

Prof. Dr. Achim Bühl

PASW18

Einführung in die moderne Datenanalyse

12., aktualisierte Auflage

PEARSON

Studium

ein Imprint von Pearson Education
München • Boston • San Francisco • Harlow, England
Don Mills, Ontario • Sydney • Mexico City
Madrid • Amsterdam



Inhaltsverzeichnis

Vorwort zur zwölften Auflage.	15
1 Installation	17
1.1 Systemvoraussetzungen von PASW Statistics 18.	17
1.2 Installation von PASW Statistics 18.	17
1.2.1 Deinstallation älterer Versionen.	18
1.2.2 Installation von PASW 18 unter Windows Vista oder XP.	18
1.3 Verknüpfung erstellen	24
1.3.1 Verknüpfung unter Windows Vista	24
1.3.2 Verknüpfung unter Windows XP.	25
1.4 Arbeitsverzeichnis einrichten.	26
1.5 Übungsdateien des Buchs downloaden.	28
2 Datenaufbereitung	31
2.1 Kodierung und Kodeplan	32
2.2 Datenmatrix.	34
2.3 Start von PASW.	35
2.4 Daten-Editor	36
2.4.1 Definition der Variablen.	36
2.4.2 Dateneingabe.	50
2.5 Speichern einer Datendatei	53
2.6 Variablendeklarationen kopieren.	53
2.7 Zusammenfügen von Datendateien.	57
2.7.1 Fallweises Zusammenfügen.	57
2.7.2 Variablenweises Zusammenfügen.	62
2.8 Einlesen bereits vorhandener Daten.	68
2.8.1 Einlesen von Daten mit Hilfe des PASW-Assistenten.	69
2.8.2 Einlesen von Daten mit Hilfe der Syntax.	71
2.9 Arbeitssitzung beenden.	73
3 PASW Statistics im Überblick	75
3.1 Auswahl einer Statistik-Prozedur.	76
3.1.1 Variablen auswählen.	79
3.1.2 Unterdialogboxen.	80
3.2 Einstellungen für den Daten-Editor.	82
3.3 Die Symbolleiste.	84
3.4 Erstellen und Editieren von Grafiken.	86
3.5 Der Viewer.	92

Inhaltsverzeichnis

3.6	Editieren von Tabellen	97
3.6.1	Der Pivot-Tabellen-Editor	98
3.6.2	Weitere Bearbeitungsmöglichkeiten	101
3.6.3	Symbole des PASW-Viewers	107
3.7	Der Syntax-Editor	108
3.8	Informationen zur Datendatei	114
3.9	Das Hilfesystem	119
3.10	Einstellungen	122
3.11	Arbeiten mit mehreren Datendateien	123
3.11.1	Kopieren und Einfügen zwischen Datendateien	125
3.11.2	Umbenennen von Daten-Sets	128
3.12	Export der Ausgabe	128
3.12.1	Statistische Ergebnisse nach Word übertragen	128
3.12.2	Statistische Ergebnisse nach Word exportieren	132
3.12.3	Diagramme nach Word übertragen	135
3.12.4	Pivot-Tabellen und Diagramme in HTML-Dokumenten verwenden	137
4	Statistische Grundbegriffe und Kennwerte	141
4.1	Voraussetzungen für die Anwendung eines statistischen Tests	141
4.1.1	Skalenniveaus	141
4.1.2	Normalverteilung	144
4.1.3	Abhängigkeit und Unabhängigkeit von Stichproben	145
4.2	Übersicht über gängige Mittelwerttests	145
4.3	Die Irrtumswahrscheinlichkeit p	146
4.4	Statistischer Wegweiser	148
4.4.1	Strukturierung, Eingabe und Überprüfung der Daten	148
4.4.2	Deskriptive Auswertung	149
4.4.3	Analytische Statistik	149
4.5	Statistische Kennwerte	150
4.5.1	Deskriptive Statistiken	152
4.5.2	Fälle zusammenfassen	155
4.5.3	Verhältnis-Statistiken	157
5	Häufigkeitsauszählungen	161
5.1	Häufigkeitstabellen	161
5.2	Ausgabe statistischer Kennwerte	162
5.3	Mediän bei gehäuften Daten	167
5.4	Formate für Häufigkeitstabellen	171
5.5	Grafische Darstellung	172
6	Datenselektion	179
6.1	Auswahl von Fällen	179
6.1.1	Einteilung der Operatoren	181
6.1.2	Relationale Operatoren	181
6.1.3	Logische Operatoren	182
6.1.4	Boolesche Algebra	182
6.1.5	Funktionen	185

