

Modellwahl bei longitudinalen Daten – am Beispiel einer Studie zur Langzeitwirkung von Knoblauch auf Fettablagerungen in Arterien

Ernst SCHUSTER

Universität Leipzig, Institut für Medizinische Informatik, Statistik und Epidemiologie (IMISE),
Liebigstraße 27, D-04103 Leipzig, Schuster@imise.uni-leipzig.de

Abstract. Die Modellauswahl durch absteigendes Vorgehen bei longitudinalen Daten wird an einem Beispiel illustriert. Dabei wird auf die Wahl der festen Effekte und der Kovarianzstruktur eingegangen. Mit WinBUGS werden solche und robustere Modelle geschätzt.

1. Einführung

Die Auswahl eines geeigneten Modells bei longitudinalen Messungen wird an Hand der in [1] erstmalig ausgewerteten Daten beschrieben. Es handelt sich dabei um eine doppelblinde randomisierte Studie zur Wirkung von Knoblauchkapseln auf arterielle Plaque-Ablagerungen, die sonografisch gemessen wurden. Das Plaque-Volumen (p_l in mm^3) wurde bei Aufnahme in die Studie sowie nach 18, 36 und 48 Monaten bestimmt. Problematisch an der Studie ist, dass sehr viele Teilnehmer die Studie schon vor der ersten Nachuntersuchung abgebrochen haben, und nur 152 von 280 Probanden die Studie nach 4 Jahren beendet haben. Von mir wurden alle Fälle einbezogen, für die mindestens zwei Plaque-Volumen vorlagen. Für die vorzeitigen Abbrecher kann meiner Meinung nach unterstellt werden, dass der Abbruchmechanismus vom Plaque-Volumen unbeeinflusst ist, was vom Verfahren unterstellt wird. Für das Plaque-Volumen als Zielgröße wird ein Modell für wiederholte Messungen mit gemischten Effekten angenommen:

$$y_i = X_i \beta + u_i + \varepsilon_i$$

dabei ist

- y_i der n_i Spaltenvektor der Plaque-Volumina (Zielgröße) von Patient i , wobei n_i 2, 3 oder 4 ist;
- X_i ist eine $n_i \times p$ Designmatrix mit den Werten der p Kovariablen;
- β ist der Vektor der Regressionskoeffizienten der festen Effekte;
- $u_i \sim N(0, \sigma^2)$ ist ein zufälliges Probandenniveau;
- $\varepsilon_i \sim N(0, W)$ ist der Vektor der korrelierten Fehler eines Probanden, wobei W die Kovarianzmatrix ist, z.B. der Dimension 4×4 für $n_i = 4$.

Zur Berechnung der Modelle wurde S-Plus 2000 [3] benutzt.

2. Aufstellung des allgemeinsten Modells und Termreduktion

Als Voraussetzung für ein absteigendes Vorgehen muss zuerst das allgemeinste untersuchte Modell aufgestellt werden. Als feste Effekte werden die Knoblauchtherapie $the(1$ nach Therapie, 0 sonst), das Geschlecht $sex(\text{Männer } 0, \text{Frauen } 1)$, das zentrierte Alter (in Jahren – 60) und die Behandlungszeit $t(\text{in Jahren})$ und deren Wechselwirkungen zugelassen. Als zufälliger Effekt wird lediglich ein individuelles Absolutglied $u_i \sim N(0, \sigma^2)$ eingeführt. Das

entspricht einer Korrelationsmatrix zwischen den Zeitpunkten mit einem (gleichen) Wert außerhalb der Hauptdiagonalen (compound symmetry). Zusätzlich wird eine allgemeine Korrelationsmatrix zwischen den Zeitpunkten modelliert. Eine schrittweise Termreduktion der festen Effekte ist in Tabelle 1 dokumentiert. Dabei ist das Ausgangsmodell mit M0 bezeichnet. Tabelle 1 enthält dabei nur p-Werte, da nach diesen die Termreduktion erfolgte.

Tabelle 1: Schrittweise Termreduktion der festen Effekte nach p-Werten.

