



# Forschungsbericht 2015

Institut für Mathematische Stochastik

# INSTITUT FÜR MATHEMATISCHE STOCHASTIK

Universitätsplatz 2, 39106 Magdeburg  
Tel. +49 (0)391 67 58651, Fax +49 (0)391 67 11172  
imst@ovgu.de

## 1. Leitung

Prof. Dr. Norbert Gaffke - geschäftsführender Leiter (bis 31.03.2015)

Prof. Dr. Claudia Kirch - geschäftsführende Leiterin (ab 01.04.2015)

Prof. Dr. Rainer Schwabe

apl. Prof. Dr. Waltraud Kahle

Dr. Heiko Großmann (seit 01.04.2015)

## 2. Hochschullehrer

Prof. Dr. Norbert Gaffke (bis 31.03.2015, seit 01.04.2015 im Ruhestand)

Prof. Dr. Jürgen Groß (seit 01.04.2015)

Prof. Dr. Claudia Kirch (seit 01.04.2015)

Prof. Dr. Cornelia Pokalyuk (seit 01.10.2015, Dorothea-Erxleben-Gastprofessur)

Prof. Dr. Rainer Schwabe

apl. Prof. Dr. Berthold Heiligers (extern)

apl. Prof. Dr. Waltraud Kahle

Professoren im Ruhestand:

Prof. Dr. rer.nat.habil. Otfried Beyer

Prof. Dr. rer.nat.habil. Gerd Christoph

Prof. Dr. rer.nat.habil. Norbert Gaffke (seit 01.04.2015)

## 3. Forschungsprofil

Mathematische Stochastik (Stochastische Prozesse): Prof. Dr. Gerd Christoph; apl. Prof. Dr. Waltraud Kahle

- Asymptotische Methoden in der Stochastik
- Untersuchungen zu Ruinwahrscheinlichkeiten bei Risiko-Prozessen
- Modellierung und Statistik von Schädigungsprozessen
- Statistische Analyse allgemeiner Ausfall-Reparatur-Prozesse
- Optimale Instandhaltung in allgemeinen Reparaturprozessen

Mathematische Stochastik (Mathematische Statistik): Prof. Dr. Norbert Gaffke

- Statistische Regressionsmodelle
- Experimental Design: Theorie und Algorithmen

- Tests und Konfidenzschranken
- Statistische Modellierung interdisziplinär

Mathematische Stochastik: Prof. Dr. Jürgen Groß

- Lineare statistische Modelle und lineare Matrixalgebra
- Psychometrische Modelle
- Methoden der Zeitreihenanalyse

Mathematische Stochastik (Mathematische Statistik): Prof. Dr. Claudia Kirch

- Zeitreihenanalyse
- Changepoint-Analyse
- Resampling-Verfahren für Zeitreihen
- Statistische Methoden im Frequenzbereich
- Sequentielle Methoden
- Funktionale/Hochdimensionale Daten

Mathematische Stochastik (Stochastische Prozesse und ihre Anwendungen): Prof. Dr. Cornelia Pokalyuk

- Stochastische Prozesse der Populationsgenetik, insbesondere Darstellung und Analyse dieser mittels Graphen, z.B. durch den ancestralen Selektionsgraphen
- Untersuchung des Effekts von Selektion auf Genealogien
- Modellierung der Evolution des Zytomegalievirus

Mathematische Stochastik (Statistik und ihre Anwendungen): Prof. Dr. Rainer Schwabe; Dr. Heiko Großmann

- Planung und Auswertung statistischer Experimente
  - Conjoint-Analyse (Psychologie, Marktforschung)
  - Intelligenzforschung (Psychologie)
  - Populationspharmakokinetik (Arzneimittelforschung)
  - Adaptive und gruppensequenzielle Verfahren
  - Diagnostische Studien mit räumlicher Datenstruktur und zeitlicher Verlaufskontrolle (Perimetrie in der Augenheilkunde)
  - Klinische Dosisfindungsstudien
  - Statistik in industriellen Anwendungen
  - Multivariate Äquivalenz und Nichtunterlegenheit

#### **4. Serviceangebot**

##### **Beratung und Unterstützung bei allen statistischen Fragestellungen**

Das Institut für Mathematische Stochastik bietet Beratung zur Planung und statistischen Auswertung von Experimenten an, insbesondere:

- zur Unterstützung von Abschlussarbeiten
- bei der Konzeption und Durchführung von Studien
- bei der Stichproben-/ Versuchsplanung, Datengewinnung und Sicherstellung der Datenqualität
- bei der Auswahl und Anwendung geeigneter Analysemethoden
- bei der Interpretation und Präsentation der Untersuchungsergebnisse

**Dieses Angebot richtet sich an ...**

- Studierende und Promovierende der Otto-von-Guericke-Universität Magdeburg (OVGU)
- Ausgenommen sind Personen, die mit dem Fachbereich Medizin assoziiert sind. (Das Universitätsklinikum bietet über das Institut für Biometrie und Medizinische Informatik Statistikberatungen an.)

