

Hans-Joachim Mittag

# Statistik

Eine interdisziplinäre Einführung

Fakultät für  
**Kultur- und  
Sozialwissen-  
schaften**

---

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung und des Nachdrucks, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (Druck, Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung der FernUniversität reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

# Inhaltsverzeichnis

<b>Vorwort</b>	<b>ix</b>
<b>I Beschreibende Statistik</b>	<b>1</b>
<b>Lernziele zu Teil I</b>	<b>2</b>
<b>1 Statistik, Daten und statistische Methoden</b>	<b>3</b>
1.1 Statistik im Alltag, in Politik und Gesellschaft . . . . .	3
1.2 Aufgaben und Teilbereiche der Statistik . . . . .	7
1.3 Methodenkompetenz als Lernziel . . . . .	9
1.4 Medienmix in der Methodenausbildung . . . . .	12
<b>2 Grundbegriffe der Statistik</b>	<b>15</b>
2.1 Statistische Einheit, Merkmal und Grundgesamtheit . . . .	15
2.2 Merkmalsklassifikationen und Skalen . . . . .	18
2.3 Operationalisierung von Merkmalen . . . . .	22
<b>3 Datengewinnung und Auswahlverfahren</b>	<b>25</b>
3.1 Erhebungsarten und Studiendesigns . . . . .	25
3.2 Stichprobenauswahl . . . . .	34
3.3 Träger amtlicher und nicht-amtlicher Statistik . . . . .	39
<b>4 Univariate Häufigkeitsverteilungen</b>	<b>41</b>
4.1 Absolute und relative Häufigkeiten . . . . .	41
4.2 Häufigkeitsverteilungen für klassierte Daten . . . . .	51
4.3 Die empirische Verteilungsfunktion . . . . .	57
<b>5 Kenngrößen empirischer Verteilungen</b>	<b>61</b>
5.1 Lagemaße . . . . .	61
5.2 Streuungsmaße . . . . .	68
5.3 Quantile und Boxplots . . . . .	73
<b>6 Konzentration von Merkmalswerten</b>	<b>79</b>
6.1 Die Lorenzkurve . . . . .	79
6.2 Konzentrationsmaße . . . . .	83
<b>7 Indikatoren</b>	<b>89</b>
7.1 Verhältniszahlen . . . . .	89
7.2 Zusammengesetzte Indexzahlen . . . . .	92

<b>8</b>	<b>Bivariate Häufigkeitsverteilungen</b>	<b>103</b>
8.1	Empirische Verteilungen diskreter Merkmale . . . . .	103
8.2	Empirische Unabhängigkeit diskreter Merkmale . . . . .	110
8.3	Empirische Verteilungen stetiger Merkmale . . . . .	118
<b>9</b>	<b>Zusammenhangsmaße</b>	<b>121</b>
9.1	Nominalskalierte Merkmale . . . . .	121
9.2	Metrische Merkmale . . . . .	126
9.3	Ordinalskalierte Merkmale . . . . .	134
<b>II</b>	<b>Wahrscheinlichkeitsrechnung und schließende Statistik</b>	<b>137</b>
	<b>Lernziele zu Teil II</b>	<b>138</b>
<b>10</b>	<b>Zufall und Wahrscheinlichkeit</b>	<b>139</b>
10.1	Grundbegriffe der Wahrscheinlichkeitsrechnung . . . . .	139
10.2	Zufallsstichproben und Kombinatorik . . . . .	146
10.3	Bedingte Wahrscheinlichkeiten . . . . .	150
10.4	Wahrscheinlichkeitsverteilungen . . . . .	156
<b>11</b>	<b>Diskrete Zufallsvariablen</b>	<b>161</b>
11.1	Wahrscheinlichkeits- und Verteilungsfunktion . . . . .	161
11.2	Kenngößen diskreter Verteilungen . . . . .	168
11.3	Die Binomialverteilung . . . . .	171
11.4	Die hypergeometrische Verteilung . . . . .	177
<b>12</b>	<b>Stetige Zufallsvariablen</b>	<b>183</b>
12.1	Dichtefunktion und Verteilungsfunktion . . . . .	183
12.2	Kenngößen stetiger Verteilungen . . . . .	187
12.3	Normalverteilung und Standardnormalverteilung . . . . .	189
12.4	$\chi^2$ -, $t$ - und $F$ -Verteilung . . . . .	197
<b>13</b>	<b>Bivariate Verteilungen</b>	<b>205</b>
13.1	Unabhängigkeit von Zufallsvariablen . . . . .	205
13.2	Kovarianz und Korrelation . . . . .	210
<b>14</b>	<b>Schätzung von Parametern</b>	<b>213</b>
14.1	Punktschätzungen und ihre Eigenschaften . . . . .	215
14.2	Punktschätzung von Erwartungswerten und Varianzen . . . . .	217
14.3	Punktschätzung von Anteilswerten . . . . .	220
14.4	Intervallschätzung für Erwartungswerte . . . . .	222
14.5	Intervallschätzung für Anteilswerte . . . . .	225

