



Der Homogenitätstest zum Vergleich zweier Erfolgsraten – ein Bayes-Ansatz im Vergleich zum klassischen χ^2 -Test

Joachim Gerß

Institut für Medizinische Informatik und Biomathematik
joachim.gerss@ukmuenster.de

Inhalt

1. Vergleich zweier Erfolgsraten mit dem χ^2 -Homogenitätstest
2. Bayes-Statistik im Beta-Binomial – Modell
3. Ein Bayes-Ansatz zum Vergleich zweier Erfolgsraten
4. Bayes-Faktoren als Ersatz zum klassischen Signifikanztest
5. Vergleich der Bayes-Ansätze mit dem χ^2 -Homogenitätstest
6. Fazit

1. Vergleich zweier „Erfolgsraten“ mit dem χ^2 -Homogenitätstest

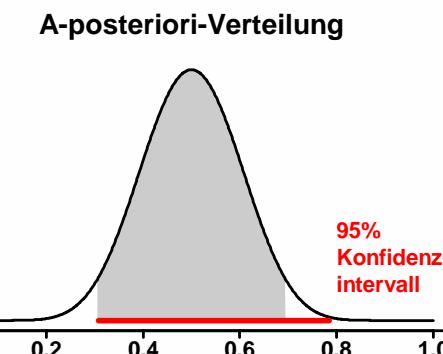
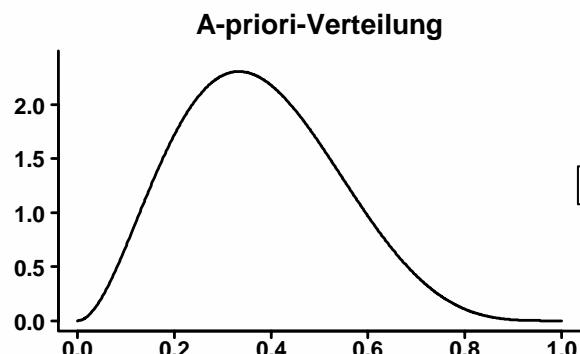
- Safety-Auswertung einer klinischen Studie (fiktive Daten)

	Nebenwirkung	keine Nebenwirkung	Σ
Plazebo	10 (=20%)	40	50
Aktive Therapie	18 (=36%)	32	50

- $H_0: p_0 = p_1$ gegen $H_1: p_0 \neq p_1$
 - $\chi^2 = 2.4306, p = 0.119 \Rightarrow$ kein Sicherheitsproblem?
- Mit welcher Wahrscheinlichkeit ist das Prüfpräparat als bedenklich anzusehen?
- Um wie viel wahrscheinlicher ist die Hypothese der Bedenklichkeit des Prüfpräparates als dessen Unbedenklichkeit?
- Wie groß ist die Beweislast für ein relevantes Sicherheitsproblem, etwa definiert als Erhöhung der Nebenwirkungsrate durch das Prüfpräparat um mehr als einen Faktor 2?

2. Frequentistische Statistik und Bayes-Statistik

- **Beispiel:** Schätzung einer unbekannten Erfolgswahrscheinlichkeit p
- **Frequentistische Statistik**
 - Der Parameter p ist eine *feste* Größe.
- **Bayes-Statistik**
 - Der Parameter p ist eine *Zufallsvariable* mit einer *Wahrscheinlichkeitsverteilung*.
→ Modellierung des *Wissens* über die Ausprägung der unbekannten Größe p



2. Das Beta-Binomial – Modell

- p : Gesuchte Erfolgswahrscheinlichkeit
 X : Beobachtete Anzahl Erfolge unter n Experimenten

- $\begin{cases} \text{Datenmodell: } X|p \sim \text{Bin}(n,p) \\ \text{A-priori – Verteilung: } f(p) \sim \text{Beta}(\alpha, \beta) \end{cases}$

- A-posteriori – Verteilung

$$\begin{aligned} f(p|x) &= \frac{f(x|p) \cdot f(p)}{f(x)} \\ &\sim f(x|p) \cdot f(p) \\ &= \binom{n}{x} p^x (1-p)^{n-x} \cdot \frac{1}{B(\alpha, \beta)} p^{\alpha-1} (1-p)^{\beta-1} \\ &\sim p^{x+\alpha-1} (1-p)^{n-x+\beta-1} \end{aligned}$$

$$\Rightarrow f(p|x) = \frac{1}{B(x+\alpha, n-x+\beta)} \cdot p^{x+\alpha-1} (1-p)^{n-x+\beta-1}, \text{ d.h. } p|x \sim \text{Beta}(\alpha+x, n-x+\beta)$$

3. Ein Bayes-Ansatz zum Vergleich zweier Erfolgsraten (1)

Kontrollgruppe:

- p_0 : Erfolgswahrscheinlichkeit

n_0 : Anzahl Patienten

X_0 : Beobachtete Anzahl Erfolge

- $\begin{cases} \text{Datenmodell: } X_0|p_0 \sim \text{Bin}(n_0, p_0) \\ \text{A-priori – Verteilung: } f(p_0) \sim \text{Beta}(1, 1) \end{cases}$
- A-posteriori – Verteilung:
 $p_0|x_0 \sim \text{Beta}(x_0+1, n_0-x_0+1)$

Aktive Therapiegruppe:

- p_1 : Erfolgswahrscheinlichkeit

n_1 : Anzahl Patienten

X_1 : Beobachtete Anzahl Erfolge

- $\begin{cases} \text{Datenmodell: } X_1|p_1 \sim \text{Bin}(n_1, p_1) \\ \text{A-priori – Verteilung: } f(p_1) \sim \text{Beta}(1, 1) \end{cases}$
- A-posteriori – Verteilung:
 $p_1|x_1 \sim \text{Beta}(x_1+1, n_1-x_1+1)$

Nichtinformative A-priori – Verteilung:
Gleichverteilung auf $[0, 1]$

3. Ein Bayes-Ansatz zum Vergleich zweier Erfolgsraten (2)

- $RR := p_1/p_0$

- $P(RR > c | x_0, x_1)$

$$= \int_{p_0, p_1 \text{ mit } p_1 > c \cdot p_0} f(p_0, p_1 | x_0, x_1) dp_0 dp_1$$

$$= \int_0^1 \left[\int_{p_0}^1 f(p_1 | x_1) dp_1 \right] f(p_0 | x_0) dp_0$$

$$= P\left(p_0 < \frac{1}{c} | x_0\right) - \sum_{i=0}^{\beta_1-1} \frac{B(\alpha_0 + \alpha_1 + i, \beta_0)}{B(\alpha_1, \beta_1) \cdot B(\alpha_0, \beta_0)} \cdot \binom{\beta_1-1}{i} \cdot (-1)^i \cdot \frac{1}{\alpha_1 + i} \cdot c^{\alpha_1 + i} \cdot P\left(Z < \frac{1}{c}\right)$$

$$\text{mit } Z \sim \text{Beta}(\alpha_0 + \alpha_1 + i, \beta_0) ,$$

$$\alpha_0 = x_0 + 1 , \quad \beta_0 = n_0 - x_0 + 1 ,$$

- $f_{RR}(c | x_0, x_1) = \frac{d}{dc} P(RR < c | x_0, x_1)$

$$\alpha_1 = x_1 + 1 , \quad \beta_1 = n_1 - x_1 + 1$$

4. Bayes-Faktoren

- $H_0: RR := p_1/p_0 \leq 1$ gegen $H_1: RR > c$
- $BF = \frac{P(H_1 | \text{Data})}{P(H_0 | \text{Data})} = \frac{P(RR > c | x_0, x_1)}{P(RR \leq 1 | x_0, x_1)}$

Strength of evidence	BF
Zero	1
Moderate	5
Moderate/strong	10
Strong	20
Very strong	40
Very strong	80

- Unterschiede zum klassischen Signifikanztest
 - Flexibilität in der Wahl der Alternative H_1
 - Metrisches und Maß der *Stärke* der Evidenz, intuitiv verständlich
 - keine Einhaltung formaler Gütekriterien (Signifikanzniveau, Power)

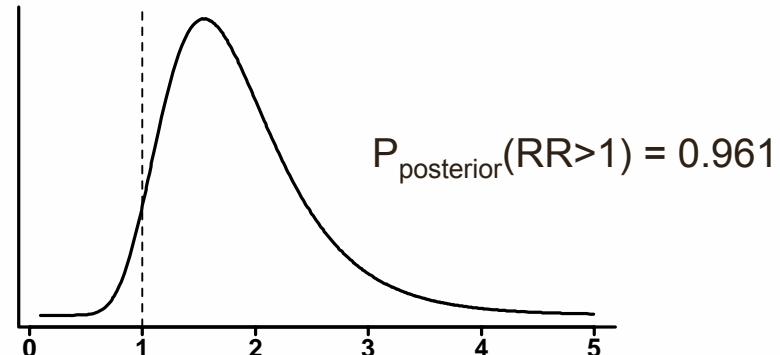
5. Vergleich der Bayes-Ansätze mit dem χ^2 -Homogenitätstest (1)

- Safety-Auswertung einer klinischen Studie (fiktive Daten)

	Nebenwirkung	keine Nebenwirkung	Σ
Plazebo	10 (=20%)	40	50
Aktive Therapie	18 (=36%)	32	50

- $H_0: p_0=p_1$ gegen $H_1: p_0 \neq p_1$
 - $\chi^2=2.4306$, $p=0.119$
- Mit welcher Wahrscheinlichkeit ist das Prüfpräparat als bedenklich anzusehen?