<http://www.statistik.ovgu.de/Statistische+Beratung.html>

## 5. Kooperationen

- Dr. Frenkel, Beer Sheva, Israel Sami Shamoan College of Engineering, Israel
- Priv.-Doz. Dr. Ekkehard Glimm, Novartis Pharma AG, Basel
- Prof. Dr. Anton Wakolbinger, Goethe-Universität Frankfurt
- Prof. Dr. Heinz Holling, Westfälische Wilhelms-Universität Münster
- Prof. Dr. Peter Pfaffelhuber, Albert-Ludwigs-Universität Freiburg

## 6. Forschungsprojekte

**Projektleiter:** Prof. Dr. Norbert Gaffke

**Projektbearbeiter:** Prof. Dr. Norbert Gaffke

**Kooperationen:** Prof. Dr. Rainer Schwabe, OVGU, FMA-IMST

**Förderer:** Fördergeber; 01.10.2015 - 28.09.2020

### **Algorithmen zum optimalen Design für lineare Regressionsmodelle.**

Im Rahmen der approximativen Design-Theorie für lineare Regressionsmodelle sollen optimale Designs algorithmisch berechnet werden (insbesondere D-optimale und I-optimale Designs). Ein universell einsetzbarer Algorithmus existiert nicht. Ob die  $\{\text{vorhandenen}\}$  Algorithmen zur Anwendung kommen können, hängt von der Komplexität des Modells ab und erfordert ggf. weiteren theoretischen Input. Im Projekt sollen unsere Quasi-Newton Methoden (s. Gaffke, Graßhoff, Schwabe, 2014) auf zwei Modellklassen angewendet werden: Zum Einen Querschnitts-Designs bei longitudinalen Daten, z.B. im Kontext von "accelerated life testing"-Untersuchungen in der Qualitätskontrolle (vgl. Weaver and Meeker, 2014). Zum anderen der Fall eines  $\{\text{endlichen}\}$  Versuchsbereichs, wobei auch Stratifizierungs- oder Kostenrestriktionen einbezogen werden. Hierfür sind in den letzten Jahren Algorithmen vom Silvey-Titterington-Torsney Typ wieder aufgegriffen worden (vgl. Harman, 2014). Diese wollen wir mit unseren Quasi-Newton Methoden kontrastieren.

Literatur:

Gaffke,N.; Graßhoff,U.; Schwabe,R.: Algorithms for approximate linear regression design applied to a first order model with heteroscedasticity.

Computational Statistics and Data Analysis 71 (2014),1113-1123.

Weaver,B.P.; Meeker, W.Q.: Methods for Planning Repeated Measures Accelerated Degradation Tests.

Applied Stochastic Models in Business and Industry 30 (2014), 658-671,

Harman,R.: Multiplicative methods for computing D-optimal stratified designs of experiments.

Journal of Statistical Planning and Inference 146 (2014), 82-94.

---

**Projektleiter:** Prof. Dr. Norbert Gaffke

**Projektbearbeiter:** Prof. Dr. Norbert Gaffke

**Förderer:** Fördergeber; 01.04.2015 - 31.03.2016

### **Die Verteilung einer nicht-parametrischen Teststatistik für den Erwartungswert.**

Im nicht-parametrischen Modell mit  $n$  i.i.d. nicht negativen Zufallsvariablen wurde von Gaffke (2005) eine Teststatistik vorgeschlagen, die auch die Konstruktion einer unteren Konfidenzschranke für den Erwartungswert erlaubt. Immer noch offen ist die Frage, ob das nominelle Konfidenzniveau eingehalten wird. Äquivalent ist die Frage, ob die Verteilung der Teststatistik stochastisch größer (oder gleich) der Standard-Rechteck-Verteilung ist, unter jeder zu Grunde liegenden Verteilung mit Erwartungswert gleich 1. Es besteht einige numerische Evidenz, dass die Antwort positiv ist. Bewiesen ist aber wenig: Nur der Fall  $n = 2$  (und der triviale Fall  $n = 1$ ), sowie die asymptotische Aussage, dass für  $n \rightarrow \infty$  die Verteilung der Teststatistik gegen die Standard-Rechteck-Verteilung konvergiert. Das erste Ziel des Projektes ist es, den Fall  $n = 3$  zu beantworten. Sollte hier die Antwort positiv sein, wird das Projekt in die Verlängerung gehen.

Literatur:

Gaffke, N. (2005): Three test statistics for a nonparametric one-sided hypothesis on the mean of a nonnegative random variable.

Mathematical Methods of Statistics 14, 451-467.

---

**Projektleiter:** Prof. Dr. Gerd Christoph

**Projektbearbeiter:** Prof. Dr. Gerd Christoph

**Kooperationen:** Prof. V.V. Ulyanov, Staatl. Lomonossov-Universität Moskau

**Förderer:** Fördergeber; 17.01.2011 - 31.12.2015

### **Exakte Fehlerberechnung bei Approximationen in Statistischen Anwendungen**

In Approximationen für den Korrelationskoeffizienten und weiteren statistischen Kerngrößen werden numerisch exakte Konstanten bestimmt, die in Anwendungen Verbesserungen z.B. für Konfidenzintervalle erlauben. Mit Hilfe von Cornish-Fisher-Entwicklungen werden Quantile einer unbekanntem Verteilung approximiert und Fehler analysiert.

---

**Projektleiter:** Prof. Dr. Gerd Christoph

**Projektbearbeiter:** Dr. Nadja Malevich

**Förderer:** Haushalt; 01.01.2011 - 31.12.2015

### **Konvergenzaussagen für zufällige Summen unabhängiger Zufallsgrößen mit schweren Flügeln**

Untersucht wird das exakte Konvergenzverhalten von Summen mit einer zufälligen Anzahl unabhängiger Zufallsgrößen, wenn die Zufallsgrößen Pareto-ähnliche Verteilungen besitzen, insbesondere wenn Erwartungswert und/oder Streuung nicht existieren.

