

Berichte aus der Statistik

**Rainer Muche,
Ralf Minkenberg (Hrsg.)**

KSFE 2013

Proceedings der 17. Konferenz der SAS®-Anwender
in Forschung und Entwicklung (KSFE)



Shaker Verlag
Aachen 2013

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Copyright Shaker Verlag 2013

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8440-2049-6

ISSN 1619-0963

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen

Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9

Internet: www.shaker.de • E-Mail: info@shaker.de

Vorwort

Der vorliegende Proceedingsband enthält die ausgearbeiteten Fassungen der Beiträge zur 17. Konferenz der SAS Anwender in Forschung und Entwicklung (KSFE). Die Konferenz fand vom 27. Februar – 01. März 2013 zum zweiten Mal nach 2007 in Ulm statt. Ausgerichtet wurde die Tagung von MitarbeiterInnen des Instituts für Epidemiologie und Medizinische Biometrie der Universität Ulm unter Leitung von Prof. Dr. Rainer Muehe in Zusammenarbeit mit dem KSFE e.V. – unterstützt durch die Firma SAS Deutschland sowie zahlreiche weitere Sponsoren wie HMS Analytical Software GmbH, iCASUS GmbH, Systematika Information Systems GmbH, viadee Unternehmensberatung GmbH, JMP Software D-A-CH, dsquare.de, KYBEIDOS Gesellschaft für Systeme zur Unternehmenssteuerung mbH, Condat AG, inVentiv Health Clinical, Altran und Springer Gabler Verlag.

Das Forum der Universität Ulm mit seinen Hörsälen war an den drei Tagen der Konferenz fast vollständig in der Hand von rund 450 Datenanalysten, Statistikern, Wissenschaftlern und Unternehmensberatern aus ganz Deutschland und dem deutschsprachigen Ausland, die ihre Erfahrungen und Kenntnisse im Umgang mit der SAS-Software teilen, diskutierten und vertieften. Für den Mittwoch meldete sich dabei knapp die Hälfte der Gäste sowohl für Tutorien zur praktischen Arbeit mit SAS-Produkten als auch für eine Zertifizierung im Umgang mit der SAS-Programmiersprache an. Diese international anerkannte Zertifizierung wurde in Ulm zum zweiten Mal im Rahmen einer KSFE-Tagung angeboten und fand wieder positiven Anklang bei den Teilnehmern, so dass einer Verstetigung nichts im Weg liegen sollte. An den beiden übrigen Tagen wurden in den Hörsälen des Forums parallel Beiträge zu aktuellen Themen wie den Neuerungen in der Software, in den klassischen Bereichen der Statistik z.B. Prognosemodellierung aber auch zu modernen Problemstellungen wie Visual and Text Mining präsentiert. Daneben gab es noch zahlreiche kleine und große „Tipps & Tricks“, die die Grundidee der Konferenz – den Austausch von Erfahrungen – unterstrichen. Weitere Diskussionsmöglichkeiten gab es natürlich auch im Rahmen der Poster-Ausstellung, in den Pausen, bei den Ständen der zahlreichen Aussteller oder auch während der Abendveranstaltung.

Ein Highlight der Konferenz war dann am Ende die Vergabe der diesjährigen KSFE Best Paper Awards an Ulrike Braisch Uni Ulm (1. Platz), Manuela Wern Schufa Holding AG (2. Platz) und Hans-Joachim Helms Uni Göttingen (3. Platz). Die Beiträge sind in diesem Tagungsband abgedruckt und können so noch einmal nachgelesen werden.

Dass die Tagung bei den Teilnehmern regen Anklang fand, zeigte sich auch in der Resonanz der Feedback-Bögen, die im Anschluss an die Tagung ausgewertet wurden. An dieser Stelle möchten wir Ihnen als Teilnehmer ein herzliches Dankeschön für die konstruktiven Rückmeldungen ausdrücken.

