

Götz Rohwer, Ulrich Pötter

Grundzüge der sozialwissenschaftlichen Statistik

Juventa Verlag Weinheim und München 2001

Inhalt

Teil I Variablen und Verteilungen

1	Gesamtheiten und Variablen	15
1.1	Einleitende Bemerkungen	15
1.2	Zur Konzeption von Gesamtheiten	17
1.3	Zur Verwendung symbolischer Schreibweisen	20
1.3.1	Notationen aus der Mengenlehre	21
1.3.2	Erläuterungen zum Funktionsbegriff	24
1.4	Statistische Variablen	26
1.4.1	Zur Konzeption von Merkmalsräumen	26
1.4.2	Definition des Variablenbegriffs	28
1.4.3	Mehrdimensionale Variablen	30
2	Variablen und Daten	33
2.1	Definition statistischer Daten	33
2.2	Einige einfache Beispiele	35
2.2.1	Merkmale von Seminarteilnehmern	35
2.2.2	Haushalte und ihr Einkommen	36
2.2.3	Erwerbsbeteiligung verheirateter Frauen	37
2.3	Reflexion von Fragestellungen	40
3	Statistische Verteilungen	43
3.1	Definition des Verteilungsbegriffs	43
3.2	Statistische Aussagen über Verteilungen	45
3.3	Eigenschaften von Häufigkeitsfunktionen	46
4	Zum Umgang mit Merkmalsräumen	48
4.1	Konzeption und Realisation	48
4.2	Numerische Repräsentation	49
4.3	Bildung neuer Merkmalsräume	51
4.3.1	Vergrößerung von Merkmalsräumen	51
4.3.2	Transformation von Merkmalsräumen	54
4.3.3	Zusammenfassung von Merkmalsräumen	55
5	Eindimensionale Verteilungen	57
5.1	Darstellung von Verteilungen	57
5.1.1	Darstellung durch Tabellen	57
5.1.2	Verteilungsfunktionen	59
5.1.3	Merkmalsmengen und Intervalle	61
5.1.4	Histogramme	63
5.1.5	Dichtefunktionen	65
5.1.6	Konstruktion von Dichtefunktionen	67

5.2	Charakterisierung von Verteilungen	72
5.2.1	Quantile	72
5.2.2	Mittelwerte	74
5.2.3	Streuungsmaße	75
5.2.4	Die Tschebyscheffsche Ungleichung	77
5.3	Bedingte Verteilungen	78
5.3.1	Definition bedingter Verteilungen	78
5.3.2	Bedingte Verteilungscharakterisierungen	81
6	Mehrdimensionale Verteilungen	83
6.1	Einleitende Bemerkungen	83
6.2	Darstellung mehrdimensionaler Verteilungen	86
6.2.1	Streuungsdiagramme	86
6.2.2	Mehrdimensionale Verteilungsfunktionen	88
6.3	Charakterisierung mehrdimensionaler Verteilungen	92
6.3.1	Mittelwerte und Varianzen	92
6.3.2	Kovarianz und Korrelation	94
6.4	Bedingte Verteilungen	97
6.4.1	Definition bedingter Verteilungen	98
6.4.2	Charakterisierungen bedingter Verteilungen	100
6.4.3	Bedingte Verteilungen als Variablen	101
6.4.4	Streuungszerlegungen	103
6.5	Mehrere eindimensionale Verteilungen	105

Teil II Regressionsrechnung

7	Grundgedanken der Regressionsrechnung	111
7.1	Regressionsfunktionen	111
7.2	Darstellung durch Tabellen	113
7.3	Weiterführende Überlegungen	115
7.4	Regressionsmodelle	119
8	Lokale Konstruktionsverfahren	124
8.1	Regression mit Mittelwerten	124
8.1.1	Erläuterung des Grundgedankens	125
8.1.2	Verwendung von Kernfunktionen	126
8.1.3	Darstellung von Streuungen	127
8.2	Regression mit Quantilen	128
8.3	Regression mit Häufigkeiten	129
9	Parametrische Konstruktionsverfahren	132
9.1	Einleitende Bemerkungen	132
9.2	Regression mit Mittelwerten	133
9.2.1	Konzeption von Modellfunktionen	133
9.2.2	Berechnung von Modellfunktionen	134

9.2.3	Lineare OLS-Regression	137
9.2.4	Beispiele	139
9.2.5	Eigenschaften des OLS-Kriteriums	141
9.2.6	Mehrere Regressorvariablen	142
9.3	Regression mit Häufigkeiten	145
9.3.1	Konzeption von Modellfunktionen	145
9.3.2	Berechnung von Modellfunktionen	147
9.3.3	Ein Beispiel	150
9.3.4	Mehrere Regressorvariablen	152
9.4	Aspekte der Modellanpassung	157
9.4.1	Allgemeine Überlegungen	158
9.4.2	Variabilität der Modellparameter	160
10	Zum Verständnis von Regressionsmodellen	166
10.1	Einleitende Bemerkungen	166
10.2	Sachverhalte und Prozesse	170
10.3	Verwendung für Voraussagen?	174
10.4	Aspekte der Modellbeurteilung	178

