

Gerhard Keller

Mathematik in den Life Sciences

Grundlagen der Modellbildung und Statistik
mit einer Einführung in die Statistik-Software R

49 Abbildungen

Verlag Eugen Ulmer Stuttgart

Inhaltsverzeichnis

| | | |
|----------|--|------------|
| 1 | Einführung | .11 |
| 1.1 | Warum Mathematik? | 11 |
| 1.2 | Vorbereitende und ergänzende Literatur. | 13 |
| 2 | Mathematische Grundbegriffe. | .15 |
| 2.1 | Zahlen. | 15 |
| 2.2 | Rechenregeln. | 16 |
| 2.3 | Zahlen als Messergebnisse. | 17 |
| 2.3.1 | Messgenauigkeit, Runden. | 17 |
| 2.3.2 | Maßeinheiten | 18 |
| 2.3.3 | Mol und Molekulargewicht | 18 |
| 2.4 | Vektoren, Matrizen. | 19 |
| 2.5 | Matrizenmultiplikation. | 21 |
| 2.6 | Zahlenfolgen. | 22 |
| 2.7 | Funktionen. | 23 |
| 2.8 | Bemerkungen zum Rechnen mit Logarithmen. | 25 |
| 2.9 | Fragen und Aufgaben. | 25 |
| 3 | Differenzieren, Ableitung | .27 |
| 3.1 | Ableitung von Funktionen einer Variablen. | 27 |
| 3.2 | Ableitungsregeln. | 29 |
| 3.3 | Integral und Stammfunktion. | 32 |
| 3.4 | Partielle Ableitungen. | 33 |
| 3.5 | Fragen und Aufgaben. | 34 |
| 4 | Grafische Darstellung von Daten und beschreibende Statistik | .35 |
| 4:1 | Datenvektoren und Datenmatrizen. | 35 |
| 4.2 | Beschreibende Statistik - Grundbegriffe. | 37 |
| 4.3 | Eindimensionale Stichproben. | 38 |
| 4.3.1 | Nominale Merkmale. | 38 |
| 4.3.2 | Metrische Merkmale | 39 |
| 4.3.3 | Statistische Kennzahlen. | 41 |
| 4.4 | Zweidimensionale Stichproben. | 43 |
| 4.5 | Lineare Regression. | 45 |
| 4.6 | Allometrie. | 48 |
| 4.7 | Fragen und Aufgaben. | 51 |
| 5 | Wachstumsmodelle: unbeschränktes Wachstum | .53 |
| 5.1 | Lineares Wachstum. | 53 |
| 5.2 | Exponentielles Wachstum - diskrete Zeit | 54 |
| 5.2.1 | Modellwahl. | 58 |
| 5.2.2 | Quadratische Abweichung. | 59 |

| | | |
|----------|--|------------|
| 5.3 | Exponentielles Wachstum - stetige Zeit | 60 |
| 5.3.1 | Von diskreter zu stetiger Zeit | 60 |
| 5.3.2 | Die Differenzialgleichung für exponentielles Wachstum in stetiger Zeit. | 61 |
| 5.3.3 | Kommentar aus der Sicht der Mathematik | 62 |
| 5.3.4 | Lineare Regression bei exponentiellem Wachstum. | 63 |
| 5.3.5 | Zusammenfassung zum exponentiellen Wachstum. | 64 |
| 5.3.6 | Exponentielles Aussterben. | 64 |
| 5.3.7 | Verdopplungszeit, Halbwertszeit | 65 |
| 5.4 | Fragen und Aufgaben. | 66 |
| 6 | Wachstumsmodelle: beschränktes Wachstum | 67 |
| 6.1 | Logistisches Wachstum. | 67 |
| 6.1.1 | Ein paar grundsätzliche Bemerkungen zum Begriff der Differenzialgleichung | 71 |
| 6.1.2 | Bemerkungen zum numerischen Lösen einer Differenzialgleichung | 72 |
| 6.1.3 | Anpassung des logistischen Modells an Daten. | 73 |
| 6.1.4 | Ein Residuenplot | 74 |
| 6.2 | Stabilisierung bei konstantem Zufluss und exponentiellem Abbau | 75 |
| 6.3 | Variationen zum logistischen Wachstum. | 76 |
| 6.3.1 | Ein logistisches Modell mit „Bejagung“. | 76 |
| 6.3.2 | Ein Modell mit zwei stabilen Gleichgewichten. | 79 |
| 6.4 | Zeitverzögerungen. | 80 |
| 6.5 | Zwei Modelle der chemischen Reaktionskinetik. | 81 |
| 6.6 | Fragen und Aufgaben. | 85 |
| 7 | Modelle der Populationsgenetik | 87 |
| 7.1 | Das Hardy-Weinberg-Modell. | 87 |
| 7.2 | Inzucht. | 91 |
| 7.3 | Selektion. | 92 |
| 7.4 | Fragen und Aufgaben. | 97 |
| 8 | Wachstumsmodelle: zwei Populationen. | 98 |
| 8.1 | Das Räuber-Beute-Modell von Lotka und Volterra. | 98 |
| 8.2 | Ein einfaches Epidemiemodell. | 103 |
| 8.3 | Fragen und Aufgaben. | 105 |
| 9 | Wahrscheinlichkeitsrechnung | 106 |
| 9.1 | Zufallsvariablen | 106 |
| 9.1.1 | Diskrete Zufallsvariablen. | 107 |
| 9.1.2 | Kontinuierliche Zufallsvariablen. | 108 |
| 9.2 | Unabhängigkeit diskreter Zufallsvariablen. | 109 |
| 9.2.1 | Bedingte Wahrscheinlichkeit und Unabhängigkeit | 110 |
| 9.2.2 | Die Binomialverteilung. | 110 |
| 9.3 | Unabhängigkeit kontinuierlicher Zufallsvariablen. | 112 |
| 9.4 | Histogramm unabhängiger Beobachtungen. | 112 |
| 9.5 | Erwartungswert und Varianz. | 113 |
| 9.5.1 | Erwartungswert und Varianz diskreter Zufallsvariablen. | 113 |

