

Werner A. Stahel

# **Statistische Datenanalyse**

**Eine Einführung  
für Naturwissenschaftler**

5., überarbeitete Auflage



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Was ist Statistische Datenanalyse? . . . . .	1
1.2	Ziele . . . . .	5
1.3	Hinweise . . . . .	7
1.4	Literatur zur angewandten Statistik . . . . .	9
<b>I</b>	<b>Beschreibende Statistik</b>	<b>11</b>
<b>2</b>	<b>Beschreibung eindimensionaler Stichproben</b>	<b>13</b>
2.1	Histogramme . . . . .	13
2.2	Einige Bezeichnungen und Begriffe . . . . .	15
2.3	Kennzahlen für eine quantitative Stichprobe . . . . .	17
2.4	Klassierte Daten . . . . .	21
2.5	Mehrere Stichproben . . . . .	25
2.6	Transformationen von beobachteten Werten . . . . .	28
2.7	Wertebereiche, Datensorten . . . . .	32
2.8	* Transformationen und Unterschiede zwischen Beobachtungen . . . . .	33
<b>3</b>	<b>Beschreibende Statistik mehrdimensionaler Daten</b>	<b>35</b>
3.1	Grafische Darstellungen für zwei zusammenhängende Größen . . . . .	35
3.2	Die Produktmomenten-Korrelation . . . . .	38
3.3	Rangkorrelation . . . . .	41
3.4	Zur Interpretation von Korrelationen . . . . .	43
3.5	Regression . . . . .	44
3.6	Multivariate Beobachtungen . . . . .	48
3.7	Zeitreihen und räumliche Daten . . . . .	55
3.8	Allgemeines zu grafischen Darstellungen . . . . .	59
3.9	Wie weiter? . . . . .	60
	Literatur zu Teil I . . . . .	60
<b>II</b>	<b>Wahrscheinlichkeitsrechnung</b>	<b>61</b>
<b>4</b>	<b>Wahrscheinlichkeit</b>	<b>63</b>
4.1	Einleitung . . . . .	63
4.2	Grundbegriffe und Grundeigenschaften . . . . .	64
4.3	Zufallsvariable . . . . .	69
4.4	Zufallszahlen . . . . .	74
4.5	Zwei Zufallsvariable, gemeinsame Verteilung . . . . .	77
4.6	Unabhängige Ereignisse und Zufallsvariable . . . . .	79
4.7	Bedingte Wahrscheinlichkeit . . . . .	82
4.8	Bedingte Verteilung . . . . .	88
4.9	Der Satz von Bayes . . . . .	90
4.10	* Was ist eine Wahrscheinlichkeit? . . . . .	92
4.11	Wie weiter? . . . . .	95

<b>5</b>	<b>Diskrete Verteilungen</b>	<b>96</b>
5.1	Binomial-Verteilung . . . . .	96
5.2	Die Poisson-Verteilung . . . . .	101
5.3	Kennzahlen . . . . .	104
5.4	Verteilungsfamilien . . . . .	108
5.5	Die multinomiale Verteilung . . . . .	111
5.6	Summen von Zufallsvariablen . . . . .	112
5.7	Zufalls-Stichproben . . . . .	117
5.8	* Gesetze der grossen Zahl . . . . .	120
5.9	* Stochastische Prozesse . . . . .	124
<b>6</b>	<b>Stetige Verteilungen</b>	<b>126</b>
6.1	Grundlagen . . . . .	126
6.2	Grundbegriffe, Exponential- und uniforme Verteilung . . . . .	127
6.3	Kennzahlen für stetige Verteilungen . . . . .	132
6.4	Transformationen von Zufallsvariablen . . . . .	134
6.5	Die Normalverteilung . . . . .	139
6.6	Die Lognormal-Verteilung . . . . .	143
6.7	* Weitere stetige Verteilungsfamilien . . . . .	146
6.8	Gemeinsame und bedingte Verteilung . . . . .	148
6.9	Unabhängige Zufallsvariable und Korrelation . . . . .	152
6.10	Funktionen von mehreren Zufallsvariablen . . . . .	154
6.11	Gauss'sche Fehler-Fortpflanzung . . . . .	160
6.12	Der Zentrale Grenzwertsatz . . . . .	162
6.13	Rückblick . . . . .	167
	Literatur zu Teil II . . . . .	168
<b>III</b>	<b>Schliessende Statistik</b>	<b>169</b>
<b>7</b>	<b>Schätzungen</b>	<b>171</b>
7.1	Drei Grundfragen der schliessenden Statistik . . . . .	171
7.2	Schätzungen für $\mathcal{B}$ , $\mathcal{P}$ und $\mathcal{N}$ . . . . .	173
7.3	Eigenschaften von Schätzungen . . . . .	175
7.4	Die Maximum-Likelihood-Methode . . . . .	180
7.5	Robuste Schätzungen . . . . .	184
<b>8</b>	<b>Tests</b>	<b>187</b>
8.1	Einführende Beispiele und Begriffe . . . . .	187
8.2	Test für eine Wahrscheinlichkeit . . . . .	192
8.3	Die Teststatistik . . . . .	195
8.4	Vorgehen bei einem statistischen Test . . . . .	197
8.5	Tests für eine Stichprobe oder zwei gepaarte Stichproben . . . . .	200
8.6	Interpretation von Testergebnissen . . . . .	206
8.7	Bemerkungen zum P-Wert . . . . .	208
8.8	Vergleich von zwei quantitativen Stichproben . . . . .	211
8.9	Macht . . . . .	217
8.10	* Asymptotische Tests und Randomisierungs-Tests . . . . .	220
8.11	Sinn und Unsinn statistischer Tests . . . . .	221