	M0	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
	p-Werte										
(Intercept)	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
sex	0.0024	0.0020	0.0020	0.0020	0.0016	0.0014	0.0013	0.0012	0.0010	0.0016	0.0014
t	0.5845	0.5839	0.5745	0.6701	-----						
the	0.9334	-----									
alter.60	0.0681	0.0672	0.0680	0.0679	0.0724	0.0847	0.0846	0.1192	0.1172	-----	
sex:t	0.0584	0.0547	0.0499	0.0534	0.0506	0.0468	0.0532	0.0495	0.0403	0.0471	0.0372
sex:alter.60	0.4846	0.4819	0.4905	0.4905	0.5012	-----					
t:alter.60	0.0295	0.0280	0.0221	0.0238	0.0264	0.0340	0.0377	0.0474	0.0805	0.3426	-----
the:alter.60	0.1570	0.1348	0.1108	0.1108	0.1117	0.1082	0.1101	-----			
t:the	0.7743	0.7033	0.7008	-----							
the:sex	0.8693	-----									
sex:t:the	0.2614	0.0379	0.0383	0.0258	0.0261	0.0263	0.0298	0.0303	0.0374	0.0383	0.0205
sex:t:alter.60	0.5124	0.4923	0.3876	0.4021	0.4157	0.6202	-----				
the:alter.60:sex	0.7927	0.8258	-----								
t:the:alter.60	0.0827	0.0643	0.0660	0.0668	0.0707	0.0689	0.0636	0.3387	-----		
LQ-Tests: p-Wert (zum vorherigen Modell + zu M0)		0.9637	0.8239	0.6997	0.6692	0.4973	0.6185	0.1093	0.3385	0.1166	0.3455
		0.9637	0.9889	0.9915	0.9937	0.9886	0.9917	0.8809	0.8642	0.7154	0.7136
AIC	6685	6681	6679	6678	6676	6674	6672	6673	6672	6672	6671
BIC	6790	6777	6770	6764	6757	6751	6745	6741	6735	6731	6726

Die letzten vier Zeilen der Tabelle 1 beschreiben wie gut die Modelle angepasst sind. In der Zeile LQ-Tests stehen die Überschreitungswahrscheinlichkeiten (p-Werte) der Likelihood-Quotienten-Tests; oben zum vorhergehenden Modell und unten zum Ausgangsmodell M0. In den folgenden beiden Zeilen folgen das Akaike Informationskriterium (AIC: $-2 \log\text{-likelihood} + p * 2$) und das Bayes Informationskriterium (BIC: $-2 \log\text{-likelihood} + p * \ln(\text{Anzahl Fälle})$). Für gemischte Modelle erfolgt die Modellvereinfachung nicht automatisch. Von Seiten der Modellanpassung (LQ-Test, AIC) erscheint neben dem von mir ausgewählten Modell M10 auch das Modell M6 geeignet, dessen zusätzliche Wechselwirkung von Alter und Therapie mir aber zu unsicher erschien. Die Termreduktion liefert das Modell M10 mit den folgenden Schätzern der festen Effekte (s. Tab. 2):

$$pl \sim -15,45 * \text{sex} + 3,50 * \text{sex:t} - 4,64 * \text{sex:t:the} + 38,95$$

Das Plaque-Volumen bei Aufnahme in die Studie ($t=0$) ist bei den Männern ($\text{sex}=0$) durchschnittlich um ca. 15 mm^3 größer als bei den Frauen ($\text{sex}=1$). Bei den Männern gibt es weder einen Zeit- noch einen Therapieeffekt. Bei den Frauen ohne Therapie steigt das Plaque-Volumen jährlich um ca. $3,5 \text{ mm}^3$ an. Dieser Anstieg wird durch die Therapie überkompensiert.

