
Diverse Ergänzungen und das Multiplizitätsproblem

Dr. Dirk Hasenclever
IMISE, Leipzig

Hasenclever@IMISE.uni-Leipzig.de

Transformation von Dichten

Sei X eine Zufallsvariable mit Dichte $f_x(x)dx$ und Verteilungsfunktion $F_x(x)$.

Sei h eine streng monotone stetig differenzierbare Transformation mit Umkehrabbildung h^{-1}

Definiere die Zufallsvariable $Y := h(X)$.

Bestimme die Dichte $f_y(y) dy$: [Transformationsformel für Integrale!]

$$F_y(y) = F_x(h^{-1}(y)) = \int_{h^{-1}(-\infty)}^{h^{-1}(y)} f_x(w) dw = \int_{-\infty}^y f_x(h^{-1}(u)) \frac{d(h^{-1})}{du}(u) du$$


Also

$$f_y(y) dy = f_x(h^{-1}(y)) \cdot \frac{d(h^{-1})}{dy}(y) dy$$

Transformation von Dichten (Bsp. Lognormalverteilung)

Beispiel :

Sei $X \sim N(\mu, \sigma^2)$

Definiere Zufallsvariable $Y := \exp(X)$

Y heißt lognormalverteilt, da $\ln(Y)$ normalverteilt ist.

Bestimme die Dichte der Lognormalverteilung :

$$F_Y(y) = F_X(\ln(y)) = \int_{\ln(0)}^{\ln(y)} f_X(w) dw = \int_{-\infty}^y f_X(\ln(u)) \frac{d(\ln(u))}{du} du = \int_{-\infty}^y f_X(\ln(u)) \frac{1}{u} du$$

!

Also

$$f_Y(y) dy = \frac{1}{y} \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(y) - \mu)^2}{2\sigma^2}\right) dy$$

Transformation durch Verteilungsfunktion

P-Werte sind unter H_0 gleichverteilt auf $[0,1]$

Sei T Zufallsvariable mit Dichte f_T und Verteilungsfunktion F_T

Definiere Zufallsvariable $P := F_T(T)$

Bestimme die Dichte von P :

$$\begin{aligned} F_P(p) &= F_T(F_T^{-1}(p)) = \int_{F_T^{-1}(0)}^{F_T^{-1}(p)} f_T(w) dw = \int_0^p f_T(F_T^{-1}(u)) \cdot \frac{dF_T^{-1}(u)}{du} du \\ &= \int_0^p f_T(F_T^{-1}(u)) \frac{1}{\left(\frac{dF_T(F_T^{-1}(u))}{dt} \right)} du = \int_0^p f_T(F_T^{-1}(u)) \frac{1}{(f_T(F_T^{-1}(u)))} du = \int_0^p du \end{aligned}$$

Ableitung der
Umkehrfunktion



Also $f_P(p) = 1$ und $P \sim \text{Uniform}(0,1)$

Das Multiplizitätsproblem

- Das Multiplizitätsproblem tritt immer dann auf, wenn mehrfach getestet werden soll
 - Mehrfaches Testen auf akkumulierenden Daten
 - Subgruppenanalysen
 - Komponentenweises Testen hochdimensionaler Daten
(z.B. Genexpressionsdaten $G=20.000$ -> Vorlesung Biometrie II)
 - Mehrfache Kontraste (z.B. Studien mit mehr als 2 Armen)

Problem der Beurteilung kumulierender Daten

- Ein Experimentator kommt mit n Beobachtungen zu einem frequentistischen Statistiker in der Erwartung mit den Daten eine Nullhypothese zum 5% Niveau ablehnen und mit diesem Resultat eine Publikation schreiben zu können.
- Der Test lehnt aber ärgerlicherweise gerade nicht ab. Der Wissenschaftler glaubt trotzdem, dass die Nullhypothese falsch ist. Er fragt: „Wie viele weitere Beobachtungen brauche ich um die Nullhypothese abzulehnen?“
- Der Statistiker sagt: „Sorry, wie viele Daten auch immer Sie zusätzlich beibringen, die Nullhypothese kann nicht mehr zum 5% Niveau verworfen werden!“ (!?!)

Zweimal Testen falls noch nicht signifikant ? I

Einfachster Fall: Ziehe Beobachtungen $X_i \sim N(0,1)$

Teste nach n_1 Beobachtungen.

Falls noch nicht signifikant, teste nach weiteren n_2 Beobachtungen.

