

QUANTITATIVE LINGUISTICS

Volume 55

Editors:

Reinhard Köhler, Burghard Rieger
Gabriel Altmann

Editorial Board:

M. V. Arapov, Moscow
J. Boy, Essen
Sh. Embleton, Toronto
R. Grotjahn, Bochum
R. G. Piotrowski, St. Petersburg
J. Sambor, Warsaw
A. Tanaka, Tokyo
M. Stubbs, Trier

Gabriel Altmann

Statistik für Linguisten

 **Wissenschaftlicher Verlag Trier**

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Altmann, Gabriel :

Statistik für Linguisten / Gabriel Altmann. -

2., verb. Aufl. -

Trier: WVT Wissenschaftlicher Verlag Trier, 1995

(Quantitative linguistics ; Vol. 55)

ISBN 3-88476-176-5

NE: GT

Umschlag: Brigitta Disseldorf

(M. Nottar, Agentur für Werbung und Design, Konz)

© WVT Wissenschaftlicher Verlag Trier, 1995

ISBN 3-88476-176-5

ISSN 0179-3616

Alle Rechte vorbehalten

Nachdruck oder Vervielfältigung nur mit
ausdrücklicher Genehmigung des Verlags

WVT Wissenschaftlicher Verlag Trier

Bergstraße 27, 54295 Trier

Postfach 4005, 54230 Trier

Tel. 0651-41503, Fax 41504

Vorwort

Das vorliegende Buch enthält die notwendigen (nicht aber unbedingt die hinreichenden) Kenntnisse, die dem Verständnis des quantitativen Teils der mathematischen Linguistik behilflich sein können. Der Stoff läßt sich in einem Semester abhandeln, wie bereits von mehreren Dozenten in unterschiedlichen Kursen erprobt wurde, und er ist so gestaltet, daß er sich auch im Selbststudium bewältigen läßt.

Es ist empfehlenswert, so viele Übungen und Aufgaben zu lösen, wie nur möglich. Zahlreiche wurden so gestaltet, daß sie zur Weiterforschung in Texten oder in der Sprache anregen. Alle Lösungen findet man am Ende des Buches.

Bei der ersten Lektüre kann man Kapitel 6 eventuell auslassen und später nachvollziehen, wenn den Lernenden der multivariate Charakter vieler sprachlicher Erscheinungen eingeholt hat.

Die erste Ausgabe dieses Buches enthielt zahlreiche Tippfehler, bedingt durch die damalige Schreibmaschinenherstellung des Manuskriptes. Zu verbindlichem Dank bin ich A. Ambros verpflichtet, der keine Mühe gescheut hat, um all die Fehler zu entdecken. Frau C. Prüin, die mir den Kummer der typographischen Vorbereitung des Manuskriptes abgenommen hat, verdanke ich wahrscheinlich die Vermeidung eines frühzeitigen Herzinfarktes, der Arbeiten dieser Art üblicherweise begleitet. Die verbliebenen Fehler nehme ich getrost auf meine eigene Kappe.

Bochum, Oktober 1995

G. A.

Inhaltsverzeichnis

Einleitung	1
1. Beschreibende Statistik	3
1.1. Ermittlung und Darstellung der Häufigkeiten	3
1.2. Das arithmetische Mittel	14
1.3. Die Varianz und die Standardabweichung	20
Aufgaben	26
2. Kombinatorik	29
2.1. Permutation ohne Wiederholung	29
2.2. Permutation mit Wiederholung	31
2.3. Permutation von Elementen, die nicht alle unterschiedlich sind	34
2.4. Variation ohne Wiederholung	36
2.5. Variation mit Wiederholung	39
2.6. Kombination ohne Wiederholung	40
2.7. Kombination mit Wiederholung	43
2.8. Binomialkoeffizienten	46
2.9. Multinomialkoeffizienten	50
Zusammenfassung der Formeln	55
Zusammenhänge der Formeln	56
Aufgaben	56
3. Wahrscheinlichkeit	61
3.1. Die Definition der Wahrscheinlichkeit	61
3.2. Axiome der Wahrscheinlichkeit	65
3.3. Totale Wahrscheinlichkeit für beliebige Ereignisse	70
3.4. Bedingte Wahrscheinlichkeit	83
3.5. Der Bayessche Satz	89
Aufgaben	91

4. Zufallsvariable	95
4.1. Diskrete Zufallsvariable	95
4.2. Stetige Zufallsvariable	102
4.3. Bedingte Verteilung	107
4.4. Funktionen von Zufallsvariablen	112
Aufgaben	126
5. Mathematische Erwartung und Momente	129
5.1. Die mathematische Erwartung	129
5.2. Momente	133
5.3. Die Tschebyschevsche Ungleichung	141
5.4. Schiefe und Exzeß	143
5.5. Die Taylorsche Formel	149
5.6. Momenterzeugende Funktionen	151
Aufgaben	157
6. Multivariate Verteilungen	161
6.1. Grundbegriffe	161
6.2. Randverteilung	166
6.3. Bedingte Verteilung	170
6.4. Momente mehrdimensionaler Zufallsvariablen	178
6.5. Funktionen zweidimensionaler Zufallsvariablen	198
Aufgaben	210
Anhang	213
Literatur	217
Lösungen der Übungen und Aufgaben	219
Index	240

Einleitung

In jedem Wissensbereich – sei es im Alltag oder in der Wissenschaft – ist Grundvoraussetzung für jede Tätigkeit, eine gedankliche Ordnung über die Phänomene und Gegenstände zu legen, mit denen man zu tun hat. Der einfachste Typ einer solchen Ordnung besteht in der Schaffung von Kategorien: Alle Objekte werden im Hinblick auf bestimmte, interessierende Kriterien als “gleich” bzw. “nicht gleich” betrachtet; anders ausgedrückt: sie werden bestimmten Klassen zugeordnet. Dies ist der grösste Ordnungstyp, er wird “qualitativ” genannt. Auf diesem Niveau ist nicht mehr möglich, als für Objekte anzugeben, ob sie zu einer bestimmten Klasse (z. B. zur Klasse der Verben oder der Flexionsendungen) gehören oder nicht. Dagegen erlauben es “quantitative” Ordnungstypen, beliebig feine Ordnungen zu schaffen, indem die Objekte im Hinblick auf die interessierenden Kriterien in eine Rangreihenfolge gebracht werden können (Ordinalskala) oder es sogar ermöglicht wird, die Unterschiede numerisch (Intervall- und Verhältnisskalen) zu erfassen und mit ihnen zu rechnen.

Je nach Anforderungen und Verwendungszweck werden entsprechend quantitative bzw. qualitative Begriffe und Verfahren bevorzugt. In einem Fall kann es genügen, z. B. Prüfungsaufgaben lediglich in schwierige und leichte zu unterscheiden, in anderen Fällen wird es notwendig sein, den Schwierigkeitsgrad auf einer Punkteskala anzugeben und damit von einem qualitativen Begriff “schwierig” zu einem quantitativen überzugehen.

In der Geschichte der Sprachwissenschaft hat man sich die meiste Zeit mit qualitativen Begriffen und Verfahren begnügt. Mathematisierung bedeutete vor allem Formalisierung dieser Begriffe mit Hilfe von Algebra, Logik etc.

Seit einigen Jahrzehnten aber nehmen quantitative Ansätze zunehmenden Raum ein – z. T. einfach, weil die entsprechende Ausbildung in der Sprachwissenschaft im Gegensatz zur heutigen Zeit früher vollkommen fehlte, z. T. aber auch, weil die Einsicht zunimmt, daß qualitative, deterministische Modelle für Sprache und Text nur sehr begrenzten Wert besitzen.

Sowohl bei der diachronischen als auch bei der synchronischen Betrachtung sprachlicher Gegebenheiten stoßen wir immer wieder auf Uneindeutigkeiten, Unschärfe, Ausnahmen von Regeln, auf Fluktuationen etc., kurz: auf deutlich spürbare Auswirkungen des stochastischen Charakters der Sprache. Hier kann nur ein Instrumentarium greifen, das es uns erlaubt, auch Tendenzen zu beschreiben, weil in den allermeisten Fällen keine absoluten Regelmäßigkeiten gefunden werden können.

Die Statistik bietet uns ein solches Instrumentarium an, indem sie uns erlaubt, Schlüsse zu ziehen, deren Zuverlässigkeit festgelegt ist und somit nachgeprüft werden kann. Dies ist auch dann möglich, wenn die zugrundegelegte Information – wie im Fall der Sprache – unvollständig bleiben muß, da die schiere Menge schriftlicher und mündlicher Sprachäußerungen weder erfaßbar noch in angemessener Zeit bearbeitbar ist und sich dabei noch ständig vergrößert.

Mit statistischen Methoden können jedoch gerade Massenerscheinungen in idealer Weise untersucht werden, da im Grunde beliebig viele Daten (wie man später sehen wird, sogar unendlich viele) mit sehr wenigen Parametern beschreibbar sind. Stellen wir Regularitäten fest, die eine Systematisierung des gewonnenen Datenmaterials zulassen, können wir induktiv Erkenntnisse über die Sprache gewinnen.

Die Statistik wird also in der Sprachwissenschaft – nicht anders als in anderen Wissenschaften – benötigt, um Sachverhalte zu beschreiben, die nichtdeterministischer Natur sind oder (praktisch oder prinzipiell) nicht vollständig erfaßbar sind (wie die Lexik einer Sprache, die sich schneller ändert, als man sie beschreiben kann).

Eine unverzichtbare Funktion kommt der Statistik in dem wissenschaftlichen Kernbereich der Hypothesenüberprüfung zu. Wissenschaftlich interessante Hypothesen können aufgrund der oben skizzierten Eigenschaften von Sprache und Text durch einzelne Beispiele weder bestätigt noch widerlegt werden. Auch hier benötigen wir statistische Methoden, um eine Hypothese anhand von Daten objektiv beurteilen zu können.

Die folgenden Kapitel geben eine Einführung in die Grundlagen der deskriptiven und der analytischen Statistik, soweit sie für die Sprachwissenschaft von Belang sind. Die Beispiele und Übungsaufgaben sollen, soweit möglich, gleichzeitig einen Eindruck von der Bedeutung der behandelten Fragen für die Praxis vermitteln. Es ist aber nicht vermeidbar, daß der Sinn mancher Formeln und Zusammenhänge erst später wirklich klar wird.

1. Beschreibende Statistik

1.1. Ermittlung und Darstellung der Häufigkeiten

Die statistische Verarbeitung von Daten erfolgt in zwei Stufen: der beschreibenden und der schließenden.

Die beschreibende, mit der wir uns in diesem Kapitel beschäftigen werden, hat die Aufgabe, in heterogenen Daten, in denen man keine klaren deterministischen Gesetze entdecken kann, eine Ordnung zu schaffen und dieses geordnete, übersichtlich dargestellte Gebilde von Daten durch sogenannte Maßzahlen oder Kenngrößen zu charakterisieren.

Das zu untersuchende linguistische oder literarische Material steht äußerst selten vollständig zur Verfügung bzw. ist sehr schwer zu ermitteln. Deswegen erhebt man lediglich *Stichproben* kleineren oder größeren Umfangs aus einer *Grundgesamtheit*, unter der man alle Texte einer Sprache, alle Texte eines literarischen Genres, das ganze Werk eines Schriftstellers, die vollständige Lexik einer Sprache usw. versteht. Man greift durch Zufall aus der Grundgesamtheit einige Elemente heraus und mißt an ihnen eine oder mehrere Eigenschaften. Aus der Gestaltung der Stichprobe zieht man dann Schlüsse über die Grundgesamtheit, ermittelt (stochastische) Tendenzen, latente Mechanismen und Zusammenhänge. Das ist die Aufgabe der Inferenz- oder schließenden Statistik.

Die Eigenschaft (Merkmal), die man untersucht, ist nicht konstant – sonst wäre die ganze Statistik nicht notwendig –, sondern zeigt an verschiedenen Elementen der Stichprobe unterschiedliche Ausprägungen. Die Faktoren, die die Veränderung der Ausprägung von Element zu Element bestimmen, sind uns unbekannt, deswegen betrachten wir die untersuchte Größe als eine *zufällige Veränderliche*, bezeichnen sie allgemein mit X , und die Werte, die sie annimmt, als x . Solche Veränderlichen (*Variable*) sind beispielsweise die Länge des Wortes im Deutschen, die Position des Akzentes in Sprachen mit freiem Akzent, die Zahl der Phoneme (Buchstaben), die hinter einem gegebenen Phonem (Buchstaben) stehen können, die Zahl der unbetonten Silben zwischen zwei betonten Silben

in Goethes Erlkönig, die phonische Ähnlichkeit zweier Verse, die Satzlänge in einem Text, die Anzahl von Adverbialphrasen in einem Satz, die Zahl von Bedeutungen eines Lemmas in einem Wörterbuch usw.

Nimmt die Veränderliche nur Werte an, die man mit Hilfe der ganzen Zahlen bezeichnen kann, so nennen wir sie *diskret*, sonst ist sie *stetig (kontinuierlich)*. Man sagt auch, daß man die Ausprägungen einer diskreten Größe auf die Menge der ganzen Zahlen abbilden kann, während man die Ausprägungen einer stetigen Größe auf die Menge der reellen Zahlen abbilden kann.

Um die Ausprägungen, d.h. die Werte einer Eigenschaft zu ermitteln, muß man eine Maßeinheit einführen, z.B. die Phonemzahl als die Maßeinheit der Wortlänge, die Millisekunde als die Maßeinheit der Lautdauer u.a. Das Aneinanderfügen von Maßeinheiten ergibt eine Meßskala. Die Ermittlung der Ausprägung erfolgt durch die Anwendung einer Meßvorschrift. Bei diskreten Zufallsgrößen geschieht dies durch Zählen, bei stetigen Größen durch *Messen* unter Anwendung eines konventionell vereinbarten Instrumentes (Uhr, Metermaß, Barometer u.a.). Bei einer diskreten Größe verfertigt man eine *Strichliste*, in die man die ermittelten Werte einzeln einträgt. Sie besteht aus einer Tabelle, in der in der ersten Spalte die natürlichen Zahlen (oft einschließlich der Null) stehen, in der zweiten Spalte für jeden ermittelten Wert ein Strich, in der dritten Spalte die Anzahl der Striche, d.h. die *absolute Häufigkeit* der Werte. Eine Tabelle dieser Art nennt man *Häufigkeitstabelle*.

Beispiel 1.1.1. Gemessene Eigenschaft (X): Wortlänge in Gretchens Monolog aus dem ersten Teil von Goethes Faust ("Meine Ruh ist hin, ..."). Maßeinheit: Anzahl der Buchstaben im geschriebenen Wort. Die Resultate der Ermittlung der Wortlängen sind in Tabelle 1.1 angeführt. ■

Bei einer stetigen Größe kann man natürlich nicht alle Werte in einer Strichliste aufschreiben, da ihre Zahl im Grunde unendlich ist und nur der Verzicht auf größere Genauigkeit sie etwas einschränken kann. Deswegen schreibt man jeden gemessenen Wert zuerst hintereinander in einer *Urliste* auf. Hat man zu viele x_i -Werte, so faßt man sie in *Klassen* zusammen, um Übersicht zu gewinnen. Man verliert dadurch etwas Information, aber die Daten sind dann leichter zu handhaben. Die Anzahl der Klassen soll zwischen 5 und 20 liegen. Jede Klasse enthält mehrere nebeneinanderliegende Werte der Veränderlichen, d.h. ein ganzes Intervall von Realisationen der Veränderlichen. Man gruppiert die Werte folgendermaßen in Klassen. Man bestimmt, wie viele Klassen (n) man haben möchte; dann findet man den größten realisierten Wert der Variablen x_{max} , den kleinsten realisierten Wert x_{min} und berechnet die *Klassenbreite* nach der Formel

Tabelle 1.1. Häufigkeitstabelle für Wortlänge, gemessen in Buchstaben, für einen Text aus Goethes Faust

Wortlänge	Strichliste	Häufigkeit der Wörter der Länge x_i
x_i		f_i
1	-	0
2		3
3		57
4		26
5		17
6		18
7		3
8		3
9		1
10		4
11		1

$$Klassenbreite = \frac{x_{\max} - x_{\min}}{n} \quad (1.1)$$

und gruppiert die betreffenden Werte x_i in eine Klasse. Die Klassengrenzen dürfen sich nicht überschneiden. Oft führt man auch bei diskreten Variablen eine Gruppierung durch, wenn die Variable sehr viele Werte annehmen kann.

Beispiel 1.1.2. Geršić (1971) hat die Dauer von kurzen unbetonten Vokalen in Batscha-Deutsch in Millisekunden an 3 Informanten gemessen. Es ergaben sich Werte etwa von 20 bis 360 Millisekunden (einige wenige Realisierungen, die unter 20 oder über 360 fielen, wurden in die Nachbarklassen eingeschlossen). Der Verfasser gruppierte die ganze Variationsbreite in 17 Klassen und bekam die Klassenbreite

$$\frac{360 - 20}{17} = \frac{340}{17} = 20.$$

Nach der Gruppierung ergab sich die Häufigkeitsverteilung in Tabelle 1.2, wo in jeder Klasse 20 Werte eingeschlossen sind. Die Intervalle 20-40, 40-60 usw. sind so zu verstehen, daß sie von links geschlossen, von rechts offen sind, d.h. $[20,40)$, $[40,60)$ usw. Das bedeutet, daß Werte kleiner als 40 in das erste Intervall gehören, Werte gleich oder größer als 40 und kleiner als 60 in das zweite usw. Es wäre geeigneter, die erste Klasse mit " < 40 " und die letzte mit " ≥ 340 " zu bezeichnen. ■

Tabelle 1.2.
Verteilung der Lautdauer der kurzen unbetonten Vokale im Batschka-Deutsch nach Geršić (1971)

Intervall der Lautdauer x_i	Häufigkeit f_i
20 - 40	468
40 - 60	958
60 - 80	840
80 - 100	601
100 - 120	310
120 - 140	156
140 - 160	70
160 - 180	41
180 - 200	30
200 - 220	17
220 - 240	6
240 - 260	10
260 - 280	3
280 - 300	2
300 - 320	1
320 - 340	2
340 - 360	1
Summe	3516

Tabelle 1.3.
Verteilung der Phonemzahl in 600 Sprachen nach Lehfeldt (1974)

Anzahl der Phoneme x_i	Anzahl der Sprachen f_i
11,5 - 16,5	15
16,5 - 21,5	98
21,5 - 26,5	124
26,5 - 31,5	118
31,5 - 36,5	101
36,5 - 41,5	72
41,5 - 46,5	33
46,5 - 51,5	23
51,5 - 56,5	9
56,5 - 61,5	7
Summe	600

Beispiel 1.1.3. Lehfeldt (1974) hat in 600 Sprachen die Anzahl der Phoneme im Inventar untersucht. Die Zahl der Phoneme variierte von 13 bis 61, und nach der Gruppierung ergab sich die Häufigkeitsverteilung, die in Tabelle 1.3 angeführt ist. ■

Da man später bei der Verarbeitung der Daten oft nur mit den Klassenmitten arbeitet, soll man die Intervalle so wählen, daß die Klassenmitten möglichst runde Zahlen sind, also keine Zahlen mit vielen Dezimalstellen, die die Rechnungen beträchtlich verlängern.

Man sieht an den Beispielen, daß

$$\sum_{i=1}^n f_i = N \quad (1.2)$$

N stellt die Summe aller Häufigkeiten, den *Stichprobenumfang*, dar, und n ist hier die Anzahl der Klassen, über die summiert wurde. Weiter bezeichnen wir

$$\hat{p}_i = \hat{p}(x_i) = \frac{f_i}{N} \quad (1.3)$$

als die *relative Häufigkeit* der Klasse x_i in der Stichprobe und

$$\hat{F}(x) = \sum_{x_i \leq x} \hat{p}(x_i) \quad (1.4)$$

als die relative Summenhäufigkeit, wobei man bis zu einer gegebenen Klasse x einschließlich summiert. Offensichtlich gilt

$$\hat{F}(x_{\max}) = \sum_{x_i \leq x_{\max}} \hat{p}(x_i) = 1 \quad (1.5)$$

Die Funktion $\hat{F}(x)$ heißt auch die *Verteilungsfunktion* der Stichprobe. Die Begriffe “relative Häufigkeit” und “Verteilungsfunktion” spielen eine große Rolle bei der Theorie der Zufallsvariablen (Kap. 4). Mit ihrer Einführung an konkreten Daten soll auf das später notwendig werdende abstrakte Verständnis vorbereitet werden.

Beispiel 1.1.4. In den Daten des Beispiels 1.1.2 liegen die Klassenmitten in den Punkten 30, 50, 70, Man berechnet sie als den Durchschnitt der unteren und der oberen Klassengrenzen. Die Werte von $\hat{p}_i$ und $\hat{F}(x)$ berechnet man aus den Formeln (1.3) und (1.4), z.B.

für $x_i = 30$

$$\text{ist } \hat{p}(x_i) = \hat{p}(30) = \frac{468}{3516} = 0,1331;$$

für $x_i = 50$

$$\text{ist } \hat{p}(50) = \frac{958}{3516} = 0,2725 \text{ usw.}$$

Die relative Summenhäufigkeit

für $x_i = 30$

$$\text{ist } \hat{F}(30) = 0,1331;$$

für $x_i = 50$

$$\text{ist } \hat{F}(50) = \hat{p}(30) + \hat{p}(50) = 0,1331 + 0,2725 = 0,4056;$$

Tabelle 1.4. Häufigkeitsverteilung der Lautdauer der kurzen unbetonten Vokale im Batschka-Deutsch nach Geršić (1971)

Intervall der Lautdauer	Klassenmitte x_i	Häufigkeit f_i	Relative Häufigkeit $\hat{p}_i$	Relative Summen-Häufigkeit $\hat{F}(X)$
20- 40	30	468	0,1331	0,1331
40- 60	50	958	0,2725	0,4056
60- 80	70	840	0,2389	0,6445
80-100	90	601	0,1709	0,8154
100-120	110	310	0,0882	0,9036
120-140	130	156	0,0444	0,9480
140-160	150	70	0,0199	0,9679
160-180	170	41	0,0117	0,9796
180-200	190	30	0,0085	0,9881
200-220	210	17	0,0048	0,9929
220-240	230	6	0,0017	0,9946
240-260	250	10	0,0028	0,9974
260-280	270	3	0,0009	0,9983
280-300	290	2	0,0006	0,9989
300-320	310	1	0,0003	0,9992
320-340	330	2	0,0006	0,9998
340-360	350	1	0,0003	1,0001
Summe		3516	1,0001	

für $x_i = 70$

ist $\hat{F}(70) = \hat{p}(30) + \hat{p}(50) + \hat{p}(70)$

$= 0,1331 + 0,2725 + 0,2389 = 0,6445$.

Alle Werte sind in Tabelle 1.4 angegeben. ■

Bemerkung: Die Tatsache, daß in der Tabelle $\Sigma \hat{p}_i = 1,0001$ ist, hängt mit der Rundung zusammen. Die *Rundung* erfolgt folgendermaßen: Von 1 bis 4 rundet man nach unten ab, von 6 bis 9 nach oben. Bei 5 rundet man nach unten ab, wenn die letzte Stelle vor der zu rundenden 5 gerade ist, und nach oben, wenn sie ungerade ist, z.B.

23,4 ~ 23

23,5 ~ 24

23,51 ~ 24

23,6 ~ 24

28,4 ~ 28

28,5 ~ 28

28,51~ 29

28,6 ~ 29. ■

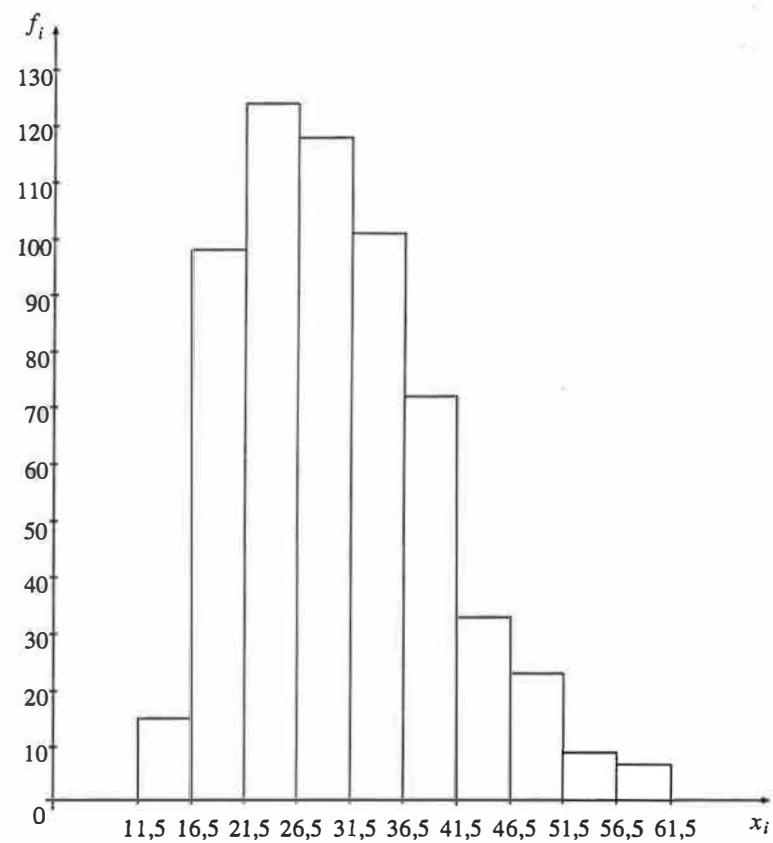


Abb. 1.1. Histogramm für die Häufigkeitsverteilung der Phonemzahl in 600 Sprachen nach Lehfeldt (1974)

Um die Häufigkeitsverteilung der Stichprobe anschaulich zu machen, pflegt man die Daten graphisch darzustellen. In einem Koordinatensystem trägt man auf der Abszisse die x -Werte und auf der Ordinate entweder die absoluten (f_i) oder die relativen ($\hat{p}_i$) Häufigkeiten auf. Um die Intervallmitte (Klassenmitte) zeichnet man ein Rechteck, dessen Breite der Breite des Intervalls und dessen Höhe der Häufigkeit entspricht. Eine solche Darstellung heißt *Häufigkeitshistogramm*.

Beispiel 1.1.5. In Abb 1.1 ist das Histogramm für die Daten von Lehfeldt (1974) aus dem Beispiel 1.1.3 gezeichnet. Auf der Ordinate sind die absoluten Häufigkeiten gekennzeichnet. ■

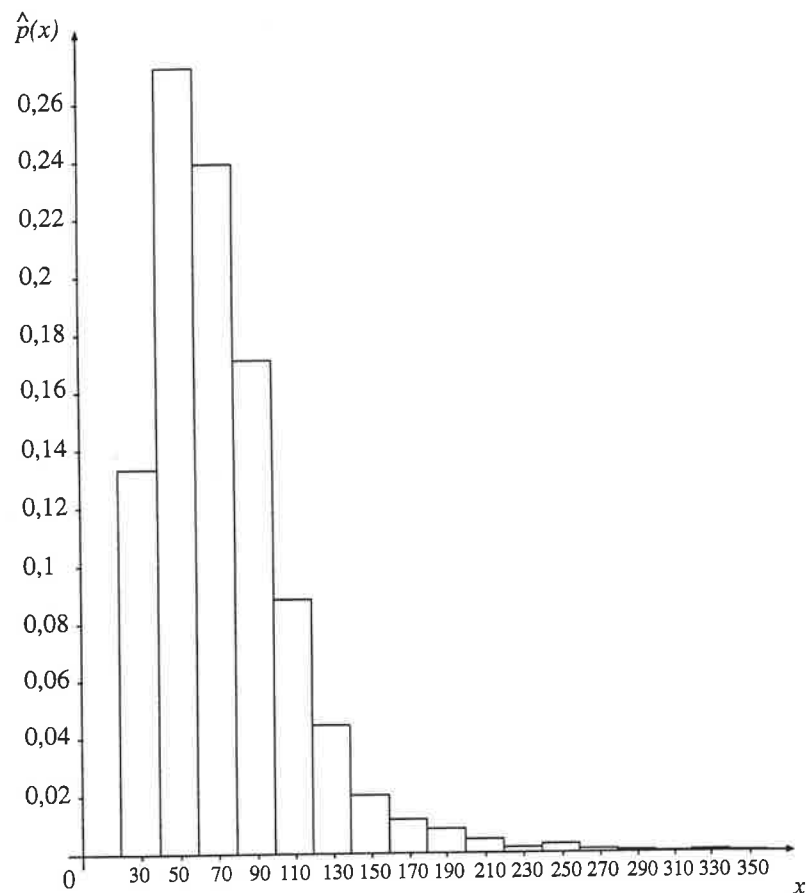


Abb. 1.2. Histogramm für die Häufigkeitsverteilung der Lautdauer nach Geršić (1971)

Beispiel 1.1.6. In Abb. 1.2 ist das Histogramm für die Daten von Geršić (1971) aus dem Beispiel 1.1.4 gezeichnet. Auf der Ordinate sind die relativen Häufigkeiten aufgetragen. ■

Nimmt die Variable nur diskrete Werte an, die man nicht gruppiert, so pflegt man die Häufigkeiten in Form von Stäbchen darzustellen. Solche Darstellungen nennt man *Stabdiagramme*.

Tabelle 1.5. Häufigkeitsverteilung der Silbenlänge indonesischer Morphemtypen.

Silbenlänge des Morphems x_i	Häufigkeit f_i	Relative Häufigkeit $\hat{p}_i$	Relative Summen-Häufigkeit $\hat{F}(x)$
1	12	0,08108	0,08108
2	35	0,23649	0,31757
3	54	0,36486	0,68243
4	39	0,26351	0,94594
5	8	0,05405	0,99999
Summe	148	0,99999	

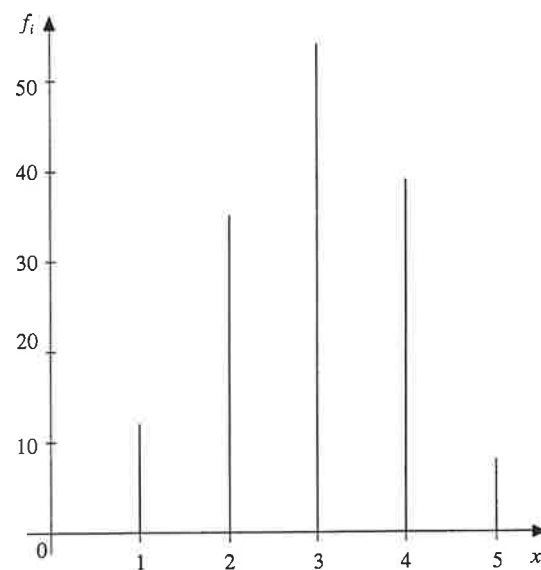


Abb. 1.3. Stabdiagramm der Häufigkeitsverteilung im Beispiel 1.1.7.

Beispiel 1.1.7. Die 148 strukturellen Morphemtypen des Indonesischen können 1 bis 5 Silben lang sein (Silbenlänge = Vokalzahl). Ihre Häufigkeiten sind in Tabelle 1.5 angegeben und das Stabdiagramm dieser Häufigkeitsverteilung in Abb. 1.3 dargestellt. ■

Verbindet man die Punkte $(x_i, \hat{p}(x_i))$ mit einer Geraden, d.h. zieht man eine Gerade von einer Klassenmitte in der Höhe $\hat{p}(x_i)$ zu der benachbarten Klassenmitte in der Höhe $\hat{p}(x_{i+1})$, so bekommt man ein *Häufigkeitspolygon*. Im Stabdiagramm verbindet man einfach die oberen Enden der Stäbe.

Beispiel 1.1.8. Die Häufigkeitspolygone der Verteilungen aus den Abbildungen 1.2 und 1.3 sind in den Abbildungen 1.4 und 1.5 dargestellt. ■

Je kleiner die Intervalle einer stetigen Variablen, desto glatter wird das Polygon.

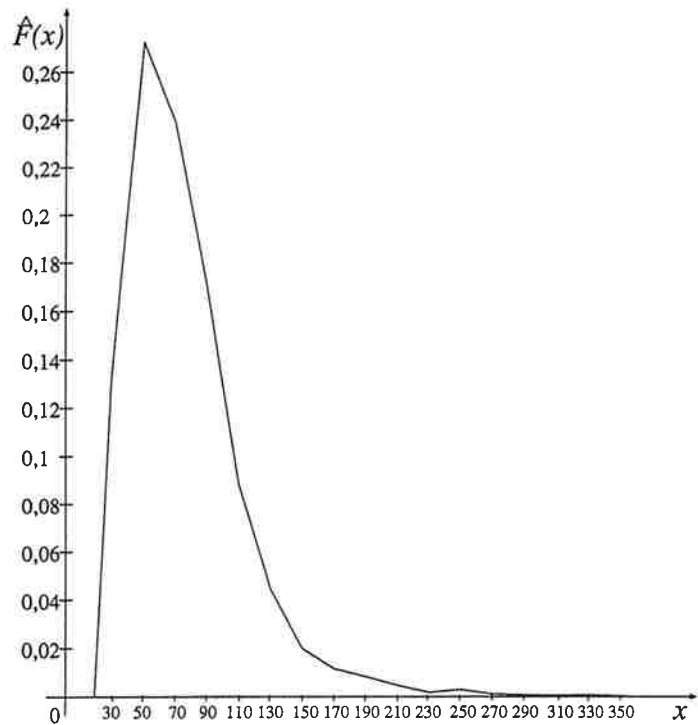


Abb. 1.4. Häufigkeitspolygon der Verteilung in Abb. 1.2.

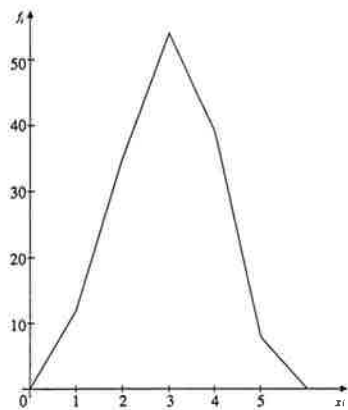


Abb. 1.5. Häufigkeitspolygon der Verteilung in Abb. 1.3.

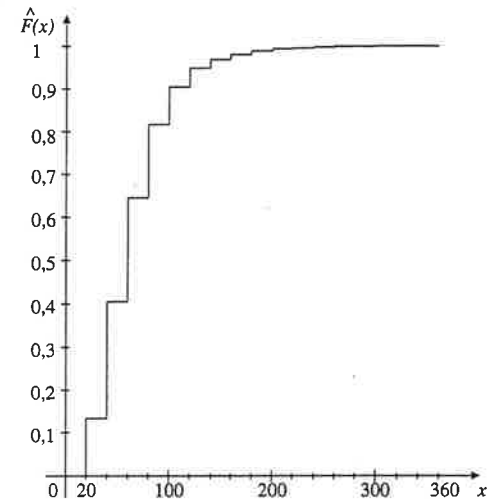


Abb. 1.6. Treppenkurve für das Beispiel 1.1.4

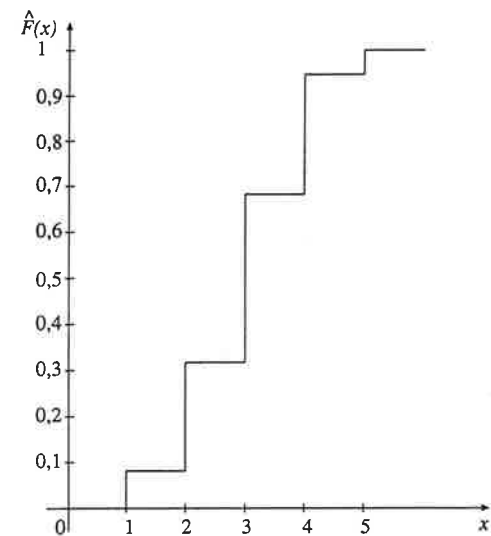


Abb. 1.7. Treppenkurve für das Beispiel 1.1.7

Auch die Summenhäufigkeiten lassen sich graphisch darstellen. Durch ihre Eintragung in ein Koordinatensystem entsteht die sogenannte *Treppenkurve*. Für die Verteilungen in den Beispielen 1.1.4 und 1.1.7 sind die Treppenkurven in den Abbildungen 1.6 und 1.7 dargestellt.

Übung 1.1.1. Die folgende Urliste enthält die Wortlänge, gemessen als die Zahl der Buchstaben, aus der ersten Strophe der "Zueignung" von Goethes Faust: 3, 4, 4, 6, 11, 9, 3, 4, 4, 5, 3, 6, 7, 7, 3, 4, 4, 7, 12, 4, 3, 4, 4, 4, 5, 4, 7, 3, 6, 4, 2, 3, 3, 2, 4, 3, 6, 3, 3, 3, 5, 3, 5, 2, 4, 6, 4, 5, 5, 4, 10, 11, 3, 11, 3, 5, 3, 9, 5.

Stellen Sie die Strichliste und die Häufigkeitsverteilung der Wortlänge auf.

Übung 1.1.2. Zeichnen Sie das Stabdiagramm der Verteilung in der Übung 1.1.1.

Übung 1.1.3. Berechnen Sie die relativen Häufigkeiten und die relativen Summenhäufigkeiten aus dem Beispiel 1.1.1.

Übung 1.1.4. Zeichnen Sie die Treppenkurve für die Verteilung in Beispiel 1.1.3.

1.2. Das arithmetische Mittel

Die meistbenutzte Kenngröße oder Charakteristik einer empirischen Verteilung ist das *arithmetische Mittel*, oft einfach *Mittelwert* oder *Durchschnitt* genannt. Man berechnet es so, daß man die ermittelten Werte der Zufallsvariablen addiert und die Summe durch die Anzahl der Summanden dividiert. Seien $x_1, x_2, x_3, \dots, x_N$ die gemessenen Werte von X , dann definieren wir das arithmetische Mittel als

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i \quad (1.6)$$

Es ist üblich, daß bei Messungen und besonders bei Zählungen derselbe Meßwert x_i nicht nur einmal, sondern öfter vorkommt. Man müßte dann den x_i -Wert so oft in die Formel (1.6) einsetzen, wie man ihn gefunden hat. Man kann die Prozedur verkürzen, indem man nicht

$$x_{i1} + x_{i2} + x_{i3} + \dots + x_{if}$$

d.h. f -mal x_i schreibt, sondern einfach $x_i f_i$, man gruppiert also gleiche Werte. So bekommt man ganz allgemein aus (1.6)

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_n f_n}{N} = \frac{1}{N} \sum_{i=1}^n x_i f_i \quad (1.7)$$

wobei $N = \sum_{i=1}^n f_i$ ist.

Tabelle 1.6. Berechnung des Mittelwertes

Länge	Häufigkeit	
x_i	f_i	$x_i f_i$
2	3	6
3	57	171
4	26	104
5	17	85
6	18	108
7	3	21
8	3	24
9	1	9
10	4	40
11	1	11
Σ	133	579

Beispiel 1.2.1. In Beispiel 1.1.1 wurde die Wortlänge in Buchstaben in einem Ausschnitt aus Goethes Faust gemessen. Wir berechnen die mittlere Wortlänge.

Lösung: Man multipliziert jeden x_i -Wert mit dem dazugehörigen f_i -Wert, um $x_i f_i$ zu bekommen, wie in Tabelle 1.6 angedeutet wird. Die Summen der f_i - und der $x_i f_i$ -Spalte setzt man in die Formel (1.7) ein. Hier ist

$$N = \sum_i f_i = 133, \quad \sum_i x_i f_i = 579, \quad \text{daher}$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n x_i f_i = \frac{579}{133} = 4,3534. \blacksquare$$

Sind die Werte von X relativ groß, so kompliziert sich jede weitere Bearbeitung der Daten. Steht kein Rechner zur Verfügung oder übersteigen die Werte (oder ihre Summe) die größte darstellbare Zahl eines Rechners, so kann man den Mittelwert etwas leichter berechnen, wenn man die ursprüngliche Zufallsvariable X transformiert. Wir definieren zunächst eine neue Variable

$$U = X - A, \quad (1.8)$$

wobei A eine beliebige Konstante ist, d.h. von jedem x_i ziehen wir die Zahl A ab. Der Durchschnitt von U ergibt sich dann als

$$\bar{u} = \frac{1}{N} \sum_{i=1}^N u_i = \frac{1}{N} \sum_{i=1}^N (x_i - A) = \frac{1}{N} \sum_{i=1}^N x_i - \frac{1}{N} N A = \bar{x} - A. \quad (1.9)$$

Aus diesem Resultat folgt, daß der wahre Durchschnitt

$$\bar{x} = \bar{u} + A \quad (1.10)$$

ist. Den gesuchten Mittelwert $\bar{x}$ erhält man also, indem man die abgezogene Konstante zu $\bar{u}$ addiert.

Bemerkung: $\sum_{i=1}^n x_i + c \neq \sum_{i=1}^n (x_i + c) = \sum_{i=1}^n x_i + nc,$

wenn $n \geq 2$, was man an einem konkreten Beispiel leicht nachprüfen kann; vgl. Übung 1.2.1.

Beispiel 1.2.2. In der Verteilung der Tabelle 1.4 ziehe man von jedem x_i -Wert die Zahl $A = 200$ ab und berechne den Mittelwert.

Lösung: Die notwendigen Zahlen sind in Tabelle 1.7 angegeben. Aus der Berechnung folgt, daß

$$\bar{u} = \frac{-440660}{3516} = -125,3299$$

woraus nach Rücktransformation

$$\bar{x} = \bar{u} + A = -125,33 + 200 = 74,67 \text{ wird. } \blacksquare$$

Tabelle 1.7. Berechnung des Mittelwertes mit der transformierten Variablen $u_i = x_i - 200$ (vgl. Beispiel 1.2.2)

Klassenmitten x_i	$u_i = x_i - 200$	f_i	$u_i f_i$
30	-170	468	-79560
50	-150	958	-143700
70	-130	840	-109200
90	-110	601	-66110
110	-90	310	-27900
130	-70	156	-10920
150	-50	70	-3500
170	-30	41	-1230
190	-10	30	-300
210	10	17	170
230	30	6	180
250	50	10	500
270	70	3	210
290	90	2	180
310	110	1	110
330	130	2	260
350	150	1	150
	Σ	3516	-440660

Die Rechenarbeit kann noch weiter reduziert werden, wenn man eine andere Transformation durchführt, nämlich

$$U = \frac{X - A}{d}, \quad (1.11)$$

wobei sowohl A als auch d beliebige Zahlen sind. Man wählt d am besten als den Unterschied zwischen x_i und x_{i+1} , d.h. als den Unterschied zweier benachbarter x_i -Werte. Man berechnet dann den Durchschnitt $\bar{u}$. Da

$$\bar{u} = \frac{1}{N} \sum_{i=1}^N u_i = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - A}{d} \right) =$$

$$= \frac{1}{Nd} \sum_{i=1}^N x_i - \frac{1}{N} \cdot N \cdot \frac{A}{d}$$

$$= \frac{\bar{x}}{d} - \frac{A}{d} = \frac{\bar{x} - A}{d}, \quad (1.12)$$

folgt daraus, daß

$$\bar{x} = d\bar{u} + A = \frac{d}{N} \sum_{i=1}^N u_i + A. \quad (1.13)$$

Beispiel 1.2.3. Berechnen Sie aus den Daten des Beispiels 1.2.2 den Mittelwert mit Hilfe der Transformation

$$u_i = \frac{x_i - 110}{20}.$$

Lösung: Wir wählen $A = 110$ und $d = 20$ aus rein rechnerischen Gründen. Die Berechnung ist in Tabelle 1.8 angegeben. Aus den Resultaten ergibt sich

$$\bar{u} = \frac{-6211}{3516} = -1,7665,$$

woraus nach (1.13) schließlich folgt:

Tabelle 1.8. Berechnung des Mittelwertes mit der transformierten Variablen $u_i = (x_i - 110)/20$ (vgl. Beispiel 1.2.3)

Klassenmitten	$u_i = \frac{x_i - 110}{20}$	f_i	$u_i f_i$
30	-4	468	-1872
50	-3	958	-2874
70	-2	840	-1680
90	-1	601	-601
110	0	310	0
130	1	156	156
150	2	70	140
170	3	41	123
190	4	30	120
210	5	17	85
230	6	6	36
250	7	10	70
270	8	3	24
290	9	2	18
310	10	1	10
330	11	2	22
350	12	1	12
	Σ	3516	-6211

$$\bar{x} = d\bar{u} + A = 20(-1,7665) + 110 = 74,67$$

wie oben. ■

Die Formel (1.13) gilt natürlich auch dann, wenn wir mit f_i arbeiten. Setzt man die Transformation $u_i = (x_i - A)/d$, dann ist

$$\begin{aligned} \bar{u} &= \frac{1}{N} \sum_{i=1}^n u_i f_i \\ &= \frac{1}{N} \sum_{i=1}^n \left(\frac{x_i - A}{d} \right) f_i \\ &= \frac{1}{Nd} \sum_{i=1}^n x_i f_i - \frac{1}{Nd} \sum_{i=1}^n A f_i = \end{aligned}$$

$$\begin{aligned} &= \frac{\bar{x}}{d} - \frac{A}{Nd} \sum_{i=1}^n f_i \\ &= \frac{\bar{x}}{d} - \frac{A}{Nd} \cdot N = \frac{\bar{x} - A}{d} \end{aligned}$$

wie in (1.13).

Der Mittelwert besitzt zwei wichtige Eigenschaften, die eine bedeutende Rolle spielen:

(1) Die Summe der Abweichungen der x_i -Werte von $\bar{x}$ ist gleich Null, d.h.

$$\sum_i (x_i - \bar{x}) = 0.$$

Beweis:
$$\sum_i (x_i - \bar{x}) = \sum_i x_i - N\bar{x} = \sum_i x_i - N \frac{1}{N} \sum_i x_i = 0,$$

da $\bar{x} = \frac{1}{N} \sum_i x_i$ aus der Definition (1.6) zu ersetzen ist. ■

(2) Die Summe der Abweichungsquadrate der x_i -Werte von einer Konstante c ist am kleinsten, wenn $c = \bar{x}$, d.h. $\sum_i (x_i - c)^2 = \text{Minimum}$, wenn $c = \bar{x}$.

Beweis: Wir setzen die erste Ableitung der Summe der Abweichungsquadrate nach c gleich 0,

$$-2 \sum_i (x_i - c) = 0.$$

Daraus folgt

$$\sum_i x_i = Nc$$

$$c = \frac{1}{N} \sum_i x_i = \bar{x}.$$

nach Definition von $\bar{x}$. ■

Übung 1.2.1. Berechnen Sie

$$\begin{array}{ll} \text{(a) } \sum_{k=1}^3 (k+1); & \text{(d) } \sum_{x=1}^3 x^2; \\ \text{(b) } \sum_{k=1}^3 k+1; & \text{(e) } \sum_{i=1}^n (x_i - k); \\ \text{(c) } \sum_{x=1}^3 (x-2); & \text{(f) } \sum_{i=1}^5 c. \end{array}$$

Übung 1.2.2. (a) Zeigen Sie, daß

$$\sum_{i=1}^n ax_i = a \sum_{i=1}^n x_i$$

(b) Zeigen Sie, daß

$$\sum_{i=1}^n (ax_i - by_i) = a \sum_{i=1}^n x_i - b \sum_{i=1}^n y_i;$$

(c) Berechnen Sie $\sum_{k=1}^4 ak$.

Übung 1.2.3. Berechnen Sie den Mittelwert der Verteilung in der Übung 1.1.1 (Wortlänge in der Zueignung aus Goethes Faust).

Übung 1.2.4. Berechnen Sie den Mittelwert der Verteilung im Beispiel 1.1.3 (Daten von Lehfeldt).

Übung 1.2.5. Berechnen Sie den Mittelwert der Meßdaten 26,7 31,5 22,6 28,2 29,2 25,3 25,7 mit Abziehen einer Konstante nach (1.8).

Übung 1.2.6. In zwei Stichproben mit $N_1 = 30$ bzw. $N_2 = 40$ hat man $\bar{x}_1 = 4,3$ bzw. $\bar{x}_2 = 6,5$ ermittelt. Wie ist der Durchschnitt der zusammengesetzten Stichprobe?

1.3. Die Varianz und die Standardabweichung

Der Mittelwert ist ein Wert, zu dem die Werte von X tendieren. Er braucht aber kein Wert zu sein, der in der Wirklichkeit zu beobachten ist, besonders dann nicht, wenn die Variable diskret ist. Wenn in den Sprachen durchschnittlich 30,3 Phoneme vorhanden sind, so bedeutet das nicht, daß in einer Sprache 30,3 Phoneme möglich sind. Bei einer Zufallsvariablen streuen die Werte um den

Mittelwert so, daß es sowohl negative als auch positive Abweichungen gibt. Die Summe dieser Abweichungen ist, wie in 1.2 gezeigt wurde, gleich Null.

Um das Ausmaß dieser Streuung zu messen, benutzt man andere Kenngrößen. Das meistbenutzte Streuungsmaß ist die durchschnittliche quadratische Abweichung vom Mittelwert, definiert als

$$S^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (1.14)$$

oder, wenn die Werte x_i öfter als einmal vorkommen und gruppiert werden, als

$$S^2 = \frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^2 f_i. \quad (1.15)$$

Diesen Ausdruck nennt man auch *Varianz*. Aus gewissen Gründen, die wir erst in der Schätzungstheorie nennen werden, benutzt man auch den Ausdruck

$$s^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (1.16)$$

bzw. bei gruppierten Werten

$$s^2 = \frac{1}{N-1} \sum_{i=1}^n (x_i - \bar{x})^2 f_i. \quad (1.17)$$

In empirischen Untersuchungen kann man in Einzelfällen nicht entscheiden, welche von diesen zwei Formeln als Schätzung der "wahren" Streuung in der Grundgesamtheit die "bessere" ist. Dazu müßte man die wahre Streuung kennen, und in diesem Fall würde sich eine Schätzung erübrigen. Man sollte diejenige Formel benutzen, die bei weiterer Bearbeitung der Daten im gegebenen statistischen Verfahren vorausgesetzt wird. Es ist leicht zu sehen, daß mit zunehmendem Stichprobenumfang N der Unterschied zwischen den beiden Ausdrücken immer geringer wird.

Die Wurzel aus S^2 bzw. aus s^2 nennt man *Standardabweichung*, die als Abweichungs- bzw. Streuungseinheit betrachtet wird.

Wenn $\bar{x}$ eine ganze Zahl ist, so ist die Berechnung der Ausdrücke $(x_i - \bar{x})^2$ ziemlich leicht, wenn aber $\bar{x}$ keine ganze Zahl ist, so erschwert sich die Berechnung erheblich. Eine einfachere Berechnungsformel bekommt man, wenn man das Binom $(x_i - \bar{x})^2$ entwickelt:

$$\begin{aligned}
 s^2 &= \frac{1}{N-1} \sum_{i=1}^n (x_i - \bar{x})^2 f_i \\
 &= \frac{1}{N-1} \sum_{i=1}^n (x_i^2 - 2\bar{x}x_i + \bar{x}^2) f_i \\
 &= \frac{1}{N-1} \left(\sum_i x_i^2 f_i - 2\bar{x} \sum_i x_i f_i + \bar{x}^2 \sum_i f_i \right) \\
 &= \frac{1}{N-1} \left(\sum_i x_i^2 f_i - 2 \frac{\sum_i x_i f_i}{N} \sum_i x_i f_i + \left(\frac{\sum_i x_i f_i}{N} \right)^2 N \right) \\
 &= \frac{1}{N-1} \left(\sum_i x_i^2 f_i - 2 \frac{(\sum_i x_i f_i)^2}{N} + \frac{(\sum_i x_i f_i)^2}{N} \right) \\
 &= \frac{1}{N-1} \left(\sum_i x_i^2 f_i - \frac{(\sum_i x_i f_i)^2}{N} \right). \quad (1.18)
 \end{aligned}$$

Man kann also die Varianz auch ohne Kenntnis des Mittelwerts berechnen. Für S^2 gilt der Ausdruck ebenfalls, man ersetzt lediglich $N-1$ durch N .

Beispiel 1.3.1. Im Beispiel 1.1.7 wurde die Verteilung der Länge indonesischer Morphemtypen angegeben (Tabelle 1.5). Wir berechnen S^2 und S .

Lösung: Die Berechnung von S^2 und S nach (1.18) erfolgt aus den Daten in Tabelle 1.9.

Tabelle 1.9. Berechnung der Varianz und der Standardabweichung

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
1	12	12	12
2	35	70	140
3	54	162	486
4	39	156	624
5	8	40	200
Σ	148	440	1462

Es ist also

$$\begin{aligned}
 S^2 &= \frac{1}{148} \left(1^2(12) + 2^2(35) + \dots + 5^2(8) - \frac{[1(12) + 2(35) + \dots + 5(8)]^2}{148} \right) \\
 &= \frac{1}{148} \left(1462 - \frac{440^2}{148} \right) = \frac{1}{148} (1462 - 1308,11) = 1,0398
 \end{aligned}$$

und daraus ist

$$S = \sqrt{1,0398} = 1,02. \blacksquare$$

Sind die Werte von X sehr groß, so werden sie durch das Quadrieren noch größer. Den Rechenaufwand kann man wiederum dadurch reduzieren, daß man eine Transformation durchführt. Benutzt man die Transformation (1.11), d.h.

$$u_i = \frac{x_i - A}{d},$$

so gilt nach (1.12) $\bar{u} = \frac{\bar{x} - A}{d}$. Danach ist

$$\begin{aligned}
 S_u^2 &= \frac{1}{N} \sum_{i=1}^n (u_i - \bar{u})^2 f_i \\
 &= \frac{1}{N} \sum_{i=1}^n \left(\frac{x_i - A}{d} - \frac{\bar{x} - A}{d} \right)^2 f_i =
 \end{aligned}$$

$$= \frac{1}{N} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{d} \right)^2 f_i$$

$$= \frac{1}{d^2 N} \sum_{i=1}^n (x_i - \bar{x})^2 f_i = \frac{S_x^2}{d^2}, \quad (1.19)$$

so daß schließlich

$$S_x^2 = d^2 S_u^2 \text{ bzw. } s_x^2 = d^2 s_u^2. \quad (1.20)$$

Beispiel 1.3.2. Wir ermitteln die Varianz s^2 aus den Daten des Beispiels 1.2.3 mit Hilfe der Transformation $u_i = (x_i - 110)/20$.

Lösung: Die Berechnung der notwendigen Zahlen erfolgt in Tabelle 1.10. Dann berechnen wir zuerst

$$s_u^2 = \frac{1}{N-1} \left(\sum_{i=1}^n u_i^2 f_i - \frac{(\sum u_i f_i)^2}{N} \right).$$

Aus Beispiel 1.2.3 kennen wir bereits $\sum u_i f_i = -6211$, also wird

$$s_u^2 = \frac{1}{3515} \left(23327 - \frac{(-6211)^2}{3516} \right)$$

$$= \frac{1}{3515} (23327 - 10971,71)$$

$$= 3,5150.$$

Daraus folgt

$$s_x^2 = d^2 s_u^2 = 400 (3,5150) = 1406,00$$

und die Standardabweichung

$$s_x = \sqrt{s_x^2} = 37,50. \blacksquare$$

Tabelle 1.10. Berechnung der Varianz mit transformierten Variablen

x_i	$u_i = \frac{x_i - 110}{20}$	f_i	$u_i^2 f_i$
30	-4	468	7488
50	-3	958	8622
70	-2	840	3360
90	-1	601	601
110	0	310	0
130	1	156	156
150	2	70	280
170	3	41	369
190	4	30	480
210	5	17	425
230	6	6	216
250	7	10	490
270	8	3	192
290	9	2	162
310	10	1	100
330	11	2	242
350	12	1	144
	Σ	3516	23327

Übung 1.3.1. Zeigen Sie, daß $s^2 = \frac{1}{N} \left(\sum_{i=1}^n x_i^2 f_i - N \bar{x}^2 \right).$

Übung 1.3.2. Finden Sie die Standardabweichung s der Meßdaten in der Übung 1.2.5. Die Daten sind

26,7; 31,5; 22,6; 28,2; 29,2; 25,3; 25,7.

[Bedenken Sie, daß $f_i = 1$ für alle x_i .]

Übung 1.3.3. Nach der Berechnung von S^2 wurde festgestellt, daß ein Wert x_f irrtümlich eingeschlossen wurde und gestrichen werden muß. Zeigen Sie, daß die korrigierte Varianz S_c^2 aus der berechneten Varianz wie folgt ermittelt werden kann:

$$S_c^2 = \frac{1}{N-1} \left(NS^2 - \frac{N(x_f - \bar{x})^2}{N-1} \right) = \frac{1}{N-1} \left(NS^2 - \frac{(Nx_f - \Sigma x)^2}{N(N-1)} \right). \quad (1.21)$$

[Bedenken Sie, daß man die Abweichungsquadrate von $\bar{x}_c$ berechnen muß, wo

$\bar{x}_c = \frac{N\bar{x} - x_f}{N-1}$; setzen Sie $NS^2 = \Sigma (x_i - \bar{x}_c + \bar{x}_c - \bar{x})^2$; dann quadrieren, $(x_f - \bar{x}_c)^2$ von

$\Sigma (x_i - \bar{x}_c)^2$ subtrahieren und ordnen.]

Übung 1.3.4. Bei der Berechnung von S aus den Daten 5, 4, 3, 8, 5 ergab sich $S^2 = 2,8$. Man hat aber festgestellt, daß der Wert 8 irrtümlich eingeschlossen wurde. Berechnen Sie S_c^2 mit Hilfe der Formel (1.12) und kontrollieren Sie das Ergebnis mit direkter Berechnung.

Aufgaben

1.1. Stellen Sie die Häufigkeitstabelle für die Daten aus dem Beispiel 1.1.3 mit Klassenmitten, relativen Häufigkeiten und relativen Summenhäufigkeiten auf.

1.2. Schreiben Sie aus

$$(a) \sum_{i=1}^n x_i (x_i - 3); \quad (b) \sum_{i=1}^n (x_i - 1)^2.$$

[Führen Sie die innere Operation aus und schreiben Sie in mehreren Summen.]

1.3. Berechnen Sie die relativen Häufigkeiten und die relativen Summenhäufigkeiten aus der Übung 1.1.1. Dann berechnen Sie s^2 und s .

1.4. Berechnen Sie

$$(a) \sum_{k=1}^n k(k-3); \quad (b) \sum_{k=1}^n (k-1)^2.$$

[Führen Sie die innere Operation aus. Benutzen Sie die Reihensummen in Anhang B1.]

1.5. Ludvíková (1972, S. 32) hat in Prosatexten die Wortlänge in Silbenzahl gezählt und bekam folgende Resultate:

Silbenzahl	x_i	1	2	3	4	5	≥ 6	Σ
Wörteranzahl	f_i	647	771	426	140	14	2	2000

Zeichnen Sie ein Stabdiagramm und berechnen Sie $\bar{x}$ und s .

1.6. Zeigen Sie, daß

$$\sum_{k=1}^n (2k-1) = n^2.$$

1.7. Berechnen Sie die mittlere Morphemtypplänge im Indonesischen aus dem Beispiel 1.1.7.

1.8. Berechnen Sie den Mittelwert und die Varianz der ersten n natürlichen Zahlen. [Benutzen Sie die Reihensummen im Anhang B1, bzw. Formel (1.18) für S^2 .]

1.9. Berechnen Sie $\bar{x}$ und s aus den Daten

x_i	20-22	23-25	26-28	29-31	32-34	35-37	38-40
f_i	2	5	9	16	8	4	1

1.10. Bei der Berechnung des Mittelwerts wurde die Häufigkeit f'_i eines bestimmten Wertes x_i falsch gelesen und muß durch die richtige Häufigkeit f_i ersetzt werden. Zeigen Sie, daß sich der richtige Mittelwert x_c aus $\bar{x}$ mit Hilfe der Formel

$$\bar{x}_c = \frac{N\bar{x} + x_i(f_i - f'_i)}{N - f'_i + f_i}$$

berechnen läßt.

1.11. In 3 Stichproben von $N_1 = 20$, $N_2 = 25$, $N_3 = 15$ wurden folgende Mittelwerte berechnet: $\bar{x}_1 = 7,15$; $\bar{x}_2 = 6,65$; $\bar{x}_3 = 7,05$. Finden Sie den Mittelwert der zusammengesetzten Stichprobe.

1.12. Eine Zufallsvariable nimmt die Werte der ersten n natürlichen Zahlen an mit Häufigkeiten, die mit den jeweiligen Werten identisch sind. Zeigen Sie, daß $\bar{x} = \frac{(2n+1)}{3}$. [Bedenken Sie, daß $\sum f_i = \sum k$ und $\sum x_i f_i = \sum k^2$.]

1.13. Berechnen Sie s^2 aus den Daten von Lehfeldt im Beispiel 1.1.3.

1.14. Eine Zufallsvariable nimmt die Werte $1, 1+d, 1+2d, \dots, 1+nd$ mit Häufigkeiten jeweils 1 an. Berechnen Sie den Mittelwert und die Varianz S^2 .

1.15. Berechnen Sie S^2 aus den Daten in Beispiel 1.2.1.

1.16. Aus zwei Stichproben von N_1 und N_2 wurden $\bar{x}_1$ und $\bar{x}_2$ bzw. S_1^2 und S_2^2 berechnet. Zeigen Sie, daß die Varianz S^2 der zusammengesetzten Stichprobe von $N_1 + N_2$ Elementen

$$S^2 = \frac{N_1 S_1^2 + N_2 S_2^2}{N_1 + N_2} + \frac{N_1 N_2 (\bar{x}_1 - \bar{x}_2)^2}{(N_1 + N_2)^2} \text{ ist.} \quad (1.32)$$

[Fügen Sie $\bar{x}_1$ bzw. $\bar{x}_2$ in S^2 ein. Ersetzen Sie den Durchschnitt $\bar{x} = \frac{N_1 \bar{x}_1 + N_2 \bar{x}_2}{N_1 + N_2}$,

ordnen Sie dann.]

1.17. Zwei Stichproben von $N_1 = 20$; $N_2 = 22$; $\bar{x}_1 = 4,3$; $\bar{x}_2 = 3,5$; $S_1^2 = 5,7$; $S_2^2 = 6,8$ wurden zusammengesetzt. Berechnen Sie $\bar{x}$ und S^2 der zusammengesetzten Stichprobe. [Benutzen Sie Formel (1.32).]

1.18. Nach der Berechnung von S^2 aus einer Stichprobe mit N Elementen hat man festgestellt, daß ein Wert x_f falsch gelesen wurde und durch den richtigen Wert x_c ersetzt werden muß. Zeigen Sie, daß die korrekte Varianz S_c^2 sich aus S^2 folgendermaßen berechnen läßt:

$$S_c^2 = S^2 + \frac{1}{N} (x_c - x_f) \left(x_c + x_f - \frac{x_c - x_f + 2 \Sigma x}{N} \right), \quad (1.33)$$

wobei Σx die ursprüngliche, nicht korrigierte Summe ist (Kapur-Saxena 1969, S.44). [Setzen Sie

$$S_c^2 = \frac{1}{N} \left(\sum x_i^2 - x_f^2 + x_c^2 - \frac{(\Sigma x_i - x_f + x_c)^2}{N} \right)$$

und ordnen Sie.]

- 1.19. Berechnen Sie S^2 aus den Zahlen 1, 2, 3, 4, 5, 6, 7, dann entfernen Sie 6 und berechnen Sie die neue Varianz S_c^2 direkt und mit Hilfe der Formel (1.21).
- 1.20. Benutzen Sie die Daten aus Aufgabe 1.19. Dann berechnen Sie die neue Varianz S_c^2 direkt und mit Hilfe der Formel (1.33) nachdem 6 durch 8 ersetzt wurde.

