

Inhaltsverzeichnis

0	Einleitung	1
1	Univariate Statistik	3
1.1	Begriffsdefinitionen	3
1.1.1	Beobachtungseinheit, Merkmal	3
1.1.2	Merkmalstypen	3
1.1.3	Skalenniveaus	4
1.1.4	Häufigkeiten	4
1.2	Grafische Darstellung	5
1.2.1	Kreisdiagramm und Stabdiagramm	5
1.2.2	Histogramm	7
1.2.3	Empirische Verteilungsfunktion	10
1.3	Ausgewählte Kenngrößen	11
1.3.1	Ausgewählte Lagemaße	12
1.3.1.1	Mittelwerte	12
1.3.1.2	Quantile, Median	12
1.3.1.3	Modalwert	13
1.3.1.4	Anmerkungen zu Lagemaßen	13
1.3.2	Streuungsmaße	14
1.3.2.1	Spannweite	14
1.3.2.2	Standardabweichung, Varianz	14
1.3.2.3	Quartilsabstand	15
1.3.2.4	Variationskoeffizient	15
1.3.2.5	Anmerkungen zu Streuungsmaßen	16
1.3.3	Box-Whisker-Plots	16
1.4	Übungen	18
1.4.1	Testaufgaben	18
1.4.2	Fragestellungen	22
2	Bivariate Statistik	27
2.1	Wertepaare, Punktwolke	27
2.2	Regression von y auf x	27

2.3	Zusammenhangsmaße	31
2.3.1	Korrelationskoeffizient	31
2.3.2	Rang-Korrelationskoeffizient	33
2.3.3	Interpretation der Ergebnisse der Regressions- bzw. Korrelationsrechnung	35
2.4	Anmerkungen	36
2.4.1	Beschreibung der internen Konsistenz	36
2.4.2	Nachweis der Gleichheit zweier Messmethoden	38
2.4.3	Regression zur Mitte	39
2.5	Kontingenztafeln	40
2.6	Multivariate Analysen	40
2.7	Übungen	42
2.7.1	Testaufgaben	42
2.7.2	Fragestellungen	44
3	Wahrscheinlichkeitsrechnung	47
3.1	Wahrscheinlichkeit und relative Häufigkeit	47
3.1.1	Additionssatz	50
3.1.2	Multiplikationssatz	50
3.1.3	Laplace-Experimente	52
3.1.4	Wahrscheinlichkeitsbaum	52
3.2	Binomialverteilung	54
3.3	Stetige Zufallsvariable und Dichtefunktion	57
3.3.1	Kenngrößen der Verteilung einer Zufallsvariablen	58
3.3.2	Standardisierung einer Zufallsvariablen	59
3.4	Zentraler Grenzwertsatz und Normalverteilung	60
3.4.1	Tabelle der Normalverteilung	65
3.4.2	Logarithmische Normalverteilung	67
3.5	Approximation der Binomialverteilung durch die Normalverteilung	68
3.6	Übungen	71
3.6.1	Testaufgaben	71
3.6.2	Fragestellungen	75
4	Bedingte Wahrscheinlichkeiten und diagnostische Tests	79
4.1	Die bedingte Wahrscheinlichkeit	79
4.2	Diagnostische Tests	81
4.2.1	Prävalenz	82
4.2.2	Sensitivität	83
4.2.3	Spezifität	83
4.2.4	Positiver Vorhersagewert	84
4.2.5	Negativer Vorhersagewert	84
4.2.6	Bewertung eines diagnostischen Tests	86

4.2.7	Likelihood Ratios	87
4.2.8	Mehrfache Tests	89
4.2.9	Receiver-Operating Characteristic	90
4.3	Übungen	96
4.3.1	Testaufgabe	96
4.3.2	Fragestellungen	97
5	Punktschätzer, Konfidenzintervalle	103
5.1	Einleitung	103
5.2	Punktschätzung	104
5.3	Intervallschätzung	105
5.4	Definition eines Konfidenzintervalls	106
5.5	Beispiele und Konstruktion von Konfidenzintervallen	107
5.5.1	$(1 - \alpha)$ -Konfidenzintervall für den Erwartungswert normalverteilter Daten mit bekannter Varianz	107
5.5.2	$(1 - \alpha)$ -Konfidenzintervall für den Erwartungswert normalverteilter Daten mit unbekannter Varianz	110
5.5.3	$(1 - \alpha)$ -Konfidenzintervall für die Erfolgswahrscheinlichkeit p einer Binomialverteilung	110
5.5.4	Asymptotisches $(1 - \alpha)$ -Konfidenzintervall für einen Parameter	112
5.6	Übungen	113
5.6.1	Testaufgaben	113
5.6.2	Fragestellungen	115
6	Testen von Hypothesen I	117
6.1	Einleitung	117
6.2	Binomialtest	118
6.3	Signifikanzniveau	119
6.4	Fehler 1. und 2. Art	122
6.5	Übungen	131
6.5.1	Testaufgaben	131
6.5.2	Fragestellungen	135
7	Testen von Hypothesen II	139
7.1	Durchführung eines Experimentes	139
7.2	Einteilung von Tests	140
7.3	Mögliche Fehlerquellen bei der Anwendung statistischer Tests	141
7.4	Problematik des multiplen Testens	142
7.5	Vierfeldertest	144
7.6	Therapiebewertung	148
7.6.1	Maßzahlen der Therapiebewertung	148

