Walter Runge · Gotthard Forbrig

Einführung in die Wahrscheinlichkeitsrechnung

für Ökonomen

Dr. Walter Runge · Prof. Dr. Gotthard Forbrig

Einführung in die Wahrscheinlichkeitsrechnung für Ökonomen

VERLAG DIE WIRTSCHAFT BERLIN

Redaktion: Dr. oec. Johannes Behr
Dipl. math. Kurt Brünecke

Verlag Die Wirtschaft, 1018 Berlin, Am Friedrichshain 22
1965 veröffentlicht · Lizenz-Nr. 122 · Druckgenehmigungs-Nr. 195/25/65
 · Umschlagentwurf: Peter Schulz, Berlin
Gesamtherstellung: Buchdruckerei Frankenstein KG, Leipzig
Es-Nr. 19 B 5

INHALTSVERZEICHNIS

VORWORT

Die ersten wahrscheinlichkeitstheoretischen Untersuchungen wurden bereits vor mehr als 300 Jahren durchgeführt. Sie bezogen sich damals auf Glücksspiele. Dieser Untersuchungsgegenstand überwiegt auch heute noch in verschiedenen Veröffentlichungen. Das kann sehr leicht zu dem Urteil verleiten, die Erkenntnisse hätten wenig praktischen Wert.

Die Wahrscheinlichkeitsrechnung ist aber in Wirklichkeit für die verschiedensten wissenschaftlichen Disziplinen unentbehrlich geworden. Ein bedeutender Teil der theoretischen Physik bedient sich der Erkenntnisse der Wahrscheinlichkeitstheorie. Auch im Bereich der Ökonomie zeichnet sich eine solche Entwicklung ab. Das neue ökonomische System der Planung und Leitung der Volkswirtschaft erfordert, alle Planungsaufgaben nach wissenschaftlichen Erkenntnissen zu lösen. Dazu gehört, wichtige ökonomische Entscheidungen soweit wie möglich durch komplexe Berechnungen zu begründen. Seit mehreren Jahren wird der Anwendung der Mathematik erhöhte Aufmerksamkeit gewidmet. Das Interesse galt zunächst solchen Methoden, die erst im Laufe des letzten Jahrzehnts bekannt geworden sind. Dies gilt beispielsweise für die lineare und nichtlineare Optimierung. Daneben gibt es jedoch zahlreiche Aufgaben, die mit Hilfe der Wahrscheinlichkeitsrechnung gelöst werden können. Im Versicherungswesen bedient man sich ihrer bereits verhältnismäßig lange. Auch die Stichprobentheorie — und in diesem Zusammenhang speziell die statistische Qualitätskontrolle — fußen auf den Erkenntnissen der Wahrscheinlichkeitsrechnung.

Es sind ständig neue Gebiete hinzugekommen. So setzt beispielsweise die optimale Lagerhaltung für Reparaturteile wahrscheinlichkeitstheoretische Überlegungen voraus. Es ist an die Theorie der Massenbedienung und an die Theorie der Spiele zu denken. Während die Theorie der Massenbedienung, insbesondere ihre Anwendung bei Warteschlangenmodellen, schon heute für die sozialistische Wirtschaft erheblichen Nutzen hat, bedarf die Anwendung der Theorie der Spiele noch weiterer Untersuchungen.

Bei richtigem Einsatz derartiger Berechnungen können erhebliche materielle und finanzielle Mittel eingespart werden. Der erfolgreiche Gebrauch derartiger Verfahren wird leider dadurch verzögert, daß insbesondere die

Ökonomen nur unzureichende mathematische Vorkenntnisse besitzen. Um diese Lücke zu schließen, wurde an verschiedenen Hochschulen und Fakultäten ein postgraduales Studium eingerichtet. An der Ingenieurökonomischen Fakultät der Universität Rostock wurden in diesem Rahmen unter anderem Fragen der Wahrscheinlichkeitsrechnung behandelt. Diese Voraussetzungen bildeten den Grundstock für die vorliegende Veröffentlichung. Damit soll die Möglichkeit geboten werden, einen größeren Interessentenkreis zu informieren. Außerdem ist diese Arbeit als Lehrmaterial für die Aneignung vertiefter mathematischer Kenntnisse gedacht.

Es wurde Wert darauf gelegt, möglichst viele ökonomische Beispiele auszuwählen. Bei den einleitenden Betrachtungen erwies es sich jedoch aus methodischen Gründen als zweckmäßig, einige Fälle zu behandeln, die Glücksspiele betreffen. Die mathematische Problematik läßt sich auf diese Weise besser erläutern.

Allen Kollegen, die uns bei der Gestaltung des Manuskripts beraten haben, sei herzlich gedankt. Insbesondere danken wir Herrn Dr. Johannes Behr und Herrn Dipl. math. Kurt Brünecke sowie Herrn Dipl. oec. Helmut Michael.

