# Evaluation und Metaanalyse

MSc Klinische Psychologie und Psychotherapie

SoSe 2024

Prof. Dr. Dirk Ostwald

| Datum | Einheit | Thema |
|---|---|---|
| 12.04.2024 | Einführung | (1) Evaluationsforschung |
| 19.04.2024 | Metanalyse | (2) Metaanalysen |
| 26.04.2024 | Metanalayse | (3) Standardisierte Effektgrößenschätzung |
| 03.05.2024 | Metaanalyse | (4) Fixed- und Random-Effects-Modelle |
| 10.05.2024 | Linear Mixed Models | (5) Linear Mixed Models (Online) |
| 17.05.2024 | Linear Mixed Models | (6) Restricted Maximum Likelihood |
| 24.05.2024 | Präsentationen | |
| 03.06.2024 | Präsentationen | |
| 07.06.2024 | Linear Mixed Models | (7) Metaanalyseanwendung |
| 17.06.2024 | Linear Mixed Models | (8) Längsschnittstudienanwendung |
| 24.06.2024 | Präsentationen | |
| 28.06.2024 | Präsentationen | |
| 08.07.2024 | Linear Mixed Models | (9) Clusteranwendung |
| 12.07.2024 | Metaanalyse | (10) Metaanalytische Selektionsmodelle |
| 19.07.2024 | Klausur | |
| Februar 2025 | Klausurwiederholungstermin | |

# (4) Fixed- und Random-Effects Modelle

Motivation

Fixed-Effect-Modell der Metaanalyse

Random-Effects-Modell der Metaanalyse

Selbstkontrollfragen

**Motivation**

Fixed-Effect-Modell der Metaanalyse

Random-Effects-Modell der Metaanalyse

Selbstkontrollfragen

A systematic review and meta-analysis of transdiagnostic cognitive behavioural therapies for emotional disorders

"In unified transdiagnostic treatments, patients with different disorders receive the same 'broadband' treatment that targets shared commonalities between these disorders."

"Across treatment formats, TD-CBT was superior to waitlist and treatment-as-usual. TD-CBT showed comparable effects to disorder-specific CBT and was superior to other active treatments for depression but not for anxiety"

"All analyses were conducted in R (v.4.3.1), using the metafor (v.4.2-0), meta (v.6.5-0) and dmetar packages (v.0.1.0). We calculated controlled effect sizes for the difference between the transdiagnostic treatment and the control conditions in main outcomes (depression and anxiety) at posttreatment (relative efficacy), using the bias-corrected Hedges' g and the 95% CI. These were calculated by subtracting the mean posttreatment score of the transdiagnostic condition from the mean score of the control condition, divided by the pooled standard deviation of both conditions. Values of 0.2, 0.5 and 0.8 of Hedges' g represent a small, moderate and large effect size, respectively."''

"**Building on previous work, we expected considerable variability and thus used a random-effects model to account for heterogeneity of included studies. We tested heterogeneity of effect sizes with the Q statistic, the $I^2$ statistic and by visual inspection of forest plots. A P value of the Q statistic below 0.05 indicates heterogeneity. $I^2$ ranges from 0 to 100%, with 25% representing low, 50% moderate and 75% high heterogeneity.**"

Schaeuffele et al. (2024)

# Motivation

Bei kleinen Einzelstudienstichproben sind empirische Effektstärken nichtzentral-$t$-verteilt

Bei großen Einzelstudienstichproben sind empirische Effektstärken approximativ normalverteilt

*Fixed-Effect-Modell der Metaanalyse*

• Empirische Effektstärken sind um eine feste wahre, aber unbekannten, Effektsstärke normalverteilt

*Random-Effects-Modell der Metaanalyse*

• Empirische Effektstärken sind um zufällige latente studienspezifische Effektsstärken normalverteilt

• Die latenten Effektstärken sind um eine feste wahre, aber unbekannte Effektstärke normalverteilt

Ziel ist immer eine möglichst gute Schätzung der wahren, aber unbekannten, Effektstärke

Beide Modelle sind spezielle Linear Mixed Models

**Motivation**

Fixed-Effect-Modell der Metaanalyse

Random-Effects-Modell der Metaanalyse

Selbstkontrollfragen

# Fixed-Effects-Modell der Metaanalyse

## Definition (Fixed-Effects Modell der Metaanalyse)

Für Studien $i = 1, ..., k$ seien $y_i$ und $\sigma_i^2$ bekannte studienspezifische standardisierte Effektstärkeschätzer und ihre Varianzschätzer, respektive. Dann ist das *Fixed-Effects Modell der Metaanalyse (FEM)* gegeben durch

$$y_i \sim N(\delta, \sigma_i^2), \tag{1}$$

wobei $\delta$ die wahre, aber unbekannte, standardisierte Effektstärke von Interesse bezeichnet.