Benutzte und weiterführende Literatur : Geršić (1971), Gradshteyn-Ryzhik (1973), Kapur-Saxena (1969), Kendall-Stuart (1969), Lehfeldt (1975), Ludvíková (1972), Neyman (1950), Weber (1961).

2. Kombinatorik

Um ihrer Aufgabe nachkommen zu können, Daten nicht nur darzustellen, sondern auch zu beurteilen, benötigt die Inferenzstatistik Bezugsgrößen und mathematische Modelle, aus denen sie diese gewinnen kann.

Das Problem, aus 49 Symbolen 6 richtig zu erraten, und der Befund, daß in einer Sprache 72% aller Silben den Vokal /a/ haben, sind nur zu beurteilen, wenn man die Wahrscheinlichkeit für das jeweilige Ereignis kennt. Wie später gezeigt wird, benötigt man für die Berechnung der Wahrscheinlichkeit eines Ereignisses die Anzahl der möglichen Ereigniskonfigurationen. Grundlage dazu ist die Kombinatorik; Permutation, Variation und Kombination liefern die Basis für die wahrscheinlichkeitstheoretische Modellierung.

2.1. Permutation ohne Wiederholung

Stellt man zwei Buchstaben A und B in allen möglichen Reihenfolgen auf, so erhält man lediglich

AB und BA,

d.h. 2 Reihenfolgen, die wir *Permutationen* nennen werden. Mit 3 Buchstaben, A, B, und C erhält man

ABC
ACB
BAC
BCA
CAB
CBA,

d.h. 6 Permutationen. Wächst die Anzahl der Buchstaben, so wächst sehr schnell auch die Anzahl der Permutationen: Bei 4 Buchstaben hätte man schon 24 Permutationen, bei 5 Buchstaben 120 usw. Man hätte viel Mühe, sie alle aufzuzählen. Deswegen muß man eine Formel finden, die es erlaubt, die Anzahl

der Permutationen von n Elementen ohne ihre Aufzählung zu berechnen. Man überlegt folgendermaßen:

Im obigen Fall kann man an die erste Stelle der Buchstabenreihe alle 3 Buchstaben setzen; an die zweite Stelle kann man nach der Fixierung des ersten Buchstabens nur noch die zwei übriggebliebenen Buchstaben stellen; z.B. kann nach A entweder B oder C stehen; an der dritten Stelle kann danach nur der letzte Buchstabe stehen. Dadurch erhält man

$$\begin{array}{ccccccc} 3 & \times & 2 & \times & 1 & = & 6 \\ \text{erste} & & \text{zweite} & & \text{dritte} & & \\ \text{Stelle} & & \text{Stelle} & & \text{Stelle} & & \end{array}$$

Permutationen, die oben aufgeführt wurden. Wiederholt man diesen Gedankengang mit 4 Elementen, so bekommt man

$$4 \times 3 \times 2 \times 1 = 24,$$

und allgemein bei n Elementen erhält man

$$n(n-1)(n-2) \dots 3 \cdot 2 \cdot 1 = n! \quad (2.1)$$

Permutationen. Die Zahl $n!$ nennt man n -Fakultät.

Bezeichnet man die Anzahl der Permutationen aus n Elementen als P_n , so erhält man die Formel

$$P_n = n! \quad (2.2)$$

und diese Formel gilt natürlich nur dann, wenn alle Elemente unterschiedlich sind.

Beispiel 2.1.1. Wie viele Wörter kann man durch Umstellung des Wortes KRAUT bilden?

Lösung: Es gibt 5 unterschiedliche Buchstaben, daher

$$P_5 = 5! = 120. \blacksquare$$

Beispiel 2.1.2. Wie viele sinnvolle und sinnlose Folgen kann man aus den Worten des Satzes DER RASEN IST GRÜN bilden?

Lösung: Es gibt 4 unterschiedliche Wörter, daher

$$P_4 = 4! = 24. \blacksquare$$

Wie man aus der Formel (2.1) direkt ansehen kann, gelten für die Fakultäten folgende Beziehungen

$$\begin{aligned} 1! &= 1 \\ n! &= n(n-1)! \\ (n+1)! &= (n+1)n! \end{aligned} \quad (2.3)$$

Die Berechnung der Fakultäten wird durch Tabelle A1 im Anhang erleichtert.

Bemerkung: Die Permutierbarkeit der Wörter im Satz ist in jeder Sprache mehr oder weniger eingeschränkt; daher kann man das Verhältnis der erlaubten und der theoretisch möglichen Wortpermutationen als ein Maß der Strenge der Wortstellung betrachten. Der Anteil der Phonempermutationen, die sinnvolle Wörter ergeben, deutet das Maß der phonematischen Wortbildung an. Je größer dieser Anteil, desto größer die Wortbildungsökonomie.

Übung 2.1.1. Berechnen Sie P_7 .

Übung 2.1.2. Wieviel ist $5(4!)$?

Übung 2.1.3. Wie viele unterschiedliche Wörter (auch sinnlose) kann man durch Umstellung der Buchstaben des Wortes SEMANTIK bilden?

Übung 2.1.4. Wie viele unterschiedliche Wörter (auch sinnlose) kann man durch Umstellung der Buchstaben von ORTE bilden? [Wie viele davon gibt es im Deutschen?]

2.2. Permutation mit Wiederholung

Wir bilden alle Permutationen der Symbole A, B, C, diesmal lassen wir aber zu, daß sich die Symbole auch wiederholen, d.h. wir bilden auch Reihenfolgen wie AAB, ACA, BBB, BBC usw. In diesem Fall kann man an die erste Stelle der Reihenfolge alle 3 Symbole setzen, an die zweite Stelle auch alle 3 und an die dritte Stelle ebenfalls alle 3, so daß man insgesamt

$$\begin{array}{ccccccc} 3 & \times & 3 & \times & 3 & = & 27 \\ \text{erste} & & \text{zweite} & & \text{dritte} & & \\ \text{Stelle} & & \text{Stelle} & & \text{Stelle} & & \end{array}$$

Permutationen mit Wiederholungen der Buchstaben bekommt. Mit 4 Buchstaben bekommt man

$$4 \times 4 \times 4 \times 4 = 4^4$$

Permutationen, und allgemein mit n Elementen erhält man

$$\underbrace{n \times n \times n \times \dots \times n}_{n\text{-mal}} = n^n$$

Permutationen. Bezeichnen wir die Anzahl der Permutationen von n Elementen mit Wiederholung als P'_n , so erhalten wir die nächste Formel

$$P'_n = n^n. \quad (2.4)$$

Beispiel 2.2.1. Wie viele Wörter kann man aus den Buchstaben des Wortes ER bilden, wenn Buchstabenwiederholung erlaubt ist?

Lösung: Es gibt 2 Buchstaben, daher

$$P'_2 = 2^2 = 4. \blacksquare$$

Beispiel 2.2.2. Wie viele verschiedene "Melodien" lassen sich durch Permutation von 5 Tönen bilden, wenn Wiederholung erlaubt ist?

Lösung: Es gibt 5 Elemente, daher

$$P'_5 = 5^5 = 3125. \blacksquare$$

Der allgemeine Permutationsfall ist derjenige, in dem die Elemente beschränkt wiederholbar sind, d.h., an bestimmten Stellen dürfen nur bestimmte Elemente stehen. Hat man 3 Buchstaben A, B, C und bildet aus ihnen Wörter so, daß man an der ersten Stelle die beiden "Konsonanten" zuläßt, an der zweiten Stelle nur den "Vokal" und an der dritten Stelle wiederum nur die beiden "Konsonanten", so erhält man insgesamt

$$\begin{array}{ccccc} 2 & \times & 1 & \times & 2 & = & 4 \\ \text{erste} & & \text{zweite} & & \text{dritte} & & \\ \text{Stelle} & & \text{Stelle} & & \text{Stelle} & & \end{array}$$

Wörter. Allgemein formuliert ist die Anzahl der Permutationen von n Elementen mit beschränkter Wiederholung, bei denen an der ersten Stelle k_1 Elemente, an der zweiten Stelle k_2 Elemente usw. bis zu der n -ten Stelle, wo k_n Elemente stehen dürfen, gleich

$$P'_n(k_1, k_2, \dots, k_n) = k_1 \cdot k_2 \cdot \dots \cdot k_n = \prod_{i=1}^n k_i. \quad (2.5)$$

Hier ist k_i die Anzahl der Elemente, die an der i -ten Stelle stehen dürfen, und Π bezeichnet das Produkt dieser Zahlen.

Beispiel 2.2.3. Aus den Buchstaben des Wortes HUND soll man Wörter derselben phonologischen Struktur KVKK bilden. Wie viele Möglichkeiten gibt es,

(a) wenn die Buchstaben nicht wiederholt werden dürfen,

(b) wenn die Wiederholung der Buchstaben erlaubt ist?

Lösung: (a) Ohne Wiederholung ergibt sich

$$3 \times 1 \times 2 \times 1 = 6$$

(b) Mit Wiederholung ergibt sich

$$3 \times 1 \times 3 \times 3 = 27. \blacksquare$$

Offensichtlich sind P_n und P'_n nur Spezialfälle der Formel (2.5). Bei P_n ist $k_1 = n, k_2 = n - 1, k_3 = n - 2$ usw. bis $k_n = 1$, oder allgemein $k_i = n - i + 1$, und bei P'_n ist $k_1 = k_2 = \dots = k_n = n$.

Eine weitere Verallgemeinerung führt zu der *Multiplikationsregel*, bei der die Anzahl der Elemente nur stellenweise bestimmt ist: Man bildet aus n Elementen Reihenfolgen, die aus r Stellen bestehen. An die erste Stelle stellt man k_1 Elemente, von denen jedes von k_2 Elementen gefolgt wird, diese wiederum einzeln k_3 Nachfolger haben usw. bis zu der letzten Stelle, wo hinter jedem von k_{r-1} Elementen k_r Elemente stehen dürfen. So ist die Anzahl aller Reihenfolgen dieser Art

$$M_n(k_1, k_2, \dots, k_r) = k_1 \cdot k_2 \cdot \dots \cdot k_r = \prod_{i=1}^r k_i. \quad (2.6)$$

Beispiel 2.2.4. Eine Sprache hat 20 Konsonanten (K) und 5 Vokale (V). Wie viele Morpheme des Typs KVKV kann man aus diesem Inventar bilden?

Lösung: An der ersten und der dritten Stelle dürfen nur Konsonanten, an der zweiten und der vierten nur Vokale stehen, daher

$$20 \times 5 \times 20 \times 5 = 10000. \blacksquare$$

Beispiel 2.2.5. Wie viele Sätze mit der Wortstellung ASV oder VAS lassen sich mit 4 Substantiven (S), 2 Adjektiven (A) und 11 Verben (V) bilden?

Lösung: Im ersten Satztyp hat man

$$2 \times 4 \times 11 = 88$$

und im zweiten

$$11 \times 2 \times 4 = 88$$

also insgesamt $88 + 88 = 176$ Möglichkeiten. ■

Übung 2.2.1. Wie viele Wörter kann man aus den Phonemen des Wortes GUT durch Umstellung bilden, wenn auch Wiederholungen erlaubt sind?

Übung 2.2.2. Wieviel ist P_7 ?

Übung 2.2.3. Wie viele Wörter derselben Länge kann man aus den Buchstaben des Wortes SEMANTIK bilden, wenn man die Buchstaben wiederholen darf?

Übung 2.2.4. Wie viele Sätze des Typs TC (T = Topic, C = Comment) kann man bilden, wenn man als T 10 Wörter und als C 4 Wörter zur Verfügung hat?

Übung 2.2.5. Wie viele Morpheme des Typs KVK lassen sich aus einem Inventar von 20 Konsonanten und 4 Vokalen bilden, wenn in der Endposition nur 7 Konsonanten vorkommen dürfen?

2.3. Permutation von Elementen, die nicht alle unterschiedlich sind

Untersuchen wir nun die Permutation der Symbole A, B, C, C, C. Wären alle drei C unterschiedlich, z.B. als C_1, C_2, C_3 bezeichnet, so bekämen wir nach (2.2) $P_5 = 5!$. Da sie aber gleich sind, verringert sich die Zahl der Permutationen. Jedesmal, wenn man die Position von A und B fixiert, z.B.

C A B C C,

kann man die drei C untereinander auf 3! Weisen umstellen, ohne daß die Permutation ihre Identität verliert. Mit C_1, C_2, C_3 hätten wir statt CABCC die Permutationen

$C_1 A B C_2 C_3$
 $C_1 A B C_3 C_2$
 $C_2 A B C_1 C_3$
 $C_2 A B C_3 C_1$
 $C_3 A B C_1 C_2$
 $C_3 A B C_2 C_1$

d.h. 3!-mal mehr. Bei Nichtunterscheidung von drei C erhalten wir also um 3!-mal weniger Permutationen, in unserem Fall demnach

$$\frac{P_5}{3!} = \frac{5!}{3!}.$$

Hätten wir sechs Symbole A, B, B, C, C, C, so müßten wir die Gesamtzahl der Permutationen, $P_6 = 6!$, dividieren durch die Anzahl möglicher Vertauschungen der beiden B untereinander (unabhängig von der Position aller übrigen) und durch die Anzahl möglicher Vertauschungen der drei C untereinander (ebenfalls unabhängig von der Position der übrigen), also durch 2! und durch 3!. Als Ergebnis bekommen wir

$$\frac{P_6}{2! 3!} = \frac{6!}{2! 3!} = 60 \text{ Permutationen.}$$

Allgemein sei n die Anzahl der Elemente, von denen k_1 von einer Sorte, k_2 von der zweiten Sorte, ..., k_r von der r -ten Sorte sind, und $k_1 + k_2 + \dots + k_r = n$. Bezeichnen wir die Anzahl der Permutationen dieser $k_1, k_2, \dots, k_r$ Elemente als $P_n^{k_1, k_2, \dots, k_r}$, so erhalten wir auf Grund des obengesagten

$$P_n^{k_1, k_2, \dots, k_r} = \frac{P_n}{k_1! k_2! \dots k_r!} = \frac{n!}{k_1! k_2! \dots k_r!}. \quad (2.7)$$

Diese Ausdrücke heißen *Multinomialkoeffizienten* (vgl. 2.9).

Beispiel 2.3.1. Wie viele Wörter lassen sich durch die Permutierung der Buchstaben des Wortes BUTTER bilden?

Lösung: Anzahl von B: $k_B = 1$

Anzahl von U: $k_U = 1$

Anzahl von T: $k_T = 2$

Anzahl von E: $k_E = 1$

Anzahl von R: $k_R = 1$

Anzahl aller Buchstaben : $n = 6$,

daher nach (2.7)

$$P_6^{1, 1, 2, 1, 1} = \frac{6!}{1! 1! 2! 1! 1!} = \frac{6!}{2!} = 360. \blacksquare$$

Beispiel 2.3.2. Wie viele Wörter lassen sich durch Permutierung des Wortes ERKENNTNIS bilden?

Lösung: $n = 10, k_E = 2, k_N = 3, k_R = k_K = k_T = k_I = k_S = 1$,

$$P_{10}^{2,3,1,1,1,1,1} = \frac{10!}{2!3!(1!)^5} = \frac{3628800}{12} = 302400. \blacksquare$$

Wenn in (2.7) alle $k_1 = k_2 = \dots = k_r = 1$, d.h. wenn alle Elemente nur einmal vorhanden sind, so wird (2.7) gleich der Formel (2.2), die wieder ein Spezialfall dieser Formel ist.

Übung 2.3.1. Wie viele Permutationen eines Satzes DET N V N DET N wären bei absolut freier Wortstellung möglich?

Übung 2.3.2. Wie viele Wörter kann man durch Umstellung der Buchstaben von MESSE bilden?

Übung 2.3.3. Wie viele Wörter in 2.3.2 enden

- (a) mit M,
- (b) mit E,
- (c) entweder mit M oder mit E? (Stellen Sie M bzw. E an die letzte Stelle und berechnen Sie dann (2.7) aus dem Rest).

Übung 2.3.4. Wie viele Wörter kann man durch Umstellung der Buchstaben des Wortes POPOCATEPETL bilden?

Übung 2.3.5. Wie viele strukturelle Morphemtypen, bestehend aus 5 Phonemen, lassen sich aus 3 Konsonanten (K) und 2 Vokalen (V) bilden? (Struktureller Typ ist z.B. KVKVK, VKKVK, usw.)

2.4. Variation ohne Wiederholung

Nehmen wir an, daß wir alle Buchstaben des Alphabets zur Verfügung haben. Aus diesem Inventar wählen wir immer 3 Buchstaben aus, die wir noch permutieren. So bekommen wir Buchstabenfolgen wie

ABC,	ACB,	BAC,	BCA,	CAB,	CBA,
ABD,	ADB,	BAD,	BDA,	DAB,	DBA,
ABE,	AEB,	BAE,	BEA,	EAB,	EBA,
.....					
ACD,	ADC,	CAD,	CDA,	DAC,	DCA,
.....					
BDE,	BED,	DBE,	DEB,	EBD,	EDB,
BDF,	BFD,	DBF,	DFB,	FBD,	FDB,
.....					
XYZ,	XZY,	YXZ,	YZX,	ZXY,	ZYX,

Solche Reihenfolgen von n Elementen, jeweils k gleichzeitig genommen und permutiert, nennen wir Variationen. Ihre Anzahl läßt sich mit derselben Überlegung berechnen wie die Anzahl der Permutationen:

Aus n Buchstaben kann man an die erste Stelle n setzen, an die zweite Stelle nur $n-1$, an die dritte Stelle nur $n-2$ usw. bis zur k -ten Stelle, wo nur $n-k+1$ Buchstaben stehen können, d.h.

$$n(n-1)(n-2)\dots(n-k+1) \quad (2.8)$$

was man auch als

$$(n-1+1)(n-2+1)(n-3+1)\dots(n-k+1)$$

schreiben kann. Diesen Ausdruck werden wir als $n_{(k)}$ bezeichnen. Um eine handliche Formel für die Berechnung der Anzahl der Variationen zu finden, multiplizieren wir (2.8) mit dem Ausdruck

$$\frac{(n-k)!}{(n-k)!}$$

wodurch sich (2.8) nicht ändert. Wir bekommen aber

$$\frac{n(n-1)(n-2)\dots(n-k+1)(n-k)!}{(n-k)!} = \frac{n!}{(n-k)!}$$

auf Grund der Beziehung (2.3). Daher ist die Anzahl der Variationen von n Elementen jeweils k genommen (als Variationen k -ter Ordnung bezeichnet) gleich

$$V_n^k = \frac{n!}{(n-k)!} \quad (2.9)$$

Beispiel 2.4.1. Eine Sprache hat 20 Phoneme. Wie viele geordnete Phonempaare lassen sich bilden, wenn gleiche Phoneme nicht nebeneinander stehen können?

Lösung: Wir bilden alle Variationen 2-ter Ordnung aus 20 Elementen, d.h.

$$n = 20$$

$$k = 2$$

und bekommen nach (2.8)

$$V_{20}^2 = 20(19) = 380$$

oder nach (2.9)

$$V_{20}^2 = \frac{20!}{(20-2)!} = \frac{20!}{18!} = 20(19) = 380. \blacksquare$$

Beispiel 2.4.2. Auf wie viele Weisen können eine kurze und eine lange Pause in einem Satz von 6 Wörtern zwischen den Wörtern plaziert werden?

Lösung: Zwischen 6 Wörtern liegen insgesamt 5 Zwischenräume. Aus diesen 5 wählt man immer 2 Positionen, auf die man die Pausen plaziert, und permutiert sie, weil die kurze Pause vor der langen und umgekehrt stehen kann. Daher

$$V_5^2 = \frac{5!}{(5-2)!} = \frac{5!}{3!} = 5(4) = 20. \blacksquare$$

Beispiel 2.4.3. In einer Sprache mit 15 Konsonanten fangen alle dreigliedrigen Konsonantenverbindungen (Cluster) mit einem /s/ an, und die Konsonanten werden in den Cluster nicht wiederholt. Wie viele Cluster dieser Art kann man bilden?

Lösung: Wenn /s/ an der ersten Stelle steht, kann man die restlichen 2 Stellen aus den restlichen 14 Konsonanten auf

$$V_{14}^2 = \frac{14!}{12!} = 182$$

Weisen besetzen, daher gibt es in der Sprache 182 (mögliche) Cluster. $\blacksquare$

Man kann leicht einsehen, daß die Permutation nur ein Spezialfall der Variation ist; denn ist $k = n$, so bekommt man

$$V_n^n = n(n-1)(n-2) \dots 3(2)1 = n! = P_n. \quad (2.10)$$

Aus dieser Beziehung folgt noch eine wichtige Tatsache. Wenn nämlich nach (2.10) und (2.9)

$$V_n^n = n! = \frac{n!}{(n-n)!}$$

dann folgt aus der Beziehung auf der rechten Seite, daß

$$(n-n)! = \frac{n!}{n!}$$

oder daß

$$0! = 1. \quad (2.11)$$

Übung 2.4.1. Berechnen Sie V_3^3 .

Übung 2.4.2. Das Alphabet einer Sprache enthält 25 Buchstaben. Jedes Autoschild (ohne Zahlen) enthält 3 Buchstaben, die nicht wiederholt werden dürfen. Wie viele Autoschilder kann man (ohne Rücksicht auf die Zahlen) bilden?

Übung 2.4.3. Eine Sprache hat 22 Konsonanten und kann insgesamt 9200 dreigliedrige Konsonantenverbindungen ($K_1K_2K_3$) ohne Wiederholung der Konsonanten bilden. Sind das alle möglichen dreigliedrigen Konsonantenverbindungen?

Übung 2.4.4. Zeigen Sie, daß $2 \cdot P_5^{2,3} = V_5^2$ [Schreiben Sie in Fakultäten aus.]

2.5. Variation mit Wiederholung

Wenn wir n Elemente haben und aus ihnen Variationen k -ter Ordnung bilden, jedoch so, daß sich die gewählten Elemente in jeder der k Positionen auch wiederholen dürfen, dann können an jeder der k Stellen n Elemente stehen, d.h.

$$\begin{array}{ccccccc} n & \times & n & \times \dots \times & n & = & n^k \\ \text{erste} & & \text{zweite} & & \text{k-te} & & \\ \text{Stelle} & & \text{Stelle} & & \text{Stelle} & & \end{array}$$

Die Anzahl der Variationen k -ter Ordnung aus n Elementen mit Wiederholung ist also

$$V_n^k = n^k \quad (2.12)$$

Beispiel 2.5.1. Wie viele zweigliedrige Konsonantenverbindungen (Cluster) kann man aus 20 Konsonanten bilden?

Lösung: Da auch Cluster wie /pp/, /bb/ usw. zugelassen sind, können an der ersten Stelle des Clusters alle 20 Konsonanten stehen, an der zweiten Stelle auch alle 20, insgesamt also

$$V_{20}^2 = 20^2 = 20(20) = 400. \blacksquare$$

Beispiel 2.5.2. Wie viele Variationen zweiter Ordnung mit Wiederholung kann man aus den Buchstaben des Wortes FENSTER bilden?

Lösung: Das Wort enthält 7 Buchstaben, zwei sind aber gleich, so daß man nur 6 unterschiedliche Buchstaben hat. Aus 6 Buchstaben lassen sich genau $V_6^2 = 6^2 = 36$

Variationen bilden.■

Beispiel 2.5.3. Die Aufeinanderfolge von Wortarten soll für eine gegebene Sprache untersucht werden. Die verwendete Grammatik gibt 10 Wortarten vor. Wie viele verschiedene Ketten der Länge 3 sind grundsätzlich möglich, wenn

- (a) Wortarten wiederholt vorkommen dürfen?
- (b) nur Ketten betrachtet werden, deren erstes Element die Wortart A ist?
- (c) nur Ketten mit A an erster Stelle betrachtet werden und Wiederholung nicht erlaubt ist?

Lösung: (a) Aus 10 Elementen bildet man Variationen 3-ter Ordnung mit Wiederholung auf $V_{10}^3 = 10^3 = 1000$ Weisen.

(b) Steht A immer an der ersten Stelle, und ist die Wiederholung möglich, dann kann man die zwei restlichen Stellen auf $V_{10}^2 = 10^2 = 100$

Weisen besetzen.
(c) Wenn Wiederholung nicht erlaubt ist und an der ersten Stelle steht A, so kann man die restlichen Stellen nur mit den restlichen 9 Buchstaben besetzen, und zwar auf

$$V_9^2 = \frac{9!}{7!} = 72 \text{ Weisen. } \blacksquare$$

- Übung 2.5.1.** Berechnen Sie V_7^2 .
Übung 2.5.2. Eine Sprache besitzt 26 Konsonanten. Alle Radikale bestehen aus drei Konsonanten. Wie viele Radikale sind in dieser Sprache möglich, wenn Konsonantenwiederholung erlaubt ist?
Übung 2.5.3. Wie viele Radikale enden mit /K/?

2.6. Kombination ohne Wiederholung

Bei den Permutationen und Variationen spielte die Reihenfolge der Elemente eine Rolle. Zwei Reihenfolgen wie ABC und BAC zählten als zwei verschiedene

Permutationen. Wenn es uns jedoch nur darum geht, aus n Elementen einige ohne Rücksicht auf ihre Reihenfolge zu wählen, dann betrachten wir alle Folgen wie ABC, ACB, BAC, BCA, CAB, CBA als identisch, als eine einzige Einheit. Diese Einheit wird als Kombination der 3 Elemente A, B, C bezeichnet. Bei den Kombinationen geht es also lediglich um die Verschiedenheit und nicht um die Reihenfolge der Elemente.

Nehmen wir beispielsweise 4 Elemente, A, B, C, D ($n = 4$) und bilden aus ihnen alle Kombinationen und Variationen von 3 Elementen ($k = 3$). Die Kombinationen sind die Zusammenstellungen von jeweils 3 Buchstaben, d.h.

- ABC
- ABD
- ACD
- BCD.

Variationen entstehen, indem jede Zusammenstellung $k!$ -mal permutiert wird, z.B. entstehen aus ABC dann noch $3!$ Variationen, und zwar

- ABC
- ACB
- BAC
- BCA
- CAB
- CBA

und ähnlich auch aus den restlichen Kombinationen. Bezeichnen wir C_n^k als die Anzahl der Kombinationen k -ter Ordnung (d.h. jeweils k), von n Elementen, so erhalten wir die gezeigte Beziehung

$$C_n^k (k!) = V_n^k$$

woraus leicht zu berechnen ist, daß

$$C_n^k = \frac{V_n^k}{k!} = \frac{n!}{k! (n - k)!} = \binom{n}{k}. \tag{2.13}$$

Der Ausdruck $\binom{n}{k}$ (lies “ n über k ”) heißt *Binomialkoeffizient*, und es ist leicht einzusehen, daß

$$\binom{n}{k} = \binom{n}{n-k}. \quad (2.14)$$

Wir definieren noch

$$\binom{n}{k} = 0, \text{ wenn } k > n \text{ oder } k < 0. \quad (2.15)$$

Beispiel 2.6.1. Ein Dichter stellte sich ein Verzeichnis von 20 Wörtern auf und nahm sich vor, in jedem Vers 3 davon zu benutzen.

(a) Wie viele Wahlmöglichkeiten hatte er für den ersten Vers?

(b) Wie viele Wahlmöglichkeiten hatte er für den zweiten Vers, wenn er die im ersten Vers gewählten Wörter nicht mehr anwenden wollte?

Lösung: (a) Aus 20 Wörtern kann man 3 auf

$$C_{20}^3 = \binom{20}{3} = \frac{20!}{3!(20-3)!} = \frac{20(19)18}{3(2)} = 1140$$

Weisen wählen.

(b) Wenn die erstgewählten Wörter nicht mehr verwendet werden dürfen, dann kann der Dichter nur noch aus 17 Wörtern 3 wählen, und zwar auf

$$C_{17}^3 = \binom{17}{3} = \frac{17!}{3!14!} = \frac{17(16)15}{3(2)} = 680$$

Weisen. ■

Beispiel 2.6.2. Aus 8 Personen sollte eine Kommission von 4 Mitgliedern gewählt werden.

(a) Wie viele Möglichkeiten gibt es insgesamt?

(b) Wie viele Möglichkeiten gibt es, daß zwei bestimmte Personen A und B gemeinsam in der Kommission sitzen

(c) Wie viele Möglichkeiten gibt es, wenn eine bestimmte Person A immer Mitglied der Kommission sein muß?

(d) Wie viele Kommissionen gibt es, in denen Person C sitzt, die Person D aber nicht?

Lösung: (a) Die Anzahl aller Möglichkeiten ist

$$C_8^4 = \frac{8!}{4!4!} = 70.$$

(b) Wenn A und B in der Kommission sitzen, kann können die restlichen 2 Stellen aus den restlichen 6 Personen auf

$$C_6^2 = \frac{6!}{2!4!} = 15$$

Weisen gewählt werden.

(c) Wenn die Person A immer gewählt werden muß, so können die restlichen 3 Plätze aus 7 Personen auf

$$C_7^3 = \frac{7!}{3!4!} = 35$$

Weisen besetzt werden.

(d) Wenn Person C in der Kommission sitzt, dann sind nur noch 3 Plätze in der Kommission frei, die jedoch nur von 6 Personen besetzt werden können, da D nicht gewählt werden darf, um nicht mit C zusammen in der Kommission zu sitzen.

Daher ist die mögliche Anzahl solcher Kommissionen gleich

$$C_6^3 = \frac{6!}{3!3!} = 20. \blacksquare$$

Übung 2.6.1. Wieviel ist C_3^4 , C_n^2 , C_3^3 ?

Übung 2.6.2. Im Beispiel 2.6.1 (b), wenn die einmal benutzten Wörter nicht mehr verwendet werden dürfen, wie ist der Unterschied zwischen der Wahlanzahl im dritten und im vierten Vers?

Übung 2.6.3. Eine Sprache hat 10 "Wortarten".

(a) Auf wie viele Weisen kann man Äußerungen von 4 Wörtern bilden, wenn jede Wortart nur einmal in der Äußerung erscheinen darf und die Wortstellung fest ist?

(b) Wie viele Äußerungen von 4 Wörtern kann man bilden, wenn in jeder Äußerung ein Nomen und ein Verbum stehen muß, jede Wortart nur einmal im Satz erscheinen darf und die Wortstellung frei ist?

2.7. Kombination mit Wiederholung

Bei den Kombinationen in 2.6 war die Anordnung der Elemente ohne Belang, nur ihr Zusammenfinden war wichtig, d.h. ABC und BCA waren eine einzige Einheit, aber die Elemente durften nicht wiederholt werden. Stellen wir uns vor, daß sich jetzt jedes Element auch noch wiederholen darf. So hätten wir als Kombinationen zweiter Ordnung (= jeweils 2) von 4 Elementen A, B, C, D mit Wiederholung folgender Reihenfolgen

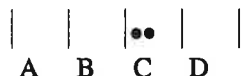
AA	BB	CC	DD
AB	BC	CD	
AC	BD		
AD			

deren Anzahl 10 ist, während ohne Wiederholung nur $\binom{4}{2} = 6$ möglich sind. Um die Zahl der Kombinationen mit Wiederholung zu berechnen, stellen wir uns 4

Urn vor, in die wir 2 Kugeln werfen. Die Urnen bezeichnen wir mit A, B, C, D. Wirft man beide Kugeln in die Urne A, so bekommt man die Folge AA, wirft man eine Kugel in die Urne B, die andere in die Urne D, so erhält man die Folge BD usw. Wir zeichnen die Urnen nebeneinander in Form von Strichen und die Kugeln als Punkte, z.B.



was die Kombination BD darstellt, oder



was die Kombination CC darstellt usw.

Wenn wir nun n Urnen haben, so werden sie als $n + 1$ Striche dargestellt, in unserem Falle 4 Urnen mit 5 Strichen. Wenn wir die Kugeln in einzelne Urnen werfen, so bilden wir eigentlich Permutationen von Kugeln und Strichen. Die Anzahl der Striche, die mit Kugeln permutiert werden dürfen, ist $n - 1$, weil aus $n + 1$ Strichen 2 an den Außenseiten fixiert sind und nicht permutiert werden dürfen, sonst würde eine Kugel außerhalb der Urnen liegen. Ist die Anzahl der Kugeln k , so haben wir die Permutationen von $n - 1 + k$ Elementen zu bilden, von denen $n - 1$ von einer Sorte (Striche) und k von der anderen Sorte (Punkte) sind. Nach der Formel (2.7) ist die Anzahl gleich

$$P_{n-1+k}^{n-1,k} = \frac{(n+k-1)!}{k!(n-1)!} = \binom{n+k-1}{k} = C_n^k. \quad (2.16)$$

Selbstverständlich ist nach (2.14)

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1}$$

Beispiel 2.7.1. Im obigen Fall haben wir 2 Kugeln ($k = 2$) in 4 Urnen ($n = 4$) geworfen. Die Anzahl der unterschiedlichen Möglichkeiten, 2 Kugeln in 4 Urnen zu werfen, ist nach (2.16) gleich

$$C_4^2 = \binom{4+2-1}{2} = \binom{5}{2} = 10.$$

Daher ist die Anzahl der Kombinationen von 4 Elementen zweiter Ordnung mit Wiederholung gleich 10. Die einzelnen Fälle haben wir oben mit den Buchstaben A, B, C, D ausgezählt. ■

Beispiel 2.7.2. Wie viele Kombinationen zweiter Ordnung von 6 Buchstaben lassen sich bilden, wenn Wiederholung erlaubt ist?

Lösung: $C_6^2 = \binom{6+2-1}{2} = \binom{7}{2} = 21.$ ■

Man sieht, daß es sich hier, genauso wie bei den einfachen Kombinationen, um eine Wahl handelt, wobei man aus einer Menge von n Elementen k wählt, jedes der n Elemente aber beliebig oft vorhanden ist. Es ist daher möglich, daß n kleiner als k ist, was bei den einfachen Kombinationen nicht der Fall ist.

Beispiel 2.7.3. (a) Wir betrachten Wortpaare einer Sprache, die 5 Kasus aufweist. Wie viele Kasuskonfigurationen gibt es?

(b) Wie viele Genuskonfigurationen mit 5 Wörtern gibt es, wenn 2 Genera unterschieden werden?

Lösung: (a) Jeder Kasus kann mit jedem anderen zusammen vorkommen, und beide Wörter können denselben Kasus haben. Daher sind es Kombinationen mit Wiederholung, und zwar

$$C_5^2 = \binom{5+2-1}{2} = \binom{6}{2} = 15.$$

(b) Hier ist $n = 2$, $k = 5$, daher

$$C_2^5 = \binom{2+5-1}{5} = \binom{6}{5} = 6.$$

Diese 6 Fälle kann man auch leicht aufzählen, nämlich entweder haben alle 5 Wörter das gleiche Genus; oder 4 haben eines und 1 das andere Genus; oder 3 haben das eine und zwei das andere Genus usw. bis zu keinem Wort des ersten und 5 Wörtern des anderen Genus, insgesamt 6, wobei die Reihenfolge belanglos ist. ■

Übung 2.7.1. Berechnen Sie C_7^2 , C_4^3 .

Übung 2.7.2. Auf wie viele Weisen können 3 Briefe in 5 Postkästen geworfen werden?

Übung 2.7.3. Stil wird oft als die Auswahl von Ausdrücken aus mehreren Möglichkeiten betrachtet. Ein Dichter teilte die Ausdrücke der Sprache in drei Kategorien und wollte in jeder Strophe seiner Gedichte 6 Ausdrücke kombiniert verwenden. Wie viele Strophen kann er schreiben, ohne daß sich die Kombination der drei Kategorien wiederholt?

2.8. Die Binomialkoeffizienten

Dem Leser ist sicherlich noch aus der Schule bekannt, daß

$$(b+a)^2 = a^2 + 2ab + b^2$$

$$(b+a)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(b+a)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

usw. Es interessiert uns, wie dieser Ausdruck allgemein aussieht, wenn der Exponent n ist ($n \in \mathbb{N}$).

Man sieht, daß die Exponenten von a schrittweise abnehmen, die Exponenten von b zunehmen und die Summe der Exponenten von a und b immer n ist. Z.B. bei $n = 3$ haben wir oben

$$a^3, a^2b^1, a^1b^2, b^3,$$

so daß man schreiben kann

$$a^k b^{n-k}.$$

Die Frage ist nur, wie man die Koeffizienten bei diesen Ausdrücken bestimmt. Wir erläutern das an einem konkreten Beispiel, wenn $n = 4$ ist. Es ist

$$(b+a)^4 = (b+a)(b+a)(b+a)(b+a)$$

Wir wissen, daß wir als Glieder der Entwicklung dieses Binoms

$$a^4, a^3b, a^2b^2, ab^3, b^4$$

bekommen. Den Koeffizienten bei z.B. a^3b berechnen wir so, daß wir ein b in einer Klammer fixieren und es mit den 3 a s in den Klammern multiplizieren. Das eine b können wir aus 4 Klammern wählen, und zwar nach 2.6 auf $\binom{4}{1}$ Weisen, denn das ist die Anzahl der Wahlmöglichkeiten für 1 Element aus 4. Daher ist der Koeffizient bei a^3b gleich $\binom{4}{1} = \binom{4}{3} = 4$. Der Koeffizient bei a^2b^2 berechnet sich so, daß man 2 b s in 2 Klammern fixiert und mit den a s in den restlichen 2 Klammern multipliziert. Das darf man aber für alle Wahlen von 2 b s tun. Man

kann nun 2 b s aus 4 genau auf $\binom{4}{2}$ Weisen wählen, daher ist der Koeffizient bei a^2b^2 gleich $\binom{4}{2} = 6$.

Daraus folgt allgemein, daß der Koeffizient bei

$$a^k b^{n-k}$$

gleich

$$\binom{n}{k} = \binom{n}{n-k}$$

ist. Aus dieser Tatsache leitet sich auch der Name "Binomialkoeffizient" ab. Allgemein kann man also die Entwicklung des Binoms schreiben

$$\begin{aligned} (b+a)^n &= \binom{n}{0} a^0 b^n + \binom{n}{1} a^1 b^{n-1} + \dots + \binom{n}{k} a^k b^{n-k} + \dots + \binom{n}{n} a^n b^0 = \\ &= \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}. \end{aligned} \quad (2.17)$$

Beispiel 2.8.1. Wir überzeugen uns, ob die Formel (2.17) für $(b+a)^3$ stimmt.

$$\text{Lösung: } (b+a)^3 = \binom{3}{0} a^0 b^3 + \binom{3}{1} a^1 b^2 + \binom{3}{2} a^2 b^1 + \binom{3}{3} a^3 b^0$$

$$= b^3 + 3ab^2 + 3a^2b + a^3.$$

Ordnet man die einzelnen Glieder nach dem abnehmenden Exponenten von a , so bekommt man die bekannte Formel. ■

Beispiel 2.8.2. Wie ist der Koeffizient bei a^9b^2 ?

Lösung: Es ist $n = 9 + 2 = 11$ und $k = 9$ (oder 2), daher

$$\binom{11}{2} = \binom{11}{9} = 55$$

und derselbe Koeffizient ist auch bei a^2b^9 . ■

Die Binomialkoeffizienten sind in Tabelle A2 des Anhangs angeführt. Es lassen sich für sie viele Beziehungen beweisen (vgl. Netto 1927), wir werden

Tabelle 2.1. Pascalsches Dreieck

0	1	1	1	1	1	1	...
1	1	2	3	4	5	6	...
2	1	3	6	10	15	21	...
3	1	4	10	20	35	56	...
4	1	5	15	35	70	126	...
5	1	6	21	56	126	252	...
...	...	...	...	...	...	...	...
	1	2	3	4	5	6	...

nur einige veranschaulichen. Tabelle 2.1 ist bekannt unter dem Namen *Pascalsches Dreieck*. Das ursprüngliche Dreieck erhält man so, daß man die Tabelle auf ihre rechte untere Ecke stellt. Die Zahlen in der Tabelle bekommt man so, daß man in die 0-te Zeile lauter Einsen einschreibt, und in den weiteren Zeilen enthält jede Zelle eine Zahl, die die Summe aller Zahlen in der darüberliegenden Zeile darstellt, wenn man sie von links bis einschließlich zu der gegebenen Zelle summiert. Beispielsweise ist in der dritten Zeile $1 + 4 + 10 + 20 = 35$, das eben unter 20 steht. Sucht man $\binom{n}{k}$, d.h. die Zahl der Kombinationen ohne Wiederholung, dann bestimmt man n und k auf der linken Seite der Tabelle, und zwar ist k die kleinere Zahl (höher liegende Zeile), n die größere (niedriger liegende Zeile). Man findet $\binom{n}{k}$ so, daß man von k waagerecht in die Tabelle und bei n von der 1 in der ersten Spalte diagonal von links nach rechts hinauf in die Tabelle eine Gerade zieht und die Zelle aufsucht, wo sie sich kreuzen.

Beispiel 2.8.2. Finden Sie $\binom{5}{2}$.

Lösung: $k = 2$ (Zeile 2), $n = 5$, d.h.

$$k = 2 \quad \text{---} \quad C_n^k = \binom{n}{k} = \binom{5}{2} = 10. \blacksquare$$

$$n = 5 \quad \nearrow$$

Liest man jedoch die Zahlen an der Kreuzung einer Zeile und einer Spalte ab, so bekommt man die Anzahl der Kombinationen von n Elementen (Spalte) k -ter Ordnung (Zeile) mit Wiederholung.

Beispiel 2.8.3. Wir suchen C_4^2 .

Lösung: An der Kreuzung von Zeile 2 und Spalte 4 finden wir

$$k = 2 \quad \text{---} \quad C_4^2 = \binom{4+2-1}{2} = \binom{5}{2} = 10. \blacksquare$$

$$n = 4 \quad \nearrow$$

Einige leichtere Identitäten folgen in den Beispielen, Übungen und Aufgaben.

Beispiel 2.8.4. Beweisen Sie die Identität

$$\binom{n-1}{k} + \binom{n-1}{k-1} = \binom{n}{k}. \quad (2.18)$$

Lösung: Man kann die Identität leicht beweisen, indem man die Ausdrücke auf der linken Seite in Fakultäten ausschreibt, nämlich

$$\begin{aligned} \binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)!}{k! (n-k-1)!} + \frac{(n-1)!}{(k-1)! (n-k)!} \\ &= \frac{(n-k) (n-1)! + k (n-1)!}{k! (n-k)!} \\ &= \frac{(n-1)! (n-k+k)}{k! (n-k)!} \\ &= \frac{n (n-1)!}{k! (n-k)!} \\ &= \frac{n!}{k! (n-k)!} = \binom{n}{k}. \blacksquare \end{aligned}$$

In Tabelle 2.1 sieht man direkt, daß jede Zahl die Summe der Zahl über ihr und links von ihr ist, was die obere Identität ausdrückt.

Beispiel 2.8.5. Beweisen Sie, daß

$$\sum_{k=0}^n \binom{n}{k} = 2^n \text{ ist!} \quad (2.19)$$

Lösung: Setzt man in

$$\sum_{k=0}^n \binom{n}{k} a^k b^{n-k} = (b+a)^n$$

beide Zahlen $b = a = 1$, so wird

$$\sum_{k=0}^n \binom{n}{k} = (1+1)^n = 2^n. \blacksquare$$

Übung 2.8.1. Wie ist der Koeffizient bei $a^6 b^3$?

Übung 2.8.2. Entwickeln Sie vollständig $(b+a)^6$.

Übung 2.8.3. Beweisen Sie, daß

$$\sum_{k=0}^n (-1)^k \binom{n}{k} = 0. \quad (2.20)$$

[Setzen Sie in $(b+a)^n$ $b = 1$ und $a = -1$.]

2.9. Die Multinomialkoeffizienten

In 2.6 haben wir festgestellt, daß man aus n Elementen k auf $\binom{n}{k}$ Weisen wählen kann, und wir bezeichneten diese Zahlen als Binomialkoeffizienten. Dasselbe Resultat erhalten wir, wenn wir fragen, auf wie viele Weisen man n Elemente in zwei Gruppen aufteilen kann, wenn die erste k und die zweite $n-k$ Elemente enthalten soll. Denn für die erste Gruppe wählen wir aus n Elementen k auf $\binom{n}{k}$ Weisen und aus dem Rest, nämlich $n-k$ Elementen "wählen" wir die zweite Gruppe auf $\binom{n-k}{n-k}$ Weisen, also insgesamt

$$\binom{n}{k} \binom{n-k}{n-k} = \binom{n}{k}.$$

Diese Prozedur läßt sich auch für mehrere Gruppen durchführen. Will man n Elemente in 3 Gruppen aufteilen, von denen die erste k_1 , die zweite k_2 und die dritte $k_3 = n - k_1 - k_2$ Elemente enthält, dann wählt man für

- die erste Gruppe k_1 Elemente aus n auf $\binom{n}{k_1}$ Weisen

- die zweite Gruppe k_2 Elemente aus $n - k_1$ auf $\binom{n-k_1}{k_2}$ Weisen und
- die dritte Gruppe k_3 Elemente aus $n - k_1 - k_2$ auf $\binom{n-k_1-k_2}{k_3}$ Weisen,

was insgesamt

$$\binom{n}{k_1} \binom{n-k_1}{k_2} \binom{n-k_1-k_2}{k_3}$$

Möglichkeiten ergibt. Allgemein ergibt die Aufteilung von n Elementen in r Gruppen, wobei die erste k_1 Elemente hat, die zweite k_2 Elemente, ..., die r -te k_r Elemente, und wobei $k_1 + k_2 + \dots + k_{r-1} + k_r = n$, insgesamt also

$$\begin{aligned} & \binom{n}{k_1} \binom{n-k_1}{k_2} \binom{n-k_1-k_2}{k_3} \dots \binom{n-k_1-k_2-\dots-k_{r-1}}{k_r} = \\ &= \frac{n!}{k_1! (n-k_1)!} \cdot \frac{(n-k_1)!}{k_2! (n-k_1-k_2)!} \cdot \frac{(n-k_1-k_2)!}{k_3! (n-k_1-k_2-k_3)!} \cdot \dots \cdot \\ & \dots \cdot \frac{(n-k_1-k_2-\dots-k_{r-1})!}{k_r! (n-k_1-k_2-\dots-k_{r-1}-k_r)!} = \\ &= \frac{n!}{k_1! k_2! \dots k_r!} = P_n^{k_1, k_2, \dots, k_r} \end{aligned} \quad (2.7a)$$

Möglichkeiten. Diese Formel haben wir bereits in 2.3 mit einer anderen Überlegung abgeleitet.

Beispiel 2.9.1. In einer Sprache L drücken 12 (eindeutige) Affixe die folgenden grammatischen Bedeutungen aus: 2 Genera, 4 Numeri und 6 Kasus. Auf wie viele Weisen könnten die 12 Affixe auf diese drei Kategorien verteilt sein?

Lösung: Es gibt $n = 12$ Elemente und drei Gruppen mit $k_1 = 2$, $k_2 = 4$, $k_3 = 6$, so daß nach (2.7a)

$$\binom{12}{2} \binom{10}{4} \binom{6}{6} = \frac{12!}{2! 4! 6!} = 13860. \blacksquare$$

Genauso, wie wir von der Aufteilung einer Menge in zwei Gruppen durch Verallgemeinerung zur Aufteilung auf mehrere Gruppen gekommen sind und aus C_n^k dadurch $P_n^{k_1, k_2, \dots, k_r}$ bekamen, müßten wir dasselbe Resultat erhalten, wenn wir statt eines Binoms $(b+a)^n$, in dessen Entwicklung die Koeffizienten C_n^k waren, ein Multinom $(a+b+c+\dots+m)^n$ entwickeln, in dem wir die Multinomialkoeffizienten vorfinden können.

Nehmen wir einfachheitshalber nur drei Elemente, d.h. ein Trinom $(c+b+a)^n$ und betrachten $b+a$ als ein Glied, dann bekommen wir

$$\begin{aligned}(c+b+a)^n &= \binom{n}{0} (b+a)^0 c^n + \binom{n}{1} (b+a)^1 c^{n-1} + \dots + \\ &+ \dots + \binom{n}{k} (b+a)^k c^{n-k} + \dots + \\ &+ \dots + \binom{n}{n} (b+a)^n c^0.\end{aligned}\quad (2.21)$$

Das allgemeine Glied ist hier

$$\binom{n}{k} (b+a)^k c^{n-k}.\quad (2.22)$$

In diesem entwickeln wir $(b+a)^k$ und bekommen

$$(b+a)^k = \binom{k}{0} a^0 b^k + \binom{k}{1} a^1 b^{k-1} + \dots + \binom{k}{r} a^r b^{k-r} + \dots + \binom{k}{k} a^k b^0.$$

Hier ist das allgemeine Glied

$$\binom{k}{r} a^r b^{k-r}$$

Wir setzen es in (2.22) ein und erhalten

$$\binom{n}{k} c^{n-k} \binom{k}{r} a^r b^{k-r} = \frac{n!}{k! (n-k)!} \cdot \frac{k!}{r! (k-r)!} a^r b^{k-r} c^{n-k}.$$

In diesen Ausdruck setzen wir

$$k-r=s$$

$$n-k=t,$$

so daß wir schließlich das allgemeine Glied der Entwicklung in (2.21) in der Form

$$\frac{n!}{r! s! t!} a^r b^s c^t \quad (2.23)$$

bekommen. Offensichtlich ist die Summe solcher Ausdrücke eben $(a+b+c)^n$, d.h.

$$(a+b+c)^n = \sum_{\substack{r,s,t=0 \\ r+s+t=n}}^n \frac{n!}{r! s! t!} a^r b^s c^t, \quad (2.24)$$

Beispiel 2.9.2. Entwickeln Sie $(a+b+c)^3$.

$$\begin{aligned}\text{Lösung: } (a+b+c)^3 &= \frac{3!}{3! 0! 0!} a^3 b^0 c^0 + \frac{3!}{0! 3! 0!} a^0 b^3 c^0 + \frac{3!}{0! 0! 3!} a^0 b^0 c^3 + \\ &+ \frac{3!}{2! 1! 0!} a^2 b^1 c^0 + \frac{3!}{2! 0! 1!} a^2 b^0 c^1 + \frac{3!}{1! 2! 0!} a^1 b^2 c^0 + \\ &+ \frac{3!}{1! 0! 2!} a^1 b^0 c^2 + \frac{3!}{0! 2! 1!} a^0 b^2 c^1 + \frac{3!}{0! 1! 2!} a^0 b^1 c^2 + \\ &+ \frac{3!}{1! 1! 1!} a^1 b^1 c^1 = \\ &= a^3 + b^3 + c^3 + 3a^2b + 3a^2c + 3ab^2 + 3ac^2 + 3b^2c + \\ &+ 3bc^2 + 6abc.\end{aligned}$$

Der Ausdruck (2.24) läßt sich auf 4, 5 oder mehr Elemente erweitern, so daß man allgemein bekommt

$$(a_1 + a_2 + a_3 + \dots + a_m)^n = \sum_{\substack{r_1, \dots, r_m=0 \\ r_1 + \dots + r_m=n}}^n \frac{n!}{r_1! r_2! r_3! \dots r_m!} \cdot a_1^{r_1} a_2^{r_2} a_3^{r_3} \dots a_m^{r_m}. \quad (2.25)$$

Aus (2.23) ist ersichtlich, daß die Glieder der Entwicklung eines Multinoms aus einer Kombination mit Wiederholung der Elemente a, b, c und einem dazu gehörigen Multinomialkoeffizienten bestehen. Das läßt sich anschaulich an einem Binom, z.B. bei

$$(b + a)^3 = a^3 + 3a^2b + 3ab^2 + b^2$$

zeigen, das aus $\begin{bmatrix} aaa & aab & abb & bbb \\ 1 & 3 & 3 & 1 \end{bmatrix}$ besteht.

Das bedeutet, daß die Anzahl der Glieder der Entwicklung eines Multinoms durch die Anzahl der Kombinationen mit Wiederholung der Elemente des Multinoms zur Ordnung des Exponenten gegeben ist.

Beispiel 2.9.3. Wie viele Glieder hat $(a + b + c)^2$?

Lösung: Die Anzahl der Elemente (a, b, c) ist $m = 3$, die Ordnung ist $n = 2$, also ist die Anzahl der Glieder gleich

$$C_m^n = C_3^2 = \binom{3+2-1}{2} = \binom{4}{2} = 6.$$

Die Glieder sind bekanntlich $a^2, b^2, c^2, 2ab, 2ac, 2bc$. Bei höheren Ordnungen ist die Aufzählung etwas langwieriger. ■

Beispiel 2.9.4. Wie viele Glieder hat $(a + b + c)^4$, und wie sind die Koeffizienten bei a^3b, ab^2c, c^4 ?

Lösung: Hier ist $m = 3, n = 4$, also

$$C_3^4 = \binom{3+4-1}{4} = \binom{6}{4} = 15 \text{ Glieder.}$$

Die Koeffizienten sind:

$$\text{bei } a^3b: \frac{4!}{3! 1! 0!} = 4$$

$$\text{bei } ab^2c: \frac{4!}{1! 2! 1!} = 6$$

$$\text{bei } c^4: \frac{4!}{0! 0! 4!} = 1. \blacksquare$$

Beispiel 2.9.5. Zeigen Sie, daß bei den Gliedern der Entwicklung eines Multinoms von m Elementen die Summe der Multinomialkoeffizienten $\sum \frac{n!}{r! s! t! \dots z!}$ gleich m^n ist.

Lösung: Ersetzt man in (2.25) $a = b = \dots = m = 1$, so wird

$$(1 + 1 + \dots + 1)^n = m^n = \sum \frac{n!}{r! s! t! \dots z!}. \blacksquare \quad (2.26)$$

Übung 2.9.1. Wie viele Glieder hat $(a + b + c + d)^5$?

Übung 2.9.2. Wie ist der Koeffizient bei a^4bc^2 ?

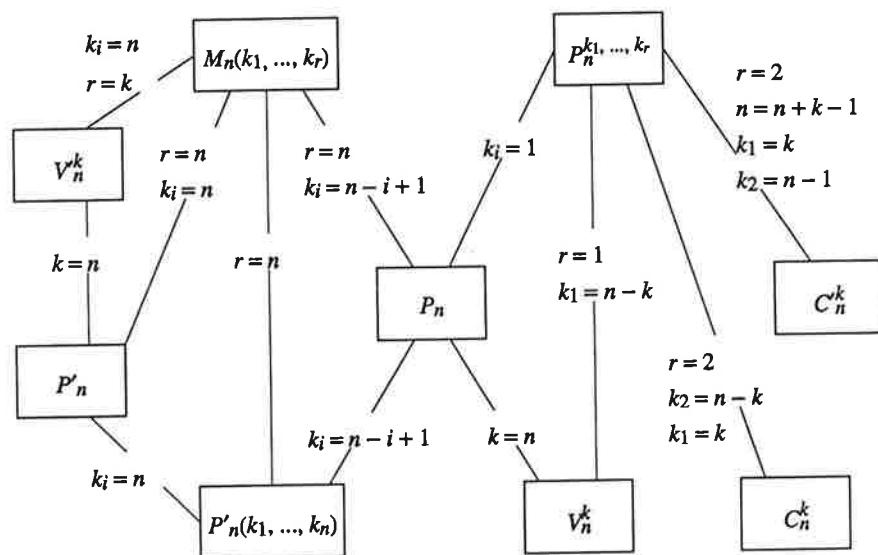
Übung 2.9.3. Auf wie viele Weisen kann man 20 Sprachen in 4 gleich große Gruppen einteilen?

Übung 2.9.4. Auf wie viele Weisen kann man 15 Konsonanten einer Sprache in 3 Gruppen einteilen, wenn in der ersten 10, in der zweiten 3 und in der dritten 2 Konsonanten sein sollen?

Zusammenfassung der Formeln

Permutation ohne Wiederholung	$P_n = n!$
Permutation mit Wiederholung	$P'_n = n^n$
Permutation mit beschränkter Wiederholung	$P'_n(k_1, k_2, \dots, k_n) = \prod_{i=1}^n k_i$
Permutation von Elementen, von denen k_1 von einer Sorte, k_2 von einer zweiten Sorte, ..., k_r von einer r -ten Sorte sind	$P_n^{k_1, k_2, \dots, k_r} = \frac{n!}{k_1! k_2! \dots k_r!}$
Variation ohne Wiederholung	$V_n^k = \frac{n!}{(n-k)!}$
Variation mit Wiederholung	$V_n^k = n^k$
Kombination ohne Wiederholung	$C_n^k = \binom{n}{k}$
Kombination mit Wiederholung	$C_n^k = \binom{n+k-1}{k}$
Multiplikationsregel	$M_n(k_1, k_2, \dots, k_r) = \prod_{i=1}^r k_i$

Zusammenhänge der Formeln



Aufgaben

- 2.1. Berechnen Sie P_6 , P'_6 , V_6^2 , $V_6'^2$, C_6^2 , $C_6'^2$, $P_6^{2,3,1}$, $P_6^{2,2,2}$, $M_6(2,3,1)$, $P'_6(6,1,2,6,4,3)$.
- 2.2. Wie viele Wörter lassen sich durch die Umstellung der Buchstaben im Wort BUCHSTABE bilden?
- 2.3. Zeigen Sie, daß $P_{k_1} P_n^{k_1, k_2} = V_n^{k_1} (k_1 + k_2 = n)$.
- 2.4. Auf wie viele Weisen kann man einen Lottoschein ausfüllen, wenn man aus 49 Zahlen 6 ankreuzen muß?
- 2.5. Wie viele Morpheme des Typs KVKV kann man bilden, wenn man 14 Konsonanten und 5 Vokale hat?
- 2.6. Zeigen Sie, daß $C_2^k = k + 1$ ist.
- 2.7. Wie viele vierbuchstabile Wörter kann man aus den Buchstaben des Wortes STEIN bilden, wenn
 - (a) Wiederholung nicht erlaubt ist,
 - (b) Wiederholung erlaubt ist?

2.8. Zur Lösung von 5 zusammenhängenden linguistischen Problemen kann man 3 Grammatikmodelle anwenden. Auf wie viele Weisen kann man die Wahl der Modelle treffen, wenn jedes Modell für die Lösung jedes Problems geeignet ist und die Reihenfolge der Anwendung irrelevant ist?

2.9. Zeigen Sie, daß $\frac{V_n^k}{P_n^k} = C_n^k$ ist.

Kann man aus dieser Beziehung den mnemotechnischen Satz ableiten: "Die Variation ist die Permutation der Kombination" ?

2.10. Wie viele Symbolketten kann man durch Umstellung von DET N ADJ N ADV V ADV V bilden?

2.11. Wie ist der Koeffizient des Gliedes $a^4 b^4$ einer binomischen Entwicklung?

2.12. Auf wie viele Weisen kann man aus 7 Seminaren 4 wählen?

2.13. Wie viele Glieder hat

(a) $(q + p)^{40}$; (b) $(p + q + r)^6$?

2.14. Beweisen Sie, daß $C_n^{k-1} + C_n^k = C_{n+1}^k$. [Schreiben Sie in Fakultäten aus und ordnen Sie dann.]

2.15. Betrachten Sie den Namen SCHMIDT.

- (a) Wie viele Wörter derselben Länge kann man aus seinen Buchstaben durch Umstellung bilden?
- (b) Wie viele Wörter davon fangen mit I an?
- (c) Wie viele fangen nicht mit I an?
- (d) In wie vielen Wörtern steht I zwischen Konsonanten, d.h. weder am Anfang noch am Ende?
- (e) Wie viele Wörter von 3 Buchstaben kann man aus dem Namen bilden, wenn Wiederholung nicht erlaubt ist?
- (f) Wie viele Wörter von 3 Buchstaben kann man bilden, wenn die Anordnung der Buchstaben irrelevant ist?
- (g) In wie vielen Wörtern von (f) steht ein I ?
- (h) Wenn die Buchstaben wiederholt werden dürfen und ihre Anordnung relevant ist, wie viele Wörter von 3 Buchstaben kann man bilden ?
- (i) Wie viele Wörter in (h) haben I in der Mitte?
- (j) In wie vielen Wörtern von (h) ist mindestens ein I ?
- (k) Wie viele Dreibuchstabenwörter kann man bilden, wenn Wiederholung erlaubt, die Anordnung aber irrelevant ist?
- (l) In wie vielen Wörtern von (k) befindet sich genau ein I?
- (m) In wie vielen Wörtern von (k) befindet sich kein I ?
- (n) Wie viele Wörter von (k) enthalten mindestens ein I ?

2.16. Zeigen Sie, daß $\frac{P_n}{P'_n} = \prod_{k=0}^{n-1} (1 - \frac{k}{n})$ [ausschreiben und gliedweise dividieren].

- 2.17. Im Indonesischen gibt es 6 Vokale und 23 Konsonanten, aus denen folgende Morphemtypen von höchstens 3 Phonemen gebildet werden können:

V	6	VV	2	VVK	14
		VK	36	VKV	82
		KV	36	VKK	7
			74	KVV	32
				KVK	39
				KKV	9
					183

Berechnen Sie, welche Proportion aller Möglichkeiten für die jeweilige Länge (L_i , $i = 1, 2, 3$) realisiert wurde. Der Ausnutzungsgrad der Morphem- oder Worttypenbildung ließe sich daher durch eine Kurve darstellen (vgl. auch Herdan 1966, 198-201). [Dividieren Sie die beobachtete Zahl der Möglichkeiten für die gegebene Länge L_1, L_2, L_3 . Nicht realisierte Typen wie KK, VVV u.a. werden nicht berücksichtigt.]

- 2.18. Überzeugen Sie sich ohne Beweis, daß aus Tabelle 2.1 folgt

$$\binom{n}{k+1} = \binom{k}{k} + \binom{k+1}{k} + \dots + \binom{n-2}{k} + \binom{n-1}{k} = \sum_{r=k}^{n-1} \binom{r}{k}$$

- 2.19. Ein Dichter hat 3 Adjektive (A), 2 Substantive (S) und 4 Verben (V) zur Verfügung.

- Wie viele Verse kann er schreiben, wenn in jedem Vers nur 3 Wörter stehen, in der Reihenfolge ASV?
- Wie viele Verse kann er schreiben, wenn in jedem Vers 3 Wörter stehen, ein Wort aus jeder Wortart, wobei die Reihenfolge frei ist?
- In (c) bis (f) werden Adjektive untereinander, Verben untereinander und Substantive untereinander nicht unterschieden. Wie viele unterschiedliche Dreiwörterverse kann er bilden, wenn die Reihenfolge ohne Belang ist, aber die Wortarten wiederholt auftreten dürfen?
- Wie viele Verse gibt es, wenn er die Wortarten nicht wiederholen darf und die Anordnung irrelevant ist?
- Wie viele Verse gibt es, wenn die Anordnung relevant ist und er die Wortarten wiederholen darf?
- Wie viele Verse gibt es, wenn die Anordnung relevant ist und die Wortarten nicht wiederholt werden dürfen?

- 2.20. Zeigen Sie, daß $\frac{V_6^2 P_8 C_6^2}{C_8^4 P_6^2 P_4 P_6} = 1$ ist. [Schreiben Sie in Fakultäten aus.]

