Ein exakter Test für die Meta-Analyse von Studien mit binären Endpunkten

Oliver Kuß, Cora Gromann

Institut für Medizinische Epidemiologie, Biometrie und Informatik,
Universität Halle-Wittenberg, Halle (Saale)

GMDS 2004,
27. September 2004, Innsbruck

Gliederung

- Situation
- Standardmodelle für Meta-Analyse
- Ein exakter Test
- Berechnung/Software
- Simulationsuntersuchungen
- Diskussion

Situation I

„Wer Meta-Analyse macht, muss nichts neu erfinden. Steht alles schon bei den Altvorderen!“ (Hartung, 2001)

Definition:
„Meta-Analyse ist die quantitative, systematische Zusammenfassung mehrerer Studien mit dem Ziel, Information zu gewinnen, die aus keiner der Einzelstudien alleine erhalten werden kann.“ (Boissel et al., 1988)

Situation II

K unabhängige, randomisierte, klinische Studien zum Vergleich zweier Therapien A und B

Dichotome Zielgröße (Heilung: ja/nein) für jeden Patient

Ergebnisse der k-ten Studie können in Vierfeldertafel dargestellt werden:

	Heilung ja	Heilung nein	
Therapie A	n_{11k}	n_{12k}	$n_{1.k}$
Therapie B	n_{21k}	n_{22k}	$n_{2.k}$
	$n_{.1k}$	$n_{.2k}$	n_k

Situation III

Der Therapie-Effekt in der k-ten Studie wird mit Hilfe des LogOdds-Ratios

$$\theta_k = \log\left(\frac{p_{Ak}(1-p_{Bk})}{p_{Bk}(1-p_{Ak})}\right) \tag{1}$$

beschrieben und mittels $\widehat{p}_{Ak} = \frac{n_{11k}}{n_{1.k}}$ und $\widehat{p}_{Bk} = \frac{n_{21k}}{n_{2.k}}$ geschätzt.

Schätze die Varianz σ_k^2 von θ_k durch

$$\widehat{\sigma}_k^2 = \frac{1}{n_{11k}} + \frac{1}{n_{12k}} + \frac{1}{n_{21k}} + \frac{1}{n_{22k}} \tag{2}$$

Standard-Modelle für Meta-Analyse I: Fixed Effects

Um den gemeinsamen Therapie-Effekt θ zu schätzen, berechne

$$\widehat{\theta}_{FEM} = \frac{\sum_{k=1}^{K} w_k \widehat{\theta}_k}{\sum_{k=1}^{K} w_k} \tag{3}$$

mit $w_k = 1/\sigma_k^2$.

Unter der Homogenität-/Fixed-Effects-Annahme

$$\theta_1 = \theta_2 = \cdots = \theta_K = \theta = 0$$

ist die Teststatistik

$$T_{FEM} = \frac{\sum_{k=1}^{K} w_k \widehat{\theta}_k}{\sqrt{\sum_{k=1}^{K} w_k}} \tag{4}$$

standardnormalverteilt.

Standard-Modelle für Meta-Analyse II: Random Effects

Weniger restriktive Annahme bzgl. der θ_k:

$$\theta_k \sim N(\theta, \tau^2)$$

Schätze gemeinsamen Therapieeffekt θ nun durch

$$\widehat{\theta}_{REM} = \frac{\sum_{k=1}^{K} w_k^* \widehat{\theta}_k}{\sum_{k=1}^{K} w_k^*} \qquad (5)$$

mit $w_k^* = 1/(\sigma_k^2 + \tau^2)$.

Schätze τ^2 durch

$$\widehat{\tau}^2 = \max\left(\frac{\sum_{k=1}^{K} w_k(\widehat{\theta}_k - \widehat{\theta}_{FEM})^2 - (K-1)}{\sum w_k - \sum w_k^2 / \sum w_k}, 0\right) \qquad (6)$$

Standard-Modelle für Meta-Analyse III: Random Effects

Teste gemeinsamen Therapieeffekt ($H_0 : \theta = 0$) mit Teststatistik

$$T_{REM} = \frac{\sum_{k=1}^{K} w_k^* \widehat{\theta}_k}{\sqrt{\sum_{k=1}^{K} w_k^*}}, \tag{7}$$

die unter H_0 ebenfalls standardnormalverteilt ist.