<b>9</b>	<b>Vertrauensintervalle</b>	<b>223</b>
9.1	Vertrauensintervalle für Binomial- und Poisson-Verteilung . . . . .	223
9.2	Die Grundeigenschaft von Vertrauensintervallen . . . . .	226
9.3	Vertrauensintervalle für Lageparameter . . . . .	228
9.4	Bootstrap und andere Resampling-Methoden . . . . .	231
9.5	Vertrauens- und andere Intervalle . . . . .	234
9.6	Schätzungen, Tests und Vertrauensintervalle im Vergleich . . . . .	235
9.7	Wo stehen wir? . . . . .	237
	Literatur zu Teil III . . . . .	237
<b>IV</b>	<b>Methoden der Datenanalyse</b>	<b>239</b>
<b>10</b>	<b>Nominale Daten</b>	<b>241</b>
10.1	Multinomiale Verteilung und Chiquadrat-Test . . . . .	241
10.2	Der Chiquadrat-Anpassungstest . . . . .	247
10.3	Der Chiquadrat-Test in Kontingenztafeln . . . . .	251
10.4	Die häufigsten Fehler beim Chiquadrat-Test . . . . .	255
<b>11</b>	<b>Überprüfung von Voraussetzungen</b>	<b>257</b>
11.1	Problemstellung . . . . .	257
11.2	Quantil-Quantil-Diagramme . . . . .	258
11.3	Anpassungstests . . . . .	261
11.4	Bedeutung von Tests zur Prüfung von Voraussetzungen . . . . .	262
11.5	Unabhängigkeit . . . . .	263
<b>12</b>	<b>Varianzanalyse</b>	<b>265</b>
12.1	Vergleich mehrerer Stichproben, einfache Varianzanalyse . . . . .	265
12.2	Multiple Vergleiche, multiple Tests . . . . .	269
12.3	Mehrere verbundene Stichproben . . . . .	271
12.4	Zweiweg-Varianzanalyse . . . . .	273
12.5	Zufällige Effekte, Varianz-Komponenten . . . . .	278
12.6	Ausblick . . . . .	281
	Literatur . . . . .	282
<b>13</b>	<b>Regression</b>	<b>283</b>
13.1	Das Modell der einfachen linearen Regression . . . . .	283
13.2	Schätzung der Parameter . . . . .	285
13.3	Tests und Vertrauensintervalle für die Koeffizienten . . . . .	288
13.4	Vertrauens- und Vorhersage-Bereiche . . . . .	290
13.5	Multiple lineare Regression . . . . .	293
13.6	Vielfalt der Modelle der multiplen linearen Regression . . . . .	296
13.7	Interpretation von Regressionskoeffizienten . . . . .	300
13.8	Residuen-Analyse . . . . .	303
13.9	Einflussreiche Beobachtungen . . . . .	309
13.10	Modellwahl . . . . .	311
13.11	Allgemeinere Modelle für stetige Zielgrößen . . . . .	317
13.12	Verallgemeinerte lineare Modelle . . . . .	321
	Literatur . . . . .	324

<b>14</b>	<b>Versuchsplanung</b>	<b>325</b>
14.1	Einleitung . . . . .	325
14.2	Allgemeine Überlegungen . . . . .	325
14.3	Versuchspläne . . . . .	330
14.4	Eine Checkliste . . . . .	332
	Literatur . . . . .	335
<b>15</b>	<b>Multivariate Statistik</b>	<b>336</b>
15.1	Mehrdimensionale Zufallsvariable . . . . .	336
15.2	Schätzung von Erwartungswert und Kovarianz-Matrix . . . . .	341
15.3	Die mehrdimensionale Normalverteilung . . . . .	343
15.4	Statistik der Normalverteilung . . . . .	346
15.5	Hauptkomponenten . . . . .	348
15.6	Diskriminanz-Analyse . . . . .	351
	Literatur . . . . .	355
<b>16</b>	<b>Zeitreihen</b>	<b>356</b>
16.1	Fragestellungen . . . . .	356
16.2	Auto-Korrelation . . . . .	357
16.3	ARMA-Modelle . . . . .	358
16.4	Statistik von Zeitreihen . . . . .	359
16.5	Vorhersage . . . . .	360
16.6	Zustandsraum-Modelle . . . . .	361
16.7	Spektralanalyse . . . . .	362
16.8	Räumliche Korrelation . . . . .	364
16.9	Regression mit Zeitreihen . . . . .	366
	Literatur . . . . .	368
<b>17</b>	<b>Stichproben-Erhebungen</b>	<b>369</b>
17.1	Einleitung . . . . .	369
17.2	Einfache Zufalls-Stichprobe . . . . .	370
17.3	Geschichtete Stichproben . . . . .	373
17.4	Weitere Stichproben-Pläne . . . . .	374
17.5	Weitere Schätzmethoden . . . . .	376
17.6	Auswertung von Umfragen . . . . .	377
17.7	Eine Checkliste . . . . .	379
	Literatur . . . . .	382
<b>18</b>	<b>Ausblick</b>	<b>383</b>
18.1	Bedeutung von Wahrscheinlichkeit-Modellen . . . . .	383
18.2	Grosse Datensätze, beschreibende Modelle . . . . .	384
18.3	Die Statistik und ihre Anwendungen . . . . .	386
	Nachwort . . . . .	387
<b>A</b>	<b>Anhang: Kurzfassung des wichtigsten Stoffes</b>	<b>389</b>
	<b>Literaturverzeichnis</b>	<b>399</b>
	<b>Sachwortverzeichnis</b>	<b>405</b>