Bemerkungen

- $y_i$ entspricht meist dem Wert von Hedges' $g_i$.
- Manchmal kann $y_i$ auch dem Wert von Cohen's $d_i$ entsprechen.
- $y_i \sim N(\delta, \sigma_i^2)$ approximiert die nichtzentrale-$t$-Verteilung von $y_i$ bei großen Stichprobenumfängen.
- $\sigma_i^2$ entspricht meist dem Wert eines Schätzers der Varianz von Hedges $g_i$.
- Manchmal kann $\sigma_i^2$ auch dem Wert eines Schätzers der Varianz von Cohen's $d_i$ entsprechen.
- Das Fixed-Effects-Modell der Metaanalyse nimmt explizit an, dass die $\sigma_i^2, i = 1, ..., k$ bekannt sind.
- Es gibt verschiedenste Möglichkeiten, die Varianz von Hedges $g_i$ bzw. Cohen's $d_i$ in Studie $i$ zu schätzen.
- Wir betrachten nur den von Hedges and Olkin (1985) und in Viechtbauer (2010) implementierten Schätzer

$$\sigma_i^2 := \hat{\mathbb{V}}_a(g_i) = \frac{n_t + n_c}{n_t n_c} + \frac{g_i^2}{2(n_t + n_c)}. \tag{2}$$

- Lin and Aloe (2021) diskutieren verschiedene Varianzschätzer für Hedges' $g_i$ bzw. Cohen's $d_i$.
- Ziel der Fixed-Effects-Modellierung der Metaanalyse ist es, $\delta$ basierend auf den $y_i$ und $\sigma_i^2$ zu schätzen.

## Theorem (FEM-Maximum-Likelihood-Effektstärkeschätzers)

Gegeben sei das Fixed-Effect-Modell der Metaanalyse mit wahrem, aber unbekannten, Effekstärkenparameter $\delta$. Dann ergibt sich der Maximum-Likelihood-Schätzer für $\delta$ zu

$$\hat{\delta} := \frac{1}{\sum_{i=1}^k w_i} \sum_{i=1}^k w_i y_i \text{ mit den Wichtungsparametern } w_i := \frac{1}{\sigma_i^2}, i = 1, ..., k. \tag{3}$$

Bemerkungen

- $\hat{\delta}$ genügt der Intuition, dass Studien mit kleiner studienspezifischer Varianz des standardisierten Effektstärkeschätzers, z.B. hervorgerufen durch einen großen studienspezifischen Stichprobenumfang, mehr zu der übergeordneten Effektstärkeschätzung beitragen sollten als solche mit hoher studienspezifischer Varianz des standardisierten Effektstärkeschätzers.
- Die Wichtungsparameter werden als reziproke Werte der studienspezifischen Varianz des standardisierten Effektschärkeschätzers auch *studienspezifische Präzisionen* des standardisierten Effektstärkeschätzers genannt.
- Sind die studienspezifischen Varianzen $\sigma_i^2$ alle identisch zu $\sigma^2$, so ergibt sich mit $w := 1/\sigma^2$

$$\hat{\delta} = \frac{1}{\sum_{i=1}^k w} \sum_{i=1}^k w y_i = \frac{w}{w \sum_{i=1}^k 1} \sum_{i=1}^k y_i = \frac{1}{n} \sum_{i=1}^k y_i. \tag{4}$$

# Fixed-Effect-Modell der Metaanalyse

Beweis

Ohne zwischen $y_i$ als Zufallsvariable oder Wert zu unterscheiden hat die Likelihood-Funktion des FEMs die Form

$$L : \mathbb{R} \to \mathbb{R}_{>0}, \delta \mapsto L(\delta) := \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2\sigma_i^2}(y_i - \delta)^2\right). \tag{5}$$

Die Log-Likelihood-Funktion des FEMs hat damit die Form

$$\ell : \mathbb{R} \to \mathbb{R}, \delta \mapsto \ell(\delta) := -\frac{k}{2}\ln 2\pi - \frac{k}{2}\ln \sigma_i^2 - \sum_{i=1}^{k} \frac{1}{2\sigma_i^2}(y_i - \delta)^2. \tag{6}$$

Die erste Ableitung der Log-Likelihood-Funktion des FEMs hat damit die Form

$$\frac{d}{d\delta}\ell(\delta) = -\sum_{i=1}^{k} \frac{1}{2\sigma_i^2} \frac{d}{d\delta}(y_i - \delta)^2 = \sum_{i=1}^{k} \frac{1}{\sigma_i^2}(y_i - \delta). \tag{7}$$

Es ergibt sich also die Maximum-Likelihood Bedingung

$$\sum_{i=1}^{k} \frac{1}{\sigma_i^2}(y_i - \hat{\delta}) = 0. \tag{8}$$

Auflösen ergibt dann

$$\sum_{i=1}^{k} \frac{1}{\sigma_i^2} y_i - \sum_{i=1}^{k} \frac{1}{\sigma_i^2}\hat{\delta} = 0 \Leftrightarrow \hat{\delta} = \frac{\sum_{i=1}^{k} \frac{1}{\sigma_i^2} y_i}{\sum_{i=1}^{k} \frac{1}{\sigma_i^2}} \Leftrightarrow \hat{\delta} = \frac{1}{\sum_{i=1}^{k} w_i} \sum_{i=1}^{k} w_i y_i. \tag{9}$$

$\square$

## Theorem (Eigenschaften des FEM-ML-Effektstärkeschätzers)

Gegeben sei das Fixed-Effect-Modell der Metaanalyse mit wahrem, aber unbekannten, Effekstärkenparameter $\delta$ und es sei $\hat{\delta}$ der Maximum Likelihood Schätzer von $\delta$. Dann gelten

(1) $\mathbb{E}(\hat{\delta}) = \delta$

(2) $\mathbb{V}(\hat{\delta}) = \frac{\sum_{i=1}^k w_i^2 \sigma_i^2}{(\sum_{i=1}^k w_i)^2}$

Bemerkungen

- Aussage (1) besagt insbesondere, dass $\hat{\delta}$ ein unverzerrter Schätzer von $\delta$ ist.
- Aussage (2) wird z.B. in der Herleitung des DerSimonian-Laird Schätzers benötigt.

