
Vorwort

Science is institutionalised scepticism.

Marc Rees, Astronomer Royal

Statistik, in meinem Verständnis, ist der formalisierte Versuch, sich nicht durch Artefakte, Zufall und Voreingenommenheit von echtem Erkenntnisgewinn abbringen zu lassen. Sie ist das systematische Zweifeln an den eigenen (und anderer) Daten. Wer als Wissenschaftler sich selbst gegenüber ehrlich ist, der will nicht einem Rauschen oder einer Tautologie¹ aufsitzen.

Jeder Datensatz hat irgendein Muster. Statistik stellt die Werkzeuge bereit zu testen, ob dieses Muster belastbar ist. Deshalb ist Churchills Ausspruch „Trau keiner Statistik, die Du nicht selbst gefälscht hast!“ ärgerlich. Er diskreditiert den Statistik-Betreiber als jemanden, der nur Unterschiede finden will. Grundvoraussetzung ist Ehrlichkeit. Wer sich selbst und andere betrügen will, sollte seine Daten erfinden; mit aufwändiger Statistik zu lügen ist selbstverständlich möglich, aber ineffizient. Nur ein selbstkritischer, zweifelnder Geist wird Statistik sinnvoll einsetzen – und einE guteR WissenschaftlerIn.² Nur für diese Menschen ist das vorliegende Machwerk gedacht.

Dieses Buch wendet sich an Laien ohne Vorkenntnisse, die, freiwillig oder gezwungenermaßen, die Grundlagen typischer statistischer Verfahren verstehen und anzuwenden lernen wollen. Der typische Leser ist ein Student im Grundstudium/BSc in einem umweltwissenschaftlichen Fachgebiet oder eine wissenschaftlich orientierte Person, die den statistischen Arbeitsbereich von der Pike auf lernen will. Die Struktur dieses Buches ist ein Wechselspiel aus Grundlagen und Umsetzungsbeispielen. Der eilige Leser kann schnell sehen, *wie* etwas umgesetzt werden kann (in den geradzahligen Kapiteln), während das Kapitel davor das Verständnis für das *Wieso* schaffen soll (in den ungeradzahligen Ka-

¹ Tautologie (Logik), eine Aussage, die, unabhängig vom Wahrheitswert der zugrunde liegenden Bestandteile, immer wahr ist (z. B.: „Es regnet oder es regnet nicht.“, <http://de.wikipedia.org/wiki/Tautologie>, 3.8.2011).

² Ich verzichte im weiteren auf die Gleichstellung von Mann und Frau in der Grammatik: das grammatikalische Geschlecht eines Wortes ist unabhängig vom Geschlecht der/des Bezeichneten.

piteln). Dieses *wieso* ist nicht durch mathematische Beweise als mehr durch praktisch Nachvollziehbarkeit motiviert.³

Eine Ausnahme bilden die Kapitel zur Wissenschaftsphilosophie und zum experimentellen Design (Kap. 13 und 14). Sie benutzen Ausdrücke, die erst im Laufe des Buchs eingeführt werden, sind aber eigentlich so grundlegend, dass sie ganz am Anfang stehen sollten. Vielleicht überfliegt der geneigte Leser diese Kapitel gleich zu Beginn und hebt sich die technischen Details für später auf.

Es gibt unzählige Quellen für statistische Verfahren. Viele Bücher (vor allem englische) sind im Literaturverzeichnis aufgeführt. Ergänzend will ich hier auf zwei „moderne“ Quellen verweisen: <http://www.khanacademy.org/#statistics>. Diese Seite bietet kurze (5–10 Minuten lange) Videocasts zu einem Thema an. Für Manchen mag diese Form der Inhaltsvermittlung bzw. -wiederholung genau geeignet sein. Ebenso sind die Wikipedia-Seiten zur Statistik häufig ausgesprochen gut (und wenn einmal nicht auf Deutsch, so doch auf Englisch). Allerdings sind sie meist sehr mathematisch und erst im unteren Teil mit Beispielen auch für Anfänger brauchbar. Einfach ’mal reinklicken.

Danksagung Dieses Buch erwuchs aus einigen Jahren der Lehrpraxis mit verschiedensten Bachelor- und MasterstudentInnen und DoktorandInnen. Während sie Statistik gelernt haben, habe ich viel über die Vermittlung von Statistik gelernt. Für die konstruktiven Rückmeldungen und für die Toleranz gegenüber experimentellen Darbietungsformen bin ich all diesen Menschen sehr dankbar. Das vorliegende Werk habe ich so an den Bachelor-StudentInnen der Studiengänge Geographie, Waldwirtschaft und Umweltnaturwissenschaften an der Universität Freiburg „getestet“ – und sie haben es für gut befunden. Danke!

Danken möchte ich auch Prof. Dr. Peter Pfaffelhuber, Freiburg, und Dr. Katharina Gerstner, Leipzig, für eine Durchsicht auf mathematisch-statistische Korrektheit. Mögliche weiterhin existierenden Fehler sind natürlich allein mir anzulasten.⁴

Schließlich danke ich meiner Familie, dass sie über Monate den Anblick ertragen hat, wie ich allabendlich über einen Klapprechner gekrümmt war. Auf diesem Rechner läuft (neben R) L^AT_EX.⁵ Beiden *open source*-Projekten bin ich extrem dankbar dafür, dass sie eine unermesslich kompetente Software entwickelt haben.

³ Ich mache auf den folgenden Seiten intensiv von Fußnoten Gebrauch. Sie konkretisieren zumeist einen Punkt oder erklären ein technisches oder mathematisches Detail. Für mich stellen sie das schriftliche Pendant zum Senken der Stimme in der Vorlesung dar: einen Einschub, der für das Verständnis nicht notwendig ist, der aber eine interessante oder für manche Situationen wichtige Information enthält.

⁴ PS: Rückmeldungen zu Fehlern, Ungenauigkeiten, Unvollständigkeiten oder auch zu Wünschen für Erweiterungen bitte an carsten.dormann@biom.uni-freiburg.de.

⁵ L^AT_EX (<http://ctan.org>) ist für die Textverarbeitung das, was R für die Statistik ist: *Simply the best*. Und wo wir gerade dabei sind: Wikipedia, Python, Windows, Excel, macOS, Google, Genstat, Stata, S-plus, Mathematica, Matlab und OpenOffice sind eingetragene Warenzeichen. Ihre Erwähnung bedeutet kein Gutheißen dieser Produkte.

Vorwort zur 2. Auflage

Diese Auflage ist vor allem eine Überarbeitung der ersten. Trotz intensiven Korrekturlesens hatten sich einige (Tipp-)Fehler eingeschlichen, die hier korrigiert wurden. Allen, die Hinweise auf Fehler gegeben haben, sei hiermit herzlich gedankt, vor allem Gregor Didenko und Jan Hanspach, die über die Hälfte aller Fehler entdeckt haben! Hinzugekommen sind neben vielen kleinen Ergänzungen etwa zur Momentenmethode oder zur Transformation der Antwortvariablen eine Box zur Residuendiagnostik für GLMs und die Berücksichtigung der *best-subset regression* bei der Modellselektion. Vielen Dank auch an Agnes Herrmann vom Springer-Verlag, die mich bei der Umsetzung des Buchs als epub begleitete.

Freiburg, 2017

Carsten F. Dormann

Parametrische Statistik

Verteilungen, maximum likelihood und GLM in R

Dormann, C.F.

2017, XXIII, 363 S. 132 Abb., 31 Abb. in Farbe.,

Softcover

ISBN: 978-3-662-54683-3