DIE VERFASSER

Rostock, 1. Oktober 1964

1. Zufällige Ereignisse

1.1. Gegenstand der Wahrscheinlichkeitsrechnung

Die Wahrscheinlichkeitsrechnung ist ein Teilgebiet der Mathematik, das die Gesetzmäßigkeiten im Bereich *zufälliger Erscheinungen* oder *Ereignisse* untersucht.

Zufällige Ereignisse sind beispielsweise die Ergebnisse beim Spiel mit einem Würfel. Es bleibt dem Zufall überlassen, ob bei einem Wurf 1, 2, 3, 4, 5 oder 6 Augen erzielt werden. Dabei ist noch keine Gesetzmäßigkeit zu erkennen. Ein anderer Sachverhalt ergibt sich jedoch, wenn in größerer Zahl mit demselben Würfel gewürfelt wird. Dann lassen sich Aussagen gewinnen, mit welcher Häufigkeit die einzelnen Augenzahlen erzielt werden.

Zufällige Erscheinungen und Gesetzmäßigkeiten schließen sich nicht aus, sondern bilden eine dialektische Einheit. Werfen wir eine Münze in die Höhe, dann fällt sie auf Grund des objektiven Wirkens des Schwerefelds der Erde auf den Boden. Das Fallen des Geldstücks auf den Boden ist also gesetzmäßig. Dagegen ist es zufällig, ob die am Boden liegende Münze mit der Zahl-Seite oder mit der Wappen-Seite nach oben zeigt (Wir schließen bei dieser Überlegung den Fall aus, daß die Münze hochgekantet steht!). Wir haben es hier also mit den zufälligen Ereignissen ‚Zahl‘ und ‚Wappen‘ zu tun. Das notwendige (gesetzmäßige) Fallen des Geldstücks auf den Boden ist untrennbar mit dem Eintreten eines der beiden zufälligen Ereignisse verbunden. Führen wir eine Serie von Würfen mit ein und demselben Geldstück durch und beobachten wir jeweils den Ausgang der angestellten ‚Versuche‘, dann ergeben sich gewisse Aussagen über das Wesen der zufälligen Ereignisse ‚Zahl‘ und ‚Wappen‘.

Wie die vorstehenden Beispiele zeigen, interessieren uns neben den Einzelereignissen besonders Serien von Einzelereignissen. Zu diesem Zweck führen wir den Begriff der *Massenerscheinung* ein. Unter einer Massenerscheinung verstehen wir einen solchen Vorgang, der in einer Gesamtheit stattfindet, die aus einer großen Anzahl von gleichberechtigten Ereignissen unter ein und demselben Komplex von Bedingungen besteht. So sind das mehrmalige Würfeln mit einem Würfel, das mehrfache Werfen mit einer Münze und die Massenproduktion einer bestimmten Erzeugnisart einige Beispiele für Massenerscheinungen. Führen wir Versuche aus und analysieren die Resultate der einzelnen Versuche, so erhalten wir Aufschluß über das Wesen der betreffenden Massenerscheinung.

Somit untersucht die Wahrscheinlichkeitsrechnung die Gesetzmäßigkeiten von Massenerscheinungen.

1.2. Der Begriff der Wahrscheinlichkeit — Der Additionssatz

Wir bezeichnen im folgenden zufällige Ereignisse einer Massenerscheinung mit großen lateinischen Buchstaben (zuweilen mit Indizes aus der Menge der natürlichen Zahlen). So seien zum Beispiel Z und W die Ereignisse „Zahl" und „Wappen" beim Werfen mit einer Münze. Die beim Würfeln mit einem Würfel möglichen zufälligen Ereignisse können wir mit A_1, A_2, A_3, A_4, A_5 und A_6 bezeichnen; dabei bedeutet allgemein A_k ($k = 1, 2, 3, 4, 5, 6$) das zufällige Ereignis „k-Augen bei einem Wurf mit einem Würfel".

Wir führen nun einige grundlegende Begriffe ein. Zunächst definieren wir das dem beliebigen Ereignis A *entgegengesetzte* oder *komplementäre* Ereignis $\bar{A}$. Das Ereignis $\bar{A}$ tritt genau dann ein, wenn das Ereignis A nicht stattfindet. Ist A das Ereignis „das Erzeugnis ist normgerecht", dann bedeutet $\bar{A}$ das Ereignis „das Erzeugnis ist nicht normgerecht". Das komplementäre Ereignis zu dem Ereignis $\bar{A}$ ist offensichtlich A selbst, d. h es gilt

