Multiples Testen

-Theorie des Multiplen Testens-

Dr. Martin Scharpenberg

MSc Medical Biometry/Biostatistics

WiSe 2019/2020

Methoden zur multiplen Fehlerkontrolle

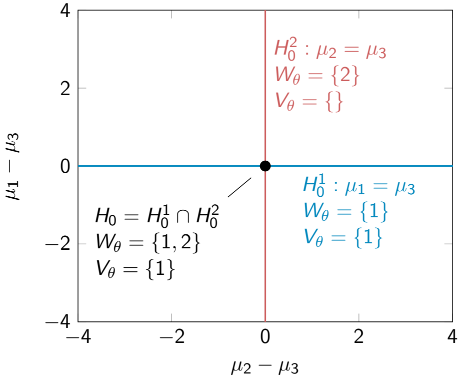
Testentscheidungsfunktion

- Ein statistischer Test kann über seine **Testentscheidungsfunktion** dargestellt werden:
$$\varphi = \begin{cases} 1 & \text{der Test verwirft } H_0 \\ 0 & \text{der Test akzeptiert } H_0 \end{cases}$$
- **Beispiele:**
Kontrasttest $\varphi_i^\alpha = \mathbf{1}_{\{T_i \geq Q^t_{N-3}(1-\alpha/2)\}}$; F-Test $\varphi_0^\alpha = \mathbf{1}_{\{F \geq Q^F_{2,N}(1-\alpha)\}}$
- Ein **multipler Test** mit den Hypothesen $H_0^1, \ldots, H_0^h$ besteht aus h Testentscheidungsfunktionen
$$\varphi = (\varphi_1, \ldots, \varphi_h)$$
wobei H_0^i verworfen wird, wenn $\varphi_i = 1$.

Wichtige Indexmengen

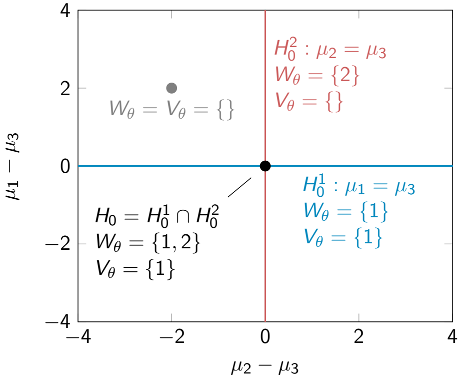
- Θ ein Parameterraum (z.B. $\theta \in \Theta = \mathbb{R}^2$, $\theta_i = \mu_i - \mu_3$, $i = 1, 2$)
- h Nullhypothesen $H_0^i \subseteq \Theta$, $i = 1, \ldots, h$. (z.B. $H_0^i = \{(\theta_1, \theta_2) : \theta_i = 0\}$)
- Wir betrachten für jedes $\theta \in \Theta$ die Indexmenge
$$W_\theta = \{i \in \{1, \ldots, h\} : \theta \in H_0^i\}$$
aller Nullhypothesen, die unter der Parameterkonstellation θ wahr sind
- Wir haben $W_\theta = \emptyset$, wenn θ in keinem H_0^i liegt
- Wir definieren zudem für jedes θ die Indexmenge
$$V_\theta = \{i \in W_\theta : H_0^i \text{ wird verworfen}\} = \{i \in W_\theta : \varphi_i = 1\}$$
der fälschlich verworfenen Nullhypothesen
- V_θ hängt von θ und den Daten ab und ist also zufällig (W_θ hängt nicht von den Daten ab)

Wichtige Indexmengen – Illustration



- Angenommen wir verwerfen H_0^1 aber nicht H_0^2, d.h. $\varphi_1 = 1$ und $\varphi_2 = 0$
- Dann ergeben sich die folgenden V_θ

Wichtige Indexmengen – Illustration



- Welche W_θ und V_θ ergeben sich für $\theta \notin H_1 \cup H_2$?
- Der graue Punkt ist ein Beispiel für $\theta \notin H_1 \cup H_2$
- Für $\theta \notin H_1 \cup H_2$ gilt $W_\theta = V_\theta = \emptyset$

