

# Vorwort

Das vorliegende Textbuch gibt eine Einführung in Fragestellungen und Methoden der mathematischen Statistik. Es basiert auf Vorlesungen, die der Autor seit 1980 regelmäßig in mathematischen Studiengängen in Aachen, Essen, Münster und Freiburg gehalten hat. Ziel dieses Kurses ist es, ausgehend von der Datenanalyse und deren Motivation darzustellen, dass mit der Verwendung stochastischer Modelle ermöglicht wird, statistische und datenanalytische Verfahren zu bewerten und zu begründen. Die Statistik wird auf diese Weise in die Entscheidungstheorie und Spieltheorie eingeordnet und damit in besonderer Weise Gegenstand der mathematischen Behandlung. Diese Eingliederung erlaubt es, die klassischen statistischen Fragestellungen wie Test- und Schätzprobleme und Konfidenzbereiche einheitlich darzustellen und statistische Auswahlkriterien wie Bayes-, Minimax- und weitere spezifische Optimalitätskriterien systematisch zu entwickeln. Durch eine Reihe von motivierenden Beispielen, z.B. zum Problem der optimalen Auswahl zur Erkennung von gefälschten Folgen und Daten, zur Identifikation und Rekonstruktion von verrauschten Bildern soll der breite Horizont statistischer Fragestellungen skizziert und die Bedeutung der Statistik als praxisrelevante und weitreichende Theorie der Entscheidungen beschreiben werden.

Klassische und auch moderne Konstruktionsprinzipien für statistische Verfahren werden behandelt und begründet. Hierzu werden zunächst eine Reihe von grundlegenden statistischen Methoden und Begriffen eingeführt. Ein zentraler Begriff ist die Suffizienz, die den Informationsgehalt einer Statistik oder  $\sigma$ -Algebra beschreibt. Weitere wichtige Prinzipien sind die Reduktion durch Invarianz oder Äquivarianz, durch Unverfälschtheit oder Erwartungstreue und die Reduktion auf geeignete Schätz- und Testklassen. ‘Highlights‘ sind dann jeweils die entscheidungstheoretische Begründung für diverse statistische Verfahren wie z.B. die Zulässigkeit des arithmetischen Mittels in Normalverteilungsmodellen, die Optimalität von  $t$ -Test,  $F$ -Test, dem Pitman-Permutationstest und dem exakten Test von Fisher oder die Optimalität von Pitman-Schätzern und von  $U$ -Statistiken.

Das Textbuch versucht neben den klassischen Themengebieten auch bis zu einem gewissen Grad in Anwendungen einzuführen, die nicht unbedingt in Standardtexten zur mathematischen Statistik zu finden sind. Gegeben wird eine Einführung in grundlegende Prinzipien der asymptotischen Statistik. Es wird in den Kapiteln zur Schätz- und Testtheorie oder zu den Konfidenzbereichen in den ganzen

Text eingebunden gezeigt, dass die Grenzwertsätze der Wahrscheinlichkeitstheorie es ermöglichen, unter sehr allgemeinen Bedingungen approximative Tests und Konfidenzbereiche zu konstruieren und Eigenschaften von Schätzverfahren zu beschreiben. So wird z.B. die asymptotische Verteilung von Maximum-Likelihood-Schätzern oder von Martingalschätzern bestimmt und der Begriff der asymptotischen Effizienz von Schätzverfahren eingeführt. Detailliert wird beispielhaft für eine Reihe nichtparametrischer Schätzverfahren die Dichteschätzung behandelt. Gegeben wird auch eine Einführung in robuste und sequentielle Tests sowie in die Statistik von Zählprozessen wie z.B. die Methode der Martingalschätzer sowie die Martingalmethode für Anpassungstests. Themen der asymptotischen Entscheidungstheorie überschreiten jedoch den Rahmen dieser Einführung.

Diese Darstellung dient insbesondere dem Ziel, in die Vielfachheit und Breite statistischer Fragestellungen einzuführen, wie etwa in das Problem der sequentiellen Statistik, mit einer möglichst geringen Anzahl von Beobachtungen effiziente Entscheidungen zu treffen oder in das Problem der robusten Statistik trotz nur approximativ zutreffender Modelle zuverlässige Entscheidungen zu konstruieren. Die Statistik von Zählprozessen ist eine methodisch wichtige Erweiterung der Statistik, mit bedeutsamen Anwendungen z.B. in der Survival-Analyse oder allgemeiner für die Statistik von zeitabhängigen Ereignissen.

Behandelt werden im Text parametrische und nichtparametrische Modelle. Es werden auch einige neuere Anwendungen der Statistik angesprochen, wie z.B. auf Bildverarbeitung und Bildrekonstruktion oder auf das Quantile hedging in der Finanzmathematik.

Ein zentrales Ziel dieses Textbuches ist es zu zeigen, dass die Mathematische Statistik ein Gebiet mit vielen besonders schönen Ideen und Methoden und überraschenden Resultaten ist. Es ist reizvoll zu sehen, wie die Auswahl statistischer Verfahren schon viel über das zugrunde liegende Modell verrät. Dieses knüpft an die Begründung des Normalverteilungsmodells mit der Optimalität des arithmetischen Mittels durch Gauß und Laplace an. Die Charakterisierung von Modellen durch Eigenschaften statistischer Verfahren ist zentrales Thema des anspruchsvollen Buches von Kagan, Linnik und Rao (1973). Überraschend ist z.B. die von Stein entdeckte Nichtzulässigkeit des arithmetischen Mittels im Normalverteilungsmodell in Dimension  $d \geq 3$ . Verbesserungen lassen sich mit Hilfe von superharmonischen Funktionen konstruieren.

