

Statistik für Verfahrenstechniker
und Chemie-Ingenieure

Jürgen Raasch

unter Mitarbeit von Wulf Alex

2010

Karlsruhe

Ausgabedatum: 1. November 2011

Jürgen Raasch
Am Kirchberg 43
76229 Karlsruhe

juergen.k.raasch@t-online.de

Wulf Alex
Rieslingweg 14
76356 Weingarten (Baden)

alex-weingarten@t-online.de

Das Skriptum ist unvollständig und kann Fehler enthalten. Für Hinweise sind wir dankbar.

Skriptum, Korrekturen und Ergänzungen:

<http://www.alex-weingarten.de/skripten/statistik/>
mit den Hyperlinks

- <http://www.alex-weingarten.de/skripten/statistik/buch.pdf>
(Ganzes Skriptum)
- <http://www.alex-weingarten.de/skripten/statistik/vorwort.pdf>
(Vorwort)
- <http://www.alex-weingarten.de/skripten/statistik/inhalt.pdf>
(Inhalt)
- <http://www.alex-weingarten.de/skripten/statistik/probe.pdf>
(Probeabschnitt)
- <http://www.alex-weingarten.de/skripten/statistik/errata.pdf>
(Errata)

Das Skriptum wird unter der GNU Free Documentation License Version 1.3 (GNU FDL 1.3) veröffentlicht. Der verbindliche englische Text der Lizenz ist unter <http://www.gnu.org/licenses/fdl-1.3.html> zu finden; eine inoffizielle deutsche Übersetzung ist unter <http://www.gnu.de/documents/> verfügbar.

Vorwort

Stochastische Vorgänge – das heißt Vorgänge, bei denen der Zufall einen mehr oder weniger großen Einfluss hat – spielen in Wirtschaft, Wissenschaft und Technik eine erhebliche Rolle. Deshalb wäre es wünschenswert, wenn Kenntnisse der mathematischen Statistik, also der mathematische Beschreibung stochastischer Vorgänge, weit verbreitet wären. Das Gegenteil ist der Fall. In vielen Ausbildungsgängen wird die mathematische Statistik eher als ein lästiges Randgebiet angesehen mit der Folge, dass Ingenieure und Vertreter ähnlicher Berufe sich zwar gut mit Differentialgleichungen auskennen, aber unsicher werden, sobald es um Fragen der mathematischen Statistik geht.

Wir beide – der eine Maschinenbauer, der andere Elektrotechniker – haben unser Berufsleben in einem Hochschulinstitut für Verfahrenstechnik verbracht und dabei in Forschung und Lehre viel mit Statistik zu tun gehabt. Als Ingenieure sehen wir die Statistik unter praktischen Gesichtspunkten, sie ist für uns ein Werkzeug. Die theoretischen Hinter-, Unter- und Abgründe überlassen wir gern den Mathematikern.

Die mathematischen Schwierigkeiten und damit auch die Voraussetzungen zum Verständnis unseres Textes halten sich in Grenzen. Es kommen ein paar Differentialquotienten und Integrale vor, aber ein großer Teil der Rechnungen geht nicht über die Schulmathematik hinaus. Die eigentlichen Hürden sind die statistischen Begriffe und Vorstellungen, die wir deshalb ausführlich und mit vielen Beispielen erläutern, auch vor dem Hintergrund, dass einige der Begriffe im Alltag ungenau und missverständlich benutzt werden.

Die Statistik gehört zu den Grundlagen vieler spezieller Wissensgebiete wie Fehler- und Ausgleichsrechnung, Versuchsplanung und Probenahme, Qualitätskontrolle, Mischungsanalyse, Spieltheorie, Risikoanalyse, Epidemiologie, Ökonometrie, Kryptologie, Meinungsforschung, Signaltheorie, nichtlineare Optimierung, Chaostheorie und weiterer. An einigen Stellen gehen wir auf derartige Themen ein. Unser Text will und kann Monografien zu diesen Fragen jedoch nicht annähernd ersetzen.

Wir beginnen mit einigen Beispielen, die das Verständnis und das Interesse für Fragestellungen der Statistik wecken sollen, sowie der Klärung einiger allgemeiner Begriffe. Es folgt ein ausführliches Kapitel zur Darstellung von Stichprobenergebnissen als Häufigkeitsverteilungen. Erst danach wird der schwierigere Begriff der Wahrscheinlichkeit eingeführt, der benötigt wird, um zum einen Grundgesamtheiten zu beschreiben und zum anderen Zusammenhänge zwischen dem Stichprobenergebnis und der zugehörigen Grundgesamtheit zu formulieren.

