Bootstrap-Methoden zur Ermittlung kritischer Werte für asymptotische FWER-Kontrolle

[Dudoit, van der Laan, Pollard: Multiple Testing. Part I Single-Step Procedures for Control of General Type-I-Error Rates]

Mathias Trabs

14.11.2010

1. Wiederholung
2. Problemstellung
3. Fehlerkontrolle und Wahl der Nullverteilung
4. Umsetzung

Wiederholung

Sei $(\Omega, \mathcal{A}, \mathcal{M}, \mathcal{H})$ ein **multiples Testproblem** mit $P \in \mathcal{M}$, $\mathcal{M}$ eine Familie von Wahrscheinlichkeitsmaßen auf $(\Omega, \mathcal{A})$, und einer Hypothesenmenge $\mathcal{H} = \{H_i : i \in I = \{1, ..., m\}\}$. Sei weiter $\varphi = (\varphi_i : i \in I)$ ein **multipler Test**.
Die (zufälligen) Anzahlen von wahren / falschen Testentscheidungen können wir darstellen als:

	Testentscheidung		
Hypothesen	0	1	
wahr	$m_0 - V(P)$	$V(P)$	m_0
falsch	$m_1 - S(P)$	$S(P)$	m_1
	$m - R(P)$	$R(P)$	m

Def.: $FWER(P) = P(V(P) > 0) = P(\bigcup_{i \in I_0} \{\varphi_i = 1\})$

Modell

Seien $X_1, ...X_n$ iid. Zufallsgrößen im $\mathbb{R}^J$,

$$X_i = (X_{ij} : j = 1, ..., J) \sim P \text{ mit } P \in \mathcal{M} \text{ unbekannt.}$$

$\Rightarrow$ Für $i \in \{1, ..., n\}$ sind $(X_{ij})_{j=1,...,J}$ **unspezifiziert korreliert**.

Wir möchten z.B. **Lokationsparameter** der Form

$$\psi(P) = (\psi_i : i = 1, ...m)$$

untersuchen.

Bsp.: Sei $X \sim P$ mit Werten in $\mathbb{R}^J$ und $Y := g(X) : \mathbb{R}^J \to \mathbb{R}^m$. Dann wählen wir

$$\psi(P) = \mathbb{E}[Y], \text{ d.h. } \psi_i = \mathbb{E}[Y_i].$$

Teststatistiken

Wir haben Statistiken $T_n = (T_{ni} : i = 1, ..., m) \in \mathbb{R}^m$ als Funktionen von $X_1, ..., X_n$ und bezeichnen deren wahre Verteilung mit $Q_n = Q_n(P)$.
Testentscheidung:

- H_i annehmen, falls $T_{ni} \leq c_i$,
- H_i ablehnen, falls $T_{ni} > c_i$,

mit den kritischen Werten $c \in \mathbb{R}^m$. Eine **multiple Testprozedur (MTP)** ist dann die (zufällige) Teilmenge $\mathcal{R}_n \subseteq I$ der abgelehnten Hypothesen. Gilt $c_i = c$ für $i = 1, ..., m$, heißt $\mathcal{R}_n$ Simultantest.

Bsp.: Y und $\psi(P)$ wie oben. Hypothesen:
$H_i = \{\psi_i(P) = \mathbb{E}[Y_i] \leq \psi_{oi}\}, i = 1, ..., m$ mit einem Nullwert $\psi_0 \in \mathbb{R}^m$.
Dann wählen wir die t-Statistiken:

$$T_{ni} = \frac{\text{Schätzer - Nullwert}}{\text{Standardfehler}} = \sqrt{n}\frac{\psi_{ni} - \psi_{0i}}{\sigma_{ni}}$$

Typ-I-Fehlermaße

Die Theorie baut auf **Fehlermaßen** $\Theta(F_{V_n}) \in [0,1]$ auf, die als Funktionen von der Verteilung der Anzahl der Typ-I-Fehler V_n definiert sind. Dabei ist F_{V_n} die Verteilungsfunktion von V_n auf $\{0, ..., m\}$.