Family wise error rate (FWER)

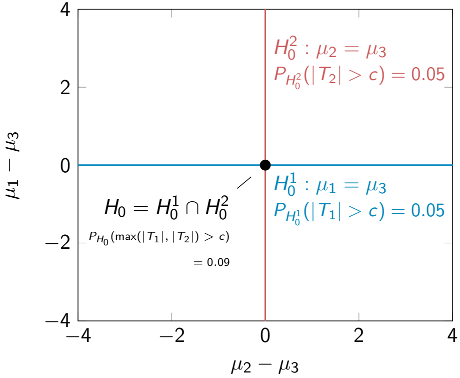
- Angenommen θ ist die wahre Parameterkonstellation
- Wir begehen genau dann **keinen** Fehler 1. Art, d.h. verwerfen **keine** wahre Nullhypothese, wenn $V_\theta = \emptyset$ bzw. $\max_{i \in W_\theta} \varphi_i = 0$
- Die *family wise error rate* ist die Wahrscheinlichkeit, mindestens eine Nullhypothese fälschlicherweise zu verwerfen:

$$\mathrm{FWER}_\theta(\varphi) = P_\theta(V_\theta \neq \emptyset) = P_\theta(|V_\theta| > 0) = P_\theta(\max_{i \in W_\theta} \varphi_i = 1)$$

  wobei $|V_\theta| = \sum_{i \in W_\theta} \varphi_i$ die Zahl der fälschlicherweise verworfenen Nullhypothesen ist
- Man spricht von *starker Kontrolle der FWER auf dem Signifkanzniveau* α, falls

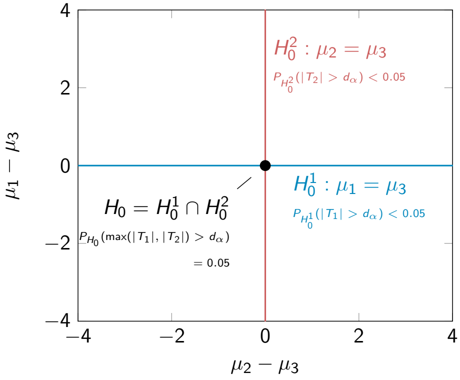
$$\sup_{\theta \in \Theta} \mathrm{FWER}_\theta(\varphi) = \sup_{\theta \in \Theta} P_\theta(|V_\theta| > 0) \leq \alpha$$

FWER mit naiven Tests



- Wenn $\theta \in H_0^1 \setminus H_0^2$ oder $\theta \in H_0^2 \setminus H_0^1$, dann ist die Fehlerrate unter Kontrolle
- Für $\theta \in H_0^1 \cap H_0^2$ ist die Fehlerrate allerdings zu groß
- Wir müssen die Testentscheidung also modifizieren

FWER beim Dunnett-Test



- Für $\alpha = 0.05$ verweden wird statt

$$c = Q^t_{19}(0.975) = 2.09$$

die kritische Grenze des Dunnett-Tests:

$$d_\alpha = 2.3649$$

- Es gilt

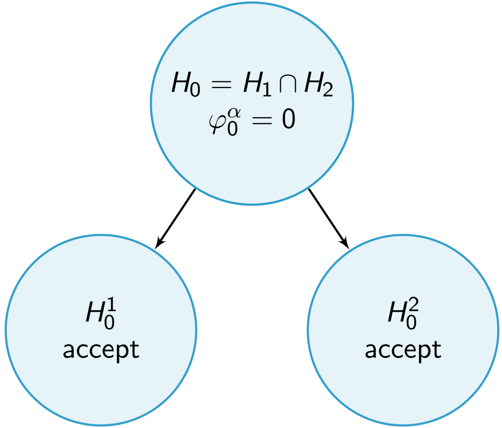
$$P_{H_i}(|T_i| \geq d_\alpha) = 0.025$$