| | | |
|-----------|--|-------------|
| 9.5.2 | Erwartungswert und Varianz kontinuierlicher Zufallsvariablen. | .115 |
| 9.6 | Normal- und Poisson-Approximation der Binomialverteilung. | .116 |
| 9.6.1 | Verteilungsfunktionen binomialverteilter Zufallsvariablen. | .116 |
| 9.6.2 | Normalapproximation der Binomialverteilung. | .117 |
| 9.6.3 | Poisson-Approximation der Binomialverteilung. | .118 |
| 9.7 | Fragen und Aufgaben. | .119 |
| 10 | Beurteilende Statistik: Testen. | .120 |
| 10.1 | Der Binomialtest. | .120 |
| 10.1.1 | Formulierung des Testproblems. | .120 |
| 10.1.2 | Durchführung des Tests. | .121 |
| 10.1.3 | Unabhängigkeit der Beobachtungen. | .122 |
| 10.2 | Chi-Quadrat-Tests. | .123 |
| 10.3 | Fragen und Aufgaben. | .129 |
| 11 | Beurteilende Statistik: Schätzen. | .131 |
| 11.1 | Schätzen von Erfolgswahrscheinlichkeiten. | .131 |
| 11.2 | Konfidenzintervall für den Erwartungswert. | .134 |
| 11.2.1 | Konfidenzintervall bei normalverteilten Beobachtungen. | .135 |
| 11.2.2 | Der Ein-Stichproben-t-Test. | .136 |
| 11.3 | Fragen und Aufgaben. | .137 |
| 12 | Beurteilende Statistik: Korrelation und Regression. | .138 |
| 12.1 | Ist der Korrelationskoeffizient signifikant von Null verschieden?. | .138 |
| 12.2 | Die statistische Beurteilung der geschätzten Regressionskoeffizienten | .140 |
| 12.3 | Vorsicht bei linearer Regression. | .142 |
| 12.4 | Fragen und Aufgaben. | .142 |
| 13 | Einführung in das Sequenz-Alignment. | .143 |
| 13.1 | Scoring-Modelle zur Bewertung von Alignments. | .143 |
| 13.1.1 | Scoring bei DNA-Alignments. | .143 |
| 13.1.2 | Scoring bei Proteinsequenz-Alignments. | .144 |
| 13.2 | Scores und Wahrscheinlichkeiten. | .146 |
| 13.3 | Der Needleman-Wunsch-Algorithmus. | .148 |
| 13.1.1 | Die Grundidee des Needleman-Wunsch-Algorithmus. | .149 |
| 13.3.2 | Eine Realisierung des Needleman-Wunsch-Algorithmus für den Vergleich zweier DNA-Sequenzen in R. | .151 |
| 13.3.3 | Beispiele zum Needleman-Wunsch-Algorithmus. | .152 |
| 13.3.4 | Der Smith-Waterman-Algorithmus. | .153 |
| 13.4 | Clustering. | .154 |
| 13.5 | Fragen und Aufgaben. | .156 |
| R | Einführung in R. | .157 |
| R1 | Erste Schritte. | .158 |
| R1.1 | R als Taschenrechner. | .158 |
| R1.2 | Eine erste Grafik. | .161 |