Dazu werden ungleichmäßige Abschätzungen in klassischen asymptotischen Entwicklungen mit stabiler Grenzverteilung wesentlich verbessert.

Anwendungen finden sich in der Finanz- und Risikotheorie.

---

**Projektleiter:** Prof. Dr. Jürgen Groß

**Projektbearbeiter:** Prof. Dr. Jürgen Groß

**Kooperationen:** Dr. Annette Möller, Georg-August-Universität Göttingen

**Förderer:** Haushalt; 01.04.2015 - 31.03.2016

### **Methoden der Zeitreihenanalyse zur statistischen Nachbereitung von Ensemble-Prognosen**

Für die Vorhersage von Wettervariablen wie etwa Temperatur, Windgeschwindigkeit oder Niederschlag werden NWP (numerical weather prediction) Modelle verwendet. Um die Unsicherheit der Vorhersage quantifizieren zu können, werden die Modelle mehrfach unter Abwandlung von Modellformulierungen und/oder Randbedingungen angewendet, woraus sich ein sogenanntes Vorhersage-Ensemble ergibt. Um eine Verbesserung des Ensembles hinsichtlich möglicher Verzerrungen und unzureichenden Dispersions-Verhaltens zu erhalten, werden unter anderem statistische Nachbereitungsverfahren verwendet (statistical postprocessing methods). Das Ziel des Projektes ist es zu untersuchen, inwieweit die Anwendung von Methoden der Zeitreihenanalyse bei der Beschreibung von Vorhersagefehlern der einzelnen Ensemble-Mitglieder einen zusätzlichen Beitrag zur verbesserten statistischen Nachbereitung leisten können.

---

**Projektleiter:** Prof. Dr. Claudia Kirch

**Projektbearbeiter:** Kerstin Reckrühm

**Förderer:** Haushalt; 01.04.2015 - 31.12.2018

**Die Detektion multipler Strukturbrüche basierend auf dem MOSUM-Verfahren**

Es existieren zwei grundlegende Verfahren zur Erkennung multipler Strukturbrüche in Zeitreihen im klassischen Modell der Erwartungswertänderung, die binäre Segmentierung und das MOSUM-Verfahren. Das Segmentierungsverfahren ist eine iterative Methode, die ausnutzt, dass Tests für Ein-Change-Point-Alternativen weiterhin Macht im Fall von multiplen Änderungen besitzen. Die zweite Methode hingegen basiert auf Statistiken, die gleitende Summen verwenden. Ein Vorteil des MOSUM-Verfahrens besteht darin, dass das Gesamtsignifikanzniveau kontrolliert werden kann. Tests und statistische Eigenschaften von Change-Point Schätzern, die auf derartige Statistiken gleitender Summen basieren, wurden von Kirch und Muhsal (2015+) im klassischen Erwartungswert-Modell detailliert untersucht. Diese Resultate sollen nun für verschiedene Change-Point Situationen verallgemeinert werden. Durch die Verwendung von MOSUM-Statistiken basierend auf Schätzfunktionen können Modelle verschiedener Parameteränderungen in ein Erwartungswert-Modell der Schätzfunktion transformiert werden. Dazu muss lediglich der globale Schätzer ermittelt werden, was einen großen Vorteil in Bezug auf den Rechenaufwand darstellt. Wir konstruieren eine entsprechende Teststatistik und analysieren ihr asymptotisches Verhalten unter der Nullhypothese und Alternativen. Weiterhin werden die zugehörigen Change-Point Schätzer hinsichtlich ihrer Konsenzeigenschaften näher untersucht.

Das Hauptproblem des MOSUM-Verfahrens besteht darin, dass die Güte dieser Methode im Wesentlichen von der Wahl der Bandbreite  $G$  abhängt. Dies erweist sich insbesondere dann als sehr problematisch, wenn die Abstände zwischen den Change-Points stark variieren. So eignen sich große Bandbreiten zur Detektion kleiner Änderungen und kleine Bandbreiten zur Erkennung großer Änderungen. Eine Lösungsmöglichkeit wurde kürzlich im Zusammenhang mit Änderungen in Punktprozessen von Messer et al. (2014) vorgeschlagen. Ein Multiskalenverfahren basierend auf MOSUM-Statistiken soll dementsprechend konstruiert und untersucht werden. Da es für dieses Verfahren bisher noch keinerlei theoretische Untersuchungen gibt, wollen wir hier zunächst bei dem einfachen Erwartungswert-Modell bleiben.

---

**Projektleiter:** Prof. Dr. Claudia Kirch

**Projektbearbeiter:** Silke Weber

**Förderer:** Fördergeber; 01.10.2013 - 30.09.2016

**Resampling-Verfahren für hochdimensionale Change-Point-Tests abhängiger Daten**

In vielen wissenschaftlichen Disziplinen werden zunehmend hochdimensionale Daten beobachtet, also solche, bei denen die Anzahl der beobachteten Variablen ähnlich groß oder sogar deutlich größer als die Länge der Stichprobe ist. Beispiele hierfür finden sich etwa in der Medizin im Zusammenhang mit bildgebenden Verfahren (funktionale Magnet-Resonanz-Tomographie, hochdimensionale oder hochauflösende EEG-Daten etc.) aber auch in den Wirtschaftswissenschaften (Hochfrequenz-Finanzdaten, Panel-Daten etc.). Klassische multivariate statistische Verfahren sind zur Analyse hierfür nicht geeignet, da sie typischerweise voraussetzen, dass die Dimension deutlich kleiner als die Stichprobenlänge ist. Daher ist die Entwicklung statistisch korrekter Verfahren für solche Daten sowie deren mathematische Analyse von zunehmendem Interesse.