Tabelle 2: Parameterschätzer (mit p-Wert in Klammern) für das vorausgewählte Modell M10 und Modelle mit anderer Kovarianzstruktur

		M10 vorausgewählt	M11 zufällige Anstiege	M12 bedingt unabhängig	M13 alle Fälle
		Schätzer (p-Wert)	Schätzer (p-Wert)	Schätzer (p-Wert)	Schätzer (p-Wert)
Feste Effekte	Konstante	38.95 ($<.0001$)	38.99 ($<.0001$)	38.93 ($<.0001$)	36.92 ($<.0001$)
	sex	-15.45 (0.0014)	-15.52 (0.0014)	-15.36 (0.0008)	-13.81 (0.0005)
	sex:t	3.50 (0.0372)	3.56 (0.0393)	3.25 (0.0377)	3.60 (0.0220)
	sex:t:the	-4.64 (0.0205)	-4.72 (0.0161)	-4.16 (0.0322)	-4.66 (0.0172)
Zufällige Effekte (Standard-Abweichung)	Konstante	19.11	19.49	18.95	18.13
	sex:t		3.62		
	Residuum	26.61	26.53	26.50	26.25
Korrelationen zwischen den Zeitpunkten	r01	0.338	0.325		0.328
	r02	-0.124	-0.135		-0.115
	r03	-0.136	-0.108		-0.133
	r12	-0.052	-0.057		-0.045
	r13	-0.059	-0.033		-0.060
	r23	-0.007	-0.005		0.001
LQ-Test zu M10: p-Werte			0.810	0.0033	
AIC		6672	6675	6679	
BIC		6726	6739	6707	

3. Auswahl der geeigneten Kovarianzstruktur

Die Parameterschätzer für das vorausgewählte Modell sind in Tabelle 2 in der ersten Spalte gezeigt. Die Hinzunahme eines individuellen Anstiegs führt auf das Modell M11. Da sich keine bessere Modellanpassung ergibt, wird das Modell abgelehnt. Bei Benutzung aller Fälle (Modell M13 in Tab. 2), also auch der Fälle, für die nur ein Plaque-Volumen vorliegt, erhält man dasselbe Modell und damit auch fast identische Parameterschätzer. Wegen der größeren Fallzahl sind AIC und BIC nicht vergleichbar und deshalb auch nicht angegeben worden. Logarithmieren des Plaque-Volumens zur Verbesserung der Anpassung an eine Normalverteilung brachte sehr ähnliche Ergebnisse, führt aber auf ein multiplikatives Modell und wurde deshalb nicht weiteruntersucht (Ergebnisse nicht dargestellt).

Betrachtet man in Modell M10 die Korrelationsmatrix, die sich bedingt auf zufälliges Kurvenniveau und feste Effekte ergibt, so legen ihre Werte (außer r01) es nahe diese Null zu setzen. Für das so gebildete Modell M12 (Tab. 2) liegt also ‚compound symmetry‘ vor. Das vereinfachte Modell M12 ist bzgl. AIC und Likelihood-Quotient schlechter angepasst. Nur BIC verbessert sich. Aber die Schätzer der festen Effekte sind sehr ähnlich. Das unterstreicht den bekannten Satz, dass bei korrekt spezifizierter Mittelwertstruktur Kleinste-Quadrate-Schätzer asymptotisch konsistent sind, auch wenn die Kovarianzstruktur fehlspezifiziert ist. Die Modelle M10 und M12 können als gleichwertig angesehen werden. Deshalb ist das einfachere Modell M12 vorzuziehen, das nun mit WinBUGS weiter untersucht wird.

4. Modellierung mit WinBUGS

Nach dem Bayes Theorem ist die a posteriori Verteilung proportional a priori Verteilung mal Likelihood-Funktion. Bei nichtinformativer a priori Verteilung ergibt die a posteriori Verteilung deshalb ähnliche Schätzungen wie die Likelihood-Funktion. Die wichtigste MCMC Methode ist der Gibbs Sampler [2], der anwendbar ist, wenn aus den bedingten Dichten jedes Parameters gegeben alle anderen Parameter und die Daten gezogen werden kann. Bei der Anwendung von WinBUGS [5] müssen diese bedingten Dichten nicht explizit angegeben werden, weil WinBUGS die Herleitung übernimmt. Nach Vorgabe von Startwerten für die Parameter zieht der Gibbs Sampler zyklisch Werte dieser eindimensionalen bedingten Verteilungen. Das Ergebnis ist eine Markov Kette, die nach Erreichen der stationären Verteilung eine Stichprobe der a posteriori Verteilung der Parameter gegeben die Daten liefert.

