Frequentistische Modellmittelung unter Berücksichtigung der Problematik fehlender Daten

Institutskolloquium, IMBE, Erlangen
22. Juli 2009

Michael Schomaker
Ludwig Maximilian Universität, Institut für Statistik

Übersicht

Modellselektion

- Modellselektion ist ein weitläufiger Begriff und umschließt:
  - ✓ Selektion von Variablen in Regressionsanalysen
  - ✓ Bestimmung der Faktoren in der Faktorenanalyse
  - ✓ Wahl der Ordnung autoregressiver Prozesse
  - ✓ Bestimmung der Anzahl der Kontrollpunkte in Bild- und Konturanalysen
  - ✓ ...

- Um Variablen zu selektieren, verwendet man in der Regressionsanalyse häufig
  - ✓ Testprozeduren (z.B. schrittweise Selektion),
  - ✓ Informationskriterien (z.B. AIC),
  - ✓ bayesianische Ansätze (z.B. BIC),
  - ✓ Vorhersagefehler (z.B. CV),
  - ✓ Minimum Description Length (MDL)
  - ✓ statistische Lernenverfahren
  - ✓ Shrinkage ...

Post Model Selection Inference

- Ist ein „Gewinnermodell“ gefunden, dann wird die Inferenz auf diesem Modell so durchgeführt, als wäre dieses *a priori* gewählt worden.

Problem:

Der Selektionsschritt ist datenbasiert; Unsicherheit bezüglich des Selektionsprozesses wird nicht berücksichtigt.

→ Konfidenzintervalle zu klein!!
→ evtl. verzerrte Parameterschätzungen
→ Tests nicht korrekt

- Die Verteilung der Parameterschätzer nach der Modellselektion (post model selection) ist nicht trivial und analytisch gar nicht bzw. sehr schwer zu bestimmen, vgl. in etwa Leeb und Pötscher (2006)

Modellmittelung

- Lösung für dieses Problem: Kombiniere mehrere „plausible“ Modelle gewichtet miteinander

- Betrachte einige Kandidatenmodelle $\mathcal{M} = \{M_1, \ldots, M_k\}$ um den Response y basierend auf verschiedenen Kombinationen der Kovariablen $X_j = (x_{1j}, \ldots, x_{nj})^t$ zu beschreiben

- Betrachte die bedingte Dichte $f(y_i|\mathbf{x}_i;\theta)$, $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})$. Dann ist

$$\hat{\hat{\theta}} = \sum_{\kappa=1}^{k} w_\kappa \hat{\theta}_\kappa$$

ein FMA-Schätzer, wobei $\hat{\theta}_\kappa$ die Parameterschätzung von θ im Modell $M_\kappa \in \mathcal{M}$ bezeichnet.

Kriteriumsbasierte Gewichte

- Einfache Möglichkeit der Gewichtung über „Plausibiltät“:

$$w_\kappa = \frac{\exp(-\frac{1}{2}\Gamma_\kappa)}{\sum_{\kappa=1}^{k}\exp(-\frac{1}{2}\Gamma_\kappa)}$$

wobei Γ_κ den Wert eines Selektionskriteriums, z.B. das AIC, in Modell M_κ bezeichnet („lower is better“).

- Die Varianz des FMA-Schätzers kann über Bootstrapping bzw. über

$$\widehat{Var}(\hat{\bar{\theta}}) = \left\{\sum_{\kappa=1}^{k} w_\kappa \sqrt{\widehat{Var}(\hat{\theta}_\kappa|M_\kappa) + (\hat{\theta}_\kappa - \hat{\bar{\theta}})^2}\right\}^2$$

geschätzt werden ⇒ Varianz innerhalb und zwischen den Modellen wird berücksichtigt!

Modellmittelung - Weitere Ansätze

- Wähle die Gewichte w_κ so, dass die zugehörige Schätzung $\hat{\bar{\theta}}$ „optimal“ ist.