Eine wichtige Rolle für die weitere statistische Analyse spielt hierbei die Frage, ob die beobachteten Daten sich zeitlich stationär verhalten, oder ob strukturelle Änderungen auftreten. Mit dieser Frage wollen wir uns in diesem Projekt genauer beschäftigen, neue Verfahren entwickeln sowie aktuelle Ansätze mittels Resampling-Verfahren verbessern. Bei letzteren handelt es sich um einen computerintensiven Ansatz, der die statistischen Eigenschaften von Tests bei kleineren Stichprobenumfängen häufig verbessert. Die mathematische Gültigkeit solcher Verfahren wurde jedoch bisher im Zusammenhang mit solchen Tests auf zeitliche Stationarität lediglich im univariaten Fall untersucht und erfordert insbesondere bei hochdimensionalen bzw. Panel-Daten, bei denen die Dimension mit der Stichprobenlänge wächst, neue Methoden.

---

**Projektleiter:** Prof. Dr. Claudia Kirch

**Projektbearbeiter:** Christina Stöhr

**Förderer:** Haushalt; 01.04.2015 - 31.03.2018

**Robuste Verfahren in der sequentiellen Change-Point Analyse basierend auf U-Statistiken**

Change-Point Analyse dient dazu, Strukturbrüche in zeitlich geordneten Datensätzen zu erkennen. Hierzu gibt es die folgenden zwei Ansätze: Einerseits verwendet das a posteriori (offline) Verfahren den vollständig beobachteten Datensatz zum Testen auf einen Change-Point. Andererseits gibt es das sequentielle (online) Verfahren, bei dem eine

sogenannte Testperiode bekannt ist und darauf aufbauend neue Daten gesammelt werden. Nach jeder neuen Beobachtung wird erneut ein Test auf eine Strukturveränderung vorgenommen. Die Entscheidung, ob eine Veränderung eingetreten ist oder nicht, wird mittels einer Teststatistik gefasst. Ist dieser größer als ein kritischer Wert, entdeckt der a-posteriori Test eine Veränderung. Im Falle des sequentiellen Verfahrens wird statt eines kritischen Wertes eine kritische Kurve verwendet. Sobald die Teststatistik über der kritischen Kurve liegt, stoppt das Verfahren und entdeckt eine Veränderung.

Robuste Verfahren in der Change-Point Analyse sind von großer Bedeutung, um auch beim Auftreten stark abweichender Beobachtungen können beispielsweise durch schiefe Verteilungen, dicke Flanken oder Ausreißer verursacht werden.

Insbesondere in der online Change-Point Analyse sind robuste Verfahren bedeutsam, da beispielsweise bei der Überwachung von Patienten- oder Maschinendaten bei Erkennung eines Strukturbruchs schnell eingegriffen werden muss. Ist das verwendete Verfahren nicht robust, können Ausreißer in den Daten leicht einen falschen Alarm auslösen. In diesem Projekt konstruieren wir robuste sequentielle Verfahren basierend auf U-Statistiken. Dazu muss zunächst die entsprechende Teststatistik konstruiert werden. Für diese Teststatistik werden dann die asymptotischen Verteilungen unter der Null- und Alternativhypothese hergeleitet. Aus den asymptotischen Verteilungen können schließlich die kritische Kurve sowie die Güte des Tests abgeleitet werden. Das Verhalten des Tests bei endlichen Stichproben wird mittels Simulationen untersucht.

---

**Projektleiter:** Prof. Dr. Rainer Schwabe

**Projektbearbeiter:** Fritjof Freise

**Förderer:** Fördergeber; 01.04.2014 - 31.01.2016

#### **Adaptives Design II**

Durch eine geeignete Wahl der Versuchsbedingungen kann in vielen statistischen Experimenten eine wesentliche Verbesserung der Analyseergebnisse bzw. eine deutliche Verringerung der Kosten für die Durchführung des Experiments erzielt werden. Liegen nichtlineare Wirkungszusammenhänge zwischen den Versuchsbedingungen und der die Zielvariable beschreibenden Wirkungsfunktion vor, ergibt sich dabei das Problem, dass die optimalen Versuchspläne, d.h. die optimale Wahl der Versuchseinstellungen, in der Regel von den unbekannt und zu schätzenden Parametern abhängen. Während dies bei einstufig geplanten Experimenten ein schier unlösbares Problem darstellt, bieten adaptive und sequenzielle Verfahren, die "on-line" die Information zuvor gemachter Beobachtungen ausnutzen, einen vielversprechenden Ansatz, um auch in solchen Situationen mit möglichst wenigen Messungen zu möglichst genauen Schätzungen zu gelangen. Derartige Verfahren sollen im Rahmen des vorliegenden Projektes entwickelt und auf ihre Eigenschaften unter realen Versuchsbedingungen untersucht werden, wobei der Schwerpunkt auf Anwendungen in sogenannten Dosis-Wirkungs-Modellen liegt, bei denen eine binäre Zielvariable, die den Erfolg oder Misserfolg einer Behandlung beschreibt und daher nur zwei Ausprägungen annehmen kann, in Abhängigkeit von der Größe ("Dosis") einer oder mehrerer erklärender Variablen untersucht wird. Neben Experimenten in der Psychophysik stellen adaptive Intelligenztests, wie sie im Projekt "Optimales Design für online generierte adaptive Intelligenztestverfahren" untersucht und weiterentwickelt werden, ein wichtiges Anwendungsgebiet dar.