- 2.21. Unter welchen Bedingungen gilt:

- $V_n^k = P_n'$
- $V_n^k = P_n$
- $P_n^{k_1, k_2, \dots, k_r} = V_n^k$
- $P_n^{k_1, k_2, \dots, k_r} = C_n^k$
- $P_n^{k_1, k_2, \dots, k_r} = P_n$

- 2.22. Manche Politiker gewöhnen sich für ihre Reden eine spezielle Verteilung von Sprechpausen innerhalb der Sätze an. Wie viele Möglichkeiten hat man, in einem Satz

- von 6 Wörtern 2 Sprechpausen
- von 10 Wörtern 3 Sprechpausen zwischen den Wörtern zu machen?

- 2.23. Die Wörter einer bestimmten Sprache, die 3 oder mehr Silben lang sind, haben immer einen Haupt- und einen Nebenzakzent.

- Auf wie viele Weisen kann man ein 3-silbiges Wort betonen?
- Auf wie viele Weisen kann man ein 3-silbiges Wort betonen, wenn die zwei Akzente nicht nebeneinander stehen dürfen?
- Berechnen Sie die Möglichkeiten von (a) und (b) für 4-, 5-, 6-, 7-, 8- und 9-silbige Wörter und stellen Sie eine allgemeine Formel auf.

- 2.24. Ein Dichter schreibt Verse, in denen immer 4 schwere, betonte Silben standen. Vor der ersten schweren Silbe dürfen 0, 1, 2 oder 3 leichte Silben stehen, genauso hinter der letzten schweren Silbe, und zwischen ihnen müssen 1, 2 oder 3 leichte Silben stehen. Wie viele metrisch unterschiedliche Verse sind unter diesen Bedingungen möglich?

- 2.25. Beweisen Sie, daß $V_n^2 + V_n^1 = n^2$ ist. [Schreiben Sie in Fakultäten aus und benutzen Sie Formel (2.3).]

- 2.26. Beweisen Sie, daß $\sum_{k=0}^n k \binom{n}{k} = n2^{n-1}$.

- 2.27. Beweisen Sie, daß $2 \cdot 4 \cdot 6 \cdot 8 \cdot \dots \cdot 2n = 2^n n!$ [Klammern Sie aus jedem Multiplikanden 2 aus.]

- 2.28. Beweisen Sie, daß $1 \cdot 3 \cdot 5 \cdot 7 \cdot \dots \cdot (2n-1) = \frac{(2n)!}{2^n n!}$ [Multiplizieren und dividieren Sie mit dem Ausdruck in der Aufgabe 2.27.]

- 2.29. Folgern Sie aus den Aufgaben 2.27 und 2.28, daß

$$\frac{2 \cdot 4 \cdot 6 \cdot \dots \cdot 2n}{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)} = \frac{(2^n n!)^2}{(2n)!}$$

- 2.30. Beweisen Sie, daß $n! = n^n \prod_{k=0}^{n-1} (1 - \frac{k}{n})$ ist. Vgl. Aufgabe 2.16; multiplizieren und dividieren Sie $n!$ mit n^n .

- 2.31. Zeigen Sie, daß $(2n)! = 2^n \prod_{k=1}^n (2k-1) \prod_{m=1}^n m$ ist. [Vgl. Aufgaben 2.27 und 2.28.]

2.32. Beweisen Sie, daß $\frac{(n! 2^n)^2}{(2n)!} = \frac{\left(\frac{n!}{n^n}\right)^2}{\prod_{k=0}^{2n-1} \left(1 - \frac{k}{2n}\right)}$ ist.

[Multiplizieren und dividieren Sie die linke Seite mit n^{2n} und ordnen Sie dann; vgl. Aufgabe 2.16.]

2.33. Beweisen Sie, daß $n \prod_{k=2}^n \left(1 - \frac{1}{k}\right) = 1$. [Schreiben Sie das Produkt explizit aus.]

Benutzte und weiterführende Literatur: Feller (1962), Harris (1966), Herdan (1966), Hogben (1956), Mosteller-Rourke-Thomas (1961), Netto (1927), Neyman (1950), Whitworth (1965).

3. Wahrscheinlichkeit

3.1. Definition der Wahrscheinlichkeit

In vielen Bereichen der Wissenschaft und des alltäglichen Lebens haben wir mit Ereignissen zu tun, deren Auftreten sich durch keine deterministischen Gesetzmäßigkeiten genau bestimmen läßt. Man sagt, daß ihr Vorkommen vom Zufall abhängt, und nennt sie *zufällige Ereignisse*. Es ist gleichgültig, ob die Ereignisse tatsächlich zufällig sind, ob sie uns wegen ihrer Unkenntnis als unregelmäßig erscheinen oder ob ihr Verhalten von einer unübersichtlichen Anzahl von Faktoren abhängt, die uns zu einem gegebenen Zeitpunkt nicht alle bekannt sind. Wir suchen lediglich nach Mitteln, mit denen wir sie beschreiben können. Die natürliche Sprache kennt eine Fülle solcher Ereignisse, bei denen es sich nur um Tendenzen handelt, die man lediglich mit dem Apparat der Wahrscheinlichkeitsrechnung erfassen kann. So stellt alles, was man unter dem Begriff "Sprachnorm" zusammenfaßt, lediglich eine Menge von Tendenzen dar, die keinem deterministischen Gesetz folgen. Findet man jedoch eine vollständig determinierte Erscheinung in der Sprache, z.B. in der Deklination, so kann man sie als Extremfall einer Tendenz betrachten.

Im weiteren werden wir Ereignisse mit Großbuchstaben $A, B, \dots$ bezeichnen und sie als Punkte im Ereignisraum S auffassen, der sie alle umfaßt. Beispielsweise bilden die 6 möglichen Ereignisse beim Wurf eines Würfels einen Ereignisraum. Die Anzahl dieser Fälle bezeichnen wir als N_A, N_B usw. und die Zahl aller möglichen Fälle (die Zahl der Elemente im ganzen Ereignisraum) als N .

- Ein *sicheres Ereignis* kommt mit absoluter Sicherheit vor.
- Ein *unmögliches Ereignis* kann nicht vorkommen.
- Zwei Ereignisse A und B *schließen sich gegenseitig aus*, wenn sie nicht gleichzeitig eintreten können. Wir nennen sie *disjunkt*.
- *Äquivalent* sind zwei Ereignisse, die immer zusammen auftreten.

- Die Wahrscheinlichkeit des Ereignisses A werden wir als $P(A)$ bezeichnen.

Man kann die Wahrscheinlichkeit als den Grad unserer Kenntnis betrachten oder als Neigung der Dinge zu einem Zustand, man kann sie entweder klassisch oder statistisch oder durch Axiome definieren. Wir werden uns nur mit den drei letzten Möglichkeiten beschäftigen.

Statistisch wird die Wahrscheinlichkeit als Limes der relativen Häufigkeit des gegebenen Ereignisses definiert, d.h. als

$$P(A) = \lim_{N \rightarrow \infty} \frac{N_A}{N}. \quad (3.1)$$

Diese Definition ist zwar praktisch, stößt aber auf mathematische Schwierigkeiten.

Beispiel 3.1.1. (Job 1980) In einer Stichprobe von 1000 georgischen Phonemen kommt /a/ 166 mal vor, daher beträgt seine relative Häufigkeit

$$\frac{N_a}{N} = \frac{166}{1000} = 0,16600.$$

Unter $N = 2000$ Phonemen ist die relative Häufigkeit von /a/

$$\frac{N_a}{N} = \frac{331}{2000} = 0,16550.$$

Unter $N = 3000$ ist sie

$$\frac{N_a}{N} = \frac{477}{3000} = 0,15900.$$

Unter $N = 7000$ ist sie

$$\frac{N_a}{N} = \frac{1178}{7000} = 0,16829.$$

Unter $N = 13000$ ist sie

$$\frac{N_a}{N} = \frac{2064}{13000} = 0,15877.$$

Unter $N = 49000$ ist sie

$$\frac{N_a}{N} = \frac{7879}{49000} = 0,16080.$$

Unter $N = 50000$ ist sie

$$\frac{N_a}{N} = \frac{8051}{50000} = 0,16102.$$

Die relative Häufigkeit stabilisiert sich allmählich mit wachsendem Stichprobenumfang, und bei hohem N ändern weitere Erhebungen wenig am Resultat. Zwar ist in jeder Stichprobe die relative Häufigkeit von /a/ etwas anders, aber der Grenzwert von N_a/N ist eine gute Schätzung der wahren Wahrscheinlichkeit, mit der /a/ im Georgischen vorkommt. ■

Klassisch definiert man die Wahrscheinlichkeit einfach als

$$P(A) = \frac{N_A}{N}, \quad (3.2)$$

d.h. als das Verhältnis der Zahl der günstigen Fälle zur Anzahl aller möglichen Fälle. Diese Definition ist zwar mit verschiedenen Mängeln behaftet, die sich in konkreten Fällen aufzeigen lassen, sie leistet aber gute Dienste bei der Lösung vieler Probleme.

Beispiel 3.1.2. In einer Urne befinden sich 20 Lose, davon sind 5 Gewinne und 15 Nieten.

(a) Man zieht ein Los. Wie ist die Wahrscheinlichkeit $P(A)$, daß man einen Gewinn gezogen hat?

(b) Man zieht 2 Lose. Wie ist die Wahrscheinlichkeit $P(A)$, einen Gewinn und eine Niete zu ziehen?

Lösung: (a) Es gibt insgesamt $N = 20$ Möglichkeiten, ein Los zu ziehen; darunter sind $N_A = 5$ Fälle günstig. Daher ist

$$P(A) = \frac{5}{20} = 0,25.$$

(b) Es gibt $N = \binom{20}{2}$ Möglichkeiten, 2 Lose aus 20 zu ziehen. Dabei existieren $\binom{5}{1}$ Möglichkeiten, aus 5 Gewinnen einen Gewinn zu ziehen, und $\binom{15}{1}$ Möglichkeiten, aus 15 Nieten eine zu wählen. Daher ist

$$N_A = \binom{5}{1} \binom{15}{1}, \text{ woraus folgt, daß}$$

$$P(A) = \frac{\binom{5}{1} \binom{15}{1}}{\binom{20}{2}} = \frac{5 \cdot 15}{190} = 0,3947. \blacksquare$$

Beispiel 3.1.3. Die Buchstaben des Wortes SEMANTIK werden auf separate Zettel geschrieben, und diese werden in eine Urne geworfen. Man zieht die Zettel nacheinander aus der Urne heraus und legt sie in eine Reihe auf den Tisch. Wie ist die Wahrscheinlichkeit $P(A)$, daß man wieder das ursprüngliche Wort bekommt?

Lösung: Man kann die Buchstaben in allen möglichen Reihenfolgen ziehen. Ihre Anzahl beträgt bekanntlich $N = 8!$. Es gibt nur eine günstige Reihenfolge, d.h. $N_A = 1$, woraus

$$P(A) = \frac{1}{8!} \text{ folgt.}$$

Alternativer Zugang: Sollen die Buchstaben in der richtigen Reihenfolge gezogen werden, dann ist die Wahrscheinlichkeit, beim ersten Zug S zu ziehen, $\frac{1}{8}$; ist S bereits gezogen, so ist die Wahrscheinlichkeit, E zu ziehen, $\frac{1}{7}$ usw. Daher ist

$$P(A) = \frac{1}{8} \cdot \frac{1}{7} \cdot \frac{1}{6} \cdot \frac{1}{5} \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{8!}. \blacksquare$$

Beispiel 3.1.4. Wir gehen von der gleichen Situation wie im vorigen Beispiel aus, legen aber jetzt nach jedem Zug den gezogenen Zettel zurück. Wie ist $P(A)$, daß man nach 8 Zügen die Buchstaben in der Reihenfolge S, E, M, A, N, T, I, K gezogen hat?

Lösung: Beim Zurücklegen ist die Zahl aller Möglichkeiten $N = 8^8$, weil man bei jedem Zug alle Buchstaben ziehen kann. Es gibt wieder nur einen günstigen Fall, also ist

$$P(A) = \frac{1}{8^8}. \blacksquare$$

Beispiel 3.1.5. Von acht Texten werden drei nach dem Zufallsprinzip für eine Untersuchung ausgewählt. Wie groß ist $P(A)$, daß sich der älteste Text in der Wahl befindet?

Lösung: Man kann aus acht Texten Dreiergruppen auf V_8^3 Weisen bilden, daher $N = 8!/5!$. Wenn sich der älteste Text in der Wahl befindet, so kann man aus den restlichen 7 Texten Zweiergruppen auf V_7^2 Weisen bilden, daher ist $N_A = 7!/5!$. Daraus ergibt sich

$$P(A) = \frac{7!/5!}{8!/5!} = \frac{1}{8}. \blacksquare$$

Übung 3.1.1. In einem Zettelkasten befinden sich ungeordnet 26 Karteikarten. Man zieht nun nacheinander die Karten heraus. Wie ist die Wahrscheinlichkeit, daß man sie in alphabetischer Reihenfolge oder in umgekehrter alphabetischer Reihenfolge zieht?

Übung 3.1.2. Wie ist die Wahrscheinlichkeit in Übung 3.1.1, wenn man 26 Züge macht und die Karten nach jedem Zug zurücklegt?

Übung 3.1.3. Die 5 Buchstaben des Wortes MESSE werden gemischt. Wie groß ist die Wahrscheinlichkeit, daß man eine Reihenfolge bekommt, in der gleiche Buchstaben nebeneinander stehen?

Übung 3.1.4. Vor einem Postamt stehen 5 Postkästen. 3 Personen werfen jeweils einen Brief in einen beliebigen Postkasten. Wie ist die Wahrscheinlichkeit, daß alle 3 Briefe in einem Postkasten liegen?

3.2. Axiome der Wahrscheinlichkeit

Beim axiomatischen Aufbau der Wahrscheinlichkeitsrechnung bedient man sich einiger Begriffe der Mengenlehre, die wir hier kurz darlegen werden. Mit dem Buchstaben S werden wir die Allmenge bezeichnen, mit $\emptyset$ die leere Menge und mit A, B beliebige nichtleere Mengen.

Teilmenge: Wenn alle Elemente der Menge A auch zu der Menge B gehören, dann ist A eine Teilmenge der Menge B , geschrieben $A \subset B$. Offenbar ist jede Menge eine Teilmenge von sich selbst, d.h. $A \subset A$. Die leere Menge ist Teilmenge jeder Menge.

Äquivalenz: Zwei Mengen A und B sind dann und nur dann äquivalent, wenn alle Elemente von A auch Elemente von B sind und umgekehrt, d.h. $A = B$ dann und nur dann, wenn $A \subset B$ und $B \subset A$.

Vereinigung oder Summe zweier Mengen A und B besteht aus allen Elementen, die zu A oder B oder zu beiden gehören. Für die Vereinigung der Mengen gelten das assoziative Gesetz

$$(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C$$

und das kommutative Gesetz

$$A \cup B = B \cup A.$$

Weiter ist

$$\begin{aligned} A \cup \emptyset &= A \\ A \cup S &= S \\ A \cup A &= A. \end{aligned} \tag{3.3}$$

Das *Produkt* oder der *Durchschnitt* zweier Mengen A und B , geschrieben AB oder $A \cap B$, besteht aus allen Elementen, die sowohl zu A als auch zu B gehören. Im Bild 3.1 wird das Produkt zweier Mengen dargestellt.

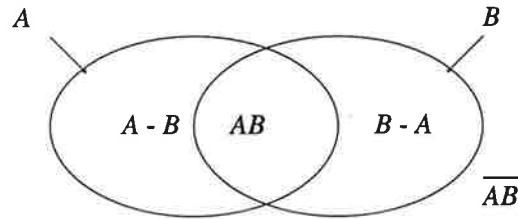


Abb. 3.1.

Für das Produkt gelten das *assoziative Gesetz*

$$(AB)C = A(BC),$$

das *kommutative Gesetz*

$$AB = BA$$

und das *distributive Gesetz*

$$A(B \cup C) = AB \cup AC. \quad (3.4)$$

Weiter gilt

$$\begin{aligned} A \emptyset &= \emptyset \\ A S &= A \\ A A &= A. \end{aligned} \quad (3.5)$$

Wenn zwei Mengen A und B keine gemeinsamen Elemente haben, dann sagt man, daß sie *disjunkt* sind. In dem Fall ist

$$AB = \emptyset$$

Das *Komplement* von A bezüglich S , bezeichnet als $\bar{A}$ ist die Menge, die alle Elemente von S enthält, die nicht zu A gehören. Daher ist

$$A \cup \bar{A} = S. \quad (3.6)$$

Weiter gilt, daß

$$\begin{aligned} \bar{\emptyset} &= S \\ \bar{S} &= \emptyset \\ \bar{\bar{A}} &= A \\ \overline{AA} &= \emptyset. \end{aligned} \quad (3.7)$$

Die *Differenz* zweier Mengen A und B besteht aus den Elementen, die zu A aber nicht zu B gehören. Im Bild 3.1 sieht man, daß

$$A - B = A - AB = A\bar{B}. \quad (3.8)$$

Weiter gilt

$$\begin{aligned} A - \emptyset &= A \\ A - S &= \emptyset \\ S - A &= \bar{A}. \end{aligned} \quad (3.9)$$

Weitere Regeln werden wir nach Bedarf einführen.

Wir werden S als den Ereignisraum betrachten und A als ein Ereignis in S . Dem Ereignis A schreiben wir die Zahl $P(A)$ zu, die wir *Wahrscheinlichkeitsfunktion* nennen und die drei Bedingungen oder Axiomen genügt, nämlich

Axiom I. $P(A) \geq 0$, d.h. $P(A)$ ist eine nichtnegative Zahl.

Axiom II. $P(S) = 1$. Hier bedeutet S das sichere Ereignis, d.h. daß überhaupt eines der möglichen Ereignisse aus dem Ereignisraum – egal welches – auftritt.

Axiom III. Sind die Ereignisse $A_1, A_2, \dots$ gegenseitig disjunkt, d.h. $A_i A_j = \emptyset$ für $i \neq j$, $i, j = 1, 2, \dots$, dann gilt

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

Aus den Axiomen folgen verschiedene Sätze, von denen wir einige beweisen werden.

(1) $P(\emptyset) = 0$.

Beweis: Nach (3.3) gilt $A = A \cup \emptyset$. Daraus folgt nach Axiom III, daß

$$P(A) = P(A \cup \emptyset) = P(A) + P(\emptyset) \Rightarrow P(A) - P(A) = P(\emptyset) \Rightarrow P(\emptyset) = 0. \blacksquare$$

(2) $P(\bar{A}) = 1 - P(A)$.

Beweis: Nach (3.6) gilt $A \cup \bar{A} = S$ und daraus folgt nach den Axiomen III und

II

$$P(A \cup \bar{A}) = P(S) \Rightarrow P(A) + P(\bar{A}) = 1 \Rightarrow P(\bar{A}) = 1 - P(A). \blacksquare$$

(3) $0 \leq P(A) \leq 1$.

Beweis: Die linke Ungleichung ist das Axiom I. Aus dem Satz (2) folgt $P(A) = 1 - P(\bar{A})$. Es ist aber laut Axiom I auch $P(\bar{A}) \geq 0$. Daher ist

$$P(A) = 1 - P(\bar{A}) \leq 1. \blacksquare$$

(4) Wenn $B \subset A$, dann $P(A) \geq P(B)$.

Beweis: Wenn $B \subset A$, dann ist $A = B + \bar{A}B$, wie aus dem Bild 3.2 ersichtlich ist, woraus

$$P(A) = P(B) + P(\bar{A}B) \geq P(B)$$

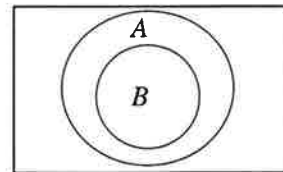


Abb. 3.2.

folgt. $\blacksquare$

Beispiel 3.2.1. Von acht Texten eines Autors wählt man zufällig drei aus. Wie groß ist die Wahrscheinlichkeit $P(A)$, daß sich sein neuester Text nicht darunter befindet?

Lösung: Man kann die Rechnung zunächst direkt durchführen. Es ist nämlich

$$N = \binom{8}{3} \text{ und } N_{\bar{A}} = \binom{7}{3},$$

weil man nur aus 7 Texten wählt, wenn einer nicht gewählt werden darf. Daher ist

$$P(\bar{A}) = \frac{\binom{7}{3}}{\binom{8}{3}} = \frac{5}{8}.$$

Man kann indirekt vorgehen, indem man zuerst $P(A)$ berechnet. Ist der neueste Text in der Wahl, so hat man nur $N_A = \binom{7}{2}$ Möglichkeiten, die anderen zwei Texte zu wählen, daher

$$P(A) = \frac{\binom{7}{2}}{\binom{8}{3}} = \frac{3}{8},$$

woraus

$$P(\bar{A}) = 1 - P(A) = 1 - \frac{3}{8} = \frac{5}{8}$$

wie oben folgt. $\blacksquare$

Beispiel 3.2.2. Aus Abb. 3.1 ist ersichtlich, daß

$$(A \cup B) = \overline{\bar{A} \bar{B}} \quad (3.10)$$

$$\text{und } \overline{AB} = \bar{A} \cup \bar{B}.$$

Diese zwei Beziehungen sind bekannt als *de Morgans Gesetze* und gelten für beliebige Mengen, d.h. $A \cup B \cup C = \overline{\bar{A} \bar{B} \bar{C}}$ usw.

Wir beweisen, daß $P(A_1 \cup A_2 \cup \dots \cup A_n) + P(\bar{A}_1 \bar{A}_2 \dots \bar{A}_n) = 1$.

Lösung: Nach (3.10) ist $P(\bar{A}_1 \bar{A}_2 \dots \bar{A}_n) = P(\overline{A_1 \cup A_2 \cup \dots \cup A_n})$, und nach Satz (2) ist $P(A_1 \cup A_2 \cup \dots \cup A_n) + P(\overline{A_1 \cup A_2 \cup \dots \cup A_n}) = 1$. $\blacksquare$

Übung 3.2.1. Seien A und B zwei beliebige Ereignisse. Drücken Sie formal folgende Ereignisse aus:

- (a) nur A kommt vor;
- (b) mindestens eines von beiden kommt vor;
- (c) beide kommen vor;
- (d) mindestens eines von beiden kommt vor;
- (e) nicht beide kommen vor.

Übung 3.2.2. Eine Urne enthält weiße, schwarze und rote Kugeln. Die Wahrscheinlichkeit, eine weiße Kugel zu ziehen, ist 0,17; die Wahrscheinlichkeit, eine schwarze zu ziehen, beträgt 0,13. Wie ist die Wahrscheinlichkeit $P(A)$, eine rote zu ziehen?

Übung 3.2.3. Beweisen Sie, daß $AB - ABC = ABC$. [Drücken Sie $AB - ABC$ als Produkt aus; benutzen Sie dann (3.10), (3.4), (3.7) und (3.5).]

3.3. Totale Wahrscheinlichkeit für beliebige Ereignisse

Das dritte Axiom der Wahrscheinlichkeit, das man unter dem Namen „Gesetz der totalen Wahrscheinlichkeit“ kennt, besagt, daß die Wahrscheinlichkeit der Summe gleich der Summe der Wahrscheinlichkeiten ist, wenn die Ereignisse disjunkt sind, d.h. wenn sie nicht gleichzeitig vorkommen können. Sind die Ereignisse jedoch nicht disjunkt, so muß man diesen Umstand in Betracht ziehen. Für zwei nicht disjunkte Ereignisse A und B (vgl. Abb. 3.1) gilt der Ausdruck

$$(5) P(A \cup B) = P(A) + P(B) - P(AB). \quad (3.11)$$

Beweis: Zunächst schreiben wir rein formal $A \cup B$ und B als Summe zweier disjunkter Ereignisse, wie aus Abb. 3.1 ersichtlich ist, nämlich

$$\begin{aligned} A \cup B &= A \cup \overline{A}B \\ B &= AB \cup \overline{A}B, \end{aligned}$$

woraus folgt:

$$\begin{aligned} P(A \cup B) &= P(A) + P(\overline{A}B) \\ P(B) &= P(AB) + P(\overline{A}B). \end{aligned}$$

Man subtrahiert die zweite Gleichung von der ersten, um $P(\overline{A}B)$ zu eliminieren, und bekommt

$$P(A \cup B) - P(B) = P(A) - P(AB),$$

woraus folgt:

$$P(A \cup B) = P(A) + P(B) - P(AB). \blacksquare$$

In Abb. 3.1 ist diese Tatsache direkt zu sehen, denn addiert man $P(A)$ und $P(B)$, so ist der Teil $P(AB)$ in jedem Summanden erhalten, wird somit zweimal berechnet und muß deswegen einmal subtrahiert werden.

Beispiel 3.3.1. In einer Sprache ergeben sich Proportionen der Wörter, die im Anlaut bzw. im Auslaut einen Vokal haben, folgendermaßen

Anlaut \ Auslaut	Vokal	Nichtvokal	Σ
Vokal	0,25	0,15	0,40
Nichtvokal	0,40	0,20	0,60
Σ	0,65	0,35	1,00

Wie ist die Wahrscheinlichkeit, daß ein zufällig dem Wörterbuch entnommenes Wort im Anlaut oder im Auslaut einen Vokal hat?

Lösung: Die Tabelle kann man auch folgendermaßen schreiben, wenn A = Vokal im Anlaut, B = Vokal im Auslaut bedeutet:

	B	$\overline{B}$	
A	$P(AB)$	$P(A\overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A}B)$	$P(\overline{A}\overline{B})$	$P(\overline{A})$
	$P(B)$	$P(\overline{B})$	1

Dann berechnet man $P(A \cup B)$ nach der obigen Formel. Addiert man $P(A) + P(B)$, so wird $P(AB)$ zweimal berücksichtigt und muß darum einmal abgezogen werden. Daher ist

$$P(A \cup B) = P(A) + P(B) - P(AB) = 0,40 + 0,65 - 0,25 = 0,80.$$

Aus der Tabelle ersieht man, daß auch die Beziehung

$$P(A \cup B) = 1 - P(\overline{A}\overline{B})$$

gilt, denn es ist $0,80 = 1 - 0,20$. ■

Formel (3.14) läßt sich ohne weiteres verallgemeinern. Nehmen wir drei Ereignisse A_1, A_2, A_3 , dann haben wir unter Anwendung von (3.14)

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P[A_1 \cup (A_2 \cup A_3)] \\ &= P(A_1) + P(A_2 \cup A_3) - P[(A_1(A_2 \cup A_3))] = \end{aligned}$$

$$\begin{aligned}
&= P(A_1) + P(A_2) + P(A_3) - P(A_2 A_3) - P(A_1 A_2 \cup A_1 A_3) \\
&= P(A_1) + P(A_2) + P(A_3) - P(A_1 A_2) - P(A_1 A_3) - \\
&\quad - P(A_2 A_3) + P(A_1 A_2 A_3) = \\
&= \sum_i P(A_i) - \sum_{i \neq j} P(A_i A_j) + P(A_1 A_2 A_3) \quad (3.15)
\end{aligned}$$

Dieser Ausdruck stellt die Wahrscheinlichkeit dar, daß *mindestens* eines der Ereignisse A_1, A_2, A_3 eintritt.

Ganz allgemein haben wir dann für n Ereignisse $A_1, A_2, \dots, A_n$, die nicht disjunkt sind,

$$\begin{aligned}
P(A_1 \cup A_2 \cup \dots \cup A_n) &= P(A_1) + P(A_2) + \dots + P(A_n) - \\
&\quad - P(A_1 A_2) - P(A_1 A_3) - \dots - P(A_{n-1} A_n) + \\
&\quad + P(A_1 A_2 A_3) + P(A_1 A_2 A_4) + \dots + P(A_{n-2} A_{n-1} A_n) - \\
&\quad - \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n) = \\
&= \sum_i P(A_i) - \sum_{i \neq j} P(A_i A_j) + \sum_{i \neq j \neq k} P(A_i A_j A_k) - \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n). \quad (3.16)
\end{aligned}$$

Wären die Ereignisse $A_1, A_2, \dots, A_n$ disjunkt, dann wäre immer $P(A_i A_j) = 0$, $P(A_i A_j A_k) = 0$ usw., so daß sich die Formel (3.16) auf Axiom III reduzieren würde.

Beispiel 3.3.2. Drei sich reimende Wortpaare (*aabbcc*) wurden vermischt und zufällig in eine Reihe gestellt. Wie groß ist die Wahrscheinlichkeit, daß mindestens ein Reimpaar zusammen bleibt?

Lösung: Wir zeigen zuerst die direkte Anwendung der Formel, dann stellen wir sie schrittweise einer anderen Überlegung gegenüber. Sei A_1 = "die beiden *a* stehen zusammen: (*aa*)", A_2 = "die beiden *b* stehen zusammen: (*bb*)", A_3 = "die beiden *c* stehen zusammen: (*cc*)". Wir suchen die Wahrscheinlichkeit $P(A_1 \cup A_2 \cup A_3)$. Da die

drei Mengen A_1, A_2, A_3 keineswegs disjunkt sind, d.h. sich gegenseitig nicht ausschließen, verwenden wir die Formel (3.16). Alle Permutationen von *aabbcc* ergeben

$$N = \frac{6!}{2! 2! 2!}$$

Reihenfolgen. Weiter ist N_{A_i} , d.h. die Zahl der Fälle, wo nur ein Paar zusammensteht, die anderen zwei permutiert werden,

$$N_{A_i} = \frac{5!}{2! 2!}$$

und da es für alle drei gilt, haben wir $\binom{3}{1}$ solcher Möglichkeiten. Daher ist

$$\sum_i N_{A_i} = \binom{3}{1} \frac{5!}{2! 2!}$$

oder

$$\sum_i P(A_i) = \frac{\binom{3}{1} \frac{5!}{2! 2!}}{\frac{6!}{2! 2! 2!}}$$

In diesen Permutationen sind jeweils auch die Fälle eingeschlossen, wo zwei und drei Paare zusammenstehen. Diese müssen abgezogen werden. Es ist

$$N_{A_i A_j} = \frac{4!}{2!}$$

und da dies wieder für alle $\langle i, j \rangle$ gilt, deren Zahl $\binom{3}{2}$ ist, wird

$$\sum_{i \neq j} N_{A_i A_j} = \binom{3}{2} \frac{4!}{2!}$$

oder

$$\sum_{i \neq j} P(A_i A_j) = \frac{\binom{3}{2} \frac{4!}{2!}}{\frac{6!}{2! 2! 2!}}.$$

Zieht man diese Zahl ab, so hat man auch alle Fälle abgezogen, wo alle Paare zusammenstehen und muß einen Fall wieder zurückgeben. Es ist

$$N_{A_1 A_2 A_3} = 3!$$

oder

$$P(A_1 A_2 A_3) = \frac{3!}{\frac{6!}{2! 2! 2!}}.$$

Wir können also schreiben

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= \sum_i P(A_i) - \sum_{i \neq j} P(A_i A_j) + P(A_1 A_2 A_3) \\ &= \frac{\binom{3}{1} \frac{5!}{2! 2!}}{\frac{6!}{2! 2! 2!}} - \frac{\binom{3}{2} \frac{4!}{2!}}{\frac{6!}{2! 2! 2!}} + \frac{3!}{\frac{6!}{2! 2! 2!}} \\ &= 1 - \frac{6}{15} + \frac{1}{15} = \frac{2}{3}. \end{aligned}$$

Eine andere Überlegung führt zu demselben Resultat und ist etwas anschaulicher:

I. Alle Paare kommen zusammen auf $3!$ Weisen vor, weil wir jedes Paar als ein Element betrachten und permutieren.

II. Nur aa und bb stehen zusammen auf $\frac{4!}{2!} - 3!$ Weisen, weil wir aa und bb jeweils als ein Element betrachten und cc als zwei. Bei der Permutierung kommen jedoch auch die 2 c zusammen, und da es $3!$ solcher Fälle gibt, muß diese Zahl abgezogen werden. Da es jedoch $\binom{3}{2}$ Möglichkeiten gibt, aus 3 Paaren jeweils Kombinationen von 2 Paaren zu bilden, erhalten wir

$$\binom{3}{2} \left[\frac{4!}{2!} - 3! \right].$$

III. Kommt z.B. nur das Paar aa zusammen vor, ergeben sich $5! / (2! 2!)$ Permutationen, von denen wir diejenigen abziehen müssen, wo auch bb zusammen vorkommt, d.h. $\frac{4!}{2!} - 3!$, und dieselbe Zahl, wo auch cc zusammen steht, also $2 (\frac{4!}{2!} - 3!)$, gleichfalls die Permutationen, wo alle drei Paare zusammen sind, also $3!$. Das gilt aber für alle 3 Buchstaben, daher haben wir

$$\binom{3}{1} \left[\frac{5!}{2! 2!} - 2 \left(\frac{4!}{2!} - 3! \right) - 3! \right],$$

und das ergibt zusammen

$$\begin{aligned} &\binom{3}{1} \left[\frac{5!}{2! 2!} - 2 \left(\frac{4!}{2!} - 3! \right) - 3! \right] + \binom{3}{2} \left[\frac{4!}{2!} - 3! \right] + 3! = \\ &= (3) \frac{5!}{2! 2!} - (3) \frac{4!}{2!} + 3! = N_{A_1 \cup A_2 \cup A_3}, \end{aligned}$$

was mit dem Zähler der oberen Formel identisch ist. Für diesen Fall kann man die Wahrscheinlichkeit auch in der Form

$$P(A_1 \cup A_2 \cup A_3) = \sum_{k=1}^3 (-1)^{k+1} \binom{3}{k} \frac{(n-k)! (2)^k}{n!} \quad (3.17)$$

zusammenfassen. ■

In der Literatur findet man oft die verkürzte Schreibweise $P(A_i) = p_i$, $P(A_i A_j) = p_{ij}$, $P(A_i A_j A_k) = p_{ijk}$ usw. und $\sum P(A_i) = S_1$, $\sum P(A_i A_j) = S_2$, $\sum P(A_i A_j A_k) = S_3$ usw.. So findet man die Formel (3.16) auch in der Form

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = S_1 - S_2 + S_3 - \dots + (-1)^{n-1} S_n. \quad (3.18)$$

Da diese Formel die Wahrscheinlichkeit angibt, daß mindestens ein Ereignis vorkommt, bezeichnet man sie oft als P_1 . Der Spezialfall für zwei Ereignisse ist

$$P(A_1 \cup A_2) = S_1 - S_2. \quad (3.19)$$

Die Verwendung der Formel (3.18) illustrieren wir noch an dem sogenannten Zuordnungsproblem (matching). Wir stellen uns vor, daß ein Paket von n (z.B. 32) Karten nebeneinander auf dem Tisch liegt; dann mischt jemand ein anderes identisches Kartenspiel gründlich durch und legt diese Karten nebeneinander unter die Karten des ersten Pakets. Die Frage lautet: Wie groß ist die Wahrscheinlichkeit, daß von den n Paaren untereinanderliegender Karten r Paare je zwei gleiche Karten enthalten? Ein Paar mit gleichen Karten werden wir als identisches Paar bezeichnen.

Die Berechnung der Wahrscheinlichkeit, daß unter n Kartenpaaren mindestens 1 Paar identisch ist, erfolgt folgendermaßen:

Die Zahl aller Zuordnungsmöglichkeiten der Karten des zweiten Pakets zu denen des ersten ist offensichtlich $N = n!$; denn betrachtet man die Karten des ersten Pakets als Stellen, auf denen die Karten des anderen Pakets permutieren können, so ergibt sich dieses Resultat.

Wenn 1 Paar aus n Paaren identisch ist, dann können die anderen $(n-1)$ Karten auf $(n-1)!$ Weisen permutieren. Das identische Paar kann man jedoch aus n auf $\binom{n}{1}$ Weisen wählen, daher ist

$$\sum_i N_{A_i} = \binom{n}{1} (n-1)! = n!$$

und

$$S_1 = \frac{\sum_i N_{A_i}}{N} = \frac{n!}{n!} = 1.$$

Sind zwei Paare identisch, die man aus n Paaren auf $\binom{n}{2}$ Weisen wählen kann, so kann der Rest auf $(n-2)!$ Weisen permutieren. Daher ist

$$\sum_{i \neq j} N_{A_i A_j} = \binom{n}{2} (n-2)! = \frac{n!}{2!}$$

oder

$$S_2 = \frac{n!}{2! n!} = \frac{1}{2!}$$

usw. Allgemein ist für r identische Paare

$$\sum_{i \neq j \neq \dots \neq r} N_{A_i A_j \dots A_r} = \binom{n}{r} (n-r)! = \frac{n!}{r!}$$

oder

$$S_r = \frac{n!}{r! n!} = \frac{1}{r!},$$

und die Zahl der günstigen Fälle beträgt

$$\begin{aligned} N_{A_1 \cup A_2 \cup \dots \cup A_n} &= n! - \frac{n!}{2!} + \frac{n!}{3!} - \dots + (-1)^{n-1} \frac{n!}{n!} \\ &= n! \left[1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{n-1} \frac{1}{n!} \right]. \end{aligned} \quad (3.20)$$

So ergibt sich die Wahrscheinlichkeit, daß mindestens ein Paar identisch ist, als

$$P_1 = P(A_1 \cup A_2 \cup \dots \cup A_n)$$

$$= \frac{N_{A_1 \cup A_2 \cup \dots \cup A_n}}{N}$$

$$= S_1 - S_2 + S_3 - \dots + (-1)^{n-1} S_n$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{n-1} \frac{1}{n!}. \quad (3.21)$$

Beispiel 3.3.3. Aus zwei Kartenspielen wurden je 10 Karten gewählt, wobei aus dem einen Spiel dieselbe Auswahl genommen wurde wie aus dem anderen, und zufällig untereinander auf den Tisch gelegt. Wie ist die Wahrscheinlichkeit, daß mindestens ein Paar identisch ist?

Lösung: Wir setzen $n = 10$ in die Formel (3.21) ein und bekommen

$$P_1 = 1 - \frac{1}{2!} + \frac{1}{3!} - \dots + \frac{1}{9!} - \frac{1}{10!} = 0,6321.$$

Bemerkung: Diese Zahl kann für große n sehr gut mit $P_1 = 1 - e^{-1}$ approximiert werden. ■

Aus diesem Resultat kann man weitere Schlüsse ziehen. Die Zahl der Permutationen, bei denen sich *keine identischen Karten* treffen, ist

$$N_{\bar{A}_1 \bar{A}_2 \dots \bar{A}_n} = N - N_{A_1 \cup A_2 \cup \dots \cup A_n}$$

$$= n! - n! \left[1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{n-1} \frac{1}{n!} \right] =$$

$$= n! \left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + (-1)^n \frac{1}{n!} \right] \quad (3.22)$$

und die Wahrscheinlichkeit, daß unter n Paaren kein identisches vorkommt, ist

$$\begin{aligned} P_{(0)} &= \frac{N - N_{A_1 \cup A_2 \cup \dots \cup A_n}}{N} = 1 - P_1 = \\ &= 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + (-1)^n \frac{1}{n!}. \end{aligned} \quad (3.23)$$

Beispiel: 3.3.4. Berechnen Sie die Wahrscheinlichkeit, daß unter 10 gleichen Karten zweier Kartenpakete kein Paar identisch ist.

Lösung: Aus dem Resultat in Beispiel 3.3.3 erhalten wir

$$P_{(0)} = 1 - P_1 = 1 - 0,6321 = 0,3697. \blacksquare$$

Weiter können wir die Wahrscheinlichkeit ableiten, daß genau r Paare identisch sind. Die r identischen Paare kann man aus n auf $\binom{n}{r}$ Weisen wählen, und die Zahl der Permutationen der restlichen $(n-r)$ Karten so, daß keine identischen Paare mehr vorhanden sind, ist nach (3.22)

$$(n-r)! \left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots + (-1)^{n-r} \frac{1}{(n-r)!} \right].$$

Daher ist die Wahrscheinlichkeit, daß genau r Paare identisch sind,

$$\begin{aligned} P_{(r)} &= \frac{\binom{n}{r} (n-r)! \left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + (-1)^{n-r} \frac{1}{(n-r)!} \right]}{n!} \\ &= \frac{1}{r!} \left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + (-1)^{n-r} \frac{1}{(n-r)!} \right] \\ &= \frac{1}{r!} \sum_{i=0}^{n-r} (-1)^i \frac{1}{i!}. \end{aligned} \quad (3.24)$$

Aus dieser Formel lassen sich für einzelne r die Formeln leicht aufstellen. Für $P_{(0)}$ gilt (3.23), weiter ist

$$\begin{aligned} P_{(1)} &= 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^{n-2}}{(n-2)!} + \frac{(-1)^{n-1}}{(n-1)!} \\ P_{(2)} &= \frac{1}{2!} \left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^{n-3}}{(n-3)!} + \frac{(-1)^{n-2}}{(n-2)!} \right] \\ P_{(3)} &= \frac{1}{3!} \left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^{n-4}}{(n-4)!} + \frac{(-1)^{n-3}}{(n-3)!} \right] \\ &\dots \dots \dots \end{aligned} \quad (3.25)$$

$$P_{(n-3)} = \frac{1}{(n-3)!} \left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} \right)$$

$$P_{(n-2)} = \frac{1}{(n-2)!} \left(1 - 1 + \frac{1}{2!} \right)$$

$$P_{(n-1)} = \frac{1}{(n-1)!} (1 - 1) = 0$$

$$P_{(n)} = \frac{1}{n!}.$$

Anmerkung: Das Ereignis, daß unter n Paaren genau $n-1$ identisch ist, ist unmöglich. Wenn $n-1$ Paare identisch sind, so muß auch das letzte Paar identisch sein. Daher $P_{(n-1)} = 0$.

Beispiel 3.3.5. Wie ist die Wahrscheinlichkeit, daß unter den 10 Paaren des Beispiels 3.3.4 genau 6 identisch sind?

Lösung: Nach (3.25) ist

$$P_{(6)} = \frac{1}{6!} \left(\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} \right) = 0,00052. \blacksquare$$

Die Formeln (3.25) kann man mit einem kleinen Kunstgriff auch in einer anderen Form darstellen. Schreibt man ebenso wie in (3.21) für $1/r! = S_r$, so ist z.B.

$$P_{(1)} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots + \frac{(-1)^{n-1}}{(n-1)!}.$$

Man multipliziert das erste Glied mit $1/1$, das zweite mit $2/2$ usw., so daß

$$\begin{aligned} P_{(1)} &= \frac{1}{1!} - \frac{2}{2!} + \frac{3}{3!} - \frac{4}{4!} + \frac{5}{5!} - \dots + \frac{(-1)^{n-1}n}{n!} \\ &= S_1 - \binom{2}{1} S_2 + \binom{3}{1} S_3 - \binom{4}{1} S_4 + \dots + (-1)^{n-1} \binom{n}{1} S_n. \end{aligned}$$

$$\begin{aligned} P_{(2)} &= \frac{1}{2!} - \frac{1}{2!} + \frac{1}{2!2!} - \frac{1}{2!3!} + \frac{1}{2!4!} + \dots + \frac{(-1)^{n-2}}{2!(n-2)!} \\ &= \frac{1}{2!} - \frac{3!}{2!3!} + \frac{4!}{2!2!4!} - \frac{5!}{2!3!5!} + \frac{6!}{2!4!6!} - \dots + \frac{(-1)^{n-1}n!}{2!(n-2)!n!} \\ &= S_2 - \binom{3}{2} S_3 + \binom{4}{2} S_4 - \binom{5}{2} S_5 + \binom{6}{2} S_6 - \dots + (-1)^n \binom{n}{2} S_n, \end{aligned}$$

und so zeigt man leicht, daß allgemein gilt:

$$P_{(r)} = S_r - \binom{r+1}{r} S_{r+1} + \binom{r+2}{r} S_{r+2} - \binom{r+3}{r} S_{r+3} + \dots + (-1)^n \binom{n}{r} S_n. \quad (3.26)$$

Mit Hilfe dieser Formel können wir wieder leicht die Wahrscheinlichkeit berechnen, daß mindestens r Paare (d.h. r oder mehr) identisch sind. Wir werden wieder schrittweise vorgehen. Die Formel für P_1 ist bereits bekannt, nämlich

$$P_1 = S_1 - S_2 + S_3 - \dots + (-1)^{n-1} S_n.$$

Die Wahrscheinlichkeit, daß 2 oder mehr Paare identisch sind, erhalten wir so, daß wir von P_1 einfach $P_{(1)}$ abziehen, d.h.

$$P_2 = P_1 - P_{(1)} = S_1 - S_2 + S_3 - S_4 + \dots + (-1)^{n-1} S_n -$$

$$- S_1 + \binom{2}{1} S_2 - \binom{3}{1} S_3 + \binom{4}{1} S_4 - \dots + (-1)^n \binom{n}{1} S_n =$$

$$= S_2 - 2S_3 + 3S_4 - \dots + (-1)^n (n-1) S_n$$

$$= S_2 - \binom{2}{1} S_3 + \binom{3}{1} S_4 - \binom{4}{1} S_5 + \dots + (-1)^n \binom{n-1}{1} S_n.$$

Auf dieselbe Weise berechnen wir noch der Anschaulichkeit halber

$$P_3 = P_2 - P_{(2)}$$

$$= S_2 - \binom{2}{1} S_3 + \binom{3}{1} S_4 - \binom{4}{1} S_5 + \dots + (-1)^n \binom{n-1}{1} S_n -$$

$$- S_2 + \binom{3}{2} S_3 - \binom{4}{2} S_4 + \binom{5}{2} S_5 - \dots + (-1)^{n-1} \binom{n}{2} S_n =$$

$$= S_2 - S_2 + \left[\binom{3}{2} - \binom{2}{1} \right] S_3 - \left[\binom{4}{2} - \binom{3}{1} \right] S_4 + \left[\binom{5}{2} - \binom{4}{1} \right] S_5 - \dots$$

$$\dots + (-1)^{n-1} \left[\binom{n}{2} - \binom{n-1}{1} \right] S_n.$$

Da nach (2.18) $\binom{r}{k} - \binom{r-1}{k-1} = \binom{r-1}{k}$, so kann man schreiben

$$P_3 = S_3 - \binom{3}{2} S_4 + \binom{4}{2} S_5 - \binom{5}{2} S_6 + \dots + (-1)^{n-1} \binom{n-1}{2} S_n.$$

So läßt sich durch Weiterrechnen mit der Rekursionsformel $P_{r+1} = P_r - P_{(r)}$ zeigen, daß sich allgemein die Wahrscheinlichkeit, daß mindestens r Paare identisch sind, ergibt als

$$P_r = S_r - \binom{r}{r-1} S_{r+1} + \binom{r+1}{r-1} S_{r+2} - \binom{r+2}{r-1} S_{r+3} + \dots + (-1)^{n-1} \binom{n-1}{r-1} S_n \quad (3.27)$$

$$\text{mit } S_r = \frac{1}{r!}, S_{r+1} = \frac{1}{(r+1)!}, \dots, S_n = \frac{1}{n!}.$$

Beispiel 3.3.6. Wie ist die Wahrscheinlichkeit, daß unter je 10 gleichen Karten zweier Kartenspiele mindestens 7 Paare identisch sind?

Lösung: Nach (3.27) ist für $n = 10, r = 7$

$$P_7 = S_7 - \binom{7}{6} S_8 + \binom{8}{6} S_9 - \binom{9}{6} S_{10} \\ = \frac{1}{7!} - \frac{7}{8!} + \frac{28}{9!} - \frac{84}{10!} = 0,0000788 \cdot \blacksquare$$

Übung 3.3.1. In einem zweisprachigen Gebiet ist die Proportion der A-sprachigen p_A und der B-sprachigen p_B , wobei keine Person beide Sprachen spricht. Wie ist die Wahrscheinlichkeit, daß aus zwei zufällig gewählten Personen

- (a) beide dieselbe Sprache sprechen,
- (b) jede eine andere Sprache spricht? (Vgl. Greenbergs Index of linguistic diversity.)

Übung 3.3.2. Ein Wort kommt in der Sprache mit der Wahrscheinlichkeit p vor. Man untersucht einen Text. Wie ist die Wahrscheinlichkeit, daß dieses Wort im Text vorkommt (d.h. mindestens einmal vorkommt)?

Übung 3.3.3. Angenommen, die Phoneme einer Sprache, von denen 5 Vokale und 15 Konsonanten sind, wären gleichwahrscheinlich. Wie wäre dann die Wahrscheinlichkeit, an einer beliebigen Stelle eines Textes einen Vokal (Ereignis A) und sodann an einer anderen Stelle einen Konsonanten zu finden (Ereignis B), also die Wahrscheinlichkeit von $C = A \cup B$?

Übung 3.3.4. Bei einem statistischen Verfahren zur Autorenbestimmung sei folgende Frage aufgetreten: Wenn 5 Texte und 5 Autoren einander zufällig zugeordnet sind, wie groß ist dann die Wahrscheinlichkeit, daß genau ein Text seinem wirklichen Autor zugeordnet wurde?

Übung 3.3.5. Wie ist die Wahrscheinlichkeit, daß man aus 8 identischen Kartenpaaren durch zufälliges Zuordnen 5 identische Paare bildet?

3.4. Bedingte Wahrscheinlichkeit

Betrachten wir zwei Ereignisse A und B , von denen wir annehmen, daß sie irgendwie zusammenhängen. Ändert sich die Wahrscheinlichkeit des Ereignisses B nach der Realisierung des Ereignisses A , so sprechen wir von der *bedingten Wahrscheinlichkeit*, die wir als

$$P(B|A)$$

bezeichnen und lesen als "die Wahrscheinlichkeit von B unter der Bedingung A ". A ist einfach eine Beschränkung, unter der man die Daten untersucht oder ein Experiment durchführt. Wenn man beispielsweise die (vokalische oder nichtvokalische) Endung der Wörter untersucht, so hat man dabei keine Beschränkung zu berücksichtigen. Untersucht man jedoch die Endung der Wörter, die mit einem Vokal anfangen, so werden nicht alle Wörter, sondern nur bestimmte in Betracht gezogen. Bezeichnen wir mit

N_A die Anzahl der Wörter mit Vokal im Anlaut

$N_{\bar{A}}$ die Anzahl der Wörter mit Konsonant im Anlaut

N_B die Anzahl der Wörter mit Vokal im Auslaut

$N_{\bar{B}}$ die Anzahl der Wörter mit Konsonant im Auslaut,

so können wir, analog zum Beispiel 3.3.1, Tabelle 3.1 aufstellen.

Tabelle 3.1.

Anlaut \ Auslaut	B	$\bar{B}$	Σ
A	N_{AB}	$N_{A\bar{B}}$	N_A
$\bar{A}$	$N_{\bar{A}B}$	$N_{\bar{A}\bar{B}}$	$N_{\bar{A}}$
Σ	N_B	$N_{\bar{B}}$	N

Will man die Proportion der Wörter, die auf Vokal enden, unter denen, die auch mit Vokal anfangen, berechnen, dann dividiert man offensichtlich

$$\frac{N_{AB}}{N_A}$$

und die Proportion der Wörter, die auf Konsonant enden, unter denen, die mit Vokal anfangen, ist offensichtlich

$$\frac{N_{A\bar{B}}}{N_A}$$

d.h. man untersucht nur die Zeile von A. Dividiert man den Zähler und den Nenner dieser Ausdrücke durch N, so wird

$$\frac{\frac{N_{AB}}{N}}{\frac{N_A}{N}} = \frac{P(AB)}{P(A)}$$

laut der klassischen Definition der Wahrscheinlichkeit, und

$$\frac{\frac{N_{A\bar{B}}}{N}}{\frac{N_A}{N}} = \frac{P(A\bar{B})}{P(A)}.$$

Daher ist die Wahrscheinlichkeit, daß ein Wort auf Vokal endet, unter der Bedingung, daß es mit Vokal anfängt, gegeben durch

$$P(B|A) = \frac{P(AB)}{P(A)}, \quad P(A) \neq 0. \quad (3.26)$$

Da $N_A = N_{AB} + N_{A\bar{B}}$, kann man auch schreiben

$$P(B|A) = \frac{N_{AB}}{N_{AB} + N_{A\bar{B}}}. \quad (3.27)$$

Diese Überlegung gilt auch dann, wenn man A unter der Bedingung B untersucht, d.h. wenn man die Spalten der Tabelle in Betracht zieht. Es gilt offensichtlich auch

$$P(A|B) = \frac{P(AB)}{P(B)}, \quad P(B) \neq 0. \quad (3.28)$$

Aus (3.26) und (3.28) folgt unmittelbar

$$P(AB) = P(A) P(B|A) = P(B) P(A|B). \quad (3.29)$$

Diesen Ausdruck nennt man auch *Multiplikationssatz*, und wir werden ihn später verallgemeinern.

Schreibt man Tabelle 3.1 in Wahrscheinlichkeiten um, wie in Tabelle 3.2, so

Tabelle 3.2.

	B	$\bar{B}$	Marginale WS
A	$P(AB)$	$P(A\bar{B})$	$P(A)$
$\bar{A}$	$P(\bar{A}B)$	$P(\bar{A}\bar{B})$	$P(\bar{A})$
Marginale WS	$P(B)$	$P(\bar{B})$	1

folgt aus (3.29) in bezug auf die marginale Wahrscheinlichkeit der Zeilen das Resultat in Tabelle 3.3.

Tabelle 3.3.

	B	$\bar{B}$	Marginale WS
A	$P(A) P(B A)$	$P(A) P(\bar{B} A)$	$P(A)$
$\bar{A}$	$P(\bar{A}) P(B \bar{A})$	$P(\bar{A}) P(\bar{B} \bar{A})$	$P(\bar{A})$
Marginale WS	$P(B)$	$P(\bar{B})$	1

Dividiert man die Zeilen durch ihre marginale Wahrscheinlichkeit, so bekommt man Tabelle 3.4

Tabelle 3.4.

	B	$\bar{B}$	Σ
A	$P(B A)$	$P(\bar{B} A)$	1
$\bar{A}$	$P(B \bar{A})$	$P(\bar{B} \bar{A})$	1
Marginale WS	$P(B)$	$P(\bar{B})$	1

Beispiel 3.4.1. Im Beispiel 3.3.1 haben wir die Wahrscheinlichkeiten des vokalischen und des konsonantischen An- und Auslauts der Wörter einer Sprache betrachtet. Sei A = Vokal im Anlaut, B = Vokal im Auslaut. Die Wahrscheinlichkeiten waren

	B	$\bar{B}$	Σ
A	0,25	0,15	0,40
$\bar{A}$	0,40	0,20	0,60
Σ	0,65	0,35	1

wobei die Zahlen den Wahrscheinlichkeiten in Tabelle 3.2 entsprechen. Wie groß ist die Wahrscheinlichkeit,

(a) daß ein Wort auf Vokal endet, wenn wir wissen (= unter der Bedingung), daß es mit Vokal anfängt,

(b) daß ein Wort auf Konsonant endet, unter der Bedingung, daß es mit Konsonant anfängt?

Lösung: (a) Laut (3.26) ist

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{0,25}{0,40} = 0,6250 ;$$

$$(b) \quad P(\bar{B}|\bar{A}) = \frac{P(\bar{A}\bar{B})}{P(\bar{A})} = \frac{0,20}{0,60} = 0,3333 . \blacksquare$$

Beispiel 3.4.2. In einer Urne sind N_A weiße und N_B schwarze Kugeln, $N_A + N_B = N$. Wir ziehen hintereinander zwei Kugeln ohne Zurücklegen. Wie groß ist die Wahrscheinlichkeit, daß die erste weiß, die zweite schwarz ist?

Lösung: Nach (3.29) ist

$$P(AB) = P(A) P(B|A).$$

Hier ist $P(A) = N_A / N$ und $P(B|A) = N_B / (N - 1)$, weil nach dem Zug einer weißen Kugel nur $N-1$ Kugeln übrigbleiben. Daher ist

$$P(AB) = \frac{N_A \cdot N_B}{N(N-1)} . \blacksquare$$

Aus Tabelle 3.4 kann man mehrere Beziehungen ableiten.

(1) Es ist immer

$$P(B|A) + P(\bar{B}|A) = 1 \quad (3.30)$$

und das gilt auch, wenn B als Bedingung steht

$$P(A|B) + P(\bar{A}|B) = 1$$

und genauso für $\bar{A}$ und $\bar{B}$ als Bedingungen. ■

(2) Wenn

$$P(B|A) = P(B|\bar{A}), \quad (3.31)$$

dann sagen wir, daß B unabhängig von A ist, weil die Bedingung keinen Einfluß auf die Wahrscheinlichkeit von B hat. Ist

$$P(B|A) \neq P(B|\bar{A})$$

dann ist B von A abhängig. ■

(3) Im Falle der Unabhängigkeit folgt aus der Beziehung (3.31)

$$P(B|A) = P(B) = P(B|\bar{A}) \quad (3.32)$$

Beweis: $P(B) = P(AB) + P(\bar{A}B)$ (vgl. Abb. 2.1, Tabelle 2.2)

$$= P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) \text{ (vgl. 2.29, Tabelle 2.3)}$$

$$= P(B|A) [P(A) + P(\bar{A})] \text{ (wegen 2.31)}$$

$$= P(B|A) . \blacksquare$$

(4) Setzt man (3.32) in (3.29) ein, so bekommt man

$$P(AB) = P(A) P(B|A) = P(A) P(B) \quad (3.33)$$

und wenn diese Beziehung gilt, so sagen wir, daß A und B gegenseitig unabhängig sind. Dieses Resultat kann man folgendermaßen interpretieren: Die Wahrscheinlichkeit, daß zwei unabhängige Ereignisse gleichzeitig (zusammen) vorkommen, ist gleich dem Produkt der Wahrscheinlichkeiten des separaten Vorkommens von jedem Ereignis. ■

Aus (2), (3) und (4) folgt, daß die Ereignisse A und B genau dann unabhängig sind, wenn eine der folgenden Bedingungen gilt

$$(a) P(B|A) = P(B), \text{ wenn } P(A) > 0$$

$$(b) P(A|B) = P(A), \text{ wenn } P(B) > 0$$

$$(c) P(AB) = P(A) P(B).$$

(5) Formel (3.29) läßt sich für n Ereignisse $A_1, A_2, \dots, A_n$ verallgemeinern. Es ist nämlich

$$\begin{aligned} P(A_1 A_2 \dots A_n) &= P[(A_1 A_2 \dots A_{n-1}) A_n] \\ &= P(A_1 A_2 \dots A_{n-1}) P(A_n | A_1 A_2 \dots A_{n-1}) \\ &= P(A_1 A_2 \dots A_{n-2}) P(A_{n-1} | A_1 A_2 \dots A_{n-2}) P(A_n | A_1 A_2 \dots A_{n-1}) \\ &= \dots \\ &= P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) \dots P(A_n | A_1 A_2 \dots A_{n-1}) \quad (3.34) \end{aligned}$$

und wenn alle Ereignisse in jeder Kombination voneinander unabhängig sind, so wird (3.34) zu

$$P(A_1 A_2 \dots A_n) = P(A_1) P(A_2) \dots P(A_n). \blacksquare \quad (3.35)$$

Beispiel 3.4.3. Die Wahrscheinlichkeit, daß auf einer Buchseite kein Druckfehler vorhanden ist, beträgt 0,99. Finden Sie die Wahrscheinlichkeit, daß auf 4 aufeinander folgenden Buchseiten keine Druckfehler zu finden sind.

Lösung: Die Druckfehler auf den einzelnen Seiten sind voneinander unabhängig. Daher ist

$$P(A_1 A_2 A_3 A_4) = P(A_1) P(A_2) P(A_3) P(A_4) = 0,99^4 = 0,96. \blacksquare$$

Übung 3.4.1. In einer Urne sind 5 Vokale und 20 Konsonanten auf Zetteln geschrieben. Jemand zieht ohne Zurücklegen 3 Zettel. Finden Sie die Wahrscheinlichkeit $P(V_1 V_2 V_3)$, d.h. daß alle drei Vokale sind.

Übung 3.4.2. Stellen Sie fest, ob die Ereignisse A und B voneinander unabhängig sind, wenn folgende Wahrscheinlichkeiten angegeben sind: $P(AB) = 0,12$; $P(\overline{A}\overline{B}) = 0,18$; $P(\overline{A}B) = 0,28$; $P(A\overline{B}) = 0,42$. [Berechnen Sie erst $P(A)$, $P(B)$, $P(\overline{A})$, $P(\overline{B})$; benutzen Sie dann 3.33.]

Übung 3.4.3. Beweisen Sie, daß $P(S|A) = 1$.

Übung 3.4.4. Beweisen Sie, daß $P(A \cup B|C) = P(A|C) + P(B|C)$, wenn $AB \neq \emptyset$.

3.5. Der Bayessche Satz

In Formel (3.29) haben wir festgestellt, daß

$$P(AB) = P(A) P(B|A) = P(B) P(A|B).$$

Dividiert man diesen Ausdruck durch $P(A)$, so bekommt man

$$P(B|A) = \frac{P(B) P(A|B)}{P(A)}. \quad (3.36)$$

In diesem Fall bestand das Ereignis A jedoch nur aus zwei Teilen $A = AB + \overline{A}B$, wie aus Tabelle 3.2 ersichtlich ist. Stellen wir uns jedoch vor, daß B mehrere Ausgänge hat, d.h. nicht nur B und $\overline{B}$, sondern $B_1, B_2, \dots, B_n$, so daß

$$A = AB_1 + AB_2 + \dots + AB_n$$

und

$$P(A) = P(AB_1) + P(AB_2) + \dots + P(AB_n). \quad (3.37)$$

Auch hier gilt für ein beliebiges $P(B_i|A)$, daß

$$P(B_i|A) = \frac{P(B_i) P(A|B_i)}{P(A)} \quad (3.38)$$

wie in Formel (3.36). Laut Multiplikationsgesetz (3.29) kann man Formel (3.37) umschreiben als

$$P(A) = P(B_1) P(A|B_1) + P(B_2) P(A|B_2) + \dots + P(B_n) P(A|B_n) \\ = \sum_{i=1}^n P(B_i) P(A|B_i). \quad (3.39)$$

Setzt man dieses Resultat in (3.38) ein, so bekommt man schließlich die Bayesche Formel

$$P(B_i|A) = \frac{P(B_i) P(A|B_i)}{\sum_{i=1}^n P(B_i) P(A|B_i)}. \quad (3.40)$$

Diese Formel wird auch als *Wahrscheinlichkeit a posteriori* bezeichnet.

Beispiel 3.5.1. Eine Zufallsstichprobe enthält Wörter aus drei Wortarten: 5 Pronomina, 7 Adjektive und 8 Substantive. Von den Pronomina sind zwei einsilbig, von den Adjektiven vier und von den Substantiven drei. Wie groß ist die Wahrscheinlichkeit, daß ein zufällig gewähltes Wort ein Pronomen, Adjektiv bzw. Substantiv ist, wenn es einsilbig ist?

Lösung: Bezeichnen wir

- B_1 - Wählen eines Pronomens
- B_2 - Wählen eines Adjektivs
- B_3 - Wählen eines Substantivs
- A - Wählen eines Einsilbers.