A-posteriori-Verteilung des RR



5. Vergleich der Bayes-Ansätze mit dem χ^2 -Homogenitätstest (1)

- Safety-Auswertung einer klinischen Studie (fiktive Daten)

	Nebenwirkung	keine Nebenwirkung	Σ
Plazebo	10 (=20%)	40	50
Aktive Therapie	18 (=36%)	32	50

- $H_0: p_0=p_1$ gegen $H_1: p_0 \neq p_1$
- $\chi^2=2.4306$, $p=0.119$

➤ Um wie viel wahrscheinlicher ist die Hypothese der Bedenklichkeit des Prüfpräparates als dessen Unbedenklichkeit?

$H_0: RR:=p_1/p_0 \leq 1$ gegen $H_1: RR > 1$

$$BF = \frac{P(H_1 | Data)}{P(H_0 | Data)} = \frac{P(RR > 1 | x_0, x_1)}{P(RR \leq 1 | x_0, x_1)} = \frac{0.961}{0.039} = 24.6 \quad (\text{strong evidence})$$

5. Vergleich der Bayes-Ansätze mit dem χ^2 -Homogenitätstest (1)

- Safety-Auswertung einer klinischen Studie (fiktive Daten)

	Nebenwirkung	keine Nebenwirkung	Σ
Plazebo	10 (=20%)	40	50
Aktive Therapie	18 (=36%)	32	50

- $H_0: p_0=p_1$ gegen $H_1: p_0 \neq p_1$
 - $\chi^2=2.4306$, $p=0.119$
- Wie groß ist die Beweislast für ein relevantes Sicherheitsproblem, etwa definiert als Erhöhung der Nebenwirkungsrate durch das Prüfpräparat um mehr als einen Faktor 2?

$H_0: RR:=p_1/p_0 \leq 1$ gegen $H_1: RR > 2$

$$BF = \frac{P(H_1 | \text{Data})}{P(H_0 | \text{Data})} = \frac{P(RR > 2 | x_0, x_1)}{P(RR \leq 1 | x_0, x_1)} = \frac{0.340}{0.039} = 8.7 \quad (\text{moderate evidence})$$

5. Vergleich der Bayes-Ansätze mit dem χ^2 -Homogenitätstest (2)

$n_0=n_1$	\hat{p}_0	\hat{p}_1	p-Wert	RR 95% KI	$P(RR>1)$	$BF(RR>1/\leq 1)$	$BF(RR>2/\leq 1)$
50	20%	36%	0.119	0.921 - 3.372	0.961	24.6	8.7
50	10%	15%	0.552	0.569 - 4.197	0.806	4.1	1.5
50	10%	20%	0.263	0.735 - 4.962	0.914	10.6	5.2
50	10%	25%	0.068	0.987 - 6.107	0.980	49.0	33.0
30	10%	20%	0.470	0.560 - 6.157	0.850	5.7	2.9
30	20%	25%	0.760	0.537 - 3.183	0.723	2.6	0.6
100	10%	20%	0.075	0.979 - 3.895	0.975	39.3	18.6
100	10%	25%	0.009	1.251 - 4.715	0.997	375.1	267.5
200	10%	20%	0.008	1.206 - 3.237	0.997	388.3	184.2
300	10%	20%	0.001	1.323 - 2.974	1.000	3305.3	1580.4

6. Fazit: Bayes-Verfahren in klinischen Studien

- Konfirmative Auswertung der primären Zielgröße einer klinischen Studie:
Anwendung klassischer Signifikanztests (Entscheidungstheorie)
- Bayes-Verfahren sind vorteilhaft in darüber hinausgehenden Fragestellungen:
 - Explorative und modellbildende Analysen
 - Safety-Analysen
 - Fallzahlschätzung
 - Adaptive Randomisierung
 - Data Monitoring in Zwischenauswertungen
- Bayes-Ansatz zum Vergleich zweier Erfolgsraten
 - Einfaches Verfahren, leicht umsetzbar und ohne viele „Stellschrauben“
 - Vorwissen kann, *aber muss nicht* (!) eingebracht werden
 - Flexible und direkte Beantwortung intuitiver Fragen („real world questions“)



Thomas Bayes
(1702-1761)

Literatur

- Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis. Boca Raton: Chapman&Hall/CRC; 1995.
- Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. New York: Wiley; 2004.
- Goodman SN. Introduction to Bayesian methods I: measuring the strength of evidence. Clinical Trials 2005; 2: 282-290.
- Louis TA. Introduction to Bayesian methods II: fundamental concepts. Clinical Trials 2005; 2: 291-294.
- Berry DA. Introduction to Bayesian methods III: use and interpretation of Bayesian tools in design and analysis. Clinical Trials 2005; 2: 295-300.