

Inhalt

Vorwort zur 2. Auflage — V

Vorwort — VII

1	Einführung in die Statistik	1
-1.1	Eingrenzungen des Begriffs „Statistik“	1
1.1.1	Komprimierende Kennwerte	1
1.1.2	Staatswissenschaftliche Disziplin	2
1.1.3	Wahrscheinlichkeitstheoretisch fundierte Datenwissenschaft	3
1.2	Grundzüge statistischer Methoden	5
1.2.1	System und Zufall	5
1.2.2	Irrtums- und Sicherheitswahrscheinlichkeiten	7
1.2.3	Deskription und Induktion	10
1.2.4	Empirischer und theoretischer Kalkül	11
1.3	Teilbereiche, Spezialgebiete und Grundlagen	13
1.3.1	Datengewinnung und Aufbereitung	14
1.3.2	Spezielle Analysemethoden und Instrumente	16
1.3.3	Wissenschaftsspezifische Ausrichtungen	19
1.3.4	Mathematische und technische Grundlagen	20
Teil 1: Deskriptive Statistik – Empirischer Kalkül		21
2	Einführung in die deskriptive Statistik	23
2.1	Daten, Datensätze und Variablen	23
2.2	Grundgesamtheit, Merkmalstypen und Skalierungsarten	26
2.3	Empirische Verteilungen und Zusammenhänge	28
3	Gesamtbeschreibung empirischer Verteilungen	31
3.1	Tabellarische Darstellungsmöglichkeiten	31
3.1.1	Elementare Begriffe und Notation	31
3.1.2	Häufigkeitstabellen bei Urlisten	33
3.1.3	Häufigkeitstabellen bei klassierten Daten	35
3.2	Grafische Darstellungsmöglichkeiten	37
3.2.1	Kreis-, Säulen- und Balkendiagramme	37
3.2.2	Stamm-Blatt-Diagramme	38
3.2.3	Histogramme (Häufigkeitsdichten)	40
3.2.4	Boxplots	44

3.3	Empirische Verteilungsfunktion (EVF)	48
3.3.1	Häufigkeitsfunktion und EVF bei Urlisten	48
3.3.2	Häufigkeitsdichtefunktion und EVF bei klassierten Daten	50
4	Spezifizierende Beschreibung empirischer Verteilungen	55
4.1	Spezifika empirischer Verteilungen	55
4.2	Lagekennwerte	58
4.2.1	Arithmetisches Mittel	59
4.2.2	Median	62
4.2.3	Modalwert	65
4.2.4	Fechner'sche Lageregeln	66
4.3	Spezielle Lagekennwerte	68
4.3.1	Arithmetisches Mittel bei gruppierten Daten	68
4.3.2	Quantile	69
4.3.3	Geometrisches Mittel	73
4.4	Streuungskennwerte	77
4.4.1	Spannweite	77
4.4.2	Mittlere absolute Abweichungen	77
4.4.3	Median absoluter Abweichungen	80
4.4.4	Varianz, Standardabweichung und Schwankungsintervalle	81
4.5	Spezielle Streuungskennwerte	85
4.5.1	Varianz bei gruppierten Daten	85
4.5.2	Quantilsabstände	88
4.5.3	Variationskoeffizient	89
4.6	Standardisierung mittels Lage und Streuung	92
4.7	Messung von Schiefe	95
4.8	Darstellung und Messung von Konzentration	97
4.8.1	Lorenz-Kurve	98
4.8.2	Gini-Koeffizient	100
4.9	Spezifische Eigenschaften empirischer Kennwerte	104
4.9.1	Minimumeigenschaft des arithmetischen Mittels	104
4.9.2	Minimumeigenschaft des Medians	106
4.9.3	Transformationseigenschaften	109
4.9.4	Robustheit	114
5	Beschreibung und Analyse empirischer Zusammenhänge	116
5.1	Zusammenhänge zwischen kategorialen Merkmalen	116
5.1.1	Kontingenztabellen	116
	Verteilungen in Kontingenztabellen	116
	Empirische Abhängigkeit und Unabhängigkeit	122
	Konzeptionelle Vorüberlegungen für ein Zusammenhangsmaß	126
5.1.2	Zusammenhangsmaße für Kontingenztabellen	128