Wahrscheinlichkeit, die wahre Nullhypothese abzulehnen:

$$\begin{aligned} pr(rej(H_0)) &= pr(rej_{test1}(H_0)) \\ &+ (1 - pr(rej_{test1}(H_0))) * pr(rej_{test2}(H_0) | \neg rej_{test1}(H_0)) \end{aligned}$$

Bedingte Power > 0

ist jedenfalls größer als: $pr(rej_{test1}(H_0)) = \alpha$

per constructionem

Zweimal Testen

falls noch nicht signifikant ? II

Einstichproben test bei bekannter Varianz: $Z = \frac{\bar{X}_n}{1/\sqrt{n}} \sim N(0,1)$
(einseitig)

Sei z_1 und z_2 die Teststatistiken aufgrund der ersten n_1 bzw. separat der n_2 Beobachtungen.

Teststatistik auf n_1+n_2 Beobachtungen:

$$Z_{1+2} = \frac{\sqrt{n_1} * z_1 + \sqrt{n_2} * z_2}{\sqrt{n_1 + n_2}}$$

$$pr(rej(H_0)) =$$

$$pr(z_1 \geq Z_\alpha) + (1 - pr(z_1 \geq Z_\alpha)) * pr(z_{1+2} > Z_\alpha \mid z_1 < Z_\alpha)$$

$$= \alpha + (1 - \alpha) \int_{-\infty}^{Z_\alpha} (1 - \Phi(\frac{Z_\alpha \sqrt{n_1 + n_2} - z' \sqrt{n_1}}{\sqrt{n_2}})) * \phi(z') dz'$$

↑
Verteilungsfunktion und Dichte der Standardnormalverteilung

Zweimal Testen

falls noch nicht signifikant ? III

Bsp.: Testet man beide Male zum Signifikanzniveau $\alpha=0.05$, dann wird **der Fehler erster Art für die Gesamtstrategie inflationiert**:

z.B. $n_1=20$ $n_2=20$ 0.080

$n_1=20$ $n_2=100$ 0.090

(numerisches Ausrechnen des Integrals)

Daher gibt es sogenannte **Sequentialpläne** für mehrfach Teststrategien auf kumulierenden Daten.

Man testet jeweils mit einem nominell niedrigeren $\alpha' < \alpha$, um insgesamt den Fehler erster Art zu kontrollieren.

z.B.: Man testet 5x nominell zum 0.0158 Niveau (Pocock), um insgesamt $\alpha=0.05$ einzuhalten.

Testen auf kumulierenden Daten

Testet man immer wieder
auf kumulierenden Daten,

so bekommt man mit
Wahrscheinlichkeit 1

irgendwann einmal ein
signifikantes Resultat,

auch wenn die
Nullhypothese wahr ist

Dies folgt aus einem „Satz
vom iterierten Logarithmus“

Tests nach Gesamt-
jeweils n Beob. Fehler 1. Art
zum 5% Niveau

1	0,05
2	0,08
3	0,11
4	0,13
5	0,14
10	0,19
20	0,25
50	0,32
100	0,37
1000	0,53
unendlich	1

Angenommen: Normalverteilt
mit bekannter Varianz

Gruppensequentielle Designs

- Anzahl der Zwischenauswertungen vorher festgelegt
- nach jeder Auswertung: Entscheidung über Fortführung
- „Signifikanzkorrektur“: Einzeltests mit einem kleineren α -Niveau; bei Endauswertung ist Gesamt- α „verbraucht“

Bsp.: Einzel- α bei 4 Zwischenauswertungen, Gesamt- $\alpha=0.05$

Auswertung Nr.	Einzel- α nach		
	O'Brien-Fleming	Pocock	Peto
1	0.00001	0.016	0.001
2	0.0013	0.016	0.001
3	0.008	0.016	0.001
4	0.023	0.016	0.001
5	0.041	0.016	0.049

Verfahrenseigenschaften von Likelihood Hypothesenvergleichen

- Wie wahrscheinlich ist es durch Likelihoodratios in die Irre geführt zu werden?

Seien H_A und H_B Hypothesen über die Zufallsgröße X mit Wahrscheinlichkeitsverteilungen $pr_A(x)$ und $pr_B(x)$.

Satz : "Falls B wahr ist, ist es unwahrscheinlich starke Evidenz für A zu beobachten."

Sei B wahr und $k > 1$, dann gilt : $pr\left(\frac{pr_A(x)}{pr_B(x)} \geq k\right) \leq \frac{1}{k}$

Beweis : Sei $S = \left\{x \text{ im Stichprobenraum} \mid \frac{pr_A(x)}{pr_B(x)} \geq k\right\}$

$$pr(S) \underset{\substack{\uparrow \\ \text{B wahr}}}{=} \sum_{x \in S} pr_B(x) \underset{\substack{\uparrow \\ \text{Def. von S}}}{\leq} \sum_{x \in S} pr_A(x) / k \underset{\substack{\uparrow \\ \sum_{x \in S} pr_A(x) \leq 1}}{\leq} \frac{1}{k} \quad \otimes$$

Likelihood-Paradigma:

Kein Multiplizitätsproblem auf akkumulierenden Daten

Es gilt sogar der Satz von Robbins :

Sei LR_n die beobachtete Likelihoodratio nach n Beobachtungen $n = 1, \dots, \infty$.