- Der Ansatz von Hansen (2007) greift dieses Prinzip auf. Der Schätzer behält seine Optimalitätseigenschaften aber nur für das lineare Modell. Außerdem ist das Ergebnis abhängig von der Anordnung der Regressoren.

- Bayesianische Ansätze über Schätzung der posteriori Wahrscheinlichkeit $p(M_\kappa|y)$
  $\rightarrow$ Bei korrekter Durchführung sehr aufwändig!

- WALS – „Weighted Averaged Least Squares“

Ein gewichtetes Akaike Kriterium

- Das AIC verliert im Kontext fehlender Daten (MAR) ggf. seine Optimalität. Hens, Aerts und Molenberghs (2006) schlagen eine Modifikation vor.

- Sei

$$\delta_i = \begin{cases} 1, & \text{für eine vollständige } i\text{-te Beobachtung } (y_i, \mathbf{x}_i) \\ 0, & \text{für eine unvollständige } i\text{-te Beobachtung } (y_i, \mathbf{x}_i) \end{cases}$$

  und sei $\pi_i = P(\delta_i = 1)$ die Wahrscheinlichkeit die i-te Zeile vollständig zu beobachten.

- Es ergibt sich ein gewichtetes Akaike Kriterium

$$\mathrm{AIC_W} = -2\sum_{i=1}^{n} \tilde{w}_i \, L_i(\hat{\theta}_W) + 2K$$

  mit $\tilde{w}_i = \delta_i / \pi_i$. $\hat{\theta}_W$ ist die gewichtete ML-Schätzung.

FMA-AIC$_W$-Schätzer

Idee: Das AIC_W kann verwendet werden, um Gewichte zur Modellmittelung für fehlende Daten zu konstruieren:

$$w_\kappa^{(adj)} = \frac{\exp(-\frac{1}{2}\text{AIC}_{\text{W},\kappa})}{\sum_{\kappa=1}^{k}\exp(-\frac{1}{2}\text{AIC}_{\text{W},\kappa})}$$

Vor- und Nachteile:

$+$	verhältnismäßig leicht zu berechnen, intuitiv plausibel Idee kann auf andere adjustierte Kriterien übertragen werden
$-$	Ad-hoc Ansatz (keine theoretische Rechtfertigung) Die Gewichte $\tilde{w}_i$ müssen geschätzt werden (z.B. mit GAM) Weitere Unsicherheit bei der Wahl eines GAMs (GCV) Effizienzverlust, da nur vollständige Daten verwendet werden

Schätze Varianz für $\hat{\hat{\theta}}_{\text{AIC}_W} = \sum_{\kappa=1}^{k} w_\kappa^{(adj)}\hat{\theta}_{W,\kappa}$:

$$\widehat{Var}(\hat{\hat{\theta}}_{\text{AIC}_W}) = \left\{\sum_{\kappa=1}^{k} w_\kappa \sqrt{\widehat{Var}(\hat{\theta}_{W,\kappa}|M_\kappa) + (\hat{\theta}_{W,\kappa} - \hat{\hat{\theta}}_{\text{AIC}_W})^2}\right\}^2$$

FMA-Imp-Schätzer (I)

- Betrachte einen imputierten Datensatz D^{imp} und den daraus resultierenden Schätzer:

$$\hat{\hat{\theta}}_{\text{imp}} = \sum_{\kappa=1}^{k} \{w_\kappa \hat{\theta}_\kappa | D^{\text{imp}}\} ,$$

- Berücksichtige *Unsicherheit bezüglich der Imputation* über M (korrekte) multiple Imputationen:

$$\hat{\hat{\theta}}_{\text{imp}}^{M} = \frac{1}{M} \sum_{m=1}^{M} \hat{\hat{\theta}}_{\text{imp}}^{(m)}$$

$\rightarrow$ „gemittelter Modellmittelungsschätzer“

FMA-Imp-Schätzer (II)

