

Multivariate Statistics in Ecology and Quantitative Genetics

Titel	Multivariate Statistics in Ecology and Quantitative Genetics	
Typ	Vorlesung mit Übung	
Credits	5 ECTS	
Lehrform/SWS	2V + 2Ü	
Sprache	Englisch	
Modulniveau	Master	
Arbeitsaufwand	Präsenzstunden	60 Stunden
	Eigenstudium	90 Stunden
	Gesamtaufwand	150 Stunden
Angestrebte Lernergebnisse	<p>Die Studierenden sind in der Lage, statistisch-bioinformatische Methoden, insbesondere aus dem Bereich der generalisierten linearen Modelle (GLMs) und der gemischten (generalisierten) linearen Modelle zur Analyse multidimensionaler Daten geeignet auszuwählen und einzusetzen. Das umfasst auch Strategien zur Auswahl von statistischen Modellen. Die Studierenden kennen spezielle Varianten der Methoden für bioinformatische Analysen genetischer Daten sowie die Querbezüge zu den Modellen und Methoden der Quantitativen Genetik. Zudem sind sie in der Lage, Programmpakete für entsprechende Analysen genomischer und ökologischer Daten anzuwenden. Des Weiteren sind sie in der Lage, die erzielten Ergebnisse der Analysen sicher zu interpretieren und kritisch zu überprüfen. Sie sind in der Lage sowohl die Ergebnisse als auch die den Methoden zugrundeliegenden Modellannahmen präzise darzustellen, auch mit Hilfe mathematischer Formeln. Sie verfügen über die Grundlagen, Publikationen zu verstehen, in denen neue Varianten der behandelten Modelle und Methoden (z.B. für spezielle Anwendungen in der Ökologie, der Quantitativen Genetik oder in den „-omics“) mathematisch beschrieben werden.</p>	

Intended Learning Outcomes	<p>The students are trained to select, apply and assess statistical-bioinformatical methods based, in particular, on generalized linear models (GLMs) and (generalized) linear mixed models, for the analysis multidimensional data sets. This includes strategies of statistical model selection. The students know variants of these models that are specialized for the bioinformatic analysis of genetic data. Furthermore they know the connection of the methodological approaches to models and methods of quantitative genetics. They are able to use software packages to carry out such analyses for genomic and ecological data. They are able to critically assess and to interpret such results appropriately. They can precisely specify methods and results of such analysis as well as the underlying model assumptions, also in terms of mathematical formulae. They have the necessary understanding and background knowledge to read and understand publications in which new variant of such models and methods are presented and mathematically specified, e.g. for special applications in ecology, quantitative genetics, or in the „-omics“.</p>
Inhalt	<p>Behandelt werden verschiedene statistische Modelle und Methoden zur Analyse multidimensionaler Daten, unter anderem generalisierte lineare Modelle (GLMs), lineare gemischte Modelle (LMMs) und generalisierte gemischte lineare Modelle (GLMM). Es geht dabei um die Grundannahmen in diesen Modellklassen und um Strategien, geeignete Methodische Ansätze sowie Modelle für gegebene Daten anzuwenden. Behandelt werden auch aktuelle Spezialanpassungen und Erweiterungen dieser Modelle, die zur Auswertung von bestimmten Arten aktueller biologischer Daten verwendet werden, etwa in der Quantitativen Genetik, der Ökologie, sowie der Genomik und der Transkriptomik.</p>
Contents	<p>We address various statistical models and methods for the analysis of multidimensional data, such as generalized linear models (GLMs), linear mixed models (LMMs) and generalized linear mixed models (GLMMs). We discuss the underlying assumptions in these model</p>

