Formelsammlung Methodenlehre

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 enhielt. 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 ($\overline{x}, s^2, s$) 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.

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

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

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

Standardabweichung

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

Interquartilsabstand

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

Schätzung für Populationsvarianz

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

Transformation

z-Standardisierung

$$z_i = \frac{x_i - \overline{x}}{s}$$
$$\overline{z} = 0$$
$$s_z = 1$$

Zentrierung

$c_i = x_i - \overline{x}$

IQ-Standardisierung

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

Standardfehler für Stichprobenverteilungen

Mittelwerte

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

MWU für US

wenn $n_A = n_B$:

$$\hat{\sigma}_{\overline{x}_\mathrm{A}-\overline{x}_\mathrm{B}} = \sqrt{\hat{\sigma}^2_{\overline{x}_\mathrm{A}} + \hat{\sigma}^2_{\overline{x}_\mathrm{B}}}$$

wenn $n_A \neq n_B$

$$\hat{\sigma}_{\overline{x}_\mathrm{A}-\overline{x}_\mathrm{B}} = \sqrt{\frac{(n_A-1)\cdot\hat{\sigma}^2_\mathrm{A} + (n_\mathrm{B}-1)\cdot\hat{\sigma}^2_\mathrm{B}}{(n_\mathrm{A}-1)+(n_\mathrm{B}-1)}(\frac{1}{n_\mathrm{A}} + \frac{1}{n_\mathrm{B}})}$$

MWU für AS

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

Binomialverteilung

$$\sigma = \sqrt{np(1-p)}$$

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

Zusammenhangsmaße

Kovarianz

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

Korrelation

Pearson-Korrelationskoeffizient
(Produkt-Moment-Korrelation)

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

Partialkorrelation

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

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

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

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

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

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

Determinationskoeffizient

$$R^2 = \frac{s^2_{\hat{y}}}{s^2_y}$$

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:

$$b_1 = \beta_1 \cdot \frac{s_y}{s_{x_1}}$$
$$b_2 = \beta_2 \cdot \frac{s_y}{s_{x_2}}$$
$$b_0 = \overline{y} - b_1 \cdot \overline{x}_1 - b_2 \cdot \overline{x}_2$$

für z-standardisierte Werte:

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

wobei:

$$\beta_1 = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{1 - r^2_{x_1x_2}}$$
$$\beta_2 = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{1 - r^2_{x_1x_2}}$$

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 $\sigma^2_{\text{Anteil}} = n \cdot p \cdot (1-p) > 9$

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

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

MWU, US

$$(\overline{x}_{\text{A}} - \overline{x}_{\text{B}}) \pm \hat{\sigma}_{\overline{x}_{\text{A}} - \overline{x}_{\text{B}}} \cdot t_{\text{df,Konfidenz}}$$

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

$$t = \frac{\overline{x} - c}{\hat{\sigma}_{\overline{x}}}$$
$$\text{df} = n - 1$$

MWU, US

$$t = \frac{\overline{x}_A - \overline{x}_B}{\hat{\sigma}_{\overline{x}_A - \overline{x}_B}}$$
$$\text{df} = (n_A - 1) + (n_B - 1)$$

MWU, AS

$$t_{\text{AS}} = \frac{\overline{x}_{\text{diff}}}{\hat{\sigma}_{\overline{x}_{\text{diff}}}}$$
$$\text{df}_{\text{AS}} = n - 1$$

ANOVA US

F-Wert

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

Populationsvarianzen

$$\hat{\sigma}^2_{\mathrm{zw}} = \frac{\mathrm{QS}_{\mathrm{zw}}}{\mathrm{df}_{\mathrm{zw}}}$$

$$\hat{\sigma}^2_{\mathrm{inn}} = \frac{\mathrm{QS}_{\mathrm{inn}}}{\mathrm{df}_{\mathrm{inn}}}$$

Quadratsummenzerlegung

$$\mathrm{QS}_{\mathrm{ges}} = \mathrm{QS}_{\mathrm{zw}} + \mathrm{QS}_{\mathrm{inn}}$$

Quadratsummen

$$\mathrm{QS}_{\mathrm{zw}} = \sum_j^k n_j \cdot (\overline{x}_j - \overline{\overline{x}})^2$$

$$\mathrm{QS}_{\mathrm{inn}} = \sum_j^k \sum_i^{n_j} (x_{ij} - \overline{x}_j)^2$$