Dann gilt :

$$\text{pr}(LR_n \geq k \text{ für ein } n \geq 1) \leq \frac{1}{k} \quad \text{für alle } k \geq 1$$

Das bedeutet, dass ein Experimentator,
der seine Studie solange fortführen will,
bis ein genehmes Resultat herauskommt,
schlechte Karten hat...

Abhängigkeit der Dateninterpretation von der „Stopping-Rule“?: Münzwurf Beispiel I

Eine Münze wird $N = 12$ mal geworfen : 9 Kopf und 3 Zahl

$H_0 : \text{pr}(\text{Kopf}) = 0.5$

Falls eine feste Fallzahl N geplant war, ist der Versuch **Binomial**.

Die Zufallsvariable ist die Anzahl k von Kopf $\in \{0, 1, \dots, 12\}$.

Likelihoodfunktion :

$$L_1(k, \theta) = \binom{N}{k} \theta^k (1 - \theta)^{N-k} = 220 * \theta^9 (1 - \theta)^3$$

Falls die Fallzahl variabel war, und die Stopp - Regel hieß : Stopp sobald 3 mal Zahl, dann ist der Versuch **Negativ - binomial**.

Der Samplespace der Zufallsvariable k (Anzahl k von Kopf) ist nun $\mathbb{R}_0 = \{0, 1, 2, 3, \dots\}$

Likelihoodfunktion :

$$L_1(k, \theta) = \binom{N+k-1}{k} \theta^k (1 - \theta)^{N-k} = 55 * \theta^9 (1 - \theta)^3$$

Abhängigkeit der Dateninterpretation von der „Stopping-Rule“?: Münzwurf Beispiel II

Fall1 (Binomialer Versuch) : $\text{pr}(k \geq 9) = \sum_{j=9}^{12} L_1(j, \theta) = 0.075$ Nicht signifikant!

Fall2 (Negativ - Binomialer Versuch) : $\text{pr}(k \geq 9) = \sum_{j=9}^{\infty} L_2(j, \theta) = 0.0325$ Signifikant!

Aber im Likelihood - Paradigma fehlt

diese Abhängigkeit von der Abbruchregel bzw. des Stichprobenraums,

da sich bei Quotientenbildung die unterschiedlichen Binomialkoeffizienten wegkürzen :

$LR(\theta = 0.75 \text{ versus } \theta = 0.5) : 4,8$ in beiden Fällen.

Evidentielle Irrelevanz des Stopp-Kriteriums

Sei $X^j = (X_1, \dots, X_j)$ eine Folge von Zufallsvariablen mit gemeinsamer Dichte $f_j(X^j/\theta)$.
Dies entspricht dem Stand einer Studie nach jeweils j Beobachtungen.
Beschreibe $\tau_j(X^j)$ die Wahrscheinlichkeit im j -ten Schritt die Studie abubrechen.

Die Likelihood dafür, dass die Studie nach n Schritten mit Ergebnis X^n stoppt, ist :

