Multiples Testen

-Simultane Testprozeduren-

Dr. Martin Scharpenberg

MSc Medical Biometry/Biostatistics

WiSe 2019/2020

Simultane Testprozeduren (STP)

- Wollen die h Nullhypothesen $H_0^1, \dots, H_0^h$ testen
- Benutze für jedes H_0^j eine Teststatistik T_j, $j = 1, \dots, h$
- Bestimmen d_α, so dass für $H_0 = H_0^1 \cap \cdots \cap H_0^h$

$$P_{H_0}(\max_{l=1}^{h} T_l \geq d_\alpha) = \alpha$$

- Verwerfe H_0^j, falls $T_j \geq d_\alpha$
- Nach Konstruktion von d_α haben wir schwache Kontrolle der FWER
- Haben wir auch starke Kontrolle, bzw. welche Bedingungen müssen dafür erfüllt sein?

Starke Kontrolle mit STP

- Für $J \subseteq J_0 = \{1, \dots, h\}$ sei wieder $H_0^J = \cap_{l \in J} H_0^l$
- Starke Kontrolle gilt, wenn

$$P_{H_0^J}(\max_{l \in J} T_l \geq d_\alpha) \leq \alpha \quad \text{für jedes } J \subseteq J_0 = \{1, \dots, h\}$$

Satz - Starke Kontrolle mit STP

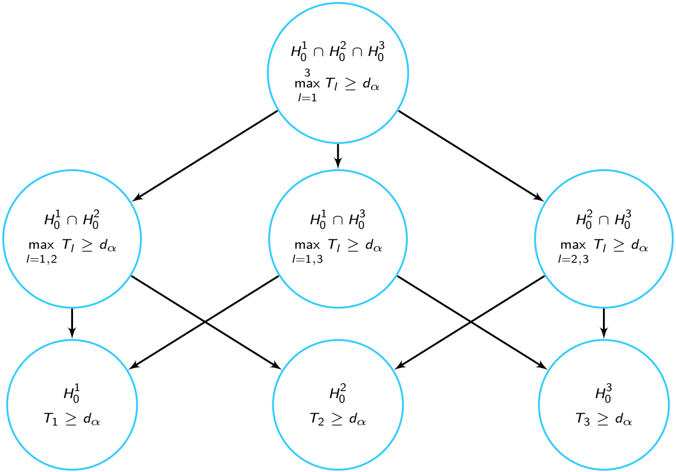
Für die eben eingeführte STP haben wir auch **starke** Kontrolle der FWER, falls

$$P_{H_0^J}(\max_{l \in J} T_l \geq d_\alpha) \leq \alpha \quad \forall J \subseteq J_0,$$

d.h. falls die Verwerfungsregel $\max_{l \in J} T_l \geq d_\alpha$ einen lokalen Niveau-α Test für alle $J \subseteq J_0$ induziert.

- *Beweis:* STP ist Abschlusstest mit $\varphi_J^\alpha = \mathbf{1}(\max_{l \in J} T_l \geq d_\alpha)$

STP als Abschlusstest (mit 3 Hypothesen)



Starke Kontrolle mit STP

- Wir haben benötigt, dass die Verwerfungsregel $\max_{I \in J} T_I \geq d_\alpha$ für H_0^J einen lokalen Niveau-α Test für alle $J \subseteq J_0$ induziert
- Definieren für $J \subseteq J_0$ die Schranke d_α^J durch

$$P_{H_0^J}(\max_{I \in J} T_I \geq d_\alpha^J) = \alpha$$

- Dann ist die Bedingung eines lokalen Niveau-α Tests von oben sicher erfüllt, falls

$$d_\alpha^J \leq d_\alpha \qquad \text{für jedes } J \subseteq J_0 = \{1, \dots h\}$$

- In diesem Fall kontrolliert die STP die FWER nach dem Satz also stark
- Werden im Folgenden ein einfaches Kriterium dafür betrachten

Einfaches Kriterium

- Angenommen $d_\alpha^J = d_{k,\alpha}$ hängt nur von $k = |J|$ ab ($|J|$ = Zahl der Hypothesen in J)
- Können dann für jedes $k \leq h$ die Grenze $d_{k,\alpha}$ durch

$$P_{H_0^1 \cap \cdots \cap H_0^k}(\max_{l=1}^{k} T_l \geq d_{k,\alpha}) = \alpha$$

  bestimmen
- Da $\max_{l=1}^{k'} T_l \leq \max_{l=1}^{k} T_l$ für $k' \leq k$, gilt:

$$d_{1,\alpha} \;\leq\; d_{2,\alpha} \;\leq\; \cdots \;\leq\; d_{h,\alpha} = d_\alpha$$

- Voraussetzungen für starke Kontrolle der FWER sind damit erfüllt!