<b>15 Statistische Testverfahren</b>	<b>229</b>
15.1 Arten statistischer Tests . . . . .	230
15.2 Grundbegriffe und Gauß-Test für Erwartungswerte . . . . .	232
15.3 Gütefunktion des Gauß-Tests . . . . .	239
15.4 Signifikanzniveau und $p$ -Wert . . . . .	242
15.5 $t$ -Test für Erwartungswerte . . . . .	244
15.6 $\chi^2$ -Test für Varianzen . . . . .	246
15.7 Zweistichproben-Tests für Erwartungswerte . . . . .	247
15.8 Unabhängigkeitstests . . . . .	249
<b>16 Das lineare Regressionsmodell</b>	<b>251</b>
16.1 Das einfache lineare Regressionsmodell . . . . .	253
16.2 KQ-Schätzung im einfachen Regressionsmodell . . . . .	255
16.3 Das Bestimmtheitsmaß . . . . .	260
16.4 Das multiple lineare Regressionsmodell . . . . .	262
16.5 KQ-Schätzung im multiplen Regressionsmodell . . . . .	266
<b>17 Grundzüge der Varianzanalyse</b>	<b>271</b>
17.1 Das Modell der einfaktoriellen Varianzanalyse . . . . .	273
17.2 Durchführung einer einfaktoriellen Varianzanalyse . . . . .	275
17.3 Ausblick auf die zweifaktorielle Varianzanalyse . . . . .	281
<b>III Anhänge</b>	<b>283</b>
<b>Lernziele zu Teil III</b>	<b>284</b>
<b>18 Grundzüge der Matrizenrechnung</b>	<b>285</b>
18.1 Grundbegriffe . . . . .	285
18.2 Operationen mit Matrizen und Vektoren . . . . .	287
18.3 Charakterisierung von Zufallsvektoren . . . . .	294
<b>19 Tabellenanhang</b>	<b>297</b>
19.1 Verteilungsfunktion der Binomialverteilung . . . . .	297
19.2 Verteilungsfunktion der Standardnormalverteilung . . . . .	304
19.3 Quantile der Standardnormalverteilung . . . . .	308
19.4 Quantile der $\chi^2$ -Verteilung . . . . .	309
19.5 Quantile der $t$ -Verteilung . . . . .	310
19.6 Quantile der F-Verteilung . . . . .	312
<b>20 Übungsaufgaben</b>	<b>317</b>
20.1 Beschreibende Statistik . . . . .	317
20.2 Wahrscheinlichkeitsrechnung und schließende Statistik . . . . .	325

---

<b>21 Lösungen zu den Übungsaufgaben</b>	<b>333</b>
21.1 Beschreibende Statistik . . . . .	333
21.2 Wahrscheinlichkeitsrechnung und schließende Statistik . .	343
<b>22 Verzeichnisse und Internet-Ressourcen</b>	<b>359</b>
22.1 Literaturverzeichnis . . . . .	359
22.2 Kommentierte Liste ausgewählter Lehrbücher . . . . .	361
22.3 Online-Ressourcen . . . . .	362
22.4 Symbolverzeichnis . . . . .	364
22.5 Autorenregister . . . . .	367
22.6 Sachregister . . . . .	369

# Vorwort

Dieser Kurs ist weitgehend identisch mit einem Lehrbuch „Statistik – eine Einführung mit interaktiven Elementen“ (MITTAG 2017). 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 heutzutage unverzichtbar.

Ursprung dieses Lehrtextes

Charakteristisch für das Werk ist die Verknüpfung mit interessanten Web-Adressen und interaktiven oder dynamischen Lernobjekten. Letztere ermöglichen es, statistische Verfahren anhand von Experimenten „auszuprobieren“ oder statistische Konzepte über tongestützte Animationen nachzuvollziehen. Die interaktiven Experimente waren anfangs alle Java-basiert (Java-Applets) und nur für Desktops konzipiert. Um die Lauffähigkeit interaktiver statistischer Experimente auch auf Smartphones und Tablets zu erreichen, wurden die Lernobjekte nach und nach auf der Basis von HTML5, Javascript und CSS neu programmiert. Dabei galt es die Touchfunktionalitäten mobiler Endgeräte und das kleinere Display von Smartphones zu berücksichtigen.

