

**Österreichische Statistiktage 2009**  
**Annual Meeting of the Austrian Statistical Society 2009**  
**Submissions**

---

1. Rainer W.A. Alexandrowicz
2. David J. Berri, Rob Simmons
3. Elmar Bittner, Andreas Nußbaumer, Wolfhard Janke, Martin Weigel
4. Angela Bohn, Ingo Feinerer, Kurt Hornik, Patrick Mair, Stefan Theußl
5. Monika Brunauer, Elisabeth Gruber
6. Karl Danzer
7. Daniel Dekic, Helmut Erb
8. Gerlinde Dinges, Martin Haitzmann
9. Antonia Egerer
10. Karl Fröschl, Ron Stockinger, Hannes Ellinger
11. Elmar Fürst
12. Richard Gisser
13. Timo Gnambis
14. Alex Gonzaga, Michael Hauser
15. Kathrin Gruber, Klaus Kubinger
16. Christine Hohensinn, Klaus D. Kubinger
17. Harald Katzmair, Ruth Pfosser
18. Johannes Klotz
19. Ingrid Koller, Sabine Strauß, Judith Glück
20. Monika Koller
21. Thomas Ledl
22. Christoph Leitner, Achim Zeileis, Kurt Hornik
23. Manuela Lenk
24. Marco Maier, Rainer W.A. Alexandrowicz
25. Ivo Ponocny, Elisabeth Ponocny-Seliger
26. Steven P. Reise
27. Kristina Riedmüller
28. Thomas Salzberger
29. Eva Schlenker
30. Tomáš Sobotka, Kryštof Zeman
31. Zsolt Speder
32. Sabine Stadler
33. Matthias Till, Johannes Kotz
34. Daniela Weber, Eva Lienbacher, Wolfgang Ziniel, Patrick Mair
35. Stefan Wiesel
36. Marcus Wurzer

## **Robustness of the LRT to compare the LLTM with the RM**

Rainer Alexandrowicz  
University of Klagenfurt, Psychology Department  
Applied Psychology and Methods Research Unit

One important way to assess fit of the Linear Logistic Test Model (LLTM) is to test it against the Rasch Model (RM) by means of a Likelihood Ratio Test (LRT). But practical application shows this LRT to become significant rather easily. This raises the question of whether it either lacks robustness (i.e. fails to maintain the chosen niveau alpha) or is overly powerful to even slightest – and therefore irrelevant – deviations of model assumptions. By means of a simulation study (performed with R) we could identify situations where the LRT reproducibly showed nonrobust results.

## **Mixing the Princes and the Paupers**

David J. Berri, Department of Economics and Finance, Southern Utah University, Cedar City, UT, USA  
Rob Simmons, Department of Economics, The Management School, Lancaster University, Lancaster, UK

This study offers a test of the relationship between pay inequality and worker performance in two field settings where performance of individuals and organisations can be observed. These settings are Major League Baseball and the National Basketball Association. Although sports leagues have been used before to test for the relationships between pay and worker performance, previous studies have tended to focus on just one sports league. Consequently the results were not easy to generalize to other sports leagues, or even beyond the world of sports. In contrast, our study covers two sports leagues with different team and roster sizes, and also differences between the role player interactions play in team performance. The more comprehensive nature of our study suggests that the inferences drawn will be more general.

Certainly the issue we examine is important to any firm that employs a team of workers. Firms must decide how to structure the pay of its employees. On the one hand, one could choose for a compressed salary structure. Such a structure could foster feelings of team cohesion, which can lead to greater levels of output. An alternative hypothesis is that wider pay dispersion induces greater effort by workers as they aspire to achieve 'star' status and a corresponding high salary. This latter hypothesis emerges from tournament theory developed within personnel economics.

The literature on pay dispersion and performance in team sports has tended to focus on team performance. Yet the mechanism by which team performance is affected by pay dispersion works through individuals' motivation and effort. Hence, we apply our empirical analysis to individual player performance and show how this varies with measures of dispersion. We model various measures of player performance in each sport with a full set of controlling explanatory variables, such as past player performance, the performance of teammates, plus non-performance indicators such as player age and injury status. Conditional on these control variables, we then test for impact of pay dispersion on player performance, which according to theory can be positive, negative or zero. Pay dispersion is captured by a number of measures including not just the Gini coefficient (standard in analysis of pay inequality), but also skewness and kurtosis of the pay distribution, to pick up the possible influence of 'superstar' outliers at the top of the pay distribution.

We have extensive data for our empirical analysis, covering 12 seasons each of baseball and basketball. The analysis of baseball is separated between pitchers and batters, reflecting the extensive division of labour in this sport. Our results suggest some support for the hypothesis that increased pay dispersion helps raise player performance. Hence, the predictions of tournament theory are supported rather than the predictions of 'pay cohesiveness' theory.

## **Football fever: goal distributions and non-Gaussian statistics**

Elmar Bittner 1, Andreas Nußbaumer 1, Wolfhard Janke 1, and Martin Weigel 2

1-Institut für Theoretische Physik, Universität Leipzig

2-Institut für Physik, Johannes Gutenberg Universität Mainz

Analyzing football score data with statistical techniques, we investigate how the not purely random, but highly co-operative nature of the game is reflected in averaged properties such as the probability distributions of scored goals for the home and away teams. As it turns out, especially the tails of the distributions are not well described by the Poissonian or binomial model resulting from the assumption of uncorrelated random events. Instead, a good effective description of the data is provided by less basic distributions such as the negative binomial one or the probability densities of extreme value statistics. To understand this behavior from a microscopical point of view, however, no waiting time problem or extremal process need be invoked. Instead, modifying the Bernoulli random process underlying the Poissonian model to include a simple component of self-affirmation seems to describe the data surprisingly well and allows to understand the observed deviation from Gaussian statistics. The phenomenological distributions used before can be understood as special cases within this framework. We analyzed historical football score data from many leagues in Europe as well as from international tournaments, including data from all past tournaments of the "FIFA World Cup" series, and found the proposed models to be applicable rather universally. In particular, here we analyse the results of the German women's premier football league and consider the two separate German men's premier leagues in the East and West during the cold war times and the unified league after 1990 to see how scoring in football and the component of self-affirmation depend on cultural and political circumstances.

Fußballfieber: Torverteilungen und nicht-Gaußsche Statistik Mit Hilfe von statistischen Methoden analysieren wir Fußballergebnisse und untersuchen, wie sich die kooperative Natur des Spiels rein zufällig, in den gemittelten Eigenschaften auswirkt, wie z. B. in der Verteilung der Anzahl der Tore der Heim- bzw. der Gastmannschaft. Es zeigt sich, dass die Verteilungen und insbesondere die Flanken nicht durch das Poisson- oder Binomialmodell beschrieben werden können, dem unkorrelierte Zufallsereignisse zu Grunde liegen. Stattdessen können die Daten mittels weniger elementarer Verteilungen, wie beispielsweise der negativen Binomialverteilung oder der Extremwertverteilung, beschrieben werden. Um dieses Verhalten mikroskopisch zu verstehen, bedarf es dabei jedoch weder eines Extremal noch eines Warteprozesses. Mittels eines modifizierten Bernoulliprozesses, der aus einem Poissonmodell und einer einfachen Komponente der Selbstverstärkung ("selfaffirmation") besteht, können die Abweichungen zur Gaußschen Statistik erklärt werden. Die bisher verwendeten phänomenologischen Verteilungen ergeben sich dann als Spezialfälle unseres Systems. Wir haben historische Fußballergebnisse vieler europäischer Ligen und internationaler Turniere analysiert, z. B. alle bis 2002 gespielten FIFA Weltmeisterschaften, und finden, dass unsere vorgeschlagenen Modelle sehr universell anwendbar sind. Im einzelnen haben wir die Ergebnisse der deutschen Frauen-Bundesliga, der höchsten Spielklassen der Männer in der BRD und DDR während des Kalten Krieges und der deutschen Bundesliga nach der Wiedervereinigung 1990 betrachtet. Dabei wurde unter anderem untersucht, ob der Selbstverstärkungseffekt von kulturellen und politischen Umständen abhängt.

## **A Network Analysis of E-Mail Texts**

Angela Bohn 1\*, Ingo Feinerer 2, Kurt Hornik 1, Patrick Mair 1, Stefan Theußl 1

1 - WU Vienna

2 - Technische Universität Wien

\* Contact author: angela.bohn@gmail.com

R is a well-known open source software for statistical computing and graphics. Hundreds of developers from all over the world create software packages via an open collaborative approach. Not only the code of R packages is open, but also part of the communication between developers and users. Developers use the R mailing list R-devel to discuss topics related to the creation of new features, while users post questions on the R-help mailing lists. Thousands of authors write dozens of e-mails daily and their findings and information is shared not only with the subscribers, but also with even more Internet users. Social network analysis (SNA) is able to reveal the writers' communication structure and find behavioral patterns. What's more, text mining (TM) allows examining the content of the great amount of e-mails. Despite the great potential, only few approaches to combine SNA and TM exist so far. In our poster, we show how such a combination, more precisely a "Network Text Analysis" of the R-help and R-devel mailing list, can help to gain even more insights into the process of open source development.

### References:

[1] Jana Diesner and Kathleen M. Carley, Using network text analysis to detect the organizational structure of covert networks, Proceedings of the North American Association for Computational Social and Organizational Science Conference, 2004.

[2] Ingo Feinerer, Kurt Hornik, and David Meyer, Text Mining Infrastructure in R, Journal of Statistical Software 25 (2008), no. 5, 1-54.

[3] Andrew McCallum, Andres Corrada-Emmanuel, and Xuerui Wang, Topic and Role Discovery in Social Networks, IJCAI '05: Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence, 2005, pp. 786-791.

[4] Stanley Wasserman and Katherine Faust, Social Network Analysis, Methods and Applications, Cambridge University Press, 1997.

Keywords: Social Network Analysis, Text Mining, Network Text Mining, R Mailing Lists

**ÖNACE 2008.**  
**Herausforderungen an die Konjunkturindikatoren.**  
**Industrie. Bau. Handel. Dienstleistungen.**

Monika Brunauer, Elisabeth Gruber  
Statistik Austria

Seit 1. Jänner 2008 ist in der Europäischen Union die revidierte Wirtschaftstätigkeitenklassifikation, NACE Rev.2, die die NACE Rev.1.1 ablöst, anzuwenden. Die Konjunkturindikatoren für die Bereiche Industrie, Bau, Handel und Dienstleistungen, die eine wichtige Informationsquelle für die Marktbeobachtung und Marktforschung darstellen, sind ab 2009 nach der neuen **ÖNACE 2008**, die österreichische Version der NACE Rev.2, zu erstellen.

Im Vortrag wird auf die Konjunkturindikatoren im Produzierenden Bereich (Produktionsindex, Umsatzindex, Auftragseingangsindex, Arbeitsinputindikatoren) sowie im Bereich Handel und Dienstleistungen (Umsatzindex, Beschäftigtenindex), die durch die Umstellungsarbeiten auf die ÖNACE 2008 betroffen sind, eingegangen. Die Anwendung dieser neuen Klassifikation sowie die Bereitstellung von zurückgerechneten Zeitreihen stellte eine große Herausforderung an die nationale Statistik dar.