- Varianzschätzung:

$$\widehat{Var}(\hat{\hat{\theta}}^{M}_{\text{imp}}) = \frac{1}{M}\sum_{m=1}^{M}\widehat{Var}(\hat{\hat{\theta}}^{(m)}_{\text{imp}}) + \frac{M+1}{M(M-1)}\sum_{m=1}^{M}(\hat{\hat{\theta}}^{(m)}_{\text{imp}} - \hat{\hat{\theta}}^{M}_{\text{imp}})(\hat{\hat{\theta}}^{(m)}_{\text{imp}} - \hat{\hat{\theta}}^{M}_{\text{imp}})'$$

$\Rightarrow$ Berücksichtigt die Unsicherheit bezüglich der Modellselektion **und** bezüglich der Imputation!

Vor- und Nachteile:

+	verhältnismäßig leicht zu berechnen, intuitiv plausibel Idee ist nicht auf das AIC beschränkt!!! Alle Unsicherheitskomponenten werden erfasst
–	Wissen über Imputationsmethoden notwendig Grenzen bei systematischen Fehlen (NMAR)

Ein Beispiel - Muskelschwäche vom Typ Duchenne (I)

- Genetisch vererbte Krankheit
- Weibliche DMD-Träger kaum beeinträchtigt, männliche Träger sterben früh!
- 208 Beobachtungen an weiblichen Patienten, wovon 75 DMD-Träger sind.
- Potentielle Kovariablen für ein (Prognose-)Modell sind: Alter (AGE), sowie die Serum Marker CK, creatine kinase, H, hemopexin, PK, pyruvate kinase, LD, lactate dehydroginase.
- Einige Beobachtungen für die (teuren!) Serum-Marker PK und LD fehlen.
- Einige zusätzliche Beobachtungen von PK und LD werden über einen MAR-Fehlendmechanismus verworfen.

Ein Beispiel - Muskelschwäche vom Typ Duchenne (II)

5 Kandidatenmodelle:

M_1: $\ln \frac{p}{1-p} = \alpha+\gamma_1\text{AGE}+\gamma_2\text{CK}+\gamma_3\text{H}+\gamma_4\text{PK}+\gamma_5\text{LD}+\gamma_6\text{CK}\cdot\text{H}+\gamma_7\text{PK}\cdot\text{LD}$

M_2: $\ln \frac{p}{1-p} = \alpha+\gamma_1\text{AGE}+\gamma_2\text{CK}+\gamma_3\text{H}+\gamma_4\text{PK}+\gamma_5\text{LD}+\gamma_6\text{CK}\cdot\text{H}$

M_3: $\ln \frac{p}{1-p} = \alpha+\gamma_1\text{AGE}+\gamma_2\text{CK}+\gamma_3\text{H}+\gamma_4\text{PK}+\gamma_5\text{LD}+\gamma_7\text{PK}\cdot\text{LD}$

M_4: $\ln \frac{p}{1-p} = \alpha+\gamma_1\text{AGE}+\gamma_2\text{CK}+\gamma_3\text{H}+\gamma_4\text{PK}+\gamma_5\text{LD}$

M_5: $\ln \frac{p}{1-p} = \alpha+\gamma_1\text{AGE}+\gamma_2\text{CK}+\gamma_3\text{H}$

- AIC-Werte (für die vollständigen Fälle der Originaldaten):
  M_1: 92.60 M_2: 92.15 M_3: 96.33 M_4: 96.04 M_5: 135.08

Gibt es ein „bestes“ Modell?

