

Multivariate Modelle der Meta-Regression: Anwendungen in der Biometrie und Methodenvergleich mittels Simulationsstudie

Anne Michels, Joachim Hartung

Fachbereich Statistik,
Universität Dortmund

GMDS 2006,
11. September 2006

Inhalt

- 1 Motivation
- 2 Daten
- 3 Multivariate Meta-Regression
- 4 Anwendung
- 5 Simulation
- 6 Zusammenfassung

Motivation

- typischerweise werden in klinischen Studien mehrere Zielgrößen betrachtet
- dies führt zu mehreren Effektgrößen in jeder Studie, die häufig miteinander korreliert sind
- Verwendung von Standardmethoden hat Informationsverlust zur Folge, Korrelation geht nicht ein

Anwendungsbeispiel: fünf Studien zur Behandlung von Paradontitis

- 5 bei Berkey *et al.* (1998) vorgestellte Studien zur Behandlung von Paradontitis
- untersucht wurde der Erfolg einer chirurgischen (S) gegenüber einer nicht chirurgischen Behandlung (NS)
- betrachtete Zielgrößen:
 - Zahnfleischtaschentiefe PD (*probing depth*)
 - Attachment-Niveaus AL (*attachment level*)
- verwendet wurde das Split-Mouth-Design

Anwendungsbeispiel: Ergebnisse der fünf Paradontitisstudien

Studie	Jahr	n^1	Taschentiefe (<i>PD</i>)			Attachment-Niveau (<i>AL</i>)			Kovarianzmatrix	
			Reduktion		Differenz $S - NS$	Zuwachs		Differenz $S - NS$	<i>PD</i>	<i>AL</i>
			<i>S</i>	<i>NS</i>		<i>S</i>	<i>NS</i>			
1	1983	14	1,37	0,90	+0,47	0,09	0,41	-0,32	$\begin{pmatrix} 0,0075 & 0,0030 \\ 0,0030 & 0,0077 \end{pmatrix}$	
2	1982	15	1,80	1,60	+0,20	-0,10	0,50	-0,60	$\begin{pmatrix} 0,0057 & 0,0009 \\ 0,0009 & 0,0008 \end{pmatrix}$	
3	1979	78	2,10	1,70	+0,40	0,63	0,75	-0,12	$\begin{pmatrix} 0,0021 & 0,0007 \\ 0,0007 & 0,0014 \end{pmatrix}$	
4	1987	89	1,54	1,28	+0,26	-0,11	0,20	-0,31	$\begin{pmatrix} 0,0029 & 0,0009 \\ 0,0009 & 0,0015 \end{pmatrix}$	
5	1988	16	1,42	0,86	+0,56	0,10	0,49	-0,39	$\begin{pmatrix} 0,0148 & 0,0072 \\ 0,0072 & 0,0304 \end{pmatrix}$	

¹ Anzahl der Patienten in der Studie

Modell nach Raudenbush *et al.* (1988)

- Erweiterung der univariaten Meta-Regression um zusätzliche Zielgrößen
- in k Studien wird eine experimentelle Gruppe E mit einer Kontrollgruppe C verglichen
- Vergleich anhand von $p = 1, \dots, P$ Zielgrößen und den resultierenden Effektschätzern $\hat{\theta}_{ip}$, $i = 1, \dots, k$
- hier $P = 2$

Modell nach Raudenbush *et al.* (1988)

- das unterstellte Modell

$$\begin{pmatrix} \hat{\theta}_{i1} \\ \hat{\theta}_{i2} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \beta_1 x_{i1} \\ \beta_2 x_{i2} \end{pmatrix}, S_i \right) \quad i = 1, \dots, k$$

kann zusammengefasst werden zu

$$\hat{\theta} \sim \mathcal{N}(X\beta, S)$$

- Designmatrix X enthält separate Intercepts und Werte der Kovariable für beide Zielgrößen
- S ist Block-Diagonalmatrix der Kovarianzmatrizen S_i ,
 $i = 1, \dots, k$

Modell nach Raudenbush *et al.* (1988)