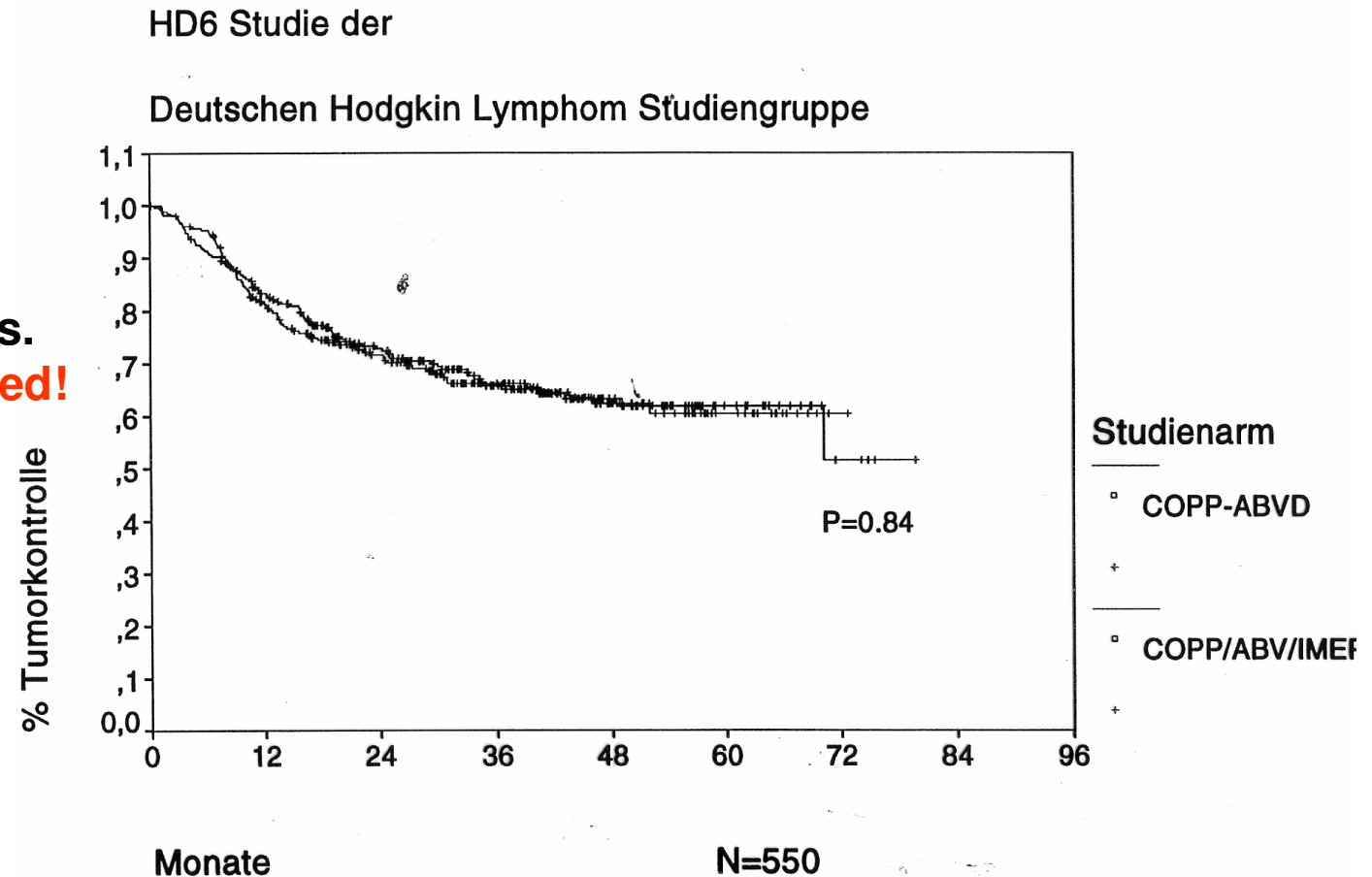
$$f(n, X^n/\theta) = \prod_{j=1}^{n-1} (1 - \tau_j(X^j)) * \tau_n(X^n) * f_n(X^n/\theta)$$

$\overbrace{\hspace{10em}}$
unabhängig von θ

Daher hängt relative Evidenz über θ NICHT von τ ab!!