In den anschließenden Kapiteln werden einige wichtige diskrete und stetige Verteilungsfunktionen vorgestellt. Deren Anzahl ließe sich beliebig vergrößern. Unsere Auswahl beschränkt sich auf solche Verteilungen, die für die im weiteren Verlauf erläuterten praktischen Anwendungen in der Verfahrenstechnik gebraucht werden.

Bei der Darstellung von Messergebnissen wie auch bei der Veröffentlichung von Umfrageergebnissen sollte die Angabe von Konfidenzintervallen selbstverständlich sein, ist es aber immer noch nicht. Dieses Thema wird deshalb in unserem Text eingehend behandelt, insbesondere was die jeweiligen Voraussetzungen betrifft. Statistische Prüfverfahren (Tests) gehören zu den wichtigsten Anwendungsgebieten der mathematischen Statistik. Es gibt unzählige spezielle Prüfverfahren. Wir beschränken uns auf die Darstellung der prinzipiellen Vorgehensweise an einem leicht verständlichen Beispiel. Das Thema Varianzanalyse haben wir beiseite gelassen, auch aus dem einfachen Grund, dass uns praktische Anwendungen sehr selten beschäftigt haben. Regression und Korrelation werden dagegen in einem eigenen Kapitel behandelt.

Im vorletzten Kapitel haben wir einige wichtigere Anwendungen der mathematischen Statistik in der Mechanischen Verfahrenstechnik, unserem ehemaligen Arbeitsgebiet, zusammengestellt. Den Abschluss bildet ein Kapitel zum Einsatz des Computers bei statistischen Rechnungen, das zum Zeitpunkt der ersten Veröffentlichung des Skriptums im Netz mehr ein Platzhalter für künftige Erweiterungen als eine Informationsquelle ist. Auch die vorangehenden Kapitel dürften in den ersten Jahren ihres Daseins im World Wide Web (WWW) manche Änderung oder Ergänzung erfahren.

Der Text geht auf ein Skriptum zu einer Vorlesung *Statistische Methoden in der Verfahrenstechnik* zurück, die der Erstautor von 1979 bis 2006 in der Universität Karlsruhe (TH) gehalten hat, und vor ihm KARL SOMMER, jetzt Weihenstephan. In der Terminologie und der Wahl der Formelzeichen passen wir uns der deutschsprachigen Wikipedia an, um das Nachschlagen zu erleichtern. Wo es angebracht erscheint, nennen wir auch die englischen Fachausdrücke. Hierbei war uns das Glossar des *International Statistical Institute* (<http://isi.cbs.nl/>) in Den Haag eine Hilfe. Erstmals im Internet (Web) wurde das Skriptum am 17. Februar 2010 in Form einer pdf-Datei veröffentlicht. Die vorliegende Fassung des Skriptums wurde auf einem PC unter Debian GNU/Linux mit Hilfe der Programme `vi`, `gnuplot`, `xfig` und `pdflatex` hergestellt.

Es liegt uns am Herzen, dass Sie unsere Ausführungen verstehen. Scheuen Sie sich nicht, uns per Email zu fragen, wenn wir uns nicht klar genug ausdrücken oder wenn Sie meinen, dass ein wichtiges Thema fehlt. Es ist auch nicht auszuschließen, dass wir uns gelegentlich irren, aber das sollte ein seltenes Ereignis sein.

Karlsruhe, Anfang 2010

Jürgen Raasch

Wulf Alex

Übersicht

1	Grundbegriffe	1
2	Häufigkeit	7
3	Wahrscheinlichkeit	23
4	Diskrete Verteilungen	63
5	Stetige Verteilungen	75
6	Konfidenzintervalle	101
7	Prüfverfahren (Tests)	123
8	Regression und Korrelation	131
9	Anwendungen	147
10	Statistisches Rechnen auf dem Computer	187
A	Zum Weiterlesen	207

Inhalt

1	Grundbegriffe	1
1.1	Kausalität und Zufall	1
1.2	Merkmale und Messverfahren	3
1.3	Mengen	4
1.4	Modelle	5
2	Häufigkeit	7
2.1	Urliste, absolute Häufigkeit	7
2.2	Relative Häufigkeit bei einer Eigenschaft	8
2.3	Relative Häufigkeit bei zwei Eigenschaften	9
2.4	Darstellung relativer Häufigkeiten	13
2.4.1	Häufigkeitsverteilung eines diskreten Merkmals	13
2.4.2	Häufigkeitsverteilung eines stetigen Merkmals	17
3	Wahrscheinlichkeit	23
3.1	Axiome der Wahrscheinlichkeitsrechnung	23
3.2	Folgerungen, Sätze	26
3.3	Wahrscheinlichkeitsverteilungen	32
3.3.1	Verteilungen einer Zufallsvariablen	32
3.3.2	Verteilungen mehrerer Zufallsvariablen	37
3.4	Erwartungswerte und Varianzen	44
3.4.1	Eindimensionale Wahrscheinlichkeitsverteilungen	44
3.4.2	Mehrdimensionale Wahrscheinlichkeitsverteilungen	49
3.4.3	Rechenregeln für Erwartungswerte und Varianzen	52
	Eindimensionale Zufallsgrößen	52
	Mehrdimensionale Zufallsgrößen	58
4	Diskrete Verteilungen	63
4.1	Diskrete Gleichverteilung	63
4.2	Binomialverteilung	63
4.3	Poisson-Verteilung	68