Ein Beispiel - Muskelschwäche vom Typ Duchenne (III)

kein DMD	AGE	CK	H	PK	LD
Minimum	20.00	15.00	34.00	2.800	66.0
1. Quartil	25.00	26.25	76.30	9.725	136.0
Median	27.00	34.00	82.50	11.900	162.0
arith. Mittel	28.81	39.13	82.95	12.146	164.6
3. Quartil	32.75	45.00	91.00	15.300	185.0
Maximum	39.00	130.00	111.50	22.700	349.0

DMD	AGE	CK	H	PK	LD
Minimum	20.00	19.0	9.00	8.30	122.0
1. Quartil	30.50	54.0	83.50	14.30	198.0
Median	35.00	101.0	91.00	19.40	245.0
arith. Mittel	38.13	187.2	86.68	23.93	256.2
3. Quartil	42.50	235.0	100.30	25.25	297.0
Maximum	61.00	1288.0	118.00	110.00	593.0

Ein Beispiel - Muskelschwäche vom Typ Duchenne (IV)

Betrachtet werden die folgenden Methoden – jeweils für Modellselektion und Modellmittelung:

1. Verwendung der Originaldaten
2. Verwendung der Complete Cases
3. Verwendung des FMA-Imp-Schätzers mit den Imputationsmethoden
   - GAMRI – generalized additive model based recursive imputation
   - MI – Multiple Imputationen mit dem R-Paket „Amelia II“
4. Verwendung des FMA-AIC$_W$-Schätzers

Resultate

	I.		AGE		CK		H	
Frequentistische Modellmittelung								
1) Original	−9.10	(8.34)	0.16	(0.05)	−0.24	(0.18)	−0.06	(0.09)
2) CC	1.85	(9.32)	0.17	(0.07)	−0.66	(0.27)	−0.20	(0.12)
3) GAMRI	−13.83	(3.31)	0.18	(0.05)	0.02	(0.03)	−0.01	(0.02)
MI	−12.61	(5.27)	0.17	(0.05)	0.02	(0.03)	−0.01	(0.08)
4) AIC_W	2.53	(9.31)	0.17	(0.07)	−0.68	(0.27)	−0.21	(0.12)
Frequentistische Modellselektion								
1) Original	−5.13	(6.71)	0.17	(0.05)	−0.30	(0.16)	−0.09	(0.08)
2) CC	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)
3) GAMRI	−13.94	(2.36)	0.17	(0.05)	0.04	(0.01)	0.00	(0.01)
MI	−13.38	(3.87)	0.17	(0.04)	0.04	(0.01)	0.01	(0.01)
4) AIC_W	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)

	PK		LD		PK×LD		CK×H	
Frequentistische Modellmittelung								
1) Original	0.23	(0.18)	0.02	(0.01)	−0.001	(0.001)	0.003	(0.002)
2) CC	0.05	(0.19)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
3) GAMRI	0.19	(0.12)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.12	(0.09)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.04	(0.21)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
Frequentistische Modellselektion								
1) Original	0.12	(0.05)	0.01	(0.01)	0.000	(0.000)	0.004	(0.002)
2) CC	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)
3) GAMRI	0.16	(0.07)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.11	(0.05)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)

Resultate – Originaldaten

	I.		AGE		CK		H	
Frequentistische Modellmittelung								
1) Original	−9.10	(8.34)	0.16	(0.05)	−0.24	(0.18)	−0.06	(0.09)
2) CC	1.85	(9.32)	0.17	(0.07)	−0.66	(0.27)	−0.20	(0.12)
3) GAMRI	−13.83	(3.31)	0.18	(0.05)	0.02	(0.03)	−0.01	(0.02)
MI	−12.61	(5.27)	0.17	(0.05)	0.02	(0.03)	−0.01	(0.08)
4) AIC_W	2.53	(9.31)	0.17	(0.07)	−0.68	(0.27)	−0.21	(0.12)
Frequentistische Modellselektion								
1) Original	−5.13	(6.71)	0.17	(0.05)	−0.30	(0.16)	−0.09	(0.08)
2) CC	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)
3) GAMRI	−13.94	(2.36)	0.17	(0.05)	0.04	(0.01)	0.00	(0.01)
MI	−13.38	(3.87)	0.17	(0.04)	0.04	(0.01)	0.01	(0.01)
4) AIC_W	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)