---

**Projektleiter:** Prof. Dr. Rainer Schwabe

**Projektbearbeiter:** Dipl.-Math. Björn Holzhauer

**Kooperationen:** Dr. Ekkehard Glimm, Novartis Pharma AG, Basel

**Förderer:** Fördergeber; 01.02.2015 - 31.01.2018

#### **Metaanalyse von unerwünschten Ereignissen in klinischen Studien basierend auf Aggregatdaten**

Zur Charakterisierung der Nebenwirkungen von medizinischen Behandlungen ist es von Interesse, die Evidenz aus mehreren klinischen Studien zu kombinieren und auch historische Daten über die Kontrollgruppe zu berücksichtigen, weil in jeder einzelnen Studie oft nur wenige unerwünschte medizinische Ereignisse auftreten. Wenn für alle Studien detaillierte Patientendaten verfügbar sind, werden üblicherweise Überlebenszeitmethoden angewandt, um das Auftreten von unerwünschten Ereignissen unter Berücksichtigung der Dauer des Beobachtungszeitraums zu analysieren. Dies ist besonders dann wichtig, falls die Beobachtungsdauer sich zwischen Behandlungsgruppen unterscheidet oder falls die Austauschbarkeit von Modellparametern zwischen Studien von verschiedener Länge angenommen wird. Traditionelle Überlebenszeitmethoden können allerdings nicht angewandt werden, wenn für einige Studien lediglich Aggregatdaten verfügbar sind. Dies stellt ein Problem in der Metaanalyse dar, da Metaanalysen meist auf Veröffentlichungen in medizinischen Fachzeitschriften beruhen, welche in aller Regel keine individuellen

Patientendaten enthalten.

Ziel dieses Projektes ist es, statistische Methoden zu entwickeln und zu evaluieren, die Überlebenszeitmetaanalysen unter Berücksichtigung historischer Studien basierend auf Aggregatdaten ermöglichen.

---

**Projektleiter:** Prof. Dr. Rainer Schwabe

**Projektbearbeiter:** Dr. Maryna Prus

**Förderer:** Haushalt; 15.10.2015 - 15.04.2016

**Optimales Design für die individuelle Vorhersage von Inter- und Extrapolation in Regressionsmodellen mit zufälligen Parametern**

Regressionsmodelle mit zufälligen Parametern werden in statistischen Anwendungsbereichen, insbesondere in den Biowissenschaften, sowie in individualisierter Medizinforschung häufig verwendet. In diesen Modellen ist neben der Schätzung des Populationsparameters die Vorhersage individueller zufälliger Parameter von Interesse. Der individualisierte Ansatz ist vor allem für die Studien, in denen nur wenige Beobachtungen pro Individuum möglich sind, von großer Bedeutung. Letzteres ist für medizinische Fragestellungen, z.B. bei Untersuchungen von Blutentnahmen bei Patienten, besonders relevant. Ziel dieses Projektes ist es, optimale oder zumindest effiziente und anwendbare Versuchspläne (Designs) für die Vorhersage von Inter- und Extrapolation individueller Wirkungsfunktionen in Regressionsmodellen mit zufälligen Parametern zu entwickeln.

---

**Projektleiter:** Prof. Dr. Rainer Schwabe

**Projektbearbeiter:** Dipl.-Math. Martin Radloff

**Kooperationen:** Dr. Hermann Kulmann, Bayer, Berlin; Dr. Thomas Schmelter, Bayer, Berlin; Dr. Tobias Mielke, Aptiv Solutions, Köln

**Förderer:** Haushalt; 01.10.2013 - 30.09.2016

**Optimales Design für Dynamische Systeme**

Viele Wirkungszusammenhänge können auch in technischen Anwendungen nur mit statistischen, d.h. zufälligen Abweichungen beobachtet werden. Diese Wirkungszusammenhänge sind oft nicht explizit darstellbar, sondern nur durch eine oder mehrere Differentialgleichungen gegeben, bei denen einige Modellparameter unbekannt sind. Anhand von beobachteten Daten sollen diese Modellparameter mittels geeigneter Schätzverfahren bestimmt werden. Die Qualität dieser Schätzung hängt im Wesentlichen auch vom Design, d.h. von der Wahl der Versuchseinstellungen und der Messzeitpunkte ab. Ziel dieses Projektes ist es, Strategien zur Bestimmung optimaler oder zumindest effizienter Designs zu entwickeln und diese zu validieren.

---

**Projektleiter:** Prof. Dr. Rainer Schwabe

**Projektbearbeiter:** Dipl. Math. Dennis Schmidt

**Kooperationen:** Prof. Dr. Waltraud Kahle, Otto-von-Guericke-Universität Magdeburg, Institut für Mathematische Stochastik

**Förderer:** Fördergeber; 01.04.2013 - 31.03.2016

**Optimales Design für statistische Modelle mit zensierten Daten**

In vielen technischen und biologischen Bereichen spielt die statistische Analyse zensierter Daten eine zunehmende Rolle. Diese Zensierungen können deterministisch (feste Studiendauer, Nachweisgrenzen) oder zufallsgesteuert (zufällige Studiendauer, zufälliges Ausscheiden aus der Studie) sein. Die beobachteten, teilweise zensierten Größen können zusätzlich von weiteren Einflussfaktoren (Behandlungen und Kovariablen) abhängen, was beispielsweise über ein "proportional hazards"-Modell beschrieben werden kann.

