

José L. Lozán • Hartmut Kausch

Angewandte Statistik für Natur- wissenschaftler

4. überarbeitete und ergänzte Auflage

2007 · Mit 55 Abbildungen, 43 Tabellen und 130 Beispielen

Wissenschaftliche Auswertungen 2007

INHALT

	Seite
Hinweise zur Benutzung dieses Buches (Siehe Innenseite des Titelblattes)	
Vorwort	5
Liste der verwendeten Symbole	13
Liste der wichtigsten Tests	14
Liste der wichtigsten Tabellen	15
Fachwörter auf deutsch, spanisch und englisch	16
1 EINFÜHRUNG	19
2 ERLÄUTERUNG EINIGER FACHBEGRIFFE	21
Statistik und Stochastik	21
Grundgesamtheit	21
Stichprobe	21
Nominale Daten	21
Ordinale Daten	22
Metrische Daten – Urliste	22
Häufigkeitsverteilung	23
Klassengröße	23
Klassenmitte (m_j)	23
»cm-below«	23
Offene Klassen	24
Zur Klassenbildung	24
Zufallsvariable	24
Toleranzbereich	24
Vertrauensbereich (= Konfidenzintervall)	25
Nullhypothese und Alternativhypothese	25
Konservative Tests	25
Einseitige und zweiseitige Tests	25
Teststärke = Trennschärfe	25
Statistische Tests	26
Korrelation und Scheinkorrelation	26
Ausreißerproblem	26
Freiheitsgrade	26
Funktionaler Zusammenhang	27
Stochastischer Zusammenhang	27
Rundungsregel, Signifikante Ziffern, Rundungsfehler	27
Quantile, Fraktile, Quartile	27
Die in der Statistik am häufigsten verwendeten Tabellen:	
Chi-Quadrat-Verteilung	28
Student-Verteilung	29
F-Verteilung	30
Studentisierte Spannweite	32

3	CHARAKTERISIERUNG VON STICHPROBEN	33
3.1	Mittelwerte	33
3.1.1	Der arithmetische Mittelwert (\bar{x})	33
	Der gewogene Mittelwert	
3.1.2	Der Zentralwert (Median) \tilde{x}	36
3.1.3	Das Dichtemittel (Modalwert) D	37
3.1.4	Der geometrische Mittelwert (G)	38
3.2	Variabilität	39
3.2.1	Variationsbreite = Spannweite (R)	40
3.2.2	Standardabweichung (s)	40
3.2.3	Klassengröße und Standardabweichung	40
3.2.4	Variabilitätskoeffizient V(%)	41
3.2.5	Standardfehler des Mittelwertes (s_x)	42
3.2.6	Interdezilbereich (I_{80})	43
3.2.7	Standardfehler des Medians (s_x)	44
3.2.8	Quartilsabstand (Q)	44
3.2.9	Box-and-Whisker-Plot	45
3.2.10	Messfehler von Bestimmungsmethoden	45
3.3	Vertrauensbereiche (Konfidenzintervalle)	47
3.3.1	Vertrauensbereich für einen Beobachtungswert	47
3.3.2	Vertrauensbereich für den wahren Mittelwert μ	47
3.4	Schiefe (S) und Exzess (E)	48
3.4.1	Schiefe	48
3.4.2	Exzess (Wölbung oder Kurtosis)	48
3.5	Randomisierung	50
3.6	Mindestanzahl der Messungen zur Schätzung des Mittelwertes	50
3.7	Mindestanzahl der Messungen zur Schätzung der Standardabweichung	51
3.8	Charakterisierung der Diversität	52
3.8.1	Diversitätsindices	52
3.8.1.1	Margalef-Index »Species richness« (D)	52
3.8.1.2	Diversitätsindex nach Shannon-Wiener (H_s)	52
3.8.1.3	Evenness (= Äquität) (E)	53
3.8.1.4	Diversitätsindex nach Heip (E_h)	53
3.8.1.5	Diversitätsindex nach Simpson (E_s)	53
3.8.1.6	Diversitätsindex nach Brillouin	53
3.8.2	Graphische Darstellung über die Dominanzstruktur	55
3.8.2.1	k-Dominanzkurve für die Artenhäufigkeit	55
3.8.2.2	k-Dominanzkurve für die Biomasse	56
3.8.2.3	k-Dominanzkurve für geometrische Häufigkeitsklassen	58
4	DARSTELLUNG VON STICHPROBEN	59
4.1	Graphische Darstellung	59
4.1.1	Kreisdiagramm	59
4.1.2	Stabdiagramm	62
4.1.3	Balkendiagramm	63
4.1.4	Liniendiagramm	64
4.2	Kurvenglättung / Korrekturfaktor	64
5.	THEORETISCHE VERTEILUNGEN: Normalverteilung	67
5.1	Dichtefunktion und Verteilungsfunktion	68
	Standardisierung: z-Transformation	68
5.2	Logarithmische Normalverteilung	72
5.3	Der zentrale Grenzwertsatz	74
5.4	Prüfung auf Nicht-Normalverteilung	74