$$\boxed{\bar{\bar{A}} = A}$$

Wir lernen jetzt Verknüpfungen von Ereignissen kennen, d. h. Operationen mit Ereignissen, die zwei oder mehreren Ereignissen ein eindeutig bestimmtes Ereignis zuordnen. Zunächst erläutern wir die „*Addition*" *zweier Ereignisse* A und B. Wir sagen: das Ereignis $A + B$ tritt ein, wenn mindestens eines der beiden Ereignisse eintritt und nennen $A + B$ die *Summe der Ereignisse* A und B. Zur Verdeutlichung dieses Begriffs betrachten wir die Massenerscheinung Würfeln mit einem Würfel. Verwenden wir die oben eingeführten Bezeichnungen, dann bedeutet $A_1 + A_2$ das Ereignis „höchstens zwei Augen" und $A_5 + A_6$ das Ereignis „wenigstens 5 Augen". Die Summe von mehr als zwei Ereignissen erklärt sich sukzessiv aus der oben eingeführten Summe von zwei Ereignissen. So ist zum Beispiel $A_1 + A_2 + A_3 + A_4 + A_5$ das Ereignis „keine sechs Augen" und damit identisch mit $\bar{A}_6$; in diesem Fall gilt also die Ereignis-Gleichung

$$A_1 + A_2 + A_3 + A_4 + A_5 = \bar{A}_6.$$

Wir führen nun die „*Multiplikation*" *zweier Ereignisse* A und B ein. Wir sagen: das Ereignis AB tritt ein, wenn die beiden Ereignisse A und B zugleich eintreten und nennen AB das *Produkt der Ereignisse* A und B. Zur Verdeutlichung dieses Begriffes betrachten wir die Massenerscheinung Würfeln mit zwei Würfeln. Mit den oben benutzten Bezeichnungen stellt dann $A_1 A_2$ das Ereignis „der eine Würfel zeigt ein Auge und der andere Würfel zeigt zwei Augen" dar. Das Produkt von mehr als zwei Ereignissen ergibt sich sukzessiv aus dem Produkt von zwei Ereignissen. So bedeutet $A_1A_1 A_1$ das Ereignis „alle drei Würfel zeigen jeweils ein Auge" bei einem Wurf mit drei Würfeln.

Wir nennen ein Ereignis *unmöglich* oder *ausgeschlossen*, wenn es in der betreffenden Massenerscheinung nicht eintreten kann. So ist das Ereignis „10 Augen bei einem Wurf mit einem Würfel" unmöglich. Ein Ereignis heißt dagegen *sicher*, wenn es stets in der betreffenden Massenerscheinung eintritt.

Beispielsweise ist das Ereignis „weniger als 7 Augen bei einem Wurf mit einem Würfel" ein sicheres Ereignis. Bezeichnen wir das unmögliche Ereignis mit U, das sichere mit S und ein beliebiges Ereignis mit A, dann gilt für die betreffende Massenerscheinung[1]

$$A + \bar{A} = S, \ A\bar{A} = U, \ AS = A, \ AU = U$$

sowie $\overline{S} = U$ und $\overline{U} = S$. Die vorstehenden Relationen lassen sich durch einfache Überlegungen beweisen.
Nach diesen notwendigen Vorbereitungen klären wir den grundlegenden Begriff der Wahrscheinlichkeit.

1.2.1. Die statistische Definition der Wahrscheinlichkeit

Wir betrachten eine bestimmte Massenerscheinung und in ihr ein gewisses zufälliges Ereignis A. Wir führen n Versuche aus und beobachten dabei $n_A$mal das Auftreten des Ereignisses A. Dann bezeichnen wir den Quotienten

$$\boxed{h_n(A) = \frac{n_A}{n}} \tag{1.1}$$

als die *relative Häufigkeit* für das Auftreten von Ereignis A innerhalb der durchgeführten Versuchsreihe.

Nehmen wir zum Beispiel als Massenerscheinung das Würfeln (mit einem Würfel) und als A das Ereignis A_1. Wir führen eine Serie von $n = 100$ Würfen durch und mögen dabei 19mal ($n_A = 19$) unser Ereignis A feststellen. Dann gilt für die relative Häufigkeit von A

$$h_n(A) = h_{100}(A) = \frac{19}{100} = 0{,}19 .$$

Wegen $0 \leqq n_A \leqq n$ gilt offenbar für jede relative Häufigkeit

$$\boxed{0 \leqq h_n(A) \leqq 1} . \tag{1.2}$$

Wegen der Zufälligkeit der Ereignisse ist $h_n(A)$ keine stabile Größe. Deshalb liefern Versuchsreihen gleichen Umfangs n unterschiedliche Werte für n_A uud gemäß (1.1) verschiedene relative Häufigkeiten für ein und dasselbe Ereignis A. Diesen Sachverhalt erläutern wir in der Stichprobentheorie ausführlicher.

