

Inhaltsverzeichnis (Stand: 29. September 2011)

Vorwort	i
I Beschreibende Statistik	1
Lernziele zu Teil I	2
1 Statistik, Daten und statistische Methoden	3
1.1 Statistik im privaten Alltag, in Politik und Gesellschaft	3
1.2 Aufgaben und Teilbereiche der Statistik	5
1.3 Methodenkompetenz als Lernziel	7
2 Grundbegriffe der Statistik	11
2.1 Statistische Einheit, Merkmal und Grundgesamtheit	11
2.2 Merkmalsklassifikationen und Skalen	14
2.3 Operationalisierung von Merkmalen	16
3 Datengewinnung und Auswahlverfahren	19
3.1 Erhebungsarten und Studiendesigns	19
3.2 Stichprobenauswahl	26
3.3 Träger amtlicher und nicht-amtlicher Statistik	29
4 Univariate Häufigkeitsverteilungen	31
4.1 Absolute und relative Häufigkeiten	31
4.2 Die empirische Verteilungsfunktion	43
5 Kenngrößen univariater empirischer Verteilungen	47
5.1 Lagemaße	47
5.2 Streuungsmaße	53
5.3 Quantile und Boxplots	57
6 Konzentration von Merkmalswerten	63
6.1 Die Lorenzkurve	63
6.2 Konzentrationsmaße	65
7 Index- und Verhältniszahlen	71
7.1 Verhältniszahlen	71
7.2 Zusammengesetzte Indexzahlen	74

8	Bivariate Häufigkeitsverteilungen	79
8.1	Darstellung empirischer Verteilungen für diskrete Merkmale	79
8.2	Empirische Unabhängigkeit diskreter Merkmale	85
8.3	Darstellung empirischer Verteilungen für stetige Merkmale	92
9	Zusammenhangsmaße	95
9.1	Nominalskalierte Merkmale	95
9.2	Metrische Merkmale	99
9.3	Ordinalskalierte Merkmale	107
II	Wahrscheinlichkeitsrechnung und schließende Statistik	109
	Lernziele zu Teil II	110
10	Zufallsvorgänge und Wahrscheinlichkeiten	111
10.1	Grundbegriffe der Wahrscheinlichkeitsrechnung	111
10.2	Zufallsstichproben und Kombinatorik	118
10.3	Bedingte Wahrscheinlichkeiten	121
11	Diskrete Zufallsvariablen	127
11.1	Wahrscheinlichkeitsfunktion und Verteilungsfunktion	127
11.2	Kenngrößen diskreter Verteilungen	133
11.3	Die Binomialverteilung	136
11.4	Die hypergeometrische Verteilung	141
12	Stetige Zufallsvariablen	147
12.1	Dichtefunktion und Verteilungsfunktion	147
12.2	Kenngrößen stetiger Verteilungen	150
12.3	Normalverteilung und Standardnormalverteilung	152
12.4	χ^2 -, t - und F -Verteilung	159
13	Bivariate Verteilungen von Zufallsvariablen	165
13.1	Unabhängigkeit von Zufallsvariablen	165
13.2	Kovarianz und Korrelation	169
14	Schätzung von Parametern	173
14.1	Punktschätzung und Eigenschaften von Schätzfunktionen	173
14.2	Schätzung von Erwartungswerten, Varianzen und Anteilswerten	176
14.3	Konfidenzintervalle für Erwartungswerte	179
15	Statistische Testverfahren	183
15.1	Arten statistischer Tests	183
15.2	Grundbegriffe und Gauß-Test für Erwartungswerte	185
15.3	t -Test für Erwartungswerte	196
15.4	χ^2 -Test für Varianzen	199
15.5	Zweistichproben-Tests für Erwartungswerte	200