Subgruppenanalyse HD6

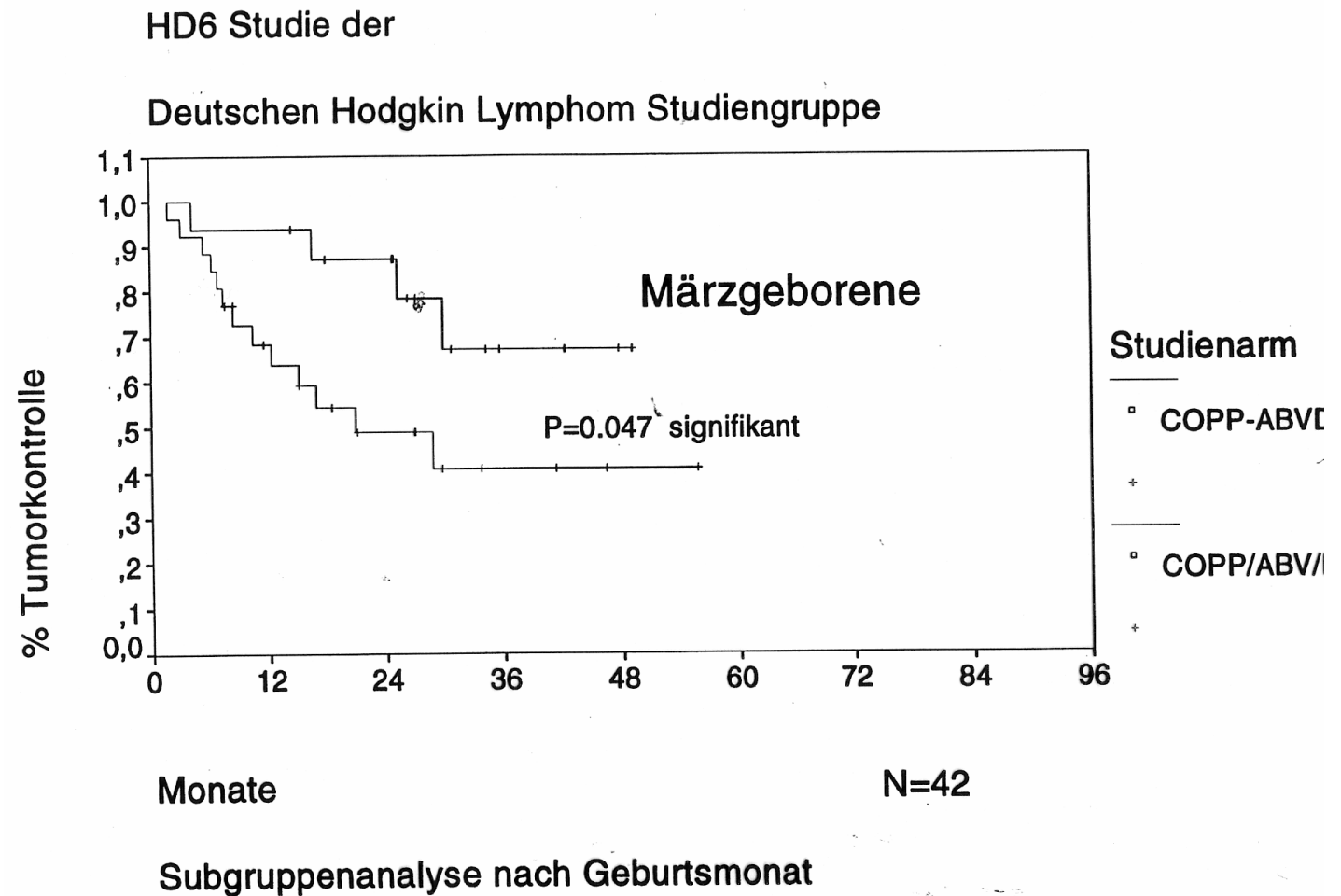
Randomisierter
Chemotherapie
Vergleich in
fortgeschrittenen
Stadien
eines
Hodgkin Lymphoms.
Keinerlei Unterschied!



Subgruppenanalyse HD6

Subgruppen
Analyse
nach
Geburtsmonat

Signifikant !?



Subgruppenanalyse HD6

HD6 Studie der Deutschen Hodgkin Lymphom Studiengruppe

Test des Therapieeffekts $p=0.84$

Subgruppenanalyse nach Geburtsmonat:

Januar	$p=0.87$
Februar	$p=0.56$
März	$p=0.047$ formal signifikant
April	$p=0.95$
Mai	$p=0.28$
Juni	$p=0.21$
Juli	$p=0.76$
August	$p=0.075$ Trend in entgegengesetzte Richtung..
September	$p=0.94$
Oktober	$p=0.91$
November	$p=0.77$
Dezember	$p=0.22$

?!

Viele unabhängige Tests

Wahrscheinlichkeit bei n unabhängigen Tests zum $\alpha=0.05$ Niveau mindestens EINEN signifikanten Subgruppeneffekt zu bekommen:

$$\Pr(\text{mindestens ein Test signifikant}) = 1 - (1 - \alpha)^n$$

n=1	0.05
n=2	0.0975
n=6	0.265
n=12	0.46
n=20	0.64
n=100	0.994

Frequentistischer Umgang mit Multiplizitätsproblem

■ Bonferroni-Adjustierung

- Ziel: Kontrolle der Wahrscheinlichkeit bei n Tests mindestens einen Fehler 1. Art zu begehen.
- $\Pr(\text{mindestens ein Test signifikant}) = 1 - (1 - \alpha)^n \doteq n \alpha$
- Einzeltests zum Niveau $\alpha' = \alpha/n \rightarrow$ Kontrolliert Gesamtfehlerwahrscheinlichkeit.

■ Zunehmend überkonservativ, wenn n groß.

■ Bonferroni-Holm Adjustierung

- etwas weniger konservatives Verfahren:
- Sortiere $p_1 < p_2 < \dots < p_n$ und
- Vergleiche p_1 mit α/n , p_2 mit $\alpha/(n-1)$, ... p_n mit α , aber stopp, sobald einmal nicht signifikant.

Multiplizitätsadjustierung ?