Die nachstehende Tabelle möge die letzten Bemerkungen am Beispiel der Massenerscheinung Würfeln (mit einem Würfel) und des Ereignisses $A = A_1$ verdeutlichen.

[1] Der Querstrich über einem Ereignissymbol gibt den Übergang zum entgegengesetzten Ereignis an.

Nr. der Serie	Anzahl der Würfe n	Anzahl des Eintretens von A n_A	Relative Häufigkeit von A $h_n(A)$
1	100	19	0,19
2	100	14	0,14
3	100	15	0,15
4	100	17	0,17

Bei zunehmender Zahl der Versuche stabilisiert sich jedoch die Größe $h_n(A)$, das heißt $h_n(A)$ nähert sich einem festen Wert, wenn die Anzahl n der Versuche genügend groß gewählt wird. Diesen festen Wert bezeichnen wir als die *Wahrscheinlichkeit* $P(A)$ für das Auftreten des Ereignisses A. Wir schreiben daher

$$\boxed{P(A) = \lim_{n \to \infty} h_n(A)}\,. \tag{1.3}$$

Das ist die *statistische Definition* der Wahrscheinlichkeit für das Auftreten des Ereignisses A. Sie ermöglicht prinzipiell, Wahrscheinlichkeiten statistisch zu erfassen.

Betrachten wir einen regelmäßigen Würfel mit homogener Dichte, so erwarten wir für die Ereignisse A_i $(i = 1, 2, \ldots, 6)$ die Wahrscheinlichkeiten

$$P(A_i) = \frac{1}{6} \, (i = 1, 2, \ldots, 6)\,.$$

Die sechs Ereignisse $A_1, \ldots, A_6$ besitzen also dieselbe Wahrscheinlichkeit, daher nennen wir sie gleichwahrscheinlich. Aus der statistischen Definition (1.3) und der Beziehung (1.2) folgt

$$\boxed{0 \leqq P(A) \leqq 1}\,. \tag{1.4}$$

Die Wahrscheinlichkeit für das Auftreten eines Ereignisses (oder kurz: die Wahrscheinlichkeit eines Ereignisses) ist demzufolge eine Zahl aus dem abgeschlossenen Intervall $[0,1]$. Für das sichere Ereignis S gilt definitionsgemäß

$$\boxed{P(S) = 1}\,. \tag{1.5}$$

Ist $S = A_1 + A_2 + \ldots + A_6$ das Ereignis „eine der Zahlen 1 bis 6 bei einem Wurf mit einem Würfel“, so ist in der Tat $P(S) = 1$. Für das unmögliche Ereignis U gilt definitionsgemäß

$$\boxed{P(U) = 0}\,. \tag{1.6}$$

Ist U das Ereignis „mindestens 7 Augen bei einem Wurf mit einem Würfel“, so wird wirklich $P(U) = 0$. Aus $P(A) = 0$ folgt aber keineswegs $A = U$. Sei zum Beispiel bei einer Gütekontrolle das Ereignis A der Anfall von Ausschuß. In einem Betrieb wird für das Ereignis A die Wahrscheinlichkeit

$P(A) = 0$ angegeben. Es darf nun nicht gefolgert werden, daß in dem Betrieb Ausschuß ausgeschlossen ist. $P(A) = 0$ bedeutet lediglich, daß die relative Häufigkeit von Ausschuß sehr klein (nahe bei null) ist.

Nennen wir ein Ereignis A *selten*, wenn $P(A) = 0$ ist, dann ist im obigen Beispiel Ausschuß ein seltenes, aber kein unmögliches Ereignis.

Nachdem wir die Wahrscheinlichkeit für das Auftreten eines Ereignisses auf statistischem Wege bestimmt haben, untersuchen wir nun die Frage „Wie groß ist die Wahrscheinlichkeit einer Summe von Ereignissen, wenn die Wahrscheinlichkeiten der einzelnen Ereignisse bekannt sind?“. Zu diesem Zweck betrachten wir zunächst *unverträgliche Ereignisse*, das heißt solche Ereignisse, die sich paarweise ausschließen. Sind A und B beispielsweise unverträglich, dann gilt offenbar $P(AB) = 0$, denn das gleichzeitige Auftreten beider Ereignisse ist ausgeschlossen.

Bei der Einstufung der Qualität von Erzeugnissen gibt es die vier möglichen Ereignisse $Q_1 \equiv$ „Gütezeichen Q“; $Q_2 \equiv$ „Gütezeichen 1“; $Q_3 \equiv$ „Gütezeichen 2“ und $Q_4 \equiv$ „ohne Gütezeichen“. Diese Ereignisse sind unverträglich, denn es ist unmöglich, daß ein Erzeugnis mehreren Qualitätsstufen gleichzeitig angehört.