Beweis

(1) Mit den Eigenschaften des Erwartungswerts gilt

$$\mathbb{E}(\hat{\delta}) = \mathbb{E}\left(\frac{1}{\sum_{i=1}^k w_i} \sum_{i=1}^k w_i y_i\right) = \frac{1}{\sum_{i=1}^k w_i} \sum_{i=1}^k w_i \mathbb{E}(y_i) = \frac{1}{\sum_{i=1}^k w_i} \delta = \delta \frac{\sum_{i=1}^k w_i}{\sum_{i=1}^k w_i} \sum_{i=1}^k = \delta. \tag{10}$$

(2) Mit den Eigenschaften der Varianz bei unabhängigen Zufallsvariablen gitl

$$\mathbb{V}(\hat{\delta}) = \mathbb{V}\left(\frac{1}{\sum_{i=1}^k w_i} \sum_{i=1}^k w_i y_i\right) = \frac{1}{(\sum_{i=1}^k w_i)^2} \mathbb{V}\left(\sum_{i=1}^k w_i y_i\right) = \frac{\sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{(\sum_{i=1}^k w_i)^2} = \frac{\sum_{i=1}^k w_i^2 \sigma_i^2}{(\sum_{i=1}^k w_i)^2}. \tag{11}$$

# Fixed-Effect-Modell der Metaanalyse

Simulation mit $k := 10$, $\delta := 1$ und $\sigma_i \sim U(1,2)$

```
set.seed(1)                                                   # reproduzierbare Ergebnisse
k                = 10                                         # Anzahl an Studien
delta            = 1                                          # wahre, aber unbekannte, Effekstärke
sigsqr_i         = rep(NaN,k)                                 # wahre, bekannte, studienspezifische Varianzenarray
y_i              = rep(NaN,k)                                 # Empirische Effekstärkenarray
for(i in 1:k){                                                # Studieniterationen
    sigsqr_i[i] = runif(1,1,2)                                # wahre, bekannte, studienspezifische Varianz
    y_i[i]      = rnorm(1, delta, sqrt(sigsqr_i[i]))}         # studienspezifischer Effektstärkeschätzer
w_i              = 1/sigsqr_i                                 # Wichtungsparameter
delta_hat        = sum(w_i*y_i)/sum(w_i)                      # Fixed-Effects-Modell-basierter Effektstärkeschätzer
```

| delta | sigma2_i | y_i | w_i |
|---|---|---|---|
| 1 | 1.27 | 0.63 | 0.79 |
| 1 | 1.91 | -0.15 | 0.52 |
| 1 | 1.94 | 1.58 | 0.51 |
| 1 | 1.06 | 0.15 | 0.94 |
| 1 | 1.69 | 0.62 | 0.59 |
| 1 | 1.50 | 1.70 | 0.67 |
| 1 | 1.38 | 1.90 | 0.72 |
| 1 | 1.21 | 1.43 | 0.82 |
| 1 | 1.27 | 0.67 | 0.79 |
| 1 | 1.38 | 2.32 | 0.72 |

```
Mittelwert der y_i                                  :   1.09
Fixed-Effects-Modell-basierter Effektstärkeschätzer :   1.08
```

# Fixed-Effect-Modell der Metaanalyse

## Theorem (Q-Statistik)

Gegeben sei das Fixed-Effects Modell der Metanalyse und es sei

$$Q := \sum_{i=1}^{k} w_i (y_i - \hat{\delta})^2 \text{ mit } \hat{\delta} := \frac{1}{\sum_{i=1}^{k} w_i} \sum_{i=1}^{k} w_i y_i \text{ und } w_i := \frac{1}{\sigma_i^2}, i = 1, ..., k. \tag{12}$$

Dann gilt approximativ für große studienspezifische Stichprobenumfänge, dass

$$Q \sim \chi^2(k-1). \tag{13}$$

Bemerkungen

- Die $Q$-Statistik ist die anhand der studienspezifischen Präzisionen gewichtete Summe der quadrierten Abweichungen der studienspezifischen Effektstärkeschätzer vom studienübergreifenden Effektstärkeschätzer. Intuitiv sprechen hohe Werte von $Q$ also gegen die Annahme einer festen Effekstärke über Studien.
- Unter der Annahme, dass die studienspezifischen Effektstärkerschätzer $y_i$ im Fixed-Effects-Modells tatsächlich auf ein und diesselbe Effektstärke $\delta$ zurückgehen, ist $Q$ approximativ $\chi^2$-verteilt. Ansätze für Beweise dieser Tatsache finden sich in Cochran (1937) und Hedges (1982).
- Ist unter $Q \sim \chi^2(k-1)$ die Wahrscheinlichkeit für einen beobachteten Wert oder einen extremeren als den beobachteten Wert der Q-Statistik klein, so spricht dies intuitiv gegen die Annahme des Fixed-Effects-Modells (nach Schaeuffele et al. (2024): "A P value of the Q statistic below 0.05 indicates heterogeneity.") Für eine vertiefte Diskussion, siehe Kulinskaya, Dollinger, and Bjørkestøl (2011b) und Kulinskaya, Dollinger, and Bjørkestøl (2011a).