Teil III Darstellung von Zustandsverläufen

11	Eine zeitliche Perspektive	183
11.1	Einleitende Bemerkungen	183
11.2	Zustandsraum und Biographieschema	184
11.3	Kalenderzeit und Prozeßzeit	187
11.4	Ansätze statistischer Begriffsbildung	189
11.4.1	Zustandsvariablen	189
11.4.2	Zustandsverteilungen	192
11.4.3	Darstellung von Übergängen	193
12	Begriffsbildungen für Episoden	196
12.1	Episoden und Verweildauern	196
12.2	Verweildauerverteilungen	197
12.3	Unvollständige Daten	200
12.4	Rechts zensierte Beobachtungen	202
12.5	Bemerkungen zum Kaplan-Meier-Verfahren	205
12.6	Links abgeschnittene Beobachtungen	207
12.6.1	Konditionale Survivorfunktionen	207
12.6.2	Illustration mit Mortalitätsdaten	209
13	Modelle für Zustandsverläufe	213
13.1	Retrospektive und temporale Ansätze	213
13.2	Modelle für Zustandswechsel	215
13.2.1	Bezugnahme auf Übergangshäufigkeiten	215
13.2.2	Ein Modellansatz für Situationen	218

13.2.3	Prozesse als Folgen von Situationen	222
13.3	Modelle für Verweildauervariablen	224
13.3.1	Einleitende Bemerkungen	224
13.3.2	Einfache Regressionsmodelle	229
13.3.3	Parametrisierung von Übergangsraten	233
13.3.4	Zeitabhängige Kovariablen	236
13.3.5	Zeitveränderliche Übergangsraten	242

Teil IV Stichproben zur Datengewinnung

14	Gesamtheiten und Stichproben	253
14.1	Reflexion der Problemstellung	253
14.2	Repräsentative Stichproben	255
14.3	Eine Problemverschiebung	258
15	Zufällige Auswahlen	260
15.1	Einleitende Bemerkungen	260
15.2	Begriffe aus der Wahrscheinlichkeitsrechnung	264
15.2.1	Zufallsgeneratoren	264
15.2.2	Aleatorische Wahrscheinlichkeit	267
15.3	Zufallsstichproben	271
15.3.1	Ein allgemeines Schema	271
15.3.2	Einfache Auswahlverfahren	274
16	Überlegungen der Stichprobentheorie	277
16.1	Einleitende Bemerkungen	277
16.2	Konzeption von Schätzfunktionen	279
16.2.1	Der gedankliche Ansatz	279
16.2.2	Inklusionsvariablen	281
16.2.3	Definition von Schätzfunktionen	284
16.3	Eigenschaften von Schätzfunktionen	285
16.3.1	Erwartungstreue Schätzfunktionen	286
16.3.2	Der mittlere quadratische Fehler	287
16.4	Beispiele für Schätzfunktionen	288
16.4.1	Schätzfunktionen für Mittelwerte	289
16.4.2	Schätzfunktionen für Häufigkeiten	291
16.4.3	Schätzfunktionen für Varianzen	292
16.5	Schätzfunktionen und Schätzwerte	295
16.5.1	Schätzfunktionen als Zufallsvariablen	295
16.5.2	Hypothetische Stichprobenverteilungen	298
16.5.3	Konfidenzintervalle	300
17	Unterschiedliche Auswahlverfahren	303
17.1	Listenbasierte Auswahlverfahren	303
17.1.1	Systematische Zufallsauswahl	303

17.1.2	Geschichtete Stichproben	306
17.1.3	Mehrstufige Auswahlverfahren	313
17.1.4	Clusterstichproben	318
17.2	Auswahlverfahren bei Umfragen	321
17.2.1	Allgemeine Bemerkungen	321
17.2.2	Flächenstichproben	323
17.2.3	Haushalte und Personen	328
17.2.4	Quotenstichproben	330
17.3	Eine kurze Zusammenfassung	332

Teil V Ungenaue und unvollständige Daten

18	Ein allgemeinerer Variablenbegriff	339
18.1	Einleitende Bemerkungen	339
18.2	Empirische Variablen	341
18.2.1	Mengenwertige Variablen	341
18.2.2	Intervallwertige Variablen	344
19	Statistik mit empirischen Variablen	346
19.1	Statistische Verteilungen	346
19.1.1	Mittlere Verteilungen	346
19.1.2	Intervallwertige Verteilungen	348
19.1.3	Selbstkonsistente Verteilungen	350
19.2	Darstellung von Verteilungen	353
19.2.1	Verteilungsfunktionen	353
19.2.2	Dichtefunktionen	356
19.3	Charakterisierung von Verteilungen	360
19.3.1	Mittelwerte	361
19.3.2	Streuungsmaße	362
19.4	Regressionsfunktionen	366
19.4.1	Lineare OLS-Regression	366
19.4.2	Lokale Mittelwertregression	371
	Literatur	375
	Symbolverzeichnis	378
	Stichwortverzeichnis	381