

Vorwort zur dritten Auflage

Das hier vorliegende Buch erscheint nun in der dritten Auflage. Damit wird uns zum einen die Gelegenheit geboten, es an die Möglichkeiten der neuesten Excel-Version anzupassen (Excel 2002), zum anderen können wir auf die zahlreichen Anregungen und Verbesserungsvorschläge eingehen, die uns freundliche und aufmerksame Leser, insbesondere auch Fachkollegen und -kolleginnen aus der Zunft der Statistiker, zugeschickt haben und für die wir uns an dieser Stelle herzlich bedanken. In diesen Zuschriften kommt mehrheitlich zum Ausdruck, dass unser Versuch, ein Statistikbuch zu schreiben, das nicht von mathematischen Formeln strotzt, begrüßt und als gelungen angesehen wird. Es versteht sich, dass wir uns über diesen Zuspruch freuen und ihn als Ansporn für weitere Bemühungen betrachten. Dass der ein oder andere gleichwohl das Buch für zu wenig anspruchsvoll ansieht, können wir verschmerzen.

Im Vorwort zur ersten Auflage, die 1996 erschien, hatten wir schon darauf aufmerksam gemacht, dass die hier vorliegende Einführung in die statistische Arbeit verknüpft werden soll mit einer Demonstration der Einsatzmöglichkeiten eines Statistikprogramms. Wir haben uns dabei für das Tabellenkalkulationsprogramm Excel entschieden, das ja eine außerordentlich weite Verbreitung erfahren hat und einfach zu bedienen ist. Im Vordergrund stand dabei nicht der Versuch, alle Details dieses Programms zu erläutern, als vielmehr, zu illustrieren, wie dieses Programm bei der Lösung von statistischen Standardaufgaben, wie sie in der Praxis immer wieder auftauchen, nutzbringend eingesetzt werden kann. Deshalb findet der Leser in diesem Buch also Folgendes:

- eine kurze einleitende Geschichte in die jeweilige Thematik
- eine knappe Darstellung statistischer Grundüberlegungen zu jedem Themenbereich
- Formeln nur dort, wo sie hingehören
- praxisnahe Beispiele
- Musterlösungen in Excel zu jeder angesprochenen statistischen Fragestellung.

Wir haben uns bemüht, das Ganze so flüssig zu schreiben, dass man auch mal ein längeres Stück lesen kann, ohne gleich den Gesamtüberblick zu verlieren.

Zum Schluss noch einige Worte zu den statistischen sowie EDV-technischen Inhalten:

Wir stützen uns bei der Software, wie schon erwähnt, auf das weit verbreitete und außerordentlich leistungsfähige Tabellenkalkulationsprogramm Excel, mit dem die wichtigsten Standardverfahren der Statistik ohne Probleme bewältigt werden können. Dabei lassen sich alle Beispiele fast ohne Einschränkungen auch noch mit den älteren Excel-Versionen behandeln.

An dieser Stelle sei aber darauf aufmerksam gemacht, dass es auch andere Softwareprogramme gibt, die speziell für statistische Auswertungsaufgaben eingesetzt werden können, so zum Beispiel das für Statistiker entwickelte Programm SPSS, das schon seit längerer Zeit auch als Windows-Version vorliegt. Gerade das Standardprogramm Excel scheint uns aber besonders geeignet, vor allem den Statistik-Pragmatiker in die Geheimnisse der Methoden einzuweihen, um ihn in die Lage zu versetzen, eigene Datenbestände mit den adäquaten Verfahren auszuwerten und zu analysieren.

Unter inhaltlichen Gesichtspunkten greifen wir – soweit die deskriptive Statistik betroffen ist – auf die Einführung von M. Tiede zurück (M. Tiede: Beschreiben mit Statistik – Verstehen,

Oldenbourg Verlag, München/Wien 2001). Bezüglich der Verfahren der induktiven Statistik orientieren wir uns an folgendem Lehrbuch: M. Tiede/W. Voß: Schließen mit Statistik – Verstehen, Oldenbourg Verlag, München/Wien 2000.

Diesem Buch ist eine CD beigelegt. Auf dieser sind die Datenbestände in Form von Excel-Tabellen abgelegt, die den einzelnen Kapiteln zugrunde liegen. Die einzelnen Dateien dieses Datenträgers können unter Excel geöffnet werden.

Sollten Sie die Daten verändern wollen, empfiehlt es sich, sie vorher auf die Festplatte zu kopieren.