# Fixed-Effect-Modell der Metaanalyse

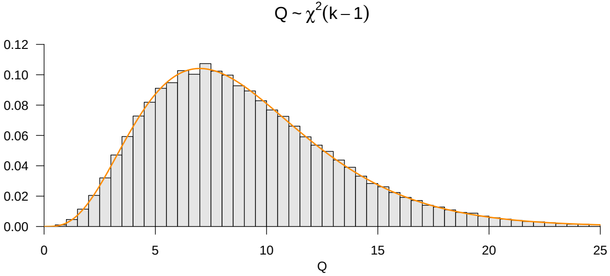
## Verteilung der Q-Statistik

$10^5$ Q-Statistikrealisierungen für $k := 10$, $\delta := 1$ und $\sigma_i \sim U(1, 2)$

```
set.seed(1)                                                 # reproduzierbare Ergebnisse
sim                = 1e5                                    # Anzahl Realisierungen
k                  = 10                                     # Anzahl an Studien
delta              = 1                                      # wahre, aber unbekannte, Effekstärke
Q                  = rep(NaN,sim)                           # Q-Statistik Array
for(s in 1:sim){                                            # Simulationsiterationen
    sigsqr_i       = rep(NaN,k)                             # wahre, bekannte, studienspezifische Varianz
    y_i            = rep(NaN,k)                             # Empirische Effekstärkenarray
    for(i in 1:k){                                          # Studieniterationen
        sigsqr_i[i] = runif(1,1,2)                          # wahre, bekannte, studienspezifische Varianz
        y_i[i]      = rnorm(1, delta, sqrt(sigsqr_i[i]))}   # empirische Effektstärke
    w_i            = 1/sigsqr_i                             # Wichtungsparameter
    delta_hat      = sum(w_i*y_i)/sum(w_i )                 # Fixed-Effects-Modell-basierter Effektstärkeschätzer
    Q[s]           = sum(w_i*(y_i - delta_hat)**2)          # Q-Statistik
}
```

## Verteilung der Q-Statistik

$10^5$ Q-Statistikrealisierungen für $k := 10$, $\delta := 1$ und $\sigma_i \sim U(1, 2)$



# Fixed-Effect-Modell der Metaanalyse

## Definition ($I^2$-Statistik)

Gegeben sei das Fixed-Effects Modell der Metanalyse und es sei

$$Q := \sum_{i=1}^k w_i (y_i - \hat{\delta})^2 \text{ mit } \hat{\delta} := \frac{1}{\sum_{i=1}^k w_i} \sum_{i=1}^k w_i y_i \text{ und } w_i := \frac{1}{\sigma_i^2}, i = 1, ..., k. \quad (14)$$

Dann ist die $I^2$-Statistik definiert als

$$I^2 := \max\left\{0, \frac{Q - (k-1)}{Q}\right\}. \quad (15)$$

Bemerkungen

- Die $I^2$-Statistik wurde von Higgins and Thompson (2002) vorgeschlagen. Der Erwartungswert einer $\chi^2(k-1)$-verteilten Zufallsvariable ist $k-1$. $I^2$ misst also in standardisierter Form und unter Annahme des Fixed-Effects-Modells approximativ, inwieweit $Q > \mathbb{E}(Q)$ ist. (nach Schaeuffele et al. (2024): "$I^2$ ranges from 0 to 100%, with 25% representing low, 50% moderate and 75% high heterogeneity.") Generell, also für beliebige Effekstärkemaße außerhalb des Fixed-Effects-Modells, muss $Q$ jedoch nicht $\chi^2(k-1)$ verteilt sein. Hoaglin (2016) gibt eine kritische Diskussion der Verwendung der $Q$- und $I^2$ Statistiken in der Metaanalyse.

Motivation

Fixed-Effect-Modell der Metaanalyse

**Random-Effects-Modell der Metaanalyse**

Selbstkontrollfragen

## Definition (Random-Effects-Modell der Metaanalyse)

Für Studien $i = 1, ..., k$ seien $y_i$ und $\sigma_i^2$ bekannte studienspezifische standardisierte Effektstärkeschätzer und ihre Varianzschätzer, respektive. Dann ist das *Random-Effects-Modell der Metaanalyse* gegeben durch

$$y_i \sim N\left(\delta_i, \sigma_i^2\right) \text{ und } \delta_i \sim N\left(\delta, \tau^2\right), \tag{16}$$

wobei

- $\delta$ die *wahre, aber unbekannte, Effektstärke*,
- $\tau^2$ die *wahre, aber unbekannte, studienübergreifende Effektstärkenvarianz* und
- $\delta_i$ die Zufallsvariable der *latentem Effektstärke der $i$ten Studie*

bezeichnen.

Bemerkungen

- Für die $y_i$ und $\sigma_i^2$ ergeben sich die gleichen Bemerkungen wie für das Fixed-Effects-Modell.
- Die Höhe der studienübergreifende Effektstärkenvarianz wird auch als *Studienheterogenität* bezeichnet.