Ist schließlich A ein beliebiges Ereignis und bedeutet $\bar{A}$ das zugehörige komplementäre (entgegengesetzte) Ereignis, dann sind A und $\bar{A}$ offensichtlich unverträglich.

Um die oben aufgeworfene Frage für unverträgliche Ereignisse zu beantworten, bedienen wir uns folgenden Beispiels: In verschiedenen Betrieben des Schiffbaus werden zur Qualitätsüberwachung der Fertigung Gütenoten für die einzelnen Arbeitsgänge erteilt. Die beste Note ist die 1, die schlechteste die 4. Da keine Zwischennoten zugelassen sind, haben wir es hier mit den unverträglichen Ereignissen $N_i \equiv$ „der Arbeitsgang erhält die Note i“ $(i = 1, 2, 3, 4)$ zu tun. Nach Einstufung aller n Arbeiten eines Monats ergeben sich folgende relative Häufigkeiten $h_n(N_i)$ für die Ereignisse N_i $(i = 1, 2, 3, 4)$:

$$h_n(N_i) = 0{,}30, \; h_n(N_2) = 0{,}52, \; h_n(N_3) = 0{,}15, \; h_n(N_4) = 0{,}03$$

Die relative Häufigkeit für das Ereignis $N_1 + N_2$, das heißt die relative Häufigkeit für das Ereignis „ein Erzeugnis wird nach der Gütenote 1 oder 2 beurteilt“ ist demzufolge $0{,}30 + 0{,}52$ oder $0{,}82$. Allgemein gilt für zwei unverträgliche Ereignisse A und B die Beziehung

$$\boxed{h_n(A + B) = h_n(A) + h_n(B)}\,. \qquad (1.7)$$

Führen wir den Grenzübergang $n \to \infty$ durch, so ergibt sich auf Grund der statistischen Definition der Wahrscheinlichkeit der Additionssatz für die Wahrscheinlichkeiten zweier unverträglicher Ereignisse A und B zu

$$\boxed{P(A + B) = P(A) + P(B)}\,. \qquad (1.8)$$

Mit zwei Beispielen belegen wir den vorstehenden Sachverhalt. Bedeuten A_1 und A_2 die schon mehrfach erwähnten Ereignisse beim Würfeln mit einem

normierten Würfel (vergleiche oben), dann ist zunächst $P(A_1 = P(A_2) = \frac{1}{6}$.
Ferner ergibt sich für $A_1 + A_2$, das heißt für das Ereignis „höchstens zwei Augen bei einem Wurf“ nach Regel (1.8) die Wahrscheinlichkeit

$$P(A_1 + A_2) = P(A_1) + P(A_2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} .$$

Ein Betrieb stellt Walzblech in Längen zu 10 m her. Die Wahrscheinlichkeiten für die Ereignisse $A \equiv$ „fehlerloses Blech“ und $B \equiv$ „Blech hat genau einen Fehler“ betragen $P(A) = 0{,}85$ und $P(B) = 0{,}09$. Das Ereignis $A + B$, nämlich das Ereignis „Blech hat nicht mehr als einen Fehler“ besitzt damit die Wahrscheinlichkeit

$$P(A + B) = P(A) + P(B) = 0{,}85 + 0{,}09 = 0{,}94 .$$

Die Aussage (1.8) läßt sich auf induktivem Wege zu dem allgemeinen *Additionssatz der Wahrscheinlichkeiten* unverträglicher Ereignisse erweitern, den wir durch den folgenden Satz ausdrücken.

Satz 1.1: Die Wahrscheinlichkeit für die Summe von endlich oder abzählbar unendlich vielen unverträglichen Ereignissen $A_1, A_2, \ldots A_k, \ldots$ ist gleich der Summe der Wahrscheinlichkeiten der Einzelereignisse $A_1, \ldots A_k$, das heißt, es besteht der Zusammenhang[1]

$$\boxed{P\left(\sum_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} P(A_i)} \cdot \qquad (1.9)$$

Zu dem vorstehenden Satz geben wir abschließend zwei Beispiele:

Wir betrachten zunächst den Verkauf von Herrenschuhen in einem bestimmten Schuhgeschäft. Für die Nachfrage nach den einzelnen Größenklassen k_i sind folgende Wahrscheinlichkeiten bekannt:

Nr. i	Größenklasse K_i	Wahrscheinlichkeit P (K_i)
1	6	0,02
2	$6^1/_2$	0,06
3	7	0,08
4	$7^1/_2$	0,10
5	8	0,14
6	$8^1/_2$	0,14
7	9	0,16
8	$9^1/_2$	0,12
9	10	0,08
10	$10^1/_2$	0,04
11	11	0,02
12	$11^1/_2$	0,02
13	12	0,02

[1] Ist $k = \infty$, dann muß die rechte Seite in (1.9) konvergent sein.