Die Wahrscheinlichkeit, eine der Wortarten zu wählen, ist $P(B_i) = \frac{1}{3}$. Zur Berechnung brauchen wir noch die bedingten Wahrscheinlichkeiten $P(A|B_i)$. Es ist

$$P(A|B_1) = \frac{2}{5}; P(A|B_2) = \frac{4}{7}; P(A|B_3) = \frac{3}{8}.$$

Setzen wir dieses Resultat in (3.40) ein, so bekommen wir

$$P(B_1|A) = \frac{P(B_1) P(A|B_1)}{\sum_{i=1}^3 P(B_i) P(A|B_i)} = \frac{\frac{1}{3} \cdot \frac{2}{5}}{\frac{1}{3} \cdot \frac{2}{5} + \frac{1}{3} \cdot \frac{4}{7} + \frac{1}{3} \cdot \frac{3}{8}} = 0,2971$$

als die Wahrscheinlichkeit dafür, auf ein Pronomen zu stoßen. Weiter ist die Wahrscheinlichkeit, ein Adjektiv zu wählen,

$$P(B_2|A) = \frac{\frac{1}{3} \cdot \frac{4}{7}}{\frac{1}{3} \cdot \frac{2}{5} + \frac{1}{3} \cdot \frac{4}{7} + \frac{1}{3} \cdot \frac{3}{8}} = 0,4244$$

und für Substantive ist die Wahrscheinlichkeit

$$P(B_3|A) = \frac{\frac{1}{3} \cdot \frac{3}{8}}{\frac{1}{3} \cdot \frac{2}{5} + \frac{1}{3} \cdot \frac{4}{7} + \frac{1}{3} \cdot \frac{3}{8}} = 0,2785.$$

Man sieht auch, daß die Summe dieser drei Wahrscheinlichkeiten gleich 1 ist. ■

Aufgaben

- 3.1. In Bochum werden Autoschilder vergeben, die vier Buchstaben enthalten, von denen die ersten zwei BO sind. Wie ist die Wahrscheinlichkeit,
 - (a) daß man ein Schild nur mit den Buchstaben B und O bekommt,
 - (b) daß man ein Schild mit BO-BO bekommt, wenn man sein Auto anmeldet?
 [Nehmen Sie an, daß im Alphabet 26 Buchstaben enthalten sind.]
- 3.2. Aus dem Namen MAIER bilden wir Wörter von 3 Buchstaben durch zufälliges Herausgreifen. Wie ist die Wahrscheinlichkeit,
 - (a) daß ein Wort nur aus Vokalen besteht,
 - (b) daß ein Wort den Buchstaben M enthält,
 - (c) daß in einem Wort das M am Anfang steht,
 - (d) daß in einem Wort 2 Vokale und 1 Konsonant sind,
 - (e) daß in einem Wort am Anfang und am Ende ein Konsonant steht und in der Mitte ein Vokal,
 - (f) daß in einem Wort am Anfang und am Ende ein Vokal steht und in der Mitte ein Konsonant?
- 3.3. Beweisen Sie, daß $P(A - BC) + P(A \bar{B} \bar{C}) = P(A\bar{B}) + P(A\bar{C})$.
- 3.4. An einer Klausur sitzen 8 Kandidaten nebeneinander, von denen 3 eine richtige Lösung abgeben. Wie ist die Wahrscheinlichkeit, daß diese 3 Personen direkt nebeneinander saßen?
- 3.5. Aus den Wörtern SEMANTIK und LINGUISTIK wird zufällig jeweils ein Buchstabe gezogen. Wie ist die Wahrscheinlichkeit, daß die Buchstaben gleich sind?

- 3.6. Die Buchstaben des Wortes SEMANTIK werden auf Zettel geschrieben und in eine Urne gelegt. Man zieht einen Zettel, schreibt den Buchstaben ab und legt den Zettel zurück. Man zieht hintereinander 5 Zettel. Wie ist die Wahrscheinlichkeit, daß kein Buchstabe zweimal gezogen wird?
- 3.7. Verallgemeinerung. In einer Urne befinden sich n unterschiedliche Buchstaben. Man zieht mit Zurücklegen k Buchstaben. Zeigen Sie, daß die Wahrscheinlichkeit, daß kein Buchstabe zweimal gezogen wird, V_n^k / V_n^k beträgt.
- 3.8. In einem Text von 1000 Sätzen hat man das Wort "Student" in 150 Sätzen gefunden und das Wort "klug" in 60 Sätzen. Wie ist die Wahrscheinlichkeit, daß man beide in einem Satz findet?
- 3.9. Die Wahrscheinlichkeit des Vorkommens eines Vokals in einer bestimmten Sprache ist $\frac{4}{7}$. Wie ist die Wahrscheinlichkeit, daß in einer zufällig gewählten Textsequenz von 10 Phonemen mindestens 8 Vokale vorkommen? [Setzen Sie die Unabhängigkeit voraus.]
- 3.10. Betrachten wir den Namen MÜLLER und bilden aus seinen Buchstaben Wörter, die 3 Buchstaben enthalten. Wie ist die Wahrscheinlichkeit,
 (a) daß in einem Wort beide L vorkommen,
 (b) daß am Anfang und am Ende des Wortes L steht,
 (c) daß in einem Wort beide L nebeneinander stehen?
 In den Aufgaben (d) bis (j) dürfen die Buchstaben wiederholt werden (= Ziehen mit Zurücklegen). Wie ist die Wahrscheinlichkeit,
 (d) daß man ein Wort mit zwei L bildet,
 (e) daß man ein Wort genau mit einem L bildet,
 (f) daß man ein Wort nur mit Vokalen bildet,
 (g) daß man ein Wort nur mit Konsonanten bildet,
 (h) daß man ein Wort mit einem Vokal und zwei Konsonanten bildet,
 (i) daß man ein Wort mit zwei Vokalen und einem Konsonanten bildet,
 (j) daß man ein Wort bildet, in dem beide L nebeneinander stehen?
- 3.11. In einer Sprache werden zweisilbige Wörter folgendermaßen gebildet: Die Wahrscheinlichkeit, daß die erste Silbe des Wortes offen ist (auf Vokal endet), ist p_1 , und die Wahrscheinlichkeit, daß die zweite Silbe offen ist, beträgt p_2 . Finden Sie die Wahrscheinlichkeit,
 (a) daß beide Silben eines Wortes offen sind,
 (b) daß beide Silben geschlossen sind,
 (c) daß die erste geschlossen, die zweite offen ist,
 (d) daß mindestens eine geschlossen ist.
- 3.12. Aus dem Wort SEMANTIK nimmt man zufällig 3 Buchstaben. Wie ist die Wahrscheinlichkeit,
 (a) daß alle drei Vokale sind,
 (b) daß alle drei Konsonanten sind,
 (c) daß es ein Vokal und 2 Konsonanten sind?

- 3.13. Aus einem Kartenspiel werden 8 Karten auf den Tisch gelegt. Aus einem anderen Spiel werden dieselben Karten gewählt und zufällig unter die ersten gelegt. Wie ist die Wahrscheinlichkeit,
 (a) daß mindestens ein Paar identisch ist,
 (b) daß genau vier Paare identisch sind?
- 3.14. Ein bestimmtes Wort A kommt in der Sprache mit der Wahrscheinlichkeit p vor. Man untersucht k Texte, von denen jeder aus n Wörtern besteht. Wie ist die Wahrscheinlichkeit,
 (a) daß das Wort A in mindestens einem Text vorkommt,
 (b) daß es in allen k Texten vorkommt?
- 3.15. Vier sich reimende Wortpaare ($aabbccdd$) wurden vermischt und zufällig in eine Reihe gestellt. Wie ist die Wahrscheinlichkeit,
 (a) daß alle Reimpaare nebeneinander stehen,
 (b) daß mindestens ein Reimpaar nebeneinander bleibt?
- 3.16. Wenn A, B, C beliebige Ereignisse sind, drücke aus:
 (a) genau ein Ereignis kommt vor;
 (b) genau zwei Ereignisse kommen vor;
 (c) alle drei kommen vor;
 (d) nur A kommt vor;
 (e) keines von ihnen kommt vor;
 (f) mindestens eins kommt vor;
 (g) nur A und B , aber nicht C kommen vor;
 (h) mindestens zwei kommen vor;
 (i) nicht mehr als zwei kommen vor.
- 3.17. Der erste Abschnitt eines Textes enthält 20 Substantive, von denen 5 Neutra sind. Zufällige Anordnung der Genera vorausgesetzt, wie ist die Wahrscheinlichkeit, daß
 (a) das erste Substantiv ein Neutrum und das zweite ein Femininum oder Maskulinum ist,
 (b) das zweite kein Neutrum ist,
 (c) das zweite ein Neutrum ist?
- 3.18. In einer Sprache gibt es 20 Konsonanten und 5 Vokale. Ein Computer bildet aus ihnen Namen der Struktur KVKVK. Wie ist die Wahrscheinlichkeit, daß sich bei einem zufällig aus der Liste gewählten Namen die Phoneme nicht wiederholen?
- 3.19. Wenn A und B unabhängig sind, dann gilt $P(AB) = P(A)P(B)$. Beweisen Sie, daß unter dieser Bedingung auch $P(\overline{A}\overline{B}) = P(\overline{A})P(\overline{B})$ gilt.
- 3.20. Fünf Studenten bereiten sich auf die Prüfung in Statistik vor. Von 10 Fragen, die bekannt sind, soll jeder Student 2 beantworten. Eine bereits gestellte Frage darf nicht mehr gestellt werden. Der Student A hat sich nur für zwei bestimmte Fragen vorbereitet. Wie ist die Wahrscheinlichkeit, daß er genau die zwei Fragen bekommt, auf die er sich vorbereitet hat?
- 3.21. Die Buchstaben des Wortes LINGUISTIK werden auf Zettel geschrieben und in eine Urne gelegt. Man zieht die Buchstaben hintereinander (ohne Zurücklegen)

heraus. Wie ist die Wahrscheinlichkeit, daß man sie in der ursprünglichen oder umgekehrten Reihenfolge herauszieht?

3.22. Ein Computer bekommt 3 Adjektive (A), 5 Substantive (S) und 4 Verben (V) und soll zufällig Verse von 3 Wörtern bilden, wobei aus jeder Gruppe nur ein Wort gewählt werden darf. Wie ist die Wahrscheinlichkeit,

(a) daß der erste Vers aus einem A, einem S und einem V besteht;

(b) daß er die Struktur ASV hat?

3.23. In einer Sprache gibt es dreimal so viele Konsonanten wie Vokale. Die Phoneme sind auf Zetteln in einer Urne. Man zieht einen Zettel. Wie ist die Wahrscheinlichkeit, daß man einen Konsonanten zieht?

3.24. Beweisen Sie, daß

$$P(\bar{A}_1 \cup \bar{A}_2 \cup \bar{A}_3) = 1 - P(A_1) P(A_2|A_1) P(A_3|A_2 A_1).$$

3.25. Das Wort A kommt in der Sprache mit der Wahrscheinlichkeit p_1 vor, das Wort B mit der Wahrscheinlichkeit p_2 . Wie ist die Wahrscheinlichkeit,

(a) daß keines von beiden in einem Text mit der Länge n vorkommt;

(b) daß nur A vorkommt;

(c) daß nur B vorkommt;

(d) daß beide vorkommen;

(e) daß mindestens eins von ihnen vorkommt;

(f) daß höchstens eins von ihnen vorkommt;

(g) daß nur eins von ihnen vorkommt (nicht beide)?

Benutzte und weiterführende Literatur: Brunk (1965), Feller (1962), Greenberg (1956), Hald (1967), Harris (1966), Job (1980), Kapur-Saxena (1970), Mood-Graybill (1963), Papoulis (1965), Ventcel'-Ovčarov, (1973), Whitworth (1965).

4. Zufallsvariable

4.1. Diskrete Zufallsvariable

Ausgänge von Experimenten brauchen nicht numerischer Art zu sein, sie können nominal bezeichnet werden, wie z.B. "Kopf" beim Werfen einer Münze, "Vorkommen von /b/" beim Zählen von Phonemen usw. Diese Ausgänge sind Elemente der Ereignismenge S , und wir betrachten sie als Elementarereignisse. Jedem Ausgang kann man jedoch eine Zahl zuschreiben und damit eine Funktion festsetzen. Der Ausgang des Experiments ist nicht fest, er läßt sich nicht eindeutig im voraus bestimmen, daher ist der Wert, den diese Funktion bei einem Experiment annimmt, sozusagen vom Zufall abhängig. Deswegen bezeichnen wir diese Funktion als eine *Zufallsvariable*. Ihr Definitionsbereich ist die Menge aller Ereignisse (Ausgänge) S , ihr Wertebereich sind reelle Zahlen.

Die Zufallsvariable X ist diskret, wenn sie nur endlich viele oder abzählbar unendlich viele Werte auf der Zahlengerade annehmen kann. Die Werte x , die die Zufallsvariable X annimmt, können positiv, negativ oder 0 sein. Die Wahrscheinlichkeit, mit der X einen bestimmten Wert x annimmt, symbolisieren wir mit

$$P(X = x) = P_x$$

und definieren die sogenannte Wahrscheinlichkeitsfunktion (WF) als

$$f(x) = \begin{cases} P(X = x) & \text{für } x = n_1, n_2, n_3, \dots \\ 0 & \text{sonst} \end{cases} \quad (4.1)$$

wobei mit n_i die Zahlen bezeichnet werden, die x durchläuft. Wir werden alle drei Bezeichnungen, $f(x)$, P_x und $P(X = x)$ alternativ benutzen. Für diese WF gelten die Axiome der Wahrscheinlichkeit, nämlich

$$(1) f(x) \geq 0.$$

$$(2) \sum_x f(x) = 1.$$

(3) Wenn A eine Teilmenge des Ereignisraumes ist, dann gilt

$$P(A) = \sum_{x \in A} f(x).$$

Schließlich gilt auch

$$(4) P(X = -\infty) = P(X = \infty) = 0 \quad (4.2)$$

Durch die WF wird die Verteilung der Zufallsvariablen vollständig festgesetzt.

Beispiel 4.1.1. Eine WF sei folgendermaßen bestimmt

$$f(x) = \begin{cases} 1/6 & \text{für } x = 0 \\ 1/3 & \text{für } x = 1 \\ 1/3 & \text{für } x = 2 \\ 1/6 & \text{für } x = 3 \\ 0 & \text{sonst.} \end{cases}$$

Stellen Sie fest, ob die Bestimmung korrekt ist, d.h., ob es sich tatsächlich um eine WF handelt.

Lösung: Es reicht festzustellen, ob die ersten zwei Bedingungen erfüllt sind. Es gilt offensichtlich, daß $f(x) \geq 0$, weil jedes $P_x > 0$ und $0 \geq 0$; weiter ist

$$\sum_x f(x) = 1/6 + 1/3 + 1/3 + 1/6 = 1,$$

so daß die Bestimmung korrekt ist. ■

Beispiel 4.1.2. Bestimmen Sie die Verteilung der Anzahl der "Köpfe" beim Werfen von 3 Münzen.

Lösung: X = Anzahl der Köpfe beim Werfen von 3 Münzen. Die Zufallsvariable kann die Werte $x = 0, 1, 2, 3$ annehmen. Da das Experiment bei jeder Münze 2 Ausgänge hat (Kopf, Zahl) besteht der Ereignisraum S aus $2 \cdot 2 \cdot 2 = 2^3$ elementaren Ereignissen, d.h. $N = 2^3$:

Z	Z	Z			
Z	Z	K	Z	K	Z
Z	K	K	K	Z	K
K	K	K			
				K	Z
				K	K
					Z

Die Zahl der Möglichkeiten, z.B. beim Wurf von 3 Münzen 1 Kopf zu bekommen, ist die "Wahl" eines Elements aus 3, was auf $\binom{3}{1}$ Weisen durchgeführt werden kann. Dasselbe gilt auch für 2 mal Kopf, nämlich $\binom{3}{2}$ usw. Die Wahrscheinlichkeit, einmal Kopf zu bekommen, ist dann

$$P(X=1) = \frac{\binom{3}{1}}{2^3},$$

und ähnlich für die anderen x . Daher definieren wir die WF als

$$f(x) = \begin{cases} \frac{\binom{3}{x}}{2^3} & \text{für } x = 0, 1, 2, 3 \\ 0 & \text{sonst.} \end{cases}$$

oder, wenn wir alle P_x ausschreiben, als

$$f(x) = \begin{cases} 1/8 & \text{für } x = 0 \\ 3/8 & \text{für } x = 1 \\ 3/8 & \text{für } x = 2 \\ 1/8 & \text{für } x = 3 \\ 0 & \text{sonst.} \end{cases}$$

Wir überzeugen uns leicht, daß

- (1) $f(x) \geq 0$, weil jedes P_x im Definitionsbereich ($x = 0, 1, 2, 3$) größer als Null ist;
- (2) $\sum_x P_x = 1$ auch für den allgemeinen Fall des Würfens von n Münzen gilt. Das folgt aus Formel (2.19) im Beispiel 2.8.5. ■

Diskrete Verteilungen lassen sich immer in Form eines Stabdiagramms darstellen. So ist die Verteilung im Beispiel 4.1.1 durch das Stabdiagramm in Abb. 4.1 dargestellt.

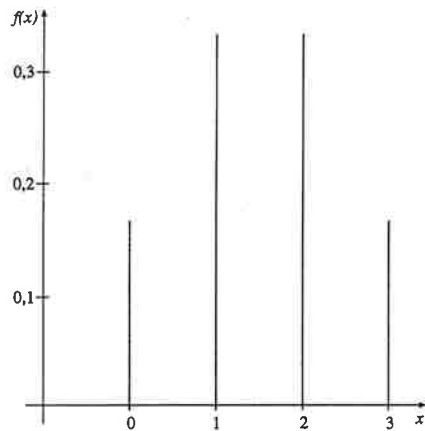


Abb. 4.1. Stabdiagramm der Verteilung in Beispiel 4.1.1.

Bisher haben wir als Ereignisse lediglich solche Mengen betrachtet, in denen X den Wert x annahm, d.h. $\{X = x\}$, mit der Wahrscheinlichkeit $P(X = x)$ (vgl. Formel 4.1). Auf Grund des Axioms III können wir jedoch auch solche Mengen als Ereignisse zulassen, in denen X einen Wert kleiner oder größer als x annimmt, d.h. Mengen $\{X \leq x\}$, $\{X < x\}$, $\{X \geq x\}$, $\{X > x\}$, die Teilmengen des Ereignisraums darstellen und deren Wahrscheinlichkeit sich durch Summierung berechnen lässt. Die Wahrscheinlichkeit des ersten dieser Ereignisse nennen wir Verteilungsfunktion der Zufallsvariablen X und bezeichnen sie als

$$F(x) = P(X \leq x) = \sum_{x' \leq x} f(x'), \quad (4.3)$$

wobei x eine feste Zahl ist. Zur Unterscheidung symbolisieren wir die laufenden Werte mit x' und den festen Wert mit x .

Für $F(x)$ gelten folgende Beziehungen:

$$(1) F(X = -\infty) = 0. \quad (4.4)$$

Beweis folgt direkt aus (4.2). ■

$$(2) F(X = \infty) = 1. \quad (4.5)$$

$$\text{Beweis: } F(X = \infty) = P(X \leq \infty) = \sum_x f(x) = P(S) = 1. \blacksquare$$

$$(3) 0 \leq F(x) \leq 1. \quad (4.6)$$

Beweis folgt direkt aus (4.4) und (4.5). ■

$$(4) \text{ Wenn } x_1 < x_2, \text{ dann ist } F(x_1) \leq F(x_2). \quad (4.7)$$

Beweis: Wenn $x_1 < x_2$, dann ist $\{X \leq x_1\} \subset \{X \leq x_2\}$, d.h. das links offene Intervall $\{X \leq x_1\}$ ist im links offenen Intervall $\{X \leq x_2\}$ enthalten, (vgl. Abb. 4.2), daher $P(X \leq x_1) \leq P(X \leq x_2)$, woraus (4.7) folgt. ■

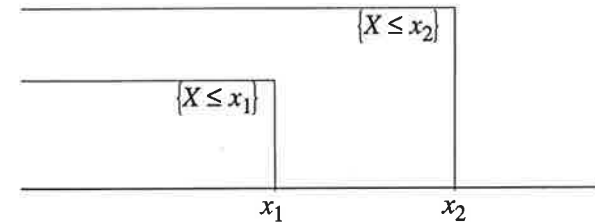


Abb. 4.2. Links offene Intervalle

$$(5) F(x) = 1 - P(X > x). \quad (4.8)$$

Beweis: $P(X \leq x) + P(X > x) = 1$, da beide Ereignisse komplementär sind. Durch einfache Umordnung folgt (4.8). ■

Beispiel 4.1.3. Betrachten wir die WF aus Beispiel 4.1.2 und berechnen

(a) ihre Verteilungsfunktion,

(b) $P(X > 1)$.

Lösung: (a) nach der Formel (4.3) haben wir

$$F(0) = P(X \leq 0) = \frac{1}{8}$$

$$F(1) = P(X \leq 1) = \frac{1}{8} + \frac{3}{8} = \frac{4}{8}$$

$$F(2) = P(X \leq 2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$$

$$F(3) = P(X \leq 3) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1,$$

woraus man die Verteilungsfunktion wie folgt darstellen kann

$$F(x) = \begin{cases} 0 & \text{für } x < 0 \text{ oder für } -\infty < x < 0 \\ \frac{1}{8} & \text{für } x = 0 \text{ oder für } 0 \leq x < 1 \\ \frac{4}{8} & \text{für } x = 1 \text{ oder für } 1 \leq x < 2 \\ \frac{7}{8} & \text{für } x = 2 \text{ oder für } 2 \leq x < 3 \\ 1 & \text{für } x = 3 \text{ oder für } x \geq 3 \end{cases}$$

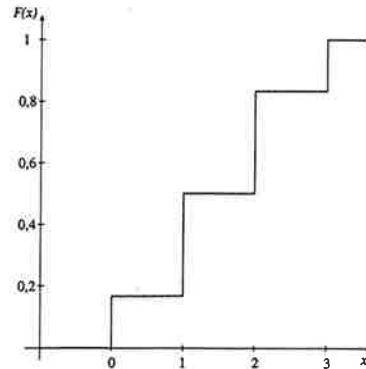


Abb. 4.3. Treppenkurve der Verteilungsfunktion in Beispiel 4.1.3.

Man sieht, daß alle Beziehungen (4.4) bis (4.8) erfüllt sind.

(b) Um $P(X > 1)$ zu berechnen, benutzen wir (4.8) und setzen $P(X > 1) = 1 - F(1) = 1 - \frac{4}{8} = \frac{4}{8}$. Dasselbe Resultat bekommen wir, wenn wir direkt summieren, nämlich

$$P(X > 1) = P(X = 2) + P(X = 3) = \frac{3}{8} + \frac{1}{8} = \frac{4}{8}. \blacksquare$$

Die graphische Darstellung der Verteilungsfunktion einer diskreten Zufallsvariablen ergibt die bereits bekannte Treppenfunktion. Für das Beispiel 4.1.3 ist $F(x)$ in Abb. 4.3 dargestellt.

Die dritte Art von Ereignissen, deren Wahrscheinlichkeit uns interessiert, sind vier unterschiedliche Intervalle $\{x_1 < X \leq x_2\}$, $\{x_1 \leq X \leq x_2\}$, $\{x_1 < X < x_2\}$, $\{x_1 \leq X < x_2\}$, wobei x_i wieder reelle Zahlen sind.

(6) Für das erste Intervall beweisen wir, daß

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1). \quad (4.9)$$

Beweis: Wir teilen die Zahlengerade in zwei disjunkte Intervalle $\{X \leq x_1\}$ und $\{x_1 < X \leq x_2\}$, deren Vereinigung offensichtlich $\{X \leq x_2\}$ ergibt, wie man sich durch einen Blick auf Abb. 4.2 leicht vergewissert. Daher ist die Wahrscheinlichkeit

$$P(x_1 < X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = F(x_2) - F(x_1). \blacksquare$$

Beispiel 4.1.4. Man sucht $P(0 < X \leq 2)$ für die Verteilung in Beispiel 4.1.3.

Lösung: Benutzt man die Verteilungsfunktion, so bekommt man sofort

$$F(2) - F(0) = \frac{7}{8} - \frac{1}{8} = \frac{6}{8}.$$

Benutzt man die WF, so muß man summieren, in diesem konkreten Fall jedoch nur zwei Summanden, und bekommt

$$P(0 < X \leq 2) = P(X = 1) + P(X = 2) = \frac{3}{8} + \frac{3}{8} = \frac{6}{8}. \blacksquare$$

Übung 4.1.1. Stellen Sie fest, ob die Funktion $f(x)$ eine WF ist, wenn sie folgendermaßen gegeben ist:

$$(a) f(x) = \begin{cases} \frac{1}{6} & \text{für } x = 3 \\ \frac{2}{6} & \text{für } x = 5 \\ \frac{3}{6} & \text{für } x = 6 \\ 0 & \text{sonst} \end{cases} \quad (b) f(x) = \begin{cases} \frac{1}{7} & \text{für } x = 1 \\ \frac{3}{7} & \text{für } x = 2 \\ \frac{2}{7} & \text{für } x = 3 \\ 0 & \text{sonst.} \end{cases}$$

Übung 4.1.2. Zeigen Sie, daß für eine diskrete Zufallsvariable gilt:

- (a) $P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1 - 1)$;
- (b) $P(x_1 < X < x_2) = F(x_2 - 1) - F(x_1)$;
- (c) $P(x_1 \leq X < x_2) = F(x_2 - 1) - F(x_1 - 1)$.

[Bedenken Sie, daß für eine diskrete Zufallsvariable $P(X < x) = P(X \leq x - 1) = F(x - 1)$.]

Übung 4.1.3. Eine WF sei folgendermaßen gegeben:

$$f(x) = \begin{cases} \frac{1}{12} & \text{für } x = -1 \\ \frac{1}{4} & \text{für } x = 0 \\ \frac{1}{6} & \text{für } x = 1 \\ \frac{1}{12} & \text{für } x = 2 \\ \frac{1}{3} & \text{für } x = 3 \\ \frac{1}{12} & \text{für } x = 4 \\ 0 & \text{sonst.} \end{cases}$$

- (a) Überprüfen Sie, ob es eine WF ist;
- (b) finden Sie die Verteilungsfunktion $F(x)$;
- (c) berechnen Sie $P(X \geq 2)$;
- (d) $F(2)$;
- (e) $P(1 < X \leq 3)$;
- (f) $P(1 \leq X < 3)$;
- (g) $F(2) - F(-1)$;
- (h) $P(X < 3)$;
- (i) $P(|X| \leq 1)$.

4.2. Stetige Zufallsvariable

Eine stetige Zufallsvariable ist ein Analogon zu der diskreten Zufallsvariablen. Sie nimmt reelle Werte an, und ihre Verteilungsfunktion $F(x)$ läßt sich mit Hilfe eines Integrals darstellen. Die WF $f(x)$ nennt man auch Dichte (oder Wahrscheinlichkeitsdichte). Eine Zufallsvariable heißt stetig, wenn sie "kontinuierlich genug" ist, daß man über ihr nicht verschwindende Integrale bilden kann. Sie darf höchstens abzählbar unendlich viele Unstetigkeitsstellen haben. Es gelten für sie die Axiome der Wahrscheinlichkeit, nämlich

$$(1) \quad f(x) \geq 0.$$

$$(2) \quad \int_{-\infty}^{\infty} f(x) dx = 1.$$

(3) Seien $A_1, A_2, \dots, A_k$ disjunkte Ereignisse, deren Vereinigung das Ereignis A darstellt, dann ist

$$\begin{aligned} P(a) &= P(A_1 \cup A_2 \cup \dots \cup A_k) = \int_A f(x) dx = \\ &= \int_{A_1} f(x) dx + \int_{A_2} f(x) dx + \dots + \int_{A_k} f(x) dx. \end{aligned}$$

Für die Verteilungsfunktion gilt

$$(1) \quad F(x) = P(X \leq x) = \int_{-\infty}^x f(x') dx'. \quad (4.10)$$

$$(2) \quad F(-\infty) = 0; \quad F(\infty) = 1; \quad 0 \leq F(x) \leq 1.$$

$$(3) \quad P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x) dx = F(x_2) - F(x_1) \quad (4.11)$$

Dieser Ausdruck stellt die Fläche unter der Kurve $f(x)$ zwischen den Punkten x_1 und x_2 dar.

(4) In einem beliebigen Punkt x ist $P(X = x) = 0$, so daß wir es hier nur mit Wahrscheinlichkeiten von Intervallen zu tun haben.

$$\text{Beweis: } \int_{x_1}^{x_1} f(x) dx = F(x_1) - F(x_1) = 0. \blacksquare$$

(5) Die Verteilungsfunktion einer stetigen Zufallsvariablen ist eine stetige, nicht fallende Funktion, d.h. eine "glatte" nicht fallende Kurve. Ein Beispiel zeigt Abb. 4.4.

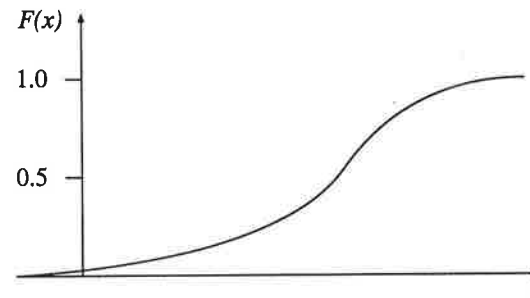


Abb. 4.4. Beispiel einer Verteilungsfunktion

$$(6) \quad F'(x) = f(x) \quad (4.12)$$

d.h. die Ableitung der Verteilungsfunktion ergibt die Dichte.

Beweis: Folgt direkt aus der Definition der Verteilungsfunktion [vgl. (4.10)].
■

$$(7) \text{ Wenn } x_1 < x_2, \text{ dann ist } F(x_1) \leq F(x_2). \blacksquare$$

Bemerkung: Im Unterschied zu einer diskreten Zufallsvariablen ist

$$P(X < x) = P(X \leq x).$$

Beweis: Offensichtlich gilt $P(X < x) \leq P(X \leq x)$, so daß es reicht zu zeigen, daß $P(X < x) < P(X \leq x)$ nicht gilt. Nehmen wir an, daß

$$P(X < x) < P(X \leq x) \quad (a)$$

gilt. In dem Fall ist

$$\frac{1}{2} [P(X \leq x) - P(X < x)] = \varepsilon > 0 \quad (b)$$

Wegen Stetigkeit gilt der einheitliche Grenzwert

$$\lim_{n \rightarrow \infty} P(X \leq x - \frac{1}{n}) = P(X \leq x),$$

d.h. man kann $n \geq n_0 \in N$ so wählen, daß

$$P(X \leq x) - P(X \leq x - \frac{1}{n}) < \varepsilon$$

oder

$$P(X \leq x - \frac{1}{n}) > P(X \leq x) - \varepsilon.$$

Setzt man ε aus (b) ein, so wird

$$\begin{aligned} P(X \leq x - \frac{1}{n}) &> P(X \leq x) - \frac{1}{2} [P(X \leq x) - P(X < x)] = \\ &= \frac{1}{2} [P(X \leq x) + P(X < x)]. \end{aligned}$$

Wegen der Annahme (a) gilt dann

$$\begin{aligned} P(X \leq x - \frac{1}{n}) &> \frac{1}{2} [P(X \leq x) + P(X < x)] > \\ &> \frac{1}{2} [P(X < x) + P(X < x)] = \\ &= P(X < x), \end{aligned}$$

d.h. $P(X \leq x - \frac{1}{n}) > P(X < x)$ für alle $n \geq n_0 \in N$.

Da aber für alle $n \in N$ wegen Monotonie von P gilt

$$P(X \leq x - \frac{1}{n}) \leq P(X < x),$$

so führt die Annahme (a) zum Widerspruch. Daher gilt die Behauptung. Aus dieser Behauptung folgt, daß im Unterschied zu einer diskreten Zufallsvariablen

$$P(x_1 < X \leq x_2) = P(x_1 \leq X \leq x_2)$$

$$= P(x_1 \leq X < x_2)$$

$$= P(x_1 < X < x_2). \blacksquare$$

Beispiel 4.2.1. Wir stellen fest, ob die angegebenen Funktionen Dichten darstellen.

(a) $f(x) = 2x$ für $1 \leq x \leq 2$ und $f(x) = 0$ sonst.

$$(b) f(x) = \begin{cases} x & 0 < x \leq 1 \\ (2-x) & 1 < x \leq 2 \\ 0 & \text{sonst.} \end{cases}$$

$$\text{Lösung: (a) } f(x) \geq 0; \int_1^2 2x \, dx = 2 \left[\frac{x^2}{2} \right]_1^2 = 2 \left(2 - \frac{1}{2} \right) = 3 \neq 1,$$

also ist es keine Wahrscheinlichkeitsdichte.

$$(b) f(x) \geq 0; \int_0^1 x \, dx + \int_1^2 (2-x) \, dx = \left[\frac{x^2}{2} \right]_0^1 + \left[2x - \frac{x^2}{2} \right]_1^2 = \frac{1}{2} + \left(4 - 2 - \frac{1}{2} \right) =$$

$= 1$, also ist es eine korrekte Dichte. $\blacksquare$

Beispiel 4.2.2. Wir berechnen die Verteilungsfunktion der Zufallsvariablen mit der Dichte im Beispiel 4.2.1(b).

Lösung: Im Intervall $-\infty < x \leq 0$ ist

$$F(x) = \int_{-\infty}^x 0 \, dx' = 0.$$

Im Intervall $0 < x \leq 1$ ist

$$F(x) = \int_{-\infty}^0 0 \, dx + \int_0^x x' \, dx' = 0 + \left[\frac{x'^2}{2} \right]_0^x = \frac{x^2}{2}.$$

Im Intervall $1 < x \leq 2$ ist

$$F(x) = \int_{-\infty}^0 0 \, dx + \int_0^1 x \, dx + \int_1^x (2-x') \, dx' = \left[\frac{x^2}{2} \right]_0^1 + \left[2x' - \frac{x'^2}{2} \right]_1^x = \frac{1}{2} + 2x - 2 - \frac{x^2}{2} + \frac{1}{2} = -1 + 2x - \frac{x^2}{2}.$$

Im Intervall $2 < x < \infty$ ist $F(x) = 1$, wie im vorigen Beispiel schon gezeigt wurde. Wir können also schreiben

$$F(x) = \begin{cases} 0 & -\infty < x \leq 0 \\ \frac{x^2}{2} & 0 < x \leq 1 \\ -1 + 2x - \frac{x^2}{2} & 1 < x \leq 2 \\ 1 & 2 < x < \infty \end{cases} \quad \blacksquare$$

Beispiel 4.2.3. Wir berechnen mit Hilfe der Verteilungsfunktion im vorigen Beispiel

- (a) $F(0,5)$; (c) $P(0,5 < X \leq 1,5)$;
(b) $F(1,5)$; (d) $P(X > 1)$.

Lösung: (a) Wenn $x = 0,5$, dann benutzen wir $F(x)$ im Intervall $0 < x \leq 1$, d.h.

$$F(0,5) = \frac{(0,5)^2}{2} = \frac{0,25}{2} = 0,125.$$

(b) Für $x = 1,5$ benutzen wir $F(x)$ im Intervall $1 < x \leq 2$, daher ist

$$F(1,5) = -1 + 2(1,5) - \frac{(1,5)^2}{2} = -1 + 3 - 1,125 = 0,875.$$

(c) $P(0,5 < X \leq 1,5) = F(1,5) - F(0,5) = 0,875 - 0,125 = 0,75$.

(d) $P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = 1 - \frac{1^2}{2} = \frac{1}{2}$. $\blacksquare$

Beispiel 4.2.4. Wir bestimmen k für die Dichte

$$f(x) = kx^2 \text{ für } 1 < x \leq 2 \text{ und } f(x) = 0 \text{ sonst.}$$

Lösung: Wenn $f(x)$ eine Dichte ist, so muß $\int_1^2 f(x) \, dx = 1$ sein. Daher ist

$$\int_1^2 kx^2 \, dx = \left[\frac{kx^3}{3} \right]_1^2 = k \left(\frac{8}{3} - \frac{1}{3} \right) = 1,$$

woraus $k = \frac{3}{7}$ folgt. Wir überzeugen uns, ob die Rechnung stimmt, indem wir integrieren

$$\int_1^2 \frac{3}{7} x^2 \, dx = \frac{3}{7} \left[\frac{x^3}{3} \right]_1^2 = \frac{3}{7} \left(\frac{8}{3} - \frac{1}{3} \right) = 1. \quad \blacksquare$$

Übung 4.2.1. Sei $f(x) = k(1 + 2x)$ für $0 \leq x < 2$ und $f(x) = 0$ sonst eine Dichte. Finden Sie k [vgl. B2].

Übung 4.2.2. Sei

$$f(x) = \begin{cases} 2x^2 & 0 \leq x < 1 \\ \frac{2}{9}x & 1 \leq x < 2 \\ 0 & \text{sonst.} \end{cases}$$

- (a) Stellen Sie fest, ob es eine Dichte ist;
(b) finden Sie die Verteilungsfunktion;
berechnen Sie
(c) $F(1,5)$; (e) $P(X \geq 1,7)$;
(d) $F(0,4)$; (f) $P(0,4 < X \leq 1,4)$.

Übung 4.2.3. Stellen Sie fest, ob $f(x) = ae^{-ax}$ für $0 < x < \infty$ und $f(x) = 0$ sonst eine Dichte ist.

4.3. Bedingte Verteilung

Genauso wie bei der bedingten Wahrscheinlichkeit kann man fragen, wie die Verteilung der Zufallsvariablen beschaffen ist, wenn wir wissen, daß sie nur in einem Intervall ihres Definitionsbereichs Werte annimmt, nämlich $f(x | X \leq x_2)$ oder $f(x | X > x_1)$ oder $f(x | x_1 < X \leq x_2)$, wobei x_1 und x_2 Konstanten sind. Die Verteilung von X ist bekannt, aber hier auf ein Intervall beschränkt. Wir bezeichnen dieses Intervall als I und berechnen zuerst die Verteilungsfunktion der Zufallsvariablen, aus der sich mit Hilfe von (4.12) die WF (Dichte) einer bedingten Verteilung durch Differenzierung berechnen läßt. Die Verteilungsfunktion ist definiert als

$$F(x | I) = P(X \leq x | I) = \frac{P(X \leq x, I)}{P(I)} \quad (4.13)$$

Wie oben angedeutet, müssen wir drei Fälle unterscheiden:

(1) Sei $I = \{X \leq x_2\}$, dann haben wir

$$F(x|I) = F(x|X \leq x_2) = P(X \leq x|X \leq x_2) = \frac{P(X \leq x, X \leq x_2)}{P(X \leq x_2)}. \quad (4.14)$$

Wenn nun $x \geq x_2$, dann ist $\{X \leq x_2\}$ nur eine Teilmenge von $\{X \leq x\}$, und der Durchschnitt beider Mengen ergibt

$$\{X \leq x, X \leq x_2\} = \{X \leq x_2\}.$$

Daher ist nach (4.14)

$$F(x|I) = \frac{P(X \leq x_2)}{P(X \leq x_2)} = 1 \quad \text{für } x \geq x_2 \quad (4.15)$$

Wenn aber $x < x_2$, dann ist

$$\{X \leq x, X \leq x_2\} = \{X \leq x\}$$

und wir haben nach (4.14)

$$F(x|I) = \frac{P(X \leq x)}{P(X \leq x_2)} = \frac{F(x)}{F(x_2)} \quad \text{für } x < x_2. \quad (4.16)$$

Die bedingte Dichte bekommen wir laut (4.12) durch Differenzierung als

$$f(x|I) = F'(x|I).$$

Differenzieren wir (4.15) und (4.16), so bekommen wir

$$f(x|X \leq x_2) = \begin{cases} \frac{f(x)}{F(x_2)} & \text{für } x < x_2 \\ 0 & \text{für } x \geq x_2 \end{cases} \quad (4.17)$$

Für diskrete Zufallsvariable ist $F(x_2) = \sum_{x \leq x_2} f(x)$, für stetige ist $F(x_2) = \int_{-\infty}^{x_2} f(x) dx$.

Beispiel 4.3.1. Die Zufallsvariable X hat die Dichte $f(x) = ae^{-ax}$ für $x > 0$ und $f(x) = 0$ sonst (Exponentialverteilung). Dabei ist $a > 0$ eine Konstante. Wir suchen die bedingte Dichte von X , wenn wir wissen, daß X nur Werte bis höchstens 5 annimmt.

Lösung: Nach (4.17) ist $x_2 = 5$, und

$$F(5) = \int_0^5 ae^{-ax} dx = 1 - e^{-5a}$$

woraus nach Einsetzen in (4.17) folgt, daß

$$f(x|X \leq 5) = \begin{cases} \frac{ae^{-ax}}{1 - e^{-5a}} & \text{für } 0 < x \leq 5 \\ 0 & \text{sonst.} \blacksquare \end{cases}$$

(2) Sei $I = \{X > x_1\}$. In dem Fall haben wir

$$F(x|I) = F(x|X > x_1) = P(X \leq x|X > x_1) = \frac{P(X \leq x, X > x_1)}{P(X > x_1)}. \quad (4.18)$$

Wenn $x \leq x_1$, dann ist offensichtlich $\{X \leq x, X > x_1\} = \emptyset$, und daraus folgt auch $F(x|I) = 0$. Wenn $x > x_1$, dann ist

$$\{X \leq x, X > x_1\} = \{x_1 < X \leq x\},$$

und man bekommt für (4.18)

$$F(x|I) = \frac{P(x_1 < X \leq x)}{P(X > x_1)} = \frac{F(x) - F(x_1)}{1 - F(x_1)}. \quad (4.19)$$

Die Ableitung von $F(x|I)$ ergibt die gesuchte Dichte als

$$f(x|X > x_1) = \begin{cases} \frac{f(x)}{1 - F(x_1)} & \text{für } x > x_1 \\ 0 & \text{sonst.} \end{cases} \quad (4.20)$$

Für diskrete Zufallsvariable ist

$$1 - F(x_1) = \sum_{x \geq x_1 + 1} f(x) ,$$

und für stetige ist

$$1 - F(x_1) = \int_{x_1}^{\infty} f(x) dx .$$

Beispiel 4.3.2. Sei die diskrete Dichte von X gegeben als

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \text{ für } x = 0, 1, 2, \dots \text{ und } f(x) = 0 \text{ sonst (Poisson-Verteilung).}$$

Wie ist die (bedingte) Dichte, wenn wir wissen, daß die Zufallsvariable nur die Werte der natürlichen Zahlen annehmen kann (d.h. den 0-Wert nicht). In der Praxis sind solche Fälle sehr häufig.

Lösung: Wir brauchen $f(x | X > 0)$ und müssen laut (4.20) $1 - F(0)$ berechnen. Es ist

$$1 - F(0) = 1 - P(X \leq 0) = 1 - P(X = 0) = 1 - \frac{e^{-\lambda} \lambda^0}{0!} = 1 - e^{-\lambda} .$$

daher ist

$$f(x | X > 0) = \begin{cases} \frac{\frac{e^{-\lambda} \lambda^x}{x!}}{1 - e^{-\lambda}} & \text{für } x = 1, 2, \dots \\ 0 & \text{sonst.} \blacksquare \end{cases}$$

(3) Sei $I = \{x_1 < X \leq x_2\}$, dann bekommen wir für (4.13)

$$F(x | I) = F(x | x_1 < X \leq x_2) = \frac{P(X \leq x, x_1 < X \leq x_2)}{P(x_1 < X \leq x_2)} \quad (4.21)$$

Wir müssen jetzt drei Fälle unterscheiden. Wenn $x \geq x_2$ ist, dann ist

$$\{X \leq x, x_1 < X \leq x_2\} = \{x_1 < X \leq x_2\}$$

und

$$F(x | I) = \frac{P(x_1 < X \leq x_2)}{P(x_1 < X \leq x_2)} = 1 \quad \text{für } x \geq x_2 \quad (4.22)$$

Wenn $x_1 \leq x < x_2$, dann ist

$$\{X \leq x, x_1 < X \leq x_2\} = \{x_1 < X \leq x\}$$

und

$$F(x | I) = \frac{P(x_1 < X \leq x)}{P(x_1 < X \leq x_2)} = \frac{F(x) - F(x_1)}{F(x_2) - F(x_1)} \quad \text{für } x_1 \leq x < x_2 . \quad (4.23)$$

Und schließlich, wenn $x < x_1$, dann ist

$$\{X \leq x, x_1 < X \leq x_2\} = \emptyset$$

Wonach

$$F(x | I) = 0 \quad \text{für } x < x_1 \quad (4.24)$$

Durch Differenzierung der Formeln (4.22) bis (4.24) bekommen wir dann

$$f(x | x_1 < X \leq x_2) = \begin{cases} \frac{f(x)}{F(x_2) - F(x_1)} & \text{für } x_1 \leq x < x_2 \\ 0 & \text{sonst.} \end{cases} \quad (4.25)$$

Beispiel 4.3.3. Sei $f(x) = ae^{-ax}$ für $x > 0$ und $f(x) = 0$ sonst. Wir suchen $f(x | 1 < X \leq 4)$.

Lösung: Wir finden erst

$$F(x_2) - F(x_1) = F(4) - F(1) = \int_1^4 f(x) dx = \int_1^4 ae^{-ax} dx = \left[-e^{-ax} \right]_1^4 = e^{-a} - e^{-4a}$$

woraus

$$f(x | 1 < X \leq 4) = \begin{cases} \frac{ae^{-ax}}{e^{-a} - e^{-4a}} & \text{für } 1 < x \leq 4 \\ 0 & \text{sonst} \end{cases}$$

folgt. ■

Übung 4.3.1. Sei $f(x) = \binom{n}{x} p^x q^{n-x}$ für $x = 0, 1, \dots, n$ und $f(x) = 0$ sonst (Binomialverteilung). Finden Sie $f(x | X > 0)$.

Übung 4.3.2. Sei $f(x) = 1$ für $0 < x \leq 1$ und $f(x) = 0$ sonst (Rechteckverteilung). Finden Sie $f(x | X \leq 0,5)$.

4.4. Funktionen von Zufallsvariablen

Manchmal interessiert man sich nicht für die Zufallsvariable selbst, deren Verteilung man kennt, sondern für eine andere, die eine Funktion der ersten ist. Würfelt man z.B. mit zwei Würfeln, so benötigt man nicht die Verteilung der Zufallsvariablen X , die der Anzahl der Augen eines Würfels entspricht, sondern die der Variablen Y , die der Summe der Augen zweier Würfel entspricht. Oder man benötigt aber die Verteilung einer sprachlichen Größe, kennt aber nur die einer anderen, deren Funktion diese ist. Ein anderes Beispiel ist folgendes: Man möchte wissen, wie die Differenz zwischen einer Zufallsvariablen und ihrem Mittelwert verteilt ist, um zu wissen, welche Abweichungen man erwarten kann (wir werden später sehen, daß die Signifikanztests der beurteilenden Statistik auf solchen Überlegungen beruhen). Diese und weitere spezielle Funktionen von Zufallsvariablen werden in Abschnitt 5.2 (Momente) eingeführt.

Um das Problem anschaulich zu machen, fangen wir mit einem konkreten Beispiel an.

Eine Zufallsvariable X hat die Dichte

$$f(x) = \begin{cases} 2x & \text{für } 0 < x < 1 \\ 0 & \text{sonst.} \end{cases} \quad (4.26)$$

Ihre Verteilungsfunktion ist, wie man leicht nachrechnet,

$$F(x) = \begin{cases} 0 & -\infty < x \leq 0 \\ x^2 & 0 < x < 1 \\ 1 & 1 \leq x < \infty \end{cases} \quad (4.27)$$

Wir transformieren nun die Zufallsvariable X in eine andere Zufallsvariable Y ,

$$Y = 3X + 1,$$

die eine Funktion von X ist, und fragen, wie die Dichte von Y aussieht. Wir zeigen zwei Berechnungsmöglichkeiten, die im Grunde identisch sind, und dann verallgemeinern wir sie und zeigen einige Spezialfälle.

I. Bezeichnen wir die Verteilungsfunktion von Y als $G(y) = P(Y \leq y)$. Wir setzen hier für Y die Funktion $3X + 1$ ein und bekommen

$$G(y) = P(Y \leq y) = P(3X + 1 \leq y) = P\left(X \leq \frac{y-1}{3}\right) = F\left(\frac{y-1}{3}\right).$$

Die Tatsache, daß $G(y) = F\left(\frac{y-1}{3}\right)$ ist, bedeutet, daß wir $G(y)$ bekommen, wenn wir in $F(x)$ in (4.27) das x durch $\frac{y-1}{3}$ ersetzen. Bemerken Sie, daß $x = \frac{y-1}{3}$ eben die Umkehrfunktion von $y = 3x + 1$ ist. Durch Einsetzung in (4.27) bekommen wir

$$G(y) = \begin{cases} 0 & -\infty < \frac{y-1}{3} \leq 0 \\ \left(\frac{y-1}{3}\right)^2 & 0 < \frac{y-1}{3} < 1 \\ 1 & 1 \leq \frac{y-1}{3} < \infty \end{cases} \quad (4.28)$$

Um daraus $g(y)$ zu bekommen, erinnern wir uns an die Tatsache, daß $F'(x) = f(x)$, d.h., wir differenzieren die Verteilungsfunktion und lösen die Intervalle in (4.28) nach y auf. Da

$$\left[\left(\frac{y-1}{3}\right)^2\right]' = 2\left(\frac{y-1}{3}\right) \frac{1}{3} = \frac{2}{9}(y-1)$$

und

$$(0)' = (1)' = 0$$

folgt daraus, daß

$$g(y) = \begin{cases} \frac{2}{9}(y-1) & 1 < y < 4 \\ 0 & \text{sonst.} \end{cases}$$

II. Die zweite Möglichkeit besteht aus einer Substitution der Zufallsvariablen unter dem Integral. Wir betrachten wieder die Dichte in (4.26). Die Wahrscheinlichkeit im Definitionsbereich, d.h. im Intervall (0,1) ist

$$P(0 < X < 1) = \int_0^1 2x \, dx. \quad (4.29)$$

Wir nehmen die Transformation $y = 3x + 1$ vor. Substituieren wir für x eine neue Variable, so müssen wir auch für dx eine Substitution vornehmen. Aus

$$y = 3x + 1 \quad (4.30)$$

finden wir die Umkehrfunktion

$$x = \frac{y-1}{3} \quad (4.31)$$

und durch Differenzieren nach y bekommen wir

$$\frac{dx}{dy} = \left(\frac{y-1}{3} \right)' = \frac{1}{3}$$

woraus

$$dx = \frac{1}{3} dy \text{ folgt.} \quad (4.32)$$

Setzt man nun (4.31) und (4.32) in (4.29) ein, unter gleichzeitiger Veränderung der Integrationsgrenzen, so bekommt man

$$\begin{aligned} P(0 < X < 1) &= \int_0^1 2x \, dx = \int_1^4 2 \left(\frac{y-1}{3} \right) \frac{1}{3} dy = \\ &= \int_1^4 \frac{2}{9} (y-1) dy = P(1 < Y < 4). \end{aligned}$$

Der Ausdruck unter dem letzten Integral ist eben $g(y)$, was mit dem oberen Resultat übereinstimmt. Die Grenzen findet man wieder so, daß man in (4.30) für x einmal 0 und einmal 1 einsetzt.

Dieses Resultat können wir nun folgendermaßen allgemein darstellen: Sei X eine Zufallsvariable mit der Dichte $f(x)$ und der Verteilungsfunktion $F(x)$; sei weiter $Y = h(X)$ eine Funktion von X und $X = h^{-1}(Y)$ die Umkehrfunktion.

Methode I. Wir unterscheiden hier zwei Fälle: (a) Y ist eine monoton wachsende Funktion, (b) Y ist eine monoton fallende Funktion.

(a) im ersten Fall ist

$$\begin{aligned} G(y) &= P(Y \leq y) = P[h(X) \leq y] = P[X \leq h^{-1}(y)] = \\ &= F[h^{-1}(y)]. \end{aligned} \quad (4.33)$$

Die Dichte $g(y)$ bekommen wir nach (4.12) aus $G(y)$ durch Differenzierung, d.h.

$$\begin{aligned} g(y) &= \frac{dG(y)}{dy} = G'(y) = \\ &= F'[h^{-1}(y)] = f[h^{-1}(y)] [h^{-1}(y)]' \end{aligned} \quad (4.34)$$

weil man erst die äußere Funktion F und dann die innere $h^{-1}(y)$ differenzieren muß.

(b) Ist Y eine monoton fallende Funktion, dann ist

$$\begin{aligned} G(y) &= P(Y \leq y) = P[h(X) \leq y] = P[X \geq h^{-1}(y)] = \\ &= 1 - F[h^{-1}(y)]. \end{aligned} \quad (4.35)$$

Durch Differenzierung bekommen wir

$$\begin{aligned}
 g(y) &= G'(y) = \left(1 - F[h^{-1}(y)]\right)' = \\
 &= -f[h^{-1}(y)] [h^{-1}(y)]'.
 \end{aligned}
 \quad (4.36)$$

Da aber die Ableitung einer monoton fallenden Funktion negativ ist, gilt

$$g(y) = f[h^{-1}(y)] \left| [h^{-1}(y)]' \right|. \quad (4.37)$$

Methode II. (a) Bei monoton wachsenden Funktionen $y = h(x)$ ersetzen wir x unter dem Integral durch $x = h^{-1}(y)$, weiter wird $dx = h^{-1}(y)' dy$, und für die Grenzen der Integration $(x_1, x_2]$ setzen wir $y_1 = h(x_1)$ und $y_2 = h(x_2)$ ein, so daß wir

$$\begin{aligned}
 P(x_1 < X \leq x_2) &= \int_{x_1}^{x_2} f(x) dx = \int_{y_1}^{y_2} f[h^{-1}(y)] [h^{-1}(y)]' dy = \\
 &= \int_{y_1}^{y_2} g(y) dy = P(y_1 < Y \leq y_2)
 \end{aligned}
 \quad (4.38)$$

bekommen. Daraus folgt offensichtlich

$$g(y) = f[h^{-1}(y)] [h^{-1}(y)]'$$

wie in (4.34).

(b) Wenn die Funktion $y = h(x)$ monoton fallend ist, so ist

$$dx = -[h^{-1}(y)]' dy,$$

weil die Ableitung einer monoton fallenden Funktion negativ ist, und $y_2 < y_1$. Daher wird

$$P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f(x) dx = \int_{y_1}^{y_2} f[h^{-1}(y)] [h^{-1}(y)]' dy =$$

$$= - \int_{y_2}^{y_1} g(y) dy = P(y_2 < Y \leq y_1). \quad (4.39)$$

Hieraus folgt offensichtlich

$$g(y) = -f[h^{-1}(y)] [h^{-1}(y)]',$$

was mit dem Resultat (4.36) übereinstimmt. Daher kann man zusammen schreiben

$$g(y) = f[h^{-1}(y)] \left| [h^{-1}(y)]' \right|. \quad (4.40)$$

Für die Berechnung von $g(y)$ genügt der letzte Ausdruck.

Beispiel 4.4.1. Wir betrachten wieder die Funktion $f(x) = 2x$ für $0 < x < 1$. Die Funktion $Y = 3X + 1$ war eine monoton wachsende Funktion. Jetzt betrachten wir die Transformation

$$Y = -3X + 1,$$

die eine monoton fallende Funktion darstellt, und suchen $g(y)$.

Lösung: Wir zeigen das Verfahren mit beiden Methoden:

$$\begin{aligned}
 \text{(I)} \quad G(y) &= P(Y \leq y) = P(-3X + 1 \leq y) = P(-3X \leq y - 1) = \\
 &= P\left(X \leq \frac{y-1}{-3}\right) = P\left(X \geq \frac{y-1}{-3}\right) = 1 - P\left(X < \frac{y-1}{-3}\right) = \\
 &= 1 - F\left(\frac{y-1}{-3}\right)
 \end{aligned}$$

woraus durch Differenzierung folgt:

$$g(y) = \left[1 - F\left(\frac{y-1}{-3}\right)\right]' = -f\left(\frac{y-1}{-3}\right) \left(-\frac{1}{3}\right) = \frac{1}{3} f\left(\frac{y-1}{-3}\right).$$

Setzen wir jetzt in $f(x) = 2x$ für $0 < x < 1$ dieses Resultat ein, so erhalten wir

$$g(y) = 2 \cdot \frac{1}{3} \left(\frac{y-1}{-3}\right) \quad \text{für } 0 < \frac{y-1}{-3} < 1$$

und schließlich

$$g(y) = -\frac{2}{9}(y-1) \quad \text{für } -2 < y < 1.$$

(II) Dasselbe erfolgt auch mit Formel (4.40), denn es ist

$$f(x) = 2x \quad 0 < x < 1$$

$$y = h(x) = -3x + 1$$

$$x = h^{-1}(y) = \frac{y-1}{-3}$$

$$\frac{dx}{dy} = \left(h^{-1}(y)\right)' = -\frac{1}{3}$$

$$\left|\frac{dx}{dy}\right| = \left|[h^{-1}(y)]'\right| = \frac{1}{3}.$$

Setzen wir dieses Resultat in (4.40) ein, so wird

$$g(y) = f[h^{-1}(y)] \left|[h^{-1}(y)]'\right| = 2 \left(\frac{y-1}{-3}\right) \frac{1}{3} = -\frac{2}{9}(y-1)$$

und das Intervall ist wieder $(-2, 1)$, was man leicht durch das Einsetzen von 0 und 1 für x in $y = -3x + 1$ berechnet. ■

Beispiel 4.4.2. Sei $f(x) = e^{-x}$ für $x > 0$. Wie lautet die Dichte von $Y = \sqrt{X}$?

Lösung: $y = \sqrt{x}$

$$x = y^2$$

$$\left|\frac{dx}{dy}\right| = 2y$$

$$g(y) = 2ye^{-y^2} \quad \text{für } y > 0. \quad \blacksquare$$

Wir führen nun einige der meistbenutzten Transformationen in allgemeiner Form auf.

(1) Die bereits behandelte lineare Transformation ist, allgemein ausgedrückt,

$$Y = aX + b. \quad (4.41)$$

Dann ist

$$y = h(x) = ax + b$$

$$x = h^{-1}(y) = \frac{y-b}{a}$$

$$\left|\frac{dx}{dy}\right| = \left|[h^{-1}(y)]'\right| = \frac{1}{|a|},$$

woraus nach Einsetzen in (4.40) folgt:

$$g(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right). \quad (4.42)$$

(2) Sei nun $Y = \frac{a}{X}$,

dann ist

$$y = h(x) = \frac{a}{x}$$

$$x = h^{-1}(y) = \frac{a}{y}$$

$$\left|\frac{dx}{dy}\right| = \left|[h^{-1}(y)]'\right| = \frac{|a|}{y^2}$$

woraus

$$g(y) = \frac{|a|}{y^2} f\left(\frac{a}{y}\right) \quad (4.43)$$

folgt.

Beispiel 4.4.3. Sei $f(x) = \frac{1}{2}(2-x)$ für $0 < x < 2$. Wir suchen die Dichte von $Y = \frac{2}{X}$.

Lösung: $y = \frac{2}{x}$

$$x = \frac{2}{y}$$

$$\left|\frac{dx}{dy}\right| = \frac{2}{y^2}$$

$$g(y) = \frac{1}{2} \left(2 - \frac{2}{y}\right) \cdot \frac{2}{y^2} = \frac{2}{y^3} (y-1) \quad \text{für } 1 < y < \infty. \quad \blacksquare$$

(3) Sei nun $Y = X^2$.

In diesem Fall ist $h(x)$ keine monotone Funktion, so daß (4.40) nicht gilt. Wenn $y < 0$, hat die Gleichung keine reelle Lösung, wenn $y > 0$, hat sie zwei Wurzeln. Dann haben wir

$$\begin{aligned} G(y) &= P(X^2 \leq y) = P(X \leq +\sqrt{y}) + P(X \leq -\sqrt{y}) = \\ &= F(+\sqrt{y}) + 1 - F(-\sqrt{y}). \end{aligned}$$

Daraus erhalten wir durch Differenzierung

$$g(y) = \begin{cases} \frac{f(\sqrt{y}) + f(-\sqrt{y})}{2\sqrt{y}} & \text{für } y > 0 \\ 0 & \text{sonst.} \end{cases} \quad (4.44)$$

In dem speziellen Fall, daß die Funktion symmetrisch ist, d.h. wenn $f(x) = f(-x)$, vereinfacht sich (4.44), und man bekommt

$$g(y) = \frac{f(\sqrt{y}) + f(-\sqrt{y})}{2\sqrt{y}} = \frac{f(\sqrt{y})}{\sqrt{y}} \quad \text{für } y > 0 \quad (4.45)$$

Beispiel 4.4.4. Sei $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ für $-\infty < x < \infty$ (Normalverteilung)

Gesucht ist die Dichte von $Y = X^2$.

Lösung: $y = x^2$
 $x = \pm \sqrt{y}$

$$\left| \frac{dx}{dy} \right| = \frac{1}{2\sqrt{y}}$$

$$\begin{aligned} g(y) &= \frac{f(\sqrt{y}) + f(-\sqrt{y})}{2\sqrt{y}} = \frac{\frac{1}{\sqrt{2\pi}} e^{-y/2} + \frac{1}{\sqrt{2\pi}} e^{-y/2}}{2\sqrt{y}} = \\ &= \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2} \quad \text{für } y > 0. \blacksquare \end{aligned}$$

Beispiel 4.4.5. Sei $f(x) = \frac{1}{9}(x^2 + 2)$ für $-1 < x < 2$.

Wir ermitteln die Dichte von $Y = X^2$.

Lösung: Die neue Verteilung muß in zwei Teilen definiert werden, weil die negativen Werte von X nur bis -1 laufen, die positiven hingegen bis 2 . Der Definitionsbereich der neuen Zufallsvariablen liegt zwischen 0 und 4 . Im Intervall $(0,1)$ werden der negative und der positive Teil summiert, im Intervall $(1,4)$ nur der positive, denn für $x > 1$ ist $f(-\sqrt{y}) \equiv 0$. Daher haben wir

$$\begin{aligned} g(y) &= \frac{f(\sqrt{y}) + f(-\sqrt{y})}{2\sqrt{y}} = \frac{\frac{1}{9}(y+2) + \frac{1}{9}(y+2)}{2\sqrt{y}} = \\ &= \frac{1}{9} y^{-1/2} (y+2) \quad \text{für } 0 < y < 1 \end{aligned}$$

und

$$g(y) = \frac{f(\sqrt{y}) + 0}{2\sqrt{y}} = \frac{\frac{1}{9}(y+2)}{2\sqrt{y}} = \frac{1}{18} y^{-1/2} (y+2) \quad \text{für } 1 \leq y < 4.$$

Es ist also

$$g(y) = \begin{cases} \frac{1}{9} y^{-1/2} (y+2) & \text{für } 0 < y < 1 \\ \frac{1}{18} y^{-1/2} (y+2) & \text{für } 1 \leq y < 4 \\ 0 & \text{sonst.} \blacksquare \end{cases}$$

Bemerkung: Nimmt die Zufallsvariable X nur positive Werte an, so hat die Gleichung $Y = X^2$ nur eine Lösung, und man kann (4.40) anwenden. Daher ist

$$g(y) = \frac{f(\sqrt{y})}{2\sqrt{y}}. \quad (4.46)$$

Beispiel 4.4.6. Die Zufallsvariable X hat die Dichte

$$f(x) = \frac{x}{\sigma^2} e^{-\left(\frac{x^2}{2\sigma^2}\right)} \quad \text{für } x > 0 \text{ (Rayleigh-Verteilung).}$$

Wir suchen die Dichte von $Y = X^2$.

Lösung: $y = x^2 \quad (x > 0)$
 $x = \sqrt{y}$

$$\left| \frac{dx}{dy} \right| = \frac{1}{2\sqrt{y}}$$

und nach (4.46) ist $g(y) = \frac{\sqrt{y}}{\sigma^2} e^{-\left(\frac{x^2}{2\sigma^2}\right)} \frac{1}{2\sqrt{y}} = \frac{1}{2\sigma^2} e^{-\left(\frac{y}{2\sigma^2}\right)}$ für $y > 0$. ■

(4) Sei $Y = |X|$.