Inhaltsverzeichnis

6.1.6	Eingabe eines Konditional-Ausdrucks.	187
6.1.7	Beispiele für Datenselektionen.	190
6.2	Ziehen einer Zufallsstichprobe.	192
6.3	Fälle sortieren.	194
6.4	Aufteilung der Fälle in Gruppen.	195
Datenmodifikation.		201
7.1	Berechnung von neuen Variablen.	201
7.1.1	Formulierung numerischer Ausdrücke.	203
7.1.2	Funktionen.	205
7.1.3	Einbindung der Syntax in den dialoggesteuerten Ablauf.	210
7.2	Erstellen von Variablen mit Hilfe des Bereichseinteilers.	212
7.3	Zählen des Auftretens bestimmter Werte.	218
7.4	Umkodieren von Werten.	221
7.4.1	Manuelles Umkodieren.	221
7.4.2	Automatisches Umkodieren.	226
7.5	Bedingte Berechnung von neuen Variablen.	229
7.5.1	Formulierung von Bedingungen.	229
7.5.2	Bildung eines Indexes.	231
7.6	Aggregieren von Daten.	236
7.7	Rangtransformationen.	240
7.7.1	Beispiel einer Rangtransformation.	240
7.7.2	Rangtypen.	242
7.8	Gewichten von Fällen.	245
7.8.1	Korrektur bei nicht gegebener Repräsentativität.	245
7.8.2	Analyse von gehäuften Daten.	251
7.9	Beispiele für die Berechnung neuer Variablen.	255
7.9.1	Erstes Beispiel: Berechnung des Benzinverbrauchs.	255
7.9.2	Zweites Beispiel: Berechnung des Datums des Ostersonntags.	256
Datenexploration.		261
8.1	Aufdeckung von Eingabefehlern.	261
8.2	Überprüfung der Verteilungsform.	262
8.3	Berechnung von Kennwerten.	262
8.4	Explorative Datenanalyse.	262
8.4.1	Analysen ohne Gruppierungsvariablen.	263
8.4.2	Analysen für Gruppen von Fällen.	270
Kreuztabellen.		275
9.1	Erstellen von Kreuztabellen.	275
9.2	Grafische Veranschaulichung von Kreuztabellen.	289
9.3	Statistiken für Kreuztabellen.	291
9.3.1	Chi-Quadrat-Test.	292
9.3.2	Korrelationsmaße.	295
9.3.3	Assoziationsmaße für nominalskalierte Variablen.	298
9.3.4	Assoziationsmaße für ordinalskalierte Variablen.	303
9.3.5	Weitere Assoziationsmaße.	304