Präsentiert werden die Arbeiten, die mit der Berechnung der Konjunkturindikatoren im Produzierenden Bereich sowie im Bereich Handel und Dienstleistungen auf Basis der neuen Wirtschaftstätigkeitenklassifikation erforderlich waren. Die Berechnungsweise der Indizes blieb dabei konzeptionell unverändert. Da für viele Anwender lange Zeitreihen von großer Bedeutung sind, bildet die Rückrechnung der Zeitreihen nach der ÖNACE 2008 einen inhaltlichen Schwerpunkt.

**Konjunkturindikatoren Produzierender Bereich**

Die für den Zeitraum von 1996 bis 2004 auf der Basis 1995 bzw. 2000 vorliegenden alten Indexreihen wurden auf die neue Wirtschaftstätigkeitenklassifikation für sämtliche Gliederungsstufen mit Hilfe von Konversionskoeffizienten auf Ebene der 3-Steller der ÖNACE umgerechnet, neu aggregiert und auf das Basisjahr 2005 umbasiert. Ab dem Berichtsmonat Jänner 2005 erfolgte eine komplette Neuberechnung der Indizes auf der Grundlage von Einzeldaten. Somit stehen nun rückwirkend bis zum Jahr 1996 Indizes auf der Basis 2005 gemäß der ÖNACE 2008 zur Verfügung.

**Konjunkturindikatoren Bereich Handel und Dienstleistungen**

Die Umsatz- und Beschäftigtenindizes wurden für den Handel bis zum Jahr 2000 und für die Dienstleistungen bis zum Jahr 2003 zurückgerechnet. Während für die Konversion der Umsatzindizes von ÖNACE 2003 auf ÖNACE 2008 ein Makroansatz auf Basis von Konversionskoeffizienten gewählt wurde, wurden die Beschäftigten auf Mikroebene nach der ÖNACE 2008 neu berechnet. Neben einigen konzeptionellen Änderungen wurde die Konjunkturstatistik Handel und Dienstleistungen auch auf die Basis 2005 umgestellt. Ziel der Präsentation ist, die Methoden und Vorgehensweisen für die Berechnung der Konjunkturindikatoren nach der ÖNACE 2008 näher zu bringen.

Stichwörter: Wirtschaftsstatistik, ÖNACE 2008, Revision.

## **Modellierung von Odds für den Wettbereich „Wer gewinnt den Rest des Spiels?“ in laufenden Fußballspielen**

Karl Danzer

Sportwetten auf laufende Fußballspiele haben in den vergangenen Jahren eine immer höhere Popularität erfahren. Aus diesem Grund ist es für Buchmacher sehr wichtig, die möglichen Ereignisse unter Berücksichtigung der Gegebenheiten, die den Spielverlauf beeinflussen, auch während eines Spiels möglichst genau vorherzusagen.

Der Vortrag beschäftigt sich mit der Entwicklung eines Modells zur Vorhersage von Wahrscheinlichkeiten für das Ereignis „Wer gewinnt den Rest des Spiels?“ in bereits laufenden Fußballspielen. Aufgrund von Daten von mehr als 30 000 Fußballspielen wird der Ausgang dieses Wettbereichs mittels Ordered-Probit-Modellen geschätzt. Als erklärende Variablen dienen einerseits Größen, die die Stärke der beiden Mannschaften widerspiegeln und andererseits Einflussfaktoren, die den aktuellen Spielverlauf beeinflussen, u.a. Spielminute und Spielstand.

Es wird dargestellt, welche Variablen einen Einfluss haben und wie sich diese auf das Modell auswirken. Zusätzlich wird untersucht, wie sich die Einflussfaktoren im Laufe eines Fußballspiels verändern. Es zeigt sich, dass der Einfluss der Spielstärke mit zunehmender Spieldauer kontinuierlich abnimmt und im Gegensatz dazu die Faktoren, die den aktuellen Spielverlauf beschreiben, zunehmen.

## **Open Source Framework in Pharmaceutical Industry**

Daniel Dekic & Helmut Erb

There are several methods in reporting statistical analysis data in pharmaceutical industry. A common simple method is to use text processing software like MS Word followed by importing tables, figures and listings, which are generated with other statistical software, usually with SAS, e.g. by using copy and paste. This method is characterized by investing a lot of time in creating the layout of the tables and figures, and is frequently subject of transcription errors.

Another reporting environment that provides a comprehensive reporting tool is R<sup>1</sup> in combination with Sweave<sup>2</sup>. Sweave allows R code to be included in a LaTeX file. Reporting can be done automatically with the advantage of a document markup language and high quality typesetting. The final result is a pdf file, which has an easy navigable table of content, list of figures and tables via hyper-references. Other documents can be embedded easily.

R is also compliant with pharmaceutical regulations, e.g. 21 CFR - Part 11<sup>3</sup> is in combination with Sweave a very powerful and impressive tool for reporting of non-interventional observational data.

In this talk we will present the implementations of R and Sweave in a contract research organization (CTM GmbH) in Vienna who uses R and Sweave for reporting of non interventional observational data. With only basic knowledge of R and LaTeX reports, which are equivalent or even superior to the old fashioned way, can be generated in much less time and with much less effort.

1 - R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>

2 - Leisch F (2002). "Sweave, Part I: Mixing R and LaTeX: A short introduction to the Sweave file format and corresponding R functions". R News 2 (3): 28-31.

3 - R: Regulatory Compliance and Validation Issues A Guidance Document for the Use of R in Regulated Clinical Trial Environments

## **Modellbasierte Ergänzung in der Konjunkturstatistik zur Darstellung der statistischen Grundgesamtheit im Produzierenden Bereich**

Gerlinde Dinges und Martin Haitzmann, Statistik Austria

Die im Jahr 1996 eingeführte EU-harmonisierte, monatliche Konjunkturstatistik im Produzierenden Bereich (KJP) stellt eine der zentralen Informationsquellen zur Beurteilung der konjunkturellen Entwicklung Österreichs und des gesamten Europäischen Wirtschafts- und Währungsraumes dar. Ihre nationalen und internationalen Verwendungszwecke sind vielfältig.

Dem Anliegen der Wirtschaft entsprechend wird die KJP als Konzentrationserhebung (Vollerhebung mit variablen Abschneidegrenzen) geführt, wodurch Klein- und Kleinstunternehmen (rund 80% der Grundgesamtheit) von statistischen Verpflichtungen weitestgehend befreit werden.

Um auch den Bedürfnissen der Datennutzer zu entsprechen und EU-Vorgaben hinsichtlich vorgegebener Repräsentanzkriterien vollständig zu erfüllen, wird ab 2009 neben der bisherigen Veröffentlichung primär-statistisch gewonnener Ergebnisse der meldepflichtigen Unternehmen und Betriebe auch eine entsprechende Darstellung für die statistische Grundgesamtheit des Produzierenden Bereiches erfolgen. Hierfür werden die monatlich erhobenen Primärdaten der KJP um die im Unternehmensregister der Statistik Austria erfassten, nicht meldepflichtigen Einheiten ergänzt. Die vorgegebenen Rahmenbedingungen lassen keine konventionelle Hochrechnung zu, weshalb in Zusammenarbeit von Unternehmensstatistikexperten und Methodikern der Statistik Austria ein modellbasierter Ansatz entwickelt wurde, welcher Merkmalszusammenhänge der monatlichen Konjunkturerhebung einbindet und Informationen aus Verwaltungsquellen effizient nutzt. Ein ähnlicher Ansatz wird seit mehreren Jahren bereits in der jährlich durchgeführten Leistungs- und Strukturstatistik angewandt, wo Ergebnisse der Primärstatistik, Informationen des Unternehmensregisters der Statistik Austria, Beschäftigtendaten der Sozialversicherung und Umsatzsteuerinformationen der Finanzbehörde in die Berechnungen eingehen. Die ungleich höhere Komplexität, welche der Konjunkturstatistik ins-besondere aufgrund der erforderlichen Ergebnisdarstellung nach Aktivitäts- und Güteransatz zugrunde liegt, gestaltet die Erstellung eines plausiblen Datenkörpers für diese statistische Grundgesamtheit jedoch als besondere Herausforderung.

Der Vortrag ist vor allem der methodischen Konzeption des Schätzmodells und dessen Umsetzung in der Praxis gewidmet, wobei auch kurz auf konzeptionelle Vorgaben, Entscheidungsfindung und mögliche Modelleffekte eingegangen wird. Darüber hinaus werden Ergebnisse und Datenqualität auch aus Sicht der Unternehmensstatistikexperten beleuchtet. Zur praktischen Veranschaulichung wird der Vortrag durch graphische Darstellungen ausgewählter Wirtschaftszweige und Merkmale der aktuellen Konjunkturstatistik 2009 ergänzt.

Keywords: Modellbasierte Schätzung, Wirtschaftsstatistik, Verwaltungsdaten.

## **Verwaltungsdaten über innergemeinschaftliche Warenverkehre und ihre Verwendbarkeit zu Imputationszwecken in Intrastat**

Antonia Egerer, Statistik Austria

Seit dem Beitritt Österreichs zur Europäischen Union basiert die Außenhandelsstatistik auf zwei unterschiedlichen Erhebungssystemen. Während der Handel mit den Ländern außerhalb der Europäischen Union weiterhin über das Zollsystem erfasst wird (Extrastat-System), werden die Daten über den innergemeinschaftlichen Warenverkehr im Rahmen einer Erhebung bei den Unternehmen (Intrastat-System) direkt von der Statistik Austria gesammelt.

Die Intrastat-Erhebung zielt darauf ab die Respondenten weitest möglich zu entlasten, wobei gleichzeitig die Qualität der Außenhandelsstatistik gewährleistet werden muss. Die Entlastung erfolgt durch die Anwendung eines Schwellensystems, das entweder eine vollständige Befreiung von der Abgabe der Intrastat-Meldung oder zumindest eine Reduktion des Umfangs der zu liefernden Informationen im Rahmen der Meldung nach sich zieht. Die sogenannte Assimilationsschwelle legt je Verkehrsrichtung die jährliche Obergrenze des Binnenhandelsumsatzes fest, welche Unternehmen von der Meldepflicht befreit. Liegt der Handel eines Unternehmens unterhalb der Assimilationsschwelle oder kommt ein Unternehmen seiner Meldepflicht nicht nach, ist die Statistik Austria auf Grund der EU-Gesetzgebung dazu verpflichtet den fehlenden Handel zu imputieren. Da das Intrastat-System eine enge Anbindung an das UmsatzsteuerSystem für den EU-Binnenhandel aufweist, erfolgt die Imputation auf Grundlage von Steuerdaten der Finanzbehörde, welche der Statistik monatlich zur Verfügung gestellt werden.