16 Das lineare Regressionsmodell	203
16.1 Das einfache lineare Regressionsmodell	204
16.2 KQ-Schätzung im einfachen Regressionsmodell	206
16.3 Das Bestimmtheitsmaß	211
16.4 Das multiple lineare Regressionsmodell	213
16.5 KQ-Schätzung im multiplen Regressionsmodell	216
17 Grundzüge der Varianzanalyse	219
17.1 Das Modell der einfaktoriellen Varianzanalyse	221
17.2 Durchführung einer einfaktoriellen Varianzanalyse	222
17.3 Ausblick auf die zweifaktorielle Varianzanalyse	227
III Anhänge	229
Lernziele zu Teil III	230
18 Grundzüge der Matrizenrechnung	231
18.1 Grundbegriffe	231
18.2 Operationen mit Matrizen und Vektoren	233
18.3 Charakterisierung von Zufallsvektoren	238
19 Tabellenanhang	241
19.1 Verteilungsfunktion der Binomialverteilung	241
19.2 Verteilungsfunktion der Standardnormalverteilung	247
19.3 Quantile der Standardnormalverteilung	249
19.4 Quantile der χ^2 -Verteilung	250
19.5 Quantile der t -Verteilung	251
19.6 Quantile der F-Verteilung	252
20 Übungsaufgaben	257
20.1 Beschreibende Statistik	257
20.2 Wahrscheinlichkeitsrechnung und schließende Statistik	265
21 Lösungen zu den Übungsaufgaben	275
21.1 Beschreibende Statistik	275
21.2 Wahrscheinlichkeitsrechnung und schließende Statistik	284
22 Verzeichnisse und Internet-Ressourcen	299
22.1 Literaturverzeichnis	299
22.2 Ausgewählte Statistiklehrbücher mit Kommentaren	301
22.3 Multimedia-Software und Internet-Ressourcen	303
22.4 Symbolverzeichnis	307
22.5 Autorenregister	308
22.6 Sachregister	310

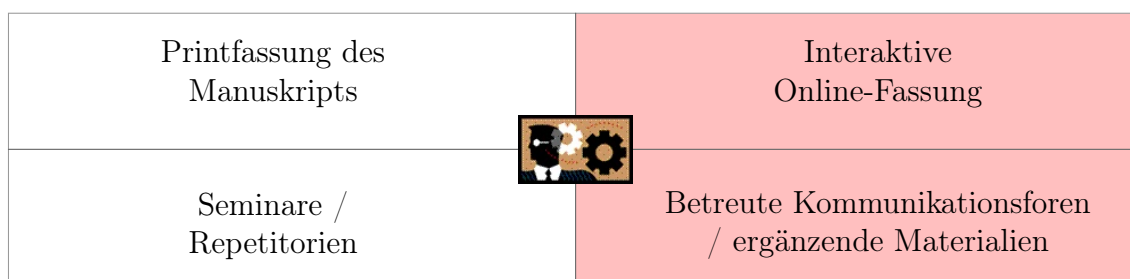
Vorwort

Dieser Kurs wird an der FernUniversität Hagen in den Bachelor-Studiengängen „Politik- und Verwaltungswissenschaft“, „Soziologie“ und „Psychologie“ im Grundstudium eingesetzt, seit dem Sommersemester 2011 auch im Studiengang „Bildungswissenschaft“. Er deckt alle Inhalte einer traditionellen Einführung in die Statistik ab, insbesondere also die beschreibende Statistik sowie Grundlagen der Wahrscheinlichkeitsrechnung und der schließenden Statistik. Die Bearbeitung des Lehrtextes soll dazu befähigen, statistische Informationen nutzen, sachadäquat interpretieren und verständlich kommunizieren zu können. Diese als **statistische Methodenkompetenz** bezeichnete Qualifikation ist in der heutigen Wissens- und Informationsgesellschaft unverzichtbar.

Herkunft dieses Lehrtextes

Der Kurs wurde mit dem Comenius-EduMedia-Siegel 2011 der Gesellschaft für Pädagogik und Information, einer Fachgesellschaft für Bildungstechnologie und Medienpädagogik, im Rahmen eines europäischen Multimedia-Wettbewerbs als exemplarisches Bildungsmedium ausgezeichnet. Er ist eine zentrale Komponente eines Medienmixes (**Blended-Learning-Ansatz**), der klassische Medien (Bücher, Präsenzlehre) und neue Medien (Internet, virtuelle Kommunikationsräume) verbindet. Der Kurs wird auch online zur Verfügung gestellt, dort als mehrfarbige pdf-Version. Die pdf-Fassung weist zahlreiche Verknüpfungen mit interessanten Web-Adressen sowie mit interaktiven oder dynamischen Elementen auf. Letztere ermöglichen es, etliche der in diesem Manuskript vorgestellten Verfahren anhand von benutzergesteuerten statistischen Experimenten unmittelbar „auszuprobieren“ oder statistische Konzepte über tongestützte Animationen nachzuvollziehen.