	Chi-Quadrat-Koeffizient	128
	Mittlere quadratische Kontingenz	131
	Kontingenzkoeffizient nach Pearson	132
	Transformationseigenschaften der Zusammenhangsmaße . . .	134
5.1.3	Grafische Analysemöglichkeiten	137
	Gestapelte und gruppierte Säulendiagramme	137
	Segmentierte Säulen- und Balkendiagramme	138
	Assoziationsplots	138
	Mosaikplots	139
	Spineplots	141
5.2	Zusammenhänge zwischen metrischen Merkmalen	142
5.2.1	Grafische Analysemöglichkeiten	142
	Streudiagramme	142
	Streudiagramm-Matrizen	144
	Hexagonalplots	144
5.2.2	Zusammenhangsmaße für metrische Merkmale	145
	Empirische Kovarianz	145
	Empirischer Korrelationskoeffizient nach Pearson	151
	Empirische Kovarianz- und Korrelationsmatrizen	159
	Korrelationskoeffizient nach Spearman	161
5.2.3	Einfache lineare Regression	165
	Zweck und allgemeine Vorgehensweise	165
	KQ-Methode (L_2 -Regression)	168
	LAD-Methode (L_1 -Regression)	175
	Kritische Punkte und Alternativen	180
5.3	Ergänzende und vertiefende Themen	183
5.3.1	Zusammenhänge in anderen Skalierungsfällen	184
	Grafische Analysemöglichkeiten	184
	Zusammenhangsmaße	186
	Spezielle Regressionsmodelle	188
5.3.2	Grafische Darstellung höherdimensionaler Daten	191
	Höherdimensionale Mosaikplots	191
	3D-Punktwolken	194
	Biplots	194
5.3.3	Wichtige Aspekte bei der Analyse empirischer Zusammenhänge	201
	Empirische Zusammenhänge und Kausalität	201
	Spezielle Formen von Kausalität	205
	Systematische Verzerrungen	207
	Zufällige Schwankungen	219

Teil 2: Wahrscheinlichkeitsrechnung – Theoretischer Kalkül	225
6 Einführung in die Wahrscheinlichkeitsrechnung	227
6.1 Wichtige Grundbegriffe und Regeln	227
6.1.1 Interpretation von Zufall und Wahrscheinlichkeiten	227
6.1.2 Elementare Mengenlehre	232
6.1.3 Axiomatischer Wahrscheinlichkeitsbegriff und Regeln	240
6.2 Rechnen mit abhängigen und unabhängigen Ereignissen	251
6.2.1 Bedingte Wahrscheinlichkeiten und Rechenregeln	251
6.2.2 Stochastische Unabhängigkeit von Ereignissen	260
6.2.3 Kalkül nach der Formel von Bayes	267
7 Theoretische Verteilungen und Abhängigkeiten	272
7.1 Zufallsvariablen und deren Wahrscheinlichkeitsverteilungen	272
7.1.1 Eindimensionale Zufallsvariablen	272
Konzept und Ausblick	272
Diskrete Zufallsvariablen und Additionskalkül	275
Stetige Zufallsvariablen und Integrationskalkül	277
Theoretische Verteilungsfunktion	283
Wahrscheinlichkeitsverteilung	288
7.1.2 Mehrdimensionale Zufallsvariablen	290
Konzept und Ausblick	290
Zweidimensionale diskrete Zufallsvariablen	291
Zweidimensionale stetige Zufallsvariablen	297
Stochastische Abhängigkeit und Unabhängigkeit	305
Höherdimensionale Wahrscheinlichkeitsverteilungen	309
7.1.3 Verteilung von Funktionen von Zufallsvariablen	319
Funktionen einer Zufallsvariable	319
Funktionen aus mehreren Zufallsvariablen	326
7.2 Theoretische Kennwerte	330
7.2.1 Kennwerte in Bezug auf Lage und Streuung	331
Erwartungswert	331
Theoretische Varianz und Standardabweichung	347
Theoretische Quantile und theoretischer Median	353
7.2.2 Kennwerte in Bezug auf Abhängigkeiten	355
Bedingte Erwartungswerte und Varianzen	355
Theoretische Kovarianz und Korrelation	365
Theoretische Regressionskoeffizienten	371
7.2.3 Spezifische Eigenschaften theoretischer Kennwerte	373
Minimumeigenschaften von Lagekennwerten	373
Wichtige Transformationseigenschaften	374
Endliche und nicht endliche theoretische Momente	377

7.3	Spezielle eindimensionale Verteilungen	379
7.3.1	Spezielle diskrete Verteilungen	380
	Elementare Kombinatorik	380
	Einpunktverteilung	383
	Bernoulli-Verteilung	383
	Binomialverteilung	385
	Poisson-Verteilung	390
7.3.2	Spezielle stetige Verteilungen	394
	Stetige Gleichverteilung	394
	Exponentialverteilung	397
	Normalverteilung	402
7.4	Verteilung stochastischer Summen und Mittelwerte	411
7.4.1	Exakte Aussagen	411
	Erwartungswerte und Varianzen	411
	Verteilungen unter bestimmten Ausgangsverteilungen	416
7.4.2	Asymptotische und approximative Aussagen	422
	Gesetz der großen Zahlen (GGZ)	422
	Zentraler Grenzwertsatz (ZGWS)	427
8	Ergänzungen und Verallgemeinerungen	438
8.1	Weitere eindimensionale Verteilungen	438
8.1.1	Weitere diskrete Verteilungen	438
8.1.2	Weitere stetige Verteilungen	440
8.1.3	Sonstige eindimensionale Verteilungen	445
8.2	Ergänzungen zu asymptotischen Aussagen	448
8.2.1	Alternative GGZ- und ZGWS-Varianten	448
8.2.2	Stochastische Konvergenz und Verteilungskonvergenz	452
8.3	Einige multivariate Verallgemeinerungen	457
8.3.1	Multivariate Abhängigkeiten	457
	Abhängigkeit und Unabhängigkeit von Zufallsvektoren	457
	Bedingte Abhängigkeit und Unabhängigkeit	461
8.3.2	Kenngrößen multivariater Abhängigkeiten	462
	Theoretische Kovarianz- und Korrelationsmatrix	462
	Bedingte Kennwerte für $n > 2$	463
8.3.3	Sonstige Verallgemeinerungen	467
Teil 3: Induktive Statistik – Verbindung von Empirie und Theorie		475
9	Einführung in die induktive Statistik	477
9.1	Modellierung und Handhabung von Schätz- und Testproblemen	477
9.2	Was versteht man unter einer Stichprobe?	481

