

Statistische Methoden der Datenanalyse

Vorlesung im Sommersemester 2002

H. Kolanoski
Humboldt-Universität zu Berlin

Inhaltsverzeichnis

Literaturverzeichnis	iii
1 Grundlagen der Statistik	3
1.1 Wahrscheinlichkeit	3
1.1.1 Definition über die Häufigkeit	3
1.1.2 Kombinatorische Definition	3
1.1.3 Axiomatische Definition der Wahrscheinlichkeit	4
1.2 Verteilungen von Zufallsvariablen	6
1.2.1 Eigenschaften von Verteilungen	7
1.2.2 Erwartungswerte	9
1.2.3 Wahrscheinlichster Wert und Median	11
1.2.4 Stichproben und Schätzwerte	11
1.3 Simulation von Verteilungen	12
1.3.1 Umkehrung der Verteilungsfunktion	13
1.3.2 ‘Hit and Miss’ Methode	14
2 Spezielle Verteilungen einer Variablen	17
2.1 Binomial-Verteilung	17
2.2 Multinomial-Verteilung	22
2.3 Poisson-Verteilung	22
2.4 Gleichverteilung	24
2.5 Normalverteilung	27
2.6 Zentraler Grenzwertsatz	32
3 Verteilungen mehrerer Variablen	33
3.1 Eigenschaften von Verteilungen mehrerer Variablen	33
3.1.1 Wahrscheinlichkeitsdichte, Verteilungsfunktion, Randverteilung	33
3.1.2 Bedingte Wahrscheinlichkeitsdichten, Selektionsschnitte	34
3.2 Erwartungswerte	35
3.3 Kovarianzmatrix	36
3.3.1 Definition und Eigenschaften der Kovarianzmatrix	36
3.3.2 Beispiel: Multi-dimensionale Gaussverteilung	36
3.3.3 Kovarianzmatrix unabhängiger Variabler	37
3.3.4 Korrelationen	39
3.4 Lineare Funktionen von mehreren Zufallsvariablen	40
3.5 Nicht-lineare Funktionen von Zufallsvariablen	42
3.5.1 Eine Funktion von einem Satz von Zufallsvariablen	42
3.5.2 Mehrere Funktionen von einem Satz von Zufallszahlen	43

3.6	Transformationen von Zufallsvariablen	45
4	Stichproben und Schätzungen	49
4.1	Stichproben, Verteilungen und Schätzwerte	49
4.2	Eigenschaften von Schätzwerten	50
4.3	Stichproben aus Normalverteilungen; χ^2 -Verteilung	52
5	Die Maximum-Likelihood-Methode	59
5.1	Das Maximum-Likelihood-Prinzip	59
5.2	Fehlerbestimmung für ML-Schätzungen	61
5.2.1	Allgemeine Methoden der Varianzabschätzung	62
5.2.2	Varianzabschätzung durch Entwicklung um das Maximum	62
5.2.3	Vertrauensintervalle und Likelihood-Kontouren	63
5.3	Eigenschaften von ML-Schätzungen	65
6	Methode der kleinsten Quadrate	67
6.1	Prinzip der Methode der kleinsten Quadrate	67
6.2	Lineare Anpassung	68
6.2.1	Anpassung der Messwerte an eine Gerade	68
6.2.2	Anpassung einer allgemeinen linearen Funktion der Parameter	71
6.3	Anpassung nicht-linearer Funktionen der Parameter	75
7	Neuronale Netze zur Datenklassifikation	79
7.1	Einleitung	79
7.1.1	Attraktivität neuronaler Modelle	79
7.2	Natürliche und künstliche neuronale Netze	80
7.2.1	Natürliche neuronale Netze	80
7.2.2	Künstliche neuronale Netze (KNN)	84
7.3	Feed-Forward-Netzwerke	86
7.3.1	Das einfache Perzeptron	87
7.3.2	Das Mehrlagen-Perzeptron	90
7.3.3	Lernen	94
7.3.4	Typische Anwendungen für Feed-Forward-Netze	98
7.3.5	BP-Lernen und der Bayes-Diskriminator	101

Literaturverzeichnis

- [1] S. Brandt: 'Datenanalyse', 4. Auflage, 1999, Spektrum Akademischer Verlag.
- [2] R.J. Barlow, 'Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences', Wiley, 1989.
- [3] V. Blobel und E. Lohrmann, 'Statistische und numerische Methoden der Datenanalyse', Teubner Studienbücher, 1998.