Realisierung eines Medienmixes



Komponenten des Medienmix

Die Experimente und Animationen stammen zum Teil aus einem großen Multimedia-Projekt „Neue Statistik“. ¹ Einige neuere Java-Applets und Animationen sind einem Anfang November 2011 erscheinenden kompakten Lernprogramm „Beschreibende Sta-

¹Das Projekt „Neue Statistik“ wurde mit öffentlichen Mitteln gefördert (s. auch unter www.fernuni-hagen.de/neuestatistik und unter www.neuestatistik.de). Im Rahmen des Projekts zeichnete die FernUniversität für statistische Experimente auf Java-Basis verantwortlich, die am Center für Digitale Systeme an der FU Berlin in Lernmodule mit Hypertextstruktur eingebettet wurden.

tistik“ auf CD entnommen.² Diese ergänzt in idealer Weise den ersten Teil dieses Kurses. Die CD hat insofern eine Brückenfunktion und eignet sich vor allem für Studierende, die noch Berührungängste gegenüber Mathematik und Statistik haben. Alle ergänzend angebotenen Materialien sind *optionale Angebote*, die aktives Lernen unterstützen und zur besseren Verständnissicherung beitragen sollen. Der vorliegende Kurs ist aber auch ohne Nutzung der Zusatzmaterialien uneingeschränkt studierbar.

Lernprogramm
„Beschreibende
Statistik“
als Einstiegshilfe

Betonung von
Interdisziplinari-
tät

In diesem Manuskript wird anhand zahlreicher Beispiele aus unterschiedlichen Anwendungsbereichen illustriert, dass die Statistik für viele Wissenschaften eine wichtige Servicefunktion erfüllt und alle Bereiche gesellschaftlichen Lebens durchdringt. Die verwendeten Beispiele – etwa die Darstellung von Ergebnissen der Nationalen Verzehrstudie II oder die Diskussion von Risiken beim Screening zur Krebsfrüherkennung – sind aktuell, relevant und motivierend. Bei der interaktiven pdf-Fassung sind die Beispiele manchmal mit Hintergrundinformationen aus Online-Ausgaben namhafter Zeitschriften verknüpft.

Einbindung von
Daten der
amtlichen
Statistik

Einige der in diesen Lehrtext integrierten Datensätze stammen aus der amtlichen Statistik, z. B. von Eurostat, dem Europäischen Amt für Statistik in Luxemburg, an dem der Autor vier Jahre als nationaler Sachverständiger tätig war. Die Daten illustrieren die Bedeutung statistischer Informationen für die Planung und das Monitoring nationaler und supranationaler Politiken. Dabei werden auch neuere Entwicklungen thematisiert und kritisch gewürdigt, etwa die zunehmende Verwendung zusammengesetzter Indizes bei internationalen Organisationen.

Bedeutung der
Icons am
Marginalienrand

Innerhalb des Kurses findet man am Marginalienrand Icons, die auf statistische **Experimente** oder tongestützte **Animationen** (Icon *Lautsprecher*) sowie auf besondere **Web-Links** hinweisen. Anklicken dieser Icons, die mit einem 3D-Effekt gestaltet wurden, führt bei der pdf-Version des Manuskripts direkt zu den genannten Elementen.



Icons mit Verlinkung: Statistisches Experiment, Animation mit Ton, Web-Link

Daneben findet man am Marginalienrand noch Icons, die nur der Orientierung dienen und nicht mit Links verknüpft sind. Diese Icons, die auf **Aufgaben**, **Lösungen** oder weiterführende **Literatur** verweisen, sind zur besseren Unterscheidung ohne 3D-Effekt ausgeführt.



Weitere Icons: Aufgabe, Lösung, Literatur

²Nähere Informationen zu dieser Software sind unter www.fernuni-hagen.de/e-statistik zu finden.

In das – inzwischen auch in einer Buchfassung vorliegende – Manuskript sind zahlreiche **Beispiele** und vereinzelt **Exkurse** integriert. Exkurse richten sich an Leser mit Interesse an detaillierteren zusätzlichen Informationen. Sie können übersprungen werden, ohne dass der rote Faden verloren geht. Bei der pdf-Fassung dieses Kurses ist es möglich, allen Verweisen auf nummerierte Gleichungen, Abbildungen, Tabellen und Aufgaben direkt per Mausclick nachzugehen. Bei der pdf-Fassung lassen sich auch die in den Text eingestreuten Web-Adressen, über grüne Rahmen sichtbar gemacht, unmittelbar ansteuern.