Insbesondere betrachten wir die **FWER**:

$$\Theta(F_{V_n}) = FWER(P) = P(V_n > 0) = 1 - F_{V_n}(0).$$

Annahmen an das Fehlermaß Θ

Seien F_1, F_2 zwei Verteilungsfunktionen auf $\{0, ..., m\}$ und
$d(F_1, F_2) := \max_{x \in I} |F_1(x) - F_2(x)|$ deren Abstand.

Wir machen folgende Annahmen an Θ:

- **(AMI)** Monotonie:

$$F_1 \geq F_2 \Rightarrow \Theta(F_1) \leq \Theta(F_2)$$

- **(ACI)** Stetigkeit bei (F_n): Sei (F_n) eine Folge von Verteilungsfunktionen auf $\{0, ..., m\}$ gegeben, dann soll für beliebige Verteilungsfunktionen (G_n) auf $\{0, ..., m\}$ gelten:

$$\lim_{n \to \infty} d(F_n, G_n) = 0 \Rightarrow \lim_{n \to \infty} (\Theta(G_n) - \Theta(F_n)) = 0$$

In den meisten Fällen genügt in der (ACI)-Annahme $(F_n) = F$, für eine Verteilungsfunktion F.

Kontrolle des Typ-I-Fehlermaßes

Definition:
Eine MTP $\mathcal{R}_n = \mathcal{R}(T_n, Q_0, \alpha)$ **kontrolliert das Niveau** $\alpha \in (0,1)$ **(strikt)**, falls

$$\Theta(F_{V_n}) \leq \alpha, \qquad (FWER(P) \leq \alpha).$$

$\mathcal{R}_n$ **kontrolliert das Niveau** $\alpha \in (0,1)$ **asymptotisch**, falls

$$\limsup_{n \to \infty} \Theta(F_{V_n}) \leq \alpha.$$

V_n hängt von der wahren Verteilung $Q_n = Q_n(P)$ der Teststatistiken T_n ab, aber Q_n ist i.A. unbekannt und muss durch ein **Nullverteilung** Q_0 geschätzt werden (um kritische Werte zu ermitteln).

Seien Teststatistiken T_n mit wahrer Verteilung Q_n, einer m-dimensionalen Nullverteilung Q_0 zur Berechnung kritischer Werte, sowie eine Niveau α gegeben.

Für die gesamte Anzahl der abgelehnten Hypothesen R und die Anzahl der abgelehnten wahren Hypothesen V schreiben wir:

$$
\begin{aligned}
R_n &= R(Q_0|Q_n) = |\mathcal{R}(T_n, Q_0, \alpha)|, && T_n \sim Q_n, \\
R_0 &= R(Q_0|Q_0) = |\mathcal{R}(T_n, Q_0, \alpha)|, && T_n \sim Q_0, \\
V_n &= V(Q_0|Q_n) = |\mathcal{R}(T_n, Q_0, \alpha) \cap I_0|, && T_n \sim Q_n, \\
V_0 &= V(Q_0|Q_0) = |\mathcal{R}(T_n, Q_0, \alpha) \cap I_0|, && T_n \sim Q_0.
\end{aligned}
$$

Road map

1. Null-Dominiertheit für das Typ-I-Fehlermaß $\Theta(F_{V_n})$:
   Wähle eine Null-Verteilung Q_0 so, dass.

$$\begin{aligned} \Theta(F_{V_n}) &\leq \Theta(F_{V_0}) && \text{[strikte Kontrolle]} \\ \limsup_{n\to\infty} \Theta(F_{V_n}) &\leq \Theta(F_{V_0}) && \text{[asymptotische Kontrolle].} \end{aligned} \tag{1}$$

2. Die Anzahl der Typ-I-Fehler ist nie größer als die gesamte Anzahl abgelehnter Hypothesen, damit

$$V_0 \leq R_0 \Rightarrow F_{V_0} \geq F_{R_0} \overset{\text{(AMI)}}{\Rightarrow} \Theta(F_{V_0}) \leq \Theta(F_{R_0})$$

3. Kontrolle des Parameters $\Theta(F_{R_0})$, bzgl. der beobachtbaren Anzahl von abgelehnten Hypothesen, unter der Null-Verteilung:

$$\Theta(F_{R_0}) \leq \alpha.$$

Null-Dominiertheit

(1) ist abhängig von Θ und gilt unter folgenden allgemeinen **Null-Dominiertheits-Bedingungen**:

- Q_0 dominiert die Verteilung $F_{V_n} : x \in \{0, ..., m\}$:

$$F_{V_n}(x) \geq F_{V_0}(x),$$
$$\liminf_{n \to \infty} F_{V_n}(x) \geq F_{V_0}(x),$$

  Insbesondere gilt dies, falls

- Q_0 dominiert die gemeinsame Verteilung Q_{n,I_0} des I_0-Vektors $(T_{ni} : i \in I_0)$:

$$Q_{n,I_0} \geq Q_{0,I_0},$$
$$\liminf_{n \to \infty} Q_{n,I_0} \geq Q_{0,I_0}.$$

Die erste Ungleichung in (1) folgt aus (AMI), für die zweite benötigen wir ebenfalls (ACI).

Konstruktion einer MTP

Schreibe für einen kritischen Wert $c \in \mathbb{R}^m$ und eine Verteilung $Q \in \{Q_0, Q_n\}$

$$R(c|Q) = \sum_{i \in I} \mathbf{1}_{\{T_{ni} > c_i\}}, \qquad T_n \sim Q,$$

$$V(c|Q) = \sum_{i \in I_0} \mathbf{1}_{\{T_{ni} > c_i\}}, \qquad T_n \sim Q.$$

Für die Null-Verteilung Q_0 auf dem $\mathbb{R}^m$ mit Randverteilungen Q_{0i} und für ein $\delta \in [0, 1]$ definieren wir außerdem den Vektor $d(Q_0, \delta)$ der δ-Quantile:

$$d(Q_0, \delta)_i = Q_{0i}^{-1}(\delta) = \inf\{z : Q_{0i}(z) \geq \delta\}, \quad i = 1, \ldots m.$$

Methode 1: common-quantil

Gegeben eine Null-Verteilung Q_0 und ein Niveau $\alpha \in (0,1)$, wähle

$$\delta_0(\alpha) = \inf\{\delta : \Theta(F_{R(d(Q_0,\delta)|Q_0)}) \leq \alpha\}.$$

Dann definieren wir die **Ein-Schritt common-quantil multiple Testprozedur** mittels der kritischen Werte

$$c(Q_0, \alpha) = d(Q_0, \delta_0(\alpha)) = (Q_{0i}^{-1}(\delta_0(\alpha)) : i = 1, ..., m),$$

welche das Typ-I-Fehlermaß $\Theta(F_{V(c(Q_0,\alpha)|Q_n)})$ zum Niveau α kontrolliert:

$$\mathcal{R}(T_0, Q_0, \alpha) = \{i : T_{ni} > c(Q_0, \alpha)_i\}.$$

Theorem 1: Asymptotische Kontrolle für die common-quantil Methode

Es existiere eine $\mathbb{R}^m$-wertige Zufallsvariable $Z \sim Q_0$, so dass für alle $c \in \mathbb{R}^m$ und $x \in \{0, ..., m\}$ gilt:

$$\liminf_{n\to\infty} P^{Q_n}\left(\sum_{i\in I_0} \mathbf{1}_{\{T_{ni}>c_i\}} \leq x\right) \geq P^{Q_0}\left(\sum_{i\in I_0} \mathbf{1}_{\{Z_i>c_i\}} \leq x\right) \quad \text{(AQ0)}$$

Oder kurz: $\liminf_n F_{V(c|Q_n)}(x) \geq F_{V(c|Q_0)}(x), \forall x$. Weiterhin erfülle die Abb. Θ die Bedinungen (AMI) und (ACI) bei $F_{V(c|Q_0)}$.
Dann kontrolliert die common-quantil Methode mit kritschen Werten $c(Q_0, \alpha) = d(Q_0, \delta_0(\alpha))$ asymptotisch das Typ-I-Fehlermaßes $\Theta(F_{V(c|Q_n)})$ zum Niveau α, d.h.

$$\limsup_{n\to\infty} \Theta(F_{V(c|Q_n)}) \leq \alpha.$$

Methode 2: common-cut-off

Gegeben eine Null-Verteilung Q_0 und ein Niveau $\alpha \in (0,1)$, wähle

$$e(Q_0, \alpha) = \inf\{c \in \mathbb{R} : \Theta(F_{R((c,..,c)|Q_0)}) \leq \alpha\}.$$

