

**Reinhold Hatzinger
Herbert Nagel**

PASW Statistics

Statistische Methoden und Fallbeispiele

Inhaltsverzeichnis

Fragestellungen und Methoden	11
Vorwort	15
Kapitel 1 Einführung	17
1.1 Konzeption des Buchs	18
1.2 Aufbau des Buchs	19
1.3 Programmversionen von PASW bzw. SPSS	20
1.4 Wie kann dieses Buch verwendet werden?	20
1.5 Typografische und andere Konventionen in diesem Buch	21
Kapitel 2 Statistische Grundbegriffe	23
2.1 Einige Beispiele	24
2.1.1 Hochrechnung (statistisches Schätzen)	24
2.1.2 Prüfen von Fragestellungen (Testen von Hypothesen)	26
2.1.3 Erstellen von Modellen (statistisches Modellieren)	28
2.2 Grundlegende Konzepte	29
2.3 Messung und Typen von Daten	31
2.4 Arten von Fragestellungen und Variablen	35
2.5 Zusammenfassung der Konzepte	38
2.6 Übungen	39
Kapitel 3 Überblick über die Bedienung von PASW	41
3.1 Einleitung	42
3.2 Fragebogen und Kodierung	43
3.3 Definieren von Variablen und Zuordnen von beschreibenden Namen	46
3.3.1 Übertragen von Definitionen	50
3.3.2 Speichern der Variablendefinitionen	52
3.4 Erfassen bzw. Eingabe der kodierten Daten, Datenkontrolle	53
3.4.1 Abspeichern der Daten	55
3.4.2 Datenkontrolle	56
3.5 Transformieren der Daten bzw. Erzeugen von neuen Variablen	57
3.5.1 Berechnen neuer Variablen	57
3.5.2 Umkodieren von Variablen	62
3.6 Auswählen von Fällen	65
3.7 Gewichten von Daten	68
3.8 Analysieren von Daten	69

3.9	Bearbeitung und Verwenden des Outputs.....	74
3.9.1	Modifikation von Tabellen.....	74
3.9.2	Modifikation von Grafiken.....	78
3.9.3	Weiterverwenden des PASW-Outputs.....	82
3.9.4	Verwenden in anderen Programmen.....	82
3.10	Einlesen externer Daten.....	84
3.10.1	Import von Excel-Dateien.....	84
3.10.2	Import von Textdateien.....	85
3.11	Verwenden der Programmsprache (SYNTAX).....	87
3.12	Hilfe zu einzelnen Prozeduren, Tutorials und Links.....	91
3.13	Zusammenfassung der Konzepte.....	93
3.14	Übungen.....	93

Teil I Kategoriale Daten 95

Kapitel 4 Eine kategoriale Variable 97

4.1	Einleitung.....	98
4.2	Kommen alle Kategorien gleich häufig vor?.....	102
4.2.1	Numerische Beschreibung.....	103
4.2.2	Grafische Beschreibung.....	105
4.2.3	Statistische Analyse der Problemstellung.....	109
4.3	Entsprechen Häufigkeiten bestimmten Vorgaben?.....	114
4.3.1	Numerische und grafische Beschreibung.....	115
4.3.2	Statistische Analyse der Problemstellung.....	121
4.4	Hat ein Prozentsatz (Anteil) einen bestimmten Wert?.....	124
4.4.1	Statistische Analyse der Problemstellung.....	126
4.5	In welchem Bereich kann man einen Prozentsatz (Anteil) erwarten? . . .	130
4.6	Zusammenfassung der Konzepte.....	135
4.7	Übungen.....	135
4.8	Anhang: Die Chi-Quadrat-Verteilung oder wie entsteht ein p-Wert?	136

Kapitel 5 Mehrere kategoriale Variablen 139

5.1	Beschreibung mehrerer kategorialer Variablen.....	140
5.1.1	Numerische Beschreibung.....	141
5.1.2	Grafische Beschreibung.....	145
5.2	Ist die Verteilung von Häufigkeiten in verschiedenen Gruppen gleich?...	147
5.3	Unterscheiden sich Anteile in zwei oder mehreren Gruppen?	151
5.3.1	Eingabe einer Tabelle.....	152
5.3.2	Vergleich der Anteile.....	153
5.3.3	Exakter Test nach Fisher.....	155

5.4	Sind zwei kategoriale Variablen unabhängig?	157
5.4.1	Datenaufbereitung	158
5.4.2	Datenbeschreibung	160
5.4.3	Unabhängigkeitstest	160
5.5	Unterscheidet sich das Risiko in zwei Gruppen?	163
5.5.1	Odds-Ratio	163
5.5.2	Odds-Ratio-Test	165
5.6	Wie kann man Veränderungen von Anteilen testen?	167
5.6.1	McNemar-Test	168
5.6.2	Cochran-Test	170
5.7	Zusammenfassung der Konzepte	173
5.8	Übungen	174