Kontrasttests mit vorgeschalteter ANOVA

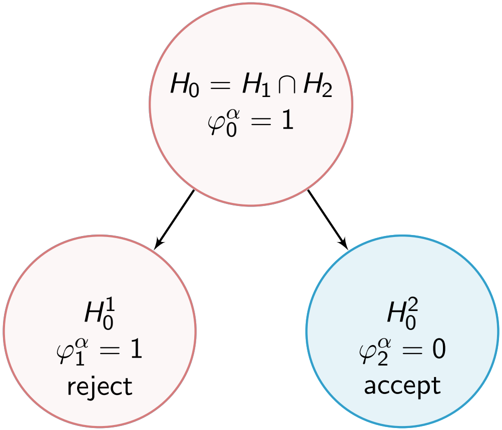
Diese Prozedur erfolgt in zwei Schritten

1. Teste $H_0 = \mu_1 = \mu_2 = \mu_3$ mit ANOVA auf Signifikanzniveau α (Testentscheidungsfunktion φ_0^α)
2. Wenn ANOVA **nicht** signifikant, dann **akzeptiere** H_0^1 und H_0^2
3. Wenn ANOVA signifikant, dann teste jedes H_0^i mit Kontrasttest (φ_i^α, $i = 1, 2$) auf Niveau α und verwerfe H_0^i, wenn auch Kontrasttest signifikant
4. Mit nur zwei Hypothesen kontrolliert diese Prozedur den multiplen Fehler 1. Art

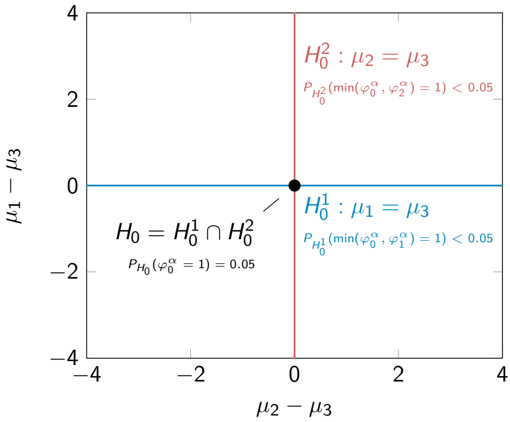
Kontrasttests mit vorgeschalteter ANOVA (φ_0^α)



Kontrasttests mit vorgeschalteter ANOVA (φ_0^α)



Fehler 1. Art bei vorgeschalteter ANOVA



Kontrasttests mit vorgeschalteter ANOVA bei 3 Hypothesen

- Vier Gruppen mit Mittelwerten $\mu_1, \mu_2, \mu_3, \mu_4$. Wir interessieren uns für die 3 Hypothesen
$$H_0^1 = \mu_1 - \mu_4, \quad H_0^2 = \mu_2 - \mu_4, \quad H_0^3 = \mu_3 - \mu_4$$
- Parameterraum: $\theta = (\theta_1, \theta_2, \theta_3) \in \Theta = \mathbb{R}^3$, $\theta_i = \mu_i - \mu_4$, $i = 1, 2, 3$
- Kann man wieder erst eine ANOVA durchführen und wenn signifikant, die drei Kontrasttests auf dem Niveau α machen?
- Falls $\theta \in H_0 = H_0^1 \cap H_0^2 \cap H_0^3$ dann gilt
$$\mathrm{FWER}_\theta = P_\theta\left(\varphi_0^\alpha = 1 \text{ und } \max_{i=1}^{3} \varphi_i^\alpha = 1\right) \leq P_\theta(\varphi_0^\alpha = 1) = \alpha$$
- Gilt $\mathrm{FWER}_\theta \leq \alpha$ für alle θ oder gibt es eine Konstellation θ, für die $\mathrm{FWER}_\theta > \alpha$ gilt?