Der Kurs ist in drei Teile gegliedert. Der erste Teil thematisiert die beschreibende Statistik, während sich der zweite Teil den Grundlagen der Wahrscheinlichkeitsrechnung und der schließenden Statistik widmet. Daneben gibt es noch diverse Verzeichnisse und Tabellen, die den dritten Teil des Manuskripts konstituieren (Anhänge).

Struktur des
Manuskripts

Dank für die Programmierung der statistischen Experimente, auf die vielfach Bezug genommen wird, gebührt vor allem Herrn M. *Günster*, Arbeitsbereich „Statistik und quantitative Methoden“ der Fakultät für Kultur- und Sozialwissenschaften, und Herrn Dr. U. *Marty*, Arbeitsbereich „Interaktive Medien“ am Lehrgebiet „Praktische Informatik I“ (Prof. Dr. G. *Schlageter* / Dr. E. *Heuel*) an der FernUniversität sowie auch Frau B. *Schweiger*, Castrop-Rauxel. Herr H. *Kusenber*g, Schwerte, hat verschiedene Grafiken mit Photoshop bearbeitet und einige Flash-Animationen erstellt. Herrn A. *Hölzle*, München, gilt Dank für seine sehr engagierte und kompetente Mitwirkung bei der Gestaltung der L^AT_EX-Umgebung für dieses Manuskript. Ohne ihn und ohne die immer wieder benötigte zusätzliche Unterstützung durch Herrn Th. *Feuerstack*, L^AT_EX-Experte am Zentrum für Medien und IT der FernUniversität, hätte dieses Manuskript kaum die jetzige Form bekommen. Ein Beispiel zu bivariaten Häufigkeitsverteilungen stammt von Herrn Heinz *Böer*, Appelhülsen. Herr Prof. Dr. K.-H. *Renner*, Herr T. *Heydasch*, Herr Dr. R. *Jansen*, Herr Dr. H.-G. *Sonnenberg* und Frau M. *Zenza-Dobbert*, alle ebenfalls an der Fakultät für Kultur- und Sozialwissenschaften der FernUniversität Hagen tätig, lieferten wertvolle Anregungen und Korrekturhinweise. Letzteres gilt auch für Frau E. *Gaus-Faltings*, Braunschweig, Herrn Dr. J. *Rothenstein*, Wuppertal, und Herrn Dr. B. *Tewes*, Katholische Universität Eichstätt.

Danksagungen

Dank gebührt ferner den Firmen und Institutionen, die kostenfrei Bildmaterial zur Verfügung gestellt haben. Die Namen dieser Unternehmen und Einrichtungen sind an der Stelle vermerkt, an der das überlassene Material eingebunden ist. Es sind zu nennen:

Fa. Böhme und Weihs Systemtechnik GmbH, Sprockhövel (Herr Dr. N. *Böhme*)
 Center für Digitale Systeme, Berlin (Herr Prof. Dr. N. *Apostolopoulos*, Herr A. *Schulz*)
 Evonik Industries AG, Standort Essen (Herr Dr. W. *Wolfes*)
 Forschungsgruppe Wahlen, Mannheim (Herr B. *Weber*)
 Fa. Q-DAS GmbH, Weinheim (Herr Dr. E. *Dietrich*)
 GfK Telecontrol AG, Schweiz (Frau Dr. T. *Hackenbruch*)
 Hessischer Rundfunk, Frankfurt (Herr C. *Bender*)
 Statistisches Amt von Malta (Herr R. *Mizzi*)
 Fa. TNS Infratest, München (Herr M. *Kögel*).

Etliche Rückmeldungen zum Kurs - etwa Fehlerhinweise, Anregungen für inhaltliche Ergänzungen oder Kommentare zur Ausgestaltung der virtuellen Lernplattform - kamen von Studierenden, denen ich hierfür sehr zu Dank verpflichtet bin. Anregungen zur Weiterentwicklung des Manuskripts sind auch in der Zukunft willkommen.

Hagen, im September 2011

Hans-Joachim Mittag

joachim.mittag@fernuni-hagen.de