Vor dem Jahr 2004 wurden Daten aus dem EU-weiten MehrwertsteuerInformationsaustausch System (MIAS; engl. VIES) für die Imputation dieses fehlenden Volumens herangezogen. Die Daten von den Unternehmen an die österreichische Finanzbehörde gesendeten Zusammenfassende Meldung (ZM) über ihre innergemeinschaftlichen Versendungen lagen quartalsweise und auf Länderebene vor. Die auch auf Quartalsebene vorliegenden innergemeinschaftlichen Eingänge hingegen wurden von den Spiegeldaten der Partnermitgliedstaaten abgeleitet und standen daher erst etwa ein halbes Jahr nach Ablauf des Berichtszeitraumes der Statistik zur Verfügung. Da Aktualität ein wichtiges Kriterium für die Außenhandelsstatistik darstellt, barg diese Quelle somit Nachteile in sich. Seit 2004 stellte man die Schätzmethode auf Daten aus der damals neu eingeführten nationalen UmsatzsteuerVoranmeldung (UVA) um. Diese Datenquelle wurde auf monatlicher Basis erhoben und war wesentlich früher als Daten aus dem MIAS verfügbar.

Da weitgehend alle umsatzsteuerpflichtigen Unternehmen eine UVA abgeben müssen, ist es möglich auf Basis dieser Quelle auf den innergemeinschaftlichen Gesamthandel zu schließen. Die EU-Verordnungen schreiben vor die Schwellenwerte für die Befreiung von der Meldepflicht in den Mitgliedsstaaten derart zu berechnen, dass ein Mindesterfassungsgrad bei den Eingängen von 95%, bei den Versendungen von 97% gewährleistet ist. Die Verfügbarkeit von Steuerdaten ist eine wichtige Voraussetzung um diesen gesetzlichen Anforderungen zu entsprechen und in weiterer Folge für die Befreiung möglichst vieler Transakteure, insbesondere KMUs, von der Meldepflicht zu sorgen.

Die Informationen auf Unternehmensebene, welche der UVA für den EU-Binnenhandel entnommen werden können, sind zum einen das Gesamthandelsvolumen sowie die jeweilige Handelsrichtung. Die Zuschätzungen werden auf Grundlage dieser Steuerdaten in einem umsatzhöhen und branchenspezifisch geschichteten Hochrechnungsverfahren in tiefster Gliederung durchgeführt, was bedeutet, dass sie auch auf Produktebene der Warennomenklatur (KN8) verfügbar sind. Die Imputation wird bis zum Vorliegen der endgültigen Ergebnisse monatlich auf Basis der neu eingegangenen Intrastat und Umsatzsteuermeldungen revidiert, um den Außenhandel auf dem aktuellsten Stand abzubilden.

Der vorliegende Vortrag skizziert die Bedeutung von Verwaltungsdaten für die Sicherstellung der Ergebnisqualität und die Entlastung der Respondenten anhand ihrer praktischen Verwendung in der Außenhandelsstatistik.

## **InfoRose – ein Modell zur navigativen Exploration statistischer Dokumentationsräume**

Karl Fröschl<sup>1</sup>, Ron Stockinger<sup>2</sup>, Hannes Ellinger<sup>3</sup>

<sup>1</sup> - Universität Wien, Institut für Scientific Computing, karl.anton.froeschl@univie.ac.at

<sup>2</sup> - IMP Nonlinear Technologies, Wien/Timişoara, <http://www.new-organizers.com>

<sup>3</sup> - Software-Engineering, kopfkraft@gmail.com

Seit nunmehr einigen Jahrzehnten (vgl. Sundgren, 1973) wird die Bedeutung der Dokumentation statistischer Datenbestände sowohl in akademischen als auch fach- und amtsstatistischen Kreisen (z.B. IMF, EUROSTAT, UN/ECE, usw.) intensiv diskutiert, wobei die Rolle der Dokumentation durch zunehmende Datenvernetzung und damit einem verbreiteten Informationszugang im Sinne besserer (Sekundär-) Nutzung statistischer Daten beständig an Relevanz gewinnt.

Neben der Entwicklung tauglicher dokumentativer Strukturen für umfangreiche, thematisch komplexe statistische Datenkörper (etwa einer nationalen oder internationalen Statistikagentur; vgl. z.B. Colledge, 1999) stellt sich allerdings auch die Frage, wie derartige formalisierte Dokumentationsbestände – die i.W. den Charakter semantischer Netzwerke, d.h. Modelle inhaltlicher Beziehungsgeflechte aufweisen (vgl. z.B. Denk et al., 2004) – nutzerseitig praktisch erschließbar sind.

In dieser Hinsicht sind im Bereich der HCI-Forschung verschiedene Ansätze zur navigativen Exploration von Netzwerkstrukturen vorgeschlagen worden (z.B. Keller und Hall, 2003), die kognitionspsychologisch von einer Verräumlichung assoziativer Strukturen ausgehen, sodass die visualisierten lokalen „Nachbarschaften“ begrifflicher Konzepte interaktiv schrittweise erkundet werden können. Ein anschauliches kommerzielles Beispiel derartiger Interaktionsformen bietet etwa „The Brain“<sup>®</sup> ([www.thebrain.com](http://www.thebrain.com)).

Der Beitrag entwickelt mit *InfoRose* das assoziative Navigationsmodell interaktiver Informationsräume unter Einführung zusätzlicher Kohärenzbedingungen der Informationsaufbereitung nach kognitionspsychologischen Kriterien weiter und bringt beispielhafte Illustrationen dieses Konzepts im Kontext der komplexen Anwendung statistischer Datendokumentation, deren spezifische Anforderungen die Nützlichkeit dieser visuellen Form des Informationszugangs verdeutlichen.

### Literatur (Auswahl)

Colledge M.J. Statistical Integration Through Metadata Management. *Int. Statistical Review* 67 (1): 79-98, 1999.

Denk M. et al. KnoWeb: An Operational Knowledge Organisation Framework. *Int. J. of Electronic Business* 2 (5): 493-507, 2004.

Keller R.M., Hall D.R. Developing Visualization Techniques for Semantics-based Information Networks. Proc. K-CAP '03 (ACM), Sanibel Island, Florida, USA, 2003.

Sundgren Bo, An Infological Approach to Data Bases. Report, Statistics Sweden, 1973.

Abstract Statistiktage 2009: 30. April 2009

Stichworte: Amtliche Statistik, Dokumentation/Metadaten, Visualisierung

## **Verkehrserhebungen von morgen: Anforderungen an die Amtliche Verkehrsstatistik**

Elmar Fürst, Statistik Austria

Verkehr ist ein Phänomen, welches jeden einzelnen direkt oder indirekt betrifft, entweder durch die eigene Teilnahme am Verkehrsgeschehen (Fahrt zur Arbeit, Einkaufen, Freizeit usw.) oder durch Konsum von Waren und Dienstleistungen für deren Erzeugung Verkehr entstanden ist. Zudem verursacht der Verkehr – neben dem großen wirtschaftlichen und gesellschaftlichen Nutzen, den er stiftet – auch eine Reihe an unerwünschten negativen Auswirkungen (Unfälle, Abgase, Staus, Lärm usw.). Folglich gibt es großen und unterschiedlich gearteten Bedarf an umfassenden, detaillierten und aktuellen Verkehrsdaten. Längerfristige (politische und wirtschaftliche) Entscheidungen basieren ebenso auf Verkehrsdaten wie kurzfristige Entscheidungen (etwa die konkrete Routenwahl aufgrund eines taktischen Verkehrslagebildes).

Aufgrund der zahlreichen unterschiedlichen Interessenlagen existiert auch eine Reihe unterschiedlicher Erhebungen im Bereich des Verkehrs, welche von verschiedensten Organisationen und Institutionen teils ad-hoc, teils regelmäßig bzw. laufend durchgeführt werden. Die Amtliche Verkehrsstatistik (genauer Verkehrsstatistiken) stellt in diesem System eine bedeutende und für die Nutzer wesentliche Datenquelle dar, auch wenn, bezogen auf die einzelnen Verkehrsträger Schiene, Straße, Wasserwege und Luft unterschiedliche rechtliche Rahmenbedingungen, Methoden, Konzepte und Erhebungsformen zum Tragen kommen und nicht der gesamte Verkehr abgebildet werden kann.

Der Nutzerkreis der Amtlichen Verkehrsstatistik ist sehr umfangreich und heterogen. Er reicht von Bundes- und Landesbehörden über Interessensvertretungen (WKO, AK, div. Vereine), Universitäten und Forschungseinrichtungen, bis hin zur internationalen Ebene (Eurostat, Vereinte Nationen, Internationales Transportforum). Auch Unternehmen, Medien und Private zählen unter anderem zu den an Verkehrsstatistik interessierten Nutzern.

Wirtschafts- und damit Verkehrssysteme sind jedoch zahlreichen Änderungen und dynamischen Entwicklungen unterworfen. Technische Weiterentwicklungen bringen ebenso neue Möglichkeiten der Datenerhebung und -verarbeitung mit sich, aber auch datenschutzrelevante Aspekte werden dadurch berührt. Dadurch verändert sich auch das Interesse und das Anforderungsprofil an Verkehrserhebungen im Allgemeinen und die Amtliche Verkehrsstatistik im Besonderen.

In einer Studie, welche das Institut für Transportwirtschaft und Logistik der WU-Wien derzeit durchführt, werden – basierend auf zahlreichen Experten- bzw. Nutzergesprächen – die Sichtweise auf die derzeitigen Erhebungen sowie die zukünftigen Anforderungen an die Verkehrsstatistiken erhoben. Im Rahmen des Vortrags sollen die wesentlichen Ergebnisse dieser Studie vorgestellt sowie Empfehlungen formuliert werden, welche Weichenstellungen zum jetzigen Zeitpunkt dazu beitragen könnten, dass die Amtlichen Verkehrsstatistiken auch in Zukunft den Wünschen und Erfordernissen der Nutzer und Anwender der Daten zu entsprechen vermögen.

## **Levels and Trends of Fertility in the Austro-Hungarian Monarchy**

Richard Gisser, Vienna Institute of Demography

The purpose of this paper is to give a background and starting point for presentations and discussion of fertility developments in Central Europe during the 20<sup>th</sup> century, in particular their communalities or diversity. It will focus on the period 1880-1910 and on differences between regions of the Danube Monarchy as building bricks for the successor states of the Austro-Hungarian Empire. The paper will also deal with contemporary data availability and with measurement issues in order to develop a link to modern approaches of fertility analysis.

## **Adaptive proficiency estimators in polytomous computer-adaptive testing**

Timo Gnambs, University of Linz

Computerized adaptive test applications seek to yield unbiased estimates of a persons' proficiency level with as few items as possible. For this purpose several proficiency estimators have been developed. In addition to (weighted) maximum likelihood estimators (ML, WML), different bayesian approaches like the modal a posteriori estimator (MAP) are currently in use. Although bayesian estimators have a couple of advantages (e.g. concerning test efficiency), they yield more biased estimates, the more the true proficiency differs from the mean of the prior distribution (shrinkage toward the mean). Raïche and Blaise (2002) recently proposed an alternative adaptive bayesian estimator (AMAP1) for dichotomous items, which does not assume a constant mean of the prior distribution (usually 0), but substitutes the proficiency estimate obtained from the previously administered items.