Kontrasttests mit vorgeschalteter ANOVA bei 3 Hypothesen

- Angenommen

$$\mu_1 = \mu_2 = \mu_4 \quad \text{aber} \quad \theta_3 = \mu_3 - \mu_4 \quad \text{groß, d.h. } \theta_3 \to \infty$$

- Daraus folgt $\gamma^2 = \sum_{j=1}^{3} n_j \alpha_j^2 / \sigma^2 \to \infty$, wobei $\alpha_j = \mu_j - \mu_0$ mit $\mu_0 = \sum_j n_j \mu_j / N$ der Gesamterwartungswert. $\Rightarrow$ Power der ANOVA $\to 1$
- Wie lautet W_θ? $W_\theta = \{1, 2\}$, d.h. H_0^1 und H_0^2 sind wahr!
- Die Wahrscheinlichkeit, dass die ANOVA signifikant wird, ist praktisch 1; die ANOVA stellt keine "Hürde" mehr da!
- Wir haben aber immer noch die zwei gültigen H_0^1 und H_0^2, die wir beide - nun ungschützt - auf dem Niveau α testen
- FWER$_\theta$ ist im Extremfall derselbe wie beim ungeschützten Testen von zwei Gruppenvergleichen $> \alpha$

Schwache Kontrolle der FWER

- Kontrasttests mit vorgeschobener ANOVA kontrollieren die FWER **nicht** stark, d.h.

$$\sup_{\theta \in \Theta} P_\theta(|V_\theta| > 0) > \alpha$$

- Allerdings kontrolliert dieses Verfahren die FWER *schwach*, d.h.

$$\sup_{\theta \in H_0} P_\theta(|V_\theta| > 0) \leq \alpha$$

- *Schwache Kontrolle der FWER* bedeutet also, dass

$$\text{FWER}_\theta \leq \alpha \quad \text{für alle} \quad \theta \in H_0 = H_0^1 \cap \cdots \cap H_0^h$$

- Wie kann man mit der ANOVA **starke Kontrolle** der FWER erreichen?

Historisches

- Fisher hat das Vorschalten der ANOVA für $H_0 : \mu_1 = \cdots = \mu_k$ beim paarweisen Vergleich aller Gruppen untereinander (*all pairwise comparisons*), also beim Testen von

$$H_0^{ij} : \mu_i = \mu_j \quad \text{für alle} \quad 1 \leq i < j \leq k \quad \left(m = \tfrac{k\,(k-1)}{2} \text{ Hypothesen}\right)$$

  vorgeschlagen. Man nennt diese Vorgehensweise *Fisher's Protected Least Significant Difference Test (PLSD-Test)*.

- Peritz (1970) schlägt eine Prozedur zum Testen aller Homogenitätshypothesen

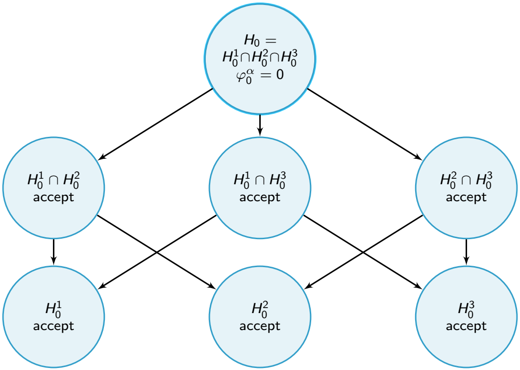
$$H_0^{i_1,\ldots,i_l} : \mu_{i_1} = \cdots = \mu_{i_l}, \qquad \{i_1, \ldots, i_l\} \subseteq \{1, \ldots, k\}$$

  vor. Marcus, Peritz und Gabriel (1976) verallgemeinern diese Prozedur zum sogenannten Abschlusstestprinzip (*closure principle* oder *closure method* oder *closed testing procedure*) auf beliebige Hypothesen.

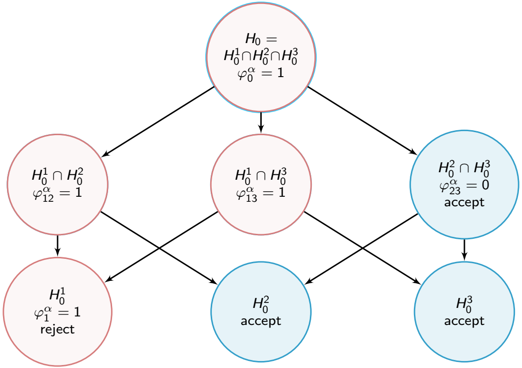
Das Abschlusstestprinzip

Abschlusstestprinzip für drei Hypothesen mit ANOVAs

(φ_0^α ANOVA für $H_0 : \mu_1 = \cdots = \mu_4$, φ_{ij}^α ANOVA für $H_0^i \cap H_0^j : \mu_i = \mu_j = \mu_4$)