7.6.2	Bewertung des Unterschiedes zweier Therapien an Hand von Konfidenzintervallen	151
7.6.3	Bewertung der Gleichwertigkeit zweier Therapien ...	153
7.7	Übungen	157
7.7.1	Fragestellungen	157
8	Testen von Hypothesen III	159
8.1	Vergleich abhängiger Stichproben	159
8.2	Vergleich unabhängiger Stichproben	165
8.3	Der Satz von Bayes als Basis für statistisches Schließen	172
8.4	Übungen	177
8.4.1	Testaufgaben	177
8.4.2	Fragestellungen	178
9	Analyse von Überlebenszeiten	181
9.1	Theoretische Überlebenskurve	181
9.2	Parametrische Modelle	183
9.3	Nichtparametrische Modelle	184
9.4	Produkt-Limit-Schätzer	184
9.5	Mediane Überlebenszeit	186
9.6	Methode der Sterbetafelanalyse	189
9.7	Vergleich von Überlebenskurven - Logrank-Test	193
9.8	Übungen	200
9.8.1	Testaufgaben	200
9.8.2	Fragestellungen	201
10	Studienplanung	203
10.1	Einleitung	203
10.2	Erhebungen	203
10.3	Experimente - Klinische Studien	204
10.3.1	Zielsetzung	204
10.3.2	Auswahl der Zielpopulation	205
10.3.3	Versuchsansatz	206
10.3.4	Randomisierung und Verblindung	209
10.3.5	Ziel- und Begleitvariable	216
10.3.6	Auswertungsstrategie	217
10.3.7	Effektmaß	219
10.3.8	Wahl des Stichprobenumfangs	220
10.3.9	Ausfälle von Beobachtungseinheiten	223
10.3.10	Unerwünschte Effekte	224
10.4	Verschiedene Aspekte der Studienplanung und -durchführung	224
10.4.1	Informations- und Wissensbeschaffung	224
10.4.2	Organisation und Dokumentation	228

10.4.3	Ethische und regulative Voraussetzungen	229
10.5	Übungen	230
10.5.1	Fragestellungen	230
11	Epidemiologie	231
11.1	Allgemeine Vorbemerkungen	231
11.1.1	Anteil, Verhältnis, Rate	231
11.2	Begriffsdefinition	232
11.3	Prävalenz und Inzidenz einer Krankheit	233
11.4	Krankheitsentwicklung	234
11.5	Statistische Tests und Assoziationsmaße bei Vierfeldertafeln	235
11.6	Einige wichtige epidemiologische Studienansätze	238
11.6.1	Kohortenstudie (Follow-up-Studie, Inzidenz-Studie, prospektive Studie, Längsschnittstudie)	238
11.6.2	Fall-Kontroll-Studie (retrospektive Studie)	245
11.6.3	Querschnitterhebung (Prävalenzstudie, survey)	249
11.6.4	Fall-Kohorten-Studie	251
11.6.5	Confounding (Vermengen)	251
11.6.6	Wechselwirkung (Interaktion, Effektmodifikation)	253
11.7	Übungen	256
11.7.1	Testaufgaben	256
11.7.2	Fragestellungen	258
12	Demographie	261
12.1	Einleitung	261
12.1.1	Verteilung der Todesfälle	261
12.1.2	Totale Todesrate	262
12.1.3	Altersspezifische Todesrate	264
12.2	Direkte Adjustierung	266
12.3	Indirekte Adjustierung	270
12.4	Vergleich der beiden Adjustierungsverfahren	273
13	Dokumentation und Informationsverarbeitung	275
13.1	Einleitung	275
13.2	Codeplan	275
13.3	Quelle der Daten	278
13.4	Datenerfassung	278
13.4.1	Erhebungsbögen	279
13.4.2	Datenbanken	281
13.4.3	Datenspeicherung	284
13.4.4	Plausibilitätskontrolle	284
13.5	Datenanalyse	286
13.5.1	EDV-gestützte statistische Analysen	286