# Random-Effects-Modell der Metaanalyse

Bemerkungen (fortgesetzt)

- Für $i = 1, ..., k$ impliziert die Definition die marginalen und die bedingten Wahrscheinlichkeitsdichtefunktionen

$$p(y_i|\delta_i) = N(y_i; \delta_i, \sigma_i^2) \text{ und } p(\delta_i) = N(\delta_i; \delta, \tau^2) \tag{17}$$

  und damit die gemeinsame WDF

$$p(y_i, \delta_i) = p(y_i|\delta_i)p(\delta_i) = N(y_i; \delta_i, \sigma_i^2)N(\delta_i; \delta, \tau^2). \tag{18}$$

- Mit den Eigenschaften der multivariaten Normalverteilung impliziert dies die Marginalverteilung

$$p(y_i) = N(y_i; \delta, \sigma_i^2 + \tau^2) \tag{19}$$

- In struktureller Form impliziert die Definition, dass

$$y_i = \delta + \xi_i + \varepsilon_i \text{ mit } \xi_i \sim N(0, \tau^2) \text{ und } \varepsilon_i \sim N(0, \sigma_i^2) \tag{20}$$

- Generativ betrachtet nimmt das Random-Effects-Modell also eine wahre, aber unbekannte, standardisierte Effektstärke $\delta$ an. Diese wahre, aber unbekannte, standardisierte Effektstärke wird in der $i$ten Studie unter dem Einfluss additivem normalverteilten Fehlers mit unbekannter Varianz $\tau^2$ als $\delta_i$ indirekt beobachtet. Direkt beobachtet wird dann mit $y_i$ eben dieses $\delta_i$ unter dem Einfluss additivem normalverteilten Fehlers mit bekannter Varianz $\sigma_i^2$. All dies entspricht einem speziellen Linear Mixed Model, wie wir an späterer Stelle sehen werden.

## Definition (DerSimonian-Laird Schätzung der REM Parameter)

Gegeben sei das Random-Effects-Modell der Metananalyse mit wahrer, aber unbekannter, standardisierter Effektstärke $\delta$ und wahrer, aber unbekannter, studienübergreifender Effektstärkevarianz $\tau^2$.

- Der *DerSimonian-Laird Schätzer von* $\tau^2$ ist definiert als

$$\hat{\tau}^2 = \frac{Q - (k-1)}{c} \text{ mit } c := \sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i}, \quad (21)$$

  wobei

$$Q := \sum_{i=1}^k w_i (y_i - \hat{\delta})^2 \text{ und } \hat{\delta} := \frac{1}{\sum_{i=1}^k w_i} \sum_{i=1}^k w_i y_i \text{ mit } w_i := \frac{1}{\sigma_i^2} \text{ für } i = 1, ..., k \quad (22)$$

  die $Q$-Statistik bzw. den ML-Schätzer des Fixed-Effects-Modells der Metaanalyse bezeichnen.

- Der *DerSimonian-Laird Schätzer-basierte Schätzer von* $\delta$ ist definiert als

$$\tilde{\delta} := \frac{1}{\sum_{i=1}^k \tilde{w}_i} \sum_{i=1}^k \tilde{w}_i y_i \text{ mit } \tilde{w}_i := \frac{1}{\hat{\tau}^2 + \sigma_i^2} \text{ für } i = 1, ..., k. \quad (23)$$

Bemerkungen

- Der DerSimonian-Laird Schätzer geht auf DerSimonian and Laird (1986) zurück.
- Es gibt viele verschiedene Schätzer für $\tau^2$ und $\delta$ im REM, vgl. Viechtbauer (2005) und Veroniki et al. (2016).
- Heutzutage wird meist eine Linear-Mixed-Modell-ReML-basierte Schätzung bevorzugt.

# Random-Effects-Modell der Metaanalyse

## Theorem (DerSimonian-Laird Schätzer von $\tau^2$)

Gegeben sei das Random-Effects-Modell der Metananalyse mit wahrer, aber unbekannter, studienübergreifender Effektstärkevarianz $\tau^2$ und es sei $\hat{\tau}^2$ der DerSimonian-Laird Schätzer von $\tau^2$. Dann gelten

(1) $\hat{\tau}^2$ ist ein Momentenmethodenschätzer von $\tau^2$ basierend auf dem Erwartungswert der $Q$-Statistik.

(2) $\hat{\tau}^2$ ist ein unverzerrter Schätzer von $\tau^2$.

Bemerkungen

Die Momentenmethode zur Schätzergewinnung besteht aus drei Schritten

(1) Schreiben der Momente als Funktionen des zu schätzenden Parameters.
(2) Ersetzen der Momente durch ihre Stichprobenäquivalente.
(3) Auflösen der Gleichung hinsichtlich des zu schätzenden Parameters.

Nach DerSimonian and Laird (1986): "(...) $Q$ ist used to derive a noniterative estimate of $\tau^2$ by equating the sample statistic with the corresponding expected value. This yields a weighted estimator (...)."

Beweis

(1) Wir bestimmen zunächst den Erwartungswert der $Q$-Statistik als Funktion der nach Voraussetzung bekannten Werte von $\sigma_1^2, ..., \sigma_k^2$ und des unbekannten Parameters $\tau^2$ (vgl. Kacker (2004), DerSimonian and Kacker (2007))

$$
\begin{aligned}
\mathbb{E}(Q) &= \mathbb{E}\left(\sum_{i=1}^k w_i(y_i - \hat{\delta})^2\right) \\
&= \sum_{i=1}^k w_i \mathbb{E}\left((y_i - \hat{\delta})^2\right) \\
&= \sum_{i=1}^k w_i \mathbb{E}\left(((y_i - \hat{\delta}) - 0)^2\right) \\
&= \sum_{i=1}^k w_i \mathbb{E}\left(((y_i - \hat{\delta}) - (\mathbb{E}(y_i) - \mathbb{E}(\hat{\delta})))^2\right) \\
&= \sum_{i=1}^k w_i \mathbb{E}\left(((y_i - \hat{\delta}) - \mathbb{E}(y_i - \hat{\delta}))^2\right) \\
&= \sum_{i=1}^k w_i \mathbb{V}\left(y_i - \hat{\delta}\right) \\
&= \sum_{i=1}^k w_i \left(\mathbb{V}(y_i) + \mathbb{V}(\hat{\delta}) - 2\mathbb{C}(y_i, \hat{\delta})\right).
\end{aligned}
\tag{24}
$$