Abschlusstestprinzip für drei Hypothesen mit ANOVAs
(φ_0^α ANOVA für $H_0 : \mu_1 = \cdots = \mu_4$, φ_{ij}^α ANOVA für $H_0^i \cap H_0^j : \mu_i = \mu_j = \mu_4$)



Das Abschlusstestprinzip (Abschlusstests)

- Wir haben h Hypothesen $H_0^i \subseteq \Theta$, $i = 1, \ldots, h$.
- Wir betrachten alle Schnitthypothesen

$$H_0^J = \cap_{j \in J} H_0^j, \qquad J \subseteq \{1, \ldots, h\}$$

- Wir legen für jedes H_0^J einen Niveau-α-Test φ_J^α fest (vorhin waren das ANOVA's), d.h., es muss gelten

$$P_\theta(\varphi_J^\alpha = 1) \leq \alpha \quad \text{für alle} \quad \theta \in H_0^J$$

- Wir verwerfen die Hypothese H_0^i, falls $\varphi_J = 1$ für alle $J \subseteq \{1, \ldots, h\}$ mit $i \in J$, d.h.,

$$\phi_i^\alpha = 1 \quad \text{wobei} \quad \phi_i^\alpha = \min_{i \in J \subseteq \{1, \ldots, h\}} \varphi_J^\alpha$$

- Prinzip benötigt i.A. Festlegung und Durchführung von $2^h - 1$ Tests (Zahl der nicht-leeren Teilmengen von $\{1, \ldots, h\}$)!

Fehlerkontrolle mit dem Abschlusstestprinzip

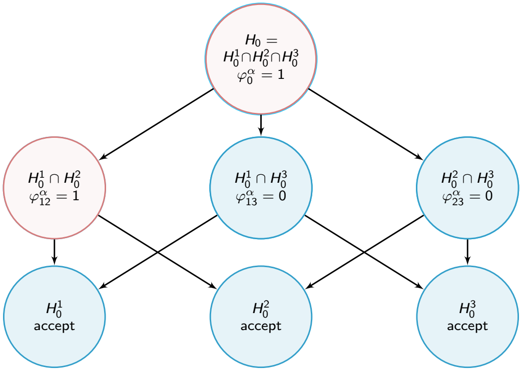
Für jedes $\theta \in \Theta$ mit $W_\theta \neq \emptyset$ gilt

$$\begin{aligned}
\mathrm{FWER}_\theta(\phi^\alpha) &= P_\theta\left(\max_{i \in W_\theta} \phi_i^\alpha = 1\right) \\
&= P_\theta\left(\max_{i \in W_\theta} \min_{i \in J \subseteq \{1,\ldots,h\}} \varphi_J^\alpha = 1\right) \\
&\leq P_\theta\left(\varphi_{W_\theta}^\alpha = 1\right) \\
&\leq \alpha
\end{aligned}$$

Also ist $\sup_\theta \mathrm{FWER}_\theta(\phi^\alpha) \leq \alpha$, d.h. familienweiser Fehler 1. Art unter Kontrolle.

Abschlusstestprinzip für 3 Hypothesen - Beispiel

(φ_0^α Niveau-α test für $H_0 = H_0^1 \cap H_0^2 \cap H_0^3$, φ_{ij}^α Niveau-α Tests für $H_0^i \cap H_0^j$)



Abschluss einer Hypothesenfamilie

- Beim Abschlusstest betrachten wir nicht nur die Familie der elementaren Hypothesen $H_0^1, \dots, H_0^h$ sondern dessen *Abschluss (closure)*

$$\mathcal{C}(H_0^1, \dots, H_0^h) = \{H_0^J = \cap_{j \in J} H_0^j : J \subseteq \{1, \dots, h\}\} \setminus \{\emptyset\}$$

Definition - $\cap$-Abgeschlossenheit

Eine Menge $\mathcal{H}$ von Hypothesen heißt $\cap$-*abgeschlossen*, falls mit $H, H' \in \mathcal{H}$ auch $H \cap H' \in \mathcal{H}$ oder $H \cap H' = \emptyset$ gilt

