

## Das Behrens-Fisher-Problem für hochdimensionale Split-Plot-Designs

**Harden, Markus**

Universitätsmedizin Göttingen, Abteilung Medizinische Statistik  
markus.harden@med.uni-goettingen.de

**HINTERGRUND:** In vielen Anwendungsbeispielen aus Biologie und Medizin werden statistische Analysemethoden benötigt, um hochdimensionale Datensätze auswerten zu können. Solche Versuchsdesigns beinhalten weniger Individuen als Messwiederholungen, die an den Individuen erhoben wurden. Unter den Annahmen normalverteilter Daten und einer endlichen Anzahl von Messwiederholungen existieren verschiedene Ansätze für Approximationen, die dieses Problem lösen. Standardverfahren wie Hotellings  $T^2$  sind aufgrund einer singulären Kovarianzmatrix nicht anwendbar. Da keine exakte Lösung dieses Problems vorliegt, basieren alle Lösungsansätze auf Approximationen, die von Geisser-Greenhouse (GG) oder Huynh-Feldt (HF) beschrieben wurden. Diese Approximationen können jedoch zu falschen Entscheidungen führen, falls die „Sphericity-Condition“ an die Kovarianzmatrizen nicht erfüllt ist. Zusätzlich wird im Falle gleicher Kovarianzmatrizen die Qualität der Approximationen durch ungleiche Stichprobenumfänge mitunter massiv verschlechtert. Die Abweichung von der zuvor festgelegten Fehlerrate 1. Art nimmt zu, wenn die Anzahl der Messwiederholungen steigt.

**METHODEN:** Basierend auf einer Momenten-Approximation von Box (1954) und Patnaik (1949), wurde die Anova-Typ-Statistik, vgl. dazu Brunner, Dette und Munk (1997), auf hochdimensionale Split-Plot-Designs mit 2 Behandlungsgruppen angewendet. Die Schätzer für das Box'sche  $\epsilon$  aus Brunner, Becker und Werner (2010) für den Ein- und Zwei-Stichprobenfall werden für Split-Plot-Designs verallgemeinert. Diese Schätzer sind dimensionsstabil, was bedeutet, dass die Qualität der Approximation nicht von der Anzahl der Messwiederholungen abhängt. Für compound symmetry (CS) und autoregressive (AR) Kovarianzstrukturen zeigen die Simulationen bessere Ergebnisse des beschriebenen Verfahrens, als die Approximation von GG und HF.

**ERGEBNISSE:** Die in diesem Vortrag vorgestellte Approximation führt zu besseren Ergebnissen als die in PROC GLM (SAS Version 9.3) enthaltenen Lösungsansätze, wie die von GG und HF. Das SAS-Makro F1\_HF\_F1 stellt daher eine Alternative zu den implementierten Approximationen dar.

**Literatur:**

Box, GEP. Some theorems on quadratic forms applied in the study of analysis of variance problems, ii. effects of inequality of variance and of correlation between errors in the two-way classification. *The Annals of Mathematical Statistics*, 25(3):pp. 484-498 (1954). ISSN 00034851.

Brunner, E, Becker, B und Werner, C. Approximate distributions of quadratic forms in high-dimensional repeated-measures designs. Technischer Bericht, Department Medizinische Statistik Georg-August-Universität Göttingen (2010).

Brunner, E, Dette, H und Munk, A. Box-type approximations in nonparametric factorial designs. *Journal of the American Statistical Association* (1997).

Geisser, S und Greenhouse, SW. An extension of box's results on the use of the f distribution in multivariate analysis. *The Annals of Mathematical Statistics*, 29(3):pp. 885-891 (1958). ISSN 00034851.

Huynh, H und Feldt, LS. Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational and Behavioral Statistics*, 1(1):69-82 (1976). doi:10.3102/10769986001001069.

Patnaik, PB. The non-central chi- and f-distribution and their applications. *Biometrika*, 36(1/2):pp. 202-232 (1949). ISSN 00063444.