5.4.1	Graphisch mit Hilfe des Wahrscheinlichkeitsnetzes	74
5.4.2	<i>Chi</i> quadrat-Test als Anpassungstest	75
5.4.3	<i>G</i> -Test als Anpassungstest	76
5.4.4	Test nach <i>Kolmogoroff-Smirnoff</i> (K-S-Test) (mod. nach <i>Kuiper</i>)	77
5.4.5	Schnelltest nach <i>David und Mitarbeitern</i>	79
5.4.6	Prüfung über die <i>Schiefe</i> (S) und den <i>Exzess</i> (E)	81
5.4.7	<i>Shapiro-Wilk</i> -Test	83
5.5	Ausreißer-Test nach Nalimov	84
5.6	Transformation zur Normalverteilung	85
5.6.1	Logarithmische Transformation	85
5.6.2	Wurzel-Transformation	85
5.6.3	Kehrwert- oder Reziproken-Transformation	85
5.6.4	Potenz-Transformation	85
5.7	Andere theoretische Verteilungen	85
5.7.1	Binomialverteilung	85
5.7.2	Poisson-Verteilung	88
6	VERGLEICH ZWEIER STICHPROBEN (Flussdiagramm)	91
	Der statistische Test	93
	Risiko I (Fehler 1. Art) Risiko II (Fehler 2. Art)	93
6.1	Vergleich zweier nicht-verbundener Stichproben	95
6.1.1	<u>Parametrische Verfahren</u>	95
6.1.1.1	Vergleich von Mittelwerten	95
	<i>t</i> -Test für große Stichproben mit gleichen Varianzen	95
	(Konfidenzgrenzen und Standardfehler für die Differenz der Mittelwerte)	96
	falls einzelne Werte vorliegen	97
	falls Werte in Klassengruppen vorliegen	98
	falls \bar{x} , s, n vorliegen	99
	<i>Welch</i> -Test für große Stichproben mit ungleichen Varianzen	100
	<i>Lord</i> -Test für kleine Stichproben mit gleichen Varianzen	101
	<i>Weir</i> -Test für kleine Stichproben mit ungleichen Varianzen	102
6.1.1.2	Prüfung auf Gleichheit der Varianzen	103
	<i>F</i> -Test	103
6.1.2	<u>Parameterfreie Verfahren</u> (Rangtests)	104
6.1.2.1	Vergleich von Mittelwerten	104
	<i>U</i> -Test für Stichproben mit gleicher Verteilungsform	104
	falls $n < m$ ist	107
	falls $n = m$ ist	108
	falls n, m größer 25 sind	109
	falls eine große Anzahl von Bindungen auftritt (Korrektur)	110
	<i>Median</i> -Test für den Vergleich zweier Stichproben	112
6.1.2.2	Prüfung auf Gleichheit der Verteilung	114
	<i>Kolmogoroff-Smirnoff-Homogenitätstest</i>	114
6.1.2.3	Prüfung auf Gleichheit der Varianzen	115
	<i>Pfanzagl</i> -Test	115
6.2	Vergleich zweier verbundener Stichproben	116
6.2.1	<u>Parametrisches Verfahren</u>	116
6.2.1.1	Vergleich zweier abhängiger Stichproben	116
	<i>t</i> -Test für den Vergleich von Paardifferenzen	116
6.2.1.2	Prüfung auf Gleichheit der Varianzen zweier abhängiger Stichproben	118
	<i>t</i> -Test für den Vergleich der Varianzen zweier verbundener Stichproben	118
6.2.2	<u>Parameterfreie Verfahren</u> (Rangtests)	119
	<i>Wilcoxon</i> -Test für den Vergleich von Paardifferenzen	119
	<i>Vorzeichen</i> -Test für den Vergleich von Paardifferenzen	122