	PK		LD		PK×LD		CK×H	
Frequentistische Modellmittelung								
1) Original	0.23	(0.18)	0.02	(0.01)	−0.001	(0.001)	0.003	(0.002)
2) CC	0.05	(0.19)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
3) GAMRI	0.19	(0.12)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.12	(0.09)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.04	(0.21)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
Frequentistische Modellselektion								
1) Original	0.12	(0.05)	0.01	(0.01)	0.000	(0.000)	0.004	(0.002)
2) CC	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)
3) GAMRI	0.16	(0.07)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.11	(0.05)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)

Resultate – Unterschiede bezüglich der Strategie

	I.		AGE		CK		H	
Frequentistische Modellmittelung								
1) Original	−9.10	(8.34)	0.16	(0.05)	−0.24	(0.18)	−0.06	(0.09)
2) CC	1.85	(9.32)	0.17	(0.07)	−0.66	(0.27)	−0.20	(0.12)
3) GAMRI	−13.83	(3.31)	0.18	(0.05)	0.02	(0.03)	−0.01	(0.02)
MI	−12.61	(5.27)	0.17	(0.05)	0.02	(0.03)	−0.01	(0.08)
4) AIC_W	2.53	(9.31)	0.17	(0.07)	−0.68	(0.27)	−0.21	(0.12)
Frequentistische Modellselektion								
1) Original	−5.13	(6.71)	0.17	(0.05)	−0.30	(0.16)	−0.09	(0.08)
2) CC	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)
3) GAMRI	−13.94	(2.36)	0.17	(0.05)	0.04	(0.01)	0.00	(0.01)
MI	−13.38	(3.87)	0.17	(0.04)	0.04	(0.01)	0.01	(0.01)
4) AIC_W	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)

	PK		LD		PK×LD		CK×H	
Frequentistische Modellmittelung								
1) Original	0.23	(0.18)	0.02	(0.01)	−0.001	(0.001)	0.003	(0.002)
2) CC	0.05	(0.19)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
3) GAMRI	0.19	(0.12)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.12	(0.09)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.04	(0.21)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
Frequentistische Modellselektion								
1) Original	0.12	(0.05)	0.01	(0.01)	0.000	(0.000)	0.004	(0.002)
2) CC	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)
3) GAMRI	0.16	(0.07)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.11	(0.05)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)

Resultate – Unterschiedliche Standardfehler

	I.		AGE		CK		H	
Frequentistische Modellmittelung								
1) Original	−9.10	(8.34)	0.16	(0.05)	−0.24	(0.18)	−0.06	(0.09)
2) CC	1.85	(9.32)	0.17	(0.07)	−0.66	(0.27)	−0.20	(0.12)
3) GAMRI	−13.83	(3.31)	0.18	(0.05)	0.02	(0.03)	−0.01	(0.02)
MI	−12.61	(5.27)	0.17	(0.05)	0.02	(0.03)	−0.01	(0.08)
4) AIC_W	2.53	(9.31)	0.17	(0.07)	−0.68	(0.27)	−0.21	(0.12)
Frequentistische Modellselektion								
1) Original	−5.13	(6.71)	0.17	(0.05)	−0.30	(0.16)	−0.09	(0.08)
2) CC	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)
3) GAMRI	−13.94	(2.36)	0.17	(0.05)	0.04	(0.01)	0.00	(0.01)
MI	−13.38	(3.87)	0.17	(0.04)	0.04	(0.01)	0.01	(0.01)
4) AIC_W	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)