Schätzung im Modell $\hat{\theta} \sim \mathcal{N}(X\beta, S)$

- linear unverzerrter Aitken-Schätzer für β

$$\beta^* = (X^T S^{-1} X)^{-1} X^T S^{-1} \hat{\theta}$$

- Kovarianzmatrix von β^* durch

$$\text{Cov}(\beta^*) = (X^T S^{-1} X)^{-1}$$

- für Schätzungen $\hat{\beta}$ und $\widehat{\text{Cov}}(\hat{\beta})$ muss die unbekannte Kovarianzmatrix S durch ihre Schätzung \hat{S} ersetzt werden

Modell nach Raudenbush *et al.* (1988)

approximative Tests im Modell $\hat{\theta} \sim \mathcal{N}(X\beta, S)$:

- $H_0 : \theta = X\beta$ kann überprüft werden mit der Teststatistik

$$H_E = (\hat{\theta} - X\hat{\beta})^T \hat{S}^{-1} (\hat{\theta} - X\hat{\beta}) \stackrel{H_0}{\sim} \chi_{2k-4}^2$$

- $H_0 : \beta = 0$ kann überprüft werden mit der Teststatistik

$$H_R = \hat{\theta}^T \hat{S}^{-1} \hat{\theta} - H_E \stackrel{H_0}{\sim} \chi_4^2$$

- $H_0 : \beta_h = 0$ kann überprüft werden mit der Teststatistik

$$Z_h = \hat{\beta}_h / \hat{\sigma}_{hh} \stackrel{H_0}{\sim} \mathcal{N}(0, 1)$$

Modell nach Raudenbush *et al.* (1988)

verwendetes Effektmaß: standardisierte Mittelwertdifferenz

- Annahme: es gilt zumindest approximativ

$$Y_{Eijp} \sim \mathcal{N}(\mu_{Eip}, \sigma_{ip}^2), \quad i = 1, \dots, k \quad p = 1, \dots, P \quad j = 1, \dots, n_{Ei}$$

$$Y_{Cijp} \sim \mathcal{N}(\mu_{Cip}, \sigma_{ip}^2), \quad i = 1, \dots, k \quad p = 1, \dots, P \quad j = 1, \dots, n_{Ci}$$

- standardisierte Mittelwertdifferenz

$$\theta_{ip} = (\mu_{Eip} - \mu_{Cip})/\sigma_{ip}, \quad i = 1, \dots, k \quad p = 1, \dots, P$$

Modell nach Raudenbush *et al.* (1988)

- üblicher, verzerrter Schätzer für θ_{ip}

$$\hat{\theta}_{ip} = (\bar{x}_{Eip} - \bar{x}_{Cip}) / \hat{\sigma}_{ip}$$

- gepoolter Varianzschätzer

$$\hat{\sigma}_{ip}^2 = \frac{(n_{Eip} - 1)\hat{\sigma}_{Eip}^2 + (n_{Cip} - 1)\hat{\sigma}_{Cip}^2}{n_i - 2}$$

- θ_{ip} ist mit $n_i = n_{Ei} + n_{Ci}$ (approximativ) unverzerrt schätzbar durch

$$\hat{\theta}_{ip}^* = \left(1 - \frac{3}{4n_i - 9}\right) \hat{\theta}_{ip}$$

Modell nach Raudenbush *et al.* (1988)

- Varianz von $\hat{\theta}_{ip}^*$:

$$\text{Var}(\hat{\theta}_{ip}^*) = \frac{n_i}{n_{Ei} n_{Ci}} + \frac{\theta_{ip}^2}{2n_i}$$

- Kovarianz zwischen den standardisierten Mittelwertdifferenzen zweier Zielgrößen:

$$\text{Cov}(\theta_{ip}, \theta_{iq}) = \rho_{ipq} \sqrt{\text{Var}(\theta_{ip})} \sqrt{\text{Var}(\theta_{iq})} \quad i = 1, \dots, k \quad p, q = 1, \dots, P$$

ρ_{ipq} ist die Korrelation zwischen den Zielgrößen p und q

Modell nach Berkey *et al.* (1998)

- Erweiterung des Modells nach Raudenbush *et al.* (1988) um den Vektor der $d_i = (d_{i1}, d_{i2})^T$ der zufälligen Effekte in Studie i :

$$\begin{pmatrix} \hat{\theta}_{i1} \\ \hat{\theta}_{i2} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \beta_1 x_{i1} \\ \beta_2 x_{i2} \end{pmatrix} + \begin{pmatrix} d_{i1} \\ d_{i2} \end{pmatrix} + \begin{pmatrix} e_{i1} \\ e_{i2} \end{pmatrix} \quad i = 1, \dots, k$$