The current simulational experiment extends this approach to the polytomous case and examines the efficiency of the adaptive proficiency estimator for the partial credit model. Additionally an alternative adaptive estimator (AMAP2) is proposed, which not only adjusts the prior mean, but as well the prior variance based on the standard error of the interim proficiency estimate. As item pool characteristics as well as test termination rules can strongly influence the performance of the proficiency estimations, the proposed estimators are studied for simulated responses to a pool of 60 items under different CAT-conditions, yielding a 5 (proficiency estimation method: ML, WML, MAP, AMAP1, AMAP2) x 3 (maximum item pool information at proficiency of: -1, 0, 1) x 3 (test termination after 5, 10 or 20 items) experimental design. Preliminary results indicate that estimators with adaptive means lead to slightly higher standard errors, but a significantly reduced bias for more extreme proficiency levels. Adaptive variances do not seem to yield comparable improvements.

## **Estimating and Forecasting Generalized Long-memory Stochastic Volatility**

Alex Gonzaga a, b

Michael Hauser b

a - Department of Physical Sciences and Mathematics, University of the Philippines, Manila, Philippines

b - Department of Statistics and Mathematics, WU Vienna, Austria

A recent approach in forecasting long-memory stochastic volatility (Deo et al.(2006)) involves the removal of slowly varying seasonality, represented by a linear combination of sines and cosines evaluated at the Fourier frequencies that show seasonal peaks. Forecasts are initially obtained from the deseasonalized series, which are then combined with the corresponding forecasts of the seasonality model to obtain forecast values of the volatility series.

In this paper, we propose an alternative approach based on the generalized stochastic volatility (GLMSV) model (Gonzaga and Hauser(2009)), a generalization of the long-memory stochastic volatility (LMSV) model of Breidt et al.(1998). The GLMSV model also represents a seasonal series with slowly varying seasonality. We propose an approximate maximum likelihood estimator of the GLMSV model using the maximal overlap discrete wavelet transform (MODWT). We show that MODWT-based estimator is more statistically efficient than the standard discrete wavelet transform in estimating the GLMSV parameters. Moreover, it is strongly consistent for all the GLMSV parameters.

We derive the analytic expression of the prediction function of the GLMSV model, and we provide the  $h$ -step prediction error. We investigate the predictive ability of the proposed method on real data – Microsoft stock intraday volatility.

### References:

- Deo R, Hurvich C, Lu Y (2006) Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment. *Journal of Econometrics* 131: 29-58
- Gonzaga A, Hauser M (2009) A wavelet-based Whittle estimator of generalized longmemory stochastic volatility, Working Paper, WU Vienna
- Breidt FJ, Crato N, de Lima P (1998) On the detection and estimation of long memory in stochastic volatility. *Journal of Econometrics* 83: 325-348.

## **Modellperformance des DifficultyPlusGuessing PL Modells im Vergleich mit 1PL und 3PL Modellen**

Kathrin Gruber, Klaus Kubinger  
Faculty of Psychology, University of Vienna

In psychologischen Leistungstests werden häufig Multiple-Choice-Items eingesetzt, ohne besonders auf das damit zusammenhängende Problem der Ratemöglichkeit zu achten, und damit darauf, dass Testpersonen relativ hohe Scores erzielen können, ohne über ein entsprechendes Fähigkeitsausmaß zu verfügen. Ein IRT-Lösungsversuch dafür stellt die Anwendung des Difficulty-plus-Guessing PL Modells (Kubinger & Draxler, 2006).

Im vorliegenden Referat werden nun umfangreiche Simulationsergebnisse, mittels R und dem Package ltm, vorgestellt, mit denen mittels Likelihood-Ratio-Test geprüft werden sollte ab welcher Rate-wahrscheinlichkeit das Difficulty-Plus-Guessing PL Modell dem 1PL und 3PL Modell, bei unterschiedlicher Stichprobengröße und Itemanzahl sowie unterschiedlicher Verteilung der Personenparameter, vorzuziehen ist.

Es wurde auch berechnet, wie oft das Modell ein „Guessing“ bei 1PL Daten entdeckt, obwohl in den Daten keines vorhanden ist und damit einen Fehler 1. Art begeht. Zusätzlich wurde auch noch der Bias der Schätzung unter bestimmten Modellannahmen geprüft.

## **On the variation of item difficulty by changing the response format**

Christine Hohensinn & Klaus D. Kubinger  
Centre of Testing and Consulting, Division of Psychological Assessment and Applied Psychometrics  
Faculty of Psychology, University of Vienna

An important aspect of ability test administration is the decision for a particular response format; there are different kinds of multiple choice formats or the open-response formats. Of course, modifying the response format might change the difficulty of an item.

The national standard tests in Austria for Language Awareness (measuring the knowledge of structure of the German language) and for Mathematics (8<sup>th</sup> grade) administer different test booklets each and use different response formats within each booklet. We consider two kinds of open-response format ("cascet" and "free" response format) and two different multiple choice formats ("1 from 6" meaning a single correct response option plus five distractors and "2 from 5" with exactly two solutions plus three distractors).

Data of the national standard tests of 2008 are analysed. In the Language awareness test for the competence of recognizing the structure of sentences and in the mathematics test for the competence of interpreting numerical results the same item contents were applied by different response formats (within different test booklets). The sample sizes of the data sets contain more than 3000 students for each test.

The impact of different response formats on the item difficulty was examined by the linear-logistic test model (LLTM). IRT analyses were conducted using eRm. Results indicate a complex interaction of response format and item content.

## **Der Ball im Netz. Netzwerkanalytische Methoden im Kontext Fußball.**

Harald Katzmair (FAS.research)  
Ruth Pfosser (FAS.research)

Beziehungen in Wirtschaft, Forschung und Politik (Kooperations-, Beteiligungs-, Wertschöpfungsbeziehungen etc.) können am besten mit dem wissenschaftlichen Instrument der Sozialen Netzwerkanalyse (SNA) untersucht werden. Unter anderem ermöglicht sie die Identifikation der für den Erfolg maßgeblichen SchlüsselspielerInnen eines bestimmten Bereichs.

SNA ist eine integrative und transdisziplinär angelegte wissenschaftliche Disziplin, die es VertreterInnen der unterschiedlichsten Forschungsbereiche erlaubt, die Probleme ihrer jeweiligen Felder in einer gemeinsamen Sprache (nämlich der Sprache der Mathematik) zu formulieren und zu bearbeiten. Die besondere Stärke der SNA liegt in ihrer Visualisierungstechnik, durch die komplexe Zusammenhänge weitaus anschaulicher dargestellt werden können als mit gängigen Statistiken und Diagrammen.

Auch ein Fußballmatch kann als Netzwerk aufgefasst werden. Die Verbindung zwischen den Spielern ist der Ball, es ist die Relation „wer passt zu wem?“, die die Struktur des Spiels und die Schlüsselspieler enthüllt. FAS.research analysiert seit Jahren (unter anderem für die Tageszeitungen *Der Standard* und die *Basler Zeitung*) die Spiele der österreichischen und der Schweizer Nationalmannschaft und beantwortet Fragen wie: Welche Strategien wurde verfolgt? Welche Spieler bilden stabile Pass-Dreiecke und bringen damit den Ball nach vorne? Betrachtungen über einen längeren Zeitraum hinweg ermöglichen Aussagen über die Entwicklung und das Potential einer Mannschaft.

Weiters untersucht FAS.research die Globalisierungsphänomene im Sport: Aus welchen Ländern stammen die Spieler der großen Fussball-Nationen? Welche Länder sind die Exporteure und Importeure von Spielern? Wie sieht das Transfernetzwerk der Clubs aus? Und schließlich: welchen Karrierepfaden folgen die einzelnen Spieler? – Die netzwerkanalytischen Antworten auf diese Fragen wurden von FAS.research in der großen Ausstellung „herz:rasen“ zur EURO 2008 im Künstlerhaus präsentiert.

Unsere Präsentation bei den Österreichischen Statistiktagen 2009 bietet einen Einblick in die verwendeten Methoden und diskutiert die vorläufigen Ergebnisse, die Mängel und Entwicklungsmöglichkeiten der Methode für die Sportanalyse im Fußball.

**Konvergenz oder Divergenz schichtspezifischer Mortalitäts- und Morbiditätsrisiken?**  
**Die Entwicklung von Lebenserwartung und Gesundheitslebenserwartung in Abhängigkeit von der höchsten abgeschlossenen Ausbildung in Österreich seit 1981.**

Johannes Klotz, Statistik Austria

Die demografische Entwicklung Österreichs war in den vergangenen Jahrzehnten charakterisiert durch einen kontinuierlichen Rückgang der Sterberaten in den mittleren und höheren Altersklassen. Als Folge davon erhöhte sich die fernere Lebenserwartung im Erwachsenenalter immer weiter. Gesundheitsdaten zeigen, dass dieser Anstieg der Lebenserwartung vor allem auf eine Zunahme der Lebenszeit in subjektiv guter Gesundheit zurückzuführen ist: Zwar ist die demografische Alterung mit einer verstärkten Prävalenz chronischer Krankheiten verbunden, diese führen aber auf Grund medizinisch-therapeutischer Fortschritte zunehmend seltener zu Einschränkungen im täglichen Leben.

Von großer sozialpolitischer Bedeutung ist in diesem Zusammenhang die Frage, ob bei der beobachteten simultanen Steigerung der Lebenserwartung überhaupt und der Lebenserwartung in subjektiv guter Gesundheit Unterschiede zwischen verschiedenen sozialen Schichten bestehen: Haben sich die Lebenserwartungsdifferenzen nach dem sozio-ökonomischen Status in den vergangenen Jahrzehnten vergrößert oder verringert? Betrifft die Steigerung der Lebenszeit in subjektiv guter Gesundheit obere und untere Schichten der Gesellschaft gleichermaßen?

Eine integrierte Nutzung von Zensus- und Mortalitätsdaten ermöglicht die Erstellung von Sterbetafeln und damit eine Schätzung der Lebenserwartung für einzelne Bildungsgruppen. Weiters stehen aus dem Mikrozensus periodisch Daten zur Beurteilung des Gesundheitszustandes abhängig von Geschlecht, Alter und Bildungsebene zur Verfügung. Mittels der Sullivan-Methode können diese Informationsquellen verbunden und so auch die Lebenserwartung in subjektiv guter Gesundheit für einzelne Bildungsgruppen geschätzt werden. Präsentiert werden Daten zum Zugewinn an Lebenserwartung überhaupt und in subjektiv guter Gesundheit in Abhängigkeit von der höchsten abgeschlossenen Ausbildung zwischen 1981 und 2001. Approximative Verfahren der Varianzschätzung (Methode von Chiang bzw. Mathers) erlauben zudem die Bestimmung der statistischen Signifikanz der Ergebnisse.