5	Stetige Verteilungen	75
5.1	Stetige Gleichverteilung	75
5.2	Eindimensionale Normalverteilung	75
5.2.1	Gewöhnliche Normalverteilung	75
5.2.2	Logarithmische Normalverteilung	87
5.2.3	Approximation der Normalverteilung durch eine Fourierreihe	89
5.3	Zweidimensionale Normalverteilung	90
5.4	Chi-Quadrat-Verteilung	93
5.5	Student-Verteilung	95
5.6	Potenzverteilung	97
5.7	Exponentielle Verteilungen	98
6	Konfidenzintervalle	101
6.1	Grundbegriffe	101
6.2	K. für den M. einer NV mit bekannter Varianz	108
6.3	K. für die V. einer NV mit bekanntem Mittelwert	113
6.4	K. für die V. einer NV mit unbekanntem Mittelwert	114
6.5	K. für den M. einer NV mit unbekannter Varianz	114
6.6	K. für die Parameter beliebiger Verteilungen	116
6.7	Beispiele für Konfidenzintervalle	116
7	Prüfverfahren (Tests)	123
7.1	Grundbegriffe	123
7.2	Durchführung eines Prüfverfahrens	124
8	Regression und Korrelation	131
8.1	Lineare Regression	131
8.1.1	Berechnung der Regressionsgeraden zu einer Stichprobe . .	131
8.1.2	Konfidenzintervalle für den Regressionskoeffizienten	136
8.1.3	Konfidenzintervalle für den Mittelwert	139
8.2	Nichtlineare Regression	141
8.3	Korrelation	142
8.3.1	Korrelationskoeffizient	142
8.3.2	Konfidenzintervalle für den Korrelationskoeffizienten	145
9	Anwendungen	147
9.1	Partikelgröße und -geschwindigkeit	147
9.1.1	Aufgabenstellung	147
9.1.2	Mathematischer Ansatz	148
9.1.3	Folgerungen	149
9.2	Koinzidenzfehler	150
9.2.1	Aufgabenstellung	150
9.2.2	Mathematischer Ansatz	151
9.2.3	Folgerungen	154

9.3	Partikelgrößenanalyse	154
9.3.1	Aufgabenstellung	154
9.3.2	Mathematischer Ansatz	155
9.3.3	Folgerungen	158
9.4	Porosität	159
9.4.1	Aufgabenstellung	159
9.4.2	Mathematischer Ansatz	160
9.4.3	Folgerungen	162
9.5	Verteilung der Abstände	162
9.5.1	Aufgabenstellung	162
9.5.2	Mathematischer Ansatz	163
9.5.3	Folgerungen	165
9.6	Mischgüte	166
9.6.1	Aufgabenstellung	166
9.6.2	Mathematischer Ansatz	166
9.6.3	Folgerungen	168
9.7	Prüfverfahren Mischtechnik	168
9.7.1	Aufgabenstellung	168
9.7.2	Mathematischer Ansatz	169
9.7.3	Folgerungen	172
9.8	Umrechnung Mengen- und Merkmalsarten	173
9.8.1	Aufgabenstellung	173
9.8.2	Mathematischer Ansatz	174
9.8.3	Folgerungen	177
9.9	Spezifische Oberfläche	178
9.9.1	Aufgabenstellung	178
9.9.2	Mathematischer Ansatz	179
9.9.3	Folgerungen	185
10	Statistisches Rechnen auf dem Computer	187
10.1	Grafische Darstellungen mittels Gnuplot	187
10.2	Tabellenkalkulation (Gnumeric)	194
10.3	Octave, Euler (in Vorbereitung)	199
10.4	Die GNU-R-Umgebung (in Vorbereitung)	199
10.5	Zufallszahlen (in Vorbereitung)	205
10.5.1	Wofür Zufallszahlen?	205
10.5.2	Erzeugung	206
10.5.3	Prüfung auf Zufälligkeit	206
A	Zum Weiterlesen	207