Während die statistische Analyse derartiger Daten schon relativ weit entwickelt ist, gibt es relativ wenig Resultate zur effizienten Planung derartiger Studien oder Experimente. Ziel des vorliegenden Projekts ist es, für eine Reihe von relevanten Modellsituationen optimale oder zumindest effiziente Designs zu charakterisieren und analytisch zu bestimmen, um Anleitungen für eine möglichst effektive Ausnutzung der vorhandenen Ressourcen beim Vorliegen zensierter Daten bereit zu stellen.

---

**Projektleiter:** Prof. Dr. Rainer Schwabe



**Projektbearbeiter:** Dipl.-Math. Marius Schmidt

**Kooperationen:** Dr. Hermann Kulmann, Bayer, Berlin; Dr. Thomas Schmelter, Bayer, Berlin; Dr. Tobias Mielke, Aptic Solutions, Köln; Priv.-Doz. Dr. Steffen Uhlig, Quo Data, Dresden; Prof. Dr. Heinz Holling, Universität Münster, Institut für Psychologie IV

**Förderer:** Haushalt; 01.10.2013 - 30.09.2016

#### **Optimales Design für verallgemeinerte lineare gemischte Modelle**

Gemischte Modelle spielen zunehmend eine wichtige Rolle nicht nur in Biowissenschaften sondern auch bei wirtschafts- und sozialwissenschaftlichen Fragestellungen, um individuelle Effekte der verschiedenen Beobachtungseinheiten als Repräsentanten einer größeren Grundgesamtheit bei der statistischen Datenanalyse berücksichtigen und erfassen zu können. Mit verallgemeinerten linearen gemischten Modellen werden Zusammenhänge für binäre ("Erfolg - Misserfolg") und diskrete Zielgrößen ("Anzahlen") beschrieben, die nicht sinnvoll durch standardmäßige lineare gemischte Modelle für metrische Daten dargestellt werden können. Für die zufälligen Effekte können dann neben normalverteilten individuellen Einflüssen auch solche aus konjugierten Familien angenommen werden, die eine explizitere Analyse erlauben. Wie in allen statistischen Analysen hängt auch hier die Qualität der Ergebnisse wesentlich vom Beobachtungs- oder Experimentaldesign, d.h. der Wahl der Beobachtungseinheiten und Beobachtungszeitpunkte, ab. Ziel dieses Projektes ist es, optimale oder zumindest effiziente Designs für verallgemeinerte lineare gemischte Modelle zu entwickeln, die sowohl normalverteilte als auch Effekte aus konjugierten Verteilungen beinhalten können, und diese zu validieren.

---

**Projektleiter:** Prof. Dr. Rainer Schwabe

**Projektbearbeiter:** Iyad Aboud, MSc

**Kooperationen:** Prof. Dr. Heinz Holling, Universität Münster, Institut für Psychologie IV

**Förderer:** Fördergeber; 01.10.2015 - 30.09.2019

#### **Optimales Design für Zähldaten**

Neben klassischen Ansätzen stetiger, metrischer oder diskreter, binärer Daten (Messungen oder Ja/Nein-Antworten) spielt in statistischen Anwendungen die Analyse von Zähldaten eine zunehmende Rolle. Derartige Beobachtungen von Anzahlen treten zum Beispiel im Transportwesen, bei der Schadstoffmessung, in der Psychologie oder in medizinischen Anwendungen auf. Klassischerweise werden Anzahlen über Modelle mit Poisson-Verteilungen beschrieben, die Äquidispersion, d.h. Gleichheit von Mittelwert und Varianz, aufweisen. In der Praxis ist diese Annahme aber zu restriktiv, so dass alternativ häufig Modelle mit Überdispersion (Vorliegen einer größeren Varianz) und/oder exzessiven Nullen (zero inflation, verstärkte Beobachtung von Nullanzahlen) verwendet werden, um die Daten adäquat zu beschreiben. Während die Datenanalyse von Zähldaten relativ weit entwickelt und entsprechende Software verfügbar ist, gibt es nur wenige Resultate zur Planung von Experimenten mit Zähldaten. Diese beschränken sich auch bisher auf den klassischen Poisson-Ansatz oder auf ein spezielles Modell der Überdispersion mit negativ-binomialverteilten Daten, das sich als Mischmodell über eine konjugierte a-priori-Verteilung der Modellparameter ergibt. Dabei spielt die Planung in Experimentalsituationen eine immens wichtige Rolle, da nur mit einer vernünftigen Auswahl der Experimentaleinstellungen die vorhandenen Ressourcen effektiv ausgenutzt sowie Kosten und Aufwand verringert werden können. Ziel dieses Projektes ist es, optimale Designs für Zähldaten zu generieren, die auch beim Vorliegen von Überdispersion und/oder exzessiven Nullen eine effiziente Auswertung der Beobachtungen ermöglichen.