Ziel ist es auch besonders, der zunehmenden Spezialisierung in der Statistik entgegenzuwirken und eine breite Orientierung über unterschiedliche Gebiete und Themenkreise der Statistik zu geben. *Wie zuverlässig sind datenanalytische Methoden, wann können die Daten für sich sprechen, was ist die Bedeutung von  $p$ -Werten, wie lassen sich Abweichungen vom Modell in die statistische Analyse einbeziehen?* Auch mit diesen und ähnlichen für die Anwendungen bedeutsamen Fragen soll sich dieser Text befassen.

Der vorliegende Kurs baut auf grundlegenden Kenntnissen der Maß- und Wahrscheinlichkeitstheorie auf. Verwendete Methoden und Resultate der Funktionalanalysis, der Spieltheorie, über analytische Funktionen und lokal kompakte

Gruppen werden im Text vorgestellt. Etwa ab dem 5.ten Studiensemester können diese Themen sowohl in Bachelor- und Master- als auch in Staatsexamensstudiengängen mit Gewinn vermittelt werden. An den Kurs lassen sich sehr gut Spezialvorlesungen z.B. über asymptotische Statistik, Regressionsanalyse, Statistik von Prozessen, Bayessche Statistik oder Survival-Analyse anschließen.

Das Buch enthält mehr Stoff, als in einer vierstündigen Vorlesung behandelt werden kann. Neben den klassischen Themen einer Statistik-Vorlesung enthält es eine Reihe von weiterführenden Darstellungen und Entwicklungen. Insbesondere sind die Themengebiete aus den Kapiteln 9–13 nicht Standardthemen. Die zentralen Begriffe und Methoden aus den Kapiteln 1–5 sowie 7.1 und 8.1 lassen sich jedoch einfach für eine Vorlesung herausfiltern und werden im Eingang der Kapitel herausgestellt. Die weiterführenden Themen und Beispielklassen können dann Anlass für abschließende Anmerkungen oder Anregungen zum Selbststudium sein. Einige exemplarische Beispielklassen werden recht ausführlich behandelt und können in einem Vorlesungskurs stark gekürzt oder ausgelassen werden. Beispiele hierfür sind etwa die Behandlung von Gibbs-Maßen und Bildrekonstruktion in Kapitel 3.3, die detaillierte Darstellung verschiedener Schätzklassen in den Kapiteln 5.2–5.5 ebenso wie in 6.5 und 8.4. Aus den weiterführenden Kapiteln 9–13 lassen sich je nach der Vorgeschichte des Kurses vielleicht ein oder zwei Themen auswählen.

Von besonderer Bedeutung für den Autor dieses Bandes waren die klassischen Darstellungen von Ferguson (1967) zur statistischen Entscheidungstheorie, von Lehmann (1959, 1983) zur Testtheorie und Schätztheorie, von Zacks (1971) zu Entscheidungstheorie und zu Bayes-Verfahren und von Witting (1966) zur Optimierungstheoretischen Behandlung von Testproblemen. Dieses frühe Werk sowie die stark erweiterte Fassung von 1985 sind das wohl einflussreichste statistische Lehrbuch im deutschsprachigen Raum. Hermann Witting verdankt der Autor dieses Bandes seine Einführung in die Mathematische Statistik sowie sein Interesse an diesem Gebiet. Bedeutsam für den Autor waren auch die mathematische Theorie der Experimente von Heyer (1973) sowie die mathematisch besonders interessante Behandlung von Charakterisierungsproblemen in Kagan, Linnik und Rao (1973). Von besonderem Interesse waren auch die weiterführenden und orientierenden Darstellungen der (asymptotischen) statistischen Entscheidungstheorie in Pfanzagl und Wefelmeyer (1982) und Strasser (1985) sowie die Darstellung der asymptotischen Statistik in Witting und Müller-Funk (1995) und Liese und Miescke (2008). Die letztgenannten Werke gehen jedoch weit über den Rahmen dieses Buches hinaus. Die Darstellung der asymptotischen Statistik in Rüschendorf (1988) schließt thematisch eng an diesen Text an.

Eine erste Ausarbeitung von Teilen dieses Vorlesungstextes wurde 1998 und 2004 von Christian Lauer erstellt. Ihm sei hiermit herzlich gedankt. Besonderen Dank schulde ich auch Monika Hattenbach für ihre vorzügliche Arbeit am Erstellen und Gestalten des nahezu kompletten Textes. Danken möchte ich auch Janine Kühn, Viktor Wolf und Swen Kiesel für das Korrekturlesen einiger Buchkapitel sowie Hans Rudolf Lerche für eine Reihe wertvoller Hinweise, insbesondere zum Kapitel über sequentielle Tests.



<http://www.springer.com/978-3-642-41996-6>

Mathematische Statistik

Rüschendorf, L.

2014, XI, 427 S. 36 Abb., Softcover

ISBN: 978-3-642-41996-6