Die Knoblauchdaten werden nun weiter mit WinBUGS untersucht. Da nur bei den Frauen Zeit- und Therapieeffekte auftraten, genügt es nur die Daten der Frauen ($i = 1, 2, \dots, 56$) zu modellieren. Das Modell ergibt sich als Spezialfall des allgemeinen Modells:

$$y_i \sim N(X_i \beta + u_i, \sigma^2 I)$$

mit den folgenden Gleichungen und Verteilungsannahmen:

$$y_i \equiv (pl_{i1}, pl_{i2}, pl_{i3}, pl_{i4})^T, \quad t = (0, 1.5, 3, 4)^T, \quad \beta \equiv (\alpha_0, \beta_1, \beta_2)^T,$$

$$X_i \beta \equiv \alpha_0 + \beta_1 t_k + \beta_2 t_k \text{ the}_i, \quad u_i \equiv \sigma_1 \alpha_i, \quad \mu_{i,k} \equiv X_i \beta + u_i,$$

$$pl_{ik} \sim \text{Normal}(\mu_{i,k}, \tau), \quad pl_{ik} \sim \text{Normal}(0, 10^{-6}) \text{ for missing } pl_{ik},$$

$$\tau \sim \text{Gamma}(10^{-3}, 10^{-3}), \quad \alpha_0, \beta_1, \beta_2 \sim \text{Normal}(0, 10^{-6}),$$

$$\sigma_1 \sim \text{Unif}(0, 100), \quad \alpha_i \sim \text{Normal}(0, 1).$$

$\text{Normal}(\mu_{i,k}, \tau)$ bedeutet hier wie in BUGS (dort allerdings als $N(\dots)$ bezeichnet) normalverteilt mit Erwartungswert $\mu_{i,k}$ und Präzision τ , dem Kehrwert der Varianz. Die Präzision τ bekommt als nichtinformativ a priori Verteilung eine konjugierte Verteilung, nämlich eine Gamma-Verteilung. Die festen Effekte sind die Konstante α_0 , der Anstieg in der Zeit β_1 und der Faktor β_2 vor dem Produkt aus Zeit und Therapie. Alle diese erhalten als nichtinformativ a priori Verteilung eine Normalverteilung mit großer Varianz. Der Faktor σ_1 vor den standardnormalverteilten Zufallsgrößen beschreibt die Standardabweichung des zufälligen Niveaus der individuellen Kurven. Als nichtinformativ a priori Verteilung von wird eine Gleichverteilung von Null bis Einhundert verwendet.

Die Abb. 1 zeigt den mit WinBUGS aufgestellten gerichteten Graphen, der mit Verteilungen und Gleichungen unterlegt ist. Tab. 3 zeigt in der zweiten Spalte das Resultat von 100 000 Iterationen nach 1 Million Iterationen „burn-in“. Der Verlauf der Parameter über die 100 000 Iterationen (nicht dargestellt) spricht sehr dafür, dass die stationäre Verteilung erreicht wurde. Die erhaltenen Ergebnisse stimmen sehr gut mit denen von S-Plus (erste Spalte von Tab. 3) überein. Ein bemerkenswerter Vorteil der MCMC Methoden ist, dass auch Vertrauensintervalle für die zufälligen Effekte angegeben werden können. Der entscheidende Vorteil der MCMC Methoden, bei den hier vorliegenden großen Fehlern, ist aber, dass auch robustere Modelle gerechnet werden können, indem z.B. die Fehler als doppelt exponentiell

angenommen werden. Die Ergebnisse (dritte Spalte Tab.3) zeigen größere Unterschiede zum Ergebnis bei Normalverteilung, untermauern aber das aufgestellte Modell.