Standard-Modelle für Meta-Analyse IV: Probleme

- Gewichte und Varianzen w_k, w_k^* und τ^2 werden als bekannt angenommen, müssen in der Realität aber geschätzt werden, d.h. Unsicherheit bei deren Schätzung wird im Verlauf der Analyse nicht berücksichtigt

- Umgang mit leeren Zellen

- Verteilungsannahmen

Ein exakter Test I: Vorüberlegung

Betrachte einzelne Vierfeldertafel:

	Heilung ja	Heilung nein	
Therapie A	n_{11}	n_{12}	$n_{1.}$
Therapie B	n_{21}	n_{22}	$n_{2.}$
	$n_{.1}$	$n_{.2}$	n

Unter der Annahme, dass alle Randsummen gegeben sind, gilt

$$Pr(n_{11} = x) = \binom{n_{.1}}{x}\binom{n_{.2}}{n_{1.} - x}/\binom{n}{n_{1.}}, \tag{8}$$

mit $\max(0, n_{1.} + n_{.1} - n) \leq x \leq \min(n_{1.}, n_{.1})$,

d.h. n_{11} ist hypergeometrisch verteilt mit Parametern $n, n_{1.}, n_{.1}$

Ein exakter Test II: Vorüberlegung

Unter der Nullhypothese der Unabhängigkeit (oder $\theta = 0$) gilt: $E(n_{11}) = \frac{n_{1.}n_{.1}}{n}$ und $Var(n_{11}) = \frac{n_{1.}n_{2.}n_{.1}n_{.2}}{n^2(n-1)}$.

Asymptotischer Test auf Unabhängigkeit:

$X^2_{MH} = \frac{(n_{11}-E(n_{11}))^2}{Var(n_{11})}$ ist unter der Nullhypothese asymptotisch χ^2_1-verteilt. Es gilt: $X^2_{MH} = \frac{n-1}{n}X^2$.

Exakter Test auf Unabhängigkeit:

Summiere die Wahrscheinlichkeiten $Pr(n_{11}) = x$ aller möglichen Vierfeldertafeln (mit gleichen Randsummen wie die beobachtete) auf, deren Test-Statistiken größer oder gleich die der beobachteten Vierfeldertafel sind (Exakter Fisher-Test, 1-seitig)

Ein exakter Test III

Verallgemeinere auf K Vierfeldertafeln:

Unter der Annahme der Unabhängigkeit in **allen** K Vierfeldertafeln (d.h. $\theta_1 = \theta_2 = \cdots = \theta_K = \theta = 0$) gilt:

Sei $S = \sum_{k=1}^{K} n_{11k}$. Dann ist

$E(S) = \sum_{k=1}^{K} E(n_{11k}) = \sum_{k=1}^{K} \frac{n_{1.k} n_{.1k}}{n_k}$,

$Var(S) = \sum_{k=1}^{K} Var(n_{11k}) = \sum_{k=1}^{K} \frac{n_{1.k} n_{2.k} n_{.1k} n_{.2k}}{n_k^2 (n_k - 1)}$.

$X_{MH}^2 = \frac{(S - E(S))^2}{Var(S)}$ ist unter der Nullhypothese asymptotisch χ_1^2-verteilt. (Mantel/Haenszel, 1959)

Ein exakter Test IV

Exakter Test auf Unabhängigkeit in **allen** K Vierfeldertafeln (d.h. $\theta_1 = \theta_2 = \cdots = \theta_K = \theta = 0$):

Summiere die Wahrscheinlichkeiten $Pr(S) = x$ aller möglichen Meta-Analysen (mit gleichen Randsummen wie die beobachtete) auf, deren Test-Statistiken größer oder gleich der der beobachteten Meta-Analyse sind.

Berechnung/Software

1. Interpretiere Meta-Analyse als logistisches Regressionsmodell mit Zielgröße Y (=Heilung) und Kovariablen X(=Therapie) und Z(=Studie)

   $\rightarrow$ Schätzung mit jeder Software, die exakte logistische Regression erlaubt (SAS PROC LOGISTIC, LogXact), ist möglich

2. Berechnung durch direktes Sampling

   (z.B. SAS-Funktion RAND('HYPER',....))