Inhaltsverzeichnis

10	Analyse von Mehrfachantworten	311
10.1	Dichotome Methode	311
10.1.1	Definition von Sets	312
10.1.2	Häufigkeitstabellen für dichotome Setvariablen	313
10.1.3	Kreuztabellen mit dichotomen Setvariablen	315
10.2	Erstellen von Ranking-Listen	319
10.3	Kategoriale Methode	323
10.3.1	Definition von Sets	325
10.3.2	Häufigkeitstabellen für kategoriale Setvariablen	325
10.3.3	Kreuztabellen mit kategorialen Setvariablen	326
10.4	Dichotome und kategoriale Methode im Vergleich	328
11	Mittelwertvergleiche	331
11.1	Vergleich von zwei unabhängigen Stichproben	332
11.2	Vergleich von zwei abhängigen Stichproben	334
11.3	Vergleich von mehr als zwei unabhängigen Stichproben	336
11.3.1	Zerlegen in Trendkomponenten	339
11.3.2	A-priori-Kontraste	339
11.3.3	A-posteriori-Tests	340
11.3.4	Weitere Optionen	341
11.4	Vergleich von mehr als zwei abhängigen Stichproben	341
11.5	Einstichproben-t-Test	343
11.6	Einbindung der Syntax in den dialoggesteuerten Ablauf	344
12	Nichtparametrische Tests	347
12.1	Vergleich von zwei unabhängigen Stichproben	348
12.1.1	U-Test nach Mann und Whitney	348
12.1.2	Moses-Test	351
12.1.3	Kolmogorov-Smirnov-Test	352
12.1.4	Wald-Wolfowitz-Test	353
12.2	Vergleich von zwei abhängigen Stichproben	354
12.2.1	Wilcoxon-Test	354
12.2.2	Vorzeichen-Test	358
12.2.3	Chi-Quadrat-Test nach McNemar	360
12.3	Vergleich von mehr als zwei unabhängigen Stichproben	361
12.3.1	H-Test nach Kruskal und Wallis	361
12.3.2	Median-Test	363
12.4	Vergleich von mehr als zwei abhängigen Stichproben	364
12.4.1	Friedman-Test	364
12.4.2	Kendalls W	366
12.4.3	Cochrans Q	367
12.5	Kolmogorov-Smirnov-Test zur Überprüfung der Verteilungsform	368
12.6	Chi-Quadrat-Einzeltest	369
12.7	Binomial-Test	373
12.8	Sequenzanalyse	375
12.9	Nichtparametrische Tests mit Hilfe des Model Viewer	376
12.9.1	U-Test nach Mann und Whitney	376
12.9.2	H-Test nach Kruskal und Wallis	382

13	Korrelationen.	385
13.1	Korrelationskoeffizient nach Pearson.	387
13.2	Rangkorrelationskoeffizienten nach Spearman und Kendall.	388
13.3	Partielle Korrelationen.	389
13.4	Distanz- und Ähnlichkeitsmaße.	392
13.5	Der Intraclass Correlation Coefficient (ICC).	395
13.6	Einbindung der Syntax in den dialoggesteuerten Ablauf.	396
14	Regressionsanalyse.	399
14.1	Einfache lineare Regression.	400
14.1.1	Berechnen der Regressionsgleichung.	401
14.1.2	Neue Variablen speichern.	403
14.1.3	Zeichnen einer Regressionsgeraden.	405
14.2	Multiple lineare Regression.	408
14.3	Nichtlineare Regression.	412
14.4	Binäre logistische Regression.	418
14.5	Multinomiale logistische Regression.	427
14.6	Ordinale Regression.	438
14.7	Probitanalyse.	446
14.8	Kurvenanpassung.	453
14.9	Gewichtsschätzung.	457
14.10	Partielle kleinste Quadrate.	462
14.11	Zweistufige kleinste Quadrate.	462
14.12	Kategoriale Regression.	462
14.12.1	Prinzip der Kategorienquantifikationen.	463
14.12.2	Zweites Beispiel: Der Untergang der Titanic.	470
14.12.3	Diskretisierung von Variablen.	477
15	Varianzanalysen.	483
15.1	Univariate Varianzanalyse.	485
15.1.1	Univariate Varianzanalyse (allgemeines lineares Modell).	486
15.1.2	Univariate Varianzanalyse nach Fisher.	493
15.1.3	Univariate Varianzanalyse mit Messwiederholung.	495
15.2	Kovarianzanalyse.	500
15.3	Multivariate Varianzanalyse.	502
15.4	Varianz-Komponenten.	504
15.5	Lineare gemischte Modelle.	508
15.5.1	Varianzanalyse mit festen Effekten.	508
15.5.2	Kovarianzanalyse mit festen Effekten.	512
15.5.3	Analyse mit festen und zufälligen Effekten.	514
15.5.4	Analyse mit wiederholten Messungen.	516
16	Diskriminanzanalyse.	521
16.1	Beispiel aus der Medizin.	521
16.2	Beispiel aus der Soziologie.	530

