

## **Modellvalidierung durch "Posterior predictive checking" und "Leave-one-out"**

G. Nehmiz, M. Könen-Bergmann

Boehringer Ingelheim Pharma GmbH & Co. KG, 88397 Biberach und 55216 Ingelheim  
gerhard.nehmiz@bc.boehringer-ingelheim.com

Es soll geprüft werden, ob ein gegebenes Modell sich an Daten hinreichend gut anpasst (Validierung). Explizite Modell-Alternativen, zwischen denen eine Auswahl getroffen werden soll (Selektion), werden jetzt nicht betrachtet. Gelman et al. (1) untersuchen die prädiktive Verteilung für einen neuen Datenpunkt, die auf der a-posteriori-Verteilung der Parameter basiert, die wiederum - über die Likelihood-Funktion - aus allen Daten abgeleitet worden ist. Diese prädiktive Verteilung könnte man mit den tatsächlich beobachteten Daten vergleichen. Wie leistungsfähig ist dieser Ansatz?

Wir untersuchen das Einweg-Varianzanalyse-Modell mit zufälligen Effekten. Marshall et al. (2) zeigen analytisch, wie die prädiktive Verteilung die Güte der Anpassung überschätzt, wenn sie, wie oben beschrieben, auf allen Datenpunkten beruht. Alternative Verfahren, die auf dem Weglassen und Vorhersagen von "etwa der Hälfte" der Datenpunkte beruhen, sind abhängig von willkürlichen Entscheidungen und daher schwer untersuchbar. Ein systematisches Verfahren ist, jeden Datenpunkt genau 1x wegzulassen. Jeder Datenpunkt wird dann wieder mit der Verteilung des prädizierten Wertes verglichen, die aberrante Lage des Datenpunktes wird mit einem p-Wert-ähnlichen Maß, dem "prädiktiven p-Wert", beschrieben.

Wir untersuchen als Beispiel wiederholte EKG-Messungen aus einer klinischen Prüfung mit gesunden Probanden, insbesondere QTcF (vgl. Camm (3)). Die a-posteriori-Verteilungen der Modellparameter werden durch MCMC ermittelt (4).

Es zeigt sich, dass das einfache Verfahren, das die Daten 2x verwendet, nicht adäquat ist. Probleme des "Leave-one-out" bleiben die Kalibrierung der prädiktiven p-Werte und die gegenseitige Maskierung von Datenpunkten.

### **Literatur**

1. Gelman A, Carlin JB, Stern HS, Rubin DB:  
Bayesian Data Analysis (2. Aufl.).  
Boca Raton / London / New York / Washington/DC: Chapman & Hall / CRC 2004.
2. Marshall EC, Spiegelhalter DJ:  
Approximate cross-validators predictive checks in disease mapping models.  
Statistics in Medicine 2003; 22: 1649-1660
3. Camm AJ:  
How does pure heart rate lowering impact on cardiac tolerability?  
European Heart Journal Supplements 2006; 8 (D): D9-D15
4. Gilks WR, Richardson S, Spiegelhalter DJ (Hg.):  
Markov Chain Monte Carlo in Practice.  
London / Weinheim / New York / Tokyo / Melbourne / Madras: Chapman & Hall 1996