Diese Gleichung hat auch nur für $y > 0$ zwei Lösungen, und es ist

$$G(y) = P(Y \leq y) = P(|X| \leq y) = P(-y \leq X \leq y) = F(y) - F(-y)$$

und daraus folgt

$$g(y) = f(y) + f(-y) \quad \text{für } y > 0 \quad (4.47)$$

Ist außerdem die Funktion symmetrisch, so gilt $F(-y) = 1 - F(y)$, woraus

$$G(y) = 2F(y) - 1$$

und

$$g(y) = 2f(y) \quad \text{für } y > 0 \quad (4.48)$$

folgt.

Beispiel 4.4.7. Sei $f(x) = \frac{1}{2}$ für $-1 < x < 1$. Wir ermitteln die Verteilungsfunktion und die Dichte von $Y = |X|$.

Lösung: Nach (4.47) bzw. (4.48) ist

$$g(y) = f(y) + f(-y) = \frac{1}{2} + \frac{1}{2} = 1$$

d.h.

$$g(y) = 1 \quad \text{für } 0 < y < 1.$$

Aus dem Resultat folgt die Verteilungsfunktion automatisch. Wir berechnen sie auch direkt, d.h.

$$F(x) = \int_{-1}^x \frac{1}{2} dx' = \frac{x+1}{2} \quad \text{für } -1 < x < 1.$$

Wegen Symmetrie bekommen wir

$$\begin{aligned} G(y) &= F(y) - F(-y) = F(y) - [1 - F(y)] = 2F(y) - 1 = \\ &= 2 \frac{y+1}{2} - 1 = y. \end{aligned}$$

Daher ist

$$G(y) = \begin{cases} 0 & -\infty < y \leq 0 \\ y & 0 < y < 1 \\ 1 & 1 \leq y < \infty. \end{cases} \blacksquare$$

(5) Ein sehr wichtiger Fall ist die sogenannte *Wahrscheinlichkeitsintegraltransformation*

$$y = F(x)$$

d.h. die Verteilungsfunktion von X selbst wird als eine neue Funktion betrachtet. Es gilt bekanntlich, daß $0 \leq F(x) \leq 1$. Durch Differenzierung beider Seiten von $y = F(x)$ erhält man

$$\frac{dy}{dx} = f(x)$$

woraus

$$dy = f(x) dx$$

folgt. Da bei den Transformationen gleichzeitig gilt (vgl. (4.38), (4.39))

$$g(y) dy = f(x) dx$$

so folgt daraus

$$g(y) = \begin{cases} 1 & \text{für } 0 \leq y \leq 1 \\ 0 & \text{sonst.} \end{cases} \quad (4.49)$$

Dasselbe kann man auch auf eine andere Weise zeigen. Wenn nämlich $y = F(x)$, dann ist

$$x = F^{-1}(y)$$

und da

$$\begin{aligned} G(y) &= P(Y \leq y) = P[F(x) \leq y] = P[X \leq F^{-1}(y)] = \\ &= F[F^{-1}(y)] = y \end{aligned}$$

ist, folgt

$$g(y) = 1 \quad \text{für } 0 \leq y \leq 1$$

wie oben. Diese Transformation zeigt, daß man jede Verteilung in eine Rechteckverteilung transformieren kann, was aber auch umgekehrt gilt, so daß man dadurch die Verteilungen ineinander überführen kann.

Beispiel 4.4.8. Sei $f(x) = 2x$ für $0 < x < 1$ und

$$g(y) = \sqrt{y} \quad \text{für } 0 < y < (3/2)^{2/3}.$$

Mit welcher Transformation kann man $f(x)$ in $g(y)$ überführen?

Lösung: Die Verteilungsfunktionen sind

$$F(x) = \int_0^x 2x' dx' = x^2$$

$$G(y) = \int_0^y \sqrt{y'} dy' = \frac{2}{3} y^{3/2}.$$

Da man beide Verteilungen durch ihre Verteilungsfunktionen in eine Rechteckverteilung überführen kann, setzen wir

$$F(x) = G(y)$$

d.h.

$$x^2 = \frac{2}{3} y^{3/2},$$

woraus

$$y = \left(\frac{3}{2}\right)^{2/3} x^{4/3}$$

folgt. Das heißt, daß man aus $f(x) = 2x$ für $0 < x < 1$ mit Hilfe der Transformation $y = (3/2)^{2/3} x^{4/3}$ die Dichte $g(y) = \sqrt{y}$ für $0 < y < (3/2)^{2/3}$ bilden kann.

Probe: Sei

$$y = \left(\frac{3}{2}\right)^{2/3} x^{4/3}$$

dann ist

$$x = \sqrt{\frac{2}{3}} y^{3/4}$$

$$\left| \frac{dx}{dy} \right| = \sqrt{\frac{2}{3}} \cdot \frac{3}{4} y^{-1/4}$$

$$g(y) = 2 \sqrt{\frac{2}{3}} y^{3/4} \sqrt{\frac{2}{3}} \cdot \frac{3}{4} y^{-1/4} = \sqrt{y}. \blacksquare$$

(6) Bei den diskreten Zufallsvariablen entfällt die Differenzierung, und man transformiert direkt in $g(y)$. Sei $Y = h(X)$, dann ist

$$g(y) = P(Y=y) = P[h(X)=y] = P[X=h^{-1}(y)] \quad (4.50)$$

Beispiel 4.4.9. Sei $f(x) = \frac{e^{-\lambda} \lambda^x}{x!}$ für $x = 0, 1, 2, \dots$ (Poisson-Verteilung).

Gesucht ist die Verteilung von $Y = 2X + 3$.

Lösung: $y = 2x + 3$

$$x = \frac{y-3}{2}$$

$$g(y) = \begin{cases} \frac{e^{-\lambda} \lambda^{(y-3)/2}}{\left(\frac{y-3}{2}\right)!} & \text{für } y = 3, 5, 7, 9, \dots \\ 0 & \text{sonst. } \blacksquare \end{cases}$$

Übung 4.4.1. Sei $f(x) = \frac{6}{5} x(x-1)$ für $1 < x < 2$. Finden Sie die Verteilung von $Y = 2X + 1$.

Übung 4.4.2. Sei $f(x) = 1$ für $0 < x < 1$ (Rechteckverteilung). Finden Sie die Verteilung von $Y = a/x$.

Übung 4.4.3. Sei $f(x) = \frac{1}{b-a}$ für $0 < a < x < b$ (Rechteckverteilung). Finden Sie die Verteilung von $Y = 2X^2$. [Benutzen Sie Formel (4.40).]

Übung 4.4.4. Sei $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ für $-\infty < x < \infty$ (Normalverteilung). Finden Sie die Verteilung von $Y = |X|$. [Benutzen Sie Formel (4.48).]

Übung 4.4.5. Sei $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ für $x = 0, 1, 2, \dots, n$ (Binomialverteilung). Finden Sie die Verteilung von $Y = X + 2$.

Aufgaben

4.1. Eine Wahrscheinlichkeitsfunktion sei definiert wie folgt

$$f(x) = \begin{cases} 1/10 & x=0 \\ 4/10 & x=1 \\ 2/10 & x=2 \\ 2/10 & x=3 \\ 1/10 & x=4 \\ 0 & \text{sonst} \end{cases}$$

(a) Stellen Sie die Verteilungsfunktion auf; berechnen Sie

- (b) $P(X > 0)$; (e) $1 - F(2)$;
(c) $P(1 < X \leq 3)$; (f) $F(3) - F(0)$;
(d) $P(X \leq 3)$; (g) $P(2 \leq X < 4)$.

4.2. Sei $f(x) = kx$ für $0 \leq x \leq \sqrt{2/3}$ und sonst $f(x) = 0$.

- (a) Finden Sie k ;
(b) berechnen Sie $F(x)$.

4.3. Eine Wahrscheinlichkeitsfunktion ist definiert wie folgt

$$f(x) = \begin{cases} k/42 & x=1 \\ k/21 & x=2 \\ k/14 & x=3 \\ k/42 & x=4 \\ 0 & \text{sonst} \end{cases}$$

- (a) bestimmen Sie k ;
(b) berechnen Sie $F(x)$.

4.4. Sei $f(x) = 4x$ für $0 \leq x \leq k$ und $f(x) = 0$ sonst.

- (a) Bestimmen Sie k ;

(b) berechnen Sie $F(x)$;

(c) berechnen Sie $P(X \leq 0,5)$.

4.5. Sei $f(x) = 2x - 1$ für $k < x < 3/2$. Bestimmen Sie k .

4.6. Sei $f(x) = 3/5 (\sqrt{x} - 1)$ für $1 < x < 4$ und $f(x) = 0$ sonst. Berechnen Sie

- (a) $F(x)$; (d) $f(x | X > 3)$;
(b) $F(3)$; (e) $P(2 < X \leq 3)$.
(c) $f(x | X \leq 2)$;

4.7. Sei $f(x) = 6/5 x(x-1)$ für $1 < x < 2$ und $f(x) = 0$ sonst.

(a) Berechnen Sie $F(x)$;

(b) Finden Sie die Dichte von $Y = \frac{5}{(X+1)}$;

(c) finden Sie die Dichte von $Y = X^2$.

4.8. Sei $f(x) = xe^{-x}$ für $x > 0$ und $f(x) = 0$ sonst. Finden Sie die Dichte von $Y = X^2$.

4.9. Sei $f(x) = 1$ für $0 < x < 1$ und $f(x) = 0$ sonst. Finden Sie die Dichte von

- (a) $Y = \sqrt{X}$; (b) $Y = e^X$.

4.10. Sei $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ für $x = 0, 1, 2, \dots, n$ und $f(x) = 0$ sonst. Finden Sie die (diskrete) Dichte von $Y = X^2$.

4.11. Sei $f(x) = 6x(1-x)$ für $0 < x < 1$ und $f(x) = 0$ sonst. Finden Sie die Dichte von $Y = \frac{2}{3X}$.

4.12. Sei $f(x) = \frac{1}{\pi(1+x^2)}$ für $-\infty < x < \infty$ (Cauchy-Verteilung).

Finden Sie die Dichte von

- (a) $Y = \frac{1}{X^2}$; (b) $Y = 3X^2$.

4.13. Man wirft zwei Würfel. Zeigen Sie, daß die Wahrscheinlichkeitsfunktion der Summe der Augenzahlen durch

$$f(x) = \begin{cases} \frac{x-1}{36} & \text{für } x = 2, 3, \dots, 6 \\ \frac{13-x}{36} & \text{für } x = 7, 8, \dots, 12 \\ 0 & \text{sonst} \end{cases}$$

gegeben ist.

4.14. Sei $f(x) = 2x$ für $0 < x < 1$ und $f(x) = 0$ sonst. Finden Sie die Dichte von

- (a) $Y = 1/2X$; (d) $Y = \sqrt{X}$;
(b) $Y = X^2/2$; (e) $Y = F(x)$.
(c) $Y = e^X$;

4.15. Sei $f(x) = 1/2(x+1)$ für $-1 < x < 1$ und $f(x) = 0$ sonst.

(a) Berechnen Sie $F(x)$;
Finden Sie die Dichte von

(b) $Y = X^2$; (c) $Y = 3X$.

4.16. Berechnen Sie aus der in Aufgabe 4.15 gegebenen Dichte $f(x | X > 0)$.

4.17. Finden Sie für die Dichte aus Aufgabe 4.14 eine Zahl a so, daß X mit gleicher Wahrscheinlichkeit größere und kleinere Werte als a annimmt.

4.18. Finden Sie für die Dichte aus Aufgabe 4.15 eine Zahl a so, daß $P(X \geq a) = 0,75$.

4.19. Sei $f(x) = \frac{\sqrt{2}}{\sqrt{\pi}} x^2 e^{-x^2/2}$ für $0 < x < \infty$ und $f(x) = 0$ sonst.

Finden Sie die Dichte von $Y = X^2$.

4.20. Finden Sie für die Dichte $f(x) = 2$ für $0 < x < 1/2$ und $f(x) = 0$ sonst

(a) eine Zahl a so, daß die Wahrscheinlichkeit, daß X größer als a ist, gleich 0,2 ist;

(b) a so, daß die Wahrscheinlichkeit, daß X kleiner als a ist, gleich 0,05 ist.

4.21. Sei $f(x) = 2$ für $0 < x < 1/2$ und $f(x) = 0$ sonst. Finden Sie die Dichte von

(a) $Y = X^2$; (d) $Y = e^X$;

(b) $Y = \sqrt{X}$; (e) $Y = 2X + 5$;

(c) $Y = \ln X$; (f) $Y = 3/2 X$.

4.22. Sei $f(x) = 2/x$ für $1 < x < \sqrt{e}$. Finden Sie die Dichte von

(a) $Y = 1/X$;

(b) $Y = X^2$;

(c) $Y = aX + b$.

Benutzte und weiterführende Literatur: Brunk (1965), Fisz (1970), Kreyszig (1968), Mathai (1967), Mood-Graybill (1963), Papoulis (1963), Wadsworth-Bryan (1960), Ventcel' (1964), Ventcel'-Ovčarov, (1973).

5. Mathematische Erwartung und Momente

5.1. Die mathematische Erwartung

Unter der mathematischen Erwartung oder einfach Erwartung der Zufallsvariablen X mit der Wahrscheinlichkeitsfunktion $f(x)$ versteht man den Ausdruck

$$E(X) = \begin{cases} \sum_x x f(x) & \text{für diskrete Zufallsvariablen} \\ \int_{-\infty}^{\infty} x f(x) dx & \text{für stetige Zufallsvariablen} \end{cases} \quad (5.1)$$

wobei die Summe bzw. das Integral endlich sein müssen. Im empirischen Bereich entspricht der erste Ausdruck eigentlich dem Mittelwert. Schätzt man nämlich aus der Stichprobe $\hat{f}(x)$ durch f_i/N , so kann man schreiben

$$\sum_x x \hat{f}(x) = \frac{1}{N} \sum_i x_i f_i = \bar{x},$$

was bereits aus (1.7) bekannt ist. Der Erwartungswert ist eigentlich der Mittelwert einer theoretischen Verteilung.

Beispiel 5.1.1. Sei X die Punktzahl beim Werfen eines Würfels, dann ist die Dichte von X

$$f(x) = \begin{cases} \frac{1}{6} & \text{für } x = 1, 2, 3, 4, 5, 6 \\ 0 & \text{sonst.} \end{cases}$$

Wie ist die Erwartung von X ?

Lösung: Nach (5.1) ist

$$E(X) = \sum_x x f(x) = \sum_{x=1}^6 x \frac{1}{6} = \frac{1}{6} \sum_{x=1}^6 x = \frac{1}{6} (21) = 3,5 . \blacksquare$$

Beispiel 5.1.2. Sei

$$f(x) = \begin{cases} x & \text{für } 0 < x \leq 1 \\ 2-x & \text{für } 1 < x < 2 \\ 0 & \text{sonst.} \end{cases}$$

Wir berechnen $E(X)$.

Lösung: Nach (5.1) ist

$$E(x) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x \cdot x dx + \int_1^2 x (2-x) dx = \left[\frac{x^3}{3} \right]_0^1 + \left[x^2 - \frac{x^3}{3} \right]_1^2 = 1 . \blacksquare$$

Bei der Berechnung der Erwartung einer Funktion von X ist es nicht nötig, X erst zu transformieren und dann die Erwartung der neuen Zufallsvariablen zu berechnen, sondern man kann sie direkt mit Hilfe von $f(x)$ ermitteln. Wenn $y = \varphi(x)$ eine Funktion der Zufallsvariablen X ist, dann ist

$$E[\varphi(x)] = \begin{cases} \sum_x \varphi(x) f(x) & \text{für diskrete Zufallsvariablen} \\ \int_{-\infty}^{\infty} \varphi(x) f(x) dx & \text{für stetige Zufallsvariablen} \end{cases} \quad (5.2)$$

wiederum unter der Bedingung, daß die Summe bzw. das Integral existieren. Diese Behauptung ist leicht zu beweisen, denn laut (4.40) ist

$$\int y g(y) dy = \int y f[h(y)] \left| [h^{-1}(y)]' \right| dy = \int y f(x) dx = \int \varphi(x) f(x) dx$$

und analog für den diskreten Fall.

Beispiel 5.1.3. Sei $f(x) = 2$ für $0 < x < 1/2$ und $f(x) = 0$ sonst und sei $Y = e^X$. Wir suchen $E(Y)$.

Lösung: Der direkte Weg führt durch (5.2) zu

$$E(Y) = E(e^X) = \int_0^{1/2} 2 e^x dx = 2 \left[e^x \right]_0^{1/2} = 2\sqrt{e} - 2 = 2(\sqrt{e} - 1) .$$

Der indirekte Weg führt über die Transformation der Zufallsvariablen X . Wenn $y = e^x$, dann ist

$$\begin{aligned} x &= \ln y \\ \left| \frac{dx}{dy} \right| &= \frac{1}{y} \\ g(y) &= \frac{2}{y} \quad \text{für } 1 < y < \sqrt{e} . \end{aligned}$$

Daraus wird dann

$$E(Y) = \int_1^{\sqrt{e}} y \cdot \frac{2}{y} dy = [2y]_1^{\sqrt{e}} = 2\sqrt{e} - 2 = 2(\sqrt{e} - 1)$$

was mit dem Resultat oben übereinstimmt. $\blacksquare$

Für die mathematische Erwartung gelten einige wichtige Sätze, deren empirische Gegenstücke aus dem Kapitel 1 bereits bekannt sind.

$$(1) \quad E(a) = a \quad (5.3)$$

d.h. die Erwartung einer Konstante ist die Konstante selbst.

$$\text{Beweis: } E(a) = \int_{-\infty}^{\infty} a f(x) dx = a \int_{-\infty}^{\infty} f(x) dx = a \cdot 1 = a .$$

Für diskrete Variablen ist der Beweis als Aufgabe gestellt. $\blacksquare$

$$(2) \quad E[a\varphi(x)] = aE[\varphi(x)] . \quad (5.4)$$

$$\text{Beweis: } E[a\varphi(x)] = \sum_x a\varphi(x)f(x) = a \sum_x \varphi(x)f(x) = aE[\varphi(x)] .$$

Für stetige Variablen ist der Beweis als Aufgabe gestellt. $\blacksquare$

$$\text{Folgerung 1. } E(aX) = aE(X) . \quad (5.5)$$

Der Beweis folgt automatisch aus (5.4), wenn man $\varphi(x) = X$ setzt. $\blacksquare$

Folgerung 2. $E(aX + b) = aE(X) + b$. (5.6)

Beweis: folgt aus (5.3) und (5.5), nämlich

$$\begin{aligned} E(aX + b) &= \int_{-\infty}^{\infty} (ax + b)f(x) dx = a \int_{-\infty}^{\infty} xf(x) dx + b \int_{-\infty}^{\infty} f(x) dx = \\ &= aE(X) + b. \blacksquare \end{aligned}$$

Folgerung 3. $E[X - E(X)] = 0$ (5.7)

Beweis: Da $E(X)$ eine Konstante ist, folgt $E[X - E(X)] = E(X) - E[E(X)] = E(X) - E(X) = 0$. ■

Beispiel 5.1.4. Sei $f(x) = \frac{2}{\sqrt{e}}$ für $\frac{1}{\sqrt{e}} < x < 1$ und $f(x) = 0$ sonst. Sei nun $Y = 3X - 1$. Gesucht ist $E(Y)$.

$$\begin{aligned} \text{Lösung: } E(Y) &= E(3X - 1) = \int_{\frac{1}{\sqrt{e}}}^1 (3x - 1) \frac{2}{x} dx = \int_{\frac{1}{\sqrt{e}}}^1 \left(6 - \frac{2}{x}\right) dx = \\ &= [6x - 2 \ln x]_{\frac{1}{\sqrt{e}}}^1 = 5 - \frac{6}{\sqrt{e}} \end{aligned}$$

Indirekt, wenn $y = 3x - 1$ ist, so ist

$$\begin{aligned} x &= \frac{y+1}{3} \\ \left| \frac{dx}{dy} \right| &= \frac{1}{3} \\ g(y) &= \frac{2}{y+1} \quad \text{für } \frac{3}{\sqrt{e}} - 1 < y < 2 \quad \text{und } g(y) = 0 \text{ sonst.} \end{aligned}$$

Dann ist

$$E(y) = \int_{\frac{3}{\sqrt{e}} - 1}^2 \frac{2y}{y+1} dy$$

Setzt man für die partielle Integration

$$\begin{aligned} u &= 2y \\ v &= \frac{1}{y+1}, \end{aligned}$$

so wird

$$\begin{aligned} \int_{\frac{3}{\sqrt{e}} - 1}^2 \frac{2y}{y+1} dy &= \left[2y \ln(y+1) \right]_{\frac{3}{\sqrt{e}} - 1}^2 - \int_{\frac{3}{\sqrt{e}} - 1}^2 2 \ln(y+1) dy \\ &= 2 \left[y + 1 - \ln(y+1) \right]_{\frac{3}{\sqrt{e}} - 1}^2 = 5 - \frac{6}{\sqrt{e}} \end{aligned}$$

wie oben (vgl. B2, Nr.17). ■

Übung 5.1.1. Sei $f(x) = \frac{6}{5}x(x-1)$ für $1 < x < 2$ und $f(x) = 0$ sonst. Finden Sie $E(X)$.

Übung 5.1.2. Sei $f(x) = 4x^3$ für $0 < x < 1$ und $f(x) = 0$ sonst. Finden Sie $E(Y)$, wenn $Y = X + 1$.

Übung 5.1.3. Sei $f(x) = \frac{1}{n}$ für $x = 1, 2, \dots, n$ (diskrete Rechteckverteilung). Finden Sie $E(X)$.

Übung 5.1.4. Sei $f(x) = \binom{n}{x} p^x q^{n-x}$ für $x = 0, 1, \dots, n$ (Binomialverteilung). Finden Sie $E(e^X)$.

Übung 5.1.5. Sei $f(x) = \frac{1}{2x^3}$ für $\frac{1}{2} < x < \infty$. Finden Sie $E(X)$.

5.2. Momente

Im weiteren werden wir uns mit Erwartungen einer speziellen Klasse der Funktionen von X beschäftigen, nämlich mit

$$\varphi(X) = (X - c)^r \quad (5.8)$$

wobei c eine beliebige Konstante ist und $r = 0, 1, 2, \dots$. Die Erwartung einer solchen Funktion, d.h.

$$E[(X - c)^r] = \begin{cases} \sum_x (x - c)^r f(x) & \text{für diskrete Zufallsvariablen} \\ \int_{-\infty}^{\infty} (x - c)^r f(x) dx & \text{für stetige Zufallsvariablen} \end{cases} \quad (5.9)$$

nennen wir das *r-te allgemeine* Moment von X bezüglich c oder das Moment *r-ter* Ordnung bezüglich c . Die Existenz der Summe bzw. des Integrals wird vorausgesetzt.

Für die Statistik sind solche Fälle von Bedeutung, in denen c einen konkreten, speziellen Wert annimmt.

(I) Wenn $c = 0$, dann nennen wir diese Erwartung das *r-te Anfangsmoment* von X und bezeichnen sie als μ'_r , d.h.

$$E(X^r) = \mu'_r = \begin{cases} \sum x^r f(x) & \text{für diskrete Zufallsvariablen} \\ \int_{-\infty}^{\infty} x^r f(x) dx & \text{für stetige Zufallsvariablen.} \end{cases} \quad (5.10)$$

Wenn $r = 1$, so ist $\mu'_1 = E(X)$, d.h. die mathematische Erwartung von X . Man findet in der Literatur oft einfach μ statt μ'_1 .

Beispiel 5.2.1. Sei $f(x) = 1/x$ für $1 < x < e$ und $f(x) = 0$ sonst. Wir suchen das erste und das zweite Anfangsmoment, d.h. μ'_1 und μ'_2 .

Lösung: $E(X) = \mu'_1 = \int_1^e x \frac{1}{x} dx = \int_1^e dx = [x]_1^e = e - 1.$

$$E(X^2) = \mu'_2 = \int_1^e x^2 \frac{1}{x} dx = \int_1^e x dx = \left[\frac{x^2}{2} \right]_1^e = \frac{e^2}{2} - \frac{1}{2} = \frac{1}{2} (e^2 - 1). \blacksquare$$

(II) Wenn $c = \mu'_1 = E(X)$, dann ergibt (5.9) das *r-te Zentralmoment* von X , bezeichnet als μ_r , d.h.

$$E[(X - \mu'_1)^r] = \mu_r = \begin{cases} \sum (x - \mu'_1)^r f(x) & \text{für diskrete Zufallsvariablen} \\ \int_{-\infty}^{\infty} (x - \mu'_1)^r f(x) dx & \text{für stetige Zufallsvariablen.} \end{cases} \quad (5.11)$$

Es ist leicht zu sehen, daß $\mu_0 = 1$, weil $[h(x)]^0 = 1$ und $\sum f(x)$ bzw. $\int f(x) dx = 1$. Weiter ist $\mu_1 = 0$, was bereits in (5.7) bewiesen wurde. Das wichtigste Zentralmoment ist μ_2 , weil

$$\mu_2 = E[(X - \mu'_1)^2] = V(X) = \sigma^2 = \begin{cases} \sum (x - \mu'_1)^2 f(x) \\ \int_{-\infty}^{\infty} (x - \mu'_1)^2 f(x) dx \end{cases} \quad (5.12)$$

die Varianz von X ist. In der empirischen Form haben wir sie unter der Bezeichnung S^2 kennengelernt. Von dort kennen wir auch schon einige Eigenschaften der Varianz, die wir in dieser neuen Terminologie beweisen werden.

$$(1) \quad V(X) = E(X^2) - [E(X)]^2 = \mu'_2 - \mu'^2_1. \quad (5.13)$$

Beweis: $V(X) = E[(X - \mu'_1)^2] = E[(X - E(X))^2] =$

$$= E(X^2 - 2X E(X) + [E(X)]^2)$$

$$= E(X^2) - 2E(X) E(X) + [E(X)]^2$$

$$= E(X^2) - [E(X)]^2 = \mu'_2 - \mu'^2_1. \blacksquare$$

$$(2) \quad V(X + b) = V(X). \quad (5.14)$$

Beweis: $V(X + b) = E[(X + b - E(X + b))^2]$

$$= E[(X + b - E(X) - b)^2]$$

$$= E[(X - E(X))^2] = V(X). \blacksquare$$

$$(3) \quad V(aX + b) = a^2 V(X). \quad (5.15)$$

$$\begin{aligned} \text{Beweis: } V(aX + b) &= E[(aX + b - E(aX + b))^2] \\ &= E[(aX + b - aE(X) - b)^2] \\ &= E[(aX - aE(X))^2] \\ &= a^2 E[(X - E(X))^2] \\ &= a^2 V(X) \quad \blacksquare \end{aligned}$$

Der Ausdruck $E[(X - c)^2]$ ist die mittlere quadratische Abweichung der Zufallsvariablen X von einer Konstanten c . Wir zeigen, daß dieser Ausdruck am kleinsten ist, wenn $c = \mu'_1$, oder daß die Varianz die kleinste mittlere quadratische Abweichung darstellt, d.h.

$$(4) \quad V(X) \leq E[(X - c)^2]. \quad (5.16)$$

$$\begin{aligned} \text{Beweis: } E[(X - c)^2] &= E[(X - \mu'_1 + \mu'_1 - c)^2] \\ &= E[(X - \mu'_1) + (\mu'_1 - c)]^2 \\ &= E[(X - \mu'_1)^2 + 2(\mu'_1 - c)(X - \mu'_1) + (\mu'_1 - c)^2] \\ &= E[(X - \mu'_1)^2] + 2(\mu'_1 - c) E(X - \mu'_1) + (\mu'_1 - c)^2 \\ &= V(X) + (\mu'_1 - c)^2 \end{aligned}$$

weil $E(X - \mu'_1) = 0$, wie bereits bewiesen wurde, so daß das mittlere Glied entfällt. Da

$$E[(X - c)^2] = V(X) + (\mu'_1 - c)^2,$$

folgt daraus, daß $E(X - c)^2 \geq V(X)$. ■

Zwischen den Anfangs- und den Zentralmomenten besteht eine nützliche Beziehung, die man bei der Berechnung der Zentralmomente häufig anwendet. Wir gehen von der Definition (5.11) aus und entwickeln das Binom für $r = 0, 1, 2, 3, 4$:

$$\mu_0 = E[(X - \mu'_1)^0] = E(1) = 1$$

$$\mu_1 = E[(X - \mu'_1)^1] = 0$$

$$\begin{aligned} \mu_2 &= E[(X - \mu'_1)^2] = E(X^2 - 2\mu'_1 X + \mu'^2_1) = \\ &= E(X^2) - 2\mu'_1 E(X) + \mu'^2_1 = \mu'_2 - \mu'^2_1 \end{aligned} \quad (5.17)$$

$$\begin{aligned} \mu_3 &= E[(X - \mu'_1)^3] = E(X^3 - 3\mu'_1 X^2 + 3\mu'^2_1 X - \mu'^3_1) = \\ &= E(X^3) - 3\mu'_1 E(X^2) + 3\mu'^2_1 E(X) - \mu'^3_1 \\ &= \mu'_3 - 3\mu'_1 \mu'_2 + 3\mu'^2_1 \mu'_1 - \mu'^3_1 \\ &= \mu'_3 - 3\mu'_2 \mu'_1 + 2\mu'^3_1 \end{aligned} \quad (5.18)$$

$$\begin{aligned} \mu_4 &= E[(X - \mu'_1)^4] = E(X^4 - 4\mu'_1 X^3 + 6\mu'^2_1 X^2 - 4\mu'^3_1 X + \mu'^4_1) = \\ &= \mu'_4 - 4\mu'_3 \mu'_1 + 6\mu'_2 \mu'^2_1 - 3\mu'^4_1 \end{aligned} \quad (5.19)$$

Allgemein kann man für ein beliebiges $r = 0, 1, 2, 3, \dots$ schreiben

$$\mu_r = E[(X - \mu'_1)^r]$$

$$= E \left[\binom{r}{0} \mu'_1{}^0 X^{r-0} - \binom{r}{1} \mu'_1{}^1 X^{r-1} + \binom{r}{2} \mu'_1{}^2 X^{r-2} - \dots \pm \binom{r}{k} \mu'_1{}^k X^{r-k} + \dots \pm \binom{r}{r} \mu'_1{}^r X^0 \right] =$$

$$= E \left[\sum_{k=0}^r \binom{r}{k} (-1)^k \mu'_1{}^k X^{r-k} \right]$$

$$= \sum_{k=0}^r \binom{r}{k} (-1)^k \mu'_1{}^k E(X^{r-k}) = \sum_{k=0}^r \binom{r}{k} (-1)^k \mu'_1{}^k \mu'_{r-k} \quad (5.20)$$

Beispiel 5.2.2. Sei $f(x) = 4x^3$ für $0 < x < 1$ und $f(x) = 0$ sonst. Gesucht sind die ersten vier Zentralmomente.

Lösung: Wir verkürzen uns die Arbeit, indem wir die Anfangsmomente berechnen und aus ihnen mit Hilfe von (5.17) bis (5.19) die Zentralmomente finden. Es ist allgemein

$$\mu'_r = \int_0^1 x^r 4 x^3 dx = 4 \int_0^1 x^{r+3} dx = \frac{4}{r+4}$$

woraus folgt:

$$\mu'_1 = \frac{4}{5}; \quad \mu'_2 = \frac{4}{6}; \quad \mu'_3 = \frac{4}{7}; \quad \mu'_4 = \frac{4}{8}.$$

Man erhält nun

$$\mu_2 = \mu'_2 - \mu'_1{}^2 = \frac{4}{6} - \left(\frac{4}{5}\right)^2 = \frac{2}{75};$$

$$\mu_3 = \mu'_3 - 3 \mu'_2 \mu'_1 + 2 \mu'_1{}^3 = \frac{4}{7} - 3 \cdot \frac{4}{6} \cdot \frac{4}{5} + 2 \left(\frac{4}{5}\right)^3 = -\frac{12}{2625};$$

$$\begin{aligned} \mu_4 &= \mu'_4 - 4 \mu'_3 \mu'_1 + 6 \mu'_2 \mu'_1{}^2 - 3 \mu'_1{}^4 \\ &= \frac{4}{8} - 4 \cdot \frac{4}{7} \cdot \frac{4}{5} + 6 \cdot \frac{4}{6} \cdot \left(\frac{4}{5}\right)^2 - 3 \left(\frac{4}{5}\right)^4 = \frac{23}{8750}. \blacksquare \end{aligned}$$

Beispiel 5.2.3. Sei $f(x) = 1/6$ für $x = 1, 2, \dots, 6$ und $f(x) = 0$ sonst (Verteilung der Punktzahl beim Werfen eines Würfels). Gesucht sind μ'_1 , μ'_2 und μ_2 .

$$\text{Lösung: } \mu'_1 = \sum_{x=1}^6 x \cdot \frac{1}{6} = \frac{1}{6} \sum_{x=1}^6 x = \frac{1}{6} \cdot \frac{6(7)}{2} = \frac{7}{2} \quad (\text{vgl. B1, Nr.2})$$

$$\mu'_2 = \sum_{x=1}^6 x^2 \cdot \frac{1}{6} = \frac{1}{6} \sum_{x=1}^6 x^2 = \frac{1}{6} \cdot \frac{6(7)(12+1)}{6} = \frac{91}{6} \quad (\text{vgl. B1, Nr.3})$$

$$\mu_2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}. \blacksquare$$

Nicht jede Verteilung muß unbedingt Momente haben. Ist der Ausdruck (5.9) nicht endlich, so sagen wir, daß X keine Momente hat.

Beispiel 5.2.4. Sei $f(x) = \frac{3}{x^2}$ für $3 < x < \infty$ und $f(x) = 0$ sonst. Wir berechnen μ'_1 und μ'_2 .

$$\text{Lösung: } \mu'_1 = \int_3^{\infty} x \cdot \frac{3}{x^2} dx = \int_3^{\infty} \frac{3}{x} dx = 3 [\ln x]_3^{\infty} = \infty$$

$$\mu'_2 = \int_3^{\infty} x^2 \cdot \frac{3}{x^2} dx = 3 \int_3^{\infty} dx = 3 [x]_3^{\infty} = \infty$$

und genauso alle höheren Momente. Daher hat diese Verteilung keine Momente. $\blacksquare$

Es gibt auch Verteilungen, die nur die niedrigeren Momente haben, während die höheren nicht vorhanden sind.

Beispiel 5.2.5. Sei $f(x) = \frac{1}{2x^3}$ für $\frac{1}{2} < x < \infty$ und $f(x) = 0$ sonst. Wir suchen μ'_1 und μ'_2 .

Lösung: In Übung 5.1.5 haben wir festgestellt, daß

$$\mu'_1 = \int_{\frac{1}{2}}^{\infty} x \cdot \frac{1}{2x^3} dx = 1$$

ist. Für μ'_2 ergibt sich jedoch

$$\mu'_2 = \int_{\frac{1}{2}}^{\infty} x^2 \cdot \frac{1}{2x^3} dx = \frac{1}{2} \int_{\frac{1}{2}}^{\infty} \frac{1}{x} dx = \frac{1}{2} [\ln x]_{\frac{1}{2}}^{\infty} = \infty.$$

Das gleiche Ergebnis erhalten wir auch für die weiteren Anfangsmomente. Es existiert somit das erste Anfangsmoment, nicht aber die höheren. ■

Die empirischen Gegenstücke zu den theoretischen Momenten bezeichnen wir mit m'_r für Anfangsmomente und m_r für Zentralfmomente. Die Definitionen sind analog, nämlich

$$m'_r = \frac{1}{N} \sum_i x_i^r f_i \quad (5.21)$$

und

$$m_r = \frac{1}{N} \sum_i (x_i - \bar{x})^r f_i \quad (5.22)$$

und die Beziehungen zwischen ihnen sind dieselben wie bei den theoretischen Momenten.

Übung 5.2.1. Sei $f(x) = 2x - 1$ für $\frac{1}{2} < x < \frac{3}{2}$ und $f(x) = 0$ sonst. Berechnen Sie μ'_1 und μ_2 .

Übung 5.2.2. Sei $f(x) = 2$ für $0 < x < \frac{1}{2}$ und $f(x) = 0$ sonst. Berechnen Sie μ'_r .

Übung 5.2.3. Beweisen Sie, daß

$$E\left(\frac{X - \mu'_1}{\sigma}\right) = 0 \text{ und } V\left(\frac{X - \mu'_1}{\sigma}\right) = 1$$

[Benutzen Sie 5.6 und 5.15]

Übung 5.2.4. In der Aufgabe 4.13 wurde die Verteilung der Punktesumme beim Werfen zweier Würfel angegeben als

$$f(x) = \frac{x-1}{36} \text{ für } x = 2, 3, \dots, 6; f(x) = \frac{13-x}{36} \text{ für } x = 7, 8, \dots, 12 \text{ und } f(x) = 0 \text{ sonst.}$$

Finden Sie μ'_1 , μ'_2 und μ_2 .

5.3. Die Tschebyschevsche Ungleichung

Eine der meistbenutzten Ungleichungen der Wahrscheinlichkeitstheorie ist die Tschebyschevsche Ungleichung. Sie behauptet, daß Abweichungen der Zufallsvariablen von ihrem Erwartungswert desto unwahrscheinlicher sind, je größer sie sind, und sie erlaubt die Abschätzung der Wahrscheinlichkeit für Abweichungen einer gegebenen Größe. Ist σ die Standardabweichung der Zufallsvariablen X , so ist die Wahrscheinlichkeit, daß sich X von dem Erwartungswert $E(X)$ um mindestens $k\sigma$ unterscheidet, oder in anderen Worten, daß X außerhalb des Intervalls $(E(X) - k\sigma, E(X) + k\sigma)$ liegt, höchstens $\frac{1}{k^2}$, d.h.

$$P[|X - E(X)| > k\sigma] \leq \frac{1}{k^2}, \text{ wobei } k > 0. \quad (5.23)$$

Beweis: Für eine stetige Zufallsvariable kann man schreiben

$$P[|X - E(X)| \geq k\sigma] = \int_{|X - E(X)| \geq k\sigma} f(x) dx. \quad (5.24)$$

Im Integrationsbereich gilt natürlich, daß

$$|X - E(X)| \geq k\sigma.$$

Dividiert man beide Seiten dieser Ungleichung durch $k\sigma$, so bekommt man

$$\frac{|X - E(X)|}{k\sigma} \geq 1,$$

woraus folgt

$$\frac{[X - E(X)]^2}{k^2 \sigma^2} \geq 1$$

und daraus automatisch auch

$$\int_{|X - E(X)| > k\sigma} \frac{[X - E(X)]^2}{k^2 \sigma^2} f(x) dx \geq \int_{|X - E(X)| > k\sigma} 1 \cdot f(x) dx$$

Die Ungleichheit beider Seiten vergrößert sich noch mehr, wenn man auf der linken Seite von $-\infty$ bis ∞ integriert, d.h.

$$\frac{1}{k^2 \sigma^2} \int_{-\infty}^{\infty} [X - E(X)]^2 f(x) dx \geq \int_{|X - E(X)| > k\sigma} f(x) dx.$$

Da das Integral auf der linken Seite jetzt genau die Definition der Varianz ist (vgl. 5.12) und die rechte Seite Formel (5.24) entspricht, so kann man schreiben

$$\frac{1}{k^2} \geq P(|X - E(X)| \geq k\sigma),$$

was die Behauptung des Satzes war. ■

Laut den Axiomen der Wahrscheinlichkeit folgt aus (5.23) sofort, daß

$$P(|X - E(X)| < k\sigma) \geq 1 - \frac{1}{k^2} \quad (5.25)$$

Da k eine beliebige Konstante ist, kann man für sie auch k/σ einsetzen. In diesem Fall wird aus (5.23) bzw. (5.25)

$$P(|X - E(X)| \geq k) \leq \frac{\sigma^2}{k^2} \quad (5.26)$$

bzw.

$$P(|X - E(X)| < k) \geq 1 - \frac{\sigma^2}{k^2}$$

Beispiel 5.3.1. Sei $P(|X - \mu'_1| \geq 2k) \leq 0,05$. Wir suchen k .

Lösung: Hier ist $\sigma = 2$. Nach (5.23) ist $0,05 = 1/k^2$, woraus

$$k^2 = \frac{1}{0,05} = 20$$

$$k = 4,47$$

d.h., die Wahrscheinlichkeit, daß X von seinem Erwartungswert um mindestens $4,47\sigma$ abweicht, ist höchstens 0,05. Hier ist 0,05 eine obere Grenze der Wahrscheinlichkeit, nicht aber die kleinste obere Grenze. ■

Beispiel 5.3.2. Sei $P(|X - \mu'_1| < k) > 0,90$. Wie lautet k , wenn $\sigma = 3$?

Lösung: Nach (5.27) ist

$$0,90 = 1 - \frac{\sigma^2}{k^2} = 1 - \frac{9}{k^2}$$

woraus

$$k = \sqrt{90} = 9,49$$

d.h., die Wahrscheinlichkeit, daß X von μ'_1 um weniger als 9,49 bei der gegebenen Standardabweichung abweicht, ist sicherlich größer als 0,90. Hier ist 0,90 eine untere Grenze der Wahrscheinlichkeit. ■

Übung 5.3.1. Sei $\sigma = 1$. Finden Sie k , wenn $P(|X - \mu'_1| < k) > 0,975$.

Übung 5.3.2. Sei $\sigma = 5$ und $P(|X - \mu'_1| \geq k) \leq 0,10$. Finden Sie k .

5.4. Schiefe und Exzeß

Aus den Momenten lassen sich zwei weitere Charakteristika einer Verteilung bilden, nämlich das Maß der *Schiefte* oder der *Asymmetrie* und das Maß des *Exzesses* oder der *Wölbung*. Nicht jede Verteilung muß unbedingt symmetrisch

sein, wovon man sich leicht überzeugt, wenn man sich einige der bereits erwähnten Verteilungen aufzeichnet.

Das Maß der Asymmetrie oder der *Schiefe* wird am häufigsten als

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}} \quad (5.28)$$

definiert. Das dritte Zentralmoment kann positiv oder negativ sein, je nachdem, auf welcher Seite die größeren Abweichungen vom Mittelwert vorhanden sind.

(a) Wenn $\gamma_1 > 0$, dann ist die Schiefe positiv, d.h., die Verteilung ist linkssteil wie in Abb. 5.1.

(b) Wenn $\gamma_1 < 0$, dann ist die Schiefe negativ, d.h., die Verteilung ist rechtssteil wie in Abb. 5.2.

(c) Wenn $\gamma_1 = 0$, dann ist die Verteilung symmetrisch um den Mittelwert.

Die "Steilheit" bzw. "Flachheit" einer Verteilung bezeichnet man als *Exzeß* und definiert ihn als

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mu_4}{\mu_2^2} - 3 \quad (5.29)$$

Man bewertet den Exzeß einer Verteilung im Vergleich zur Normalverteilung, wo er 3 beträgt, daher wird in (5.29) der Betrag von 3 abgezogen.

(a) Wenn $\gamma_2 > 0$, dann ist die Verteilung steil und schmal.

(b) Wenn $\gamma_2 < 0$, dann ist die Verteilung flach und breit.

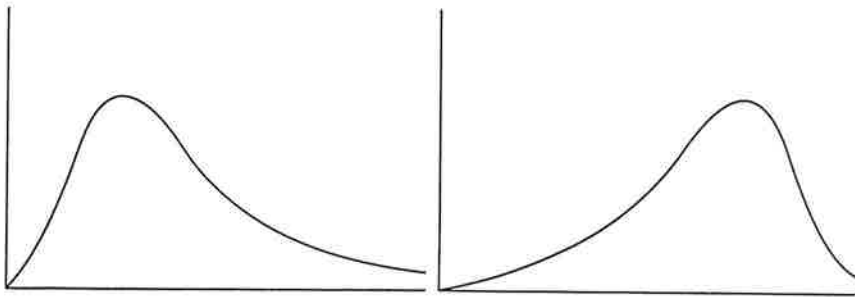


Abb. 5.1. Linkssteile Verteilung

Abb. 5.2. Rechtssteile Verteilung

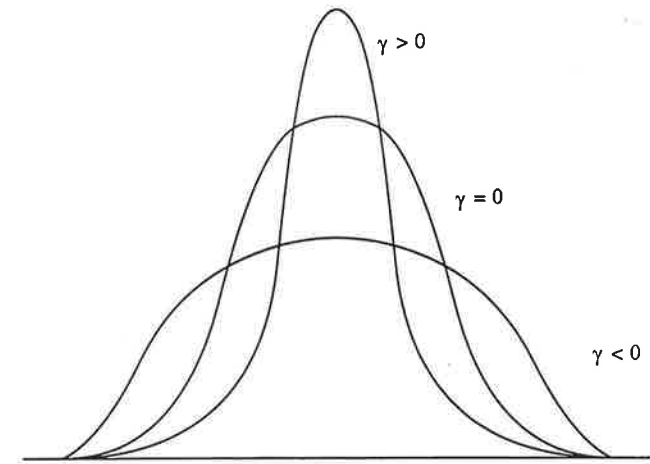


Abb. 5.3. Exzeß

(c) Wenn $\gamma_2 = 0$, dann sind die Steilheit und die Gipfelbreite der Verteilung normal, d.h. gleich denen einer Normalverteilung. Alle drei Typen sind in Abb. 5.3 gezeichnet.

Wenn eine Verteilung keine höheren Momente hat, dann ist die Berechnung der Schiefe und des Exzesses auf die angegebene Weise nicht möglich. Bei den empirischen Verteilungen ist sie jedoch immer möglich.

Beispiel 5.4.1. Wir berechnen die Schiefe und den Exzeß der Verteilung mit der Dichte $f(x) = 4x$ für $0 < x < 1/\sqrt{2}$ und $f(x) = 0$ sonst.

Lösung: Zur Berechnung brauchen wir die ersten vier Anfangsmomente. Es ist

$$\begin{aligned} \mu'_r &= \int_0^{1/\sqrt{2}} x^r 4x \, dx = 4 \int_0^{1/\sqrt{2}} x^{r+1} \, dx = \left[\frac{4x^{r+2}}{r+2} \right]_0^{1/\sqrt{2}} = \\ &= \frac{1}{2^{r/2-1}} (r+2), \end{aligned}$$

woraus

$$\mu'_1 = \frac{\sqrt{2}}{3}; \quad \mu'_2 = \frac{1}{4}; \quad \mu'_3 = \frac{1}{5\sqrt{2}}; \quad \mu'_4 = \frac{1}{12}$$

und

$$\mu_2 = \frac{1}{4} - \frac{2}{9} = 0,0278$$

$$\mu_3 = \frac{1}{5\sqrt{2}} - 3 \cdot \frac{1}{4} \cdot \frac{\sqrt{2}}{3} + \frac{2 \cdot 2^{3/2}}{27} = -0,0027$$

$$\mu_4 = \frac{1}{12} - 4 \cdot \frac{1}{5\sqrt{2}} \cdot \frac{\sqrt{2}}{3} + 6 \cdot \frac{1}{4} \cdot \frac{2}{9} - \frac{3 \cdot 4}{81} = 0,0018.$$

Setzt man diese Resultate in (5.28) und (5.29) ein, so bekommt man

$$\gamma_1 = \frac{-0,0027}{(0,0278)^{3/2}} = -0,5825,$$

was eine negative Schiefe, d.h. Rechtssteilheit bedeutet. Das wird evident, wenn man sich die Verteilung aufzeichnet. Schließlich ist

$$\gamma_2 = \frac{0,0018}{(0,0278)^2} - 3 = -0,6709,$$

d.h. die Verteilung ist etwas flach. ■

Beispiel 5.4.2. Sei $f(x) = xe^{-x}$ für $0 < x < \infty$ und $f(x) = 0$ sonst (Gammaverteilung). Wie sind die Schiefe und der Exzeß der Verteilung?

Lösung: Es ist

$$\mu'_r = \int_0^{\infty} x^{r+1} e^{-x} dx = (r+1)! \quad [\text{vgl. B2, Nr. 15}]$$

Dieses Integral wird später ausführlicher behandelt, jetzt reicht uns das Resultat

$$\mu'_1 = 2! = 2$$

$$\mu'_2 = 3! = 6$$

$$\mu'_3 = 4! = 24$$

$$\mu'_4 = 5! = 120$$

woraus

$$\mu_2 = 2; \quad \mu_3 = 4; \quad \mu_4 = 24$$

und schließlich

$$\gamma_1 = \frac{4}{2^{3/2}} = 1,4142,$$

was eine Rechtsasymmetrie bedeutet, und

$$\gamma_2 = \frac{24}{4} - 3 = 3,$$

was einen ziemlich großen Exzeß bedeutet. ■

Beispiel 5.4.3. Gesucht sind die Schiefe und der Exzeß der empirischen Verteilung in Tabelle 1.5.

Lösung: Wir berechnen die ersten vier Anfangsmomente tabellarisch (vgl. Tabelle 5.1).

Dann bekommen wir

$$m'_1 = \frac{1}{N} \sum_{i=1}^5 x_i f_i = \frac{440}{148} = 2,9730$$

$$m'_2 = \frac{1}{N} \sum_{i=1}^5 x_i^2 f_i = \frac{1462}{148} = 9,8784$$

$$m'_3 = \frac{1}{N} \sum_{i=1}^5 x_i^3 f_i = \frac{5246}{148} = 35,4459$$

Tabelle 5.1. Berechnung der Anfangsmomente

x_i	f_i	$x_i f_i$	$x_i^2 f_i$	$x_i^3 f_i$	$x_i^4 f_i$
1	12	12	12	12	12
2	35	70	140	280	560
3	54	162	486	1458	4374
4	39	156	624	2496	9984
5	8	40	200	1000	5000
Σ	148	440	1462	5246	19930

$$m'_4 = \frac{1}{N} \sum_{i=1}^5 x_i^4 f_i = \frac{19930}{148} = 134,6622,$$

woraus

$$m_2 = 9,8784 - (2,9730)^2 = 1,0397$$

$$m_3 = 35,4459 - 3 (9,8784) 2,9730 + 2 (2,9730)^3 = -0,1045$$

$$m_4 = 134,6622 - 4 (35,4459) 2,9730 + 6 (9,8784) (2,9730)^2 - 3 (2,9730)^4 = 2,6452$$

und schließlich

$$\gamma_1 = \frac{m_3}{m_2^{3/2}} = \frac{-0,1045}{(1,0397)^{3/2}} = -0,0986,$$

was eine Linkssymmetrie bedeutet, und

$$\gamma_2 = \frac{m_4}{m_2^2} - 3 = \frac{2,6452}{(1,0397)^2} - 3 = -0,5530$$

was eine gewisse Flachheit bedeutet. ■

Übung 5.4.1. Sei $f(x) = x$ für $0 < x < 1$, $f(x) = 2 - x$ für $1 \leq x < 2$ und $f(x) = 0$ sonst.

(a) Zeigen Sie, daß das r -te Anfangsmoment $\mu'_r = \frac{2(2^{r+1} - 1)}{(r+1)(r+2)}$ ist [bedenken Sie,

daß $2^{r+3} - 2^{r+2} = 2^{r+2}$];

(b) berechnen Sie die Schiefe und (c) den Exzeß.

Übung 5.4.2. Berechnen Sie die Schiefe und den Exzeß der Verteilung $f(x) = 2x$ für $0 < x < 1$ und $f(x) = 0$ sonst.

Übung 5.4.3. Berechnen Sie die Schiefe und den Exzeß der empirischen Verteilung

x_i	0	1	2	3
f_i	3	22	10	5

5.5. Die Taylorsche Formel

Es gibt eine Möglichkeit, die Momente einer Zufallsvariablen ganz allgemein zu berechnen. Zu diesem Zweck benötigen wir zunächst ein mathematisches Hilfsmittel: die Taylorsche Formel. Die Methode der Reihenentwicklung nach Taylor hat sehr viele nützliche Anwendungen, von denen wir eine spezielle Variante für unsere Zwecke benutzen werden.

Eine Funktion $h(x)$ kann man durch ein (möglicherweise unendliches) Polynom n -ten Grades ausdrücken:

$$h(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots + a_n x^n, \quad (5.30)$$

womit man eine Form erhält, die die Funktion leicht berechenbar macht, wenn man die Koeffizienten kennt.

Um nun die Koeffizienten a_i zu finden, differenziert man (5.30) n -mal und bekommt

$$h'(x) = a_1 + 2 a_2 x + 3 a_3 x^2 + \dots + n a_n x^{n-1}$$

$$h''(x) = 2 a_2 + 3 \cdot 2 a_3 x + \dots + n(n-1) a_n x^{n-2}$$

$$h'''(x) = 3 \cdot 2 \cdot 1 a_3 + \dots + n(n-1)(n-2) a_n x^{n-3}$$

.....

$$h^{(n)}(x) = n(n-1)(n-2) \dots 3 \cdot 2 \cdot 1 a_n.$$

Setzt man $x = 0$, so wird

$$h(0) = a_0$$

$$h'(0) = a_1$$

$$h''(0) = 2 a_2 = 2! a_2$$

$$h'''(0) = 3 \cdot 2 \cdot 1 a_3 = 3! a_3$$

.....

$$h^{(n)}(0) = n(n-1)(n-2) \dots 3 \cdot 2 \cdot 1 a_n = n! a_n.$$

Aus diesen Ausdrücken berechnet man die Koeffizienten a_i als

$$a_0 = h(0)$$

$$a_1 = h'(0)$$

$$a_2 = \frac{h''(0)}{2!}$$

$$\dots\dots$$

$$a_n = \frac{h^{(n)}(0)}{n!}.$$

Setzt man diese Resultate in (5.30) ein, so erhält man

$$h(x) = h(0) + \frac{h'(0)}{1!}x + \frac{h''(0)}{2!}x^2 + \frac{h'''(0)}{3!}x^3 + \dots + \frac{h^{(n)}(0)}{n!}x^n \quad (5.31)$$

Diese Reihe kann natürlich bis ins Unendliche gehen, und sie ist nur ein Spezialfall der Taylorschen Formel für $x = 0$.

Beispiel 5.5.1. Wir stellen e^x mit Hilfe der Taylorschen Formel als eine unendliche Reihe dar.

Lösung: Alle Ableitungen von e^x sind

$$\left[e^x \right]^{(k)} = e^x.$$

Setzt man dieses Resultat in (5.31) ein, so bekommt man

$$e^x = e^0 + \frac{e^0}{1!}x + \frac{e^0}{2!}x^2 + \frac{e^0}{3!}x^3 + \dots = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad \blacksquare$$

Beispiel 5.5.2. $\ln(1-x)$ soll mit Hilfe eines Polynoms ausgedrückt werden.

Lösung: Zuerst finden wir die einzelnen Ableitungen und setzen $x = 0$:

$$[\ln(1-x)]' = -\frac{1}{1-x} \rightarrow h'(0) = -1$$

$$[\ln(1-x)]'' = -\frac{1}{(1-x)^2} \rightarrow h''(0) = -1$$

$$[\ln(1-x)]''' = -\frac{1 \cdot 2}{(1-x)^3} \rightarrow h'''(0) = -2!$$

$$[\ln(1-x)]^{(4)} = -\frac{1 \cdot 2 \cdot 3}{(1-x)^4} \rightarrow h^{(4)}(0) = -3!$$

$$[\ln(1-x)]^{(k)} = -\frac{(k-1)!}{(1-x)^k} \rightarrow h^{(k)}(0) = -(k-1)!$$

Verwendet man diese Resultate, so bekommt man schließlich

$$\ln(1-x) = \ln(1-0) - 1x - \frac{1x^2}{2!} - \frac{2!x^3}{3!} - \frac{3!x^4}{4!} - \dots =$$

$$= -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \dots \quad \blacksquare \quad (5.32)$$

Übung 5.5.1. Zeigen Sie, daß

$$e^{tx} = 1 + tx + \frac{t^2 x^2}{2!} + \frac{t^3 x^3}{3!} + \dots \quad (5.33)$$

Dieses Resultat wird im nächsten Paragraphen benutzt.

Übung 5.5.2. Zeigen Sie, daß

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \quad (5.34)$$

$$[\text{Bedenken Sie, daß } [\ln(1+x)]^{(k)} = \frac{(-1)^{k-1} (k-1)!}{(1+x)^k}.]$$

5.6. Momenterzeugende Funktionen

In 5.1. haben wir bereits festgestellt, daß man die Erwartung einer Funktion der Zufallsvariablen X auch direkt, ohne vorherige Transformationen berechnen kann (vgl. Formel 5.2). Es gibt einige wichtige Funktionen von Zufallsvariablen, die in der Wahrscheinlichkeitsrechnung eine große Bedeutung haben. Wir werden uns hier nur mit der Funktion e^{tx} beschäftigen. Wenn X eine Zufallsvariable mit der Dichte $f(x)$ ist, dann heißt die Erwartung von e^{tx} die momenterzeugende Funktion von X , wobei t als eine reelle Konstante betrachtet wird, d.h.

$$E(e^{tx}) = M_X(t) = \begin{cases} \sum e^{tx} f(x) & \text{für diskrete Zufallsvariablen} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & \text{für stetige Zufallsvariablen} \end{cases} \quad (5.35)$$

Nehmen wir beispielsweise den stetigen Fall und entwickeln unter dem Integral e^{tx} in eine Taylorsche Reihe. Die Ableitungen nach x ergeben

$$[e^{tx}]' = e^{tx} t, \text{ woraus } h'(0) = t$$

$$[e^{tx}]'' = e^{tx} t^2, \text{ woraus } h''(0) = t^2$$

....

$$[e^{tx}]^{(k)} = e^{tx} t^k, \text{ woraus } h^{(k)}(0) = t^k.$$

Nach Einsetzung in (5.31) folgt daraus, daß

$$e^{tx} = 1 + tx + \frac{t^2 x^2}{2!} + \frac{t^3 x^3}{3!} + \dots$$

Setzt man dieses Resultat in (5.35) ein, so erhält man

$$\begin{aligned} M_x(t) &= \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_{-\infty}^{\infty} \left(1 + tx + \frac{t^2 x^2}{2!} + \frac{t^3 x^3}{3!} + \dots\right) f(x) dx = \\ &= \int_{-\infty}^{\infty} f(x) dx + t \int_{-\infty}^{\infty} x f(x) dx + \frac{t^2}{2!} \int_{-\infty}^{\infty} x^2 f(x) dx + \frac{t^3}{3!} \int_{-\infty}^{\infty} x^3 f(x) dx + \dots \\ &= 1 + t E(X) + \frac{t^2}{2!} E(X^2) + \frac{t^3}{3!} E(X^3) + \dots \\ &= 1 + t \mu'_1 + \frac{t^2}{2!} \mu'_2 + \frac{t^3}{3!} \mu'_3 + \dots \end{aligned} \quad (5.36)$$

Es ist leicht einzusehen, daß diese Funktion nach t differenzierbar ist, wenn sie existiert. Differenziert man (5.36) nach t , so bekommt man

$$M'_x(t) = \mu'_1 + t \mu'_2 + \frac{t^2}{2!} \mu'_3 + \frac{t^3}{3!} \mu'_4 + \dots$$

Setzt man nun $t = 0$, so entfallen alle Glieder dieser Reihe außer dem ersten, weil alle eine Potenz von t enthalten, und es ist

$$[M'_x(t)]_{t=0} = \mu'_1.$$

Die zweite Ableitung ergibt

$$M''_x(t) = \mu'_2 + t \mu'_3 + \frac{t^2}{2!} \mu'_4 + \dots$$

und für $t = 0$ wird es zu

$$[M''_x(t)]_{t=0} = \mu'_2.$$

Allgemein bedeutet das, daß die r -te Ableitung der momenterzeugenden Funktion nach t in $t = 0$ das r -te Anfangsmoment ergibt, d.h.

$$[M_x^{(r)}(t)]_{t=0} = E(X^r) = \mu'_r. \quad (5.37)$$

Wenn $M_x(t)$ existiert und die direkte Berechnung der Momente langwierig ist, kann man diese Tatsache für die Berechnung der Momente anwenden.

Beispiel 5.6.1. Sei $f(x) = e^{-x}$ für $0 < x < \infty$ und $f(x) = 0$ sonst.

(a) Gesucht sind μ'_1 , μ'_2 und μ'_3 direkt;

(b) wie lautet die MEF? μ'_1 , μ'_2 und μ'_3 sollen mit ihrer Hilfe berechnet werden.

Lösung: (a) Nach B2, Nr.15 ist

$$\int_0^{\infty} x^p e^{-bx} dx = \frac{\Gamma(p+1)}{b^{p+1}}.$$

Daher wird

$$\mu'_1 = \int_0^{\infty} x e^{-x} dx = \frac{\Gamma(2)}{1^2} = 1! = 1$$

$$\mu'_2 = \int_0^{\infty} x^2 e^{-x} dx = \frac{\Gamma(3)}{1^3} = 2! = 2$$

$$\mu'_3 = \int_0^{\infty} x^3 e^{-x} dx = \frac{\Gamma(4)}{1^4} = 3! = 6.$$

(b) Die MEF ist

$$\begin{aligned} M_X(t) &= E(e^{tx}) = \int_0^{\infty} e^{tx} e^{-x} dx = \int_0^{\infty} e^{-x(1-t)} dx = \left[\frac{-e^{-x(1-t)}}{1-t} \right]_0^{\infty} = \\ &= \frac{1}{1-t} = (1-t)^{-1}. \end{aligned}$$

Daraus ergibt sich

$$\begin{aligned} \mu'_1 &= [M'_X(t)]_{t=0} = [(1-t)^{-1}]'_{t=0} = [1(1-t)^{-2}]_{t=0} = 1 \\ \mu'_2 &= [M''_X(t)]_{t=0} = [1(1-t)^{-2}]'_{t=0} = [1 \cdot 2(1-t)^{-3}]_{t=0} = 2! = 2 \\ \mu'_3 &= [M'''_X(t)]_{t=0} = [1 \cdot 2(1-t)^{-3}]'_{t=0} = [1 \cdot 2 \cdot 3(1-t)^{-4}]_{t=0} = 3! = 6, \end{aligned}$$

was mit dem obigen Resultat übereinstimmt. ■

Beispiel 5.6.2. Sei $f(x) = 1/6$ für $x = 1, 2, \dots, 6$ und $f(x) = 0$ sonst (diskrete Rechteckverteilung, z.B. die Verteilung der Punktzahl beim Werfen eines Würfels). Wir suchen $E(X)$ und $E(X^2)$ und berechnen sie direkt und mit Hilfe der MEF.