- unter entsprechender Normalverteilungsannahme gilt:

$$\begin{pmatrix} \hat{\theta}_{i1} \\ \hat{\theta}_{i2} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \beta_1 x_{i1} \\ \beta_2 x_{i2} \end{pmatrix}, D + S_i \right) \quad i = 1, \dots, k$$

Modell nach Berkey *et al.* (1998)

iterative Schätzung mit generalisierter Kleinster-Quadrate-Methode:

- 1 Anpassung eines Modells mit festen Effekten
- 2 Berechnung der Residuen $(\hat{\theta} - X\hat{\beta})$ und Umordnung zu einer $(k \times 2)$ -dimensionalen Matrix
- 3 erste Schätzung der Interstudienkovarianzmatrix D mittels

$$\hat{D}_{GLS} = \frac{(\hat{\theta} - X\hat{\beta})_{k \times 2}^T (\hat{\theta} - X\hat{\beta})_{k \times 2}}{(k - 2)} - \frac{1}{k} \sum_{i=1}^k \hat{S}_i$$

- 4 bei erneuter Berechnung der Schritte 1-3 werden die Kovarianzmatrizen \hat{S}_i durch die Summen $\hat{D}_{GLS} + \hat{S}_i$ ersetzt

Modell nach Berkey *et al.* (1998)

iterative Schätzung mit multivariater Maximum-Likelihood-Methode:

- gilt zumindest approximativ $\hat{\theta}_i \sim \mathcal{N}(X_i\beta, D + S_i)$ $i = 1, \dots, k$
- multivariaten Likelihood

$$L = \prod_{i=1}^k |(D + S_i)|^{-1/2} \exp\{-(\hat{\theta}_i - X_i\beta)^T (D + S_i)^{-1} (\hat{\theta}_i - X_i\beta) / 2\}$$

- Likelihood-Gleichungen

$$\hat{\beta} = \left(\sum_{i=1}^k X_i^T (\hat{D} + \hat{S}_i)^{-1} X_i \right)^{-1} \sum_{i=1}^k X_i^T (\hat{D} + \hat{S}_i)^{-1} \hat{\theta}_i$$

$$\hat{D}_{MML} = \sum_{i=1}^k \left[\hat{D} (\hat{D} + \hat{S}_i)^{-1} (\hat{\theta}_i - X_i \hat{\beta}) (\hat{\theta}_i - X_i \hat{\beta})^T (\hat{D} + \hat{S}_i)^{-1} \hat{D} + \hat{D} - \hat{D} (\hat{D} + \hat{S}_i)^{-1} \hat{D} \right] / k$$

Anwendung auf die fünf Parodontitisstudien

- aufgrund eines durchgeführten Homogenitätstests nur Modell mit zufälligen Effekten
- zur Schätzung wurde sowohl die GLS als auch die MML Methode verwendet

Ergebnisse im bivariaten Modell mit zufälligen Effekten

Schätzung	Zielgröße	Schätzer	Standard- abweichung	\hat{D}
GLS	<i>PD</i>	0,374	0,064	$\begin{pmatrix} 0,015 & 0,008 \\ 0,008 & 0,022 \end{pmatrix}$
	<i>AL</i>	-0,346	0,074	
MML	<i>PD</i>	0,346	0,050	$\begin{pmatrix} 0,007 & 0,009 \\ 0,009 & 0,026 \end{pmatrix}$
	<i>AL</i>	-0,338	0,080	

Null- hypothese	Test- statistik	Wert der Teststatistik		p -Wert	
		GLS	MML	GLS	MML
$H_0 : \beta = 0$	H_R	89,714	148,673	< 0,0001	< 0,0001
$H_0 : \alpha_1 = 0$	Z_1	5,810	6,862	< 0,0001	< 0,0001
$H_0 : \alpha_2 = 0$	Z_2	4,664	4,239	< 0,0001	< 0,0001

Ergebnisse im bivariaten Modell mit zufälligen Effekten

- positives Vorzeichen bei der Zahnfleischtaschentiefe
→ chirurgische Behandlung wirkt sich besser aus
- negatives Vorzeichen beim Attachment-Niveau
→ nicht chirurgische Behandlung wirkt sich besser aus
- Wahl der Therapieform hängt davon ab, welches Symptom stärker ausgeprägt ist und eine vorrangige Behandlung benötigt
- Ziel: möglichst große Wirkung auf beide Zielgrößen