7	VARIANZANALYSE (Vergleich mehrerer Stichproben)	123
	(Flussdiagramm)	
7.1	Einfache Varianzanalyse	124
7.1.1	Vergleich mehrerer nicht-verbundener Stichproben	126
7.1.1.1	<u>Parametrische Verfahren</u>	126
	Signifikanzprüfung	126
	<i>F-Test</i> für mehrere normalverteilte Stichproben	126
	falls einzelne Werte vorliegen	127
	falls Werte in Klassengruppen vorliegen	128
	falls \bar{x} , s, n vorliegen	130
	falls mehrere Variationsursachen bestehen	131
	Prüfung auf Gleichheit mehrerer Varianzen	132
	<i>Bartlett-Test</i>	132
	<i>Cochran-Test</i>	134
	Prüfung mehrerer Verteilungen auf Homogenität	135
	<i>Informationsstatistik-Test</i>	135
	Multipler Vergleich von Mittelwerten	136
	<i>Tukey-Test</i>	136
	<i>Student-Newman-Keuls-Test</i>	139
7.1.1.2	<u>Parameterfreie Verfahren</u> (Rangtests)	140
	Signifikanzprüfung	140
	<i>H-Test</i> für mehrere Stichproben mit gleicher Verteilungsform	140
	falls drei Stichproben vorliegen	142
	falls mehr als drei Stichproben vorliegen	143
	falls eine große Anzahl von Bindungen auftritt	144
	<i>Erweiterter Median-Test</i> für Stichproben ungleicher Verteilungsform	145
	Prüfung auf Gleichheit mehrerer Varianzen	147
	<i>Pfanzagl-Test</i>	147
	Multipler Vergleich von Mittelwerten	148
	<i>Nemenyi-Test</i>	148
	falls die Stichproben gleich besetzt sind	149
	falls die Stichproben ungleich besetzt sind	150
7.1.2	Vergleich mehrerer verbundener Stichproben	151
7.1.2.1	<u>Parametrische Verfahren</u>	151
	Signifikanzprüfung	151
	<i>F-Test</i> für mehrere verbundene Stichproben	151
7.1.2.2	<u>Parameterfreie Verfahren</u> (Rangvarianzanalyse)	152
	Signifikanzprüfung	152
	<i>Friedman-Test</i>	152
	falls vier oder weniger Stichproben vorliegen	154
	falls mehr als vier Stichproben vorliegen	155
	falls eine große Anzahl von Bindungen auftritt	156
	Multipler Vergleich von Mittelwerten	157
	<i>Wilcoxon-Wilcox-Test</i>	157
7.2	Zweifache Varianzanalyse	159
7.2.1	Einfache Besetzung	159
	Signifikanzprüfung	160
	Multipler Vergleich von Mittelwerten	161
	<i>Tukey-Test</i>	161
7.2.2	Mehrfache Besetzung	165
	Signifikanzprüfung	167
	Multipler Vergleich von Mittelwerten	167
	<i>Tukey-Test</i>	167

7.3	Dreifache Varianzanalyse	170
7.3.1	Einfache Besetzung (Modell I)	170
	Signifikanzprüfung	170
	Multipler Vergleich von Mittelwerten	174
	<i>Tukey-Test</i>	174
7.3.2	Das lateinische Quadrat	175
7.3.3	Varianzanalyse: Versuche mit mehreren Faktoren auf verschiedenen Stufen . .	177
	Hauptwirkungen und Wechselwirkungen	178
	<i>Yates-Tafel</i>	179
	Signifikanzprüfung	180
8	PRÜFUNG VON ABHÄNGIGKEITEN	181
8.1	Prüfung auf Abhängigkeit stetiger Zufallsvariabler	181
	<i>Regression und Korrelation</i>	181
8.1.1	Einfache Regression (<u>parametrisch</u>)	181
8.1.1.1	<u>Die »ordinary« oder »predictive« Regressionsgerade (Modell I)</u>	182
	Schätzung der Regressionsparameter	183
	Prüfung auf Korrelation	186
	• <i>F-Test</i>	188
	• <i>Korrelationskoeffizient r</i>	188
	• <i>Bestimmtheitsmaß B = r²</i>	189
	• <i>Prüfung der Hypothese $\Gamma = 0$</i>	190
	• <i>Prüfung der Hypothese $\beta = 0$</i>	190
	Überprüfung der Voraussetzungen	194
	• <i>Prüfung auf Linearität</i>	194
	• <i>Prüfung auf Autokorrelation der Residuen (Durbin-Watson-Test)</i>	198
	Standardabweichungen, Vertrauens- und Toleranzbereiche	195
8.1.1.2	<u>Die mittlere »Geometrical-Mean«-Regressionsgerade (Modell II)</u>	199
	Schätzung der Regressionsparameter	203
	Standardabweichung, Vertrauens- und Toleranzbereiche (Ellipsen)	204
	Orthogonale Regressionskoeffizienten b_o , $1/b_o$	204
	Länge der Halbachsen der Ellipse	204
8.1.1.3	<u>Vergleich des Regressionskoeffizienten mit einem theoretischen Wert</u>	208
8.1.1.4	<u>Linearisierbare Regressionsfunktionen</u>	209
	Linearisierende Transformationen	209
8.1.2	Einfache Regression (<u>parameterfrei</u>)	214
	<i>Spearman-Rangkorrelation</i>	214
	<i>Kendall-Rangkorrelation</i>	216
8.2	Prüfung auf Abhängigkeiten nicht-stetiger Zufallsvariabler	218
8.2.1	<i>Chiquadrat-Test</i> zur Prüfung auf Abhängigkeit zweier Stichproben	218
	(Korrektur nach <i>Yates</i>)	
8.2.2	<i>G-Test</i> zur Prüfung auf Abhängigkeit zweier Stichproben	221
8.2.3	<i>Chiquadrat-Test</i> zur Homogenitätsprüfung zweier oder mehrerer Stichproben.	222
	Formel von <i>Brandt & Snedecor</i> für zwei Stichproben	223
8.2.4	<i>G-Test</i> zur Homogenitätsprüfung zweier oder mehrerer Stichproben	224
8.2.5	<i>Chiquadrat-Test</i> zum Vergleich von Verteilungen	225