	PK		LD		PK×LD		CK×H	
Frequentistische Modellmittelung								
1) Original	0.23	(0.18)	0.02	(0.01)	−0.001	(0.001)	0.003	(0.002)
2) CC	0.05	(0.19)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
3) GAMRI	0.19	(0.12)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.12	(0.09)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.04	(0.21)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
Frequentistische Modellselektion								
1) Original	0.12	(0.05)	0.01	(0.01)	0.000	(0.000)	0.004	(0.002)
2) CC	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)
3) GAMRI	0.16	(0.07)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.11	(0.05)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)

Resultate – Wenig Imputationsunsicherheit

	I.		AGE		CK		H	
Frequentistische Modellmittelung								
1) Original	−9.10	(8.34)	0.16	(0.05)	−0.24	(0.18)	−0.06	(0.09)
2) CC	1.85	(9.32)	0.17	(0.07)	−0.66	(0.27)	−0.20	(0.12)
3) GAMRI	−13.83	(3.31)	0.18	(0.05)	0.02	(0.03)	−0.01	(0.02)
MI	−12.61	(5.27)	0.17	(0.05)	0.02	(0.03)	−0.01	(0.08)
4) AIC_W	2.53	(9.31)	0.17	(0.07)	−0.68	(0.27)	−0.21	(0.12)
Frequentistische Modellselektion								
1) Original	−5.13	(6.71)	0.17	(0.05)	−0.30	(0.16)	−0.09	(0.08)
2) CC	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)
3) GAMRI	−13.94	(2.36)	0.17	(0.05)	0.04	(0.01)	0.00	(0.01)
MI	−13.38	(3.87)	0.17	(0.04)	0.04	(0.01)	0.01	(0.01)
4) AIC_W	1.19	(7.57)	0.17	(0.07)	−0.66	(0.25)	−0.20	(0.11)

	PK		LD		PK×LD		CK×H	
Frequentistische Modellmittelung								
1) Original	0.23	(0.18)	0.02	(0.01)	−0.001	(0.001)	0.003	(0.002)
2) CC	0.05	(0.19)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
3) GAMRI	0.19	(0.12)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.12	(0.09)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.04	(0.21)	0.03	(0.02)	0.000	(0.001)	0.009	(0.003)
Frequentistische Modellselektion								
1) Original	0.12	(0.05)	0.01	(0.01)	0.000	(0.000)	0.004	(0.002)
2) CC	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)
3) GAMRI	0.16	(0.07)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
MI	0.11	(0.05)	0.02	(0.01)	0.000	(0.000)	0.000	(0.000)
4) AIC_W	0.09	(0.11)	0.03	(0.01)	0.000	(0.000)	0.009	(0.003)

Michael Schomaker

Fazit

- Unsicherheit bezüglich der Modellselektion kann leicht berücksichtigt werden. Einfachste Möglichkeit: Kriteriumsbasierte Gewichte!
- Fehlende Daten ⇒ Multiple Imputationen können Variabilität korrekt erfassen
- Fehlende Daten ⇒ Imputationsansatz erlaubt Flexibilität
- Fehlende Daten ⇒ Gewichtungsansatz mit AIC_W sehr interessant; kann aber zu neuen Unsicherheiten führen

Literatur

- Buckland, S.T., Burnham, K.P., Augustin, N.H. (1997) *Model selection: an integral part of inference*. Biometrics 53:603–618.
- Hansen, B.E. (2007) *Least squares model averaging*. Econometrica 75:1175–1189.
- Hens, N., Aerts, M., Molenberghs G. (2006) *Model selection for incomplete and design based samples*. Statistics in Medicine 25:2502–2520.
- Hjort, L., Claeskens, G. (2003) *Frequentist model average estimators*. Journal of the American Statistical Association 98:879–945.
- Leeb, H., Pötscher, B.M. (2006) *Can one estimate the conditional distribution of post-model-selection estimators?* Annals of Statistics 34:2554-2591.
- Schomaker, M., Wan, A.T.K., Heumann, C. (2009) *Frequentist model averging with missing observations*. Computational Statistics and Data Analysis, mimeo.