---

**Projektleiter:** Prof. Dr. Rainer Schwabe

**Projektbearbeiter:** Dipl.-Wirt.-Math. Maryna Prus

**Förderer:** Haushalt; 01.01.2011 - 30.09.2015

#### **Optimales Design zur individuellen Anpassung in gemischten Modellen**

In der Arzneimittelentwicklung besteht vorrangiges Interesse an Charakteristika einer Zielpopulation, um ein möglichst allgemein wirkendes Präparat auf den Markt bringen zu können. Neuere Entwicklungen zielen jedoch auch auf einen individualisierten Ansatz. Hierzu ist es notwendig die Charakteristika einzelner Individuen basierend sowohl auf den individualspezifischen Beobachtungen als auch auf den Populationseigenschaften möglichst genau zu spezifizieren. Letzteres ist insbesondere dann von Bedeutung, wenn aus ethischen oder technischen Gründen nur sehr wenig (invasive) Beobachtungen je Individuum gemacht werden können. Für diese Fragestellung sollen optimale Designs generiert werden, die eine effiziente Auswertung der Beobachtungen ermöglichen.

---

**Projektleiter:** apl. Prof. Dr. Waltraud Kahle  
**Förderer:** Haushalt; 01.01.2011 - 30.12.2015

### **Optimale Instandhaltung in Abnutzungsprozessen**

Wir betrachten einen Wiener Prozeß mit Drift als Abnutzungsmodell. Ein Ausfall tritt ein, wenn der Abnutzungsprozess erstmalig ein vorgegebenes Niveau  $h$  erreicht. Die zufällige Zeit bis zum Ausfall ist dann invers Gauss-verteilt.

Zur vorbeugenden Instandhaltung wird der Abnutzungsprozess regelmäßig kontrolliert. Wenn zu einem dieser Inspektionszeitpunkte die Abnutzung ein festgelegtes Maß  $a$  überschritten hat, wird das Bauteil vorbeugend durch ein neues ersetzt.

Dabei entstehen 3 Arten von Kosten:

- Inspektionskosten,
- Kosten einer vorbeugenden Instandhaltung,
- Ausfallkosten.

Inhalt des Projektes ist es, sowohl optimale Zeitintervalle zwischen den Inspektionen zu bestimmen, als auch eine optimalen Grenze  $a$  für die vorbeugenden Instandhaltung festzulegen.

---

**Projektleiter:** Dr. Heiko Großmann  
**Projektbearbeiter:** Bairu Zhang, MSc  
**Förderer:** Fördergeber; 06.01.2014 - 06.01.2018

### **Funktionale Datenanalyse von Ganganalyse-Daten**

Bestimmte neurologische Erkrankungen beeinträchtigen die Gehfähigkeit der betroffenen Individuen. In diesem Projekt werden Verfahren der funktionalen Datenanalyse entwickelt, um Daten zu analysieren, die mit Hilfe bildgebender Verfahren in einem Ganglabor bei Kindern und Jugendlichen erhoben werden. Im angewandten Teil des Projekts wird unter anderem untersucht, wie sich bestimmte medizinische Hilfsmittel (Orthesen) auf das Gehverhalten auswirken.

---

**Projektleiter:** Dr. Heiko Großmann  
**Projektbearbeiter:** Mohammad Zakir Hossain, MSc  
**Förderer:** Fördergeber; 01.11.2012 - 01.11.2016

### **Schätzverfahren für gemischte verallgemeinerte lineare Modelle bei kleinen Stichprobenumfängen**

Verallgemeinerte lineare Modelle mit festen und zufälligen Effekten bieten eine elegante Möglichkeit zur Modellierung abhängiger Beobachtungen. Bei der Schätzung der Modellparameter wird in der Regel angenommen, dass die zufälligen Parameter eine multivariate Normalverteilung besitzen. In diesem Projekt wird ein alternativer und speziell für kleine Stichprobenumfänge geeigneter Ansatz betrachtet, bei dem zwar, wie üblich, die bedingte Verteilung der abhängigen Variable bei gegebenen Werten der zufälligen Parameter zur Exponentialfamilie gehört, die Verteilung der zufälligen Effekte jedoch aus Randomisierungsüberlegungen abgeleitet ist. Für das sich ergebende semiparametrische Modell wird ein Schätzalgorithmus entwickelt. Weiterhin wird in Simulationsstudien numerisch untersucht, wie sich Verletzungen der Normalverteilungsannahme auf die Schätzungen auswirken.

## **7. Eigene Kongresse, wissenschaftliche Tagungen und Exponate auf Messen**

- Prof. Dr. Rainer Schwabe: 12th Workshop on Stochastic Models, Statistics and Their Applications, Organisation der Sektion "Methods and Applications of Designed Experiments", 2015, Wrocław
- Prof. Dr. Rainer Schwabe: 8th International Workshop on Simulation, Organisation der Sektion "Optimal Design in Item Response Theory", 2015, Wien

## **8. Veröffentlichungen**

### **Begutachtete Zeitschriftenaufsätze**

**Dörpinghaus, Meik; Gaffke, Norbert; Imhof, Lorens A.; Mathar, Rudolf**

A log-det inequality for random matrices

In: SIAM journal on matrix analysis and applications. - Philadelphia, Pa: Soc, Bd. 36.2015, 3, S. 1164-1179;

[Imp.fact.: 1,590]

**Großmann, Heiko**

Automating the analysis of variance of orthogonal designs

In: Computational statistics & data analysis. - Amsterdam: Elsevier Science, Bd. 70.2014, S. 1-18;

[Imp.fact.: 1,400]

**Jentsch, Carsten; Kirch, Claudia**

How much information does dependence between wavelet coefficients contain?