Realisierung eines Medienmixes

Verknüpfung des Lehrbuchs mit mobilen Endgeräten über QR-Codes



*Plattformunabhängige Lernobjekte im Einsatz auf mobilen Endgeräten*

Die plattformunabhängigen und in einer Web-App zusammengefassten Lernobjekte wurden zunächst *englischsprachig* realisiert. Die App ist unter <http://www.mittag-statistik.de/app/> frei zugänglich. Ihr wurde das Comenius-EduMedia-Siegel 2014 der Gesellschaft für Pädagogik, Information und Medien verliehen. Die Elemente dieser App sind mit dem vorliegenden Kurs verknüpft.





Bedeutung der Icons  
am Seitenrand

Inzwischen gibt es unter der Adresse <https://www.hfh-fernstudium.de/statistik-app/> auch eine erweiterte *deutschsprachige* Fassung der Web-App mit kurzen Handhabungshinweisen für die einzelnen interaktiven Elemente. Die deutschsprachige Statistik-App wurde mit dem Comenius-EduMedia-Siegel 2015, dem Innovationspreis Bildung 2015 des Bundesverbands für Bildung, Wissenschaft und Forschung e. V. und mit dem Gütesiegel 2016 des Verbands ausgezeichnet.

Wo Videos mit Ton, statistische Experimente sowie besondere Web-Links zu finden sind, wird am Rand durch runde Icons sichtbar gemacht:



*Icons mit Verlinkung: Video mit Ton für Desktops, plattformunabhängig einsetzbares Video mit Ton, interaktives Java-Applet für Desktops, plattformunabhängig einsetzbares interaktives Lernobjekt, externer Web-Link*



Bei den plattformunabhängigen Videos und interaktiven Lernobjekten sind zusätzlich *QR-Codes* (engl.: *Quick Response*) platziert, damit man sie direkt von der Printfassung des Lehrbuchs mit einem mobilen Endgerät aufrufen kann. Anhand der runden Icons und QR-Codes sowie der zahlreichen in den Text eingestreuten Web-Links erkennt man, wo es Zusatzangebote gibt, die sich nur bei Nutzung des Computers erschließen lassen. Lediglich die über QR-Codes hervorgehobenen Objekte kann man sofort von der Printfassung aufrufen, indem man den Code mit einem Smartphone oder – besser – mit einem Tablet einscannt. Um den Mehrwert der anderen Zusatzangebote zu nutzen, wechselt man von der Print- zur e-Kursfassung.

Innerhalb dieses Manuskripts findet man noch quadratische Icons, die nur der Orientierung dienen und nicht mit Links verknüpft sind. Diese Icons verweisen auf die *Kapitelvorschau* am Anfang eines Kapitels sowie auf *Aufgaben*, *Lösungen* oder ergänzende *Literatur*:



*Weitere Icons: Vorschau, Aufgabe, Lösung, Literatur*

Von den Web-Links, die aus dem Lehrbuch hinausführen (externe Links), sind nur die interessantesten durch ein rundes Icon am Marginalienrand



betont. Die meisten externen Links sind beim e-Kurs anhand magentafarbener Schrift hervorgehoben, Links innerhalb des Manuskripts (Sprünge zu Aufgaben oder zu Nummern von Gleichungen, Abbildungen oder Tabellen) erscheinen in blauer Schrift.

Die jetzige Kursfassung unterscheidet sich von der vorausgegangenen Version durch die Einbindung neuer plattformunabhängiger Lernobjekte und durch zahlreiche Ergänzungen und Aktualisierungen. Neu sind z. B. einige Lernobjekte zum Themenkreis „Schätzen und Testen“ und zur Visualisierung amtlicher Daten. Etliche Beispiele und Exkurse wurden aktualisiert, z. B. zum demografischen Wandel. Es wurden auch Grafiken verbessert und neue Fotos aus der Praxis eingefügt.

Was ist neu?

Die e-Version dieses Kurses ist im Gegensatz zur Printfassung nicht nur mehrfarbig, sondern auch interaktiv. Die interaktiven oder dynamischen Elemente sind durch Anklicken der runden Icons am Seitenrand erreichbar. Auch die in den Text eingestreuten Web-Links zu Online-Ausgaben von Zeitschriften müssen hier nur angeklickt werden. Der e-Kurs ist ein Hypertext, bei dem man Querverweisen zu Gleichungen, Abbildungen oder Aufgaben per Mausklick nachgehen oder vom Stichwortregister zu den dort aufgeführten Schlüsselbegriffen springen kann.