Lösung: Im Beispiel 5.1.1 haben wir festgestellt, daß $E(X) = \mu'_1 = 3,5$, und im Beispiel 5.2.3 war $E(X^2) = \mu'_2 = 91/6$. die MEF ist

$$M_X(t) = \sum_{x=1}^6 e^{tx} \frac{1}{6} = \frac{1}{6} \sum_{x=1}^6 e^{tx}.$$

Hier ist es einfacher, die Summe nicht zu berechnen, sondern so zu belassen. Dann wird

$$M'_X(t) = \frac{1}{6} \sum_{x=1}^6 x e^{tx}$$

und

$$[M'_X(t)]_{t=0} = \frac{1}{6} \sum_{x=1}^6 x = \frac{1}{6} \cdot \frac{6(7)}{2} = 3,5.$$

Weiter ist

$$M''_X(t) = \frac{1}{6} \sum_{x=1}^6 x^2 e^{tx}$$

und

$$[M''_X(t)]_{t=0} = \frac{1}{6} \sum_{x=1}^6 x^2 = \frac{1}{6} \cdot \frac{6(7)(12+1)}{6} = \frac{91}{6}. \blacksquare$$

Die Transformation der Zufallsvariablen ergibt auch transformierte MEF. Zwei Fälle sind besonders wichtig.

$$(1) M_{X+a}(t) = e^{ta} M_X(t) \quad (5.38)$$

$$\text{Beweis: } M_{X+a}(t) = E(e^{t(X+a)}) = E(e^{tX+ta}) = E(e^{tX} e^{ta})$$

$$= e^{ta} E(e^{tX}) = e^{ta} M_X(t). \blacksquare$$

Folgerung: Setzt man $a = -\mu'_1$, so ist

$$M_{X-\mu'_1}(t) = e^{-t\mu'_1} M_X(t). \blacksquare \quad (5.39)$$

Mit Hilfe dieses Ausdruckes lassen sich die Zentralmomente berechnen, denn

$$\left[M^{(r)}_{X-\mu'_1}(t) \right]_{t=0} = \mu_r.$$

Beispiel 5.6.3. Im Beispiel 5.6.1 haben wir festgestellt, daß die MEF für $f(x) = e^{-x}$ für $0 < x < \infty$ und $f(x) = 0$ sonst $(1-t)^{-1}$ war. Gesucht ist die MEF von $Y = X - 1$ (1 war nämlich μ'_1) und daraus μ_2 und μ_3 .

Lösung: Wenn $M_X(t) = (1-t)^{-1}$, dann wird nach (5.39) $M_{X-1}(t) = e^{-t} (1-t)^{-1}$. Daraus ist dann

$$M'_{X-1}(t) = -e^{-t}(1-t)^{-1} + e^{-t}(1-t)^{-2}$$

und

$$\mu_1 = [M'_{X-1}(t)]_{t=0} = -1 + 1 = 0.$$

Weiter ist

$$M''_{X-1}(t) = e^{-t}(1-t)^{-1} - e^{-t}(1-t)^{-2} - e^{-t}(1-t)^{-2} + e^{-t}2(1-t)^{-3}$$

und

$$\mu_2 = [M''_{X-1}(t)]_{t=0} = 1 - 1 - 1 + 2 = 1.$$

Schließlich ist

$$\begin{aligned} M'''_{X-1}(t) = & -e^{-t}(1-t)^{-1} + e^{-t}(1-t)^{-2} + e^{-t}(1-t)^{-2} - \\ & - 2e^{-t}(1-t)^{-3} + e^{-t}(1-t)^{-2} - 2e^{-t}(1-t)^{-3} - \\ & - 2e^{-t}(1-t)^{-3} + 3 \cdot 2e^{-t}(1-t)^{-4}, \end{aligned}$$

woraus

$$\mu_3 = [M'''_{X-1}(t)]_{t=0} = -1 + 1 + 1 - 2 + 1 - 2 - 2 + 6 = 2.$$

Dieselben Resultate bekommt man, wenn man die Zentralmomente mit Hilfe der Anfangsmomente $\mu'_1 = 1$, $\mu'_2 = 2$, $\mu'_3 = 6$ berechnet (vgl. Beispiel 5.5.1). ■

$$(2) M_{aX+b}(t) = e^{tb} M_X(ta) \quad (5.41)$$

$$\begin{aligned} \text{Beweis: } M_{aX+b}(t) &= E(e^{t(aX+b)}) = E(e^{taX+tb}) = E(e^{taX} e^{tb}) = e^{tb} E(e^{taX}) = \\ &= e^{tb} M_X(ta) \end{aligned}$$

Auch hier ist ein Spezialfall wichtig, nämlich

$$M\left(\frac{X-\mu'_1}{\sigma}\right)(t) = e^{-\frac{t\mu'_1}{\sigma}} \cdot M_X\left(\frac{t}{\sigma}\right), \quad (5.42)$$

was automatisch nach Einsetzung von $a = \frac{1}{\sigma}$ und $b = -\frac{\mu'_1}{\sigma}$ folgt.

So transformierte Variablen sind "normiert", und ihre Erwartung ist gleich 0 und ihre Varianz gleich 1 (vgl. Übung 5.2.3). ■

(3) Schließlich kann man zeigen, daß die momenterzeugende Funktion einer Summe von n unabhängigen Zufallsvariablen das Produkt der momenterzeugenden Funktionen einzelner Zufallsvariablen ist, d.h.

$$M_{X_1+X_2+\dots+X_n}(t) = \prod_{i=1}^n M_{X_i}(t) \quad (5.43)$$

$$\begin{aligned} \text{Beweis: } M_{X_1+X_2+\dots+X_n}(t) &= E[e^{t(X_1+X_2+\dots+X_n)}] \\ &= E(e^{tX_1} e^{tX_2} \dots e^{tX_n}) \\ &= E(e^{tX_1}) E(e^{tX_2}) \dots E(e^{tX_n}) \\ &= \prod_{i=1}^n M_{X_i}(t) \quad \blacksquare \end{aligned}$$

Übung 5.6.1. Sei $f(x) = \binom{n}{x} p^x q^{n-x}$ für $x = 0, 1, \dots, n$ und $f(x) = 0$ sonst.

Zeigen Sie, daß $M_X(t) = (q + p e^t)^n$ und berechnen Sie daraus μ'_1 .

Übung 5.6.2. Setzen Sie in der Übung 5.6.1 $Y = X - \mu'_1$ und zeigen Sie, daß $M_Y(t) = e^{-t\mu'_1} (q + p e^t)^n$. Berechnen Sie aus diesem Ausdruck μ_2 .

Aufgaben

- 5.1. Beweisen Sie, daß $E(c) = c$ für eine diskrete Zufallsvariable.
- 5.2. Beweisen Sie, daß $E[c h(x)] = c E[h(x)]$ für eine stetige Zufallsvariable.
- 5.3. Finden Sie für jede der folgenden Verteilungen μ'_i :
 - (a) $f(x) = 4x$ für $0 \leq x \leq \frac{1}{\sqrt{2}}$
 - (b) $f(x) = x e^{-x}$ für $0 < x < \infty$

(c) $f(x) = 6x(1-x)$ für $0 < x < 1$

(d) $f(x) = \frac{6}{5}x(x-1)$ für $1 < x < 2$

(e) $f(x) = \sqrt{x}$ für $0 < x < (\frac{3}{2})^{\frac{2}{3}}$

(f) $f(x) = 2x - 1$ für $\frac{1}{2} < x < \frac{3}{2}$

(g) $f(x) = \frac{1}{2}(x+1)$ für $-1 < x < 1$

(h) $f(x) = \frac{1}{18}(x+3)$ für $-3 < x < 3$

(i) $f(x) = \frac{1}{2(b-a)\sqrt{2x}}$ für $2a < x < 2b$

(j) $f(x) = \frac{1}{2\sqrt{x}}$ für $0 < x < 1$

5.4. Berechnen Sie μ_2 für alle Verteilungen in Aufgabe 5.3.

5.5. Sei $\mu_3(x)$ das dritte Zentralmoment der Zufallsvariablen X , und sei $Y = aX$. Zeigen Sie, daß $\mu_3(y) = a^3 \mu_3(x)$.

5.6. Berechnen Sie die Schiefe und den Exzeß von $f(x) = 1$ für $0 < x < 1$ und $f(x) = 0$ sonst.

5.7. Sei $f(x) = \frac{1}{b-a}$ für $a < x < b$ und $f(x) = 0$ sonst. Berechnen Sie

(a) μ'_r ; (c) die Schiefe;

(b) μ_2, μ_3, μ_4 ; (d) den Exzeß.

5.8. Zeigen Sie, daß $\mu_4 \geq \mu_2^2$.

5.9. Der Ausdruck

$$E|X - \mu'_1|^r$$

heißt *r-tes absolutes Moment* bezüglich μ'_1 . Speziell heißt $E|X - \mu'_1|$ die *mittlere Abweichung*. Berechnen Sie die mittlere Abweichung, wenn $f(x) = 4x^3$ für $0 < x < 1$ und $f(x) = 0$ sonst.

5.10. Zeigen Sie für die Verteilung in Aufgabe 5.9, daß $\sigma > E|X - \mu'_1|$.

5.11. Zeigen Sie allgemein, daß $\sigma \geq E|X - \mu'_1|$ ist.

5.12. Sei $\mu'_4(x)$ das vierte Anfangsmoment der Zufallsvariablen X . Sei $Y = aX + b$. Finden Sie $\mu'_4(y)$. [Setzen Sie $E[(aX + b)^4]$ und entwickeln Sie.]

5.13. Sei $f(x) = \frac{2}{x}$ für $\frac{1}{\sqrt{e}} < x < 1$

(a) Finden Sie $E(\frac{X^2}{2})$;

(b) Finden Sie die Dichte von $Y = \frac{X^2}{2}$ und daraus $E(Y)$.

5.14. Sei $f(x) = \frac{1}{2\sqrt{x}}$ für $0 < x < 1$.

(a) Finden Sie das r -te Anfangsmoment von X^2 ;

(b) finden Sie die Dichte von $Y = X^2$ und zeigen Sie, daß das r -te Anfangsmoment von Y identisch ist mit (a).

5.15. Sei $f(x) = 4x$ für $0 < x < \frac{1}{\sqrt{2}}$ und $f(x) = 0$ sonst.

(a) Finden Sie $E(e^X)$;

(b) finden Sie die Dichte von $Y = e^X$ und daraus $E(Y)$ und vergleichen Sie mit (a).

5.16. Der Median (M) ist derjenige Wert, der die Verteilung in zwei gleiche Teile unterteilt, d.h. $P(X \leq M) = P(X \geq M) = \frac{1}{2}$ (bei diskreten Zufallsvariablen ist es etwas problematischer). Finden Sie den Median für folgende Verteilungen:

(a) $f(x) = 4x^3$ für $0 < x < 1$

(b) $f(x) = \frac{1}{2x^3}$ für $\frac{1}{2} < x < \infty$

(c) $f(x) = \frac{1}{x}$ für $1 < x < e$

(d) $f(x) = \frac{a}{x^2}$ für $a < x < \infty$

(e) $f(x) = x$ für $0 < x < 1$ und $f(x) = 2 - x$ für $1 < x < 2$.

5.17. Sei $f(x) = \frac{3^5}{24} x^4 e^{-3x}$ für $0 < x < \infty$ und $f(x) = 0$ sonst.

(a) Finden Sie μ'_r ;

(b) Finden Sie die momenterzeugende Funktion;

(c) berechnen Sie μ'_1 und μ'_2 aus der momenterzeugenden Funktion.

5.18. Der Modus einer stetigen Verteilung (x_M) ist der häufigste Wert, d.h. das lokale Maximum der Dichtekurve, und er wird berechnet aus $f'(x_M) = 0$, wenn $f''(x_M) < 0$. Finden Sie den Modus von $f(x) = x e^{-x}$ für $0 < x < \infty$. [Setzen Sie $f'(x) = 0$ und berechnen Sie x .]

5.19. Sei $f(x) = x^2 e^{-x^2}$ für $0 < x < \infty$. Berechnen Sie

(a) μ'_r [vgl. B 2];

(b) $\mu_2 \left[\Gamma\left(\frac{5}{2}\right) = \frac{3}{2} \cdot \frac{1}{2} \Gamma\left(\frac{1}{2}\right); \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \right]$;

(c) den Modus.

5.20. Finden Sie die obere Grenze der Wahrscheinlichkeit $P|X - \mu'_1| \geq k \sigma$, wenn

- (a) $k = 1$;
- (b) $k = 2$;
- (c) $k = 3$.

5.21. Sei $X = aY + b$. Zeigen Sie, daß

- (a) $\gamma_1(y) = \gamma_1(x)$ und
- (b) $\gamma_2(y) = \gamma_2(x)$.

Benutzte und weiterführende Literatur: Cramér (1966), Fichtenholz (1971), Fisz (1970), Gnedenko (1968), Kapur-Saxena (1970), Mood-Graybill (1963), Ventcel'-Ovčarov (1973).

6. Multivariate Verteilungen

6.1. Grundbegriffe

Bisher haben wir uns nur mit Verteilungen einer Zufallsvariablen beschäftigt, es ist aber möglich, auch Verteilungen mehrerer Zufallsvariablen zu konstruieren, d.h. solche, bei denen die Wahrscheinlichkeitsfunktion eine Funktion zweier oder mehrerer Veränderlichen ist. Ein Beispiel aus der Linguistik ist die Verteilung der Wortlänge, wenn sie gleichzeitig als Funktion der Phonem- und der Silbenzahl betrachtet wird.

Wir werden uns hier nur mit Wahrscheinlichkeitsfunktionen zweier Zufallsvariablen beschäftigen und werden sie als $f(x,y)$ bezeichnen. Analog zu eindimensionalen Zufallsvariablen müssen auch hier die Bedingungen

$$f(x,y) \geq 0$$

und

$$F(\infty, \infty) = \begin{cases} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1 \\ \sum_x \sum_y f(x,y) = 1 \end{cases} \quad (6.1)$$

erfüllt sein.

Die Verteilungsfunktion ist auch analog, nämlich

$$F(x,y) = P(X \leq x, Y \leq y) = \begin{cases} \int_{-\infty}^y \int_{-\infty}^x f(x',y') dx' dy' \\ \sum_{x' \leq x} \sum_{y' \leq y} f(x',y') \end{cases} \quad (6.2)$$

Die Wahrscheinlichkeit, daß beide Zufallsvariablen in einem bestimmten Intervall Werte annehmen, wird

$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1), \quad (6.3)$$

was in Abb. 6.1 anschaulich dargestellt wird. Um die weiße Fläche zu berechnen, die das in (6.3) angegebene Intervall darstellt, betrachten wir zunächst die ganze Fläche $F(x_2, y_2)$, von der wir die Fläche $F(x_1, y_2)$, bezeichnet mit Rechtsschraffur ($///$), und die Fläche $F(x_2, y_1)$, bezeichnet mit Linksschraffur ($\backslash \backslash \backslash$), abziehen müssen. Da wir dabei die Fläche $F(x_1, y_1)$, bezeichnet sowohl mit Links- als auch mit Rechtsschraffur ($\times \times \times$), zweimal abgezogen haben, müssen wir sie einmal wieder hinzufügen. Daraus resultiert Formel (6.3).

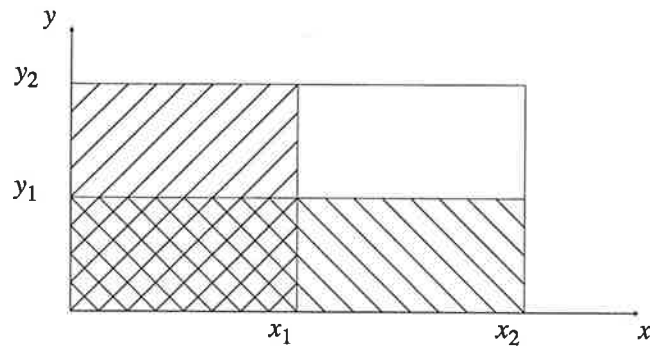


Abb. 6.1. Schema zur Berechnung eines zweidimensionalen Intervalls

Beispiel 6.1.1. Sei

$$f(x, y) = \begin{cases} 1/20 & \text{für } x = 0, y = 0 \\ 3/20 & \text{für } x = 0, y = 1 \\ 6/20 & \text{für } x = 1, y = 0 \\ 2/20 & \text{für } x = 1, y = 1 \\ 5/20 & \text{für } x = 2, y = 0 \\ 3/20 & \text{für } x = 2, y = 1 \\ 0 & \text{sonst} \end{cases}$$

eine zweidimensionale diskrete Wahrscheinlichkeitsfunktion. Wir finden die Verteilungsfunktion $F(x, y)$.

Lösung: Die WF kann man auch tabellarisch darstellen:

$y \backslash x$	0	1
0	$1/20$	$3/20$
1	$6/20$	$2/20$
2	$5/20$	$3/20$

Die Verteilungsfunktion berechnet man durch einfache Addition. Für ein gegebenes $F(x, y)$ addiert man zu der Wahrscheinlichkeit in der Zelle (x, y) alle Zellen links, schräg nach links oben und oberhalb von (x, y) . So kann man vollständig schreiben

	$y < 0$	$0 \leq y < 1$	$y \geq 1$
$x < 0$	0	0	0
$0 \leq x < 1$	0	$1/20$	$4/20$
$1 \leq x < 2$	0	$7/20$	$12/20$
$x \geq 2$	0	$12/20$	1

oder in der üblichen Form

$$F(x, y) = \begin{cases} 0 & \text{für } x < 0 \text{ oder } y < 0 \\ 1/20 & \text{für } 0 \leq x < 1, 0 \leq y < 1 \\ 4/20 & \text{für } 0 \leq x < 1, y \geq 1 \\ 7/20 & \text{für } 1 \leq x < 2, 0 \leq y < 1 \\ 12/20 & \text{für } 1 \leq x < 2, y \geq 1 \\ 12/20 & \text{für } x \geq 2, 0 \leq y < 1 \\ 1 & \text{für } x \geq 2, y \geq 1. \blacksquare \end{cases}$$

Beispiel 6.1.2. Sei $f(x, y) = k(x + y)$ für $0 < x < 1$, $1 < y < 2$ und $f(x, y) = 0$ sonst. Wir bestimmen k .

Lösung: Wir benutzen Gleichung (6.1) und berechnen

$$\begin{aligned} k \int_0^1 \int_1^2 (x + y) dy dx &= k \int_0^1 \left[xy + \frac{y^2}{2} \right]_1^2 dx = k \int_0^1 \left(x + \frac{3}{2} \right) dx = \\ &= k \left[\frac{x^2}{2} + \frac{3x}{2} \right]_0^1 = k \left(\frac{1}{2} + \frac{3}{2} \right) = 2k. \end{aligned}$$

Da nach (6.1) $2k = 1$, folgt daraus

$$k = \frac{1}{2}. \blacksquare$$

Beispiel 6.1.3. Finden Sie die Verteilungsfunktion für die Dichte $f(x,y) = \frac{1}{2}(x+y)$ für $0 < x < 1, 1 < y < 2$ und $f(x,y) = 0$ sonst.

Lösung: Für das Intervall $-\infty < x \leq 0, -\infty < y \leq 1$ ist $F(x,y) = 0$.

Für $0 < x < 1, 1 < y < 2$ wird

$$\begin{aligned} F(x,y) &= \frac{1}{2} \int_0^x \int_1^y (x' + y') dy' dx' = \frac{1}{2} \int_0^x \left[x' y' + \frac{y'^2}{2} \right]_1^y dx' = \\ &= \frac{1}{2} \int_0^x \left(x' y - x' + \frac{y^2}{2} - \frac{1}{2} \right) dx' = \frac{1}{2} \left(\frac{x'^2 y}{2} - \frac{x'^2}{2} + \frac{x' y^2}{2} - \frac{x'}{2} \right) \Big|_0^x = \\ &= \frac{1}{4} x (y-1)(x+y+1). \end{aligned}$$

Für $x \geq 1, 1 < y < 2$ wird lediglich in das vorherige Resultat $x = 1$ eingesetzt, so daß

$$F(x,y) = \frac{1}{4} (y-1)(y+2).$$

Für $0 < x < 1, y \geq 2$ wird wiederum nur $y = 2$ eingesetzt, und es wird

$$F(x,y) = \frac{1}{4} x (x+3)$$

und für $x \geq 1, y \geq 2$ ist dann

$$F(x,y) = 1.$$

Zusammenfassend kann man schreiben

$$F(x,y) = \begin{cases} 0 & \text{für } x \leq 0, y \leq 1 \\ \frac{1}{4} x (y-1)(x+y+1) & \text{für } 0 < x < 1, 1 < y < 2 \\ \frac{1}{4} (y-1)(y+2) & \text{für } x \geq 1, 1 < y < 2 \\ \frac{1}{4} x (x+3) & \text{für } 0 < x < 1, y \geq 2 \\ 1 & \text{für } x \geq 1, y \geq 2. \blacksquare \end{cases}$$

Beispiel 6.1.4. Sei $f(x,y) = \frac{1}{2}(x+y)$ für $0 < x < 1, 1 < y < 2$ und $f(x,y) = 0$ sonst. Finden Sie $P(\frac{1}{2} < X < 1, \frac{3}{2} < Y < 2)$

(a) durch Integration,

(b) mit Hilfe von (6.3).

Lösung:

$$\begin{aligned} \text{(a) } P(\tfrac{1}{2} < X < 1, \tfrac{3}{2} < Y < 2) &= \frac{1}{2} \int_{\frac{1}{2}}^1 \int_{\frac{3}{2}}^2 (x+y) dy dx = \frac{1}{2} \int_{\frac{1}{2}}^1 \left[xy + \frac{y^2}{2} \right]_{\frac{3}{2}}^2 dx = \\ &= \frac{1}{2} \int_{\frac{1}{2}}^1 \left(2x - \frac{3x}{2} + 2 - \frac{1}{2} \cdot \frac{9}{4} \right) dx \\ &= \frac{1}{2} \int_{\frac{1}{2}}^1 \left(\frac{x}{2} + \frac{7}{8} \right) dx = \frac{1}{2} \left[\frac{x^2}{4} + \frac{7x}{8} \right]_{\frac{1}{2}}^1 = \\ &= \frac{1}{2} \left(\frac{1}{4} - \frac{1}{4} \cdot \frac{1}{4} + \frac{7}{8} - \frac{7}{8} \cdot \frac{1}{2} \right) \\ &= \frac{10}{32} = 0,3125. \end{aligned}$$

(b) Aufgrund von $F(x,y)$ im Beispiel 6.1.3 haben wir

$$\begin{aligned} F(1,2) - F\left(\frac{1}{2}, 2\right) - F\left(1, \frac{3}{2}\right) + F\left(\frac{1}{2}, \frac{3}{2}\right) &= \\ &= 1 - \frac{1}{4} \cdot \frac{1}{2} \left(\frac{1}{2} + 3 \right) - \frac{1}{4} \left(\frac{3}{2} - 1 \right) \left(\frac{3}{2} + 2 \right) + \frac{1}{4} \cdot \frac{1}{2} \left(\frac{3}{2} - 1 \right) \left(\frac{1}{2} + \frac{3}{2} + 1 \right) \\ &= 1 - 0,4375 - 0,4375 + 0,1875 = 0,3125. \blacksquare \end{aligned}$$

Übung 6.1.1. Bestimmen Sie k , wenn $f(x,y) = ke^{-(x+y)}$ für $x, y > 0$ und $f(x,y) = 0$ sonst.

Übung 6.1.2. Sei $f(x,y) = 2e^{-x-y}$ für $0 < x < y < \infty$ und $f(x,y) = 0$ sonst. Stellen Sie fest, ob es sich um eine Wahrscheinlichkeitsfunktion handelt.

Übung 6.1.3. Sei $f(x,y) = 4xy$ für $0 < x < 1, 0 < y < 1$ und $f(x,y) = 0$ sonst. Finden Sie $F(x,y)$.

Übung 6.1.4. Sei $f(x,y)$ tabellarisch wie nebenstehend gegeben:

(a) Finden Sie $F(x,y)$;

(b) Berechnen Sie $P(0 \leq X < 1, 2 \leq Y \leq 3)$.

	y		
x \ y	1	2	3
0	$\frac{2}{15}$	$\frac{4}{15}$	$\frac{3}{15}$
1	$\frac{3}{15}$	$\frac{1}{15}$	$\frac{2}{15}$

Übung 6.1.5. Sei $f(x,y) = e^{-(x+y)}$ für $x, y > 0$ und $f(x,y) = 0$ sonst. Berechnen Sie $P(X \leq 1, Y \geq 1)$.

6.2. Randverteilung

Stellen wir uns eine diskrete zweidimensionale Verteilung vor. Die Wahrscheinlichkeiten können tabellarisch angegeben werden wie in Tabelle 6.1. In der Tabelle befinden sich die einzelnen Wahrscheinlichkeiten $f(x,y) = P(X=x, Y=y)$. Es ist nicht unbedingt nötig, daß die Zufallsvariablen mit 1 anfangen. Sie können auch negative Werte oder 0 annehmen.

Daher ist es oft üblich, die Zufallsvariablen noch mit einem Index als x_i bzw. y_i zu kennzeichnen, was wir hier jedoch unterlassen werden. An den Rändern der Tabelle stehen die Summen der einzelnen Zeilen bzw. Spalten, d.h.

$$f(x,.) = \sum_y f(x,y)$$

bzw.

$$f(.,y) = \sum_x f(x,y) \quad (6.4)$$

Tabelle 6.1. Schema einer zweidimensionalen diskreten Verteilung

$X \setminus Y$	1	2	3	...	y	...	n	$f(x,.)$
1	$f(1,1)$	$f(1,2)$	$f(1,3)$	...	$f(1,y)$	...	$f(1,n)$	$f(1,.)$
2	$f(2,1)$	$f(2,2)$	$f(2,3)$	...	$f(2,y)$	...	$f(2,n)$	$f(2,.)$
3	$f(3,1)$	$f(3,2)$	$f(3,3)$	...	$f(3,y)$	...	$f(3,n)$	$f(3,.)$
...	...	...	...	...	...	...	...	...
x	$f(x,1)$	$f(x,2)$	$f(x,3)$	...	$f(x,y)$	...	$f(x,n)$	$f(x,.)$
...	...	...	...	...	...	...	...	...
m	$f(m,1)$	$f(m,2)$	$f(m,3)$	...	$f(m,y)$	...	$f(m,n)$	$f(m,.)$
$f(.,y)$	$f(.,1)$	$f(.,2)$	$f(.,3)$	...	$f(.,y)$	...	$f(.,n)$	1

die nun auch eine Verteilung darstellen, nämlich die Randverteilung von X bzw. Y . Offensichtlich ist $f(x,.) \geq 0, f(.,y) \geq 0$ und $\sum_x f(x,.) = \sum_y f(.,y) = 1$. Die Randverteilungen wollen wir definieren als

$$f_1(x) = \begin{cases} \int_{-\infty}^{\infty} f(x,y) dy & \text{für stetige Zufallsvariablen} \\ \sum_y f(x,y) = f(x,.) & \text{für diskrete Zufallsvariablen} \end{cases} \quad (6.5)$$

und

$$f_2(y) = \begin{cases} \int_{-\infty}^{\infty} f(x,y) dx & \text{für stetige Zufallsvariablen} \\ \sum_x f(x,y) = f(.,y) & \text{für diskrete Zufallsvariablen} \end{cases} \quad (6.6)$$

Sucht man also die Randverteilung X , so muß man über Y summieren bzw. integrieren und umgekehrt.

Beispiel 6.2.1. Finden Sie die Randverteilung von X , wenn $f(x,y) = 4xy$ für $0 < x, y < 1$ und $f(x,y) = 0$ sonst.

$[0 < x, y < 1$ stellt eine verkürzte Schreibweise für $0 < x < 1, 0 < y < 1$ dar]

Lösung: Nach (6.5) ist

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y) dy = \int_0^1 4xy dy = \left[\frac{4xy^2}{2} \right]_0^1 = 2x \text{ für } 0 < x < 1.$$

Man integriert lediglich im Definitionsbereich, da der Integrand außerhalb des Definitionsbereiches sowieso als 0 definiert wurde. ■

Beispiel 6.2.2. Sei $f(x,y)$ folgendermaßen gegeben

$$\begin{aligned} f(0,1) &= 1/10, & f(0,2) &= 2/10, \\ f(1,1) &= 3/10, & f(1,2) &= 4/10, \\ f(x,y) &= 0 \text{ sonst.} \end{aligned}$$

Finden Sie die Randverteilung von Y .

Lösung: Die Randverteilung von Y wird durch die Spaltensummen dargestellt, nämlich

$$f(0,1) + f(1,1) = \frac{1}{10} + \frac{3}{10} = \frac{4}{10} \text{ für } y = 1,$$

$$f(0,2) + f(1,2) = \frac{2}{10} + \frac{4}{10} = \frac{6}{10} \text{ für } y = 2,$$

d.h.

$$f_2(y) = \begin{cases} \frac{2}{5} & \text{für } y = 1 \\ \frac{3}{5} & \text{für } y = 2 \\ 0 & \text{sonst.} \blacksquare \end{cases}$$

Jede Randverteilung hat natürlich auch ihre Verteilungsfunktion, die wir analog mit $F_1(x)$ bzw. $F_2(y)$ bezeichnen und definieren als

$$F_1(x) = F(x, \infty) = \begin{cases} \int_{-\infty}^x f_1(x') dx' = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x', y) dy dx' \\ \sum_{x' \leq x} f(x', \cdot) = \sum_{x' \leq x} \sum_y f(x', y) \end{cases} \quad (6.7)$$

bzw.

$$F_2(y) = F(\infty, y) = \begin{cases} \int_{-\infty}^y f_2(y') dy' = \int_{-\infty}^y \int_{-\infty}^{\infty} f(x, y') dx dy' \\ \sum_{y' \leq y} f(\cdot, y') = \sum_{y' \leq y} \sum_x f(x, y') \end{cases} \quad (6.8)$$

Beispiel 6.2.3. Sei $f(x, y) = \frac{1}{2}(x + y)$ für $0 < x < 1$, $1 < y < 2$ und $f(x, y) = 0$ sonst. Berechnen Sie

(a) $F_1(x)$ und

(b) $F_2(y)$.

Lösung: (a) Der Hauptteil der Verteilungsfunktion ergibt nach (6.7)

$$\begin{aligned} F_1(x) &= \frac{1}{2} \int_0^x \int_1^2 (x' + y) dy dx' = \frac{1}{2} \int_0^x \left[x'y + \frac{y^2}{2} \right]_1^2 dx' = \frac{1}{2} \int_0^x \left(x' + \frac{3}{2} \right) dx' = \\ &= \frac{1}{2} \left[\frac{x'^2}{2} + \frac{3x'}{2} \right]_0^x = \frac{1}{4} x(x + 3) \text{ für } 0 < x < 1. \end{aligned}$$

Daher ist

$$F_1(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ \frac{1}{4} x(x + 3) & \text{für } 0 < x < 1 \\ 1 & \text{für } x \geq 1. \end{cases}$$

Diese Formel haben wir bereits im Beispiel 6.1.3 angetroffen, so daß es leicht ist, die Zusammenhänge von $F(x, y)$, und $F_1(x)$ bzw. $F_2(y)$ zu ersehen.

(b) Der Hauptteil von $F_2(y)$ ergibt

$$\begin{aligned} F_2(y) &= \frac{1}{2} \int_1^y \int_0^1 (x + y') dx dy' = \frac{1}{2} \int_1^y \left[\frac{x^2}{2} + xy' \right]_0^1 dy' = \left[\frac{y'}{2} + \frac{y'^2}{2} \right]_1^y = \\ &= \frac{1}{2} \left(\frac{y}{2} - \frac{1}{2} + \frac{y^2}{2} - \frac{1}{2} \right) = \frac{1}{4} (y - 1)(y + 2) \text{ für } 1 < y < 2. \end{aligned}$$

Daraus ist

$$F_2(y) = \begin{cases} 0 & \text{für } y \leq 1 \\ \frac{1}{4} (y - 1)(y + 2) & \text{für } 1 < y < 2 \\ 1 & \text{für } y \geq 2 \end{cases}$$

was mit dem Resultat in Beispiel 6.1.3 übereinstimmt. ■

Übung 6.2.1. Sei $f(x, y) = 8x(1 - 2y)$ für $0 < x < 1$, $0 < y < \frac{1}{2}$ und $f(x, y) = 0$ sonst. Finden Sie

(a) $f_1(x)$;

(b) $f_2(y)$.

Übung 6.2.2. Sei $f(x, y) = 3x$ für $0 < y < x < 1$ und $f(x, y) = 0$ sonst. Finden Sie

(a) $f_1(x)$;

(b) $f_2(y)$.

[Beim Integrieren sind die Grenzen $(y,1)$ für x bzw. $(0,x)$ für y .]

Übung 6.2.3. Sei $f(x,y)$ folgendermaßen gegeben:

$$f(1,3) = \frac{3}{16}; \quad f(1,4) = \frac{2}{16}; \quad f(1,5) = \frac{5}{16};$$

$$f(2,3) = \frac{1}{16}; \quad f(2,4) = \frac{4}{16}; \quad f(2,5) = \frac{1}{16};$$

und $f(x,y) = 0$ sonst. Finden Sie

(a) $f_1(x)$;

(b) $f_2(y)$.

6.3. Bedingte Verteilung

Dividiert man in Tabelle 6.1 in der Zeile x alle $f(x,y)$ durch ihre Summe $f(x, \cdot)$, d.h. die gemeinsame Verteilung von X und Y durch die Randverteilung von X , so ergeben sich in den einzelnen Zellen neue Größen $f(x,y)/f(x, \cdot)$, die immer größer oder gleich 0 sind und deren Summe 1 ergibt, so daß man sie auch als Wahrscheinlichkeiten auffassen kann. Sie stellen eine neue Verteilung dar, nämlich die bedingte Verteilung von Y unter der Bedingung, daß $X = x$ ist. Aufgrund der Definition der bedingten Wahrscheinlichkeit können wir nun die bedingte Dichte der diskreten Zufallsvariablen Y unter der Bedingung $X = x$ definieren als

$$g(y|x) = \frac{P(X=x, Y=y)}{P(X=x)} = \frac{f(x,y)}{f(x, \cdot)} = \frac{f(x,y)}{f_1(x)}. \quad (6.9)$$

Entsprechend erhält man für die Spalten von Tabelle 6.1 die bedingte Dichte von X unter der Bedingung $Y = y$ als

$$g(x|y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{f(x,y)}{f(\cdot, y)} = \frac{f(x,y)}{f_2(y)}.$$

Für eine stetige Zufallsvariable gilt analog

$$g(y|x) = \frac{f(x,y)}{f_1(x)} \quad (6.11)$$

bzw.

$$g(x|y) = \frac{f(x,y)}{f_2(y)}. \quad (6.12)$$

Die Verteilungsfunktion dieser Zufallsvariablen erhalten wir analog zu der bereits bekannten Definition der Verteilungsfunktion als

$$F(y|x) = \begin{cases} \int_{-\infty}^y f(y'|x) dy' = \frac{\int_{-\infty}^y f(x,y') dy'}{f_1(x)} \\ \sum_{y' \leq y} f(y'|x) = \frac{\sum_{y' \leq y} f(x,y')}{f_1(x)} \end{cases} \quad (6.13)$$

bzw.

$$F(x|y) = \begin{cases} \int_{-\infty}^x f(x'|y) dx' = \frac{\int_{-\infty}^x f(x',y) dx'}{f_2(y)} \\ \sum_{x' \leq x} f(x'|y) = \frac{\sum_{x' \leq x} f(x',y)}{f_2(y)} \end{cases} \quad (6.14)$$

Es läßt sich leicht zeigen, daß $F(\infty|x) = \int_{-\infty}^{\infty} f(y|x) dy = \int_{-\infty}^{\infty} f(x|y) dx = F(\infty|y) = 1$,

denn es ist z.B.

$$F(\infty|y) = \int_{-\infty}^{\infty} f(y|x) dy = \frac{\int_{-\infty}^{\infty} f(x,y) dy}{f_1(x)} = \frac{f_1(x)}{f_1(x)} = 1$$

und analog für $F(\infty|y)$.

Beispiel 6.3.1. Eine diskrete Verteilung ist folgendermaßen angegeben:

$$\begin{array}{lll} f(-1,1) = \frac{3}{21}; & f(-1,2) = \frac{4}{21}; & f(-1,3) = \frac{1}{21}; \\ f(0,1) = \frac{1}{21}; & f(0,2) = \frac{1}{21}; & f(0,3) = \frac{3}{21}; \\ f(1,1) = \frac{2}{21}; & f(1,2) = \frac{5}{21}; & f(1,3) = \frac{1}{21}. \end{array}$$

Berechnen Sie die bedingte Verteilung

(a) von Y , wenn $X = -1$;

(b) von X , wenn $Y = 2$.

Lösung: (a) für $f(y|x)$ brauchen wir die Randverteilung von X . Es ist

$$f_1(x) = f(x, \cdot) = \begin{cases} \frac{8}{21} & \text{für } x = -1 \\ \frac{5}{21} & \text{für } x = 0 \\ \frac{8}{21} & \text{für } x = 1. \end{cases}$$

Für uns ist relevant $f(-1, \cdot) = \frac{8}{21}$. Durch diese Zahl dividieren wir die Zahlen der ersten Zeile und bekommen

$$f(y|x = -1) = \begin{cases} \frac{\frac{3}{21}}{\frac{8}{21}} = \frac{3}{8} & \text{für } y = 1 \\ \frac{\frac{4}{21}}{\frac{8}{21}} = \frac{4}{8} & \text{für } y = 2 \\ \frac{\frac{1}{21}}{\frac{8}{21}} = \frac{1}{8} & \text{für } y = 3 \\ 0 & \text{sonst.} \end{cases}$$

(b) Für diesen Fall brauchen wir $f_2(y)$. Durch Summierung ergibt sich

$$f_2(y) = f(\cdot, y) = \begin{cases} \frac{6}{21} & \text{für } y = 1 \\ \frac{10}{21} & \text{für } y = 2 \\ \frac{5}{21} & \text{für } y = 3 \end{cases}$$

woraus wieder

$$f(x|y = 2) = \begin{cases} \frac{\frac{4}{21}}{\frac{10}{21}} = \frac{4}{10} & \text{für } x = -1 \\ \frac{\frac{1}{21}}{\frac{10}{21}} = \frac{1}{10} & \text{für } x = 0 \\ \frac{\frac{5}{21}}{\frac{10}{21}} = \frac{5}{10} & \text{für } x = 1 \\ 0 & \text{sonst.} \blacksquare \end{cases}$$

Beispiel 6.3.2. Sei $f(x,y) = x + y$ für $0 < x, y < 1$ und $f(x,y) = 0$ sonst. Finden Sie

(a) $f(y|x)$;

(b) $f(y|x = \frac{1}{2})$.

Lösung: (a) Die Randverteilung von X ist

$$f_1(x) = \int_0^1 (x+y) dy = \left[xy + \frac{y^2}{2} \right]_0^1 = x + \frac{1}{2} \quad \text{für } 0 < x < 1,$$

woraus

$$f(y|x) = \frac{x+y}{x+\frac{1}{2}} \quad \text{für } 0 < y < 1, 0 < x < 1.$$

(b) Wenn $x = \frac{1}{2}$, so wird

$$f(y|x = \frac{1}{2}) = \frac{\frac{1}{2} + y}{\frac{1}{2} + \frac{1}{2}} = y + \frac{1}{2} \quad \text{für } 0 < y < 1. \blacksquare$$

Beispiel 6.3.3. Finden Sie für das vorige Beispiel

(a) $F(y|x)$;

(b) $F(y|x = \frac{1}{2})$.

$$\text{Lösung: (a) } F(y|x) = \int_0^y \frac{x+y'}{x+\frac{1}{2}} dy' = \left[\frac{xy' + \frac{y'^2}{2}}{x+\frac{1}{2}} \right]_0^y = \frac{2xy + y^2}{2x+1}.$$

Die bedingte Verteilungsfunktion lautet folglich

$$F(y|x) = \begin{cases} 0 & \text{für } y \leq 0 \\ \frac{2xy + y^2}{2x+1} & \text{für } 0 < y < 1, 0 < x < 1 \\ 1 & \text{für } y \geq 1. \end{cases}$$

(b) In die obere Formel setzen wir $x = \frac{1}{2}$ ein und erhalten

$$F(y|x = \frac{1}{2}) = \begin{cases} 0 & \text{für } y \leq 0 \\ \frac{1}{2} y (1+y) & \text{für } 0 < y < 1 \\ 1 & \text{für } y \geq 1. \blacksquare \end{cases}$$

Beispiel 6.3.4. Sei $f(x,y) = 8xy$ für $0 < y < x < 1$ und $f(x,y) = 0$ sonst. Finden Sie

(a) $f(y|x)$;

(b) $f(x|y)$.

Lösung: (a) $f_1(x) = \int_0^x 8xy \, dy = \left[4xy^2 \right]_0^x = 4x^3$ für $0 < x < 1$

und daraus ist

$$f(y|x) = \frac{8xy}{4x^3} = \frac{2y}{x^2} \quad \text{für } 0 < y < x, 0 < x < 1.$$

(b) $f_2(y) = \int_y^1 8xy \, dx = \left[4x^2y \right]_y^1 = 4y - 4y^3 = 4y(1 - y^2)$ für $0 < y < 1$,

woraus

$$f(x|y) = \frac{8xy}{4y(1 - y^2)} = \frac{2x}{1 - y^2} \quad \text{für } y < x < 1, 0 < y < 1. \blacksquare$$

Die Bedingung muß nicht unbedingt auf einen Punkt beschränkt sein, sondern kann auch durch ein Intervall gegeben werden. Die Formeln leitet man leicht ab, wenn man bedenkt, daß

$$\begin{aligned} F(y|x_1 < X \leq x_2) &= \frac{P(x_1 < X \leq x_2, Y \leq y)}{P(x_1 < X \leq x_2)} \\ &= \frac{\int_{x_1}^{x_2} \int_y^{x_2} f(x,y') \, dx \, dy'}{\int_{x_1}^{x_2} \int_{x_1}^{x_2} f(x,y') \, dx \, dy'} \\ &= \frac{\int_{x_1}^{x_2} f_1(x) \, dx}{\int_{x_1}^{x_2} f_1(x) \, dx} \\ &= \frac{F(x_2, y) - F(x_1, y)}{F_1(x_2) - F_1(x_1)}. \end{aligned}$$

Die Ableitung dieses Ausdrucks nach y ergibt

$$f(y|x_1 < X \leq x_2) = \frac{\int_{x_1}^{x_2} f(x,y) \, dx}{\int_{x_1}^{x_2} f_1(x) \, dx}.$$

So kann man gleichzeitig auch für diskrete Zufallsvariablen definieren

$$f(y|x_1 < X \leq x_2) = \begin{cases} \frac{\sum_{x=x_1+1}^{x_2} f(x,y)}{\sum_{x=x_1+1}^{x_2} f(x,.)} = \frac{\sum_{x=x_1+1}^{x_2} f(x,y)}{F_1(x_2) - F_1(x_1)} \\ \frac{\int_{x_1}^{x_2} f(x,y) \, dx}{\int_{x_1}^{x_2} f_1(x) \, dx} = \frac{\int_{x_1}^{x_2} f(x,y) \, dx}{F_1(x_2) - F_1(x_1)} \end{cases} \quad (6.15)$$

$$F(y|x_1 < X \leq x_2) = \begin{cases} \frac{\sum_{y' \leq y} \sum_{x=x_1+1}^{x_2} f(x,y')}{\sum_{x=x_1+1}^{x_2} \sum_{y' \leq y} f(x,y')} = \frac{\sum_{y' \leq y} \sum_{x=x_1+1}^{x_2} f(x,y')}{F_1(x_2) - F_1(x_1)} \\ \frac{\int_{-\infty}^y \int_{x_1}^{x_2} f(x,y') \, dx \, dy'}{\int_{-\infty}^y \int_{x_1}^{x_2} f(x,y') \, dx \, dy'} = \frac{\int_{-\infty}^y \int_{x_1}^{x_2} f(x,y') \, dx \, dy'}{F_1(x_2) - F_1(x_1)} \end{cases} \quad (6.16)$$

Beispiel 6.3.5. Finden Sie für die Verteilung in Beispiel 6.3.2

(a) $f(y|\frac{1}{3} < X < \frac{1}{2})$;

(b) $F(y|\frac{1}{3} < X < \frac{1}{2})$.

Lösung: Im Beispiel 6.3.2 war $f(x,y) = x + y$ für $0 < x, y < 1$ und $f_1(x) = x + \frac{1}{2}$ für $0 < x < 1$. Daraus folgt

$$(a) \quad f(y | \frac{1}{3} < X < \frac{1}{2}) = \frac{\int_{\frac{1}{3}}^{\frac{1}{2}} (x+y) dx}{\int_{\frac{1}{3}}^{\frac{1}{2}} (x + \frac{1}{2}) dx} = \frac{\left[\frac{x^2}{2} + xy \right]_{\frac{1}{3}}^{\frac{1}{2}}}{\left[\frac{x^2}{2} + \frac{x}{2} \right]_{\frac{1}{3}}^{\frac{1}{2}}} =$$

$$= \frac{1}{11} (5 + 12y) \quad \text{für } 0 < y < 1.$$

$$(b) \quad F(y | \frac{1}{3} < X < \frac{1}{2}) = \int_0^y \frac{1}{11} (5 + 12y') dy'$$

$$= \frac{1}{11} y (5 + 6y) \quad \text{für } 0 < y < 1. \blacksquare$$

Analog zu der Definition der Unabhängigkeit zweier Ereignisse betrachten wir die Zufallsvariablen X und Y als *unabhängig*, wenn

$$f(x,y) = f_1(x) f_2(y), \quad (6.17)$$

d.h. wenn ihre gemeinsame Verteilung das Produkt ihrer Randverteilungen ist. Wenn zwei Zufallsvariablen unabhängig sind, dann ist

$$f(x|y) = \frac{f(x,y)}{f_2(y)} = \frac{f_1(x) f_2(y)}{f_2(y)} = f_1(x),$$

und analog ist

$$f(y|x) = \frac{f(x,y)}{f_1(x)} = \frac{f_1(x) f_2(y)}{f_1(x)} = f_2(y).$$

Beispiel 6.3.6. Sei $f(x,y) = 8x(1-2y)$ für $0 < x < 1$, $0 < y < \frac{1}{2}$ und $f(x,y) = 0$ sonst. Stellen Sie fest, ob X und Y unabhängig sind.

Lösung: In Übung 6.2.1 haben wir festgestellt, daß

$$f_1(x) = 2x \quad \text{für } 0 < x < 1$$

$$f_2(y) = 4(1-2y) \quad \text{für } 0 < y < \frac{1}{2}.$$

Daraus folgt

$$f_1(x) f_2(y) = 2x \cdot 4(1-2y) = 8x(1-2y) = f(x,y),$$

also sind sie unabhängig. $\blacksquare$

Beispiel 6.3.7. Sei $f(x,y) = 3x$ für $0 < y < x < 1$ und $f(x,y) = 0$ sonst. Stellen Sie fest, ob X und Y unabhängig sind.

Lösung: In Übung 6.2.2 haben wir festgestellt, daß

$$f_1(x) = 3x^2 \quad \text{für } 0 < x < 1$$

und

$$f_2(y) = \frac{3}{2} (1 - y^2) \quad \text{für } 0 < y < 1.$$

Da

$$f_1(x) f_2(y) = \frac{9}{2} x^2 (1 - y^2) \neq f(x,y)$$

ist, schließen wir, daß X und Y nicht unabhängig sind. $\blacksquare$

Beispiel 6.3.8. Beweisen Sie, daß $F(x,y) = F_1(x) F_2(y)$, wenn X und Y unabhängig sind.

Lösung: Für eine stetige Verteilung ist

$$F(x,y) = \int_{-\infty}^x \int_{-\infty}^y f(x',y') dy' dx' = \int_{-\infty}^x \int_{-\infty}^y f_1(x') f_2(y') dy' dx' =$$

$$= \int_{-\infty}^x f_1(x') dx' \int_{-\infty}^y f_2(y') dy' = F_1(x) F_2(y). \blacksquare$$

Beispiel 6.3.9. Sei eine diskrete Wahrscheinlichkeitsfunktion $f(x,y)$ folgendermaßen gegeben:

$$f(0,1) = \frac{1}{8}; \quad f(0,2) = \frac{2}{8};$$

$$f(1,1) = \frac{3}{8}; \quad f(1,2) = \frac{2}{8}.$$

Stellen Sie fest, ob X und Y unabhängig sind.

Lösung: Wir berechnen zuerst die Randverteilungen

$$f_1(x) = \begin{cases} \frac{3}{8} & \text{für } x = 0 \\ \frac{5}{8} & \text{für } x = 1 \end{cases}; \quad f_2(y) = \begin{cases} \frac{4}{8} & \text{für } y = 1 \\ \frac{4}{8} & \text{für } y = 2 \end{cases}.$$

Wenn alle $f(x, \cdot) f(\cdot, y) = f(x,y)$ wären, dann wären X und Y unabhängig. Da aber schon

$$f(0, \cdot) f(\cdot, 1) = \frac{3}{8} \cdot \frac{4}{8} = \frac{3}{16} \neq \frac{1}{8} = f(0,1),$$

erübrigt sich die weitere Rechnung, weil X und Y danach nicht unabhängig sind. $\blacksquare$

Übung 6.3.1. Sei $f(x,y) = \frac{1}{2}(x+y)$ für $0 < x < 1$, $1 < y < 2$ und $f(x,y) = 0$ sonst. Finden Sie

- (a) $f(x|y)$; (c) $f(x|y < \frac{3}{2})$.
(b) $F(x|y)$;

Übung 6.3.2. Sei $f(x,y) = \frac{1}{2}$ für $0 < x < 1$, $0 < y < 2$ und $f(x,y) = 0$ sonst. Stellen Sie fest, ob X und Y unabhängig sind.

Übung 6.3.3. Sei

$$f(x,y) = \frac{2(x+y)}{1+2y} \text{ und } f_2(y) = \frac{1}{4}(1+2y) \text{ für } 0 < x < 1 \text{ und } 1 < y < 2.$$

Berechnen Sie

$$P\left(\frac{1}{4} < \frac{X}{2} < \frac{1}{2}\right).$$

[Berechnen Sie erst $f(x,y)$ laut Definition 6.12 und daraus $P(\frac{1}{2} < X < 1)$.]

6.4. Momente mehrdimensionaler Zufallsvariablen

Analog zu eindimensionalen Zufallsvariablen definieren wir das $r+q$ -te Anfangsmoment einer zweidimensionalen Zufallsvariablen (X,Y) mit der Wahrscheinlichkeitsfunktion $f(x,y)$ als

$$\mu'_{rq} = E(X^r Y^q) = \begin{cases} \sum_x \sum_y x^r y^q f(x,y) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^r y^q f(x,y) dx dy \end{cases} \quad (6.18)$$

unter der Bedingung, daß die Summe bzw. das Integral absolut konvergieren. Die Zahl $r+q$ ist die Ordnung des Moments. Man sieht, daß $x^r y^q$ eine Funktion von X und Y ist, sagen wir $h(x,y)$, und im Grunde könnte man auf diese Weise die Erwartung einer beliebigen Funktion $h(x,y)$ definieren. In der Statistik spielen nur die Momente erster und zweiter Ordnung eine bedeutende Rolle. Da sich die Ordnung des Moments aus der Summe von r und q bildet, haben wir zwei Momente erster Ordnung, nämlich μ'_{10} und μ'_{01} . Man sieht sofort, daß

$$\mu'_{10} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x,y) dy dx = \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x,y) dy \right] dx = \int_{-\infty}^{\infty} x f_1(x) dx = E(X) \quad (6.19)$$

ist, d.h. die Erwartung der Randverteilung von X . Analog gilt natürlich auch

$$\mu'_{01} = E(Y).$$

Beispiel 6.4.1. Sei $f(x,y) = 2$ für $0 < x < y < 1$ und $f(x,y) = 0$ sonst. Berechnen Sie

- (a) μ'_{10}
(b) μ'_{01} .

Lösung:

$$(a) \mu'_{10} = E(X) = \int_0^1 \int_0^y 2x dx dy = \int_0^1 \left[x^2 \right]_0^y dy = \int_0^1 y^2 dy = \left[\frac{y^3}{3} \right]_0^1 = \frac{1}{3};$$

$$(b) \mu'_{01} = E(Y) = \int_0^1 \int_0^y 2y dx dy = \int_0^1 [2xy]_0^y dy = \int_0^1 2y^2 dy = \left[\frac{2y^3}{3} \right]_0^1 = \frac{2}{3}. \blacksquare$$

Es gibt drei Momente der zweiten Ordnung, nämlich

$$\mu'_{20} = E(X^2 Y^0) = E(X^2)$$

$$\mu'_{02} = E(X^0 Y^2) = E(Y^2) \quad (6.20)$$

$$\mu'_{11} = E(X^1 Y^1) = E(XY)$$

und es ist offensichtlich, daß $\mu'_{00} = 1$.

Beispiel 6.4.2. Berechnen Sie die Momente zweiter Ordnung im vorigen Beispiel.

Lösung:

$$(a) \mu'_{20} = E(X^2) = \int_0^1 \int_0^y 2x^2 dx dy = \int_0^1 \frac{2y^3}{3} dy = \frac{1}{6};$$

$$(b) \mu'_{02} = E(Y^2) = \int_0^1 \int_0^y 2y^2 dx dy = \int_0^1 y^3 dy = \frac{1}{4};$$

$$(c) \mu'_{11} = E(XY) = \int_0^1 \int_0^y 2xy dx dy = \int_0^1 y^3 dy = \frac{1}{4}. \blacksquare$$

Die zentralen Momente definieren wir als

$$\mu_{rq} = E[(X - \mu'_{10})^r (Y - \mu'_{01})^q] = \begin{cases} \sum_x \sum_y (x - \mu'_{10})^r (y - \mu'_{01})^q f(x, y) \\ \int \int_{-\infty-\infty}^{\infty \infty} (x - \mu'_{10})^r (y - \mu'_{01})^q f(x, y) dx dy. \end{cases} \quad (6.21)$$

Für die einzelnen Ordnungen ergeben sich folgende wichtige Spezialfälle:

$$\begin{aligned} \mu_{00} &= 1 \\ \mu_{10} &= E(X - \mu'_{10}) = E(X) - \mu'_{10} = 0 \\ \mu_{01} &= E(Y - \mu'_{01}) = E(Y) - \mu'_{01} = 0, \end{aligned}$$

was aus (6.19) folgt. Die Zentralmomente zweiter Ordnung lassen sich aus Anfangsmomenten berechnen. Es bestehen folgende Beziehungen:

$$\begin{aligned} \mu_{20} &= E[(X - \mu'_{10})^2] = E(X^2) - \mu'^2_{10} = \mu'_{20} - \mu'^2_{10} = \sigma_x^2 \\ \mu_{02} &= E[(Y - \mu'_{01})^2] = E(Y^2) - \mu'^2_{01} = \mu'_{02} - \mu'^2_{01} = \sigma_y^2, \end{aligned} \quad (6.22)$$

wobei man die zweite Gleichung mit Hilfe von (5.13) leicht bekommt.

Das Zentralmoment μ_{11} heißt *Kovarianz* und wird alternativ als

$$\text{Cov}(X, Y) = \sigma_{xy}$$

bezeichnet. Es läßt sich auch mit Hilfe von Anfangsmomenten ausdrücken, denn es ist

$$\begin{aligned} \mu_{11} &= E[(X - \mu'_{10})(Y - \mu'_{01})] \\ &= E(XY - X\mu'_{01} - Y\mu'_{10} + \mu'_{10}\mu'_{01}) \\ &= E(XY) - \mu'_{01}E(X) - \mu'_{10}E(Y) + \mu'_{10}\mu'_{01} \\ &= \mu'_{11} - \mu'_{01}\mu'_{10} - \mu'_{10}\mu'_{01} + \mu'_{10}\mu'_{01} = \end{aligned}$$

$$= \mu'_{11} - \mu'_{10}\mu'_{01} = \sigma_{xy} = \text{Cov}(X, Y). \quad (6.23)$$

Mit Hilfe der Zentralmomente läßt sich eine weitere sehr wichtige Größe, nämlich der Korrelationskoeffizient ρ_{xy} , definieren als

$$\begin{aligned} \rho_{xy} &= \frac{E[(X - \mu'_{10})(Y - \mu'_{01})]}{\sqrt{E[(X - \mu'_{10})^2]} \sqrt{E[(Y - \mu'_{01})^2]}} = \frac{\mu_{11}}{\sqrt{\mu_{20}\mu_{02}}} = \frac{\sigma_{xy}}{\sigma_x\sigma_y} = \\ &= \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}, \end{aligned} \quad (6.24)$$

der in der Korrelationsstatistik eine bedeutende Rolle spielt. Die Kovarianz läßt sich demnach auch als

$$\sigma_{xy} = \sigma_x \sigma_y \rho_{xy}$$

schreiben.

Beispiel 6.4.3. Berechnen Sie die Zentralmomente zweiter Ordnung und den Korrelationskoeffizienten, wenn $f(x, y) = 2$ für $0 < x < y < 1$ und $f(x, y) = 0$ sonst.

Lösung: Im Beispiel 6.4.1 und 6.4.2 haben wir festgestellt, daß

$$\mu'_{10} = \frac{1}{3}; \quad \mu'_{01} = \frac{2}{3}; \quad \mu'_{20} = \frac{1}{6}; \quad \mu'_{02} = \frac{1}{2}; \quad \mu'_{11} = \frac{1}{4}.$$

Daraus berechnen wir

$$\begin{aligned} \mu_{20} &= \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{6} - \frac{1}{9} = \frac{3-2}{18} = \frac{1}{18}; \\ \mu_{02} &= \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}; \\ \mu_{11} &= \frac{1}{4} - \frac{1}{3} \cdot \frac{2}{3} = \frac{1}{36}. \end{aligned}$$

Setzt man diese Resultate in (6.24) ein, so bekommt man

$$\rho_{xy} = \frac{\frac{1}{36}}{\sqrt{\frac{1}{18}}\sqrt{\frac{1}{18}}} = \frac{1}{2}. \blacksquare$$

Die bedingten Momente definieren wir folgendermaßen

$$E(X^r | Y = y) = \begin{cases} \sum_x x^r f(x|y) = \sum_x \frac{x^r f(x,y)}{f_2(y)} \\ \int_{-\infty}^{\infty} x^r f(x|y) dx = \int_{-\infty}^{\infty} \frac{x^r f(x,y)}{f_2(y)} dx \end{cases} \quad (6.25)$$

und ähnlich für $E(Y^q | X = x)$. Als Bedingung kann jedoch nicht nur ein einzelner Wert, sondern auch ein ganzes Intervall angegeben werden. In diesem Falle definieren wir

$$E(X^r | y_1 < Y \leq y_2) = \begin{cases} \sum_x \sum_{y=y_1+1}^{y_2} \frac{x^r f(x,y)}{F_2(y_2) - F_2(y_1)} \\ \int_{-\infty}^{\infty} \int_{y_1}^{y_2} \frac{x^r f(x,y) dy dx}{F_2(y_2) - F_2(y_1)} \end{cases} \quad (6.26)$$

und ähnlich für die bedingten Momente von Y . Auf ähnliche Weise kann man natürlich auch die bedingte Erwartung einer Funktion beider Zufallsvariablen definieren als

$$E(X^r Y^q | y_1 < Y \leq y_2) = \begin{cases} \sum_x \sum_{y=y_1+1}^{y_2} \frac{x^r y^q f(x,y)}{F_2(y_2) - F_2(y_1)} \\ \int_{-\infty}^{\infty} \int_{y_1}^{y_2} \frac{x^r y^q f(x,y) dy dx}{F_2(y_2) - F_2(y_1)} \end{cases} \quad (6.27)$$

Beispiel 6.4.4. Sei $f(x,y) = 8xy$ für $0 < y < x < 1$ und $f(x,y) = 0$ sonst. Finden Sie

- (a) $E(YX)$;
- (b) $E(Y | X = 1/2)$;
- (c) $E(Y | 1/3 < X \leq 1/2)$.

Lösung: In Beispiel 6.3.4 haben wir festgestellt, daß $f(y|x) = 2y/x^2$ für $0 < y < x$, $0 < x < 1$. Daher ist nach (6.25)

$$(a) E(YX) = \int_0^1 \int_0^x \frac{y \cdot 2y}{x^2} dy = \frac{1}{x^2} \int_0^x 2y^2 dy = \frac{1}{x^2} \left[\frac{2y^3}{3} \right]_0^x = \frac{2}{3} x.$$

(b) Wenn $X = 1/2$, dann ist nach (6.25)

$$E(Y | X = 1/2) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}.$$

Man kann jedoch auch direkt rechnen, indem man erst $X = 1/2$ einsetzt und dann integriert; es ist nämlich $f(y|x = 1/2) = \frac{2y}{1/4} = 8y$, woraus

$$E(Y | X = 1/2) = \int_0^{1/2} 8y^2 dy = \left[\frac{8y^3}{3} \right]_0^{1/2} = \frac{8}{3} \cdot \frac{1}{8} = \frac{1}{3}.$$

(c) Da nach Beispiel 6.3.4 $f_1(x) = 4x^3$ ist, so ist

$$F_1(1/2) - F_1(1/3) = \frac{1}{16} - \frac{1}{81} = \frac{65}{1296}.$$

Setzt man das in die Formel für die bedingte Erwartung von Y ein (vgl. 6.26), so wird

$$E(Y | 1/3 < X \leq 1/2) = \frac{1296}{65} \int_{1/3}^{1/2} \int_0^x 8xy^2 dy dx = \frac{1296}{65} \int_{1/3}^{1/2} \frac{8x^4}{3} dx = 0,28855. \blacksquare$$

Beispiel 6.4.5. Sei eine diskrete Zufallsvariable $f(x,y)$ folgendermaßen angegeben:

$$\begin{array}{lll} f(-1,1) = 3/21; & f(-1,2) = 4/21; & f(-1,3) = 1/21; \\ f(0,1) = 1/21; & f(0,2) = 1/21; & f(0,3) = 3/21; \\ f(1,1) = 2/21; & f(1,2) = 5/21; & f(1,3) = 1/21. \end{array}$$

Berechnen Sie

- (a) $E(X | Y = 2)$;
- (b) $E(X | 1 < Y \leq 3)$.

Lösung:(a) In Beispiel 6.3.1 haben wir festgestellt, daß

$$f(x | y = 2) = \begin{cases} 4/10 & \text{für } x = -1 \\ 1/10 & \text{für } x = 0 \\ 5/10 & \text{für } x = 1 \\ 0 & \text{sonst.} \end{cases}$$

$$\text{Daher ist } E(X|Y=2) = \sum_x f(x|y=2) = -1 \cdot \frac{4}{10} + 0 \cdot \frac{1}{10} + 1 \cdot \frac{5}{10} = \frac{1}{10};$$

(b) Die Randverteilung von $f_2(y)$ ergab

$$f_2(y) = \begin{cases} \frac{6}{21} & \text{für } y=1 \\ \frac{10}{21} & \text{für } y=2 \\ \frac{5}{21} & \text{für } y=3 \\ 0 & \text{sonst.} \end{cases}$$

$$\text{Daraus folgt } F_2(3) = 1 \text{ und } F_2(1) = \frac{6}{21}, \text{ so daß } F_2(3) - F_2(1) = 1 - \frac{6}{21} = \frac{15}{21}.$$

Dann ist

$$\begin{aligned} E(X|1 < Y \leq 3) &= \sum_{x=-1}^1 \sum_{y=2}^3 \frac{x f(x,y)}{\frac{15}{21}} \\ &= \frac{21}{15} \left[-1 \cdot \frac{4}{21} + (-1) \cdot \frac{1}{21} + 0 \cdot \frac{1}{21} + 0 \cdot \frac{3}{21} + 1 \cdot \frac{5}{21} + 1 \cdot \frac{1}{21} \right] \\ &= \frac{21}{15} \cdot \frac{1}{21} = \frac{1}{15}. \blacksquare \end{aligned}$$

Wir beweisen nun einige wichtige Sätze über die genannten Charakteristika.

$$(1) \quad E(X + Y) = E(X) + E(Y). \quad (6.28)$$

Beweis: Für eine stetige Zufallsvariable ist

$$\begin{aligned} E(X + Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f(x,y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x,y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x,y) dx dy \\ &= E(X) + E(Y). \blacksquare \end{aligned}$$

Beispiel 6.4.6. Sei $f(x,y) = 8xy$ für $0 < y < x < 1$ und $f(x,y) = 0$ sonst. Zeigen Sie, daß $E(X + Y) = E(X) + E(Y)$.

$$\text{Lösung: } E(X + Y) = \int_0^1 \int_0^x (x + y) 8xy dy dx = \frac{4}{3}$$

$$E(X) = \int_0^1 \int_0^x x(8xy) dy dx = \frac{4}{5}$$

$$E(Y) = \int_0^1 \int_0^x y(8xy) dy dx = \frac{8}{15}.$$

Das Resultat stimmt, da

$$\frac{4}{3} = \frac{4}{5} + \frac{8}{15}. \blacksquare$$

Obwohl wir uns in diesem Kapitel nur mit zweidimensionalen Zufallsvariablen beschäftigen, wird dieser Satz hier auf n Dimensionen verallgemeinert, da wir ihn später brauchen werden.