## **Item response models for measuring change: A review of different models and their application to real data.**

Ingrid Koller, Sabine Strauß & Judith Glück,  
University of Klagenfurt, Department of Developmental and Educational Psychology

Item response models for measuring change offer interesting possibilities especially for analyzing item-specific intervention effects. Applying classical methods (analysis of variance) change is analyzed on score level. If the items change in different directions you cannot detect this in the data. With Item response models for longitudinal research like the Linear Logistic Test Model (LLTM), the Linear Logistic Model with Relaxed Assumptions (LLRA), hybrid LLRA, and Mixed Rasch models (MRM) it is possible to measure change on person-group level and item level (e.g. Glück & Spiel, 1997, 2007). For example, it is possible to check whether the change is the same across all items or different across groups of items. In this session we want to give a short overview of item response models for longitudinal research.

The application of the models and the discussion about the models are illustrated by means of a real data set: first evaluation for the Dynamic Spatial Ability Test in Augmented Reality (d\*star; FWF Project P19265, 2006-2009) with a complex design: pretest, macro-adaptive training, and two versions of posttests (including link items). The data are analyzed with the package eRm (Hatzinger & Rusch, 2009; Mair & Hatzinger 2007a, 2007b), which is implemented in R 2.9.0 (R-Development core team, 2009). The discussion focused on conditions, advantages, and disadvantages of the various models.

### References:

- Glück, J. & Spiel, C. (1997). Item Response- Modelle für Meßwiederholungsdesigns: Anwendung und Grenzen verschiedener Ansätze. *Methods of Psychological Research Online*, 2 (1). Available: <http://www.pabst-bublishers.de/mpr>.
- Glück, J. & Spiel, C. (2007). Using item response models to analyze change: Advantages and limitations. In A.D. Ong & M.H.M. van Dulmen (Eds.), *Oxford Handbook of Methods in Positive Psychology* (349-361). Oxford: Oxford University Press. [M/E].
- Hatzinger, R. & Rusch, T. (2009). *IRT models with relaxed assumptions in eRm: A manual-like instruction*. *Psychology Science Quarterly*, 51 (1), 87-120.
- Mair, P. & Hatzinger, R. (2007a). CML based estimation of extended Rasch models with the eRm package in R. *Psychology Science*, 49, 26-43.
- Mair, P. & Hatzinger, R. (2007b). Extended Rasch Modeling: The eRm Package for the Application of IRT Models in R. *Journal of Statistical Software [online-serial]*, 20 (9). Available: <http://www.jstatsoft.org/v20/i09>.
- R Development Core Team (2009). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. (URL <http://www.R-project.org>).

## **Alternative measurement of psychological phenomena – advances and applications in neuromarketing**

Monika Koller, WU Vienna, Institute for Marketing Management

Marketing science has always been concerned with the investigation of (psychological) phenomena that are relevant in 'real world' purchase situations. If psychological constructs are to be addressed empirically, sound measures of these constructs are of utmost importance. Very recently, marketing research has recognized the potential of applying neuroscientific methods as an alternative approach in this regard. Methods like functional magnetic resonance imaging (fMRI), electroencephalography (EEG) or magnetoencephalography (MEG) offer a promising avenue to investigate and better understand human emotions and cognitive processes. The major advantages of these techniques, such as an alternative possibility of investigating unconscious phenomena and emotions and a better handling of response bias or measurement effects, are quite evident from a scientific point of view. Psychophysiological information complements traditional methods of investigating human behavior and especially human decision-making. The so called 'black-box' of the consumer is no longer an unreachable sphere'.

Neuroimaging techniques can contribute to fundamental research in terms of a better conceptual understanding of latent constructs in consumer behavior. They explore physiological correlates and thereby contribute to content and construct validity of traditional measures in the social sciences in general and in marketing research in particular. This knowledge will help to overcome the obstacles traditional quantitative marketing research has to deal with and will help to establish sound and valid measures to capture e.g. cognitions, emotions or attitudes.

However, we assume that these approaches cannot fully replace conventional questionnaire-based measurement. In practice, only the latter can provide quantifications of latent variables on a large scale in 'real world'-situations. Psychophysiological knowledge on consumer behavior perfectly fits in a mixed-methods-approach by contributing to both exploratory qualitative research and addressing content validity issues of a latent construct as well as building the base for establishing appropriate measures by alternative measurement models such as e.g. advanced IRT models to overcome the weaknesses that are involved in classical test theory.

## **Wählerstromanalyse – Einfluss der Gemeindecluster auf das Ergebnis**

Thomas Ledl, Universität Wien

Die Methode der Wählerstromanalyse benötigt in zahlreichen Ansätzen als Voraussetzung die Unterteilung des gesamten Wahlgebietes in Teilgebiete, innerhalb derer alle Wähler einer bestimmten Partei bei der vergangenen Wahl bei der neuen Wahl bezogen auf die zur Verfügung stehenden Wahlentscheidungen dieselbe Wahrscheinlichkeitsverteilung haben, was zu in etwa gleich hohen prozentuellen Wählerströmen in den Gemeinden des entsprechenden Wahlgebiets führt. Diese Forderung ist aber nie direkt überprüfbar, weil die Schätzung dieser Wahrscheinlichkeiten gerade die Fragestellung selbst ist.

Bisherige Verfahren verwendeten zumeist eine Gruppierung nach Bundesländern oder nach Wahlergebnissen (Prozentsätzen) bei vergangenen Wahlen. Aufgrund von unterschiedlichen Basis-Modellen verschiedener Anwender ist mit den bisherigen Ergebnissen ein Vergleich der Güte der verschiedenen verwendeten Clustereinteilungen jedoch nicht möglich.

In diesem Vortrag werden die Auswirkungen der Wahl verschiedener Clustereinteilungen auf die Schätzungen und die prognostische Güte an einem Modell gegenübergestellt.

## **Predicting the winner of sports tournaments**

Christoph Leitner, Achim Zeileis, Kurt Hornik  
Department of Statistics and Mathematics, WU Vienna

Every year a lot of great sports tournaments take place all over the world and the ubiquitous question is: "Who is going to win the tournament?"

Here, we investigate a novel method for forecasting tournament outcomes. Unlike many well-established methods it is not based on historical data (such as participants' performances in previous competitions) but on "prospective" data. More precisely, our method employs quoted winning odds from a selection of bookmakers—a type of data that is abundant and easily available due to the increasing popularity of online sports betting.

The new approach utilizes a mixed-effects model for the unknown "true" log-odds for winning the tournament, capturing effects associated with the participants, the bookmakers, as well as potential further variables. In addition to the forecast of the winning odds, methods for assessing the tournament performance and the participants' strengths are discussed.

For illustration, the techniques are applied to the prediction of the EURO 2008 and the UEFA Champions League 2008/09. For the former, our method forecast the correct final and outperformed the FIFA/Coca Cola World rating and the Elo rating.

Keywords: Sports forecasting, tournament, odds, mixed-effects model.

## **Finanzausgleich – jährliche „Mini“-Registerzählungen**

Manuela Lenk, Statistik Austria

Das im Dezember 2007 erlassene Finanzausgleichsgesetz 2008 sieht eine Verwendung der Ergebnisse der Probezählung für die einwohnerzahlabhängige vorläufige Zuweisung von Finanzmitteln an die Gemeinden ab dem Finanzjahr 2009 vor. Es war daher notwendig, die Wohnsitzanalyse, welche Karteileichen und sonstige Fehler im Zentralen Melderegister identifiziert und für Zählungszwecke (unabhängig vom Meldewesen) eliminiert, flächendeckend für ganz Österreich durchzuführen.

Die im November 2008 veröffentlichte vorläufige Volkszahl für das Finanzjahr 2009 wurde mit den durch die Bevölkerungsstatistik fortgeschriebenen Ergebnissen der Probezählung 2006 ermittelt. Im Jahr 2009 findet zum Stichtag 31.10.2008 eine sogenannte „Mini“-Registerzählung statt. Die im November 2009 veröffentlichte endgültige Volkszahl für den Stichtag 31.10.2008 wird für das Finanzjahr 2009 revidierte Zahlen und für das Finanzjahr 2010 bereits endgültige Zahlen liefern.

Das grundsätzliche Procedere der „Mini“-Registerzählung, die Zusammenführung der verschiedenen Register unterscheidet sich nicht von der Probe bzw. der eigentlichen Registerzählung. Es tritt lediglich an Stelle der Wohnsitzanalyse mit Personenrecherche ein mathematisches Modell, das auf den Erfahrungen der Wohnsitzanalyse im Zuge der Probezählung basiert.

Die Anforderungen an dieses Modell sowie die Methode der synthetischen Wohnsitzanalyse werden vorgestellt und diskutiert.

## **Catching Position Effects Red-Handed: Recent Modeling Strategies and Application**

Marco J. Maier, Rainer W. A. Alexandrowicz  
University of Klagenfurt, Psychology Department  
Applied Psychology and Methods Research Unit

Position effects in psychological tests are by no means a new phenomenon and have been investigated since the 1950s. In a nutshell: depending on an item's position within a psychological test, not only item difficulty per se (unconditional item difficulty) has to be taken into account, but also its position's influence.

Throughout history, various modeling strategies have been used, ranging from simple 2-group designs to elaborate IRT methods. One prominent approach is to use a LLTM for estimating both unconditional and conditional item difficulty components.

A Rasch-homogeneous computer-based reasoning-test has been administered in 12 groups that feature the same items but in varying arrangements (Latin Square). The application of the LLTM shows a slightly U-shaped pattern of the position effects, yet remains inconclusive altogether. Modeling position effects in this fashion, when item processing times are unrestricted and thus have a tremendous variation is doubtful.

Therefore we switched to a different modeling approach using a reformulation of the LLTM in terms of a generalized linear random coefficient multilevel model with items stacked within respondents. This model allows for the inclusion of continuous predictor variables, such as the item starting times in a kind of growth curve model (technically a polynomial) to account for the position influence. This model fits the data much better, revealing a clear-cut U-shaped effect of time of the item difficulties.

The two approaches clearly have their assets and drawbacks: The LLTM approach allows for the CML estimation method, which makes no distributional assumptions concerning the ability and allows for the application of a likelihood ratio test to assess model fit. But due to the inclusion of position effects in a discrete fashion, these cannot be convincingly detected. On the other hand, the generalized linear mixed model approach offers more flexibility in the inclusion of arbitrarily scaled effects, but as it applies the MML estimation method the advantages of the CML method are lost.

In our case, the latter approach proved more informative: We were able to detect aberrations of item difficulties due to test taking time, which would render the psychological test invalid when applied e.g. in an adaptive testing environment. A key component in the discussion is whether the methodological ends (new insight into position effects) justify the means (abandoning CML estimation).

## **Geburtstageeffekte in den Datensätzen der amtlichen Statistik. Eine Replikation der „Akte Astrologie“ von Gunter Sachs.**

Ivo Ponocny, Elisabeth Ponocny-Seliger  
Modul University Vienna

Vor 12 Jahren verblüffte Gunter Sachs die Öffentlichkeit mit seinem Bestseller „Die Akte Astrologie“, welcher den Nachweis eines Zusammenhangs zwischen Sternzeichen und menschlichem Verhalten liefern sollte.