Teil II Metrische Daten 177

Kapitel 6	Eine metrische Variable	179
6.1	Wie kann man die Verteilung einer metrischen Variablen beschreiben? ..	180
6.1.1	Klassifizieren, Tabellen und Histogramme	181
6.1.2	Maßzahlen zur Beschreibung der Verteilung	187
6.1.3	Boxplot	194
6.1.4	Ausreißer	196
6.2	In welchem Bereich kann man den Mittelwert in einer Grundgesamtheit erwarten?	197
6.3	Ist ein Mittelwert in der Grundgesamtheit anders als eine bestimmte Vorgabe?	201
6.4	Folgt eine metrische Variable einer bestimmten Verteilung?	203
6.4.1	Q-Q-Plot	204
6.4.2	Kolmogorov-Smirnov-Test	207
6.4.3	Anpassungstest mit χ^2 -Verteilung	209
6.5	Anhang: Stem-and-Leaf-Plot	211
6.6	Zusammenfassung der Konzepte	213
6.7	Übungen	213

Kapitel 7	Mehrere metrische Variablen	215
7.1	Wie stark ist der Zusammenhang zwischen zwei metrischen Variablen? .	216
7.1.1	Grafische Beschreibung	216
7.1.2	Korrelationskoeffizient nach Pearson	218
7.1.3	Korrelationskoeffizient nach Spearman	221
7.2	Welche Form hat der Zusammenhang zwischen zwei Variablen?	222
7.3	Lässt sich der Wert einer Variablen anhand des Werts einer zweiten vorhersagen?	227
7.3.1	Punktprognosen	227

	7.3.2	Intervallprognosen	229
7.4		Kann der Zusammenhang einer mit mehreren Variablen beschrieben werden?.....	230
	7.4.1	Multiple lineare Regression	231
	7.4.2	Bestimmtheitsmaß und F-Test	231
	7.4.3	Koeffizienten	233
	7.4.4	Prognose	234
	7.4.5	Dummyvariablen	234
	7.4.6	Kategoriale Variablen	237
	7.4.7	Überprüfung der Annahmen	240
	7.4.8	Typische Verletzungen von Modellannahmen	243
7.5		Unterscheiden sich Mittelwerte zu zwei oder mehreren Zeitpunkten? ...	246
	7.5.1	Grafische Beschreibung	247
	7.5.2	Analyse der Fragestellung	248
7.6		Zusammenfassung der Konzepte	250
7.7		Übungen	250

Teil III Kategoriale und metrische Daten 253

Kapitel 8		Metrische und kategoriale Variablen	255
8.1		Unterscheiden sich die Mittelwerte in zwei Gruppen?.....	256
	8.1.1	Grafische Beschreibung	256
	8.1.2	Analyse der Fragestellung	259
8.2		Unterscheidet sich die Lage einer Variablen zwischen zwei Gruppen? ...	261
	8.2.1	Grafische Beschreibung	261
	8.2.2	Analyse der Fragestellung	262
8.3		Unterscheiden sich die Mittelwerte mehrerer Gruppen?	265
	8.3.1	Grafische und numerische Beschreibung	265
	8.3.2	Analyse der Fragestellung	267
	8.3.3	Post-hoc-Tests	269
8.4		Unterscheidet sich die Lage einer Variablen zwischen mehreren Gruppen?	270
8.5		Wie wirken zwei kategoriale Variablen kombiniert auf Mittelwerte? ...	273
	8.5.1	Numerische und grafische Beschreibung	273
	8.5.2	Analyse der Fragestellung	275
8.6		Hängen Chancen von einer oder mehreren Variablen ab?	280
	8.6.1	Logistische Regression	280
	8.6.2	Logistische Regression mit mehreren erklärenden Variablen	283
8.7		Unterscheiden sich Chancen zwischen zwei oder mehreren Gruppen? ..	287
8.8		Zusammenfassung der Konzepte	289
8.9		Übungen	290

Teil IV	Multivariate Daten	293
Kapitel 9	Dimensionsreduktion	295
9.1	Kann man die Komplexität multidimensionaler Daten reduzieren?	296
9.1.1	Grundlagen der Hauptkomponentenanalyse	297
9.1.2	Anwendung der Hauptkomponentenanalyse	307
9.2	Weiterverwenden der Ergebnisse einer Hauptkomponentenanalyse	316
9.3	Übungen	320
9.4	Anhang: Extraktion der Hauptkomponenten für zwei Variablen	321
Kapitel 10	Gruppierung von Beobachtungen	325
10.1	Wie entdeckt man Gruppen ähnlicher Beobachtungen?	326
10.1.1	Distanz- und Ähnlichkeitsmaße	327
10.1.2	Clusterbildung	328
10.1.3	Outputteile	329
10.2	Wie findet man Cluster in den Variablen?	334
10.3	Wie können kategoriale Variablen in eine Clusteranalyse einbezogen werden?	336
10.3.1	Zweistufige Clusteranalyse	336
10.3.2	Outputteile	337
10.4	Zusammenfassung der Konzepte	343
10.5	Übungen	343
Literaturverzeichnis		345
Index		346