Beweis (fortgeführt)

Dabei gilt nachdem Theorem zu den Eigenschaften des ML-Effektstärkeschätzers im FEM, dass

$$\mathbb{V}(\hat{\delta}) = \frac{\sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{(\sum_{i=1}^k w_i)^2}. \tag{25}$$

Mit der Eigenschaft der Kovarianz bei linear-affinen Transformationen sowie der Unabhängigkeit der $y_i$ gilt ferner

$$\begin{aligned}
\mathbb{C}(y_i, \hat{\delta}) &= \mathbb{C}\left(y_i, \frac{1}{\sum_{j=1}^k w_i} \sum_{j=1}^k w_j y_j\right) \\
&= \frac{1}{\sum_{j=1}^k w_j} \mathbb{C}\left(\sum_{j=1}^k w_j y_j, y_i\right) \\
&= \frac{\sum_{j=1}^k w_j \mathbb{C}\left(y_j, y_i\right)}{\sum_{j=1}^k w_j} \\
&= \frac{w_i \mathbb{C}(y_i, y_i)}{\sum_{j=1}^k w_j} \\
&= \frac{w_i \mathbb{V}(y_i)}{\sum_{j=1}^k w_j}.
\end{aligned} \tag{26}$$

# Random-Effects-Modell der Metaanalyse

Beweis (fortgeführt)

Es ergibt sich also

$$
\begin{aligned}
\mathbb{E}(Q) &= \sum_{i=1}^k w_i \left(\mathbb{V}(y_i) + \mathbb{V}(\hat{\delta}) - 2\mathbb{C}(y_i, \hat{\delta})\right) \\
&= \sum_{i=1}^k w_i \left(\mathbb{V}(y_i) + \frac{\sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{(\sum_{i=1}^k w_i)^2} - 2\,\frac{w_i \mathbb{V}(y_i)}{\sum_{i=1}^k w_i}\right) \\
&= \sum_{i=1}^k w_i \mathbb{V}(y_i) + \sum_{i=1}^k w_i \left(\frac{\sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{(\sum_{i=1}^k w_i)^2}\right) - 2\sum_{i=1}^k w_i \frac{w_i \mathbb{V}(y_i)}{\sum_{i=1}^k w_i} \\
&= \sum_{i=1}^k w_i \mathbb{V}(y_i) + \frac{\sum_{i=1}^k w_i \sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{(\sum_{i=1}^k w_i)^2} - 2\sum_{i=1}^k w_i \frac{w_i \mathbb{V}(y_i)}{\sum_{i=1}^k w_i} \\
&= \sum_{i=1}^k w_i \mathbb{V}(y_i) + \frac{\sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{\sum_{i=1}^k w_i} - 2\sum_{i=1}^k w_i \frac{w_i \mathbb{V}(y_i)}{\sum_{i=1}^k w_i} \\
&= \sum_{i=1}^k w_i \mathbb{V}(y_i) + \frac{\sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{\sum_{i=1}^k w_i} - 2\,\frac{\sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{\sum_{i=1}^k w_i} \\
&= \sum_{i=1}^k w_i \mathbb{V}(y_i) - \frac{\sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{\sum_{i=1}^k w_i}\,.
\end{aligned}
\tag{27}
$$

Beweis (fortgeführt)

Mit $\mathbb{V}(y_i) = \tau^2 + \sigma_i^2$ und $w_i = 1/\sigma_i^2$ergibt sich schließlich

$$
\begin{aligned}
\mathbb{E}(Q) &= \sum_{i=1}^k w_i \mathbb{V}(y_i) - \frac{\sum_{i=1}^k w_i^2 \mathbb{V}(y_i)}{\sum_{i=1}^k w_i} \\
&= \sum_{i=1}^k w_i (\tau^2 + \sigma_i^2) - \frac{\sum_{i=1}^k w_i^2 (\tau^2 + \sigma_i^2)}{\sum_{i=1}^k w_i} \\
&= \tau^2 \sum_{i=1}^k w_i + \sum_{i=1}^k w_i \sigma_i^2 - \tau^2 \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i} - \frac{\sum_{i=1}^k w_i^2 \sigma_i^2}{\sum_{i=1}^k w_i} \\
&= \tau^2 \left( \sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2 \sigma_i^2}{\sum_{i=1}^k w_i} \right) + \left( \sum_{i=1}^k w_i \sigma_i^2 - \frac{\sum_{i=1}^k w_i^2 \sigma_i^2}{\sum_{i=1}^k w_i} \right) \\
&= \tau^2 \left( \sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2 \sigma_i^2}{\sum_{i=1}^k w_i} \right) + \left( \sum_{i=1}^k \frac{\sigma_i^2}{\sigma_i^2} - \frac{\sum_{i=1}^k \frac{\sigma_i^2}{(\sigma_i^2)^2}}{\sum_{i=1}^k \frac{1}{\sigma_i^2}} \right) \\
&= \tau^2 \left( \sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2 \sigma_i^2}{\sum_{i=1}^k w_i} \right) + \left( \sum_{i=1}^k 1 - 1 \right) \\
&= \tau^2 \left( \sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2 \sigma_i^2}{\sum_{i=1}^k w_i} \right) + (k-1).
\end{aligned}
\tag{28}
$$