In: Journal of the American Statistical Association: JASA; the premier journal of statistical science. - Abingdon: Taylor & Francis, 2015; <http://dx.doi.org/10.1080/01621459.2015.1093945>;

[Imp.fact.: 3,052]

**Kirch, Claudia; Muhsal, Birte; Ombao, Hernando**

Detection of changes in multivariate time series with application to EEG data

In: Journal of the American Statistical Association: JASA; the premier journal of statistical science. - Abingdon: Taylor & Francis, Bd. 110.2015, 511, S. 1197-1216;

[Imp.fact.: 3,052]

**Kirch, Claudia; Tadjuidje Kamgaing, Joseph**

On the use of estimating functions in monitoring time series for change points

In: Journal of statistical planning and inference: JSPI. - Amsterdam: North-Holland Publ. Co, Bd. 161.2015, S. 25-49;

[Imp.fact.: 0,675]

**Schmidt, Dennis; Schwabe, Rainer**

Erratum to: On optimal designs for censored data

In: Metrika: international journal for theoretical and applied statistics. - Berlin: Springer, Bd. 78.2015, 3, S. 259;

[Imp.fact.: 0,517]

**Schmidt, Dennis; Schwabe, Rainer**

On optimal designs for censored data

In: Metrika: international journal for theoretical and applied statistics. - Berlin: Springer, Bd. 78.2015, 3, S. 237-257;

[Imp.fact.: 0,450]

**Schmidt, Marius; Schwabe, Rainer**

Optimal cutpoints for random observations

In: Statistics: a journal of theoretical and applied statistics. - London [u.a.]: Taylor & Francis, Bd. 49.2015, 6, S. 1366-1381;

[Imp.fact.: 0,532]

**Soumaya, Moudar; Gaffke, Norbert; Schwabe, Rainer**

Optimal design for multivariate observations in seemingly unrelated linear models

In: Journal of multivariate analysis: JMVA. - Orlando, Fla: Acad. Press, Bd. 142.2015, S. 48-56;

[Imp.fact.: 0,934]

***Nicht begutachtete Zeitschriftenaufsätze***

**Alonso Cabrera, Jesús; Schwabe, Rainer**

Optimal design in the presence of random block effects at baseline

In: Magdeburg: Univ., Fak. für Mathematik, 2015; 14 S.: graph. Darst. - (Preprint / Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg; 2015,17);

**Graßhoff, Ulrike; Holling, Heinz; Schwabe, Rainer**

Optimal design for the Rasch Poisson-Gamma model

In: Magdeburg: Univ., Fak. für Mathematik, 2015; 8 S.: graph. Darst. - (Preprint / Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg; 2015,14);

**Großmann, Heiko**

Partial-profile choice designs for estimating main and interaction effects of two-level attributes from paired comparison data

In: Magdeburg: Univ., Fak. für Mathematik, 2015; 24 S. - (Preprint / Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg; 2015,15);

**Prus, Maryna; Schwabe, Rainer**

Interpolation and extrapolation in random coefficient regression models - optimal design for prediction

In: Magdeburg: Univ., Fak. für Mathematik, 2015; 8 S.: graph. Darst. - (Preprint / Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg; 2015,13);

**Radloff, Martin; Schwabe, Rainer**

Invariance and equivariance in experimental design for nonlinear models

In: Magdeburg: Univ., Fak. für Mathematik, 2015; 8 S. - (Preprint / Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg; 2015,16);

**Schmidt, Dennis; Schwabe, Rainer**

Optimal design for multiple regression with information driven by the linear predictor

In: Magdeburg: Univ., Fak. für Mathematik, 2015; 13 S. - (Preprint / Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg; 2015,19);

**Soumaya, Moudar; Gaffke, Norbert; Schwabe, Rainer**

Optimal design for multivariate observations in seemingly unrelated linear models

In: Magdeburg: Univ., Fak. für Mathematik, 2015; 11 S. - (Preprint / Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg; 2015,5);

**Buchbeiträge**

**Graßhoff, Ulrike; Holling, Heinz; Schwabe, Rainer**

Poisson model with three binary predictors - When are saturated designs optimal?

In: Stochastic Models, Statistics and Their Applications. - New York, NY [u.a.]: Springer, S. 75-81, 2015 - (Springer proceedings in mathematics & statistics; 122);

**Grossmann, Heiko; Schwabe, Rainer**

Design for discrete choice experiments

In: Dean, Angela.: Handbook of design and analysis of experiments. - Hoboken: CRC Press, S. 787-832, 2015 - (CRC Handbooks of Modern Statistical Methods; 7);

**Soumaya, Moudar; Schwabe, Rainer**

On the impact of correlation on the optimality of product-type designs in SUR models

In: Stochastic Models, Statistics and Their Applications: Wrocław, Poland, February 2015; [papers collected on the occasion of the 12th Workshop on Stochastic Models, Statistics and Their Applications]. - Cham [u.a.]: Springer, S. 159-167 - (Springer Proceedings in Mathematics & Statistics; 122);

**Dissertationen**

**Malevich, Nadja; Christoph, Gerd [Gutachter]**

Approximations and asymptotic expansions for sums of heavy-tailed random variables. - Magdeburg, Univ., Fak. für Mathematik, Diss., 2015; 150 S.: graph. Darst.;

**Prus, Maryna; Schwabe, Rainer [Gutachter]**

Optimal designs for the prediction in hierarchical random coefficient regression models. - Magdeburg, Univ., Fak. für Mathematik, Diss., 2015; V, 87 S.: graph. Darst.;