| | | |
|-------|--|------|
| R2 | Grundlegende Begriffe | .162 |
| R2.1 | Variablen | .162 |
| R2.2 | Folgen | .163 |
| R2.3 | Die erzeugten Objekte | .165 |
| R3 | Funktionen, Nullstellen, Maxima, Minima, R-Hilfe | .166 |
| R3.1 | Funktionen | .166 |
| R3.2 | Funktionsgraphen | .167 |
| R3.3 | Hilfe in der Dokumentation | .168 |
| R3.4 | Nullstellen, Maxima und Minima | .169 |
| R4 | Funktionen mehrerer Variablen, der Workspace von R | .171 |
| R4.1 | Funktionen mehrerer Variablen | .171 |
| R4.2 | Wie funktioniert R im Hintergrund? | .172 |
| R5 | Vektoren, Matrizen, der Dateneditor | .174 |
| R5.1 | Vektoren | .174 |
| R5.2 | Matrizen | .176 |
| R6 | Matrizenmultiplikation, Dotplots | .180 |
| R6.1 | Matrizenmultiplikation | .180 |
| R6.2 | Der Befehl <code>outer()</code> | .182 |
| R6.3 | Eine Vergleichstabelle für Sequenzvergleiche und ein Dotplot | .183 |
| R7 | Datensätze, R Commander, beschreibende Statistik | .185 |
| R7.1 | Der R Commander | .185 |
| R7.2 | Datensätze | .186 |
| R7.3 | Speichern von Programmen und Objekten | .189 |
| R7.4 | Lineare Regression und Korrelation | .190 |
| R8 | Datenim- und -export, Grafikexport | .193 |
| R8.1 | Erstellen und Einlesen eigener Datensätze | .193 |
| R8.2 | Grafik-Export | .195 |
| R9 | Exponentielles Wachstum und Abklingen | .197 |
| R9.1 | Zinseszins mit R als Taschenrechner | .197 |
| R9.2 | Exponentielles Wachstum - US-Bevölkerungsdaten | .197 |
| R9.3 | Exponentieller Abbau - Medikamentenabbau im Körper | .198 |
| R10 | Nichtlineare Regression | .200 |
| R10.1 | Logistisches Wachstum - US-Bevölkerungsdaten | .201 |
| R10.2 | Biexponentielles Abklingen | .201 |
| R10.3 | Michaelis-Menten-Funktion | .202 |
| R11 | Binomial-, Normal- und Poisson-Verteilung | .203 |
| R11.1 | Die Binomialverteilung | .203 |
| R11.2 | Die Normalverteilung | .204 |
| R11.3 | Die Poisson-Verteilung | .205 |
| R11.4 | Plotten von Dichten und Verteilungsfunktionen | .205 |

| | | |
|-------|--|------------|
| R12 | Binomialtest und Chi-Quadrat-Tests. | 206 |
| R12.1 | Die Binomialverteilung und der Binomialtest | 206 |
| R12.2 | /-Tests. | 208 |
| R13 | Schätzen und Testen bei normalverteilten Beobachtungen. | 209 |
| R13.1 | Konfidenzintervalle bei normalverteilten Beobachtungen | 209 |
| R13.2 | Ein-Stichproben-t-Test | 211 |
| R13.3 | Zwei-Stichproben-t-Test - verbundene Strichproben. | 212 |
| R13.4 | Statistik zur linearen Regression. | 213 |
| R14 > | Sequence Alignment,. | 214 |
| R14.1 | Die Datenbank Genbank. | 214 |
| R14.2 | Die Bereitstellung von Sequenzen für R. | 215 |
| R14.3 | Needleman-Wunsch-Algorithmus für Protein-Sequenzen. | 216 |
| R14.4 | Gleichzeitiger Vergleich mehrerer Sequenzen | 217 |
| | Verzeichnisse. | 219 |
| | R-Codes zu ausgewählten Abbildungen. | 219 |
| | Literatur. | 225 |
| | Sachregister. | 227 |
| | Index der R-Befehle. | 231 |