(2) Seien $X_1, X_2, \dots, X_n$ Zufallsvariablen mit den Erwartungen $E(X_1), E(X_2), \dots, E(X_n)$, und sei $U = X_1 + X_2 + \dots + X_n$. Dann ist

$$E(U) = \sum_{i=1}^n E(X_i). \quad (6.29)$$

Beweis: Als Folgerung des ersten Satzes kann man die Summe schrittweise auf jeweils zwei Summanden aufteilen, so daß

$$\begin{aligned} E(U) &= E\left(\sum_i X_i\right) \\ &= E(X_1 + X_2 + \dots + X_n) = \end{aligned}$$

$$= E(X_1) + E(X_2 + \dots + X_n)$$

$$= \dots$$

$$= E(X_1) + E(X_2) + \dots + E(X_n)$$

$$= \sum_i E(X_i) .$$

Dieser Satz behauptet, daß die Erwartung einer Summe die Summe der Erwartungen ist. Ähnliches gilt auch, wenn U eine lineare Funktion von X_i ist:

Sei $U = a_1X_1 + a_2X_2 + \dots + a_nX_n$, dann ist offensichtlich

$$E(U) = E\left(\sum_i a_i X_i\right) = \sum_i E(a_i X_i) . \blacksquare \quad (6.30)$$

(3) Wenn X und Y unabhängig sind, dann ist

$$E(XY) = E(X) E(Y) . \quad (6.31)$$

Beweis: Für eine diskrete Zufallsvariable ist

$$\begin{aligned} E(XY) &= \sum_x \sum_y xy f(x,y) = \sum_x \sum_y xy f(x,.) f(.,y) = \\ &= \sum_x x f(x,.) \sum_y y f(.,y) = E(X) E(Y) . \blacksquare \end{aligned}$$

(4) Wenn X und Y unabhängig sind, dann ist

$$\text{Cov}(X,Y) = \mu_{11} = 0. \quad (6.32)$$

Beweis: $\text{Cov}(X,Y) = E[X - E(X)] [Y - E(Y)] = E(XY) - E(X) E(Y).$

Falls X und Y unabhängig sind, dann gilt Satz (3), wonach $E(XY) = E(X) E(Y)$, so daß $\text{Cov}(X,Y) = 0$. ■

(5) Für die Varianz der Summe gilt der Satz

$$V(X \pm Y) = \begin{cases} V(X) + V(Y) \pm 2\text{Cov}(X,Y) , & \text{allgemein} \\ V(X) + V(Y) , & \text{wenn } X \text{ und } Y \text{ unabhängig sind.} \end{cases} \quad (6.33)$$

Beweis: Sei $U = X + Y$, dann ist

$$V(X+Y) = V(U) = E[U - E(U)]^2 = E(U^2) - [E(U)]^2 =$$

$$= E[(X+Y)^2] - [E(X+Y)]^2$$

$$= E(X^2 + 2XY + Y^2) - [E(X) + E(Y)]^2$$

$$= E(X^2) + 2E(XY) + E(Y^2) - [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2$$

$$= (E(X^2) - [E(X)]^2) + (E(Y^2) - [E(Y)]^2) + 2[E(XY) - E(X)E(Y)]$$

$$= V(X) + V(Y) + 2 \text{Cov}(X,Y) .$$

Damit ist der allgemeine Teil des Satzes bewiesen. Wenn nun X und Y unabhängig sind, dann folgt aus (6.32), daß $\text{Cov}(X,Y) = 0$ und $V(X+Y) = V(X) + V(Y)$, was den speziellen Teil beweist. Für den zweiten Fall, die Differenz, haben wir

$$V(X-Y) = E[(X-Y)^2] - [E(X-Y)]^2$$

$$= E(X^2 - 2XY + Y^2) - [E(X) - E(Y)]^2$$

$$= (E(X^2) - [E(X)]^2) + (E(Y^2) - [E(Y)]^2) - 2[E(XY) - E(X)E(Y)] =$$

$$= V(X) + V(Y) - 2\text{Cov}(X, Y),$$

und für unabhängige Zufallsvariablen wird analog $V(X - Y) = V(X) + V(Y)$. ■

Beispiel 6.4.7. Sei $f(x, y) = 2$ für $0 < x < y < 1$ und $f(x, y) = 0$ sonst. Zeigen Sie, daß $V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$.

Lösung: Im Beispiel 6.4.3 haben wir festgestellt, daß

$$V(X) = \mu_{20} = \frac{1}{18}; \quad V(Y) = \mu_{02} = \frac{1}{18}; \quad \text{Cov}(X, Y) = \mu_{11} = \frac{1}{36}.$$

Daraus ergibt sich

$$V(X) + V(Y) + 2\text{Cov}(X, Y) = \frac{1}{18} + \frac{1}{18} + 2 \cdot \frac{1}{36} = \frac{1}{6}.$$

Die direkte Berechnung von $V(X + Y)$ ergibt wegen

$$E(X + Y) = E(X) + E(Y) = \frac{1}{3} + \frac{2}{3} = 1$$

(vgl. Beispiel 6.4.3)

$$\begin{aligned} V(X + Y) &= \int_0^1 \int_0^y (x + y - [E(X) + E(Y)])^2 \cdot 2 \, dx \, dy \\ &= \int_0^1 \int_0^y 2(x + y - 1)^2 \, dx \, dy = \frac{1}{6}, \end{aligned}$$

was mit dem obigen Resultat übereinstimmt. ■

Beispiel 6.4.8. Das Resultat in (6.33) illustrieren wir noch anhand eines empirischen Beispiels. An 4 Objekten wurden zwei Eigenschaften X und Y gemessen. Es ergaben sich folgende Werte

	Objekte			
	a	b	c	d
x	4	5	8	3
y	9	7	3	11
$u = x + y$	13	12	11	14
$z = x - y$	-5	-2	5	-8

Zeigen Sie, daß

$$S_{x \pm y}^2 = S_x^2 + S_y^2 \pm 2 S_{xy}.$$

Lösung:

$$\begin{aligned} S_x^2 &= \frac{1}{4} \left(\sum_i x_i^2 - \frac{(\sum_i x_i)^2}{4} \right) = \frac{1}{4} (114 - 100) = 3,5 \\ S_y^2 &= \frac{1}{4} (260 - 225) = 8,75. \end{aligned}$$

Wegen $\bar{x} = 5$ und $\bar{y} = 7,5$ bekommen wir

$$\begin{aligned} S_{xy} &= \frac{1}{4} \sum_i (x_i - 5)(y_i - 7,5) \\ &= \frac{1}{4} (\sum_i x_i y_i - 5 \sum_i y_i - 7,5 \sum_i x_i + 4(5)7,5) \\ &= \frac{1}{4} (128 - 150 - 150 + 150) = -5,5. \end{aligned}$$

Daher ist

$$S_{x+y}^2 = S_x^2 + S_y^2 + 2 S_{xy} = 3,5 + 8,75 + 2(5,5) = 1,25$$

und

$$S_{x-y}^2 = S_x^2 + S_y^2 - 2 S_{xy} = 3,5 + 8,75 + 2(5,5) = 23,25.$$

Berechnet man die Varianzen für die Summe von X und Y und für die Differenz von X und Y , d.h. für die Zahlen in der unteren Hälfte der obigen Tabelle, so bekommt man

$$\begin{aligned} S_{x+y}^2 &= S_u^2 = \frac{1}{4} \left(\sum_i u_i^2 - \frac{(\sum_i u_i)^2}{4} \right) = \frac{1}{4} (630 - 625) = 1,25 \\ S_{x-y}^2 &= S_z^2 = \frac{1}{4} \left(\sum_i z_i^2 - \frac{(\sum_i z_i)^2}{4} \right) = \frac{1}{4} (118 - 25) = 23,25 \end{aligned}$$

was mit dem obigen Resultat übereinstimmt. ■

Auch dieser Satz hat eine wichtige Verallgemeinerung, die wir gleich in Form einer linearen Funktion zeigen.

(6) Seien $X_1, X_2, \dots, X_n$ Zufallsvariablen mit den Erwartungen $E(X_1), E(X_2), \dots, E(X_n)$ und den Varianzen $V(X_1) = \sigma_1^2, V(X_2) = \sigma_2^2, \dots, V(X_n) = \sigma_n^2$, und sei $U = a_1 X_1 + a_2 X_2 + \dots + a_n X_n$. Dann ist

$$V(U) = \begin{cases} \sum_i^n a_i^2 \sigma_i^2 + 2 \sum_{i < j}^n \sum_j^n a_i a_j \sigma_{ij} \\ \sum_i^n a_i^2 \sigma_i^2, \text{ wenn die } X_i \text{ voneinander unabhängig sind.} \end{cases} \quad (6.34)$$

Beweis: Aus Satz (2) folgt, daß $E(U) = a_1 E(X_1) + \dots + a_n E(X_n)$. Daher ist

$$U - E(U) = a_1 X_1 + a_2 X_2 + \dots + a_n X_n - a_1 E(X_1) - a_2 E(X_2) - \dots - a_n E(X_n)$$

$$= a_1 [X_1 - E(X_1)] + a_2 [X_2 - E(X_2)] + \dots + a_n [X_n - E(X_n)] .$$

Daraus folgt, daß

$$\begin{aligned} V(U) &= E([U - E(U)]^2) \\ &= E(\{a_1[X_1 - E(X_1)] + a_2[X_2 - E(X_2)] + \dots + a_n[X_n - E(X_n)]\}^2) \\ &= E\{a_1^2 [X_1 - E(X_1)]^2 + a_2^2 [X_2 - E(X_2)]^2 + \dots + a_n^2 [X_n - E(X_n)]^2 + \\ &\quad + 2a_1a_2 [X_1 - E(X_1)] [X_2 - E(X_2)] + \\ &\quad + 2a_1a_3 [X_1 - E(X_1)] [X_3 - E(X_3)] + \\ &\quad + \dots + 2a_{n-1}a_n [X_{n-1} - E(X_{n-1})] [X_n - E(X_n)]\} = \\ &= a_1^2 V(X_1) + a_2^2 V(X_2) + \dots + a_n^2 V(X_n) + 2a_1a_2 \text{Cov}(X_1, X_2) + \\ &\quad + 2a_1a_3 \text{Cov}(X_1, X_3) + \dots + 2a_{n-1}a_n \text{Cov}(X_{n-1}, X_n) = \end{aligned}$$

$$= \sum_i^n a_i^2 V(X_i) + 2 \sum_{i < j}^n \sum_j^n a_i a_j \text{Cov}(X_i, X_j)$$

$$= \sum_i^n a_i^2 \sigma_i^2 + 2 \sum_{i < j}^n \sum_j^n a_i a_j \sigma_{ij} .$$

Wenn die Zufallsvariablen voneinander unabhängig sind, dann fällt die Summe der Kovarianzen weg, und es gilt

$$V(U) = \sum_i^n a_i^2 \sigma_i^2 . \blacksquare$$

Bemerkung: Man kann alternativ schreiben

$$2 \sum_{i < j}^n \sum_j^n a_i a_j \sigma_{ij} = \sum_{i \neq j}^n \sum_j^n a_i a_j \sigma_{ij} .$$

Beispiel 6.4.9. Als Beispiel zeigen wir hier einen speziellen Fall, der später eine wichtige Rolle spielen wird. Seien $X_1, \dots, X_n$ unabhängige Zufallsvariablen, die alle die gleiche Verteilung haben. Finden Sie $V(\bar{X})$, d.h. die Varianz des Mittelwertes.

Lösung: Da alle X_i gleich verteilt sind, ist $V(X_i) = \sigma^2$. Weiter ist

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n}X_1 + \frac{1}{n}X_2 + \dots + \frac{1}{n}X_n,$$

so daß $a_i = \frac{1}{n}$. Dann ist

$$\begin{aligned} V(\bar{X}) &= V\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{1}{n^2} V(X_1 + X_2 + \dots + X_n) = \\ &= \frac{1}{n^2} \sum_i^n V(X_i) = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n} . \blacksquare \end{aligned}$$

(7) Für den Korrelationskoeffizienten gilt

$$\rho_{ax+b, cy+d} = \pm \rho_{xy}. \quad (6.35)$$

Beweis: $\rho_{ax+b, cy+d} = \frac{E([aX+b - E(aX+b)] [cY+d - E(cY+d)])}{\sqrt{V(aX+b)} \sqrt{V(cY+d)}}.$

Benutzt man die Sätze (5.6) und (5.15), so bekommt man sofort

$$\rho_{ax+b, cy+d} = \frac{ac E([X - E(X)][Y - E(Y)])}{\pm ac \sqrt{V(X)} \sqrt{V(Y)}} = \pm \rho_{xy},$$

woraus folgt, daß der Korrelationskoeffizient bei einer Verschiebung aller X - und Y -Werte um eine Konstante und ihre Multiplikation mit einer Konstanten seinen Wert nicht ändert, d.h. daß er gegenüber einer linearen Transformation invariant ist. ■

(8) Für den Wertebereich von ρ beweisen wir, daß

$$-1 \leq \rho \leq 1 \quad (6.36)$$

Beweis: Da eine additive und eine multiplikative Konstante ρ nicht verändern, transformieren wir X und Y auf

$$U = \frac{X - E(X)}{\sqrt{V(X)}}, \quad Z = \frac{Y - E(Y)}{\sqrt{V(Y)}}.$$

Aus Übung 5.2.3 ist bereits bekannt, daß $E(U) = E(Z) = 0$ und $V(U) = V(Z) = 1$. Daher ist

$$\rho_{uz} = \frac{\text{Cov}(U, Z)}{\sqrt{V(U)} \sqrt{V(Z)}} = \text{Cov}(U, Z).$$

Nun betrachten wir anhand von (6.33) den immer nichtnegativen Ausdruck (die Varianz ist immer ≥ 0)

$$V(U + Z) = V(U) + V(Z) + 2\text{Cov}(U, Z) = 1 + 1 + 2\rho_{uz} = 2 + 2\rho_{uz} \geq 0,$$

woraus folgt, daß

$$\begin{aligned} 2\rho_{uz} &\geq -2 \\ \rho_{uz} &\geq -1, \end{aligned}$$

was die linke Ungleichung von (6.37) beweist. Weiter ist laut (6.33)

$$V(U - Z) = V(U) + V(Z) - 2\text{Cov}(U, Z) = 1 + 1 - 2\rho_{uz} = 2 - 2\rho_{uz} \geq 0,$$

woraus folgt, daß

$$\begin{aligned} -2\rho_{uz} &\geq -2 \\ -\rho_{uz} &\geq -1 \\ \rho_{uz} &\leq 1, \end{aligned}$$

was die rechte Ungleichung von (6.36) beweist. Daher gilt

$$-1 \leq \rho \leq 1. \quad \blacksquare$$

(9) Für die bedingte Erwartung gilt

$$E_x[E_y(Y|X)] = E(Y), \quad (6.37)$$

d.h., die Erwartung von Y unter der Bedingung X , wenn X alle seine Werte durchläuft, ist wieder die Erwartung von Y . Das erste E betrifft also X , das zweite Y .

Beweis:
$$\begin{aligned} E_x[E_y(Y|X)] &= \int_{-\infty}^{\infty} E_y(Y|X) f_1(x) dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \frac{f(x, y)}{f_1(x)} f_1(x) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy = E(Y) \text{ laut Definition. } \blacksquare \end{aligned}$$

Auch mehrdimensionale Zufallsvariablen haben momenterzeugende Funktionen. Wir definieren sie für eine zweidimensionale Zufallsvariable als

$$M_{X,Y}(t_1, t_2) = E(e^{t_1 X + t_2 Y}) \quad (6.38)$$

Aus dieser Funktion kann man die Momente auf dieselbe Weise berechnen wie bei der eindimensionalen Zufallsvariablen. Ebenso wie in (5.31) entwickeln wir (6.38) in eine Taylorreihe mit Hilfe partieller Ableitungen in der Umgebung von (0,0). Es ist

$$\begin{aligned} h(t_1, t_2) = e^{t_1 x + t_2 y} &= h(0,0) + \frac{\partial h(0,0)}{\partial t_1} \cdot t_1 + \frac{\partial h(0,0)}{\partial t_2} \cdot t_2 + \\ &+ \frac{\partial^2 h(0,0)}{2! \partial t_1^2} \cdot t_1^2 + \frac{\partial^2 h(0,0)}{2! \partial t_2^2} \cdot t_2^2 + \\ &+ \frac{\partial^2 h(0,0)}{1! 1! \partial t_1 \partial t_2} \cdot t_1 t_2 + \dots + \\ &+ \frac{\partial^{r+q} h(0,0)}{r! q! \partial t_1^r \partial t_2^q} \cdot t_1^r t_2^q + \dots \end{aligned}$$

Die erste Ableitung von $h(x,y)$ nach t_1 ergibt

$$\frac{\partial h(x,y)}{\partial t_1} = x e^{t_1 x + t_2 y},$$

woraus

$$\left. \frac{\partial h(x,y)}{\partial t_1} \right|_{t_i=0} = x e^{0x+0y} = x.$$

Ähnlich ergibt sich

$$\left. \frac{\partial h(x,y)}{\partial t_2} \right|_{t_i=0} = y e^{0x+0y} = y$$

und

$$\left. \frac{\partial^2 h(x,y)}{\partial t_1 \partial t_2} \right|_{t_i=0} = x y e^{0x+0y} = xy \text{ usw.}$$

Setzt man diese Ausdrücke in $h(x,y)$ ein, so bekommt man

$$\begin{aligned} h(x,y) = e^{t_1 x + t_2 y} &= 1 + t_1 x + t_2 y + \frac{(t_1 x)^2}{2!} + \frac{(t_2 y)^2}{2!} + \frac{t_1 t_2 xy}{1! 1!} + \dots + \\ &+ \frac{(t_1 x)^r (t_2 y)^q}{r! q!} + \dots \end{aligned}$$

Analog zu (5.36) kann man also schreiben

$$\begin{aligned} M_{X,Y}(t_1, t_2) &= E(e^{t_1 X + t_2 Y}) \\ &= E[1 + t_1 X + t_2 Y + \frac{(t_1 X)^2}{2!} + \frac{(t_2 Y)^2}{2!} + \frac{t_1 t_2 XY}{1! 1!} + \\ &+ \dots + \frac{(t_1 X)^r (t_2 Y)^q}{r! q!} + \dots] = \\ &= 1 + t_1 E(X) + t_2 E(Y) + \frac{t_1^2}{2!} E(X^2) + \frac{t_2^2}{2!} E(Y^2) + \frac{t_1 t_2}{1! 1!} E(XY) + \\ &+ \dots + \frac{t_1^r t_2^q}{r! q!} E(X^r Y^q) + \dots = \\ &= 1 + t_1 \mu'_{10} + t_2 \mu'_{01} + \frac{t_1^2}{2!} \mu'_{20} + \frac{t_2^2}{2!} \mu'_{02} + \frac{t_1 t_2}{1! 1!} \mu'_{11} + \dots + \\ &+ \frac{t_1^r t_2^q}{r! q!} \mu'_{rq} + \dots \end{aligned}$$

Die einzelnen Ableitungen dieser Funktion nach t_i ($i = 1, 2$) ergeben dann die einzelnen Anfangsmomente, nämlich

$$\begin{aligned} \left. \frac{\partial^r M_{X,Y}(t_1, t_2)}{\partial t_1^r} \right|_{t_i=0} &= \mu'_{r0} \\ \left. \frac{\partial^q M_{X,Y}(t_1, t_2)}{\partial t_2^q} \right|_{t_i=0} &= \mu'_{0q} \\ \left. \frac{\partial^{r+q} M_{X,Y}(t_1, t_2)}{\partial t_1^r \partial t_2^q} \right|_{t_i=0} &= \mu'_{rq} \end{aligned} \quad (6.39)$$

Beispiel 6.4.10. Sei $f(x, y) = e^{-x-y}$ für $x, y > 0$ und $f(x, y) = 0$ sonst. Finden Sie

- (a) die momenterzeugende Funktion;
- (b) $E(X)$;
- (c) $E(X^2)$;
- (d) $E(XY)$;
- (e) $E(X^2 Y^2)$.

Lösung: (a) Für die momenterzeugende Funktion erhalten wir

$$\begin{aligned} M_{X,Y}(t_1, t_2) &= \int_0^\infty \int_0^\infty e^{t_1 x + t_2 y} e^{-x-y} dx dy \\ &= \int_0^\infty \int_0^\infty e^{-x(1-t_1)} e^{-y(1-t_2)} dy dx \\ &= \int_0^\infty \frac{e^{-y(1-t_1)}}{1-t_2} dy \\ &= \frac{1}{(1-t_1)(1-t_2)} = (1-t_1)^{-1} (1-t_2)^{-1}, \end{aligned}$$

was man leicht mit Hilfe von Anhang B2, Nr. 8 berechnet.

(b) Um $E(X)$ zu berechnen, setzen wir

$$\frac{\partial M_{X,Y}(t_1, t_2)}{\partial t_1} = 1 (1-t_1)^{-2} (1-t_2)^{-1}.$$

Setzt man nun $t_1 = t_2 = 0$, so bekommt man

$$E(X) = 1.$$

(c) Für $E(X^2)$ differenzieren wir nochmals nach t_1 und bekommen

$$\frac{\partial^2 M_{X,Y}(t_1, t_2)}{\partial t_1^2} = 2 (1-t_1)^{-3} (1-t_2)^{-1},$$

woraus nach Gleichsetzung von $t_1 = t_2 = 0$ folgt

$$E(X^2) = 2$$

Es ist leicht zu sehen, daß $E(X^r) = r!$ ist.

(d) Für $E(XY)$ differenzieren wir den Ausdruck in (b) nach t_2 und bekommen

$$\frac{\partial M_{X,Y}(t_1, t_2)}{\partial t_1 \partial t_2} = 1 (1-t_1)^{-2} 1 (1-t_2)^{-2},$$

woraus $E(XY) = 1$ folgt.

(e) Für $E(X^2 Y^2)$ differenzieren wir schließlich den Ausdruck in (c) zweimal nach t_2 und bekommen

$$\frac{\partial^4 M_{X,Y}(t_1, t_2)}{\partial t_1^2 \partial t_2^2} = 2(1-t_1)^{-3} 2(1-t_2)^{-3},$$

woraus $E(X^2 Y^2) = 2 \cdot 2 = 4$ folgt.

Auch hier ist leicht zu sehen, daß $E(X^r Y^q) = r!q!$. Die Richtigkeit des Resultats kann man durch direkte Berechnung der Momente leicht nachprüfen (benutzen Sie Anhang B2, Nr. 15). ■

Übung 6.4.1. Sei $f(x, y) = 6(x-y)$ für $0 < y < x < 1$ und $f(x, y) = 0$ sonst. Berechnen Sie μ_{11} . [Berechnen Sie $E(X)$, $E(Y)$, $E(XY)$ und daraus μ_{11} .]

Übung 6.4.2. Sei $f(x, y) = 3x$ für $0 < y < x < 1$ und $f(x, y) = 0$ sonst. Berechnen Sie

(a) μ'_{20} ;

(b) μ'_{02} .

Übung 6.4.3. Sei $f(x,y) = \frac{1}{2}(x+y)$ für $0 < x < 1$, $1 < y < 2$ und $f(x,y) = 0$ sonst. Berechnen Sie ρ_{xy} .

Übung 6.4.4. Sei $f(x,y) = 2e^{-x-y}$ für $0 < x < y < \infty$ und $f(x,y) = 0$ sonst. Berechnen Sie

(a) $E(X)$;

(b) $M_{X,Y}(t_1, t_2)$ und daraus $E(X)$.

6.5. Funktionen zweidimensionaler Zufallsvariablen

Funktionen mehrdimensionaler Zufallsvariablen spielen in der Statistik eine bedeutende Rolle. Man kann durch Transformationen neue Verteilungen ableiten, die in der Inferenzstatistik verwendet werden. Die Ableitung ist in den wichtigsten Fällen glücklicherweise nicht allzu schwierig, es ist aber aus Erfahrung bekannt, daß dieses Kapitel im allgemeinen trotzdem ziemlich große Schwierigkeiten verursacht. Es ist deswegen empfehlenswert, diesem Paragraphen besondere Aufmerksamkeit zu widmen.

Wir werden uns nur mit den geläufigsten Fällen, nämlich mit dem Quotienten, mit dem Produkt, mit der Summe und mit der Differenz beschäftigen und beschränken uns auf zweidimensionale Zufallsvariablen. Eine ausführliche Behandlung findet der Leser bei Wadsworth und Bryan (1960), Papoulis (1965), Harris (1966) u.a.

Allgemein sei (X,Y) eine zweidimensionale stetige Zufallsvariable mit der Dichte $f(x,y)$. Nun betrachten wir zwei stetige Funktionen von X und Y

$$U = h_1(X,Y)$$

$$V = h_2(X,Y),$$

die partielle Ableitungen haben. Wir verfahren analog zum eindimensionalen Fall und ersetzen $f(x,y) dx dy$ unter dem zweifachen Integral

$$\iint f(x,y) dx dy$$

durch Ausdrücke, in denen nur U und V vorkommen. Die Umkehrfunktionen ergeben sofort

$$X = h_1^{-1}(U,V)$$

$$Y = h_2^{-1}(U,V).$$

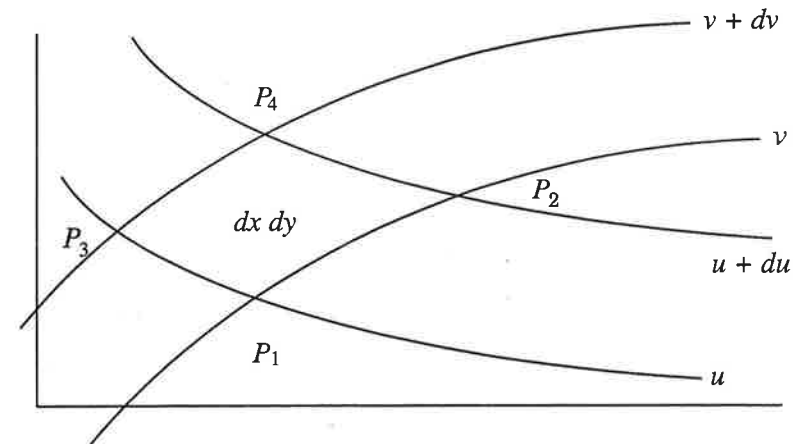


Abb. 6.2.

Um nun $dx dy$ zu ersetzen, betrachten wir die Fläche $P_1P_2P_3P_4$ in der Abbildung 6.2, die durch die Linien u , $u + du$, v , $v + dv$ gebildet wird. Wir führen eine Transformation des Koordinatensystems durch, wobei die Fläche $P_1P_2P_3P_4$ invariant bleibt. Die Koordinaten der vier Eckpunkte ergeben sich aus den Umkehrfunktionen, die, falls nötig, in eine Taylorreihe entwickelt werden, wobei die Glieder höherer Ordnung verschwindend klein werden und vernachlässigt werden können. So bekommen wir für die einzelnen Punkte die Koordinaten

P_1 :

$$x_1 = h_1^{-1}(u,v)$$

$$y_1 = h_2^{-1}(u,v)$$

$$(6.40a)$$

P_2 :

$$x_2 = h_1^{-1}(u + du, v)$$

$$y_2 = h_2^{-1}(u + du, v)$$

$$= h_1^{-1}(u,v) + \frac{\partial h_1^{-1}(u,v)}{\partial u} \cdot du$$

$$= h_2^{-1}(u,v) + \frac{\partial h_2^{-1}(u,v)}{\partial u} \cdot du$$

$$(6.40b)$$

P_3 :

$$\begin{aligned} x_3 &= h_1^{-1}(u, v + dv) & y_3 &= h_2^{-1}(u, v + dv) \\ &= h_1^{-1}(u, v) + \frac{\partial h_1^{-1}(u, v)}{\partial v} \cdot dv & &= h_2^{-1}(u, v) + \frac{\partial h_2^{-1}(u, v)}{\partial v} \cdot dv \end{aligned} \quad (6.40c)$$

P_4 :

$$\begin{aligned} x_4 &= h_1^{-1}(u + du, v + dv) & y_4 &= h_2^{-1}(u + du, v + dv) \\ &= h_1^{-1}(u, v) + \frac{\partial h_1^{-1}(u, v)}{\partial u} \cdot du + \frac{\partial h_1^{-1}(u, v)}{\partial v} \cdot dv & &= h_2^{-1}(u, v) + \frac{\partial h_2^{-1}(u, v)}{\partial u} \cdot du + \frac{\partial h_2^{-1}(u, v)}{\partial v} \cdot dv \end{aligned} \quad (6.40d)$$

Das Flächenelement $P_1P_2P_3P_4$ ist so klein, daß wir die Linien, auf denen die Koordinaten liegen, zwischen den Punkten als Geraden betrachten können, wobei $P_1P_2 = P_3P_4$ und $P_1P_3 = P_2P_4$. Das Flächenelement bildet also ein Parallelogramm, das aus zwei gleichen Dreiecken besteht. Es reicht daher, den Flächeninhalt des Dreiecks $P_1P_2P_3$ zu berechnen. Der Flächeninhalt des Parallelogramms ist dann das Doppelte davon. Aus der analytischen Geometrie ist bekannt, daß die Fläche dieses Dreiecks als die Determinante

$$D = \frac{1}{2} \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

gegeben ist. Wir nehmen gleich das Doppelte, subtrahieren die erste Spalte von der zweiten und der dritten, so daß

$$dx \, dy = \begin{vmatrix} 1 & 0 & 0 \\ x_1 & x_2 - x_1 & x_3 - x_1 \\ y_1 & y_2 - y_1 & y_3 - y_1 \end{vmatrix} = \begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{vmatrix}.$$

Setzt man jetzt die Ausdrücke aus (6.40a-c) ein, so wird

$$dx \, dy = \begin{vmatrix} \left[h_1^{-1}(u, v) + \frac{\partial h_1^{-1}(u, v)}{\partial u} du \right] - h_1^{-1}(u, v) & \left[h_1^{-1}(u, v) + \frac{\partial h_1^{-1}(u, v)}{\partial v} dv - h_1^{-1}(u, v) \right] \\ \left[h_2^{-1}(u, v) + \frac{\partial h_2^{-1}(u, v)}{\partial u} du \right] - h_2^{-1}(u, v) & \left[h_2^{-1}(u, v) + \frac{\partial h_2^{-1}(u, v)}{\partial v} dv - h_2^{-1}(u, v) \right] \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial h_1^{-1}(u, v)}{\partial u} du & \frac{\partial h_1^{-1}(u, v)}{\partial v} dv \\ \frac{\partial h_2^{-1}(u, v)}{\partial u} du & \frac{\partial h_2^{-1}(u, v)}{\partial v} dv \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial h_1^{-1}(u, v)}{\partial u} & \frac{\partial h_1^{-1}(u, v)}{\partial v} \\ \frac{\partial h_2^{-1}(u, v)}{\partial u} & \frac{\partial h_2^{-1}(u, v)}{\partial v} \end{vmatrix} \cdot du \, dv$$

$$= \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \cdot du \, dv \quad (6.41a)$$

$$= \left| \frac{\partial x}{\partial u} \cdot \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \cdot \frac{\partial y}{\partial u} \right| \cdot du \, dv. \quad (6.41b)$$

Die Determinante von (6.41a) wird als *Jacobische Funktionaldeterminante* oder auch als *Jacobian* bezeichnet. Bezeichnet man den Jacobian mit J , so bekommt man

$$\iint f(x, y) \, dx \, dy = \iint f[h_1^{-1}(u, v), h_2^{-1}(u, v)] |J| \, du \, dv$$

d.h. die neue Zufallsvariable (U, V) hat die Dichte

$$g(u,v) = f[h_1^{-1}(u,v), h_2^{-1}(u,v)] |J| . \quad (6.41c)$$

Gewöhnlich braucht man nur die Randverteilung von U , die als

$$g(u) = \int_{D_v} f[h_1^{-1}(u,v), h_2^{-1}(u,v)] |J| dv \quad (6.42)$$

zu berechnen ist, wobei die Integration über das Intervall D_v verläuft, das von Fall zu Fall bestimmt werden muß.

Bei einer diskreten Zufallsvariablen (X,Y) gewinnen wir die Dichte von $U = h(X,Y)$ durch einfache Summation aller $f(x,y)$ -Werte, für die $h(x,y) = u$ ergibt, d.h.

$$g(u) = P(U = u) = \sum_{\substack{x \\ h(x,y) = u}} \sum_y f(x,y) . \quad (6.43)$$

Sind die Zufallsvariablen X und Y unabhängig, so kann man schreiben

$$g(u) = \begin{cases} \int_{D_v} f_1[h_1^{-1}(u,v)] f_2[h_2^{-1}(u,v)] |J| dv \\ \sum_{\substack{x \\ h(x,y) = u}} \sum_y f(x,y) . \end{cases} \quad (6.44)$$

Dieses allgemeine Verfahren wenden wir jetzt auf die oben genannten Fälle (Quotienten, Produkte u.a.) an. Andere Transformationen folgen jedoch demselben Schema.

(1) *Quotient*. Beim Quotienten von X und Y wählen wir

$$\begin{aligned} U &= h_1(X,Y) = \frac{X}{Y} \\ V &= h_2(X,Y) = Y . \end{aligned} \quad (6.45)$$

Die Umkehrfunktionen sind

$$\begin{aligned} X &= h_1^{-1}(U,V) = UV \\ Y &= h_2^{-1}(U,V) = V . \end{aligned} \quad (6.46)$$

Der Jacobian ergibt

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = 1(v) - 0(u) = v . \quad (6.47)$$

Wir setzen die neue Dichte zusammen und bekommen

$$g(u,v) = f[h_1^{-1}(u,v), h_2^{-1}(u,v)] |J| = f(uv, v) |v| . \quad (6.48)$$

Da wir die Verteilung des Quotienten $U = X/Y$ brauchen, berechnen wir die Randverteilung als

$$g(u) = \int_{-\infty}^{\infty} f(uv, v) |v| dv . \quad (6.49)$$

Dies bedeutet, daß man in der ursprünglichen Dichte $f(x,y)$ jedes x durch uv und jedes y durch v ersetzt, die neue Funktion mit $|v|$ multipliziert und schließlich nach v integriert. Das Resultat ergibt die Verteilung des Quotienten.

Beispiel 6.5.1. Sei $f(x,y) = a^2 e^{-a(x+y)}$ für $0 < x, y < \infty$ und $f(x,y) = 0$ sonst ($a > 0$ konstant). Finden Sie die Dichte von $U = X/Y$.

Lösung: Nach (6.45) wird

$$\left. \begin{aligned} u &= x/y \\ v &= y \end{aligned} \right\} \Rightarrow \left. \begin{aligned} x &= uv \\ y &= v \end{aligned} \right\} \Rightarrow |J| = |v| .$$

Setzen wir diese Werte in $f(x,y)$ ein, so erhalten wir

$$g(u) = a^2 \int_0^{\infty} v e^{-v(au+a)} dv = a^2 \frac{\Gamma(2)}{(au+a)^2} = (u+1)^{-2} \text{ für } 0 < u < \infty ,$$

was man leicht mit Hilfe von Anhang B2 nachvollzieht. Man überzeugt sich leicht, daß $g(u)$ tatsächlich eine korrekte Dichte ist, denn $F(\infty) = 1$ und $g(u) \geq 0$, wenn $0 < u < \infty$.

(2) **Produkt.** Beim Produkt von X und Y wählen wir

$$\begin{aligned} U &= XY \\ V &= Y, \end{aligned} \quad (6.50)$$

woraus die Umkehrfunktion

$$\begin{aligned} X &= \frac{U}{V} \\ Y &= V \end{aligned} \quad (6.51)$$

und der Jacobian

$$J = \begin{vmatrix} \frac{1}{v} & -\frac{u}{v^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{v}. \quad (6.52)$$

Daraus folgt die neue Dichte

$$g(u, v) = f\left(\frac{u}{v}, v\right) \frac{1}{|v|} \quad (6.53)$$

und schließlich

$$g(u) = \int_{-\infty}^{\infty} f\left(\frac{u}{v}, v\right) \frac{1}{|v|} dv. \quad (6.54)$$

Beispiel 6.5.2. Sei $f(x, y) = xe^{-x(1+y)}$ für $x, y > 0$ und $f(x, y) = 0$ sonst. Finden Sie die Dichte von $U = XY$.

Lösung: Nach (6.50) ist

$$\left. \begin{aligned} u &= xy \\ v &= y \end{aligned} \right\} \Rightarrow \left. \begin{aligned} x &= \frac{u}{v} \\ y &= v \end{aligned} \right\} \Rightarrow |J| = \frac{1}{v}$$

und daraus

$$g(u, v) = \frac{1}{v} \left(\frac{u}{v} \right) e^{-\frac{u}{v}(1+v)} = u v^{-2} e^{-\frac{u}{v} - u} \text{ für } u, v > 0.$$

Die Dichte von U wird dann

$$g(u) = u e^{-u} \int_0^{\infty} v^{-2} e^{-\frac{u}{v}} dv.$$

Um leichter zu integrieren, transformieren wir noch $z = 1/v$, woraus

$$v = \frac{1}{z}, \quad \left| \frac{dv}{dz} \right| = \frac{1}{z^2}.$$

Setzen wir diese Größe in $g(u)$ ein, so wird

$$g(u) = u e^{-u} \int_0^{\infty} z^2 e^{-uz} \frac{1}{z^2} dz = u e^{-u} \int_0^{\infty} e^{-uz} dz = u e^{-u} \left[\frac{-e^{-uz}}{u} \right]_0^{\infty} = e^{-u} \text{ für } u > 0. \blacksquare$$

Beispiel 6.5.3. Sei eine diskrete Zufallsvariable folgendermaßen angegeben:

$$\begin{array}{lll} f(1,1) = \frac{3}{21} & f(1,2) = \frac{4}{21} & f(1,3) = \frac{1}{21} \\ f(2,1) = \frac{1}{21} & f(2,2) = \frac{1}{21} & f(2,3) = \frac{3}{21} \\ f(3,1) = \frac{2}{21} & f(3,2) = \frac{5}{21} & f(3,3) = \frac{1}{21}. \end{array}$$

Man finde die Dichte von $U = XY$.

Lösung: Bei diskreten Zufallsvariablen entfällt der Jacobian. Man berechnet lediglich

$$P(U = u) = \sum_{\substack{x \ y \\ x \cdot y = u}} f(x, y),$$

d.h. es ist immer $xy = u$. Da $x = 1, 2, 3$ und $y = 1, 2, 3$, ergeben ihre Produkte $u = 1, 2, 3, 4, 6, 9$. So wird

$$P(U = 1) = f(1,1) = \frac{3}{21}$$

$$P(U = 2) = f(1,2) + f(2,1) = \frac{4}{21} + \frac{1}{21} = \frac{5}{21}$$

$$P(U = 3) = f(1,3) + f(3,1) = \frac{1}{21} + \frac{2}{21} = \frac{3}{21}$$

$$P(U=4) = f(2,2) = 1/21$$

$$P(U=6) = f(2,3) + f(3,2) = 3/21 + 5/21 = 8/21$$

$$P(U=9) = f(3,3) = 1/21.$$

Zusammen kann man schreiben

$$g(u) = \begin{cases} 3/21 & u=1 \\ 5/21 & u=2 \\ 3/21 & u=3 \\ 1/21 & u=4 \\ 8/21 & u=6 \\ 1/21 & u=9 \\ 0 & \text{sonst.} \end{cases}$$

(3) *Summe.* Bei der Summe von X und Y wählen wir

$$\begin{aligned} U &= X + Y \\ V &= Y. \end{aligned} \quad (6.55)$$

Die Umkehrfunktionen sind

$$\begin{aligned} X &= U - V \\ Y &= V \end{aligned} \quad (6.56)$$

und der Jacobian ist

$$J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1 \quad |J| = 1.$$

Daraus folgt die Dichte

$$g(u,v) = f(u-v, v) \quad (6.57)$$

und schließlich

$$g(u) = \int_{-\infty}^{\infty} f(u-v, v) dv. \quad (6.58)$$

Diesen Ausdruck nennt man auch *Faltung von Zufallsvariablen*. Der Anfänger stößt bei der Bestimmung der Integrationsgrenzen besonders hier auf große Schwierigkeiten.

Beispiel 6.5.4. Sei $f(x,y) = e^{-(x+y)}$ für $x,y > 0$ und $f(x,y) = 0$ sonst. Finden Sie die Dichte von $U = X + Y$.

Lösung: Nach (6.55) ist

$$\begin{cases} u = x + y \\ v = y \end{cases} \Rightarrow \begin{cases} x = u - v \\ y = v \end{cases} \Rightarrow |J| = 1$$

so daß

$$g(u,v) = e^{-u} \text{ für } u > 0, 0 < v < u.$$

Dann ist

$$g(u) = e^{-u} \int_0^u dv = e^{-u} [v]_0^u = ue^{-u} \text{ für } u > 0.$$

Aus (6.56) ist ersichtlich, daß v nur bis u laufen kann, denn wenn es größer wäre, dann wäre $x < 0$, und in diesem Fall ist $f(x,y) = 0$. ■

Beispiel 6.5.5. Sei $f(x,y)$ folgendermaßen gegeben:

$$\begin{array}{lll} f(1,1) = 1/15 & f(1,2) = 3/15 & f(1,3) = 3/15 \\ f(2,1) = 2/15 & f(2,2) = 4/15 & f(2,3) = 2/15 \end{array}$$

Finden Sie die Dichte von $U = X + Y$.

Lösung: Für einzelne Werte von U summiert man diejenigen $f(x,y)$, bei denen $x+y = u$, d.h.

$$g(u) = \sum_{\substack{x \\ x+y=u}} \sum_y f(x,y).$$

Dann ist

$$g(u) = \begin{cases} 1/15 & \text{für } u=2 \\ 2/15 + 3/15 = 5/15 & \text{für } u=3 \\ 4/15 + 3/15 = 7/15 & \text{für } u=4 \\ 2/15 & \text{für } u=5 \\ 0 & \text{sonst.} \end{cases}$$

(4) Differenz. Bei der Differenz von X und Y wählen wir

$$\begin{aligned} U &= X - Y \\ V &= X. \end{aligned} \quad (6.59)$$

Die Umkehrfunktionen sind

$$\begin{aligned} X &= V \\ Y &= V - U \end{aligned} \quad (6.60)$$

und der Jacobian ist

$$J = \begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} = 1$$

d.h. $|J| = 1$.

Daraus folgt die Dichte

$$g(u, v) = f(v, v - u) \quad (6.61)$$

und schließlich

$$g(u) = \int_{-\infty}^{\infty} f(v, v - u) dv. \quad (6.62)$$

Beispiel 6.5.6. Sei $f(x, y) = a^2 e^{-1(x+y)}$ für $x, y > 0$ und $f(x, y) = 0$ sonst. Es handelt sich hier eigentlich um zwei unabhängige Exponentialverteilungen $f(x) = ae^{-ax}$ und $f(y) = ae^{-ay}$. Finden Sie die Dichte von $U = X - Y$.

Lösung: Nach (6.59) setzen wir

$$\begin{cases} u = x - y \\ v = x \end{cases} \Rightarrow \begin{cases} x = v \\ y = v - u \end{cases} \Rightarrow |J| = 1,$$

wonach jetzt wegen Unabhängigkeit

$$g(u, v) = f(v, v - u) = f_1(v) f_2(v - u) = ae^{-av} ae^{-a(v-u)}.$$

Wir müssen hier zwei Fälle unterscheiden, nämlich $u < 0$ und $u > 0$.

(a) Sei $u < 0$. Dann ist wegen $y = v - u > 0$ immer $v > u$, so daß v von 0 bis ∞ laufen kann, d.h.

$$g(u) = \int_0^{\infty} ae^{-av} ae^{-a(v-u)} dv = \int_0^{\infty} a^2 e^{au} e^{-2av} dv = \frac{a}{2} e^{au}.$$

(b) Sei $u > 0$. Dann ist $y = v - u > 0$ nur, wenn $v > u$, daher kann v nur von u bis ∞ laufen, d.h.

$$g(u) = \int_u^{\infty} a^2 e^{-a(2v-u)} dv = a^2 e^{au} \int_u^{\infty} e^{-2av} dv = \frac{a}{2} e^{au}.$$

Zusammengeschrieben ergibt sich als

$$g(u) = \frac{a}{2} e^{-a|u|} \quad \text{für } -\infty < u < \infty$$

die sogenannte Laplace-Verteilung. ■

Beispiel 6.5.7. Man wirft zwei Würfel. Berechnen Sie die Dichte von $U = X - Y$, wenn " X = Punktzahl auf dem ersten Würfel" und " Y = Punktzahl auf dem zweiten Würfel".

Lösung: Die Wahrscheinlichkeit einer beliebigen Zahlenkombination auf beiden Würfeln ist $f(x, y) = f(x) f(y) = 1/36$.

Die neue Zufallsvariable kann die Werte von $-5, -4, \dots, 0, \dots, 4, 5$ annehmen, da $1 - 6 = -5, \dots, 6 - 1 = 5$ ist. Wir berechnen

$$\begin{aligned} P(U = -5) &= f(1, 6) = 1/36 \\ P(U = -4) &= f(1, 5) + f(2, 6) = 1/36 + 1/36 = 2/36 \\ P(U = -3) &= f(1, 4) + f(2, 5) + f(3, 6) = 3/36 \\ P(U = -2) &= f(1, 3) + f(2, 4) + f(3, 5) + f(4, 6) = 4/36 \\ P(U = -1) &= f(1, 2) + f(2, 3) + f(3, 4) + f(4, 5) + f(5, 6) = 5/36 \\ P(U = 0) &= f(1, 1) + f(2, 2) + f(3, 3) + f(4, 4) + f(5, 5) + f(6, 6) = 6/36 \\ P(U = 1) &= f(2, 1) + f(3, 2) + f(4, 3) + f(5, 4) + f(6, 5) = 5/36 \\ P(U = 2) &= f(3, 1) + f(4, 2) + f(5, 3) + f(6, 4) = 4/36 \\ P(U = 3) &= f(4, 1) + f(5, 2) + f(6, 3) = 3/36 \\ P(U = 4) &= f(5, 1) + f(6, 2) = 2/36 \\ P(U = 5) &= f(6, 1) = 1/36, \end{aligned}$$

was man in der üblichen Form als

$$g(u) = \begin{cases} 1/36 & \text{für } u = \pm 5 \\ 2/36 & \text{für } u = \pm 4 \\ 3/36 & \text{für } u = \pm 3 \\ 4/36 & \text{für } u = \pm 2 \\ 5/36 & \text{für } u = \pm 1 \\ 6/36 & \text{für } u = 0 \\ 0 & \text{sonst} \end{cases}$$

darstellen kann. ■

Übung 6.5.1. Sei $f(x,y) = 1$ für $0 < x, y < 1$ und $f(x,y) = 0$ sonst. Finden Sie die Dichte von $U = XY$. [Bedenken Sie, daß wegen $X = U/Y \vee \geq u$ sein muß.]

Übung 6.5.2. Man wirft zwei Münzen und betrachtet die Zufallsvariablen "X = Anzahl der Köpfe auf der ersten Münze", "Y = Anzahl der Köpfe auf der zweiten Münze". Berechnen Sie die Dichte von

(a) $U = X + Y$;

(b) $U = X - Y$.

[Bedenken Sie, daß $x = 0, 1, y = 0, 1, f(x,y) = f(x)f(y)$.]

Übung 6.5.3. Sei $f(x,y) = 2e^{-x-y}$ für $0 < x < y < \infty$. Finden Sie die Verteilung von $U = X/Y$.

Aufgaben

6.1. Sei $f(x,y) = k(4 - x - y)$ für $0 < x < 1, 1 < y < 3$ und $f(x,y) = 0$ sonst. Berechnen Sie k .

6.2. Beweisen Sie, daß $F(\infty | y) = 1$ für diskrete und stetige Zufallsvariablen.

6.3. Sei (X,Y) eine zweidimensionale diskrete Zufallsvariable. Beweisen Sie, daß $\mu_{01}' = E(Y)$.

6.4. Beweisen Sie, daß

(a) $\rho_{ax,by} = \pm \rho_{x,y}$;

(b) $\rho_{x+a, y+b} = \rho_{x,y}$.

6.5. Beweisen Sie, daß $E(XY) = E(X)E(Y)$ für eine stetige Zufallsvariable, wenn X und Y unabhängig sind.

6.6. Sei $f(x,y)$ folgendermaßen gegeben:

$f(0,2) = 3/10$

$f(0,3) = 1/10$

$f(1,2) = 2/10$

$f(1,3) = 4/10$.

Berechnen Sie

(a) $f_1(x)$;

(d) $f(x|y = 2)$;

(b) $f_2(y)$;

(e) stellen Sie fest, ob X und Y unabhängig sind.

(c) μ_{10}', μ_{01}' ;

6.7. Sei $f(x,y) = 1/3 (4 - x - y)$ für $0 < x < 1, 1 < y < 3$ und $f(x,y) = 0$ sonst. Berechnen Sie

(a) $F(x,y)$;

(b) $f_1(x); f_2(y)$;

(c) $F_1(x); F_2(y)$;

(d) stellen Sie fest, ob X und Y unabhängig sind.

6.8. Beweisen Sie, daß $\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$.

6.9. Eine zweidimensionale Zufallsvariable ist gegeben als $f(x,y) = k(2 - x)$ für $0 < x < 2, 0 < y < 1$ und $f(x,y) = 0$ sonst. Berechnen Sie

(a) k ;

(g) $f(y|x = 1)$;

(b) $F(x,y)$;

(h) $f(y|1/2 < x < 3/2)$;

(c) $f(x|y)$;

(i) $E(X^T Y^Q)$;

(d) $f(y|x)$;

(j) $\mu_{10}', \mu_{20}', \mu_{01}', \mu_{02}', \mu_{11}'$;

(e) $F(x|y)$;

(k) $\rho_{x,y}$;

(f) $F(y|x)$;

(l) $E(X^2 | Y = 1/2)$.

6.10. Beweisen Sie für eine diskrete Zufallsvariable, daß

$$E(X + Y) = E(X) + E(Y).$$

6.11. Sei $f(x,y) = 1/2$ für $0 < x < 1, 0 < y < 2$ und $f(x,y) = 0$ sonst. Berechnen Sie die Dichte von $U = XY$.

6.12. Sei $f(x,y) = 4x(1 - y)$ für $0 < x, y < 1$ und $f(x,y) = 0$ sonst. Bestimmen Sie die Dichte von $U = XY$.

6.13. Sei $f(x,y) = 6(x - y)$ für $0 < y < x < 1$ und $f(x,y) = 0$ sonst. Finden Sie die Dichte von $U = Y/X$.

6.14. Fortsetzung. Finden Sie die Dichte von $U = X/Y$.

6.15. Sei $f(x,y) = 1/3 (4 - x - y)$ für $0 < x < 1, 1 < y < 3$ und $f(x,y) = 0$ sonst. Finden Sie alle Momente erster und zweiter Ordnung und daraus $\rho_{x,y}$.

6.16. Sei $f(x,y)$ folgendermaßen gegeben:

$f(0,2) = 3/10$

$f(0,3) = 1/10$

$f(1,2) = 2/10$

$f(1,3) = 4/10$.

Finden Sie die Dichte von

(a) $U = XY$;

(b) $U = X - Y$.

6.17. Sei $f(x,y) = 6(x - y)$ für $0 < y < x < 1$ und $f(x,y) = 0$ sonst. Finden Sie

(a) $f(y|x)$;

(c) $E(Y|X)$.

(b) $f(y|x = 1/2)$;

6.18. Sei $f(x,y) = 4xye^{-(x+y)}$ für $x, y > 0$. Berechnen Sie

(a) die momenterzeugende

(c) μ_{01}' ;

Funktion und aus dieser

(d) μ_{11}' ;

(b) μ_{10}' ;

(e) daraus μ_{11} .

6.19. Fortsetzung. Finden Sie die Dichte von $U = \frac{X}{Y}$.

6.20. Sei $f(x,y) = \frac{1}{3}(4-x-y)$ für $0 < x < 1$, $1 < y < 3$ und $f(x,y) = 0$ sonst. Berechnen Sie $P(0 < X < 0,5, 1 < Y < 2)$.

Literatur: Brunk (1965); Fisz (1970); Harris (1966); Kendall-Stuart (1969); Mathai (1967); Mood-Graybill (1963); Papoulis (1965); Wadsworth-Bryan (1960).

Anhang

A1. Fakultäten der Zahlen von 1 bis 10

n	$n!$
1	1
2	2
3	6
4	24
5	120
6	720
7	5040
8	40320
9	362880
10	3628800

A2. Binomialkoeffizienten $\binom{n}{k}$

$n \backslash k$	1	2	3	4	5	6	7	8	9	10
0	1	1	1	1	1	1	1	1	1	1
1	1	2	3	4	5	6	7	8	9	10
2		1	3	6	10	15	21	28	36	45
3			1	4	10	20	35	56	84	120
4				1	5	15	35	70	126	210
5					1	6	21	56	126	252
6						1	7	28	84	210
7							1	8	36	120
8								1	9	45
9									1	10
10										1

B1. Verzeichnis von Summen

$$1. \sum_{k=0}^n \binom{n}{k} = \binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{n} = 2^n$$

$$2. \sum_{k=1}^n k = 1 + 2 + \dots + n = \frac{n(n+1)}{2}$$

$$3. \sum_{k=1}^n k^2 = 1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

$$4. \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots = e^x$$

B2. Verzeichnis einiger Integrale

$$1. \int c \, dx = cx$$

$$2. \int x \, dx = \frac{x^2}{2}$$

$$3. \int x^n \, dx = \frac{x^{n+1}}{n+1}$$

$$4. \int \frac{x}{a+bx} \, dx = \frac{1}{b^2} [a+bx - a \ln(a+bx)]$$

$$5. \int \frac{x}{(a+bx)^n} \, dx = \frac{1}{b^2} \left[\frac{1}{(n-1)(a+bx)^{n-1}} - \frac{1}{(n-2)(a+bx)^{n-2}} \right];$$

$n \neq 1, 2, 3$

$$6. \int \frac{dx}{x(a+bx)} = \frac{2}{\sqrt{ab}} \operatorname{arctg} \sqrt{\frac{bx}{a}}$$

$$7. \int e^{-x} \, dx = -e^{-x}$$

$$8. \int e^{-ax} \, dx = \frac{-e^{-ax}}{a}$$

$$9. \int a e^{-ax} \, dx = -e^{-ax}$$

$$10. \int x e^{-x} \, dx = -x e^{-x} - e^{-x}$$

$$11. \int x e^x \, dx = x e^x - e^x$$

$$12. \int x^2 e^x \, dx = x^2 e^x - 2(x e^x - e^x)$$

$$13. \int x e^{ax} \, dx = \frac{x e^{ax}}{a} - \frac{e^{ax}}{a^2}$$

$$14. \int x^2 e^{ax} \, dx = \frac{x^2 e^{ax}}{a} - \frac{2x e^{ax}}{a^2} + \frac{2 e^{ax}}{a^3}$$

$$15. \int_0^{\infty} x^{a-1} e^{-bx} \, dx = \frac{\Gamma(a)}{b^a} = \frac{(a-1)!}{b^a}$$

$$16. \int_0^{\infty} x^c e^{-ax^b} \, dx = \frac{\left(\frac{c+1}{b} - 1\right)!}{b a^{\frac{c+1}{b}}}$$

$$17. \int \ln x \, dx = x(\ln x - 1)$$

B3. Einige Ableitungsformeln

y	y'
c	0
cx	c
x ⁿ	n x ⁿ⁻¹
e ^x	e ^x
e ^{cx}	c e ^{cx}
a ^x	a ^x ln a
ln x	$\frac{1}{x}$

y	y'
$\log_a x$	$\frac{1}{x \ln a}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$u + v$	$u' + v'$
uv	$uv' + u'v$
$\frac{u}{v}$	$\frac{u'v - uv'}{v^2}$
$g[f(x)]$	$g'[f(x)]f'(x)$

Literatur

- Brunk, H. D.**, *An introduction to mathematical statistics*. Waltham, Blaisdell 1965 (2).
- Cramér, H.**, *Mathematical methods of statistics*. Princeton, University Press 1966.
- Feller, W.**, *An introduction to probability theory and its applications*. Vol. I. New York, Wiley 1962 (2).
- Fichtenholz, G. M.**, *Differential- und Integralrechnung*. B. I. Berlin, VEB Deutscher Verlag der Wissenschaften 1971 (6).
- Fisz, M.**, *Wahrscheinlichkeitsrechnung und mathematische Statistik*. Berlin, VEB Deutscher Verlag der Wissenschaften 1970 (5).
- Geršić, S.**, *Mathematisch-statistische Untersuchungen zur phonetischen Variabilität, am Beispiel von Mundarten aus der Batschka*. Göppingen, Kümmerle 1971.
- Gnedenko, B. W.**, *Lehrbuch der Wahrscheinlichkeitsrechnung*. Berlin, Akademie-Verlag 1968 (5).
- Gradshteyn, I. S., Ryzhik, I. M.**, *Tables of integrals, series, and products*. New York - London, Academic Press 1973 (7).
- Greenberg, J. H.**, *The measurement of linguistic diversity*. Language 32, 1956, 109-115.
- Hald, A.**, *Statistical theory with engineering applications*. New York, Wiley 1967.
- Harris, B.**, *Theory of probability*. Reading, Addison-Wesley, 1966.
- Herdan, G.**, *The advanced theory of language as choice and chance*. Berlin - Heidelberg - New York, Springer 1966.
- Hogben, L.**, *Zahl und Zufall*. München, Oldenbourg 1956.
- Job, M.**, *Persönliche Mitteilung*. 1980.
- Kapur, J. N., Saxena, H. C.**, *Mathematical statistics*. Delhi, Chand 1970 (5).
- Kendall, M. G., Stuart, A.**, *The advanced theory of statistics*. Vol. I. London, Griffin 1969 (3).

- Kreyszig, E., *Statistische Methoden und ihre Anwendungen*. Göttingen, Vandenhoeck-Rupprecht 1968 (3).
- Lehfeldt, W., *Die Verteilung der Phonemzahl in den natürlichen Sprachen*. *Phonetica* 31, 1975, 274-287.
- Ludvíková, M., *Quantitative syllable analysis of words in Czech*. *Prague Studies in Mathematical Linguistics* 3, 1972, 27-34.
- Mathai, A. M., *Introduction to statistical mathematics*. Delhi, Chand 1967.
- Mood, A. M., Graybill, F. A., *Introduction to the theory of statistics*. New York, McGraw-Hill 1963 (2).
- Mosteller, F., Rourke, R. E. K., Thomas, G. B. Jr., *Probability with statistical applications*. Reading, Addison-Wesley 1961.
- Netto, E., *Lehrbuch der Combinatorik*. New York, Chelsea 1927 (2).
- Neyman, J., *First course in probability and statistics*. New York, Holt, Rinehart, Winston 1950.
- Pappoulis, A., *Probability, random variables, and stochastic processes*. New York McGraw-Hill 1965.
- Ventcel', Je. S., *Teorija verojatnostej*. Moskva, Nauka 1964.
- Ventcel', Je. S., Ovčarov, L. A., *Teorija verojatnostej*. Moskva, Nauka 1973 (2).
- Wadsworth, G. P., Bryan, J. G., *Introduction to probability and random variables*. New York, McGraw-Hill 1960.
- Weber, E., *Grundriß der biologischen Statistik*. Jena, Fischer 1961 (4).
- Whitworth, W. A., *Choice and chance*. New York-London, Hafner 1965 (5).

Lösungen der Übungen und Aufgaben

Kapitel 1

Übungen

1.1.1.	x_i	2	3	4	5	6	7	8	9	10	11	12	Σ
	f_i	3	16	16	8	5	4	0	2	1	3	1	59

1.2.1. (a) 9 ;

(d) 14 ;

(b) 7 ;

(e) $\sum_{i=1}^n x_i - nk$;

(c) 0 ;

(f) 5c.

1.2.2. (a) $\sum_{i=1}^n ax_i = ax_1 + ax_2 + \dots + ax_n = a(x_1 + x_2 + \dots + x_n) = a \sum_{i=1}^n x_i$;

(b) $\sum_{i=1}^n (ax_i - by_i) = \sum_{i=1}^n ax_i - \sum_{i=1}^n by_i = a \sum_{i=1}^n x_i - b \sum_{i=1}^n y_i$ laut (a) ;

(c) 10a.

1.2.3. $\bar{x} = 289/59 = 4,90$.

1.2.4. 30,3167.

1.2.5. 27,03.

1.2.6. $\frac{N_1 \bar{x}_1 + N_2 \bar{x}_2}{N_1 + N_2} = 5,56$.