Dann definieren wir die **Ein-Schritt common-cut-off multiple Testprozedur** mittels des kritischen Wertes $e(Q_0, \alpha)$ durch

$$c(Q_0, \alpha) = (e(Q_0, \alpha), ..., e(Q_0, \alpha)),$$

welche das Typ-I-Fehlermaß $\Theta(F_{V(c(Q_0,\alpha)|Q_n)})$ zum Niveau α kontrolliert:

$$\mathcal{R}(T_0, Q_0, \alpha) = \{i : T_{ni} > c(Q_0, \alpha)_i\}.$$

Common-qunatil vs. common-cut-off

- Beide Methoden sind äquivalent, falls $(T_{ni})_{i=1,\ldots,m}$ unter Q_0 identisch verteilt sind.
- Unterschiede in: Balance, Güte und technischer Umsetzbarkeit.
- Wird Q_0 durch Resampling geschätzt (bootstrap) tendiert die common-quantil Methode zur größerer Sensibilität gegenüber der Anzahl der Resampling-Schritte und der Diskretheit der geschätzten Null-Verteilung.

Theorem 2: Allgemeine Konstuktion der Null-Verteilung

Es seien $\lambda_0 \in \mathbb{R}^m$ und $\tau_0 \in \mathbb{R}_0^m$ so gegeben, dass gilt

$$\limsup_{n\to\infty} \mathbb{E}[T_{ni}] \leq \lambda_0 \quad \text{und} \quad \limsup_{n\to\infty} Var(T_{ni}) \leq \tau_{0i}, \quad i \in I_0.$$

Definiere $\nu_i = \sqrt{\min\left(1, \frac{\tau_{0i}}{Var(T_{ni})}\right)}$ und einen Zufallsvektor verschobener und skalierter Teststatistiken

$$Z_{ni} = \nu_i(T_{ni} + \lambda_{0i} - \mathbb{E}[T_{ni}]), \quad i = 1, ..., m.$$

Falls $Z_n \xrightarrow{w} Z \sim Q_0 = Q_0(P)$, dann gilt für $c \in \mathbb{R}^m, x \in \{0, ..., m\}$

$$\liminf_{n\to\infty} P^{Q_n}\left(\sum_{i\in I_0} \mathbf{1}_{\{T_{ni}>c_i\}} \leq x\right) \geq P^{Q_0}\left(\sum_{i\in I_0} \mathbf{1}_{\{Z_i>c_i\}} \leq x\right)$$

Damit gilt (AQ0) für die Nullverteilung Q_0 und Theorem 1 ist anwendbar.

Diskusion von Theorem 2

- Bei einer zusammengesetzten Hypothese H_i wird λ_{0i} am Schwellenwert bestimmt.
- $\lambda_0 \in \mathbb{R}^m$ zur Erzeugung von Statistiken $(Z_{ni})_{i \in I_0}$ die stochastisch größer sind als die $(T_{ni})_{i \in I_0}$ und daher gegen eine Verteilung konvergieren, die (AQ0) erfüllt.
- $\tau_0 \in \mathbb{R}_0^m$ zur Vermeidung einer degenerierter asymptotischer Nullverteilung und unendlicher kritischer Werte
- λ_0, τ_0 hängen nur von den Randverteilungen der wahren Verteilung von T_n ab
- hängen λ_0, τ_0 vom unbekannten P ab, so können sie durch konsistente Schätzer ersetzt werden.
- τ_0 ist für FWER-Kontrolle nicht zwingend nötig.

Bootstrap-Schätzung der Nullverteilung

Schätze wahre Verteilung P aus den Daten $X_1, ..., X_n$ durch $P_n^\star$
$\Rightarrow$ Generiere **bootstrap-sample**: n iid. Realisierungen $X_1^\sharp, ..., X_n^\sharp \sim P_n^\star$.
$\Rightarrow$ Erzeuge Teststatistik $(T_i^{\sharp n})_{i=1,...,m}$ aus bootstrap-sample
$\Rightarrow$ Berechne entsprechend Theorem 2

$$Z_i^{\sharp n} = \sqrt{\min(1, \frac{\tau_{0i}}{Var_{P_n^\star}(T_i^{\sharp n})})}(T_i^{\sharp n} + \lambda_{0i} - \mathbb{E}_{P_n^\star}[T_i^{\sharp n}]), \quad i = 1, ..., m.$$