Am Lager befinden sich nur noch Schuhe der Größen 6, $6^1/_2$, 7 und $7^1/_2$. Wie groß ist die Wahrscheinlichkeit dafür, daß die Wünsche eines beliebigen Käufers, der in diesem Schuhgeschäft einkaufen möchte, befriedigt werden können? Gesucht ist offenbar die Wahrscheinlichkeit für das Ereignis $K_1 + K_2 + K_3 + K_4$. Nach Satz 1.1 erhalten wir diese zu

$$P(K_1 + K_2 + K_3 + K_4) = P(K_1) + P(K_2) + P(K_3) + P(K_4) = 0{,}26\,.$$

Für das zweite Beispiel betrachten wir wieder die Massenerscheinung Würfeln mit einem normierten Würfel und berechnen die Wahrscheinlichkeiten für die Ereignisse $A = A_1 + A_3 + A_5$ und $B = A_1 + A_2 + A_3 + A_4$. Das erste Ereignis A ist das Ereignis „ungerade Augenzahl bei einem Wurf", das zweite Ereignis B ist äquivalent dem Ereignis „nicht mehr als 4 Augen bei einem Wurf". Nach Satz 1.1 ergibt sich fast von selbst das Ergebnis zu

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{2}{3}.$$

Wir fragen nun nach der Wahrscheinlichkeit für das Ereignis $A + B$. Wenden wir auf $A + B$ den Additionssatz an, so wird

$$P(A + B) = \frac{1}{2} + \frac{1}{3} = \frac{7}{6}.$$

Das Ergebnis ist offenbar sinnlos, da eine Wahrscheinlichkeit nie die Zahl 1 übertrifft. Wir gelangen zu dem richtigen Resultat, wenn wir bedenken, daß $A + B$ gleich dem Ereignis „keine 6 Augen bei einem Wurf" ist. Daher wird $P(A + B) = P(A_1 + \ldots + A_5) = \frac{5}{6}$. Die Anwendung des Satzes 1.1 zur Berechnung von $P(A + B)$ führte zu einem falschen Ergebnis, weil A und B verträglich sind, denn eine ungerade Augenzahl kann auch kleiner als 5 sein. Damit sind die Voraussetzungen des obigen Satzes verletzt. Diese Überlegung zeigt, daß die Unverträglichkeit der Ereignisse im Satz 1.1 eine notwendige Voraussetzung ist.

1.2.2. Die axiomatische Definition der Wahrscheinlichkeit

Im täglichen Sprachgebrauch hat der Begriff der Wahrscheinlichkeit keine genaue Bedeutung. Bisweilen wird zum Ausdruck gebracht, daß ein zufälliges Ereignis mit 50 Prozent Wahrscheinlichkeit eintrifft. Solche Bewertungszahlen sind meist nur gefühlsmäßig angegeben. Da aber die Wahrscheinlichkeitsrechnung eben mit diesem zentralen Begriff operiert, muß die Wahrscheinlichkeit als Bewertungszahl für Ereignisse präziser gefaßt werden, als es bisher geschehen ist.

Zu diesem Zweck geben wir im folgenden die Begriffsbestimmung von

KOLMOGOROFF [1][1]. Diese Definition besteht aus einem sogenannten Axiomensystem, das die Wahrscheinlichkeit eines Ereignisses mittels ihrer Eigenschaften erklärt.

1. Axiom: Jedem zufälligen Ereignis A ist eine bestimmte nicht negative Zahl $P(A)$ zugeordnet. Diese Zahl heißt die Wahrscheinlichkeit für das Ereignis A.
2. Axiom: Die Wahrscheinlichkeit des sicheren Ereignisses S ist 1.
3. Axiom: Für unverträgliche Ereignisse $A_1, A_2, \ldots$ gilt

$$P(A_1 + A_2 + \ldots) = P(A_1) + P(A_2) + \ldots$$

Die bei der obigen axiomatischen Definition aufgetretenen Eigenschaften des Wahrscheinlichkeitsbegriffs sind uns schon aus dem vorangegangenen Abschnitt geläufig. Alle anderen Eigenschaften dieses Begriffs lassen sich aus den drei Axiomen herleiten. Damit beinhalten diese Axiome das Minimum an Forderungen, die wir an einen vernünftigen Begriff der Wahrscheinlichkeit als Bewertungszahl für zufällige Ereignisse stellen müssen.
Wir leiten nun aus den Axiomen weitere, zum Teil bereits bekannte Aussagen über Wahrscheinlichkeiten ab.
Satz 1.2: Für das komplementäre Ereignis $\bar{A}$ eines Ereignisses A gilt

$$\boxed{P(\bar{A}) = 1 - P(A)}\,. \tag{1.10}$$

Beweis: A und $\bar{A}$ sind unverträglich. Ferner gilt $A + \bar{A} = S$. Nach Axiom 2 ist dann $P(A + \bar{A}) = 1$. Wegen Axiom 3 folgt weiter $P(A) + P(\bar{A}) = 1$ oder die Behauptung.