16.3	Beispiel aus der Biologie	538
16.4	Diskriminanzanalyse mit drei Gruppen	540
17	Reliabilitätsanalyse.	545
17.1	Richtig-falsch-Aufgaben	546
17.2	Stufen-Antwort-Aufgaben	553
18	Faktorenanalyse.	555
18.1	Rechenschritte und Verfahrenstypen der Faktorenanalyse.	555
18.2	Explorative Faktorenanalyse.	556
18.2.1	Beispiel aus der Soziologie.	556
18.2.2	Beispiel aus der Psychologie.	564
18.3	Konfirmatorische Faktorenanalyse.	574
18.3.1	Beispiel aus der Freizeitforschung.	574
18.3.2	Grafische Darstellung des Rechnens mit Faktorwerten . . .	582
18.3.3	Beispiel aus der Medienwissenschaft	586
18.4	Das Rotationsproblem	590
19	Clusteranalyse.	593
19.1	Das Prinzip der Clusteranalyse.	594
19.2	Hierarchische Clusteranalyse.	598
19.2.1	Hierarchische Clusteranalyse mit zwei Variablen.	598
19.2.2	Hierarchische Clusteranalyse mit mehr als zwei Variablen.	603
19.2.3	Hierarchische Clusteranalyse mit vorgeschalteter Faktorenanalyse.	606
19.3	Ähnlichkeits- und Distanzmaße.	610
19.3.1	Intervallskalierte (metrische) Variablen.	610
19.3.2	Häufigkeiten.	613
19.3.3	Binäre Variablen.	614
19.4	Fusionierungsmethoden	615
19.5	Clusteranalyse für hohe Fallzahlen (Clusterzentrenanalyse).	616
19.6	Die Two-Step-Clusteranalyse.	622
19.6.1	Die Two-Step-Clusteranalyse per Dialogboxen der Version 17.	624
19.6.2	Die Two-Step-Clusteranalyse per Syntax.	628
19.6.3	Die Two-Step-Clusteranalyse per Model Viewer.	645
20	Klassifikationsanalyse.	659
20.1	Einführendes Beispiel aus der Geschichtswissenschaft	661
20.1.1	Erstellen einer Analysedatei	661
20.1.2	Erzeugung und Interpretation eines Baumdiagramms	663
20.1.3	Interpretation der Vorhersagewerte.	669
20.1.4	Arbeiten mit dem Baumeditor.	672
20.2	Vertiefungsbeispiel aus dem Bereich der Wirtschaftswissenschaften	676
20.2.1	Erstellen einer Analysedatei	677

20.2.2	Erzeugung und Interpretation eines Baumdiagramms	678
20.2.3	Erleichterung beim Erstellen der finalen Tabelle	684
20.3	Der CHAID-Algorithmus als Klassifikationsmethode	688
20.3.1	Erstellen einer Analysedatei	689
20.3.2	Erzeugung und Interpretation eines Baumdiagramms	690
20.3.3	Ansichten eines Baumdiagramms und Navigation durch den Baum	694
20.3.4	Analyse der finalen Segmente	699
20.4	Der Exhaustive-CHAID-Algorithmus als Klassifikationsmethode	703
20.4.1	Erstellen einer Analysedatei	704
20.4.2	Erzeugung eines Baumdiagramms	705
20.4.3	Betrachtung des Baummodells und der Baumtabelle	707
20.4.4	Gewinnzusammenfassung, Risiko und Klassifikation	710
20.4.5	Vorhergesagte Werte	712
20.4.6	Analyse der finalen Segmente	713
20.5	Der CRT-Algorithmus als Klassifikationsmethode	715
20.5.1	Der binäre Algorithmus in vergleichender Betrachtung	716
20.5.2	Vertiefungsbeispiel zum CRT-Algorithmus	722
20.6	Der QUEST-Algorithmus als Klassifikationsmethode	730
20.6.1	Erstellen einer Analysedatei	731
20.6.2	Erzeugung und Interpretation eines Baumdiagramms	734
20.6.3	Analyse der Vorhersagewerte	737
20.6.4	Analyse der Endknoten	738
20.6.5	Dichotomisierung der Zielvariablen	740
20.6.6	Analyse einzelner Parteien	744
20.7	Die Hilfeoption des Baumeditors	749
21	Loglineare Modelle	751
21.1	Eine typische Anwendungssituation	751
21.2	Das Prinzip der loglinearen Modelle	754
21.3	Überblick über die loglinearen Modelle	756
21.4	Hierarchisches loglineares Modell	757
21.5	Allgemeines loglineares Modell	766
21.6	Logit-loglineares Modell	772
22	Überlebens- und Ereignisdatenanalyse	779
22.1	Sterbetafeln	780
22.1.1	Einführende Beispiele aus der Medizin	780
22.1.2	Vertiefende Beispiele aus der Soziologie	789
22.2	Kaplan-Meier-Methode	796
22.3	Regressionsanalyse nach Cox	801
22.3.1	Beispiel aus der Medizin	801
22.3.2	Beispiel aus der Ökonomie	807
22.4	Cox-Regression mit zeitabhängigen Kovariaten	808
23	Multidimensionale Skalierung	813
23.1	Das Prinzip der MDS	814
23.2	Beispiel aus dem Marketing-Bereich	819