Motivation

- im univariaten Modell mit festen Effekten hält der Test zur Überprüfung von $H_0 : \theta = 0$ das vorgegebene Niveau nicht ein (Hartung and Knapp, 2001)
- Ist dies auch im bivariaten Modell mit festen Effekten der Fall?
- betrachtet wird das tatsächliche Niveau der Tests
 - H_R zur Überprüfung von $H_0 : \beta = 0$
 - Z_h zur Überprüfung von $H_0 : \beta_h = 0$

bei $k = 3$ zu kombinierenden Studien und vorgegebenem Niveau $\alpha = 0,05$

- Effektmaß: standardisierte Mittelwertdifferenz

Annahmen

- 10.000 Wiederholungen des Modells mit zwei Zielgrößen ohne Kovariable

$$\begin{pmatrix} \hat{\theta}_{i1} \\ \hat{\theta}_{i2} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, S_i \right) \quad i = 1, \dots, 3$$

- erzeugte Zufallszahlen aus

$$\begin{pmatrix} Y_{tij1} \\ Y_{tij2} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma \right) \quad t = E, C \quad i = 1, \dots, 3, \quad j = 1, \dots, n_i$$

- mit Kovarianzmatrix $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$

verwendete Stichprobenpläne bei $k = 3$ zu kombinierenden Studien

- fünf verschiedenen Korrelationen zwischen den Zielgrößen:
 $\rho = 0; 0,05; 0,1; 0,25; 0,5$
- sechs verschiedene Stichprobenpläne:

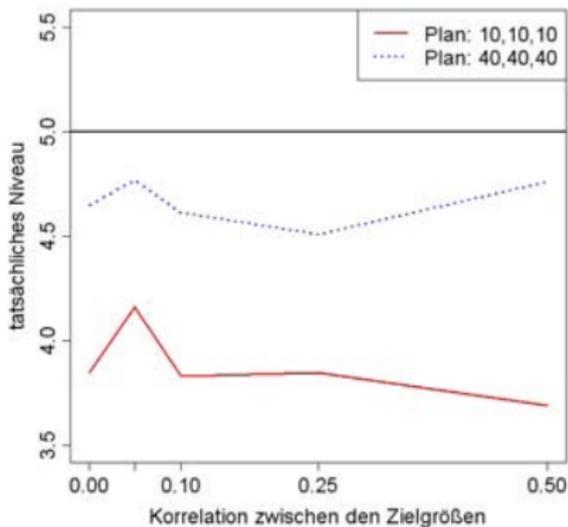
Plan	Eigenschaft	(n_1, n_2, n_3)
1	balanciert	(10,10,10)
2		(40,40,40)
3	moderat	(10,12,14)
4	unbalanciert	(40,48,56)
5	unbalanciert	(10,15,25)
6		(40,60,100)

Ergebnisse der Simulationsstudie im Modell mit festen Effekten:

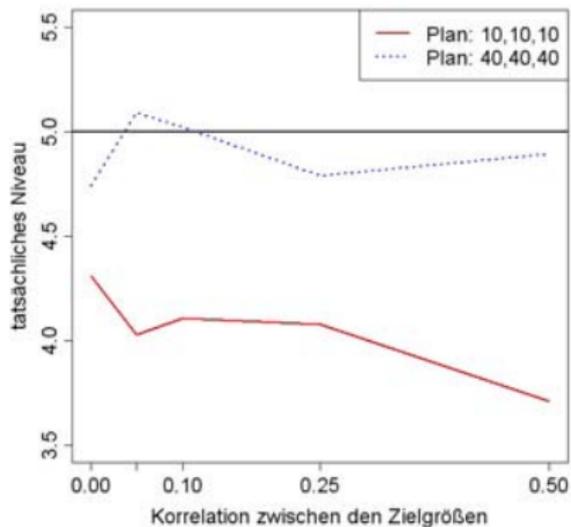
Tatsächliches Niveau von H_R , Z_1 und Z_2 in %