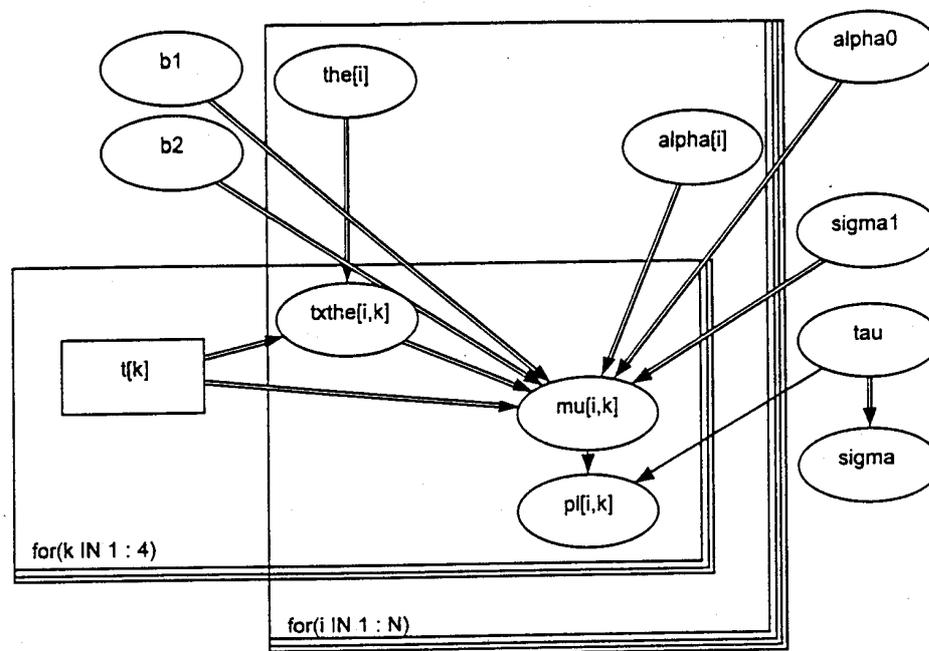


Abbildung 1: Graphisches Modell der Knoblauchdaten der Frauen.

Tabelle 3: Ergebnisse von S-Plus und BUGS.

		S-Plus bedingt unabhängig	BUGS- Modell Normalv.	BUGS- Modell Doppeltexp
		Schätzer (95%-Konfidenz- intervall)	Schätzer (95%-Vertrauens- intervall)	Schätzer (95%-Vertrauens- intervall)
Feste Effekte	Konstante	23.57 (16.66, 30.49)	23.58 (16.64, 30.52)	19.53 (15.38, 23.92)
	t (Jahre)	3.34 (0.51, 6.17)	3.32 (0.51, 6.16)	2.02 (0.30, 3.82)
	t:therapie	-4.33 (-7.75, -0.91)	-4.30 (-7.73, -0.89)	-3.234 (-5.33, -1.14)
Zufällige Effekte	Konstante	14.56	14.99 (9.87, 20.63)	9.972 (7.02, 13.67)
	Residuum	24.45	24.70 (22.16, 27.79)	19.36 (16.74, 22.56)

Literatur

- [1] Koscielny J, Klüßendorf D, Latza R, Schmitt R, Radtke H, Siegel G, Kiesewetter H: The antiatherosclerotic effect of Allium sativum. *Atherosclerosis* 1999; 144: 237-249.
- [2] Robert, C.P., Casella, G. (1999) *Monte Carlo Statistical Methods*, New York: Springer.
- [3] S-PLUS 2000 Guide to Statistics (1999), MathSoft, Seattle.

- [4] Schuster, E.: Markov Chain Monte Carlo Methods for Handling Missing Covariates in Multiple Regression Models. In: Friedl, H. Berghold, A. and Kauermann, G. (eds.) *Statistical Modelling – Proceedings of the 14th International Workshop on Statistical Modelling, Graz, Austria, July 19-23, 1999*, 651-655.
- [5] Spiegelhalter DJ, Thomas A, Best NG. *WinBUGS Version 1.2 User Manual*. Cambridge, U.K., 1999.