	<p>classes and strategies to select model classes and to apply models to given data sets. We cover also special variants and extensions of these models for particular types of biological data sets, e.g. from quantitative genetics, ecology, genomic and transcriptomics.</p>
Prüfung	<p>Prüfungsleistung (benotet): -Klausur: 90 min</p> <p>Wiederholungsklausur zu Ende des Semesters. Details werden zu Beginn des Moduls bekannt gegeben.</p> <p>In der Klausur weisen die Studierenden nach, inwieweit sie die vorgestellten Modelle und Methoden verstanden haben, komprimiert wiedergeben und anwenden sowie auf verwandte Problemstellungen übertragen können. In der Klausur werden 7 bis 9 Aufgaben gestellt, die eine eigenständige Anwendung der Modelle und Methoden aus der Vorlesung erfordern (wie z.B. Datenanalyseergebnisse interpretieren, Modellannahmen mit Hilfe mathematischer Formeln präzise darstellen, Anhand von Visualisierungen und numerischen Analysen von Daten Strategien für verbesserte Modellanpassungen vorschlagen).</p>
Examination	<p>Written exam: 90 min</p> <p>Make-up exam at the end of the semester, details will be announced at the beginning of the semester.</p> <p>In the exam the students show to what extent they have understood the models and methods covered in the course, and how well they can concisely specify them, apply them to data and transfer them to related problems. The exam will consist of 7 to 9 assignments, which will require the application of models and methods covered in the course, e.g. by interpreting results of data analyses, specifying model assumptions with mathematical formulae, interpreting visual and numerical representations of data to propose improvements in the model.</p>
Literatur/Literature	<ul style="list-style-type: none"> • Mixed effects models and extensions in ecology with R / Alain F. Zuur, Elena N. Ieno, N. J. Walker, Anatoly A. Saveliev, Graham M. Smith. (Springer 2009) • Modern applied statistics with S / W. N.

	<p>Venables ; B. D. Ripley. - 4. ed. (Springer 2002)</p> <ul style="list-style-type: none"> • A Guide to QTL Mapping with R/qtl / Karl W. Broman & Saunak Sen (Springer 2009) • Analysing ecological data / Alain F. Zuur ; Elena N. Ieno ; Graham M. Smith. (Springer 2007) • Genetics and analysis of quantitative traits / Michael Lynch ; Bruce Walsh. (Sinauer 1998) • Statistical Genetics of Quantitative Traits / Wu, Ma & Casella. (Springer 2007) • Bioinformatics and Computational Biology Solutions Using R and Bioconductor / R. Gentleman, V.J. Carey, W. Huber, R.A. Irizarry, S. Dudoit (Eds.). (Springer 2005) • Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics / D. Sorensen, D. Gianola (Springer 2002) • Mixed effects models in S and S-PLUS / Jose C. Pinheiro ; Douglas M. Bates. (Springer, 2004)
Medienformen	Beamer-Präsentation, Tafelpräsentation, Handout, Arbeit am Computer
Media	slides show, blackboard presentation, handouts, computer lab
Lehr- und Lernmethode	<p>Lösen von Übungsaufgaben, Präsentation und Diskussion der Lösungen, Erproben von Software am Computer mit echten und simulierten Daten. Vorlesung, Tutorübung, Rechnerübung, Aufgaben zum Selbststudium.</p> <p>Das Modul besteht aus einer Vorlesung und Übungen in kleinen Gruppen (als Tutor- und Rechnerübungen). In den Aufgaben zum Selbststudium, die freiwillig abzugeben sind, analysieren die Studierenden die in der Vorlesung vorgestellten Methoden und Modelle. Mit echten Daten und Simulationsstudien erproben die Studierenden in den Rechnerübungen unter welchen Bedingungen die behandelten Methoden bzw ihre Implementationen in verschiedenen Software-Pakete (z.B. die R-Pakete lmer, qtl, vegan, DESeq, mcmcglmm) sinnvolle Ergebnisse liefern. Durch das Lösen von Aufgaben, die sich auf die behandelten theoretischen Konzepte beziehen vertiefen die Studierenden ihr Verständnis für die Inhalte der Vorlesung. Die Lösungen werden in den Übungen diskutiert. Dabei wird auch das Präsentieren von Datenanalyseergebnissen geübt.</p>
Teaching and Learning	The module consists of a lecture and exercise sessions

Methods	<p>(Übung) in small groups, in which solutions of assignments will be presented and discussed. In the assignments, which are optional for the students, the students can improve their understanding of the models and methods from the lecture by analyzing them theoretically and by exploring computer software, in which these methods are implemented, with real and simulated data. With such computer exercises the students will explore under what conditions the methods (or their implemetations in software packages such as the R packages lmer, qtl, vegan, DESeq and mcmcglmm) lead to reasonable results. By solving assignments that deal with the theoretical concepts covered in the lecture, the students improve their understanding of the contents of the lecture. The solutions will be discussed in the exercise sessions, such that also the presentation of results of data analyses will be practiced.</p>
Turnus	Sommersemester
Modulverantwortlicher	Prof. Dr. Dirk Metzler
Dozenten	Prof. Dr. Dirk Metzler