Plan	Test	Korrelation ρ					Plan	Test	Korrelation ρ				
		0	0,05	0,1	0,25	0,5			0	0,05	0,1	0,25	0,5
1	H_R	3,85	4,16	3,83	3,85	3,69	2	H_R	4,65	4,77	4,61	4,51	4,76
	Z_1	4,31	4,03	4,11	4,08	3,71		Z_1	4,74	5,09	5,02	4,79	4,89
	Z_2	4,33	4,07	4,05	3,77	3,87		Z_2	4,42	4,81	4,82	4,58	4,48
3	H_R	4,12	4,18	4,26	4,19	3,82	4	H_R	4,65	4,70	4,56	4,81	4,90
	Z_1	4,49	4,61	4,10	4,15	4,32		Z_1	4,83	4,88	4,68	5,10	4,89
	Z_2	4,41	4,32	4,34	4,40	3,75		Z_2	5,15	4,78	4,82	5,01	4,90
5	H_R	4,07	4,53	4,06	4,52	4,22	6	H_R	4,94	4,68	4,68	4,58	4,94
	Z_1	4,39	4,57	4,27	4,82	4,31		Z_1	5,07	4,75	4,94	4,42	4,92
	Z_2	4,41	4,64	4,63	4,49	4,17		Z_2	4,72	5,02	4,89	4,87	4,62