23.3	Ähnlichkeiten aus Daten erstellen	821
23.4	Multidimensionale Skalierung und Faktorenanalyse	829
24	Korrespondenzanalyse	833
24.1	Einfache Korrespondenzanalyse	834
24.1.1	Das Prinzip der einfachen Korrespondenzanalyse	835
24.1.2	Beispiel einer Produktpositionierung	845
24.1.3	Das Seriationsproblem in der Archäologie	851
24.2	Multiple Korrespondenzanalyse mit Nominalvariablen	854
24.2.1	Erstes Beispiel: Produktpositionierung im Marketingbereich	854
24.2.2	Zweites Beispiel: Visualisierung der Variablenzusammenhänge	860
24.2.3	Drittes Beispiel: Darstellung der Kategorienquantifikationen	866
24.3	Multiple Korrespondenzanalyse mit beliebigen Variablen	871
24.3.1	Erstes Beispiel: Alle Variablen numerisch	871
24.3.2	Zweites Beispiel: Numerische und nominalskalierte Variablen	877
24.4	Kanonische Korrespondenzanalyse	883
25	Conjoint-Analyse	899
25.1	Zielsetzung	899
25.2	Vorstellung eines Beispiels	901
25.3	Erstellung eines orthogonalen Designs	901
25.4	Die Ausgabe des orthogonalen Designs	906
25.5	Die Prozedur CONJOINT	909
26	Berichte und Gruppenwechsel	917
26.1	Zeilenweise Berichte	917
26.1.1	Erstellen eines einfachen Berichts	917
26.1.2	Zweistufiger Gruppenwechsel	922
26.1.3	Dreistufiger Gruppenwechsel	923
26.1.4	Berichts-Layout	927
26.2	Spaltenweise Berichte	932
26.3	Komprimierte Berichtsausgabe	938
26.4	Übungsaufgaben	941
27	Diagramme	943
27.1	Balkendiagramm	943
27.1.1	Einfaches Balkendiagramm: Darstellung von Häufigkeiten	943
27.1.2	Einfaches Balkendiagramm: Kennwerte einer metrischen Variablen	947
27.1.3	Gruppiertes Balkendiagramm	953
27.1.4	Gestapeltes Balkendiagramm	955

27.2	Liniendiagramm	958
27.2.1	Einfaches Liniendiagramm	959
27.2.2	Mehrfaches Liniendiagramm	961
27.3	Flächendiagramm	963
27.3.1	Einfaches Flächendiagramm	963
27.3.2	Gestapeltes Flächendiagramm	965
27.4	Kreisdiagramm	966
27.5	Streudiagramm	968
27.6	Histogramm	973
27.6.1	Einfaches Histogramm	973
27.6.2	Gestapeltes Histogramm	975
27.7	Hoch-Tief-Diagramme	976
27.7.1	Einfache Hoch-Tief-Schluss-Diagramme	976
27.7.2	Einfache Bereichsbalken	979
27.7.3	Gruppierte Hoch-Tief-Schluss-Diagramme	980
27.7.4	Differenzliniendiagramme	980
27.8	Boxplot	981
27.8.1	Einfacher Boxplot	981
27.8.2	Gruppiertes Boxplot	983
27.9	Doppelachsen	984
27.9.1	Zwei Y-Achsen mit kategorialer X-Achse	984
27.9.2	Zwei Y-Achsen mit metrischer X-Achse	987
27.10	Erstellen eines Diagramms aus einer Pivot-Tabelle	988
28	Weiterführende Literatur	991
29	Index	995