Satz 1.3: Das unmögliche Ereignis U hat die Wahrscheinlichkeit 0.

Beweis: Aus Satz 1.2 folgt mit $A = S$ und $\bar{A} = \bar{S} = U$ sofort $P(U) = 1 - P(S)$. Nun ist aber $P(S) = 1$ nach Axiom 2. Damit ist alles bewiesen.

Satz 1.4: Für jedes Ereignis A besteht die Beziehung

$$\boxed{0 \leqq P(A) \leqq 1}\,.$$

Beweis: Nach Axiom 1 ist zunächst $0 \leqq P(A)$. Wir zeigen nun $P(A) \leqq 1$ indirekt. Wäre nämlich $P(A) > 1$, dann würde gemäß Satz 1.2 $P(\bar{A}) = 1 -$. $P(A) < 0$ folgen. Dann wäre aber $\bar{A}$ ein Ereignis mit negativer Wahrscheinlichkeit; dieser Umstand widerspricht der Aussage des Axioms 1. Folglich muß $P(\bar{A}) \leqq 1$ sein.

Wir betrachten nun gewisse Ereignisse $A_1, A_2 \ldots, A_k$ einer Massenerscheinung. Wir sagen: Die Ereignisse $A_1, \ldots, A_k$ bilden ein *vollständiges System von Ereignissen,* wenn bei jedem Versuch genau eines der Ereignisse $A_1, \ldots, A_k$ als Ausgang eintreten muß. Ein vollständiges System von Ereignissen ist damit ein System unverträglicher Ereignisse.

[1] Ziffern in eckigen Klammern verweisen auf das Literaturverzeichnis am Schluß dieses Buches.

Beispiele für vollständige Systeme von Ereignissen lassen sich leicht angeben. A und $\bar{A}$ bilden stets ein solches System. Ferner die Ereignisse A_1, $A_2, \ldots, A_6$ beim Würfeln mit einem Würfel, die Ereignisse $K_1, K_2, \ldots K_{13}$ beim Kauf von Herrenschuhen, die Ereignisse $Q_1, \ldots, Q_4$ bei der Qualitätseinstufung von Erzeugnissen und schließlich die Ereignisse $N_1, \ldots, N_4$ bei der Benotung von Arbeitsgängen.

Wir verallgemeinen nun Satz 1.2. zu

Satz 1.5: Bilden $A_1, A_2, \ldots, A_k$ ein vollständiges System von Ereignissen, so gilt

$$\boxed{\sum_{i=1}^{k} P(A_i) = 1} \,. \tag{1.11}$$

Beweis: Definitionsgemäß ist zunächst $\sum\limits_{i=1}^{K} A_i = S$. Weiter sind die Ereignisse A_i $(i = 1, 2, \ldots, k)$ paarweise unverträglich. Daher folgt aus den Axiomen 2 und 3 unmittelbar die Behauptung.

1.2.3. *Die klassische Definition der Wahrscheinlichkeit*

Wir nennen ein System von Ereignissen $A_1, A_2, \ldots, A_n$ ein *System von Chancen*, wenn für alle diese Ereignisse die gleiche Möglichkeit besteht, bei einem Versuch als Ergebnis zu erscheinen. Wir geben ein vollständiges (im Sinne von 1.2.2) System von Chancen vor und betrachten ein beliebiges Ereignis A in der betreffenden Massenerscheinung. Läßt sich A als Summe von genau m Chancen des Systems $A_1, \ldots, A_n$ darstellen, dann besagt die *klassische Definition* für die Wahrscheinlichkeit $P(A)$ von A einfach

$$\boxed{P(A) = \frac{m}{n}} \tag{1.12}$$

oder in Worten: Die Wahrscheinlichkeit für das Eintreten des Ereignisses A ist gleich dem Verhältnis aus der Anzahl der für A günstigen Chancen und der Anzahl der Gesamtchancen (möglichen Chancen für A) [2].