- $\mathcal{C}(H_0^1, \dots, H_0^h)$ ist die kleinste $\cap$-*abgeschlossene* Familie von Hypothesen (Teilmengen von Θ), die $H_0^1, \dots, H_0^h$ enthalten

Abschluss einer Hypothesenfamilie - Beispiele

- (a) $\mathcal{C}(H_0^1, \ldots, H_0^h)$.
- (b) Zwei Gruppen (E und C) mit Erwartungswerten μ_E und μ_C. Wir fixieren den Nicht-Unterlegenheits- margin $\delta > 0$. Dann ist die Familie

$$H_0^1 : \mu_E - \mu_C \leq -\delta \quad \text{und} \quad H_0^2 : \mu_E - \mu_C \leq 0$$

$\cap$-abgeschlossen, denn $H_0^1 \cap H_0^2 = H_0^1$.

Kohärenz

Definition - Kohärenz (Gabriel, 1969)

Ein multipler Test $\varphi_1, \ldots, \varphi_h$ für die Hypothesen $H_0^1, \ldots, H_0^h \subseteq \Theta$ heißt *kohärent*, falls

$$\text{für alle } 1 \leq i, j \leq h \text{ mit } H_0^i \subseteq H_0^j \text{ gilt:} \quad \varphi_j = 1 \Rightarrow \varphi_i = 1$$

- *Bemerkungen:*
  - $\varphi_j = 1 \Rightarrow \varphi_i = 1$ genau dann, wenn immer $\varphi_j \leq \varphi_i$
  - Ein nicht-kohärenter multipler Test kann zu logisch inkonsistenten Verwerfungen führen. Daher sollten nur kohärente Tests verwendet werden
  - Jeder Abschlusstest ist kohärent

Kohärenz - Beispiel 1

- Wir betrachten wieder die Hypothesen
$$H_0^1 : \mu_E - \mu_C \leq -\delta \quad \text{und} \quad H_0^2 : \mu_E - \mu_C \leq 0$$
(nun) für einen normalverteilten Endpunkt $Y_j \sim N(\mu_j, \sigma^2)$, $j \in \{E, C\}$
- Fallzahlen: $N = n_E + n_C$; krit. Grenze des t-Tests: $c_\alpha = t_{N-2}(1-\alpha)$
- **Beispiel 1:** Die Hypothesen H_0^1 und H_0^2 werden in der selben Population (z.B. PP) getestet:
$$\varphi_1^\alpha = \begin{cases} 1 & \frac{\bar{Y}_E - \bar{Y}_C + \delta}{\hat{\sigma}\sqrt{n_E^{-1} + n_C^{-1}}} \geq c_\alpha \\ 0 & \text{sonst} \end{cases}, \quad \varphi_2^\alpha = \begin{cases} 1 & \frac{\bar{Y}_E - \bar{Y}_C}{\hat{\sigma}\sqrt{n_E^{-1} + n_C^{-1}}} \geq c_\alpha \\ 0 & \text{sonst} \end{cases}$$
- Da $H_0^1 \subseteq H_0^2$ und $\varphi_1^\alpha \geq \varphi_2^\alpha$ ist der multiple Test $\varphi^\alpha = (\varphi_1^\alpha, \varphi_2^\alpha)$ kohärent.

Kohärenz - Beispiel 2

- Wir betrachten wieder die Hypothesen

$$H_0^1 : \mu_E - \mu_C \leq -\delta \quad \text{und} \quad H_0^2 : \mu_E - \mu_C \leq 0$$