13.6	Ergebnispräsentation und Publikation	289
13.7	Codierungssysteme	290
13.7.1	TNM-System	291
13.7.2	ICD-Schlüssel	292
13.7.3	SNOMED	292
13.7.4	Andere Schlüsselssysteme für klinische Studien	293
Anhang A:	Formelsammlung	294
	Algebraische Ausdrücke	295
	Lineare Interpolation	296
	Grafische Veranschaulichung einiger Funktionen	297
Anhang B:	Rechenblätter	302
Anhang C:	Lösungen zu MC-Fragen	308
Literatur		311
	Primärliteratur	311
	Begleitende Literatur	315
	Weiterführende Literatur	316
	Internetadressen	317
Symbole		318
Index		323

Beispielverzeichnis

1.1	Kreisdiagramm der relativen Häufigkeiten des Merkmals Geschlecht von $n = 55$ Probanden	5
1.2	Stabdiagramm der absoluten Häufigkeit des Merkmals "Anzahl der Kinder in den Familien" von $n = 55$ Probanden	6
1.3	Histogramm der relativen Häufigkeit des Merkmals "systolischer Blutdruck" von $n = 55$ Probanden	8
1.4	Empirische Verteilungsfunktion der relativen Häufigkeit des Merkmals "Blutzuckerkonzentration" bei $n = 52$ Probanden	10
1.5	Median und Quartile der "Blutzuckerkonzentrationen" von $n = 52$ Probanden	12
1.6	Modalwert der Blutzuckerkonzentrationen von $n = 52$ Probanden	13
1.7	Standardabweichung und Spannweite der "Blutzuckerkonzentrationen" von $n = 52$ Probanden	15
1.8	Box-Whisker-Plot des Merkmals "Cholesterin" von $n = 40$ Probanden	16
2.1	Regression des log-Adrenalins auf die Zeit	29
2.2	Korrelationskoeffizient zwischen log-Adrenalin und Zeit	32
2.3	Rang-Korrelationskoeffizient zwischen Adrenalinkonzentra- tion und Zeit	34
2.4	Kontingenztafel zur Beschreibung des Zusammenhangs zwischen dem Rauchverhalten und dem Geschlecht	40
3.1	Relative Häufigkeiten des Geschlechts von Neugeborenen in sieben Kliniken	48
3.2	Multiplikationssatz bei relativen Häufigkeiten	51
3.3	Wahrscheinlichkeiten für die möglichen Kombinationen von Blutgruppen bei zwei Personen	52

3.4	Wahrscheinlichkeit für $k = 4$ Universalspender unter 5 Blutspendern (Binomialverteilung)	54
3.5	Berechnung des Binomialkoeffizient für $n = 5$ und $k = 4$	56
3.6	Erwartungswert und Varianz der Binomialverteilung mit $n = 5$ und $p = 0.4$	59
3.7	Standardisierung des Risikogeburtsgewichtes für Knaben und Mädchen	60
3.8	Wahrscheinlichkeiten für Risikogeburten bei Knaben und Mädchen	65
4.1	Bewertung eines (Screening-)Tests zur Diagnosestellung HIV	85
4.2	ROC-Analyse der Glaskörper-Fluorometrie für die Diagnosestellung Präretinopathie	91
5.1	Konfidenzintervalle für den Erwartungswert einer Normalverteilung mit bekannter Varianz	108
5.2	Stichprobenumfang für ein Konfidenzintervall vorgegebener Breite für den Erwartungswert einer Normalverteilung mit bekannter Varianz	109
6.1	Präferenz von Orangengeschmack (Idee des Binomialtest) . . .	118
6.2	Einseitiger bzw. zweiseitiger p -Wert sowie Testentscheidung des Binomialtest	121
6.3	Verteilung einer möglichen Alternative $p = 0.8$ für den Binomialtest mit $n = 11$	124
6.4	Power vs. Stichprobenumfang und zweiseitige 5%- Ablehnbereiche für den Binomialtest $H_0 : p = 0.5$ vs $H_1 : p = 0.8$	127
7.1	Non-Hodgkin-Lymphome in Abhängigkeit vom Stadium	144
7.2	Non-Hodgkin-Lymphome in Abhängigkeit vom Stadium: Testentscheidung des χ^2 -Test	146
7.3	Therapiebewertung im Rahmen einer kontrollierten klinischen Studie (PVR-Studie)	148
7.4	Masszahlen (OR, ARR, RRR) zur Therapiebewertung im Rahmen einer kontrollierten klinischen Studie (PVR-Studie)	149
7.5	Number needed to treat zur Therapiebewertung im Rahmen einer kontrollierten klinischen Studie (PVR-Studie)	150
7.6	Bewertung des Therapieeffekts anhand des Konfidenzintervalls für die Differenz der Erfolgsraten im Rahmen einer kontrollierten klinischen Studie (PVR-Studie)	151
7.7	Gleichwertigkeit zweier medikamentöser ophthalmologischer Therapien	154
8.1	Schlafverlängerung unter medikamentöser Therapie	159
8.2	Vergleich der durchschnittlichen Schlafverlängerung in Abhängigkeit zweier medikamentöser Therapien (t -Test)	164