Simulationsuntersuchung

Verfahren: T_{FEM}, T_{REM}, X^2_{MH}, β_{EXACT}

Unterscheide: H0/H1 (Behandlungseffekt), FEM/REM

Design: $n_{1.k} = n_{2.k} \equiv 20$, $k = 5, 10, 20$, 1000 Läufe, SAS

Simulationsuntersuchung H0/FEM

p_A	p_B	k	Empirische Güte T_{FEM}	T_{REM}	X^2_{MH}	β_{EXACT}
0.2	0.2	5	0.029	0.027	0.045	0.032
0.2	0.2	10	0.027	0.025	0.044	0.037
0.2	0.2	20	0.026	0.026	0.043	0.038

Simulationsuntersuchung H1/FEM

			Empirische Güte			
p_A	p_B	k	T_{FEM}	T_{REM}	X^2_{MH}	β_{EXACT}
0.2	0.25	5	0.096	0.084	0.139	0.106
0.2	0.25	10	0.172	0.155	0.232	0.198
0.2	0.25	20	0.304	0.279	0.388	0.359
0.2	0.3	5	0.320	0.275	0.385	0.334
0.2	0.3	10	0.547	0.506	0.627	0.583
0.2	0.3	20	0.872	0.861	0.915	0.902

Simulationsuntersuchung H0/REM

			Empirische Güte			
p_A	p_B	k	T_{FEM}	T_{REM}	X^2_{MH}	β_{EXACT}
0.2 (± 0.05)	0.2 (± 0.05)	5	0.060	0.057	0.087	0.070
0.2 (± 0.05)	0.2 (± 0.05)	10	0.104	0.093	0.14	0.117
0.2 (± 0.05)	0.2 (± 0.05)	20	0.164	0.156	0.230	0.215

Simulationsuntersuchung H1/REM

p_A	p_B	k	Empirische Güte T_{FEM}	T_{REM}	X^2_{MH}	β_{EXACT}
0.2 (± 0.05)	0.25 (± 0.05)	5	0.137	0.123	0.185	0.154
0.2 (± 0.05)	0.25 (± 0.05)	10	0.217	0.198	0.285	0.253
0.2 (± 0.05)	0.25 (± 0.05)	20	0.366	0.357	0.436	0.419
0.2 (± 0.05)	0.3 (± 0.05)	5	0.343	0.299	0.413	0.355
0.2 (± 0.05)	0.3 (± 0.05)	10	0.548	0.522	0.606	0.579
0.2 (± 0.05)	0.3 (± 0.05)	20	0.762	0.742	0.810	0.794

Diskussion I: Vorteile

Es existiert eine exakte Version des Mantel-Haenszel-Tests, dessen Berechnung mit SAS ist möglich

- Keine Probleme mit Leerzellen

- Schätzung von Gewichten und Varianzen nicht nötig

- Keine Verteilungsannahmen, Studieneffekte werden „herausbedingt“

Diskussion II: Vorteile

- (Moderate) Robustheit gegen Verletzung der Homogenitäts-annahme

- Dieser Ansatz liefert auch exakte Schätzer für das Odds Ratio und exakten Homogenitätstest (Zelen, 1971)

Diskussion III: Nachteile

- Rechenzeit (nur für exakte log. Regression)
- Keine Meta-Regression möglich
- Asymptotischer (Mantel-Haenszel-) Test ist eigentlich besser

Ausblick

- Umfangreichere Simulation ($n_{i.k} > 20$)
- „Mid-p“-Methode zur p-Wert-Berechnung
- Direktes Sampling der Test-Statistik

Literatur

Birch MW, (1965): Maximum Likelihood in Three-way Contingency Tables. Journal of the Royal Statistical Society, Series B, 25, 220-233.

Agresti A, (1992): A Survey of Exact Inference for Contingency Tables (with Discussion). Statistical Science, 7, 131-177.

Lloyd J, (1999): Statistical Analysis of Categorical Data. Wiley & Sons, New York.

Mantel N, Haenszel W, (1959): Statistical Aspects of the Analysis of Data from Retrospective Disease. Journal of the National Cancer Institute, 22, 719-748.

Zelen M, (1971): The Analysis of Several 2x2 Contingency Tables. Biometrika, 58, 129-137.