9.3	Formale Unterscheidungsebenen	493
10	Statistisches Schätzen	497
10.1	Punktschätzung	497
10.1.1	Schätzer und ihre Gütekriterien	497
10.1.2	Erwartungstreue	504
10.1.3	Erwartete quadratische Abweichung (MSE)	509
10.1.4	Konsistenz	511
10.1.5	Effizienz	519
10.2	Intervallschätzung	522
10.2.1	Was versteht man unter einem Konfidenzintervall?	522
10.2.2	Konfidenzintervalle für Erwartungswerte	528
10.2.3	Konfidenzintervalle für Erwartungswertdifferenzen	535
10.2.4	Weitere Konfidenzintervalle	545
10.2.5	Adäquatheit bestimmter Modellannahmen	550
10.3	Schätzmethoden	556
10.3.1	Momentenmethode	557
10.3.2	Maximum-Likelihood-Methode	560
10.3.3	Weitere Schätzmethoden	566
11	Statistisches Testen	568
11.1	Was versteht man unter einem Test?	568
11.1.1	Einführende Beispiele	568
11.1.2	Grundstruktur und Durchführung	581
11.1.3	Zusammenhang zur Intervallschätzung	584
11.2	Wichtige Aspekte beim Testen	587
11.2.1	Hypothesenwahl und Fehlerarten	587
11.2.2	Irrtumswahrscheinlichkeiten und Güte	591
11.2.3	p -Werte	604
11.2.4	Signifikanz vs. Relevanz	607
11.3	Ausgewählte Testverfahren	609
11.3.1	Tests über Erwartungswerte	609
11.3.2	Tests über Erwartungswertdifferenzen	613
11.3.3	Nichtparametrische χ^2 -Tests	618
	Allgemeiner Überblick	618
	χ^2 -Anpassungstest	618
	χ^2 -Unabhängigkeitstest	627
11.3.4	Weitere Tests	633
	Tests über Anteilswerte	633
	Tests auf Unkorreliertheit und Unabhängigkeit	640
	Sonstige Tests und Testmethoden	644
11.3.5	Allgemein zu beachtende Punkte	648

12 Das lineare Regressionsmodell	649
12.1 Einfaches lineares Regressionsmodell	649
12.1.1 Grundmodell und KQ-Methode	650
Statistisches Grundmodell	650
Herleitung der KQ-Schätzer	651
Eigenschaften der empirischen KQ-Regression	654
Bestimmtheitsmaß und Standardfehler der Regression	659
12.1.2 Modellannahmen und theoretische KQ-Regression	662
Ensembles von Modellannahmen	662
Theoretische KQ-Regression	670
12.1.3 Verteilungstheoretische Grundlagen	676
Verteilungen der KQ-Schätzer	676
Konsistenz und Effizienz der KQ-Schätzer	686
Schätzung der Varianzen der KQ-Schätzer	687
Verteilungen der Inferenzstatistiken	690
12.1.4 Schätzen und Testen	691
Konfidenzintervalle und Tests	691
Spezialfall: Binärer Regressor	694
Adäquatheit bestimmter Modellannahmen	697
Fallbeispiel 1: Bewässerung und Wachstum	704
Fallbeispiel 2: Klassengröße und Lernerfolg	709
12.2 Einführung in das multiple lineare Regressionsmodell	716
12.2.1 Partielle lineare KQ-Regression	716
Empirische partielle Regression	716
Theoretische partielle Regression	723
Verbindung von Empirie und Theorie	725
12.2.2 Multiple lineare KQ-Regression	727
Empirische multiple Regression	727
Theoretische multiple Regression	731
Verbindung von Empirie und Theorie	733
Statistische Modelle und Inferenz	733
12.2.3 Fallbeispiele	742
Fallbeispiel 2 fortgesetzt: Determinanten des Lernerfolgs	742
Fallbeispiel 3: Gewicht und Geschlecht	745
Fallbeispiel 4: Binäre Regressoren und ANOVA-Modelle	746

Anhang: Tabellen — 752

Stichwortverzeichnis — 755

Literatur — 767