Zu "Neuronale Netze":

- [4] D.E.Rumelhart and J.L.McClelland: '*Parallel Distributed Processing*', MIT Press 1984 (9.Aufl. 1989).
- [5] J.Hertz, A.Krogh and R.G.Palmer: '*Introduction to the Theory of Neural Computation*', Addison-Wesley Publishing Company, 1991.
- [6] R.Brause: '*Neuronale Netze*', Teubner Verlag 1991.
- [7] R.Hecht-Nielsen: '*Neurocomputing*', Addison-Wesley Publishing Company, 1987.
- [8] H.Ritter, T.Martinetz und K.Schulten: '*Neuronale Netze. Eine Einführung in die Neuroinformatik selbstorganisierender Netze*', Addison-Wesley Publishing Company, 1991.
- [9] G.E.Hinton: '*Wie Neuronale Netze aus Erfahrung lernen*', Spektrum der Wissenschaft, Nov. 1992.
- [10] T.Kohonen: '*Self-Organization and Associative Memory*', Springer Verlag, 3.Auflage 1989.
- [11] A.Zell: '*Simulation Neuronaler Netze*', Addison-Wesley, 1.Auflage 1994.
- [12] Scientific American: '*The Brain*', Vol. 241, Sept. 1979.
- [13] Spektrum der Wissenschaft, Nov. 1992.

1.15 (31619) Statistische Methoden der Datenanalyse

Lehrende(r):

Prof. Dr. H. Kolanoski

Art der Veranstaltung:

Vorlesung + Übung

Umfang:

VL: 1 SWS UE: 1 SWS

Zielgruppe:

Diplom und Lehramt Physik, Grundstudium ab 3. Semester oder Hauptstudium

Einordnung:

Die Vorlesung vermittelt das Grundwissen auf dem Gebiet der statistischen Behandlung von Daten, das jede/r Experimentalphysiker/in (spätestens) bei Anfertigung der Diplom- oder Studienarbeit benötigt.

Fakultative Lehrveranstaltung bzw. wahlfreier Kurs im WPF Elementarteilchenphysik.

Voraussetzung:

Mathematische Grundkenntnisse wie sie in den ersten Semestern des Grundstudiums erworben werden, insbesondere Vektor- und Matrizenrechnung, Diff.- und Integralrechnung. Erfahrung im Programmieren und dem Umgang mit Computern ist wünschenswert.

Grobgliederung:

- Grundbegriffe
- Wahrscheinlichkeitsverteilungen
- Fehlerfortpflanzung
- Simulation nach der Monte-Carlo-Methode
- Anpassungsrechnung
- Test von Hypothesen
- Mehrdimensionale Wahrscheinlichkeitsverteilungen und Korrelationen
- Entfaltungsalgorithmen
- Multidimensionale Datenanalyse

Literatur:

- **S. Brandt:** "Datenanalyse", BI Wissenschaftsverlag
- **R.J. Barlow:** "Statistics: A Guide to the Use of Statistical [NEWLINE] Methods in the Physical Sciences", Wiley, 1989
- **H.Kolanoski:** "Stat.Methoden der Datenanalyse (Vorlesung SS 2000)", <http://eeh01.physik.hu-berlin.de/stat/statistics.html>

Leistungsnachweis:

Bedingungen für Leistungsnachweis (nicht erforderlich !):

1) mindestens 50 Prozent der Gesamtpunkte in den Übungsaufgaben,