# Random-Effects-Modell der Metaanalyse

Beweis (fortgeführt)

Das Stichprobenäquivalent des Erwartungswerts der $Q$-Statistik ist der Mittelwert der $Q$-Statistik, allerdings bei nur einer Beobachtung. Gleichsetzen des Erwartungswerts von $Q$ mit seinem Stichprobenäquivalent ergibt also folgende Bedingung für den Momentenmethodenschätzer von $\tau^2$

$$Q = \mathbb{E}(Q) \Leftrightarrow Q = \hat{\tau}^2 \left( \sum_{i=1}^{k} w_i - \frac{\sum_{i=1}^{k} w_i^2 \sigma_i^2}{\sum_{i=1}^{k} w_i} \right) + (k-1) \tag{29}$$

Auflösen nach $\hat{\tau}^2$ ergibt dann den Momentenmethodenschätzer von $\tau^2$

$$\hat{\tau}^2 \left( \sum_{i=1}^{k} w_i - \frac{\sum_{i=1}^{k} w_i^2 \sigma_i^2}{\sum_{i=1}^{k} w_i} \right) + (k-1) = Q \Leftrightarrow \hat{\tau}^2 = \frac{Q-(k-1)}{c} \text{ mit } c := \sum_{i=1}^{k} w_i - \frac{\sum_{i=1}^{k} w_i^2 \sigma_i^2}{\sum_{i=1}^{k} w_i} \tag{30}$$

(2) Wir verzichten auf einen Beweis.

□

Simulation mit $k := 10$, $\delta := 1$, $\tau^2 := 0.1$, $\sigma_i^2 \sim U(0, 0.2)$

```
set.seed(2)                                                             # reproduzierbare Ergebnisse
k                = 10                                                   # Anzahl an Studien
delta            = 1                                                    # wahre, aber unbekannte, Effektstärke
tausqr           = .1                                                   # Between-Study Varianzparameter
sigsqr_i         = rep(NaN,k)                                           # studienspezifische Varianzparameterarray
delta_i          = rep(NaN,k)                                           # Latente Effektstärken Array
y_i              = rep(NaN,k)                                           # Empirische Effektstärken Array
for(i in 1:k){                                                          # Studieniterationen
    delta_i[i]   = rnorm(1,delta, sqrt(tausqr))                         # latente Studieneffektstärke
    sigsqr_i[i] = runif(1,.0,.2)                                        # bekannte studienspezifische Varianz
    y_i [i]      = rnorm(1,delta_i[i],sigsqr_i[i])}                     # empirische Effektstärke
w_i              = 1/sigsqr_i                                           # Wichtungsparameter des FEM Schätzers
delta_hat        = sum(w_i*y_i)/sum(w_i )                               # FEM ML-Effektstärkeschätzer
Q                = sum((w_i*(y_i-delta_hat)^2))                         # Q-Statistik
I2               = (Q - (k-1))/Q                                        # I^2-Statistik
c                = sum(w_i) - (sum(w_i**2)/sum(w_i))                    # c für DerSimonian-Laird
tau_hat_dl       = (Q - (k-1))/c                                        # DerSimonian-Laird tau^2 Schätzer
w_i_dl           = 1/(sigsqr_i + tau_hat_dl)                            # Wichtungsparameter des REM Schätzers
delta_hat_dl     = sum(w_i_dl*y_i)/sum(w_i_dl)                          # DerSimonian-Laird Effekstärkenschätzer
```

# Random-Effects-Modell der Metaanalyse

Simulation mit $k := 10$, $\delta := 1$, $\tau^2 := 0.1$, $\sigma_i^2 \sim U(0, 0.2)$

| delta | tau2 | delta_i | sigma2_i | y_i | w_i |
|---|---|---|---|---|---|
| 1 | 0.1 | 0.72 | 0.11 | 0.61 | 8.72 |
| 1 | 0.1 | 1.50 | 0.17 | 1.49 | 6.00 |
| 1 | 0.1 | 1.04 | 0.15 | 0.90 | 6.57 |
| 1 | 0.1 | 1.33 | 0.05 | 1.33 | 22.14 |
| 1 | 0.1 | 1.13 | 0.17 | 0.96 | 5.97 |
| 1 | 0.1 | 0.99 | 0.07 | 1.12 | 14.00 |
| 1 | 0.1 | 0.27 | 0.16 | 0.45 | 6.17 |
| 1 | 0.1 | 1.10 | 0.06 | 1.13 | 17.55 |
| 1 | 0.1 | 1.66 | 0.02 | 1.64 | 43.45 |
| 1 | 0.1 | 1.26 | 0.07 | 1.26 | 14.32 |

```
Q-Statistik                                     : 16.3747
I^2-Statistik                                   :  45.04
DerSimonian-Laird-Schätzer von tau^2            :  0.0604
Maximum-Likelihood-Effektstärkeschätzer des FEM :  1.261
DerSimonian-Laird-Effektstärkeschätzer des REM  :  1.1671
```

# Random-Effects-Modell der Metaanalyse

Simulation mit $k := 10$, $\delta := 1$, $\tau^2 := 0.1$, $\sigma_i^2 \sim U(0, 0.2)$