Der Šidák-Test als STP

- Seien $T_1, \ldots, T_h$ stochastische unabhängige Teststatistiken für die Hypothesen $H_0^1, \ldots, H_0^h$
- Angenommen, alle T_l haben dieselbe stetige Null-Verteilungsfunktion F_0
- Dann hängt die gemeinsame Null-Verteilung von $\{T_j : j \in J\}$ nur von $k = |J|$ ab
- Für jede kritische Grenze d gilt dann

$$P_{H_0^1 \cap \cdots \cap H_0^k}(\max_{l=1}^{k} T_l \geq d) = 1 - \prod_{l=1}^{k} P_{H_0^l}(T_l < d) = 1 - F_0(d)^k$$

- Und damit

$$d_{k,\alpha} = F_0^{-1}\left(\sqrt[k]{1-\alpha}\right) \leq d_\alpha = F_0^{-1}\left(\sqrt[h]{1-\alpha}\right)$$

Der Šidák-Test als STP

- Man nennt den Test mit $\varphi_I^\alpha = \mathbf{1}(T_I \geq d_\alpha)$ *Šidák-Test*
- Offensichtlich ist nun

$$d_{1,\alpha} \;\leq\; d_{2,\alpha} \;\leq\; \cdots \;\leq\; d_{h,\alpha} = d_\alpha$$

- Der Šidák-Test kontrolliert die FWER also stark
- **Beachte:** Dies gilt nur, wenn die Teststatistiken stochastisch unabhängig sind und dieselbe stetige Null-Verteilung haben!

Der Šidák-Test über p-Werte

- Seien $p_1, \dots, p_h$ stoch. unabhängig, unter Nullhyp. gleichverteilt (meist definiert als: $p_l = 1 - F_0(T_l)$)
- Gemeinsame Null-Verteilung von $\{p_j : j \in J\}$ hängt nur von $k = |J|$ ab
- Für jedes lokale Niveau α' (verwerfe H_0^l falls $p_l \leq \alpha'$) gilt dann:

$$P_{H_0^1 \cap \dots \cap H_0^k}(\min_{l=1}^{k} p_l \leq \alpha') = 1 - \prod_{l=1}^{k} P_{H_0^l}(p_l > \alpha') = 1 - (1 - \alpha')^k$$

- Wir kontrollieren die FWER stark auf den Niveau α, falls

$$\alpha' = 1 - \sqrt[h]{1 - \alpha}$$

Many-to-One Vergleich (1)

- Angenommen wir wollen wieder die Erwartungswerte in k Gruppen vergleichen
- Angenommen die Gruppen sind gleich groß: $n_1 = \cdots = n_k = n$
- Gruppe k ist Kontrollgruppe. Testen $H_0^j : \mu_j - \mu_k = 0$ mit der Teststatistik

$$T_j = \left(\bar{Y}_j - \bar{Y}_k\right) / \left(\hat{\sigma}\sqrt{n_j^{-1} + n_k^{-1}}\right) = \left(\bar{Y}_j - \bar{Y}_k\right) / \left(\hat{\sigma}\sqrt{2 \cdot n_k^{-1}}\right)$$

- Man kann zeigen, dass unter diesen Bedingungen und falls H_0^i und H_0^j gelten,

$$\mathrm{Corr}(\bar{Y}_i - \bar{Y}_k, \bar{Y}_j - \bar{Y}_k) = 1/2$$

  gilt
- Für jedes $J \subseteq \{1, \ldots, k-1\}$ hängt die gemeinsame H_0^J-Verteilung der T_j, $j \in J$, nur von $|J|$ ab

Many-to-One Vergleich (2)

- Wissen: Dunnett-Test kontrolliert die FWER stark, wenn alle Gruppen gleich groß sind
- Starke Kontrolle der FWER bei ungleich großen Gruppen?
- **Problem bei ungleich großen Gruppen:** Gemeinsame Verteilung von $\{T_j : j \in J\}$ hängt von mehr als nur $|J|$ ab, denn:
- die Korrelation hängt von den beteiligten Gruppengrößen ab:

$$\rho_{ij} = \mathrm{Corr}(\bar{Y}_i - \bar{Y}_k, \bar{Y}_j - \bar{Y}_k) = 1/\sqrt{(1+\frac{n_k}{n_i})(1+\frac{n_k}{n_j})},\ 1 \leq i, j \leq k-1$$

Teilmengen-Pivotalität

Definition - Teilmengen-Pivotalität

Die Bedinung der *Teilmengen-Pivotalität* gilt, falls für alle $J \subseteq J_0$ die Verteilung von

$$\{T_j : j \in J\}$$

nur abhängig von θ_j für $j \in J$ und unabhängig von θ_i für $i \notin J$ ist

- *In Worten:* Die gemeinsame Verteilung von $\{T_j : j \in J\}$ wird nicht von Parametern der anderen Nullhypothesen beeinflusst
- **Beispiel:** *Familie der Paarvergleiche T_{ij} bei ungleich großen Gruppen:* Die gemeinsame Verteilung einer Menge J von T_{ij}'s hängt nur von den Fallzahlen n_j und Erwartungswerten μ_j der beteiligten Gruppen und der gemeinsamen Varianz σ^2 ab

Fehlerkontrolle bei Teilmengen-Pivotalität

- Unter Teilmengen-Pivotalität gilt für jedes $J \subseteq J_0$

$$P_{H_0^J}(\max_{l \in J} T_l \geq d_\alpha) = P_{H_0}(\max_{l \in J} T_l \geq d_\alpha)$$

- und damit

$$P_{H_0^J}(\max_{l \in J} T_l \geq d_\alpha) \quad \leq \quad P_{H_0}(\max_{l=1}^{h} T_l \geq d_\alpha) = \alpha$$

- Damit haben wir starke Kontrolle der FWER!

Schlussfolgerungen

- Dunnett- und Tukey-Test kontrollieren FWER stark auch bei ungleich großen Gruppen
- Können mit STP-Prinzip bel. Ausgangsmengen von Paarvergleichen bei beliebigen Fallzahlen testen
- Teilmengen-Pivotalität gilt auch für Kontraste
- D.h. wir können mit STP-Prinzip eine beliebige endliche Ausgangsmenge von Kontrasten testen