1.3.1. $S^2 = \frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^2 f_i = \frac{1}{N} [\sum x_i^2 f_i - 2\bar{x} \sum x_i f_i + \bar{x}^2 \sum f_i] =$

$= \frac{1}{N} [\sum x_i^2 f_i - 2\bar{x}N\bar{x} + \bar{x}^2 N] = \frac{1}{N} [\sum x_i^2 f_i - N\bar{x}^2]$. Oder man multipliziere den zweiten Ausdruck in der Klammer in (1.18) mit N/N und das Resultat folgt.

$$1.3.2. \left\{ \frac{1}{6} \left[5164,36 - \frac{(189,2)^2}{7} \right] \right\}^{1/2} = \sqrt{8,4257} = 2,90.$$

$$\begin{aligned} 1.3.3. NS^2 &= \sum (x_i - \bar{x})^2 = \sum (x_i - \bar{x}_c + \bar{x}_c - \bar{x})^2 = \\ &= \sum (x_i - \bar{x}_c)^2 + 2(\bar{x}_c - \bar{x}) \sum (x_i - \bar{x}_c) + N(\bar{x}_c - \bar{x})^2 \\ &= [\text{wegen } \sum (x_i - \bar{x}_c) = N(\bar{x} - \bar{x}_c)] \\ &= \sum (x_i - \bar{x}_c)^2 + 2N(\bar{x}_c - \bar{x})(\bar{x} - \bar{x}_c) + N(\bar{x}_c - \bar{x})^2 \\ &= \sum (x_i - \bar{x}_c)^2 - 2N(\bar{x}_c - \bar{x})^2 + N(\bar{x}_c - \bar{x})^2 \\ &= \sum (x_i - \bar{x}_c)^2 - N(\bar{x}_c - \bar{x})^2 \\ \Rightarrow \sum (x_i - \bar{x}_c)^2 &= NS^2 + N(\bar{x}_c - \bar{x})^2. \text{ Durch Abziehen des falschen Quadrats } \\ (x_f - \bar{x}_c)^2 \text{ von beiden Seiten und Dividieren durch } N-1 \text{ bekommt man} \\ S_c^2 &= \frac{\sum (x_i - \bar{x}_c)^2 - (x_f - \bar{x}_c)^2}{N-1} = \frac{NS^2 + N(\bar{x}_c - \bar{x})^2 - (x_f - \bar{x}_c)^2}{N-1}. \end{aligned}$$

Setzt man auf der rechten Seite der Gleichung $\bar{x}_c = (N\bar{x} - x_f)/(N-1)$ ein, so folgt das Resultat nach algebraischer Umordnung.

$$1.3.4. 0,6875$$

Aufgaben

$$1.2. (a) \sum x_i^2 - 3 \sum x_i; \quad (b) \sum x_i^2 - 2 \sum x_i + n.$$

$$1.3. s^2 = 5,8860; \quad s = 2,43.$$

$$1.4. (a) \frac{n(n+1)(2n-8)}{6}; \quad (b) \frac{n(n-1)(2n-1)}{6}.$$

$$1.5. \bar{x} = 2,05; \quad s = 0,94.$$

$$1.7. 2,97.$$

$$1.8. \bar{x} = \frac{n+1}{2}; \quad S^2 = \frac{n^2-1}{12}.$$

$$1.9. \bar{x} = 29,6; \quad s = 4,02.$$

$$1.11. \bar{x} = 6,92.$$

$$1.13. s^2 = 89,18.$$

$$1.14. \bar{x} = 1 + \frac{dn}{2}; \quad S^2 = \frac{d^2 n(n+2)}{12}.$$

$$1.15. S^2 = 3,27.$$

1.16. Für die zusammengesetzte Stichprobe ist

$$\begin{aligned} S^2 &= \frac{1}{N_1 + N_2} \sum_{i=1}^{N_1 + N_2} (x_i - \bar{x})^2 = \frac{1}{N_1 + N_2} \left[\sum_{i=1}^{N_1} (x_{i1} - \bar{x})^2 + \sum_{i=N_1+1}^{N_1+N_2} (x_{i2} - \bar{x})^2 \right] = \\ &= [\bar{x}_1 \text{ und } \bar{x}_2 \text{ einschieben}] \\ &= \frac{1}{N_1 + N_2} \left[\sum_{i=1}^{N_1} (x_{i1} - \bar{x}_1 + \bar{x}_1 - \bar{x})^2 + \sum_{i=N_1+1}^{N_1+N_2} (x_{i2} - \bar{x}_2 + \bar{x}_2 - \bar{x})^2 \right] \\ &= \frac{1}{N_1 + N_2} [\sum (x_{i1} - \bar{x}_1)^2 + 2(\bar{x}_1 - \bar{x}) \sum (x_{i1} - \bar{x}_1) + N_1(\bar{x}_1 - \bar{x})^2 + \\ &\quad + \sum (x_{i2} - \bar{x}_2)^2 + 2(\bar{x}_2 - \bar{x}) \sum (x_{i2} - \bar{x}_2) + N_2(\bar{x}_2 - \bar{x})^2] = \\ &= [\text{da } \sum (x_{i1} - \bar{x}_1) = \sum (x_{i2} - \bar{x}_2) = 0] \\ &= \frac{1}{N_1 + N_2} [N_1 S_1^2 + N_1(\bar{x}_1 - \bar{x})^2 + N_2 S_2^2 + N_2(\bar{x}_2 - \bar{x})^2] \\ &= \frac{N_1 S_1^2 + N_2 S_2^2}{N_1 + N_2} + \frac{N_1(\bar{x}_1 - \bar{x})^2 + N_2(\bar{x}_2 - \bar{x})^2}{N_1 + N_2} \\ &= [\text{nach Einsetzen von } \bar{x} = \frac{N_1 \bar{x}_1 + N_2 \bar{x}_2}{N_1 + N_2} \text{ und Ordnen}] \\ &= \frac{N_1 S_1^2 + N_2 S_2^2}{N_1 + N_2} + \frac{N_1 N_2 (\bar{x}_1 - \bar{x}_2)^2}{(N_1 + N_2)^2}. \end{aligned}$$

$$1.17. \bar{x} = 3,88; \quad S^2 = 6,44.$$

$$\begin{aligned} 1.18. S_c^2 &= \frac{1}{N} \left[\sum x_i^2 - x_f^2 + x_c^2 - \frac{(\sum x_i - x_f + x_c)^2}{N} \right] \\ &= \frac{1}{N} \left[\sum x_i^2 - x_f^2 + x_c^2 - \frac{(\sum x_i)^2}{N} - \frac{2(x_c - x_f) \sum x_i}{N} - \frac{(x_c - x_f)^2}{N} \right] \\ &= \frac{1}{N} \left(\sum x_i^2 - \frac{\sum x_i^2}{N} \right) + \frac{1}{N} \left[x_c^2 - x_f^2 - \frac{(x_c - x_f)^2 + 2(x_c - x_f) \sum x_i}{N} \right] \\ &= S^2 + \frac{1}{N} (x_c - x_f) \left(x_c + x_f - \frac{x_c - x_f + 2 \sum x_i}{N} \right). \end{aligned}$$

$$1.19. S^2 = 4; \quad S_c^2 = 3,8889.$$

$$1.20. S^2 = 4; \quad S_c^2 = 5,6327.$$

Kapitel 2

Übungen

- 2.1.1. 5040.
 2.1.2. $5! = 120$.
 2.1.3. $8! = 40320$.
 2.1.4. $4! = 24$.
 2.2.1. $3^3 = 27$.
 2.2.2. 7^7 .
 2.2.3. 8^8 .
 2.2.4. 40.
 2.2.5. 560.
 2.3.1. $\frac{6!}{3!2!} = 60$.
 2.3.2. $\frac{5!}{2!^2} = 30$.
 2.3.3. (a) 6;
 (b) 12;
 (c) 18.
 2.3.4. 9979200.
 2.3.5. 10.
 2.4.1. 210.
 2.4.2. 13800.
 2.4.3. 9240.
 2.5.1. 49.

Aufgaben

- 2.1. $6!$; 6^6 ; 30; 36; 15 21; 60; 90; 6; 864.
 2.2. $\frac{9!}{2!} = 181440$.
 2.4. $\binom{49}{6} = 13983816$.
 2.5. 4900.
 2.7. (a) 120; (b) 625.
 2.8. $C_3^5 = 21$.
 2.10. 5040.
 2.11. $\binom{8}{4} = 70$.
 2.12. 35.
 2.13. (a) 41; (b) 28.

- 2.5.2. $26^3 = 17576$.
 2.5.3. $26^2 = 676$.
 2.6.1. 5 ; $\frac{n(n-1)}{2}$; 1.
 2.6.2. $\binom{14}{3} - \binom{11}{3} = 199$.
 2.6.3. (a) 210;
 (b) $\binom{8}{2} 4! = 672$.
 2.7.1. 28; 15.
 2.7.2. 35.
 2.7.3. 28.
 2.8.1. 84.
 2.8.2. $a^6 + 6a^5b + 15a^4b^2 + 20a^3b^3 + 15a^2b^4 + 6ab^5 + b^6$.
 2.9.1. $\binom{4+5-1}{5} = 56$.
 2.9.2. $\frac{7!}{4!1!2!} = 105$.
 2.9.3. $\frac{20!}{5!^4}$.
 2.9.4. $\frac{15!}{10!3!2!} = 30030$.

- 2.15. (a) 5040;
 (b) 720;
 (c) 4320;
 (d) 3600;
 (e) 210;
 (f) 35;
 (g) 15;
 (h) 343;
 (i) 49;

(j) Anzahl aller Möglichkeiten: 7^3 ;

Wörter mit keinem I: 6^3 , Wörter mit mindestens einem I:
 $7^3 - 6^3 = 127$; auch direkte Berechnung möglich;

- (k) 84;
 (l) 21;
 (m) 56;
 (n) $C_7^3 - C_6^3 = 84 - 56 = 28$ vgl.

Aufgabe 2.14.

- 2.17. Anzahl der Möglichkeiten: $L_1 = 6$, $L_2 = 6$ ($6 + 6$) ($23 + 23$) ($6 + 6$) = 312,
 $L_3 = 6$ (6) (23) ($3 + 23$) (23) (6) (3) = 12006;
 Anzahl der realisierten: $L_1 = 6$, $L_2 = 74$, $L_3 = 183$.
 Daraus $P_1 = \frac{6}{6} = 1$, $P_2 = \frac{74}{312} = 0,2372$, $P_3 = \frac{183}{12006} = 0,0152$; die Möglichkeiten KK, KKK, VVV, sind nicht gerechnet worden.
- 2.19. (a) $C_3^1 C_2^1 C_4^1 = 24$; (d) $C_3^3 = 1$;
 (b) $24 (3!) = 144$; (e) $V_3^3 = 27$;
 (c) $C_3^3 = 10$; (f) $P_3 = V_3^3 = 3! = 6$.
- 2.21. (a) Wenn $k = n$;
 (b) wenn $k = n$;
 (c) wenn $r = 1$, $k_1 = n - k$;
 (d) wenn $r = 2$, $k_1 = k$, $k_2 = n - k$;
 (e) wenn $k_i = 1$.
- 2.22. (a) $C_5^2 = 10$; (b) $C_9^3 = 84$.
- 2.23. (a) Man kann folgendermaßen verfahren: Aus 3 Silben wählt man 2 für Betonung auf $\binom{3}{2}$ Weisen. Da die Akzente unterschiedlich sind, muß man sie noch permutieren, d.h. $\binom{3}{2} 2! = V_3^2 = 6$. Daher ist für dreisilbige Wörter V_3^2 , für viersilbige V_4^2 usw., allgemein für n -silbige V_n^2 ;
- (b) wenn die Akzente nicht nebeneinander stehen dürfen, dann setzt man sie zunächst nebeneinander, die beiden betonten Silben betrachtet man als eine Silbe und berechnet die Zahl der Möglichkeiten. Sie beträgt offensichtlich C_2^1 , weil man aus 2 Silben eine auf C_2^1 Weisen wählen kann. Da die Akzente unterschiedlich sind, muß man die C_2^1 Möglichkeiten noch $2!$ -mal permutieren (der Hauptakzent darf vor oder hinter dem Nebenzakzent stehen), daher $2C_2^1$. Diese Zahl zieht man von V_3^2 ab, so daß man $V_3^2 - 2C_2^1 = 6 - 4 = 2$

bekommt, was leicht nachzuprüfen ist. Allgemein bekommt man für ein n -silbiges Wort dann $V_n^2 - 2C_{n-1}^1$;

(c) Es ist $V_n^2 - 2C_{n-1}^1 = \frac{n!}{(n-2)!} - \frac{2(n-1)!}{1!(n-2)!} = \frac{(n-1)!}{(n-3)!} = V_{n-1}^2$; daher ist die

Zahl der Möglichkeiten, zwei unterschiedliche Akzente auf ein n -silbiges Wort zu stellen, gleich V_{n-1}^2 . Daraus ergibt sich

für $n = 4$: $V_3^2 = 6$; für $n = 5$: $V_4^2 = 12$; für $n = 6$: $V_5^2 = 20$;

für $n = 7$: $V_6^2 = 30$; für $n = 8$: $V_7^2 = 42$; für $n = 9$: $V_8^2 = 56$ usw.

2.24. Vor der ersten schweren Silbe gibt es vier Möglichkeiten, leichte Silben zu plazieren (nämlich keine, eine, zwei oder drei), hinten ebenfalls und dazwischen jeweils 3 (1, 2 oder 3), daher $4(3)3(3)4 = 432$.

2.26.
$$\sum_{k=0}^n k \binom{n}{k} = \sum_{k=1}^n k \frac{n!}{k!(n-k)!}$$

$$= [\text{wenn } k=0, \text{ dann ist der ganze Ausdruck } 0]$$

$$= \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!}$$

$= [\text{setzen Sie } n-1 = t \text{ und } k-1 = s, \text{ woraus}]$

$$= n \sum_{s=0}^{t+1} \frac{t!}{s!(t-s)!} = n \sum_{s=0}^t \binom{t}{s} = n2^t = n2^{n-1}$$

$[s \text{ läuft nur bis } t, \text{ denn bei } s = t+1 \text{ ist nach (1.15)} \binom{t}{s} = 0].$

2.27. Aus $2 \cdot 4 \cdot 6 \cdot \dots \cdot 2n$ 2^n ausklammern, so daß $2^n(1 \cdot 2 \cdot 3 \cdot \dots \cdot n) = 2^n n!$

2.28. $1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1) = 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1) \cdot \frac{2 \cdot 4 \cdot 6 \cdot \dots \cdot 2n}{2 \cdot 4 \cdot 6 \cdot \dots \cdot 2n} =$

$$= \frac{(2n)!}{2 \cdot 4 \cdot 6 \cdot \dots \cdot 2n} = \frac{(2n)!}{2^n n!} [\text{laut Aufgabe 2.27}].$$

2.30. $n! = \frac{n! n^n}{n^n}$

$$= \frac{n^n \cdot n(n-1)(n-2)(n-3)\dots 3 \cdot 2 \cdot 1}{n \cdot n \cdot n \cdot \dots \cdot n}$$

$$= n^n \cdot 1 \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{n-1}{n}\right) = n^n \prod_{k=0}^{n-1} \left(1 - \frac{k}{n}\right).$$

2.31. $(2n)! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot 2n = 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1) \cdot 2 \cdot 4 \cdot 6 \cdot \dots \cdot 2n$; wir
können schreiben $1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1) = \prod_{k=1}^n (2k-1)$ und $2 \cdot 4 \cdot 6 \cdot \dots \cdot 2n =$

$$= 2^n n! = 2^n \prod_{m=1}^n m, \text{ woraus das Resultat folgt.}$$

2.32.
$$\frac{(n! 2^n)^2}{(2n)!} = \frac{(n!)^2 2^{2n}}{(2n)!} \cdot \frac{n^{2n}}{n^{2n}} = \frac{(n!)^2}{n^{2n}} \cdot \frac{1}{\frac{2n(2n-1)(2n-2)\dots 3 \cdot 2 \cdot 1}{2n(2n)2n \cdot \dots \cdot 2n}} =$$

$$= \left(\frac{n!}{n^n}\right)^2 \frac{1}{1(1-\frac{1}{2})(1-\frac{2}{2})\dots(1-\frac{2n-1}{2n})}$$

$$= \left(\frac{n!}{n^n}\right)^2 \frac{1}{\prod_{k=0}^{2n-1} (1-\frac{k}{2n})}.$$

2.33. $n \prod_{k=2}^n (1 - \frac{1}{k}) = n \prod_{k=2}^n \frac{k-1}{k} = n \left[\frac{(n-1)!}{n!} \right] = 1$

Kapitel 3

Übungen

3.1.1. $2/26!$.

3.1.2. $2/26^{26}$.

3.1.3. $N = 5!/2!2!1!1!; N_A = 3!; P(A) = 3!2!2!5! = 1/5$.

3.1.4. $N = \binom{5+3-1}{3} = 35; N_A = \binom{5+1-1}{1} = 5; P(A) = 5/35 = 1/7$.

3.2.1. (a) $\overline{AB}$; (d) $\overline{\overline{A}\overline{B}}$; $\overline{\overline{A}\overline{B}}$;
(b) $A+B$; (e) $\overline{AB} = \overline{A+B}$.
(c) $\overline{AB}$;

3.2.2. $P(\overline{A}) = P(W) + P(\overline{S}) = 0,17 + 0,13 = 0,30$
 $\Rightarrow P(A) = 1 - P(\overline{A}) = 0,70$.

3.2.3. $\overline{AB - ABC} = \overline{AB(\overline{ABC})} = \overline{AB(\overline{A+B+C})} =$
 $= \overline{AAB + AB\overline{B} + AB\overline{C}} = \overline{ABC}$ weil $\overline{AA} = \emptyset$ und $A\emptyset = \emptyset$.

3.3.1. (a) $p_A^2 + p_B^2$; (b) $1 - p_A^2 - p_B^2 = 1 - \sum_i p_i^2 (i = A, B)$.

3.3.2. $1 - (1-p)^n = 1 - P(\overline{A})$.

$$3.3.3. P(A) = P(A) + P(B) - P(AB) = \frac{5}{20} + \frac{15}{20} - \frac{75}{400} = \frac{13}{16}.$$

$$3.3.4. \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} = 0,3750.$$

$$3.3.5. P_{(5)} = 0,0028.$$

$$3.4.1. P(V_1 V_2 V_3) = P(V_1) P(V_2|V_1) P(V_3|V_1 V_2) = \frac{5}{25} \cdot \frac{4}{24} \cdot \frac{3}{23} = 0,0043.$$

3.4.2. Unabhängig.

$$3.4.3. P(S|A) = \frac{P(SA)}{P(A)} = \frac{P(A)}{P(A)} = 1.$$

$$3.4.4. P(A + B|C) = \frac{P[(A+B)C]}{P(C)} = \frac{P(AC + BC)}{P(C)} = \frac{P(AC)}{P(C)} + \frac{P(BC)}{P(C)} = P(A|C) + P(B|C).$$

$$3.5.1. (a) 0,1414; \quad (b) 0,3030$$

Aufgaben

$$3.1. (a) 2^2_{26^2} = 0,005917; \quad (b) \frac{1}{26^2} = 0,001479.$$

$$3.2. (a) \frac{1}{10}; \quad (d) \frac{3}{5};$$

$$(b) \frac{3}{5}; \quad (e) \frac{1}{10};$$

$$(c) \frac{1}{5}; \quad (f) \frac{1}{5}.$$

$$3.3. P(A - BC) + P(A \bar{B} \bar{C}) = P(\overline{ABC}) + P(A \bar{B} \bar{C}) = P[A(\bar{B} + \bar{C})] + P(A \bar{B} \bar{C}) = P(A\bar{B} + A\bar{C}) + P(A \bar{B} \bar{C}) = P(A\bar{B}) + P(A\bar{C}) - P(A \bar{B} \bar{C}) + P(A \bar{B} \bar{C}) = P(A\bar{B}) + P(A\bar{C}).$$

$$3.4. \frac{3}{28}.$$

$$3.5. N = 8 \cdot 10 = 80; \quad N_A = 7; \quad P(A) = \frac{7}{80}.$$

$$3.6. N = 8^5; \quad N_A = 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4; \quad P(A) = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{8^5} = 0,205078.$$

$$3.7. N = n^k = V_n^k; \quad N_A = n(n-1) \dots (n-k+1) = \frac{n!}{(n-k)!} = V_n^k; \quad P(A) = \frac{V_n^k}{V_n^k}.$$

$$3.8. \frac{150}{1000} \cdot \frac{60}{1000} = \frac{9}{1000}.$$

$$3.9. 10 \text{ Vokale mit der Wahrscheinlichkeit } \left(\frac{4}{7}\right)^{10}; 9 \text{ Vokale mit der Wahrscheinlichkeit } \left(\frac{4}{7}\right)^9 \left(\frac{3}{7}\right) \binom{10}{1}; 8 \text{ Vokale mit der Wahrscheinlichkeit } \left(\frac{4}{7}\right)^8 \left(\frac{3}{7}\right)^2 \binom{10}{2};$$

$$P(A_8 + A_9 + A_{10}) = \frac{(4^{10} + 4^9 \cdot 3 \cdot 10 + 4^8 \cdot 3^2 \cdot 45)}{7^{10}} = \frac{4^8 \cdot 541}{7^{10}}.$$

$$3.10. N = \text{Wörter ohne L} + \text{Wörter mit einem L} + \text{Wörter mit zwei L} =$$

$$= \binom{4}{3} 3! + \binom{4}{2} 3! + \binom{4}{1} 3!_{2!} = 72;$$

$$(a) N_A = \binom{4}{1} 3!_{2!} = 12; \quad P(A) = \frac{12}{72} = \frac{1}{6};$$

$$(b) N_A = \binom{4}{1} = 4; \quad P(A) = \frac{4}{72} = \frac{1}{18};$$

$$(c) N_A = 4 \binom{2}{1} = 8; \quad P(A) = \frac{8}{72} = \frac{1}{9};$$

$$(d) N = 5^3 = 125; \quad N_A = \binom{4}{1} 3!_{2!} = 12; \quad P(A) = \frac{12}{125};$$

$$(e) N_A = 4 \cdot 3!_{2!} + 6 \cdot 3! =$$

= bei zwei gleichen Buchst. + bei zwei verschiedenen Buchst. außer L =

$$= 12 + 36 = 48;$$

$$P(A) = \frac{48}{125};$$

$$(f) N_A = 2^3 = 8; \quad P(A) = \frac{8}{125};$$

$$(g) N_A = 3^3 = 27; \quad P(A) = \frac{27}{125};$$

(h) die Berechnung geht auch direkt, es ist aber anschaulicher zu unterscheiden, ob 2 verschiedene oder 2 gleiche Konsonanten im Wort stehen: bei ungleichen Konsonanten wählt man den Vokal auf $\binom{2}{1}$ Weisen, die Konsonanten auf $\binom{3}{2}$ Weisen und permutiert alle auf $3!$ Weisen, d.h. $\left[\binom{2}{1} \binom{3}{2}\right] 3! = 36$; bei gleichen Konsonanten wählt man den Vokal auf $\binom{2}{1}$ Weisen, den ersten Konsonanten auf $\binom{3}{1}$ Weisen, den zweiten auf 1 Weise, da er identisch sein muß und man permutiert alle auf $3!_{2!}$ Weisen, d.h. $\left[\binom{2}{1} \binom{3}{1}\right] 3!_{2!} = 18$, daher

$$N_A = 36 + 18 = 54; \quad P(A) = \frac{54}{125};$$

$$(i) N_A = (1 \cdot 3) 3! + \left[\binom{2}{1} 1 (3)\right] 3!_{2!} = 18 + 18 = 36; \quad P(A) = \frac{36}{125};$$

$$(j) N_A = 8; \quad P(A) = \frac{8}{125}.$$

$$3.11. (a) p_1 p_2;$$

$$(b) (1 - p_1)(1 - p_2);$$

$$(c) (1 - p_1)p_2;$$

$$(d) \text{erste geschlossen oder zweite geschlossen oder beide geschlossen} = (1 - p_1)p_2 + p_1(1 - p_2) + (1 - p_1)(1 - p_2) = 1 - p_1 p_2, \text{ was direkt aus (a) folgt.}$$

$$3.12. N = \binom{8}{3} = 56;$$

$$(a) N_A = 1; \quad P(A) = \frac{1}{56};$$

$$(b) N_A = \binom{5}{3} = 10; \quad P(A) = \frac{5}{28};$$

$$(c) N_A = \binom{3}{1} \binom{5}{2} = 30; \quad P(A) = \frac{15}{28}.$$

$$3.13. (a) P_1 = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!} - \frac{1}{6!} + \frac{1}{7!} - \frac{1}{8!} = 0,6321 \approx 1 - e^{-1};$$

$$(b) P_{(4)} = \frac{1}{4!} (1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}) = 0,015625.$$

$$3.14. P(\bar{A}) = (1 - p)^n (1 - p)^n \dots (1 - p)^n [k\text{-mal}] = (1 - p)^{nk} \text{ woraus } P(A) = 1 - (1 - p)^{nk};$$

(b) Vorkommen in einem Text der Länge n : $1 - (1 - p)^n$ woraus in allen k Texten

$$P(A) = [1 - (1 - p)^n]^k.$$

3.15. (a) $N = \frac{8!}{2!^4}$; $N_A = 4!$; $P(A) = \frac{4! \cdot 2!^4}{8!} = \frac{1}{105}$;

(b) $S_1 = \binom{4}{1} \frac{\frac{7!}{2! \cdot 2! \cdot 2!}}{N}$, $S_2 = \binom{4}{2} \frac{\frac{6!}{2! \cdot 2!}}{N}$, $S_3 = \binom{4}{3} \frac{\frac{5!}{2!}}{N}$, $S_4 = \frac{4!}{N}$;

$$P(a) = S_1 - S_2 + S_3 - S_4 = 1 - \frac{3}{7} + \frac{2}{21} - \frac{1}{105} = \frac{69}{105}.$$

3.16. (a) $\overline{A} \overline{B} \overline{C} + \overline{A} \overline{B} C + \overline{A} B \overline{C}$; (f) $A + B + C$;

(b) $\overline{A} B \overline{C} + \overline{A} B C + \overline{A} B \overline{C}$; (g) $\overline{A} B \overline{C}$;

(c) $\overline{A} B C$; (h) $\overline{A} B + A C + B C$;

(d) $\overline{A} \overline{B} \overline{C}$; (i) $\overline{A} B \overline{C} = \overline{A} + \overline{B} + \overline{C}$.

(e) $\overline{A} \overline{B} \overline{C} = A + B + C$;

3.17. (a) $\frac{5}{20} \cdot \frac{15}{19} = \frac{15}{76}$;

(b) $P(N_1) P(K_2|N_1) + P(K_1) P(K_2|K_1) = \frac{5}{20} \cdot \frac{15}{19} + \frac{15}{20} \cdot \frac{14}{19} = \frac{3}{4}$;

(c) $P(N_1) P(N_2|N_1) + P(K_1) P(N_2|K_1) = \frac{5}{20} \cdot \frac{4}{19} + \frac{15}{20} \cdot \frac{5}{19} = \frac{1}{4}$.

3.18. $N = 20 \cdot 5 \cdot 20 \cdot 5 \cdot 20$; $N_A = 20 \cdot 5 \cdot 19 \cdot 4 \cdot 18$;

$$P(A) = \frac{(20 \cdot 5 \cdot 19 \cdot 4 \cdot 18)}{(20 \cdot 5 \cdot 20 \cdot 5 \cdot 20)} = \frac{171}{250}.$$

3.19. $P(\overline{A} \overline{B}) = P(A + B) = 1 - P(A) - P(B) + P(AB) =$
 $= 1 - P(A) - P(B) + P(A) P(B) = [1 - P(A)] [1 - P(B)] = P(\overline{A}) P(\overline{B}).$

3.20. Man teilt die 5 Gruppen à 2 auf $\binom{10}{2} \binom{8}{2} \binom{6}{2} \binom{4}{2} \binom{2}{2} = \frac{10!}{2!^5}$ Weisen und kann sie den Studenten auf 5! Weisen zuordnen, $N = \frac{10! 5!}{2!^5}$; hat A seine 2 Fragen, so kann man den Rest auf $N_A = \frac{8! 4!}{2!^4}$ den anderen zuordnen, daher $P(A) = \frac{8! 4! 2!}{10! 5!} = \frac{1}{225}$.

3.21. $N = \frac{10!}{3! 1!^7}$; $N_A = 2$; $P(A) = 2 \cdot \frac{3!}{10!}$.

3.22. $N = V_{12}^3 = 1320$;

(a) $N_A = 3 \cdot 5 \cdot 4 \cdot 3! = 360$; $P(A) = \frac{3!}{11!}$;

(b) $N_A = 3 \cdot 5 \cdot 4 = 60$; $P(A) = \frac{1}{22}$.

3.23. $N = V + 3V = 4V$; $N_A = 3V$; $P(A) = \frac{3V}{4V} = \frac{3}{4}$.

3.24. $P(\overline{A}_1 + \overline{A}_2 + \overline{A}_3) = P(\overline{A}_1 \overline{A}_2 \overline{A}_3) = 1 - P(A_1 A_2 A_3) =$
 $= 1 - P(A_1) P(A_2|A_1) P(A_3|A_1 A_2).$

3.25. (a) $P(\overline{A} \overline{B}) = [(1 - p_1)(1 - p_2)]^n$;

(b) $P(\overline{A} \overline{B}) = [1 - (1 - p_1)]^n (1 - p_2)^n$;

(c) $P(\overline{A} \overline{B}) = (1 - p_1)^n [1 - (1 - p_2)^n]$;

(d) $P(AB) = [1 - (1 - p_1)^n] [1 - (1 - p_2)^n]$;

(e) $P(A + B) = 1 - P(\overline{A} \overline{B}) = 1 - (1 - p_1)^n (1 - p_2)^n$;

(f) $P(\overline{A} \overline{B}) = P(\overline{A} + \overline{B}) = 1 - P(AB) = 1 - [1 - (1 - p_1)^n] [1 - (1 - p_2)^n]$;

(g) $P(\overline{A} \overline{B}) + P(\overline{A} \overline{B}) = (1 - p_1)^n + (1 - p_2)^n - 1(1 - p_1)^n (1 - p_2)^n$

Kapitel 4

Übungen

4.1.1. (a) ja;

(b) nein.

4.1.2. (a) $P(x_1 \leq X \leq x_2) = P(X \leq x_2) - P(X < x_1) = F(x_2) - F(x_1 - 1)$;

(b) $P(x_1 < X < x_2) = P(X < x_2) - P(X \leq x_1) = F(x_1 - 1) - F(x_1)$;

(c) $P(x_1 \leq X < x_2) = P(X < x_2) - P(X < x_1) = F(x_2 - 1) - F(x_1 - 1).$

4.1.3. (a) ja;

$$(b) F(x) = \begin{cases} 0 & \text{für } -\infty < x < -1 \\ \frac{1}{12} & \text{für } -1 \leq x < 0 \\ \frac{1}{3} & \text{für } 0 \leq x < 1 \\ \frac{1}{2} & \text{für } 1 \leq x < 2 \\ \frac{7}{12} & \text{für } 2 \leq x < 3 \\ \frac{11}{12} & \text{für } 3 \leq x < 4 \\ 1 & \text{für } 4 \leq x < \infty \end{cases}$$

(c) $P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1) = 1 - F(1) = 1 - \frac{1}{2} = \frac{1}{2}$;

(d) $F(2) = \frac{7}{12}$;

(e) $P(1 < X \leq 3) = F(3) - F(1) = \frac{11}{12} - \frac{1}{2} = \frac{5}{12}$;

(f) $P(1 \leq X < 3) = F(3 - 1) - F(1 - 1) = F(2) - F(0) = \frac{7}{12} - \frac{1}{3} = \frac{1}{4}$;

(g) $F(2) - F(-1) = \frac{7}{12} - \frac{1}{12} = \frac{1}{2}$;

(h) $P(X < 3) = F(2) = \frac{7}{12}$;

(i) $P(|X| \leq 1) = P(-1 \leq X \leq 1) = F(1) - F(-2) = F(1) = \frac{1}{2}$.

4.2.1. $k = \frac{1}{6}$.

4.2.2. (a) ja;

$$(b) F(x) = \begin{cases} 0 & -\infty < x < 0 \\ (\frac{2}{3}) x^3 & 0 \leq x < 1 \\ \frac{1}{9}(5 + x^2) & 1 \leq x < 2 \\ 1 & 2 \leq x < \infty \end{cases}$$

(c) $F(1,5) = \frac{1}{9}(5 + 1,5^2) = 0,8056$;

(d) $F(0,4) = \frac{2(0,4)^3}{3} = 0,0427$;

(e) 0,1233 ;

(f) $0,7733 - 0,0427 = 0,7306$.

4.2.3. Ja, $f(x) \geq 0$ und $\int_0^{\infty} a e^{-ax} dx = 1$.

4.3.1. $P(X > 0) = 1 - P(X \leq 0) = 1 - P(0) = 1 - \binom{n}{0} p^0 q^{n-0} = 1 - q^n$, woraus

$$f(x|X > 0) = \frac{\binom{n}{x} p^x q^{n-x}}{1 - q^n} \quad \text{für } x = 1, 2, \dots, n.$$

4.3.2. $F(0,5) = \int_0^{0,5} dx = 0,5$, woraus $f(x|X \leq 0,5) = \frac{1}{0,5} = 2$ für $0 < x \leq 0,5$.

4.4.1. $g(y) = \frac{3}{20} (y-1)(y-3)$ für $3 < y < 5$ und $g(y) = 0$ sonst.

4.4.2. $g(y) = \frac{|a|}{y^2}$ für $a < y < \infty$ und $g(y) = 0$ sonst.

4.4.3. $g(y) = \frac{1}{2(b-a)\sqrt{2y}}$ für $2a^2 < y < 2b^2$ und $g(y) = 0$ sonst.

4.4.4. $g(y) = \frac{2}{\sqrt{2\pi}} e^{-y^2/2}$ für $0 < y < \infty$ und $g(y) = 0$ sonst.

4.4.5. $g(y) = \binom{n}{y-2} p^{y-2} (1-p)^{n-y+2}$ für $y = 2, 3, 4, \dots, n+2$ und $g(y) = 0$ sonst.

Aufgaben

$$4.1. \quad (a) F(x) = \begin{cases} 0 & \text{für } -\infty < x < 0 \\ 1/10 & \text{für } 0 \leq x < 1 \\ 5/10 & \text{für } 1 \leq x < 2 \\ 7/10 & \text{für } 2 \leq x < 3 \\ 9/10 & \text{für } 3 \leq x < 4 \\ 1 & \text{für } 4 \leq x < \infty \end{cases}$$

$$\begin{aligned} (b) & 9/10; \\ (c) & 4/10; \\ (d) & 9/10; \\ (e) & 3/10; \\ (f) & 8/10; \\ (g) & 4/10. \end{aligned}$$

4.2. (a) $k = 3$;

$$(b) F(x) = \begin{cases} 0 & \text{für } -\infty < x < 0 \\ 3/2 x^2 & \text{für } 0 \leq x \leq \sqrt{2/3} \\ 1 & \text{für } \sqrt{2/3} < x < \infty \end{cases}$$

4.3. (a) $k = 6$;

$$(b) F(x) = \begin{cases} 0 & \text{für } -\infty < x < 1 \\ 1/7 & \text{für } 1 \leq x < 2 \\ 3/7 & \text{für } 2 \leq x < 3 \\ 6/7 & \text{für } 3 \leq x < 4 \\ 1 & \text{für } 4 \leq x < \infty \end{cases}$$

4.4. (a) $k = \sqrt{1/2}$;

$$(b) F(x) = \begin{cases} 0 & \text{für } -\infty < x < 0 \\ 2x^2 & \text{für } 0 \leq x \leq \sqrt{1/2} \\ 1 & \text{für } \sqrt{1/2} < x < \infty \end{cases}$$

(c) 0,5.

4.5. $k = 1/2$.

$$4.6. \quad (a) F(x) = \begin{cases} 0 & \text{für } -\infty < x \leq 1 \\ 1/5 (2x^{3/2} - 3x + 1) & \text{für } 1 < x < 4 \\ 1 & \text{für } 4 \leq x < \infty \end{cases}$$

(b) 0,4785;

$$(d) \frac{3(\sqrt{x}-1)}{5 \cdot 0,5215};$$

$$(c) \frac{3(\sqrt{x}-1)}{5 \cdot 0,1314};$$

(e) 0,3471.

$$4.7. \quad (a) F(x) = \begin{cases} 0 & \text{für } -\infty < x \leq 1 \\ 1/5 (2x^3 - 3x^2 + 1) & \text{für } 1 < x < 2 \\ 1 & \text{für } 2 \leq x < \infty \end{cases}$$

$$(b) g(y) = \frac{6}{y^4} (5-y)(5-2y) \quad \text{für } 5/3 < y < 5/2;$$

$$(c) g(y) = \frac{3}{5} (\sqrt{y}-1) \quad \text{für } 1 < y < 4.$$

4.8. $g(y) = 1/2 e^{-y^{1/2}}$ für $0 < y < \infty$.

4.9. (a) $g(y) = 2y$ für $0 < y < 1$;

$$(b) g(y) = 1/y \quad \text{für } 1 < y < e.$$

4.10. $\left(\frac{n}{\sqrt{y}}\right) p^{\sqrt{y}} (1-p)^{n-\sqrt{y}}$ für $y = 0, 1, 4, 9, \dots, n^2$.

4.11. $g(y) = \frac{8(3y-2)}{9y^4}$ für $2/3 < y < \infty$.

4.12. (a) $g(y) = \frac{1}{\pi \sqrt{y} (y+1)}$ für $0 < y < \infty$;

$$(b) g(y) = \frac{\sqrt{3}}{\pi \sqrt{y} (y+3)} \quad \text{für } 0 < y < \infty.$$

- 4.14. (a) $g(y) = \frac{1}{2}y^3$ für $\frac{1}{2} < y < \infty$;
 (b) $g(y) = 2$ für $0 < y < \frac{1}{2}$;
 (c) $g(y) = 2y^{-1} \ln y$ für $1 < y < e$;
 (d) $g(y) = 4y^3$ für $0 < y < 1$;
 (e) $g(y) = 1$ für $0 < y < 1$.
- 4.15. (a) $F(x) = \begin{cases} 0 & \text{für } -\infty < x \leq -1 \\ \frac{1}{4}(x+1)^2 & \text{für } -1 < x < 1 \\ 1 & \text{für } 1 \leq x < \infty \end{cases}$;
 (b) $g(y) = \frac{1}{2}\sqrt{y}$ für $0 < y < 1$;
 (c) $g(y) = \frac{1}{18}(y+3)$ für $-3 < y < 3$.
- 4.16. $f(x|X > 0) = \frac{2}{3}(x+1)$ für $0 < x < 1$.
- 4.17. $a = \sqrt{2} - 1$.
- 4.18. $a = 0,5$.
- 4.19. $g(y) = \frac{1}{\sqrt{2\pi}} y^{\frac{1}{2}} e^{-\frac{1}{2}y}$ für $0 < y < \infty$.
- 4.20. (a) $a = 0,4$; (b) $a = 0,025$.
- 4.21. (a) $g(y) = \frac{1}{4\sqrt{y}}$ für $0 < y < \frac{1}{4}$;
 (b) $g(y) = 4y$ für $0 < y < \frac{1}{4\sqrt{2}}$;
 (c) $g(y) = 2e^y$ für $-\infty < y < -\ln 2$;
 (d) $g(y) = \frac{2}{y}$ für $1 < y < \sqrt{e}$;
 (e) $g(y) = 1$ für $5 < y < 6$;
 (f) $g(y) = \frac{3}{y^2}$ für $3 < y < \infty$.
- 4.22. (a) $g(y) = \frac{2}{y}$ für $\frac{1}{\sqrt{e}} < y < 1$;
 (b) $g(y) = \frac{1}{y}$ für $1 < y < e$;
 (c) $g(y) = \frac{2}{(y-b)}$ für $a+b < y < a\sqrt{e}+b$

Kapitel 5

Übungen

- 5.1.1. 1,7.
 5.1.2. $\frac{9}{5}$.
 5.1.3. $\frac{n+1}{2}$.
 5.1.4. $(q+pe)^n$.
 5.1.5. 1.
 5.2.1. $\mu'_1 = \frac{7}{6}$; $\mu'_2 = \frac{1}{18}$.

5.2.2. $1 / [2^r (r+1)]$.

5.2.3. $E(\frac{x-\mu'_1}{\sigma}) = E(X - \mu'_1) / \sigma = [\text{laut (5.7)}] = \frac{1}{\sigma} 0 = 0$;

$$V(\frac{x-\mu'_1}{\sigma}) = [\text{laut (5.15)}] = \frac{V(X - \mu'_1)}{\sigma^2} =$$

$$= [\text{laut (5.14)}] = V(X) / \sigma^2 = \sigma^2 / \sigma^2 = 1.$$

5.2.4. $\mu'_1 = 7$; $\mu'_2 = 54,83$; $\mu_2 = 5,83$.

5.3.1. $k = 6,32$.

5.3.2. $k = 15,8$.

5.4.1. (b) 0; (c) -0,6.

5.4.2. $\gamma_1 = -0,5657$; $\gamma_2 = -0,6000$.

5.4.3. $\gamma_1 = 0,5349$; $\gamma_2 = -0,2937$.

5.6.1. Vgl. Kap. 5.6.

5.6.2. Vgl. Kap. 5.6.

Aufgaben

5.1. $E(c) = \sum c p(x) = c \sum p(x) = c \cdot 1 = c$.

5.2. $E[ch(x)] = \int ch(x) f(x) dx = c \int h(x) f(x) dx = c E[h(x)]$.

5.3. (a) $\frac{1}{2^{r/2-1}(r+2)}$;

(b) $(r+1)!$;

(c) $\frac{6}{(r+2)(r+3)}$;

(d) $\frac{6 [2^{r+2}(r+1)+1]}{5(r+2)(r+3)}$;

(e) $\frac{3^{2r/3+1}}{2^{2r/3}(2r+3)}$;

(f) $\frac{e^{r+1}(2e+1)+1}{2^{r+1}(r+1)(r+2)}$;

(g) $\frac{1}{r+1}$ wenn r gerade und $\frac{1}{(r+2)}$ wenn r ungerade;

(h) $\frac{3^r}{r+1}$ wenn r gerade und $\frac{1}{r+2}$ wenn r ungerade;

(i) $\frac{2^r (b^{r+1/2} - a^{r+1/2})}{(2r+1)(b-a)}$;

$$(j) \frac{1}{2r+1}.$$

- 5.4. (a) $1/36$;
(b) 2;
(c) $1/20$;
(d) $1/20$;
(e) 0,15140;
(f) $1/6$;

- (g) $7/18$;
(h) 2;
(i) $\frac{4(b^{5/2} - a^{5/2})}{5(b-a)} - \frac{4(b^{3/2} - a^{3/2})^2}{9(b-a)^2}$;
(j) $4/45$.

5.5. $\mu_3(y) = \mu_3(ax) = E([aX - E(aX)]^3) = a^3 E([X - E(X)]^3) = a^3 \mu_3(x).$

5.6. $\gamma_1 = 0$; $\gamma_2 = -1, 2.$

5.7. (a) $\mu'_r = \frac{b^{r+1} - a^{r+1}}{(r+1)(b-a)};$

(b) $\mu_2 = \frac{(b-a)^2}{12}$; $\mu_3 = 0$; $\mu_4 = \frac{(b-a)^4}{80}$;

(c) 0;

(d) -1, 2.

5.8. $\sum_i (x_i - \mu'_1)^4 p(x_i) \geq [\sum_i (x_i - \mu'_1)^2 p(x_i)]^2$. Bezeichnen wir $x_i - \mu'_1 = y_i$, $p(x_i) = p_i$ und multiplizieren die linke Seite mit $1 = \sum_i p_i$. Dann ist $\sum_i y_i^4 p_i \sum_i p_i = (\sum_i y_i^4 p_i) (\sum_i p_i) \geq (y_1^2 p_1 + y_2^2 p_2 + \dots + y_n^2 p_n)^2$.

Daraus ergibt sich weiter

$$\begin{aligned} & y_1^4 p_1^2 + y_2^4 p_2^2 + \dots + y_n^4 p_n^2 + p_1 p_2 (y_1^4 + y_2^4) + p_1 p_3 (y_1^4 + y_3^4) + \dots \\ & + p_{n-1} p_n (y_{n-1}^4 + y_n^4) \geq \\ & \geq y_1^4 p_1^2 + y_2^4 p_2^2 + \dots + y_n^4 p_n^2 + 2p_1 p_2 y_1^2 y_2^2 + 2p_1 p_3 y_1^2 y_3^2 + \dots + \\ & + 2p_{n-1} p_n y_{n-1}^2 y_n^2. \end{aligned}$$

Einiges hebt sich auf beiden Seiten auf und es bleibt übrig $p_1 p_2 (y_1^4 + y_2^4) - 2p_1 p_2 y_1^2 y_2^2 + p_1 p_3 (y_1^4 + y_3^4) - 2p_1 p_3 y_1^2 y_3^2 + \dots + p_{n-1} p_n (y_{n-1}^4 + y_n^4) - 2p_{n-1} p_n y_{n-1}^2 y_n^2 = \sum_{i \neq j} p_i p_j (y_i^2 - y_j^2)^2 \geq 0$, was immer gilt, weil $p_i \geq 0$

und die Quadrate immer positiv sind.

5.9. 0,1311.

5.10. $\mu_2 = 2/75$, woraus $\sigma = 0,1633 \geq 0,1311$.

5.11. $\sigma^2 = \sum_i (x_i - \mu'_1)^2 p(x_i) \geq [\sum_i (x_i - \mu'_1) p(x_i)]^2$.

Sei $x_i - \mu'_1 = y_i$, $p(x_i) = p_i$, $1 = \sum_i p_i$, dann ist $\sigma^2 = \sum_i y_i^2 p_i \sum_i p_i = (\sum_i y_i^2 p_i) (\sum_i p_i) \geq (y_1^2 p_1 + y_2^2 p_2 + \dots + y_n^2 p_n)^2$,

woraus folgt $y_1^2 p_1^2 + y_2^2 p_2^2 + \dots + y_n^2 p_n^2 + p_1 p_2 (y_1^2 + y_2^2) + p_1 p_3 (y_1^2 + y_3^2) + \dots + p_{n-1} p_n (y_{n-1}^2 + y_n^2) \geq y_1^2 p_1^2 + y_2^2 p_2^2 + \dots + y_n^2 p_n^2 + 2p_1 p_2 y_1 y_2 + 2p_1 p_3 y_1 y_3 + \dots + 2p_{n-1} p_n y_{n-1} y_n$

also ist $p_1 p_2 (y_1^2 + y_2^2) - 2p_1 p_2 y_1 y_2 + p_1 p_3 (y_1^2 + y_3^2) - 2p_1 p_3 y_1 y_3 + \dots$

$\dots + p_{n-1} p_n (y_{n-1}^2 + y_n^2) - 2p_{n-1} p_n y_{n-1} y_n =$

$= \sum_{i \neq j} p_i p_j (y_i - y_j)^2 \geq 0$, was immer gilt.

5.12. $\mu'_4(y) = \mu'_4(ax + b) = E[(aX + b)^4] =$

$$= E(a^4 X^4 + 4a^3 b X^3 + 6a^2 b^2 X^2 + 4ab^3 X + b^4) = \sum_{k=0}^4 \binom{4}{k} a^{4-k} b^k \mu'_{4-k}(x).$$

5.13. (a) $E(x^2/2) = (e-1)/2e$;

(b) $g(y) = 1/y$ für $1/2e < y < 1/2$; $E(Y) = (e-1)/2e$.

5.14. (a) $E(x^{2r}) = 1/(4r+1)$;

(b) $g(y) = 1/4y^{3/4}$ für $0 < y < 1$; $E(Y) = 1/(4r+1)$.

5.15. (a) $E(e^X) = 2\sqrt{2} [\sqrt{2} + e^{1/\sqrt{2}} (1 - \sqrt{2})]$;

(b) $g(y) = \frac{4 \ln y}{y}$ für $1 < y < e^{1/\sqrt{2}}$; $E(Y) = E(e^X)$.

5.16. (a) $1/2^{1/4}$;

(d) 2a;

(b) $1/\sqrt{2}$;

(e) 1.

(c) $e^{1/2}$;

5.17. (a) $\frac{(r+4)!}{4! 3^r}$;

(b) $3^5 (3-t)^{-5} = (1-1/3)^{-5}$;

(c) $\mu'_1 = 5/3$; $\mu'_2 = 30/9$.

5.18. $x_M = 1$.

5.19. (a) $\mu'_r = \frac{1}{2} \Gamma\left(\frac{r+3}{2}\right)$;

$$(b) \mu_2 = \frac{3\sqrt{\pi} - 2}{8};$$

(c) 1.

5.20. (a) 1;

(b) 0,25;

(c) 0,11.

$$\begin{aligned} 5.21. (a) \gamma_1(x) &= \frac{\mu_3(x)}{\mu_2^{3/2}(x)} = \frac{E(X - \mu'_1)^3}{[E(X - \mu'_1)^2]^{3/2}} \cdot \frac{E(aY + b - a\mu'_1 - b)^3}{[E(aY + b - a\mu'_1 - b)^2]^{3/2}} = \\ &= \frac{a^3 E(Y - \mu'_1)^3}{a^3 [E(Y - \mu'_1)^2]^{3/2}} = \frac{\mu_3(y)}{\mu_2^{3/2}(y)}. \end{aligned}$$

Der Beweis für (b) verläuft analog mit anderen Exponenten.

Kapitel 6

Übungen

6.1.1. $k = 1$.

6.1.2. Ja.

$$6.1.3. F(x, y) = \begin{cases} 0 & x, y \leq 0 \\ x^2 y^2 & 0 < x, y < 1 \\ x^2 & 0 < x < 1, y \geq 1 \\ y^2 & x \geq 1, 0 < y < 1 \\ 1 & x, y \geq 1 \end{cases}$$

6.1.4. (a)

	$1 \leq y < 2$	$2 \leq y < 3$	$y \geq 3$
$0 \leq x < 1$	$\frac{2}{15}$	$\frac{6}{15}$	$\frac{9}{15}$
$x \geq 1$	$\frac{5}{15}$	$\frac{10}{15}$	1

(b) $\frac{7}{15}$.

6.1.5. $(e-1)/e^2$.

6.2.1. (a) $f_1(x) = 2x, 0 < x < 1$;

(b) $f_2(y) = 4(1-2y), 0 < y < \frac{1}{2}$.

6.2.2. (a) $f_1(x) = 3x^2, 0 < x < 1$;

(b) $f_2(y) = \frac{3}{2}(1-y^2), 0 < y < 1$.

$$6.2.3. (a) f_1(x) = \begin{cases} \frac{10}{16} & x = 1 \\ \frac{6}{16} & x = 2 \end{cases}; \quad (b) f_2(y) = \begin{cases} \frac{4}{16} & y = 3 \\ \frac{6}{16} & y = 4 \\ \frac{6}{16} & y = 5 \end{cases}$$

6.3.1. (a) $f(x|y) = \frac{2(x+y)}{1+2y}, 0 < x < 1, 1 < y < 2$;

$$(b) F(x|y) = \frac{x(x+2y)}{1+2y}, 0 < x < 1, 1 < y < 2;$$

$$(c) f(x|y < \frac{3}{2}) = (4x+5)/7.$$

6.3.2. $f_1(x) = 1$; $f_2(y) = \frac{1}{2}$; unabhängig.

6.3.3. 0,5625.

6.4.1. $\mu_{11} = \frac{1}{80}$.

6.4.2. $\mu'_{20} = \frac{3}{5}$; $\mu'_{02} = \frac{1}{5}$.

6.4.3. $\mu'_{10} = \frac{13}{24}$; $\mu'_{01} = \frac{37}{24}$; $\mu'_{11} = \frac{5}{6}$;

$\mu'_{20} = \frac{3}{8}$; $\mu'_{02} = \frac{59}{24}$;

$\mu_{20} = \frac{47}{576}$; $\mu_{02} = \frac{47}{576}$;

$\rho = -0,0213$

6.4.4. (a) $\frac{1}{2}$;

$$(b) 2(1-t_2)^{-1}(2-t_1-t_2)^{-1}$$

6.5.1. $g(u) = -\ln u, 0 < u < 1$.

$$6.5.2. (a) g(u) = \begin{cases} \frac{1}{4} & u = 0 \\ \frac{1}{2} & u = 1 \\ \frac{1}{4} & u = 2 \end{cases}; \quad (b) g(u) = \begin{cases} \frac{1}{4} & u = -1 \\ \frac{1}{2} & u = 0 \\ \frac{1}{4} & u = 1 \end{cases}$$

6.5.3. $g(u) = 2(u+1)^{-2}, 0 < u < 1$.

Aufgaben

6.1. $k = \frac{1}{3}$.

$$6.2. F(\infty|y) = \sum_x \frac{f(x,y)}{f(\cdot,y)} = \frac{f(\cdot,y)}{f(\cdot,y)} = 1;$$

$$F(\infty|y) = \int_{-\infty}^{\infty} \frac{f(x,y) dx}{f_2(y)} = \frac{f_2(y)}{f_2(y)} = 1.$$

6.3. $\mu'_{01} = \sum_y \sum_x y f(x,y) = \sum_y y \sum_x f(x,y) = \sum_y y f(\cdot,y) = E(Y)$.

$$\begin{aligned} 6.4. (a) \rho_{ax,by} &= \frac{E\{[aX - E(aX)][bY - E(bY)]\}}{\sqrt{V(aX)} \sqrt{V(bY)}} \\ &= \frac{ab E\{[X - E(X)][Y - E(Y)]\}}{\pm ab \sqrt{V(X)} \sqrt{V(Y)}} = \pm \rho_{x,y}; \\ (b) \rho_{x+a,y+b} &= \frac{E\{[X+a - E(X+a)][Y+b - E(Y+b)]\}}{\sqrt{V(X+a)} \sqrt{V(Y+b)}} \\ &= \frac{E\{[X - E(X)][Y - E(Y)]\}}{\sqrt{V(X)} \sqrt{V(Y)}} = \rho_{x,y}. \end{aligned}$$

$$6.5. E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x,y) dx dy = [\text{wegen Unabhängigkeit}] =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_1(x) f_2(y) dx dy$$

$$= \int_{-\infty}^{\infty} x f_1(x) dx \int_{-\infty}^{\infty} y f_2(y) dy = E(X) E(Y).$$

$$6.6. (a) f_1(x) = \begin{cases} 4/10 & x=0 \\ 6/10 & x=1 \end{cases}; \quad (d) f(x|y=2) = \begin{cases} 3/5 & x=0 \\ 2/5 & x=1 \end{cases};$$

$$(b) f_2(y) = \begin{cases} 5/10 & y=2 \\ 5/10 & y=3 \end{cases}; \quad (e) \text{nein.}$$

$$(c) \mu'_{10} = 3/5; \quad \mu'_{01} = 5/2;$$

$$6.7. (a) F(x,y) = \begin{cases} 0 & x \leq 0, \quad y \leq 1 \\ 1/6 x (y-1) (7-x-y) & 0 < x < 1, \quad 1 < y < 3 \\ 1/3 (4-x) & 0 < x < 1, \quad y \geq 3 \\ 1/6 (y-1) (6-y) & x \geq 1, \quad 1 < y < 3 \\ 1 & x \geq 1, \quad y \geq 3 \end{cases};$$

$$(b) f_1(x) = \frac{2}{3} (2-x), \quad 0 < x < 1; \quad f_2(y) = \frac{1}{6} (7-2y), \quad 1 < y < 3;$$

$$(c) F_1(x) = \frac{x}{3} (4-x), \quad 0 < x < 1; \quad F_2(y) = \frac{1}{6} (y-1) (6-y), \quad 1 < y < 3;$$

(d) nein.

$$6.8. \text{Cov}(aX, bY) = E\{[aX - E(aX)] [bY - E(bY)]\} =$$

$$= ab E\{[X - E(X)] [Y - E(Y)]\} = ab \text{Cov}(X, Y).$$

$$6.9. (a) k = 1/3;$$

$$(b) F(x,y) = \begin{cases} 0 & x, y \leq 0 \\ 1/12 xy (8-xy) & 0 < x < 2, \quad 0 < y < 1 \\ 1/12 x (8-x) & 0 < x < 2, \quad y \geq 1 \\ 1/3 y (4-y) & x \geq 2, \quad 0 < y < 1 \\ 1 & x \geq 2, \quad y \geq 1 \end{cases}$$

$$(c) f(x|y) = \frac{2-xy}{2(2-y)}, \quad 0 < x < 2, \quad 0 < y < 1;$$

$$(d) f(y|x) = \frac{2(2-xy)}{4-x}, \quad 0 < y < 1, \quad 0 < x < 2;$$

$$(e) F(x|y) = \frac{x(4-xy)}{4(2-y)}, \quad 0 < x < 2, \quad 0 < y < 1;$$

$$(f) F(y|x) = \frac{y(4-xy)}{4-x}, \quad 0 < y < 1, \quad 0 < x < 2;$$

$$(g) f(y|x=1) = \frac{2}{3} (2-y), \quad 0 < y < 1;$$

$$(h) f(y|1/2 < x < 3/2) = \frac{2}{3} (2-y), \quad 0 < y < 1;$$

$$(i) E(X^r Y^q) = \frac{2^{r+2} (r+q+3)}{3(r+1)(r+2)(q+1)(q+2)};$$

$$(j) \mu'_{10} = 8/9; \quad \mu'_{20} = 10/9; \quad \mu'_{01} = 4/9; \quad \mu'_{02} = 5/18; \quad \mu'_{11} = 10/27;$$

$$(k) \rho = -2/13;$$

$$(l) E(X^2|Y=1/2) = 10/9.$$

$$6.10. \Sigma_y \Sigma_x (x+y) f(x,y) = \Sigma_y \Sigma_x x f(x,y) + \Sigma_y \Sigma_x y f(x,y)$$

$$= \Sigma_x x \Sigma_y f(x,y) + \Sigma_y y \Sigma_x f(x,y)$$

$$= \Sigma_x x f(x, \cdot) + \Sigma_y y f(\cdot, y) = E(X) + E(Y).$$

$$6.11. g(u) = 1/2 (\ln 2 - \ln u), \quad 0 < u < 2.$$

$$6.12. g(u) = 4(1-u+u \ln u), \quad 0 < u < 1.$$

$$6.13. g(u) = 2(1-u), \quad 0 < u < 1.$$

$$6.14. g(u) = \frac{2(u-1)}{u^3}, \quad 1 < u < \infty.$$

$$6.15. \rho = 0,1860.$$

$$6.16. (a) g(u) = \begin{cases} 4/10 & u=0 \\ 2/10 & u=2 \\ 4/10 & u=3 \end{cases}; \quad (b) g(u) = \begin{cases} 1/10 & u=-3 \\ 7/10 & u=-2 \\ 2/10 & u=-1 \end{cases}$$

$$6.17. (a) f(y|x) = \frac{2(x-y)}{x^2}, \quad 0 < y < x < 1;$$

$$(b) f(y|x=1/2) = 4(1-2y), \quad 0 < y < 1/2;$$

$$(c) E(Y|X) = 2/3.$$

$$6.18. (a) M_{X,Y}(t_1, t_2) = 4(2-t_2)^{-2} (1-t_1)^{-2}; \quad (d) \mu'_{11} = 2;$$

$$(b) \mu'_{10} = 2;$$

$$(e) \mu_{11} = 0.$$

$$(c) \mu'_{01} = 1;$$

$$6.19. g(u) = 24 u (u+2)^{-4}, \quad 0 < u < \infty.$$

$$6.20. P(0 < X < 0,5, \quad 1 < Y < 2) = 3/8.$$

Index

A

Abweichung
mittlere 136, 158
Asymmetrie, s. Schiefe

B

Batschka-Deutsch 5, 6, 8
Bayesscher Satz 89
Formel (Wahrscheinlichkeit a posteriori) 90
Binomialkoeffizient 41, 46
Binomialverteilung 112, 126, 127, 133, 157
Brunk, H. D. 94, 128, 212
Bryan, J. G. 128, 212

C

Cauchy-Verteilung 127
Cluster 38
Cramér, H. 160

D

Dichte 102
Durchschnitt, s. Mittel

E

Ereignisraum 61
Ereignisse
abhängige 87
äquivalente 61
disjunkte 61

günstige 63
sich ausschließende (s. auch disjunkte) 61
sichere 61
unabhängige 87, 88
unmögliche 61
zufällige 61
Erwartung 129, 134
bedingte, einer Fkt. zweier Zufallsvariablen 182, 186
Sätze 131
Exponentialverteilung 109, 112, 155
Exzeß 143

F

$F(x)$, s. Verteilungsfunktion
Fakultät 30
Faltung von Zufallsvariablen 207
Feller, W. 60, 94
Fichtenholz, G. M. 160
Fisz, M. 128, 160, 212
Funktionen
eindimensionaler Zufallsvariablen 112
zweidimensionaler Zufallsvariablen 198

G

Gammaverteilung 146
Georgisch 62
Geršić, S. 5, 6, 8, 10, 28
Gnedenko, B. W. 160
Gradshteyn, I. S. 28
Graybill, F. A. 94, 128, 160, 212
Greenberg, J. H. 82, 94
Grundgesamtheit 3

H

Hald, A. 94
Harris, B. 60, 94, 212
Häufigkeit
absolute 4
relative 7, 63
Häufigkeitshistogramm 9
Häufigkeitspolygon 11
Häufigkeitstabelle 4
Häufigkeitsverteilung 6
Herdan, G. 58, 60
Hogben, L. 60

I

Indonesisch 11, 58

J

Jacobian 201
Job, M. 62, 94

K

Kapur, J. N. 27, 28, 94, 160
Kendall, M. G. 28, 212
Klassen 4
Klassenbreite 4
Kombination
mit Wiederholung 43, 48, 53, 55
ohne Wiederholung 40, 48, 55
Kombinatorik 29

Korrelationskoeffizient 181, 191
Kovarianz 180
Kreyszig, E. 128

L

Laplace-Verteilung 209
Lehfeldt, W. 6, 9, 20, 27, 28
Ludvíková, M. 26, 28

M

Mathai, A. M. 128, 212
Median 159
Mengen
äquivalente 65
Differenz von 67
disjunkte 66
Durchschnitt von 66
Komplement von 67
Produkt von 66
Summe von (s. auch Vereinigung von) 65
Teil- 65
Vereinigung von 65
Messen 4
Mittel, s. auch Mittelwert
arithmetisches 14
Mittelwert 14, 129
Eigenschaften 19
Modus 159
Momente 133
absolute 158
allgemeine 134
Anfangs- 134, 137, 138, 153, 156, 178, 196
bedingte 182
mehrdimensionaler Zufallsvariablen 178
Zentral- 134, 137, 138, 155, 156, 180

momenterzeugende Funktion
eindimensionaler Zufallsvariablen 151
mehrdimensionaler Zufallsvariablen 193

Mood, A. M. 94, 128, 160, 212
de Morgans Gesetze 69
Mosteller, F. 60
Multinomialkoeffizient 35, 50
Multiplikationsregel 33, 55
Multiplikationssatz 85

N
 n über k , s. Binomialkoeffizient
Netto, E. 47, 60
Neyman, J. 28, 60
Normalverteilung 120, 126
Normierung, s. Zufallsvariable, normierte

O
Ovčarov, L. A. 94, 128, 160

P
Papoulis, A. 94, 128, 212
Pascalsches Dreieck 48
Permutationen 29
mit beschränkter Wiederholung 32, 55
mit Wiederholung 31, 55
ohne Wiederholung 29, 55
von Elementen, d. nicht alle unterschiedsch. sind 34, 55
Poisson-Verteilung 110, 125

R
Randverteilung 166
Rayleigh-Verteilung 122
Rechteckverteilung 112, 124, 126, 128, 130, 133, 154

Reim 72, 93
Rourke, R. E. K. 60
Rundung 8
Ryzhik, I. M. 28

S
Saxena, H. C. 27, 28, 94, 160
Schiefe 143
Stabdiagramm 10, 97
Standardabweichung 20, 21
Stichproben 3
-umfang 6
Streuung, s. Varianz
Strichliste 4
Stuart, A. 28, 212
Summenhäufigkeit 13

T
Taylorsche Formel 149
Thomas, G. B. Jr. 60
Treppenkurve 13
Tschebyschevsche Ungleichung 141

U
Umkehrfunktion 113, 114, 115
unendliche Reihe 150
Urliste 4

V
Varianz 20, 135, 136
Definition der 21
Variation
mit Wiederholung 39, 55
ohne Wiederholung 36, 42, 55
Ventcel', Je. S. 94, 128, 160
Verteilung
bedingte 107, 170
mehrdimensionale 161

Verteilungsfunktion 7, 98, 161, 163, 171

W
Wadsworth, G. P. 128, 212
Wahrscheinlichkeit 61
a posteriori 90
Axiome der 65, 95
bedingte 83
Definition der 61
klassische Definition der 63
statistische Definition der 62
totale 70
Wahrscheinlichkeitsdichte, s. Dichte
Wahrscheinlichkeitsfunktion 67, 95
Wahrscheinlichkeitsintegraltransformation 123

Weber, E. 28
Whitworth, W. A. 60, 94
Wölbung, s. Exzess
Wortstellung 33
feste 43
freie 36, 43
Maß für die Strenge der 31

Z
Zufallsvariable 3, 95
diskrete 4, 95
Funktionen von 112
normierte 157
stetige 4, 102
unabhängige 176
Zufallsvariable, kontinuierliche, s. stetige