Freiheitsgrade

$$\mathrm{df}_{\mathrm{zw}} = k - 1$$

$$\mathrm{df}_{\mathrm{inn}} = \sum_j^k (n_j - 1) = N - k$$

Notation

- k: Anzahl der Bedingungen
- n_j: Anzahl der Messwerte pro Bedingung j
- N: Gesamtzahl der Messwerte

ANOVA AS

F-Wert

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

Populationsvarianzen

$$\hat{\sigma}^2_{\mathrm{UV}} = \frac{\mathrm{QS}_{\mathrm{UV}}}{\mathrm{df}_{\mathrm{UV}}}$$

$$\hat{\sigma}^2_{\mathrm{res}} = \frac{\mathrm{QS}_{\mathrm{res}}}{\mathrm{df}_{\mathrm{res}}}$$

Quadratsummen

$$\mathrm{QS}_{\mathrm{ges}} = \sum_j^k \sum_i^n (x_{ji} - \overline{\overline{x}})^2$$

$$\mathrm{QS}_{\mathrm{UV}} = \sum_j^k n \cdot (\overline{x}_j - \overline{\overline{x}})^2$$

$$\mathrm{QS}_{\mathrm{Person}} = \sum_i^n k \cdot (\overline{x}_i - \overline{\overline{x}})^2$$

$$\mathrm{QS}_{\mathrm{res}} = \mathrm{QS}_{\mathrm{ges}} - \mathrm{QS}_{\mathrm{UV}} - \mathrm{QS}_{\mathrm{Person}}$$

Quadratsummenzerlegung

$$\mathrm{QS}_{\mathrm{ges}} = \mathrm{QS}_{\mathrm{UV}} + \mathrm{QS}_{\mathrm{Person}} + \mathrm{QS}_{\mathrm{res}}$$

Freiheitsgrade

$$\mathrm{df}_{\mathrm{gesamt}} = N - 1$$

$$\mathrm{df}_{\mathrm{UV}} = k - 1$$

$$\mathrm{df}_{\mathrm{Person}} = n - 1$$

$$\mathrm{df}_{\mathrm{res}} = (k - 1)(n - 1)$$

Notation

- k: Anzahl der Bedingungen
- n: Anzahl Messwerte pro Bedingung
- N: Gesamtzahl der Messwerte
- $\overline{\overline{x}}$: Gesamtmittelwert

Zweifaktorielle Varianzanalyse

Hinweis: gilt nur bei balancierten Designs (gleiche Stichprobengröße)

F-Werte

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

Populationsvarianzen

$$\hat{\sigma}^2_A = \frac{\text{QS}_A}{\text{df}_A}$$
$$\hat{\sigma}^2_B = \frac{\text{QS}_B}{\text{df}_B}$$
$$\hat{\sigma}_{A\times B} = \frac{\text{QS}_{A\times B}}{\text{df}_{A\times B}}$$
$$\hat{\sigma}^2_{\text{inn}} = \frac{\text{QS}_{\text{inn}}}{\text{df}_{\text{inn}}}$$

Freiheitsgrade

$$\text{df}_{\text{ges}} = N - 1$$
$$\text{df}_{\text{inn}} = k \cdot m(n-1)$$
$$\text{df}_{\text{inn}} = \sum_j^k \sum_l^m (n_{jl} - 1) = N - k$$
$$\text{df}_{\text{A}} = k - 1$$
$$\text{df}_{\text{B}} = m - 1$$
$$\text{df}_{\text{AxB}} = (k-1)(m-1)$$

Quadratsummen

$$\text{QS}_{\text{ges}} = \text{QS}_A + \text{QS}_B + \text{QS}_{A\times B} + \text{QS}_{inn}$$
$$\text{QS}_{\text{ges}} = \sum_j^k \sum_l^m \sum_i^n (x_{ijl} - \overline{\overline{x}})^2$$
$$\text{QS}_{\text{inn}} = \sum_j^k \sum_l^m \sum_i^n (x_{ijl} - \overline{x}_{jl})^2$$
$$\text{QS}_A = n \cdot m \sum_j^k (\overline{x}_{j\cdot} - \overline{\overline{x}})^2$$
$$\text{QS}_B = n \cdot k \sum_l^m (\overline{x}_{\cdot l} - \overline{\overline{x}})^2$$
$$\text{QS}_{\text{AxB}} = \text{QS}_{\text{ges}} - \text{QS}_{\text{A}} - \text{QS}_{\text{B}} - \text{QS}_{\text{inn}}$$

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
- $\overline{\overline{x}}$: Gesamtmittelwert
- x_j: Mittelwert für Faktor A
- x_i: Mittelwert für Faktor B

