \cdot n_B}}$$

$$r = \sqrt{\frac{t_{US}^2}{t_{US}^2 + df}}$$

Letzteres gilt auch für F , wenn df_{zw} 1 sind, da in diesem Fall $F = t^2$.

Vereinfachung, wenn $n_A = n_B$:

$$d = \frac{2 \cdot t_{US}}{\sqrt{df}}$$

$$g = \frac{2 \cdot t_{US}}{\sqrt{n}}$$

t-Test: AS

$$g = \frac{t_{AS}}{\sqrt{n}}$$

$$d = \frac{t_{AS}}{\sqrt{df}}$$

Kontrastanalyse US

aus F-Test

$$r_{\text{effectsize}} = \sqrt{\frac{F_{\text{Kontrast}}}{F_{zw} \cdot df_{zw} + df_{inn}}}$$

$$r_{\text{contrast}} = \sqrt{\frac{F_{\text{Kontrast}}}{F_{\text{Kontrast}} + df_{inn}}} = \sqrt{\frac{t_{\text{Kontrast}}^2}{t_{\text{Kontrast}}^2 + df}}$$

Kontrastanalyse AS

aus t-Test

$$g = \frac{t}{\sqrt{n}}$$

 χ^2 -Tests

$$w = \sqrt{\sum_{i=1}^k \frac{(P_{b,i} - P_{e,i})^2}{P_{e,i}}}$$

aus Ergebnis χ^2 -Test:

$$w = \sqrt{\frac{\chi^2}{N}}$$

aus 2 dichotomen Variablen

$$w = \sqrt{\frac{\chi^2}{N}} = \phi$$

ANOVA

US

$$\eta^2 = \frac{F \cdot df_{zw}}{F \cdot df_{zw} + df_{inn}}$$

$$\eta^2 = \frac{QS_{zw}}{QS_{ges}}$$

AS

$$\eta^2 = \frac{QS_{UV}}{QS_{ges}}$$

$$\eta_p^2 = \frac{QS_{UV}}{QS_{UV} + QS_{res}}$$

mehrfaktoriell

$$\eta^2 = \frac{QS_{\text{Effekt}}}{QS_{ges}}$$

$$\eta_p^2 = \frac{QS_{\text{Effekt}}}{QS_{\text{Effekt}} + QS_{inn}}$$

Effektgrößen aus Rohwerten

$$d = \frac{\bar{x} - c}{\sigma_x}$$

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(\sigma_1^2 + \sigma_2^2)}{2}}}$$

$$g = \frac{\bar{x} - c}{s_x}$$

$$g = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

$$s = \sqrt{\frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}}$$

$$g = \frac{\bar{L}}{\hat{\sigma}_L}$$

$$r_{\text{effectsize}} = \frac{\frac{1}{n} \sum_i^n (x_i - \bar{x}) \cdot (\lambda_i - \bar{\lambda})}{s_x \cdot s_\lambda}$$

$$r_{\text{contrast}} = \frac{QS_{\text{Kontrast}}}{QS_{\text{Kontrast}} + QS_{\text{Nicht-Kontrast}}}$$

$$r_{\text{alerting}}^2 = \frac{QS_{\text{Kontrast}}}{QS_{\text{zw}}}$$

Odds Ratio (OR)

$$OR = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a \cdot d}{b \cdot c}$$

	Risikofaktor	ohne Risikofaktor
Krankheit	a	b
Keine Krankheit	c	d

Effektgrößen aus Effektgrößen

Abstandsmaße aus Abstandsmaßen

$$d = g \cdot \sqrt{\frac{n}{df}}$$

$$g = d \cdot \sqrt{\frac{df}{n}}$$

Korrelationen aus Abstandsmaßen

$$r = \frac{d}{\sqrt{d^2 + \frac{1}{p \cdot q}}}$$

$$r = \sqrt{\frac{g^2(n_A \cdot n_B)}{g^2(n_A \cdot n_B) + (n_A + n_B)df}}$$

Abstandsmaße aus Korrelationen

$$d = \frac{r}{\sqrt{1 - r^2}} \cdot \sqrt{\frac{1}{p \cdot q}}$$

$$g = \frac{r}{\sqrt{1 - r^2}} \cdot \sqrt{\frac{(n_A + n_B) \cdot df}{n_A \cdot n_B}}$$

wobei

$$p = \frac{n_A}{n_A + n_B}$$

und

$$q = \frac{n_B}{n_A + n_B}$$

Meta-Analyse

Mittlere, an der Stichprobe gewichtete, Effektgröße: $\bar{r} = \frac{\sum_i^n (N_i r_i)}{\sum_i^n N_i}$

Bedingte Wahrscheinlichkeit

Die bedingte Wahrscheinlichkeit von A unter der Bedingung B wird notiert als $P(A|B)$ und berechnet sich durch:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Hierbei ist $P(A \cap B)$ die Wahrscheinlichkeit, dass sowohl A als auch B eintreten.

Produktregel

Die Produktregel beschreibt die Wahrscheinlichkeit des gemeinsamen Eintretens von zwei Ereignissen A und B :

$$P(A \cap B) = P(A|B) \cdot P(B)$$

Alternativ kann die Produktregel auch in umgekehrter Reihenfolge geschrieben werden:

$$P(A \cap B) = P(B|A) \cdot P(A)$$

Summenregel

Die Summenregel gibt die Wahrscheinlichkeit des Eintretens mindestens eines von zwei disjunkten Ereignissen an:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Wahrscheinlichkeitsrevision (Bayes)

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(\neg A) \cdot P(B|\neg A)}$$

Literatur

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *The Journal of Abnormal and Social Psychology*, 65(3), 146. <https://doi.org/10.1037/h0045186>

Sedlmeier, P., & Renkewitz, F. (2018). *Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler*: Pearson