DerSimonian-Laird-Effektstärkeschätzung im Random-Effects-Modell mit metafor (Viechtbauer (2010))

```
library(metafor)                                          # R Paket
res = rma(yi = y_i, vi = sigsqr_i, method = "DL")         # Daten und DerSimonian-Laird ("DL") Methode
print(res)                                                # Ergebnisausgabe
```

```
Random-Effects Model (k = 10; tau^2 estimator: DL)

tau^2 (estimated amount of total heterogeneity): 0.0604 (SE = 0.0660)
tau (square root of estimated tau^2 value):      0.2458
I^2 (total heterogeneity / total variability):   45.04%
H^2 (total variability / sampling variability):  1.82

Test for Heterogeneity:
Q(df = 9) = 16.3747, p-val = 0.0595

Model Results:

estimate      se    zval    pval   ci.lb   ci.ub
  1.1671  0.1204  9.6899  <.0001  0.9310  1.4032  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Motivation

Fixed-Effect-Modell der Metaanalyse

Random-Effects-Modell der Metaanalyse

**Selbstkontrollfragen**

# Selbstkontrollfragen

1. Geben Sie die Definition des Fixed-Effects-Modells der Metaanalyse wieder und erläutern Sie es.
2. Geben Sie das Theorem zum Fixed-Effects-Modell-Maximum-Likelihood-Effektstärkeschätzer wieder.
3. Geben Sie das Theorem zur $Q$-Statistik wieder.
4. Geben Sie die Definition der $I^2$-Statistik wieder.
5. Geben Sie die Definition des Random-Effects-Modells der Metaanalyse wieder und erläutern Sie es.
6. Geben Sie das Theorem zur DerSimonian-Laird Schätzung der Random-Effects-Modell Parameter wieder.

# Referenzen I


Cochran, W. G. 1937. "Problems Arising in the Analysis of a Series of Similar Experiments." *Journal of the Royal Statistical Society Series B: Statistical Methodology* 4 (1): 102–18. https://doi.org/10.2307/2984123.

DerSimonian, Rebecca, and Raghu Kacker. 2007. "Random-Effects Model for Meta-Analysis of Clinical Trials: An Update." *Contemporary Clinical Trials* 28 (2): 105–14. https://doi.org/10.1016/j.cct.2006.04.004.

DerSimonian, Rebecca, and Nan Laird. 1986. "Meta-Analysis in Clinical Trials." *Controlled Clinical Trials* 7 (3): 177–88. https://doi.org/10.1016/0197-2456(86)90046-2.

Hedges, Larry V. 1982. "Fitting Categorical Models to Effect Sizes from a Series of Experiments." *Journal of Educational Statistics* 7 (2): 119. https://doi.org/10.2307/1164961.

Hedges, Larry V., and Ingram Olkin. 1985. *Statistical Methods for Meta-Analysis*. Academic Press.

Higgins, Julian P. T., and Simon G. Thompson. 2002. "Quantifying Heterogeneity in a Meta-Analysis." *Statistics in Medicine* 21 (11): 1539–58. https://doi.org/10.1002/sim.1186.

Hoaglin, David C. 2016. "Misunderstandings about $Q$ and 'Cochran's $Q$ Test' in Meta-Analysis." *Statistics in Medicine* 35 (4): 485–95. https://doi.org/10.1002/sim.6632.

Kacker, Raghu N. 2004. "Combining Information from Interlaboratory Evaluations Using a Random Effects Model." *Metrologia* 41 (3): 132–36. https://doi.org/10.1088/0026-1394/41/3/004.

Kulinskaya, Elena, Michael B. Dollinger, and Kirsten Bjørkestøl. 2011a. "Testing for Homogeneity in Meta-Analysis I. The One-Parameter Case: Standardized Mean Difference." *Biometrics* 67 (1): 203–12. https://doi.org/10.1111/j.1541-0420.2010.01442.x.

———. 2011b. "On the Moments of Cochran's $Q$ Statistic Under the Null Hypothesis, with Application to the Meta-Analysis of Risk Difference." *Research Synthesis Methods* 2 (4): 254–70. https://doi.org/10.1002/jrsm.54.

# Referenzen II


Lin, Lifeng, and Ariel M. Aloe. 2021. "Evaluation of Various Estimators for Standardized Mean Difference in Meta-Analysis." *Statistics in Medicine* 40 (2): 403–26. https://doi.org/10.1002/sim.8781.

Schaeuffele, Carmen, Laura E. Meine, Ava Schulz, Maxi C. Weber, Angela Moser, Christina Paersch, Dominique Recher, et al. 2024. "A Systematic Review and Meta-Analysis of Transdiagnostic Cognitive Behavioural Therapies for Emotional Disorders." *Nature Human Behaviour* 8 (3): 493–509. https://doi.org/10.1038/s41562-023-01787-3.

Veroniki, Areti Angeliki, Dan Jackson, Wolfgang Viechtbauer, Ralf Bender, Jack Bowden, Guido Knapp, Oliver Kuss, Julian PT Higgins, Dean Langan, and Georgia Salanti. 2016. "Methods to Estimate the Between-Study Variance and Its Uncertainty in Meta-Analysis." *Research Synthesis Methods* 7 (1): 55–79. https://doi.org/10.1002/jrsm.1164.

Viechtbauer, Wolfgang. 2005. "Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model." *Journal of Educational and Behavioral Statistics* 30 (3): 261–93. https://doi.org/10.3102/10769986030003261.

———. 2010. "Conducting Meta-Analyses in *R* with the **Metafor** Package." *Journal of Statistical Software* 36 (3). https://doi.org/10.18637/jss.v036.i03.