Basis des Buches waren empirische Studien, hauptsächlich mit Daten aus amtlichen statistischen Daten, welche fast durchwegs signifikante Resultate lieferten. Von wissenschaftlich-skeptischer Seite wurden diese Schlussfolgerungen aufgrund diverser methodischer oder interpretatorischer Mängel heftig kritisiert. Insbesondere ließ das Buch offen, welche anderen Sachverhalte zu Geburtstageeffekten führen können bzw. welche Geburtstageeffekte sich in der amtlichen Statistik nun wirklich niederschlagen.

Nun wurde mit Förderung durch die *Gesellschaft für die wissenschaftliche Untersuchung paranormaler Phänomene* eine Replikation etlicher in der Akte Astrologie beschriebener Untersuchungen anhand von Auswertungen aus Daten der österreichischen amtlichen Statistik vorgenommen, je nach Fragestellung mit kleineren oder größeren Populationen als in der „Akte Astrologie“ (zwischen 60.000 und 6.000.000 Personen, Berichtszeiträume zwischen 2004 und 2006). Diese Auswertungen waren sekundärstatistischer Natur, mit Schwerpunkten auf den Themen Bildungsabschlüsse, Eheschließungen, Scheidungen, Selbstmord, Todesursachen, Trends bei der Verteilung der Geburten und ähnlichen Variablen. Zur Absicherung der Interpretationen wurden auch Daten aus der PISA-Studie der OECD herangezogen.

Generell konnte bestätigt werden, dass die Effektgrößen, so sie nicht null sind, extrem klein ausfallen. Die von Sachs errechneten Signifikanzen wurden großteils nicht beobachtet, zumindest nicht unter Verwendung angemessener statistischer Verfahren. Allerdings ließen sich in einigen Teilbereichen Geburtstageeffekte zweifelsfrei nachweisen, teilweise solche, die bereits in der Akte Astrologie beschrieben wurden, aber auch noch weitere darüberhinaus. Speziell ließen sich – wie stellenweise in der wissenschaftlichen Literatur bereits beobachtet oder vermutet – relative Alterseffekte belegen, speziell im Bildungsbereich, aber auch Auswirkungen des saisonalen Charakters der Tätigkeit bestimmter Bevölkerungsgruppen. Statistisch wirksame (sehr kleine) Effekte der Stellung des Geburtstags im Jahreskreis können mithilfe des österreichischen Datensatzes somit eindeutig bestätigt werden. (Die Art dieser Einflüsse bietet allerdings wenig Veranlassung, darin einen Beleg für die Richtigkeit astrologischer Thesen zu sehen.) Weiters bestätigen sich einige Beobachtungen aus der Literatur in Bezug auf Wochentageeffekte und Zusammenhänge zwischen Geburts- und Sterbemonaten.

## **The Emergence of Item Response Theory Models**

Steven P. Reise, UCLA

Item response theory (IRT) measurement models were originally developed in the context of large scale cognitive testing (e.g., abilities, achievement, or aptitude testing). It is fair to state that measurement research in this domain is now dominated by IRT as opposed to traditional classical test theory methods.

In this keynote address I describe the problems in cognitive measurement that can potentially be addressed by IRT methods. However, the main focus of the presentation is how IRT psychometrics has emigrated to other content domains, in particular, personality, psychopathology, and patient reported health outcomes assessment. The talk will not focus on technical/statistical problems per se (e.g., what is the best fit index, or parameter estimation method), but rather on the conceptual issues that arise in applying IRT models to typical performance as opposed to maximal performance types of constructs (e.g., the problems of specifying an item pool, and the problem of non-continuous constructs).

Specifically, I will review recent work applying IRT in psychopathology assessment and emphasize how IRT methods potentially provide unique and potentially more valid information in this domain. In this context I will emphasize the conceptual problems that must be addressed in applying IRT in this domain. For example, the nature of psychopathology is such that high trait level individuals (i.e., individuals at high levels of psychopathology) are not expected to universally display any particular symptom. As I will argue, this and other conceptual problems creates some interesting challenges for fitting IRT models. I will also review recent work on a large scale project called PROMIS that is attempting to develop common measures of important health outcomes constructs such as pain, fatigue, and anxiety. PROMIS is the largest IRT study ever conducted outside of cognitive assessment, and has revealed many interesting challenges, but also many opportunities to increase the quality of assessment and ultimately basic research.

## **Liebe – emotional, aber digital** **Partnerwahl und Partnerschaft im Internet**

Kristina Riedmüller (Diplomarbeit)

Das Hauptziel der Arbeit war festzustellen, ob theoretische Konzepte der Partnerwahl, die es für die Partnerwahl im realen Leben gibt, auch auf die Partnerwahl im Internet anwendbar sind.

Es gibt verschiedene theoretische Konzepte, die versuchen zu erklären, aufgrund welcher Kriterien eine Person als Partner gewählt wird. Dabei werden der soziobiologische Ansatz, die Ähnlichkeits-Hypothese, die Komplementaritäts-Theorie und die Equity-Theorie vorgestellt. Es wird diskutiert, wie sich das jeweilige Konzept auf die Partnerwahl im Internet umlegen lässt.

Die theoretischen Konzepte wurden mit Daten aus einer Online-Umfrage überprüft. Als Methoden wurden T-Test, Korrelations-Koeffizienten, Kappa-Koeffizient und Regressionsanalyse angewendet. Die gängigen Theorien, die es über die Partnerwahl gibt, konnten anhand des Datensatzes bestätigt werden. Sie lassen sich auch auf die Partnerwahl im Internet anwenden. Nicht nur bei der Offline-Partnersuche, sondern auch bei Personen, die ihren Partner im Internet kennen gelernt haben, ist es der Fall, dass Frauen mehr Wert auf die gute berufliche Stellung des Partners legen und für Männer die physische Attraktivität wichtiger ist (sozio-biologischer Ansatz). Jedoch ist Personen, die ihren Partner Online getroffen haben, die physische Attraktivität weniger wichtig, als jenen, die ihren Partner im realen Leben kennen gelernt haben. Sowohl Offline- als auch Online-Paare weisen überdurchschnittliche Homogamie in den Merkmalen Bildung, Einkommen Religionsbekenntnis und Alter auf. Dabei ist die Ähnlichkeit von Online-Paaren etwas schwächer ausgeprägt. Auch in Bezug auf kulturelle Interessen, Freizeit und Lifestyle weisen Paare Ähnlichkeiten auf. Die Equity-Theorie konnte nicht bestätigt werden. Personen mit einem gleichwertigen Partner sind nicht am zufriedensten.

## **Measurement of Latent Variables in Marketing – Current Practices and Future Avenues**

Thomas Salzberger, WU Vienna, Department of Marketing

### Extended abstract:

In scientific research in marketing, quantitative theories and approaches have become more and more popular over the past decades. Not only in consumer behaviour research but also in business-to-business research and other fields of application, reference to latent variables has proliferated. Whether this development reflects progress in the discipline heavily depends on the quality and the trustworthiness of measures of latent variables.

Well into the 1970s, measurement in marketing was often based on single-item measures. The analysis of multi-item scales was mostly confined to reliability assessment. However, the relationship of the observed responses and the concept of interest was not investigated. Accordingly, in 1978, Jacoby (1978, p.91, emphasis in the original) lamented that “most of our measures are only measures because someone says that they are, not because they have been shown to satisfy standard criteria”. One year later, Churchill (1979) established a ‘Paradigm for Developing Better Measures of Marketing Constructs’. Single-item measures were superseded by multi-item scales. Mere face validity was replaced by content validity, criterion-related and construct validity. In other words, the era of measurement by fiat came to an end and the quality of measurement became a subject of empirical investigation.

The shift from true score theory (classical test theory, Lord and Novick, 1968) to the congeneric measurement model (Jöreskog, 1971) and the widespread application of factor analysis went by in a surprisingly unspectacular way. As pointed out by Borsboom (2005), the concept of a latent variable is alien to true score theory but of vital importance in the con-generic model, where the latent variable is explicitly accounted for. However, classical test theory and its key concepts (like reliability) are still considered the theoretical basis of measurement in marketing. Even though confirmatory factor analysis has contributed to an increase in the quality of measures, emphasis is still focused on item inter-correlations and covariances.

While the methodology to analyse the relationship of observed responses and latent constructs has been refined, the definition of measurement has been unaltered during the past 60 years. According to Stevens (1946, 1951), measurement “involves rules for assigning numbers to objects to represent quantities of attributes” (Nunnally, 1967, p.2). Thus, measurement is achieved by assignment. In other words measurement is decreed and it is hard to see how measurement can fail as long as some arbitrary rules are observed. Consequently, the subsequent estimation of reliability and the application of exploratory and/or confirmatory factor analysis are actually not concerned with the question whether measurement has been achieved or not but merely assess to what degree the measures appear acceptable. The fact that fit of the data and a factor analytic model, which specifies reflective indicators, is not regarded as essential for measurement becomes obvious when alternative measurement approaches based on formative indicators are suggested (Diamantopoulos and Winklhofer, 2001; Rossiter, 2002; MacKenzie, 2003; and Jarvis, MacKenzie, and Podsakoff, 2003). According to proponents of such models, the choice of reflective versus formative indicators is not arbitrary. Consequently, a measurement model can be ill-specified. Although several criteria (flow of causality, common theme, dimensionality, etc.) have been proposed to distinguish between these variants of indicators, no criterion actually allows for an empirical investigation of the suitability of either model. The reason, in the end, is that neither model addresses the pivotal problem of measurement properly, namely whether or not measurement has actually been achieved. While the reflective indicator model at least allows for a latent variable (albeit in a somewhat speculative way), the formative model essentially adapts a model for predication (without any reference to a latent variable) in an unjustified way and passes it on as a measurement model.

Any improvement of measurement in marketing has to start with a reconsideration of the concept of a latent variable. A latent variable, or a construct, implies a special sort of scientific theory which claims that a particular variable actually exists as a quantitative property. Therefore, we need empirical evidence that this hypothesis is plausible.

Measurement understood as a task of assignment does not contribute to testing this hypothesis; it rather obscures the essence of measurement. The same applies to sophisticated analyses, like confirmatory factor analysis, based on presumed measurement. We, consequently, have to adapt our methodology in a way that the hypothesis of a latent variable is properly addressed. Particularly helpful in this respect is Michell's (1997) distinction between the scientific task of measurement and the instrumental task of measurement. The scientific task is concerned with the actual existence of a latent variable, which is an ontological claim. A quantitative variable implies a particular structure, the additive structure of quantity. It is important to realise that "[t]here is no logical necessity that any attribute should have this kind of structure" Michell (1997, p.358). In other words, evidence is needed that this is the case. Otherwise measurement is impossible and any operations carried out nevertheless lead to pseudo-quantification. It goes without saying that the scientific task of measurement cannot be addressed using Stevens definition, because assignment is incompatible with scientific discovery. Michell (1997, p.358) defines measurement as "the estimation or discovery of the ratio of some magnitude of a quantitative attribute to a unit of the same attribute". This definition parallels the understanding of measurement in the natural sciences. The subsequent instrumental task of measurement is concerned with devising concrete measurement instruments. The mainstream approach to measurement in marketing has been concerned with the instrumental task but effectively neglected the scientific task completely.