9	MULTIVARIATE VERFAHREN	227
9.1	Multiple lineare Regression	228
9.1.1	Schätzung der Regressionsparameter	229
9.1.2	Prüfung auf Korrelation	233
	<i>F-Test</i>	233
	<i>Bestimmtheitsmaß B</i>	233
9.1.3	Überprüfung der Voraussetzungen	235
	Prüfung auf Nicht-Normalverteilung der Residuen	235
	Konstanz der Variabilität über die n-untersuchten Perioden	235
	Prüfung auf Autokorrelation der Residuen (<i>Durbin-Watson-Koeffizient</i>)	235
9.1.4	Prüfung auf Redundanz	237
9.2	Der partielle Korrelationskoeffizient	239
9.3	Kovarianzanalyse (Vergleich mehrerer Regressionsgeraden)	242
9.3.1	Prüfung auf Parallelität	242
9.3.2	Prüfung des Abstands der Regressionsgeraden von null	243
	falls zwei Regressionsgeraden vorliegen	244
	falls mehr als zwei Regressionsgeraden vorliegen	246
9.4	Eigenwertbasierende Methoden der multivariaten Statistik	249
9.4.1	Hauptkomponentenanalyse (PCA)	252
9.4.2	Faktorenanalyse	255
9.4.2.1	Faktorenanalyse nach der Hauptkomponentenmethode	255
9.4.2.2	Faktorenanalyse nach der Hauptachsenmethode	259
9.4.3	Kanonische Korrelationsanalyse und Redundanzanalyse	261
9.4.4	Diskriminanzanalyse	265
9.5	Gradientenanalyse	268
9.6	Ähnlichkeits-Distanz-Methoden	273
9.6.1	Clusteranalyse	276
9.6.2	Mehrdimensionale Skalierung (MDS)	279
10	PROBITANALYSE	281
10.1	Probit-Transformation	281
10.2	Probit-Regression nach der graphischen Methode	283
10.3	Probit-Regression nach der Maximum-Likelihood-Methode	285
	Schätzung von Gewichtskoeffizienten W	285
	Schätzung der Rechenprobits	286
	Schätzung der korrigierten Probit-Regression	287
	Prüfung auf Anpassung der korrigierten Probit-Regression	288
	Schätzung des L_{50} -Wertes	288
	Schätzung der Standardabweichung des L_{50} -Wertes	288
	Schätzung der Standardabweichung der Regressionssteigung b	288
10.4	Vergleich von L_{50}-Werten	289
11	LOGITANALYSE	293
11.1	Logit-Transformation	293
	Berechnung der empirischen Logit-Regression	293
11.2	Logit-Regression nach der Maximum-Likelihood-Methode	293
	Ermittlung der provisorischen Logits L_0	293
	Korrektur der empirischen Logit-Regression mit Gewichtskoeffizienten W	294
	Schätzung des L_{50} -Wertes und dessen Standardabweichung sL_{50}	294
	Prüfung auf Güte der Anpassung der korrigierten Logit-Regression	296
12	ZITIERTE LITERATUR	297
13	SACHVERZEICHNIS	300