There are sound scientific arguments against blanket use of the Type I approach. The mere fact that a question is asked shouldn't have an effect on the answer to another question. Allowing it to have an effect is inconsistent with the likelihood principle (Berger and Wolpert, 1984; Berry, 1987) and also with my understanding of the scientific method. Roughly speaking, the likelihood principle says that the only aspect of an experiment that is relevant for inference is the observed data. The likelihood principle does not prohibit the *observations* on one variable from affecting a conclusion about a possibly related variable, even one not being addressed by the analysis, but it does mean that the *act* of analyzing those observations is irrelevant.

To see that adjustments for multiplicities may be inconsistent with the scientific method, consider two investigators, A and B. They start with the same knowledge base and design the same experiment and get *identical data*. Investigator A analyzes a large number of subgroupings, including one based on patients' left ventricular (LV) function, and suitably adjusts the significance levels because of the multiple tests. As a result of such adjustment, investigator A finds no significant interaction between treatment and LV function. Investigator B looks at only one subgrouping—LV function—and so does not have to adjust the significance level. Investigator B can claim a significant treatment effect for patients with LV dysfunction. The only difference between their studies is what they analyze after collecting the data. Drawing different conclusions from identical studies seems to me inconsistent with the scientific method.

Donald Berry (1990)

CAVE: Subgruppenanalyse

- Subgruppenanalysen – **vor allem ungeplante**
– sollten mit allergrößter Vorsicht zur Kenntnis genommen werden und erfordern in der Regel eine Bestätigung an unabhängigen Daten.
- Gründe
 - ▶ Multiplizitätsproblem (Wer viel anschaut findet auch was..)
 - ▶ In der Regel unzureichende Power zur Beurteilung von Interaktionstermen:

Standardfehler von Interaktionstermen I

Betrachte als einfachsten Fall zwei Zufallsvariable mit gleicher bekannter Varianz :

$$X_i \sim \text{Normal}(\mu_i, \sigma^2) \quad i = 1, 2$$

als Endpunkt eines randomisierten Therapievergleichs.

Gegeben zwei Stichproben des Umfangs n :

$$\text{se}^2(\hat{\mu}_1 - \hat{\mu}_2) = \frac{\sigma^2}{n} + \frac{\sigma^2}{n} = \frac{2 \cdot \sigma^2}{n}$$

Also gilt für den Standardfehler des Schätzers $(\hat{\mu}_1 - \hat{\mu}_2)$
für die Differenz $(\mu_1 - \mu_2)$ der Erwartungswerte :

$$\text{se}(\hat{\mu}_1 - \hat{\mu}_2) = \sqrt{2} \cdot \text{se}(\hat{\mu}_i)$$

Die Differenzbildung verursacht also einen Faktor $\sqrt{2}$ im Standardfehler.

Standardfehler von Interaktionstermen II

Das Patientenkollektiv werde nun in zwei komplementäre Subgruppen A und $\bar{A}$ aufgeteilt mit relativem Anteil q bzw. (1 - q).

$$\text{se}^2(\hat{\mu}_1 - \hat{\mu}_2)|_A = \frac{2 \cdot \sigma^2}{q \cdot n} ; \quad \text{se}^2(\hat{\mu}_1 - \hat{\mu}_2)|_{\bar{A}} = \frac{2 \cdot \sigma^2}{(1 - q) \cdot n}$$

FRAGE :

Unterscheiden sich die Therapie - Unterschiede zwischen den Subgruppen?

Differenz von Differenzen geschätzt mit reduzierter Fallzahl!

$$\text{se}^2((\hat{\mu}_1 - \hat{\mu}_2)|_A - (\hat{\mu}_1 - \hat{\mu}_2)|_{\bar{A}}) = \frac{2 \cdot \sigma^2}{n} \left(\frac{1}{q} + \frac{1}{1 - q} \right)$$

Standardfehler von Interaktionstermen III

q	0.5	0.3	0.2	0.1
$1/q + 1/(1-q)$	4	4.75	6.25	11.1
Faktor in se	2	2.18	2.5	3.33

Interaktionsschätzung: weiterer Faktor >2 für se
in Abhängigkeit von dem Anteil der kleineren Untergruppe.

Falls Größenordnung des Interaktionsterms $\sim$
ursprüngliche Differenz, dann nötige Fallzahl
zur Schätzung des Interaktionsterms: $n_{I=} [1/q + 1/(1-q)] * n$ (!)
wobei n die Fallzahl für einfache Differenzschätzung ist.