The criteria for the selection of a measurement model that adequately captures the scientific task as well as the instrumental task of measurement follow from cancellation conditions (Karabatsos, 2001) based on axiomatic representational measurement (Krantz et al, 1971) and additive conjoint measurement (Luce and Tukey, 1964). Item response theory (IRT, Embretson and Reise, 2000) offers a range of models which explicitly account for a latent variable and allow for a non-linear relationship between manifest responses and a latent variable. Within IRT, the Rasch model (Rasch, 1960) stands out as it complies with single and double cancellation (Karabatsos, 2001), which are requirements of quantity.

Which avenue measurement in marketing will follow in the future, remains to be seen. Over the past ten years, measurement has continuously been on the methodological research agenda and a matter of debate. While some have promoted the formative measurement model intensively (e.g., MacKenzie, 2003; and Jarvis, MacKenzie, and Podsakoff, 2003), otherseven advocate a revival of single-item measures (Bergkvist and Rossiter, 2007) for latent variables under some conditions. By contrast, an increasing body of research has been based on the Rasch model for measurement (e.g., Salzberger et al., 2009; Ganglmair-Wooliscroft, 2007; Koller and Salzberger, 2007; Salzberger and Sinkovics, 2006; Ewing et al., 2005; Ganglmair and Lawson, 2003a, 2003b) and a comprehensive framework for the measurement of latent variables in marketing based on the Rasch model has been proposed (Salzberger, 2009).

#### References:

- Bergkvist, L., Rossiter, J. 2007. 'The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs', *Journal of Marketing Research* Vol. XLIV, 175-184.
- Borsboom, D. (2005), 'Measuring the Mind: Conceptual Issues in Contemporary Psychometrics', Cambridge: Cambridge University Press.
- Churchill, G. A. 1979, 'A paradigm for developing better measures of marketing constructs', *Journal of Marketing Research* XVI (February), 64-73.
- Diamantopoulos, A., Winklhofer, H. M. 2001, 'Index Construction with Formative Indicators: An Alternative to Scale Development', *Journal of Marketing Research* 38 (May), 269-277.
- Embretson, S.E. and S.P. Reise (2000), 'Item Response Theory for Psychologists, Mahwah', New Jersey: Lawrence Erlbaum Associates.

- Ewing, M., T. Salzberger and R. Sinkovics (2005), 'An Alternate Approach to Assessing Cross-Cultural Measurement Equivalence in Advertising Research', *Journal of Advertising*, 34 (1), 17-36.
- Ganglmair-Wooliscroft, A. (2007), 'A Comparison of Affective Response to Consumption in Two Contexts', *der markt*, 46 (1+2), 50-60.
- Ganglmair, A. and R. Lawson (2003a), 'Advantages of Rasch Modelling for the Development of a Scale to Measure Affective Response to Consumption', *European Advances in Consumer Research*, 6, 162-168.
- Ganglmair, A.; Lawson, R. (2003b), 'Measuring Affective Response to Consumption Using Rasch Modelling', *Journal of Customer Satisfaction, Dissatisfaction and Complaining Behavior*, 16, 198-21.
- Jacoby, J., 1978, 'Consumer Research: A State of the Art Review', *Journal of Marketing*, 42 (April), 87-96.
- Jarvis, C. B., MacKenzie, S. B., and Podsakoff, P. M., 2003, 'A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research', *Journal of Consumer Research*, 30 (September), 199-218.
- Jöreskog, K.G. (1971), 'Statistical Analyses of Sets of Congeneric Tests', *Psychometrika*, 36, 109-133.
- Karabatsos, G., 2001, 'Understanding Rasch Measurement: The Rasch Model, Additive Conjoint Measurement, and New Models of Probabilistic Measurement Theory', *Journal of Applied Measurement*, 2 (4), 389-423.
- Koller, M., Salzberger, T. (2007), 'Cognitive Dissonance as a Relevant Construct Throughout the Decision-Making and Consumption Process - An Empirical Investigation Related to a Package Tour', *Journal of Customer Behaviour*, Vol 6 (3), 217-227.
- Krantz, D. H., R.D. Luce, P. Suppes, and A. Tversky (1971), 'Foundations of Measurement, Vol. I: Additive and Polynomial Representations', New York: Academic Press.
- Lord, F. M., and Novick, M.R., 1968, 'Statistical Theories of Mental Test Scores', Reading (Mass.): Addison-Wesley.
- Luce, R.D., and Tukey, T.W., 1964, 'Simultaneous conjoint measurement: a new type of fundamental measurement', *Journal of Mathematical Psychology*, 1, 1-27.
- MacKenzie, S. B., 2003, 'The Dangers of Poor Construct Conceptualization', *Journal of the Academy of Marketing Science*, 31 (3), 323 - 326.
- Michell, J., 1997, 'Quantitative Science and the Definition of Measurement in Psychology', *British Journal of Psychology*, 88, 355-383.
- Nunnally, J. C., 1967, 'Psychometric Theory', New York: McGraw Hill.
- Rasch, G. (1960), 'Probabilistic Models for Some Intelligence and Attainment Tests', Copenhagen: Danish Institute for Educational Research, expanded edition (1980) with foreword and afterword by B.D. Wright. Chicago: The University of Chicago Press.
- Salzberger, T. (2009, forthcoming), 'Measurement of Latent Variables in Marketing - An Alternative Framework', Edward Elgar: Cheltenham, UK and Northampton, MA, USA.
- Salzberger, T. and R. Sinkovics (2006), 'Reconsidering the Problem of Data Equivalence in International Marketing Research - Contrasting Approaches Based on CFA and the Rasch model for Measurement', *International Marketing Review*, 23 (4), 390-417.
- Salzberger, T., Holzmüller, H.H., Souchon, A. (2009), 'Advancing the Understanding of Construct Validity and Cross-National Comparability - Illustrated by a Five-Country Study of Corporate Export Information Usage', *Advances in International Marketing*, Volume 20 'New Challenges to International Marketing', 321-360.
- Stevens, S.S. (1946), 'On the Theory of Scales of Measurement', *Science*, 103, 667-680.
- Stevens, S.S. (1951), 'Mathematics, Measurement, and Psychophysics', in S.S. Stevens (ed.) *Handbook of Experimental Psychology*, New York, NY.: Wiley, 1-49.

## **„Frauen als Stille Reserve im Ingenieurbereich – eine ökonometrische Analyse“**

Eva Schlenker  
Statistik & Ökonometrie, Universität Hohenheim; D

Angesichts der demographischen Veränderungen wird die Zahl der Erwerbspersonen in Deutschland in den kommenden Jahrzehnten dramatisch sinken. Und schon heute fehlt qualifiziertes Personal, besonders im Ingenieurbereich. Die Politik begegnet diesen Entwicklungen mit einem Spektrum an Kampagnen auf dem Feld der Familien- und Bildungspolitik. Ein Ansatzpunkt ist die in Deutschland geringe Erwerbsquote von Frauen, insbesondere Müttern. Die zentrale Fragestellung ist nun, ob und in welchem Umfang Frauen im Ingenieurbereich ein ungenutztes Potential an Arbeitskräften bilden. Es soll außerdem eine Größenordnung der Stillen Reserve an Ingenieuren abgeschätzt werden. Dabei werden die Schätzungen für Nicht- bzw. in Teilzeit Erwerbstätige und für fachfremd tätige Arbeitskräfte durchgeführt.

Die empirische Untersuchung basiert auf dem Mikrozensus des Jahres 2006. Der Datensatz wird nach Personen mit akademischem Abschluss an einer Universität oder Fachhochschule im Ingenieurbereich sowie einer Altersobergrenze von 56 Jahren gefiltert. Anschließend werden die Daten in drei Untersuchungsgruppen analysiert: Männer sowie Frauen mit und ohne Partner im Haushalt. Die abhängigen Variablen sind der Grad der Erwerbstätigkeit (Nicht-Erwerb, Teilzeit, Vollzeit) sowie die Adäquatheit der ausgeübten Tätigkeit (fachfremd, fachlich adäquat). Zur Analyse werden Probit-Modelle für ordinale und binäre Daten verwendet. Für den Grad der Erwerbstätigkeit werden außerdem semi-nonparametrische Schätzungen eingesetzt. Als unabhängige Variablen werden neben soziodemographischen Variablen (Alter, Wohnsitz in Ostdeutschland) haushaltsbezogene Variablen wie die Kinderzahl, das Alter des jüngsten Kindes, das Einkommen des Partners und der Familienstand verwendet.

Die empirischen Arbeiten belegen große Unterschiede zwischen den untersuchten Personengruppen. Ingenieure scheinen allgemein im Alter in geringerem Umfang und seltener fachlich adäquat zu arbeiten. Im Gegensatz dazu muss bei den Unterschieden zwischen Ost- und Westdeutschland zwischen Männern und Frauen differenziert werden: Während ostdeutsche Frauen häufiger erwerbstätig sind als ihre westdeutschen Kolleginnen, ist der Effekt bei den Männern entgegengesetzt. Ingenieure arbeiten in Westdeutschland häufiger Vollzeit als in den östlichen Bundesländern. Bezüglich der Haushaltsvariablen zeigt sich, dass Frauen vor allem den Grad ihrer Erwerbstätigkeit stark an ihren Familien orientieren: Je mehr Kinder eine Ingenieurin hat und je jünger diese sind, desto seltener ist sie erwerbstätig. Außerdem schränkten Ingenieurinnen ihre Erwerbstätigkeit eher zu Gunsten eines einkommensstarken Partners ein. Zur Abschätzung einer Stillen Reserve an Arbeitskräften werden nun auf Basis der Modelle individuelle Wahrscheinlichkeiten für den Grad der Erwerbstätigkeit prognostiziert, anschließend gemittelt und auf die Gesamtzahlen hochgerechnet. Die Hochrechnungen ergeben, dass insgesamt 605.400 Ingenieure fachfremd arbeiten, 113.600 von ihnen sind Frauen. Wird der Grad der Erwerbstätigkeit betrachtet, sind 76.000 Ingenieure nicht erwerbstätig. Weitere 71.400 Personen arbeiten in Teilzeit. Von diesen 147.400 Personen in Nichterwerb oder Teilzeit sind 44 Prozent weiblich. Ingenieurinnen, die im Ingenieurwesen allgemein stark unterrepräsentiert sind, stellen damit eine wichtige Größe in der Stillen Reserve dar. Neben den Ingenieurinnen hat sich in der vorliegenden Arbeit gezeigt, dass es eine zweite große Gruppe bei den Ingenieuren der Stillen Reserve gibt: Männer ab Mitte Vierzig. Unternehmen bevorzugen bei der Einstellung jedoch noch immer junge Ingenieure.