Darüber hinaus finden Sie auf der CD-ROM eine Formelsammlung sowie Übungsaufgaben, wie sie an der Bochumer Universität bearbeitet werden. Um das Buch abzurunden, haben wir den gesamten Buchtext als PDF-Datei gespeichert.

Zum Schluss möchten wir denjenigen danken, die dieses Buch tatkräftig unterstützt haben. Hier ist vor allem Frau Dr. Wuschel-Monka zu nennen. Nicht nur weil einer der Autoren das Glück hat, mit ihr liiert zu sein, sondern weil sie mit zahlreichen Anregungen und kritischen Einwänden die statistischen Inhalte des Buches zu einer Einheit gebracht hat. Unseren Dank an die Kritiker der ersten beiden Auflagen haben wir schon weiter oben abgestattet.

Herrn Prof. Dr. Manfred Tiede danken wir für die Bereitstellung der Formelsammlung auf der CD, Frau Ekaterina Serova und Herrn Dr. Stefan Karduck für die redaktionelle Überarbeitung des Textes.

Dr. Michael Monka

Prof. Dr. Werner Voß

Bonn, Hattingen im Januar 2002

Inhaltsverzeichnis

Vorwort.....	5
1 Was überhaupt ist Statistik?	13
1.1 Die Statistik lügt	13
1.2 Zielsetzungen	14
1.3 Zum Begriff der Statistik	15
1.4 Anwendungsbereiche	17
1.5 Methodengruppen	19
1.6 Grundbegriffe	21
2 Excel – Grundlagen	23
2.1 Im Schweiß meines Angesichts	23
2.2 Statistik und Excel	24
2.3 Der Start von Excel	25
2.4 Ein erstes Beispiel	27
2.5 Veränderungen	32
2.6 Mathematische Berechnungen	36
2.7 Drucken	45
2.8 Diagramme	46
2.9 Daten importieren	50
3 Sortieren, Gruppieren, Klassifizieren.....	53
3.1 Der Lottospieler	53
3.2 Wie man statistische Daten gewinnt	54
3.3 Von der Urliste zur Häufigkeitsverteilung	60
3.4 Zeitreihen	70
3.5 Kreuztabellen	70
4 Mittelwerte	75
4.1 Wo bitte ist die Mitte?	75
4.2 Wozu braucht man Mittelwerte?	75
4.3 Das arithmetische Mittel	77
4.4 Der häufigste Wert (Modus)	84
4.5 Der Zentralwert (Median)	87
4.6 Das geometrische Mittel	88

5	Streuungsmaße.....	91
5.1	Am Frühstückstisch.....	91
5.2	Spannweite.....	91
5.3	Mittlere lineare Abweichung.....	93
5.4	Standardabweichung.....	95
5.5	Quartile und Semiquartilsabstand	98
6	Konzentrationsmaße.....	103
6.1	Isabella.....	103
6.2	Was bedeutet Konzentration?	103
6.3	Herfindahl-Index.....	104
6.4	Maß von Lorenz/Münzner	106
6.5	Die Lorenzkurve	107
6.6	Der Lorenzkoeffizient.....	112
7	Grafische Darstellungen.....	117
7.1	Chinesische Tusche.....	117
7.2	Die Vorzüge grafischer Darstellungen	117
7.3	Stabdiagramm	119
7.4	Histogramm	124
7.5	Tortendiagramm.....	133
7.6	Die grafische Darstellung von Zeitreihen	134
7.7	Streudiagramm.....	138
8	Indexberechnungen	145
8.1	Alles wird immer teurer	145
8.2	Gliederungszahlen, Messziffern, Wachstumsraten	145
8.3	Umbasierung und Verkettung	150
8.4	Preisindex	153
8.5	Mengenindex	159
8.6	Wertindex	162
9	Regressionsrechnung (bivariat).....	163
9.1	Auf dem Tennisplatz.....	163
9.2	Zielsetzungen	163
9.3	Mathematische Grundlagen	169
9.4	Die Methode der kleinsten Quadrate	171
9.5	Beispiel zur Regressionsrechnung	174
9.6	Fallstricke	180