-> **Interaktionsschätzer meist instabil.**

Subgruppenanalyse problematisch (zusätzlich zum Multiplizitätsproblem)

Bayesianische Subgruppen-Analyse

Fall disjunkter Subgruppen

- Struktureller Prior:
 - Subgruppen hinreichend ähnlich um Zusammenfassung in einer Studie zu rechtfertigen.
 - Keine Subgruppe ausgezeichnet
 - Heterogenität der Subgruppeneffekte plausibel, aber in unklarer Größenordnung

Bayesianische Subgruppen-Analyse II

Fall disjunkter Subgruppen

r disjunkte Untergruppen.

Sparate Effektschätzer $x_j \sim N(\theta_j, se_j^2)$ $j = 1, \dots, r$
(se_j^2 sei aufgrund der Fallzahlen annähernd bekannt.)

Struktureller Prior :

Globaler Studieneffekt $\theta \sim N(0, 10^6)$ mit vagem Prior (Nicht informativ).

Untergruppeneffekte $\theta_j \sim N(\theta, 1/\tau)$ streuen um Studieneffekt mit Heterogenität $1/\tau$.

Für den Homogenitätsparameter τ vaguer Prior $\tau \sim \text{Gamma}(10^{-3}, 10^{-3})$
(Nicht informativ).

Bayesianische Subgruppen-Analyse III

Fall disjunkter Subgruppen

Likelihood :

$$\text{pr}((x_j) \mid \theta, (\theta_j), \tau) = \text{pr}((x_j) \mid (\theta_j)) \cdot \text{pr}((\theta_j) \mid \theta, \tau) \cdot \text{pr}(\theta) \cdot \text{pr}(\tau)$$

Hochdimensionale Posteriorverteilung : $\text{pr}(\theta, (\theta_j), \tau \mid (x_j))$

Marginale Posteriorverteilung für θ_j :

$$\text{pr}(\theta_j \mid (x_j)) = \int \int \int_{\theta \quad \tau \quad (\theta_i)_{i \neq j}} \text{pr}(\theta, (\theta_j), \tau \mid (x_j))$$

WINBUGS

- WINGUGS frei zum Herunterladen <http://www.mrc-bsu.cam.ac.uk/bugs>
- Allgemeines Programm zur numerischen Berechnung von Bayesianischen hierarchischen Modellen.
- Grundidee:
 - ▶ Mathematischer Trick ([Monte Carlo Markov Chains](#)) zum Samplen aus der hoch-dimensionalen Posterior-Verteilung. (Vgl. Biometrie III)
 - ▶ Dann Approximation der (marginalen) Posterior-Verteilungen durch Deskription der gesampteten Werte.

Bayesian Subgroup analysis HD6

```
model
{ for (i in 1:12) {
    prez[i]<-1/se[i]/se[i]
    x[i]~dnorm(theta[i],prez[i])
    theta[i]~dnorm(global,tau)
}
global~dnorm(0,1.0E-06)
tau~dgamma(1.0E-3,1.0E-3)
}
```

Data

```
list(x = c(-0.077,-0.368,1.090,-0.030,0.634, -0.535,0.156, -1.267, 0.066, -0.060,0.151, 0.829 ),
     se = c(0.486, 0.647, 0.507, 0.505,0.594, 0.435, 0.501, 0.760, 0.817,0.575,0.529,0.691 ))
```