## **Long-term fertility trends in Austria, the Czech Republic and Slovakia**

Tomáš Sobotka and Kryštof Zeman, Vienna Institute of Demography

This contribution compares long-term fertility developments in Austria, the Czech Republic and Slovakia, three central European countries that belonged to the Austro-Hungarian Empire until its demise in 1918. First, we analyse fertility histories of women born since the beginning of the 20<sup>th</sup> century up to the late 1960s and notice a convergence in family size among Czech and Slovak women and their strong orientation towards a two-child family model that took place during the existence of Czechoslovakia. We also notice lower fertility and higher childlessness among Austrian women. Second, we look at long term changes of period fertility level and timing after the World War II and show how the cycle of baby booms and busts in Austria differed from the cycles experienced in the Czech Republic and Slovakia, where also a shift to a very early childbearing pattern took place. Third, we contrast rapid fall in fertility, massive postponement of childbearing and a recent modest recovery in total fertility rates in the Czech Republic and Slovakia after 1990 with considerably more stable fertility trends in Austria. In many respects, fertility trends in Slovakia follow those in the Czech Republic with an apparent time lag of several years. Finally, we compare the most recent fertility trends after 2000 and outline major emerging similarities and differences between these neighboring countries. In all the three countries, the completed cohort fertility of women born in the mid- to late-1970s can reach similar levels of 1.6-1.8, but their parity distributions will show persistent differences.

## **Fertility trends in Hungary before and after 1989.**

Zsolt Speder, Director, Hungarian Demographic Institute

This paper will summarize trends in childbearing behavior in Hungary since the 1960s. It will particularly focus on the recent fertility decline and try to discuss the factors which caused this trend. Finally, the paper will discuss the role of policy interventions and attempt an outlook for Hungarian fertility).

## **Psychologische Daten und Datensätze in der statistischen Empirie.**

Sabine Stadler

Die Psychologie und konkret die Testpsychologie haben in großen Umfang in Österreich Fuß gefasst und aufbauend auf den Kenntnissen große Handlungsfelder entwickelt. Ausgehend von den zahlreichen Studien die auf Universitäten erarbeitet werden, gebe ich einen kurzen Überblick über die wichtigsten großen Unternehmen und ihre Merkmale als auch ihre großen psychometrischen Datensätze. Oft resümiert das Berufsfeld unter Marktforschung, in keiner Hinsicht werden die zahlreichen besonderen Daten gespeichert und in komprimierter Form veröffentlicht.

Die GfK Dr. Fessel Institute haben sich im gesamten Osteuropa und GUS eingerichtet und die Aufgabe transkulturelle Studien, Umfragen und Marktforschung zu betreiben. Unter dem Titel healthy-typology, als auch growth via knowledge werden zahlreiche Studien zum körperlichen und seelischen Wohlbefinden erarbeitet.

In health typology erfahren wir ob

- in Rumänien man oft oder nie zum Arzt geht
- in Polen niedrige oder hohe Medikamentenpreise auf dem Markt sind
- in der Slowakei in irgendeiner Form neue Operationsmethoden existieren etc.

Auch werden repräsentative Studien von Wien ausgemacht, die osteuropäischen Regierungen zahlen gut und sind stabil.

Die soziokulturellen Trends haben zahlreiche unmittelbarer Folge, gem. mit dem Zukunftsinstitut von Dr. Horx hat sich eine besondere Kultur von TREND-Reports 2009, Life-style Studien sowie zu den Themen.

Finanzmarkt und Versicherungsmarkt

Handel

Konsumgüter

Sozial- und Medienforschung

Touristik

entwickelt und dabei werden die Studien online zur Verfügung gestellt.

Ein zweites Institut, Gallup Dr. Kamarsin, arbeitet auf den österreichischen Markt in Marketing und Marktforschung und insbesondere in der Werbeforschung. Der meist gesehene Werbespot, das meist gekaufte Produkt, das sind die Aufträge die Dr. Kamarsin, ganz im Sinne der US -Forschung durchführt. Die Reports sind nur online abrufbar und sind eine Zusammenfassung der Marktforschungsergebnisse.

Dr. Kamarsin arbeitet für die freie Wirtschaft, auch diese zahlt gut. Es sind die den Bereichen stabile und wenig expandierende Märkte mit guten Einkommen vorhanden, Österreich ist im Vergleich zum restlichen Europa, sowie alle empirischen Methoden und Umfragen europäische Mittelklasse bis Schlusslicht, weil keine Öffnung und Kooperation mit anderen Auftraggebern.

Literatur:

GfK Dr. Fessel, WEBPAGE

Institut Dr. Kamarsin, WEBPAGE

## **Strategien zur Analyse der kleinräumigen Verteilung von Armut in Österreich.**

Matthias Till und Johannes Klotz  
Statistik Austria

Seit dem Jahr 2003 führt die Bundesanstalt Statistik Österreich jährlich die Erhebung EUSILC zu Einkommen und Lebensbedingungen auf Basis von EU-Verordnungen durch. Hauptzweck dieser Erhebung ist die Berechnung von Indikatoren zur sozialen Eingliederung bzw. die Berichterstattung über die Verteilung der Haushaltseinkommen und das Ausmaß der Armutsgefährdung in der Europäischen Union.

Um die soziale Lage der Gesamtbevölkerung einigermaßen zuverlässig einschätzen zu können, sieht die Verordnung für die Republik Österreich eine Mindeststichprobengröße von 4.500 Haushalten vor. Die regionale Verteilung von Armut oder die Situation besonders exponierter Bevölkerungsgruppen kann auf dieser Grundlage jedoch nur sehr rudimentär beschrieben werden.

Als eine mögliche Strategie präsentiert dieser Beitrag deshalb modellbasierte Schätzungen für die Verteilung der Haushaltseinkommen und das Ausmaß der Armutsgefährdung. Dabei werden auf Grundlage der vorhandenen Beobachtungen in EUSILC zunächst jene Merkmale identifiziert, von denen die finanzielle Lage eines Haushaltes abhängt. Dazu zählen etwa die Zahl, Geschlecht, Bildung, berufliche Stellung und Beschäftigungsausmaß der Erwerbspersonen im Haushalt sowie die Anzahl der Kinder und ökonomisch nicht aktiven Personen. Diese Merkmale werden nun in ein logistisches Regressionsmodell eingeführt, mit dem die Wahrscheinlichkeit von Armutsgefährdung geschätzt wird. Aus der Kombination der signifikanten Parameter ergibt sich ein differenziertes Schichtungsmodell von in Hinblick auf Armutsgefährdung relativ homogenen Bevölkerungsgruppen. Die auf diese Weise geschätzten Wahrscheinlichkeiten können nun auf beliebige Abbildungen der Grundgesamtheit projiziert werden, solange die dem Modell zugrundeliegenden Merkmale identifiziert werden können. Viele der für Armutsgefährdung maßgeblichen Merkmale werden beispielsweise in der Arbeitskräfteerhebung erfasst. Aufgrund des größeren Stichprobenumfangs sind auf dieser Basis differenziertere Schätzungen möglich. Noch detailliertere Schätzungen sind möglich, wenn Ergebnisse der Volkszählung herangezogen werden.

Neben diesem modellbasierten Ansatz sind mittelfristig auch andere Strategien zur Verbesserung direkter Schätzungen denkbar. Beispielsweise indem durch Befragung erhobene Merkmale in der Arbeitskräfteerhebung angereichert werden durch Verwaltungsdaten über Einkommensbezüge. Durch die Befragung werden dabei Haushaltsbezüge hergestellt, die der Lebensrealität entsprechen. Eine weitere Alternative wäre es, den Haushaltsbezug bereits auf der Ebene von Verwaltungsdaten zu rekonstruieren und auf dieser Grundlage personelle Einkommen zu Haushaltseinkommen zu aggregieren. Die Vor- und Nachteile bzw. notwendigen rechtlichen Rahmenbedingungen für die jeweiligen Strategien werden in diesem Beitrag zur Diskussion gestellt.

## **Statistical Challenges in Search Engine Marketing - Schuhplattler Go Google AdWords**

Daniela Weber, Department of Statistics and Mathematics, WU Vienna  
Eva Lienbacher, Institute for Retailing and Marketing, WU Vienna  
Wolfgang Ziniel, Institute for Marketing Management, WU Vienna  
Patrick Mair, Department of Statistics and Mathematics, WU Vienna

Search engines are more than just portals or information tools. They are agents of a transformation that is making the business environment more transparent, and thus, potentially more competitive. This new environment is creating opportunities but also challenges for businesses of every stripe. The success of sponsored search has radically affected the way people interact with the information, websites, and services on the web. During our participation at the Google Online Marketing Challenge 2009, a challenge where student teams can gain experiences with sponsored search marketing, we will take out several campaigns to reach new clients and new attendees. The differentiation between clients and attendees requires variable AdWords strategies, which will be analyzed in a statistical point of view in retrospect and also during the campaign. Optimizing our campaigns we applied predictive methods of web usage mining within the context of search engine marketing, e.g. web-traffic-analysis, click stream-analysis as well as quantitative analysis of page views.

## **Segmenting and structuring markets using Configural Frequency Analysis (CFA)**

Stefan Wiesel, Institute for Retailing and Marketing, WU Vienna

Despite segmenting and structuring markets are core concepts in marketing theory and practice, most of the existing statistic approaches require metric data. The purpose of our research is to introduce a non-parametric approach to segmenting and structuring markets: Configural Frequency Analysis (CFA). The approach is applied to standard market research data representing the Austrian grocery retail market. While segmentation is demonstrated by using functional CFA (fCFA), structuring is carried out by more-sample CFA (mCFA). Mosaic Plots and Correspondence Analysis are used to visualize the findings.

## **Development of a Microanalytical Simulation Model to predict the long-term evolution of employment biographies in Austria Implementation of the Demographics Module**

Marcus Wurzer, Department für Statistik und Mathematik, WU Vienna

The well-known problems of decreasing birth rates and population ageing represent a major challenge for the Austrian pension system, just like in many other countries. It is expected that the group of pensioners will grow steadily in the future, while the proportion of people that support them - the taxpayers - will shrink.

Microsimulation provides a valuable tool to identify the impact of various policy measures. The main advantage in comparison to macro models that work with aggregate data is that the consequence of a certain measure can be evaluated directly on the level of the individual persons. Therefore, it is not only possible to predict cross-sectional data (e.g., the distribution of age groups in 2050), but also to simulate lifecourses of people, providing longitudinal outcomes. Due to the stochastic nature of the Microsimulation Model, it is easy to simulate different scenarios in order to get answers to 'What if...' - questions.

Traditionally, the whole Microsimulation Model is divided into different modules (see, e.g., SESIM, APPSIM, DYNAMOD etc.), each of them approaching a number of events that can happen to a person during its lifecourse. One of the most important modules is the Demographics Module, simulating death, birth, marriage, migration etc. This module also constitutes the first one of the Microsimulation Prototype that has been finished. The author will talk about the definition of the submodules, the statistical modeling of the events, the implementation into the statistical programming language R and the evaluation of the results, using projections of the Statistik Austria as reference.