$\Rightarrow$ Schätzung der Verteilung von $(Z_i^{\sharp n})_{i=1,...,m}$ durch empirische Verteilungsfunktion über B bootstrap-samples
$\Rightarrow$ Approximation von $Q_0(P)$ (aus Theorem 2)

Methode 3: Bootstrap-Schätzung der Nullverteilung

1. Erzeuge B bootstrap samples $\{X_{1,b}^{\sharp}, ..., X_{n,b}^{\sharp}\}$ für $b = 1, ..., B$ mit $X_{i,b}^{\sharp} \sim P_n^{\star}, i = 1, ..., n, b = 1, ..., B$
2. Berechne für jedes bootstrap sample die Teststatistiken $T_{\cdot,b}^{\sharp n} = (T_{i,b}^{\sharp n} : i = 1, ..., m)$, so dass wir eine $m \times B$-Matrix $T^{\sharp n} = (T_{i,b}^{\sharp n})$ erhalten.
3. Berechne zeilenweise Erwartungswerte und Varianzen in der Matrix $T^{\sharp n}$ um $\mathbb{E}[T_{ni}]$ und $Var(T_{ni}), i = 1, ..., m$, zu schätzen.
4. Erzeuge $m \times B$-Matrix $Z^{\sharp n} = (Z_{i,b}^{\sharp n})$ durch zeilenweises Verschieben und Skalieren von $T^{\sharp n}$
5. Die bootstrap Schätzung Q_{0n} der Nullverteilung Q_0 aus Theorem 2 erhalten wir als empirische Verteilung der Spalten $Z_{\cdot,b}^{\sharp n}$ der Matrix $Z^{\sharp n}$.

Methode 4: Bootstrap-Schätzung der common-quantil kritischen Werte

1. Wende Methode 3 an um die Matrix $Z^{\sharp n}$ und die geschätze Nullverteilung Q_{0n} zu ermitteln.
2. Die bootstrap common-quantil cut-offs sind die Zeilenquantile der Matrix $Z^{\sharp n}$, also die δ-Quantile des B-Vektors $Z^{\sharp n}_{i,\cdot}$:

$$d(Q_{0n,i}, \delta) = Q^{-1}_{0n,i}(\delta) = \inf \left\{ z : \frac{1}{B} \sum_{b=1}^{B} \mathbf{1}_{\{Z^{\sharp n}_{i,b} \leq z\}} \geq \delta \right\}, \quad i = 1, ..., m$$

3. Für einen Test zum Niveau $\alpha \in (0, 1)$, wird δ gewählt als

$$\delta_{0n}(\alpha) = \inf\{\delta : \Theta(F_{R(d(Q_{n0,\delta})|Q_{0n})}) \leq \alpha\}.$$

4. FWER: (min-P)
   1. p-Wert-Matrix $P^{\sharp n}$ bestimmen durch Ersetzten der Einträge in $Z^{\sharp n}$ durch deren zeilenweise Ordnungszahlen (groß zu klein).
   2. Wähle in jeder Spalte von $P^{\sharp n}$ den kleinsten p-Wert.
   3. $(1 - \delta_{0n}(\alpha))$ ist das α-Quantil dieses B-Vektors der kleinsten p-Werte.

Methode 5: Bootstrap-Schätzung der common-cut-offs

1. Wende Methode 3 an um die Matrix $Z^{\sharp n}$ und die geschätze Nullverteilung Q_{0n} zu ermitteln.
2. Berechne den gemeinsamen kritischen Werte aus Q_{0n} entsprechend

$$c(Q_{0n}, \alpha) = e(Q_{0n}, \alpha) = \inf\{c \in \mathbb{R} : \Theta(F_{R((c,..,c)|Q_{0n})}) \leq \alpha\}$$

3. FWER: (max-T)
   1. Bestimme in jeder Spalte von $Z^{\sharp n}$ den größten Wert.
   2. $e(Q_{0n}, \alpha)$ ist das $(1-\alpha)$-Quantil des B-Vektors der größten Werte.