WINBUGS: <http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml>

Marginal posterior estimates of parameters

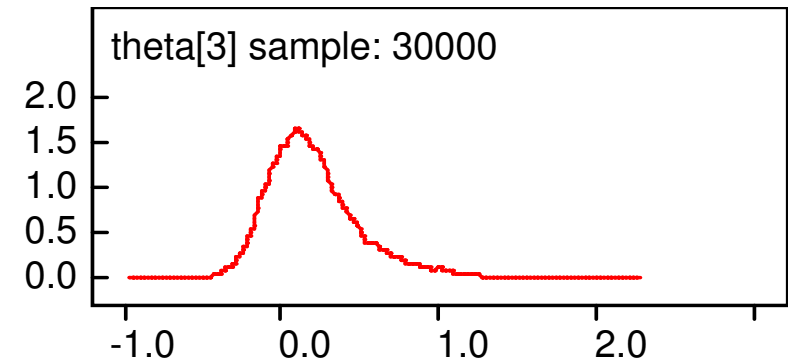
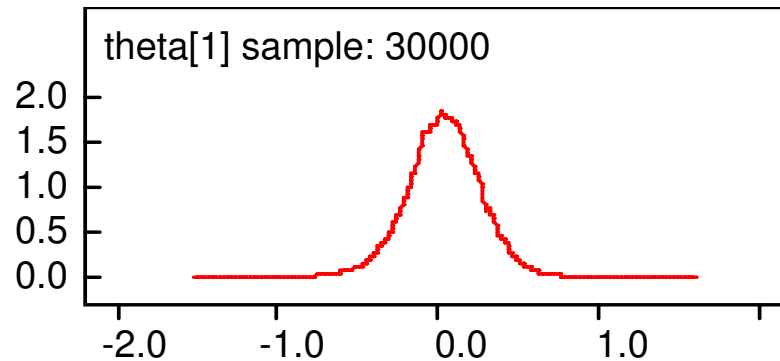
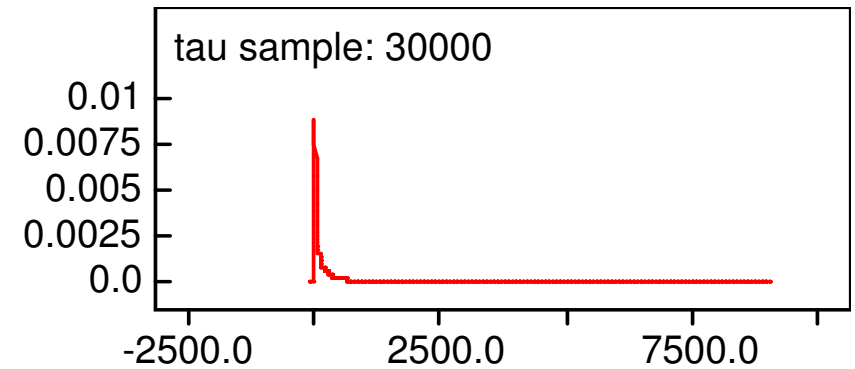
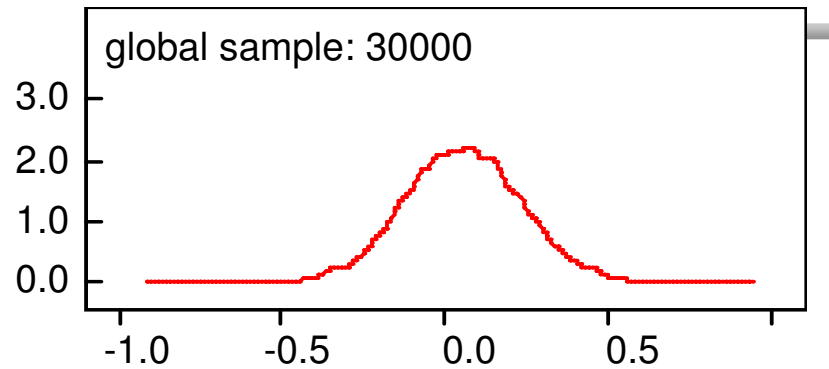
Naïve Schätzer in rot. Modell Schätzer in grün

node↓	X[l]	mean	sd	se[l]	2.5%	5.0%	median	95.0%	97.5%
global		0.05445	0.1805		-0.3017	-0.2376	0.05372	0.3511	0.413
tau		204.2	406.2		2.39	3.351	51.17	942.3	1350.0
theta[1]	-0.077	0.03135	0.245	0.486	-0.4805	-0.3703	0.03487	0.4166	0.5046
theta[2]	-0.368	0.008101	0.266	0.647	-0.5786	-0.4345	0.0196	0.413	0.5
theta[3]	1.090	0.2067	0.307	0.507	-0.2587	-0.194	0.158	0.7934	0.989
theta[4]	-0.030	0.04024	0.246	0.505	-0.4587	-0.3568	0.04021	0.4378	0.5269
theta[5]	0.634	0.1237	0.273	0.594	-0.3465	-0.2664	0.09756	0.6096	0.7531
theta[6]	-0.535	-0.05213	0.258	0.435	-0.6539	-0.5214	-0.02952	0.3243	0.3938
theta[7]	0.156	0.06891	0.245	0.501	-0.4036	-0.3125	0.06292	0.4758	0.5836
theta[8]	-1.267	-0.05831	0.313	0.760	-0.8469	-0.633	-0.01848	0.3561	0.4325
theta[9]	0.066	0.0547	0.275	0.817	-0.5023	-0.3732	0.0535	0.4889	0.6163
theta[10]	-0.060	0.03932	0.256	0.575	-0.4841	-0.3724	0.04098	0.4454	0.5425
theta[11]	0.151	0.06898	0.251	0.529	-0.4138	-0.3233	0.06196	0.4832	0.5925
theta[12]	0.829	0.129	0.287	0.691	-0.356	-0.274	0.102	0.6349	0.8113

→
Shrinking to
the global mean

←
Information
borrowing across subgroups

Marginal posterior distributions



“normale” Subgruppe

“shrunk estimate” in
Ausreissersubgruppe