Kontrastanalyse US

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

oder

$$F = t^2$$
$$|t_{\text{Kontrast}}| = \sqrt{F_{\text{Kontrast}}}$$

Populationsvarianz

$$\hat{\sigma}^2_{\text{Kontrast}} = \frac{\text{QS}_{\text{Kontrast}}}{\text{df}_{\text{Kontrast}}}$$
$$\hat{\sigma}^2_{\text{inn}} = \frac{\text{QS}_{\text{inn}}}{\text{df}_{\text{inn}}}$$
$$\text{oder} \quad \hat{\sigma}^2_{\text{inn}} = \frac{\sum_{j=1}^k \hat{\sigma}^2_j}{k}$$

Quadratsummen

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

Freiheitsgrade

$$\text{df}_{\text{Kontrast}} = 1$$
$$\text{df}_{\text{inn}} = N - k$$

Notation

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

Kontrastanalyse für AS

(ohne Subgruppen)

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

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

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

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

Notation

$\overline{L}$: 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

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

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

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

für zwei Variablen

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

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

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

χ^2-Unabhängigkeitstest

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

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

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

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. $U = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 - 1)}{2} - T_1$
   $U' = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - = n_1 \cdot n_2 - U$
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

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

Manhatten-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-Tests

Einstichprobenfall

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

wenn $n_\mathrm{A} = n_\mathrm{B}$:

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

Zwei unabhängige Stichproben

$$g = t_\mathrm{US} \cdot \sqrt{\frac{n_\mathrm{A} + n_\mathrm{B}}{n_\mathrm{A} \cdot n_\mathrm{B}}}$$

wenn $n_\mathrm{A} = n_\mathrm{B}$:

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

$$d = t_\mathrm{US} \cdot \frac{n_\mathrm{A} + n_\mathrm{B}}{\sqrt{\mathrm{df}} \cdot \sqrt{n_\mathrm{A} + n_\mathrm{B}}}$$

$$r = \sqrt{\frac{t^2_\mathrm{US}}{t^2_\mathrm{US} + \mathrm{df}}}$$

Letzteres gilt auch für F, wenn df_zw 1 sind, da in diesem Fall $F = t^2$.

Zwei abhängige Stichproben

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

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

Kontrastanalyse US

aus F-Test

$$r_\mathrm{effectsize} = \sqrt{\frac{F_\mathrm{Kontrast}}{F_\mathrm{zw} \cdot \mathrm{df}_\mathrm{zw} + \mathrm{df}_\mathrm{inn}}}$$

$$r_\mathrm{contrast} = \sqrt{\frac{F_\mathrm{Kontrast}}{F_\mathrm{Kontrast} + \mathrm{df}_\mathrm{inn}}} = \sqrt{\frac{t^2_\mathrm{Kontrast}}{t^2_\mathrm{Kontrast} + \mathrm{df}}}$$

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

$$\eta^2 = \frac{F \cdot \mathrm{df}_\mathrm{zw}}{F \cdot \mathrm{df}_\mathrm{zw} + \mathrm{df}_{inn}}$$

$$\eta^2 = \frac{\mathrm{QS}_\mathrm{zw}}{\mathrm{QS}_\mathrm{ges}}$$

AS

$$\eta^2 = \frac{\mathrm{QS}_\mathrm{UV}}{\mathrm{QS}_\mathrm{ges}}$$

$$\eta^2_p = \frac{\mathrm{QS}_\mathrm{UV}}{\mathrm{QS}_\mathrm{UV} + \mathrm{QS}_\mathrm{res}}$$

mehrfaktoriell

$$\eta^2 = \frac{\mathrm{QS}_\mathrm{Effekt}}{\mathrm{QS}_\mathrm{ges}}$$

$$\eta^2_p = \frac{\mathrm{QS}_\mathrm{Effekt}}{\mathrm{QS}_\mathrm{Effekt} + \mathrm{QS}_\mathrm{inn}}$$

Effektgrößen aus Rohwerten

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

$$d = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{(\sigma_1^2 + \sigma_2^2)}{2}}}$$

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

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

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

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

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

$$r^2_{\text{alerting}} = \frac{\text{QS}_{\text{Kontrast}}}{\text{QS}_{\text{zw}}}$$

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

Odds Ratio (OR)

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

	Risikofaktor	ohne Risikofaktor
Krankheit	a	b
Keine Krankheit	c	d

Effektgrößen aus Effektgrößen

Abstandsmaße aus Abstandsmaßen

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

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

Korrelationen aus Abstandsmaßen

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

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

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

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

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