10	Zusammenhangsrechnung.....	183
10.1	Das Klassenbuch.....	183
10.2	Korrelations- und Determinationskoeffizient.....	183
10.3	Rangkorrelation	194
10.4	Zusammenhangsmaße für Nominaldaten.....	196
10.5	Der Alleskönner.....	201
11	Trendfunktionen und Trendprognosen	203
11.1	Statistik lügt?	203
11.2	Zielsetzungen.....	203
11.3	Glättung einer Zeitreihe.....	206
11.4	Linearer Zeitreihentrend.....	209
11.5	Exponentieller Trend.....	216
11.6	Saisonale Schwankungen.....	217
12	Grundlagen der Wahrscheinlichkeitstheorie.....	225
12.1	Pfälzer Leberwurst.....	225
12.2	Ereignisse	226
12.3	Kombinatorik.....	228
12.4	Kombinatorische Praxisaufgaben	236
12.5	Bestimmung der Wahrscheinlichkeit	238
13	Wahrscheinlichkeitsverteilungen.....	245
13.1	Mensch ärgere Dich nicht!.....	245
13.2	Grafische Darstellung von Wahrscheinlichkeiten.....	246
13.3	Der Additionssatz	247
13.4	Der Multiplikationssatz	250
13.5	Zwei weitere Sätze.....	253
13.6	Zufallsvariable	256
13.7	Diskrete Wahrscheinlichkeitsverteilungen.....	258
13.8	Stetige Wahrscheinlichkeitsverteilungen	260
13.9	Erwartungswert und Varianz von Zufallsvariablen.....	265
14	Spezielle Wahrscheinlichkeitsverteilungen.....	269
14.1	Das neue Fahrrad	269
14.2	Vorbemerkung	270
14.3	Binomialverteilung	270
14.4	Normalverteilung.....	281
14.5	Multinomialverteilung	289
14.6	Hypergeometrische Verteilung	290
14.7	Poisson-Verteilung	293
14.8	Student-t-Verteilung	295

14.9	Chi-Quadrat-Verteilung	295
14.10	F-Verteilung.....	296
15	Stichproben und Stichprobenverteilungen.....	297
15.1	Zufall?.....	297
15.2	Auswahlverfahren für Stichproben	297
15.3	Bewusste Auswahlen	299
15.4	Zufällige Auswahlen	300
15.5	Stichprobenverteilungen	306
15.6	Das zentrale Grenzwerttheorem.....	310
15.7	Stichprobenverteilungen wichtiger Maßzahlen.....	311
15.8	Notationen	321
16	Der parametrische Hypothesentest	323
16.1	An der Theke	323
16.2	Beispiel: Der Zigarettentest	324
16.3	Mittelwerttest.....	326
16.4	Entscheidungsfehler.....	331
16.5	Weitere Parametertests	339
16.7	Die Güte eines Tests	347
17	Hochrechnungen	351
17.1	Modifizierte Bundestagswahl	351
17.2	Aufgaben der Schätzstatistik.....	351
17.3	Punktschätzverfahren.....	352
17.4	Intervallschätzungen	357
18	Tests bei kleinen Stichprobenumfängen.....	363
18.1	Die sparsamste Lösung	363
18.2	Ausgangslage	363
18.3	Anteilswerttest	364
18.4	Anteilswertdifferenzentest (Fisher-Test).....	366
18.5	Mittelwerttest.....	369
18.6	Mittelwertdifferenzentest.....	370
18.7	Varianzquotiententest	372
18.8	Mittelwerte aus mehr als zwei Stichproben	374

19	Anpassungstestverfahren.....	381
19.1	Total normal?.....	381
19.2	Aufgaben von Anpassungstests.....	381
19.3	Chi-Quadrat-Anpassungstest	384
19.4	Fisher-Test und Chi-Quadrat-Homogenitätstest	395
19.5	Kolmogoroff/Smirnow-Test.....	397
20	Testen statistischer Zusammenhänge	401
20.1	Die Streitfrage.....	401
20.2	Chi-Quadrat-Unabhängigkeitstest (Nominaldaten).....	401
20.3	Unabhängigkeitstest bei Ordinaldaten	407
20.4	Test des Korrelationskoeffizienten r (metrische Daten)	409
20.5	Test des Regressionskoeffizienten	412
20.6	Schätzen der Regressionsgeraden	414
21	Multiple Regression und partielle Korrelation.....	417
21.1	Vom Sinn der Aufklärung.....	417
21.2	Drittvariablen.....	418
21.3	Partielle Korrelationsrechnung	420
21.4	Multiple lineare Regression	423
21.5	Multipler Determinationskoeffizient.....	426
21.5	Nicht-metrische Daten	427
22	Weitere Verfahren.....	429
22.1	Der Kollege	429
22.2	Was fehlt?	430
22.3	Zusätzliche Testverfahren.....	430
22.4	Multivariate Verfahren	434
	Excel-Funktionen.....	441
	Glossar.....	447
	Register.....	465