  für einen normalverteilten Endpunkt $Y_j \sim N(\mu_j, \sigma^2)$, $j \in \{E, C\}$
- Fallzahlen: $N = n_E + n_C$; krit. Grenze des t-Tests: $c_\alpha = t_{N-2}(1-\alpha)$
- **Beispiel 2:** Nun werden die Hypothesen H_0^1 und H_0^2 in **verschiedenen** Populationen getestet (z.B. FAS und PP)
- Dann ist $\varphi_1^\alpha < \varphi_2^\alpha$ möglich und die t-Tests (unadjustiert) liefern **keinen** kohärenten Test
- Allerdings, ist $(\tilde{\varphi}_1, \tilde{\varphi}_2)$ mit $\tilde{\varphi}_1 = \varphi_1^\alpha$ und $\tilde{\varphi}_2 = \min(\varphi_1^\alpha, \varphi_2^\alpha)$ kohärent
- Bei diesem Test verwerfen wir H_0^2 nur, wenn H_0^1 verworfen wurde. (Hierarchischer Test zur Sequenz $H_0^1 \to H_0^2$)

Fehlerkontrolle bei kohärenten Tests

Definition - lokales Niveau

$\varphi = (\varphi_1, \ldots, \varphi_h)$ ein multipler Test für $H_0^1, \ldots, H_0^h \subseteq \Theta$. Wenn für alle $i = 1, \ldots, h$

$$\sup_{\theta \in H_0^i} P_\theta(\varphi_i = 1) \leq \alpha$$

dann hat der multipler Test *das lokale Niveau* α

Beispiel: Ungeschützte Kontrasttests auf dem Niveau α

Fehlerkontrolle bei kohärenten Tests

Satz - Starke Kontrolle der FWER

Es sei $\varphi = (\varphi_1, \ldots, \varphi_h)$ ein multipler Test für $H_0^1, \ldots, H_0^h$ zum lokalen Niveau α. Ist die Familie $\mathcal{H} = \{H_0^1, \ldots, H_0^h\}$ $\cap$-abgeschlossen und φ kohärent, dann ist

$$\sup_{\theta \in \Theta} \mathrm{FWER}_\theta(\varphi) \leq \alpha \ .$$

Der Beweis ist fast derselbe wie beim Abschlusstestprinzip (Übung).

Beispiele: Aus dem Satz folgt, dass der familienweise Fehler I. Art in Beispiel 1 mit φ und Beispiel 2 mit $\tilde{\varphi}$ kontrolliert ist.

Kohärenz bei Fehlerkontrolle

Satz - Verbesserung durch kohärente Tests

Jeder multiple Test $\varphi = (\varphi_1, \dots, \varphi_h)$ für $H_0^1, \dots, H_0^h \subseteq \Theta$ mit starker Kontrolle der FWER kann durch einen kohärenten multiplen Test $\tilde{\varphi} = (\tilde{\varphi}_1, \dots, \tilde{\varphi}_h)$ mit starker Kontrolle der FWER verbessert werden, d.h. für diesen Test gilt immer (für jede Stichprobe)

$$\tilde{\varphi}_i \geq \varphi_i \quad \text{für alle} \quad i = 1, \dots, h \ .$$

Beweis. Definiere für alle $i = 1, \dots, h$ den Test

$$\tilde{\varphi}_i = \max_{j : H_0^j \supseteq H_0^i} \varphi_j \ .$$

Klarerweise gilt immer $\tilde{\varphi}_i \geq \varphi_i$. Aus starker Kontrolle der FWER folgt für alle θ:

$$P_\theta(\max_{i \in W_\theta} \tilde{\varphi}_i = 1) = P_\theta(\max_{i \in W_\theta} \max_{j : H_0^j \supseteq H_0^i} \varphi_j = 1) = P_\theta(\max_{i \in W_\theta} \varphi_i = 1) \leq \alpha \ ,$$

da aus $i \in W_\theta = \{j : \theta \in H_0^j\}$ und $H_0^j \supseteq H_0^i$ auch $j \in W_\theta$ folgt. Kohärenz folgt aus der Definition von $\tilde{\varphi}$.

Erweiterung auf Abschluss

Satz - Erweiterung kohärenter Tests

Es sei $\varphi = (\varphi_1, \dots, \varphi_h)$ ein kohärenter multipler Test für $H_0^1, \dots, H_0^h \subseteq \Theta$ mit starker Kontrolle der FWER. Dann kann φ zu einem kohärenten multiplen Test auf dem Abschluss

$$\mathcal{H} = \mathcal{C}(H_0^1, \dots, H_0^h)$$

mit starker Kontrolle der FWER ausgedehnt werden.