8.3	Gesamtprotein im Glaskörper bei männlichen PDR-Patienten (<i>U</i> -Test).....	166
9.1	Produkt-Limit-Schätzer und medianes rezidivfreies Überleben nach radiologischer Therapie für Morbus Hodgkin Patienten.....	186
9.2	Sterbetafelschätzer und medianes rezidivfreies Überleben nach Standardtherapie für Morbus Hodgkin.....	191
9.3	Logrank-Test für den Vergleich der Standardtherapie vs. neue Therapie an Hand des rezidivfreien Überlebens bei Morbus Hodgkin.....	197
10.1	Entwicklung einer Fragestellung zum Therapievergleich bei Patienten mit proliferativer Vitreoretinopathie (1991).....	204
10.2	Definition einer Zielpopulation zum Therapievergleich bei Patienten mit Blasenstörung (1988).....	206
10.3	Mögliche Ziel- und Begleitvariablen zu einer Melanom-Studie	217
10.4	Formulierung von Ziel- und Begleitvariablen im Rahmen einer Therapiestudie bei Patienten mit Blasenstörung (1988)	217
10.5	Mögliche Auswertungsstrategie zu einer Melanom-Studie (I)	217
10.6	Auswertungsstrategie im Rahmen einer Therapiestudie bei Patienten mit Blasenstörung (1988).....	218
10.7	Mögliche Auswertungsstrategie zu einer Melanom-Studie (II)	218
10.8	All-Randomized-Patients-Analyse im Rahmen einer Morbus Crohn-Studie.....	223
10.9	Themensuchmaschinen.....	228
11.1	Der Begriff "Rate" in den Wissenschaften.....	232
11.2	Odds-Ratio zur Beschreibung des Zusammenhangs zwischen Exposition und Erkrankung.....	237
11.3	Maße (RR, RD, OR) zur Beschreibung des Zusammenhangs zwischen Risiko und Erkrankung in Kohortenstudien (fiktiv)	241
11.4	Konfidenzintervall für RR zur Beschreibung des Zusammenhangs zwischen Risiko und Erkrankung in Kohortenstudien (fiktiv).....	243
11.5	Populationsattributierbares Risiko (PAR) in Kohortenstudien (fiktiv).....	244
11.6	Fall-Kontroll-Studie zur Untersuchung der Assoziation zwischen dem Auftreten von Myokardinfarkt und dem Gebrauch oraler Kontrazeptiva.....	248
11.7	Odds-Ratio und Prävalenz-Ratio in Querschnittstudien.....	250
11.8	Rauchen als Confounder bei der Ätiologie des Ösophaguskarzinoms.....	252
11.9	Wechselwirkung zwischen Rauchen und Asbestbelastung bei Lungenkrebs-erkrankung.....	254

12.1	Beschreibung der Häufigkeit von Todesfällen in den USA ...	261
12.2	Totale Mortalitätsrate der Todesfälle in den USA	263
12.3	Altersspezifische Todesraten der Todesfälle in den USA	265
12.4	Direkt adjustierte Krebsmortalität lediger vs. verheirateter Frauen	268
12.5	Mortalitätsraten und indirekt adjustierte Krebstodesraten in den USA (1929–1931)	270
13.1	Auszug aus einem Codierungsplan für eine Erhebung (fiktiv)	276
13.2	Relationale Datenbanken	282