Vorzüge des e-Kurs

<p><b>Printfassung</b> mehrfarbig; Link zu interaktiven Experimenten und Videos über QR-Codes</p>	<p><b>e-Kurs</b> mehrfarbiger Hypertext, interaktive Experimente und Videos integriert</p>
---	--



Unterschiede zwischen Kurs und e-Kurs

Der Kurs ist in drei Teile gegliedert. Der erste Teil widmet sich der beschreibenden Statistik, der zweite Teil der Wahrscheinlichkeitsrechnung und der schließenden Statistik. Der dritte Teil (Anhänge) umfasst Aufgaben, Lösungen, statistische Tabellen und diverse Verzeichnisse. Das Werk illustriert anhand zahlreicher Beispiele, dass die Statistik alle Bereiche gesellschaftlichen Lebens durchdringt. Die verwendeten Beispiele und Exkurse sind manchmal mit Hintergrundinformationen aus Online-Ausgaben überregionaler Zeitschriften verknüpft. Es gibt viele Bezüge zur amtlichen Statistik. Exkurse können übersprungen werden.

Struktur des Kurses

Betonung von  
Interdisziplinarität

Dank für die Programmierung der interaktiven Experimente gebührt Herrn A. *Michel*, Hagen. Unterstützung kam auch von Herrn B. *Wallenborn* (Hilfe bei IT-Problemen), Herrn Th. *Feuerstack* (Unterstützung bei Fragen zur L<sup>A</sup>T<sub>E</sub>X-Programmierung) sowie Herrn S. *Senicer* (Realisierung von Lehrvideos), alle Hagen. Frau A. *Dirks*, Hamburger Fern-Hochschule,

Danksagungen

konzipierte die Icons am Seitenrand. Herr Prof. Dr. R. *Münnich*, Trier, stellte Grafiken zur Armutsgefährdung zur Verfügung.

Dank gebührt ferner den nachstehenden Firmen und Institutionen, die kostenfrei Fotos oder andere Materialien zur Verfügung gestellt haben:

- Fa. Böhme und Weihs GmbH, Sprockhövel (Herr Dr. N. *Böhme*),
- DeStatis, Wiesbaden (Herr Dr. G. *Brückner*),
- Evonik Industries AG, Standort Essen (Herr Dr. W. *Wolfes*),
- GfK, Nürnberg (Herr S. *Heller* und Herr R. *Nicklas*),
- Hessischer Rundfunk, Frankfurt (Herr C. *Bender*),
- JMP / SAS, Köln (Herr Dr. V. *Kraft*),
- Kennesaw State University, USA (Dr. J. *McNeill*),
- Q-DAS GmbH, Weinheim (Herr Dr. E. *Dietrich*),
- Fa. STATCON, Witzenhausen (Herr B. *Schäfer*),
- TNS Infratest, München (Herr M. *Kögel*),
- Volkswagen AG, Wolfsburg (Herr T. *Cramm*).

Wetter / Ruhr, im Oktober 2017

*Hans-Joachim Mittag*

[joachim.mittag@fernuni-hagen.de](mailto:joachim.mittag@fernuni-hagen.de)

# Teil I

## Beschreibende Statistik



# Lernziele zu Teil I

Nach der Bearbeitung des ersten Teils dieses Manuskripts sollten Sie

- wissen, warum statistische Methodenkompetenz heutzutage als Schlüsselqualifikation gilt;
- zentrale Aufgaben und Anwendungsfelder der Statistik kennen;
- mit wichtigen Grundbegriffen der Statistik vertraut sein (z. B. Merkmale und Merkmalstypen);
- alternative Ansätze zur Gewinnung von Daten und zur Entnahme von Stichproben kennen;
- Datensätze für ein Merkmal unter Verwendung geeigneter Grafiken visualisieren können;
- in der Lage sein, Lage- und Streuungsparameter empirischer Verteilungen zu berechnen;
- wissen, wie sich Merkmalskonzentration messen und visualisieren lässt;
- den Einsatzzweck von Indikatoren sowie einige Beispiele für Indikatoren benennen können;
- in der Lage sein, Datensätze für zwei Merkmale anhand von Kontingenztafeln oder, bei stetigen Merkmalen, anhand von Streudiagrammen darzustellen;
- Maße kennen, mit denen sich ein Zusammenhang zwischen zwei Merkmalen quantifizieren lässt.