tatsächliches Niveau der Tests in Abhängigkeit der vorgegebenen Korrelation

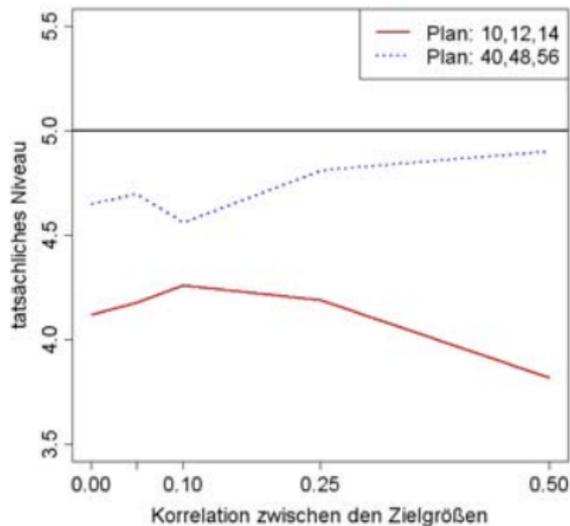


(a) Teststatistik H_R

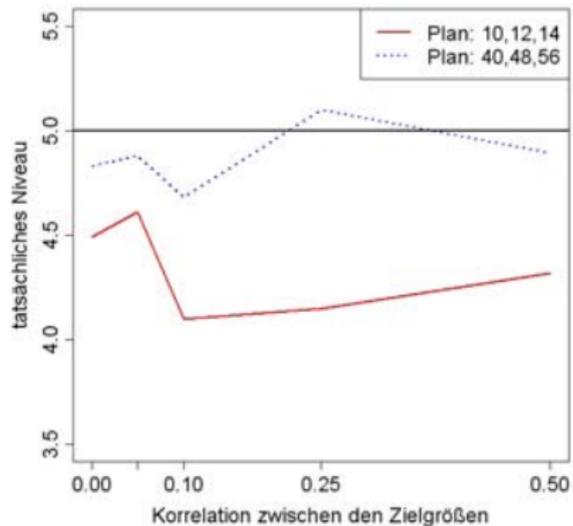


(b) Teststatistik Z_1

tatsächliches Niveau der Tests in Abhängigkeit der vorgegebenen Korrelation

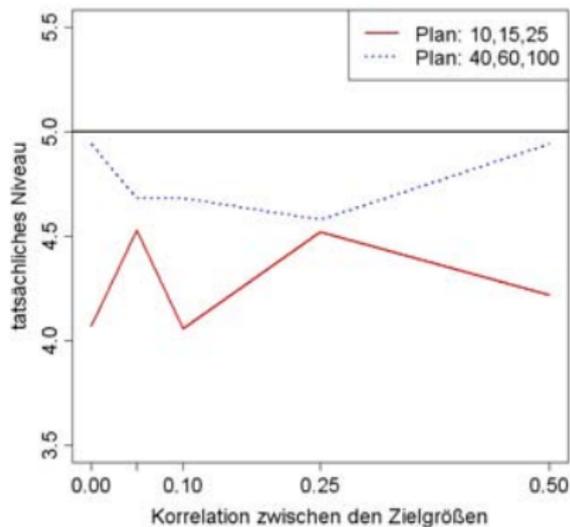


(c) Teststatistik H_R

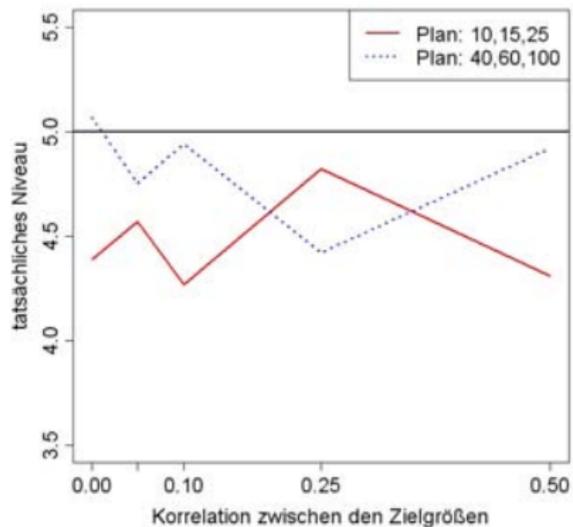


(d) Teststatistik Z_1

tatsächliches Niveau der Tests in Abhängigkeit der vorgegebenen Korrelation



(e) Teststatistik H_R



(f) Teststatistik Z_1

Ergebnisse der Simulationsstudie

- Ergebnisse nicht abhängig von der unterstellten Korrelation
- bei kleinen Stichprobenumfängen konservativ
- bei großen Stichprobenumfängen wird das Niveau annähernd eingehalten
- Ergebnisse bei großen Stichprobenumfängen unabhängig vom Stichprobenplan

weitere Ergebnisse

- Simulationsstudie für $k = 6$ und 9 Studien
 - vergleichbare Ergebnisse
- Simulationsstudie für die absolute Mittelwertdifferenz
 - unabhängig von der unterstellten Korrelation
 - bei kleinen Stichprobenumfängen liberal
 - bei großen Stichprobenumfängen halten sie das Niveau annähernd ein

weitere Ergebnisse

- Simulationsstudie im Modell mit zufälligen Effekten
 - für die Interstudienkovarianzmatrix

$$D = \begin{pmatrix} \tau^2 & 0 \\ 0 & \tau^2 \end{pmatrix}$$

werden die Heterogenitätsparameter $\tau^2 = 0,01, 0,1$ und 1 betrachtet

- Verwendung einer gestutzten Version von \hat{D}_{GLS}
- Korrelation $\rho = 0,5$
- 5.000 Wiederholungen

weitere Ergebnisse

- Ergebnisse für die absolute Mittelwertdifferenz
 - bei einer geringen Anzahl zu kombinierender Studien überschreiten alle Tests das vorgegebene Niveau deutlich, insbesondere bei hohem Heterogenitätsparameter
 - erhöhte Anzahl der Studien führt zu besseren Ergebnissen
- Ergebnisse für die standardisierte Mittelwertdifferenz
 - bei einer geringen Anzahl zu kombinierender Studien leicht konservative Ergebnisse für kleinen Heterogenitätsparameter
 - ansonsten auch deutlich antikonservative Ergebnisse
 - erhöhte Anzahl der Studien führt zu besseren Ergebnissen

Zusammenfassung

- Simulationsergebnisse des multivariaten Tests sind mit den bei Hartung and Knapp (2001) vorgestellten univariaten Ergebnissen vergleichbar
- im Fall der standardisierten Mittelwertdifferenz sind die klassischen Tests anwendbar
- im Fall der absoluten Mittelwertdifferenz ergeben sich zu viele ungerechtfertigte Signifikanzen

Literatur

- Berkey, C., Hoaglin, D., Antczak-Bouckoms, A., Mosteller, F., and Colditz, G. (1998). Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine*, **17**, 2537–2550.
- Hartung, J. and Knapp, G. (2001). On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Statistics in Medicine*, **20**, 1771–1782.
- Raudenbush, S., Becker, B., and Kalaian, H. (1988). Modeling multivariate effect sizes. *Psychological Bulletin*, **103**, 111–120.