Beweis. Definiere für H_0^J, $J \subseteq \{1, \dots, h\}$, den Test $\varphi_J = \max_{i \in J} \varphi_i$. Aus der starken Kontrolle der FWER folgt für alle $\theta \in H_0^J$

$$P_\theta(\varphi_J = 1) \leq P_\theta(\max_{i \in W_\theta} \varphi_i = 1) \leq \alpha \qquad (\text{da } \theta \in H_0^J \iff J \subseteq W_\theta)$$

D.h. die Erweiterung von φ auf $\mathcal{H}$ hat lokales Niveau α. Kohärenz folgt aus der Definition der Erweiterung. Aus dem Satz der vorletzten Folie folgt starke Kontrolle der FWER.

Folgerung aus den letzten zwei Sätzen

- Kohärenz ist nicht nur eine Frage der Logik sondern auch eine Frage der Effizienz
- Es macht überhaupt keinen Sinn nicht-kohärente Tests mit starker Kontrolle der FWER zu betrachten
- Nach dem letzten Satz können wir uns zudem auf kohärente multiple Tests für den Abschluss

$$\mathcal{C}(H_0^1, \ldots, H_0^h) = \{H_0^J = \cap_{j \in J} H_0^j : J \subseteq \{1, \ldots, h\}\}$$

mit lokalem Niveau α beschränken. Das sind, in einem gewissen Sinne, gerade die Abschlusstests

Kohärenz von Abschlusstests

- Wir können einen Abschlusstest $(\varphi_J^\alpha)_{J\subseteq\{1,\ldots,h\}}$ für $H_0^1,\ldots,H_0^h$ immer auf den Abschluss
$$\mathcal{C}(H_0^1,\ldots,H_0^h)$$
erweitern, indem wir $H_0^J=\cap_{j\in J}H_0^j$, $J\subseteq\{1,\ldots,h\}$, verwerfen, falls $\phi_J^\alpha=1$ mit
$$\phi_J^\alpha=\min_{J\subseteq J'\subseteq\{1,\ldots,h\}}\varphi_{J'}^\alpha\ .$$
- **Bemerkungen:**
  - Dieser Test auf $\mathcal{C}(H_0^1,\ldots,H_0^h)$ ist per Definition kohärent
  - Falls schon $(\varphi_J^\alpha)_{J\subseteq\{1,\ldots,h\}}$ köhärent ist, dann gilt
  $$\phi_J^\alpha=\varphi_J^\alpha \text{ für alle } J\subseteq\{1,\ldots,h\}$$
  Kohärente und Abschlusstests sind also identisch
  - Wir schreiben später auch $\varphi_{H_0^J}$ und $\phi_{H_0^J}$ statt φ_J und ϕ_J

Konsonanz (Gabriel, 1969; Brannath & Bretz, 2010)

Definition - Konsonanz

Gegeben seien die Hypothesen

$$H_0^1, \ldots, H_0^h \ .$$

Ein multipler Test ϕ auf dem Abschuss $\mathcal{H} = \mathcal{C}(H_0^1, \ldots, H_0^h)$ heißt *konsonant*, falls für alle $H \in \mathcal{H}$

$$\{\phi_H = 1\} = \cup_{H_0^i \supseteq H}\{\phi_i = 1\} \ ,$$

d.h. mit jedem $H \in \mathcal{H}$ wird auch mindestens ein $H_0^i \supseteq H$ verworfen.

Bemerkungen zur Konsonanz

- Diese Definition folgt eher Brannath & Bretz (2010). Sie ist intuitiver und schwächer als die in Gabriel (1969). Die Definitionen stimmen jedoch in vielen Fällen überein
- Konsonanz ist eine wünschenswerte aber nicht zwingend notwendige Eigenschaft
- Nicht jeder Abschlusstest ist konsonant. Ein Beispiel ist der Abschlusstest mit ANOVA's
- *Wir werden später sehen:* Konsonanz erlaubt Abkürzungen im Algorithmus des Abschlusstestprinzips und vereinfacht die Durchführung von Abschlusstests wesentlich.