Bedanken möchten wir uns auch für die sehr gute Zusammenarbeit und die wertvolle Unterstützung während der Vorbereitung und der Durchführung der Konferenz beim KSFE-Organisationskomitee. Die Weitergabe der Erfahrungen aus vorausgehenden Konferenzen war dabei besonders hilfreich.

Weiterhin möchten wir Frau Carina Ortseifen vom Universitätsrechenzentrum Heidelberg und Frau Henriette Höhle vom KSFE e.V. für die professionelle Erstellung dieses Konferenzbandes danken. Nicht zuletzt möchten wir auch allen beteiligten Mitarbeiterinnen und Mitarbeitern des Instituts für Epidemiologie und Medizinische Biometrie der Uni Ulm für ihren Einsatz vor, während und nach der Konferenz unseren Dank aussprechen. Besonderer Dank geht dabei vor allem (in alphabetischer Reihenfolge) an Jens Dreyhaupt, Beate Einsiedler, Andrea Elsässer, Iris Lichtblau, Marianne Meule, Karin Schiefele und Sabrina Sufeida. Durch ihr großes Engagement wurde der Erfolg dieser Konferenz erst möglich.

Abschließend möchten wir Sie noch auf die 18. Konferenz der SAS Anwender in Forschung und Entwicklung an der Universität Göttingen hinweisen. Veranstalter im kommenden Jahr ist die Abteilung für Medizinische Statistik unter der Leitung von Prof. Dr. Tim Friede und der KSFE e.V. Vielleicht sind ja auch Sie bei der 18. KSFE 2014 in Göttingen mit dabei, um Ihr Fachwissen an andere Teilnehmer weiter zugeben!

Ulm, April 2013

Prof. Dr. Rainer Muche
Universität Ulm
Tagungsleiter

Ralf Minkenber
KSFE e.V.
Vorsitzender

Inhalt

<i>H. Fangerau, M. Martin</i>	1
Medizin und Statistik: Einblicke in die Geschichte einer schwierigen Beziehung	
<i>H. Armes</i>	11
Smarte Integration von R im SAS Data Step	
<i>S. Beimel</i>	27
Rechnen mit der Wahrheit und If 0 Then Set	
<i>T. Bluhmki, J. Habeck, F. Fleischer</i>	35
Die stochastische Verteilung der intra-individuellen Varianz in 2x2x2 Cross-Over Bioverfügbarkeitsstudien	
<i>U. Braisch, R. Muehe</i>	49
PROC LOGISTIC: Warum sind die Koeffizienten nicht mit den Odds Ratios konsistent?	
<i>T. Bruckner, L. Uhlmann, A. Deckert</i>	63
Quantilregression mit SAS	
<i>G. Büchele</i>	73
Wahlfach an der HS Ulm „WF-STUPA: Studienplanung, -durchführung und -auswertung in der praktischen Anwendung“ mit SAS	
<i>A. Büchse, A. Zenk</i>	83
Multiple Mittelwertvergleiche nach Student-Newman-Keuls in PROC MIXED	
<i>A. Deckert</i>	97
Programmierung anpassungsfähiger Makros durch Datensatzzerlegung am Beispiel eines erweiterten Bubble-Plots	
<i>J. Dreyhaupt, S. Sufeida, R. Muehe, J.-M. Steinacker</i>	111
Power- und Fallzahlabschätzungen für hierarchische und longitudinale Studien	
<i>E. Endri</i>	121
Timeline Figure - Grafische Darstellung eines Subjekts in der klinischen Forschung	
<i>E. Endri</i>	127
Studienauswertung per Knopfdruck. Code-Generierung direkt vom Statistical Analysis Plan - Ist das möglich?	

<i>K. Fink, E. Molz, A. Hungele, R. Ranz, M. Grabert, R. Holl</i> Automatisches Erstellen und Verschicken eines Benchmarking-Reports zur Qualitätssicherung dokumentierter Diabetesdaten mittels SAS 9.3	135
<i>S. Fink</i> Auswertung stratifizierter Studien mittels des Cox-Modells im Falle von seltenen Events	147
<i>K. Flunkert, S. Ernst</i> Beyond the Borders: Einbinden externer Shapefiles zur Erweiterung kartografischer Darstellungen	159
<i>L. Fortwengel</i> Leserliche Grafiken für Präsentationen	173
<i>B. Gigic, K. Buck, C. M. Ulrich</i> SAS DATA Step - Optionen vs. Anweisungen	177
<i>M. Harden</i> Das Behrens-Fisher-Problem für hochdimensionale Split-Plot-Designs	189
<i>H.-J. Helms, N. Benda, J. Zinserling, T. Friede</i> Auswertung von Dosis-Findungs-Studien mit aktiver Kontrolle in SAS 9.3	201
<i>W. Hering, R. T. Wigand, N. Agarwal</i> PROC IML vs. PROC REG: Erfolgsprognostik in Online Rollenspielen (MMORPG)	213
<i>B. Hientzsch, G. Lückel, J. Müller, N. Tambascia</i> Unabhängige Doppelprogrammierung - das Nonplusultra der Validierung?	225
<i>T. Hochgürtel, T. Lösch</i> Analysen mit der DRG-Statistik – Herausforderungen und Lösungsansätze	237
<i>L. Hupperz, F. Rohlmann, B. Einsiedler, R. Muche</i> Untersuchung zum Balanceverhalten der stratifizierten Blockrandomisierung - Eine Lösung mit SAS Makros -	249
<i>B. P. Jäger, C. Schüler, K.-E. Biebler, P. E. Rudolph</i> Was man in SAS Genetics vergeblich sucht: Allelfrequenzschätzungen bei Dominanz von Allelen mittels EM-Algorithmus	261
<i>R. Kaluscha, J. Holstiege, G. Krischak</i> Proc SQL: Passthrough-Facility für effizientes Datenmanagement bei komplexen großen Datenbeständen	277

<i>R. Kaluscha, S. Jankowiak, G. Krischak</i>	285
Full Model Selection mit vielen unabhängigen Variablen: Ein Beispiel für SAS-Tuning bei komplexen rechenintensiven Aufgaben	
<i>K. Lang</i>	291
SAS in den Ernährungswissenschaften – Berechnung der Nährstoffaufnahme bei Kindern und Jugendlichen	
<i>A. Menrath</i>	301
Schöne neue Welt - So können Sie fehlende SAS-Funktionalitäten mit PROC FCMP nachrüsten	
<i>S. Meyer, F. Fritz, C. Weinhardt</i>	311
Modellierung von Optionspreisen mit PROC FCMP	
<i>C. Ortseifen</i>	321
Der Datenschnitt - Ein mächtiges Werkzeug innerhalb der SAS-Umgebung	
<i>H. Ramroth</i>	337
Wo ist Mister X? oder: Welche Variablen sind (wie) belegt?	
<i>S. Reimann</i>	341
Geht nicht? Gibt's nicht! Daten lesen mit dem Data Step	
<i>J. Schmidtke, W. Mönkemeyer, K. Schmidt</i>	357
Statistische Auswertungen für Anwender ohne SAS Programmierkenntnisse	
<i>S. Schneider, H. Schmidli, T. Friede</i>	369
Fallzahlplanung für klinische Studien mit Zähldaten	
<i>G. Svolba</i>	379
Rechtzeitig mit SAS ein Bild über die Qualität der Analysedaten erhalten	
<i>G. Svolba</i>	389
Grundlagen und Anwendungen von Strukturgleichungsmodellen und das zugehörige SAS/JMP Benutzerinterface	
<i>T. Vogelmann, T. Schubert</i>	395
Matching mit den vorhandenen Anwendungsroutinen des SAS Enterprise Guide	
<i>P. R. Warnat</i>	409
Neuerungen in SASUnit, insbesondere Ermittlung der Testabdeckung	

<i>A. Welz, R. Bender</i>	417
Punkt- und Intervallschätzung adjustierter NNT-Maße im Cox-Modell mit Hilfe von SAS	
<i>M. Wern, S. Daum</i>	427
PROC FREQ erklärt (mit) Fußball	