Erläutern wir den Sachverhalt an einem Beispiel. Wir betrachten abermals die Massenerscheinung Würfeln, aber diesmal mit 4 normierten Würfeln und fragen nach der Wahrscheinlichkeit für das Ereignis $A \equiv$ „genau 7 Augen bei einem Wurf". Aus der Kombinatorik ist bekannt, daß die Anzahl der verschiedenen Kombinationen von i Elementen zur j-ten Klasse mit Wiederholung $\binom{i+j-1}{j}$ beträgt. Demzufolge gibt es bei einem Wurf mit 4 normierten Würfeln (wegen $i = 6$ und $j = 4$) genau $\binom{9}{4} = 126$ verschiedene gleichmögliche Ausgänge. Unser vollständiges System von Chancen besteht also aus 126 Chancen, das heißt es ist $n = 126$. Für das Ereignis A kommen offenbar nur die Ereignisse $A_1 =$ „$1 + 1 + 1 + 4$“, $A_2 =$ „$1 + 1 + 2 + 3$“ und $A_3 =$ „$1 + 2 + 2 + 2$“ in Frage, das heißt es gilt $A = A_1 + A_2 + A_3$. Unser Ereignis A läßt sich damit als Summe von genau 3 Chancen darstellen. Gemäß obiger Definition (1.12) erhalten wir

$$P(A) = \frac{3}{126} = \frac{1}{42}.$$

Wir zeigen nun, daß unser statistischer Wahrscheinlichkeitsbegriff unter den hier erforderlichen Einschränkungen mit dem klassischen Begriff identisch ist. Das vollständige System von Chancen ist ein vollständiges System gleichwahrscheinlicher Ereignisse. Wegen $\sum\limits_{i=1}^{n} A_i = S$ und $P(S) = 1$ wird $\sum\limits_{i=1}^{n} P(A_i) = 1$. Nun ist aber wegen der Gleichwahrscheinlichkeit $P(A_1) = \ldots = P(A_n)$ und daher

$$P(A_i) = \frac{1}{n} \quad (i = 1, 2, \ldots, n).$$

Ist A die Summe von m der Ereignisse $A_1, \ldots, A_n$, dann ergibt sich für $P(A)$ nach Satz 1.1 die Summe von m Einzelwahrscheinlichkeiten, die alle gleich $\frac{1}{n}$ sind. Mithin folgt

$$P(A) = m \cdot \frac{1}{n} = \frac{m}{n};$$

das ist aber gerade die Beziehung (1.12).

Der klassischen Definition gebührt in zweifacher Hinsicht Kritik. Einmal geht sie von der Gleichmöglichkeit bestimmter Elementarereignisse aus und liefert nur für solche Ereignisse Wahrscheinlichkeiten, die eine Darstellung als Summe dieser Elementarereignisse zulassen. Zum anderen setzt die Gleichmöglichkeit die Endlichkeit des Systems der Chancen (Elementarereignisse) voraus. Es gibt aber in der Praxis vollständige Systeme von unendlich vielen Elementarereignissen, so daß auch diese Einengung des Wahrscheinlichkeitsbegriffs nicht gerechtfertigt erscheint.

Wir haben in diesem Abschnitt drei Definitionen für den Begriff der Wahrscheinlichkeit eines Ereignisses kennengelernt. Vom mathematischen Standpunkt aus ist der zweiten, also der axiomatischen Definition der Vorrang zu geben, da diese Begriffsbestimmung die Grundlage der modernen Theorie der Wahrscheinlichkeitsrechnung bildet. Die zweite Definition hat jedoch den Nachteil, daß sie eine direkte (unmittelbare) Erfassung (Berechnung) von Wahrscheinlichkeiten realer Ereignisse nicht zuläßt. Wir können gemäß dieser Definition lediglich mittelbar — auf dem Wege über bereits ermittelte Wahrscheinlichkeiten — Wahrscheinlichkeiten praktischer Ereignisse erfassen, indem wir die aus der axiomatischen Begriffsbestimmung resultierenden — und oben zum Teil hergeleiteten — Regeln anwenden. Die klassische Definition läßt dagegen die Ermittlung von Wahrscheinlichkeiten (über die entsprechenden relativen Häufigkeiten) zu. Gemäß (1.3) führt dieser Weg jedoch nur zu einem Näherungswert für die gesuchte Wahrscheinlichkeit, dessen Genauigkeit durch die Stichprobentheorie angegeben wird. Unter den erforderlichen Voraussetzungen ergeben sich die Wahrscheinlichkeiten nach der klassischen Definition numerisch exakt, wobei die Kombinatorik erfolgreiche Anwendung findet. Auf Grund ihres eingeschränkten Gültigkeitsbereiches verwenden wir die klassische Definition im folgenden jedoch nur selten. Wir verwenden deshalb vorwiegend die statistische Definition der Wahrscheinlichkeit [2].

1.3. Der Multiplikationssatz

1.3.1. Der Begriff der bedingten Wahrscheinlichkeit

Um den Begriff der bedingten Wahrscheinlichkeit zu prägen, gehen wir von einem Beispiel aus: In einem Betrieb werden Damenstrümpfe hergestellt. Für den Betrieb ist das Ereignis $A \equiv$ „ein Strumpf gehört zur 1. Wahl" von großem Interesse. Dieses Ereignis A kann nun unter verschiedenen Bedingungen eintreten, einmal unter der Bedingung $B_1 \equiv$ „alle hergestellten Strümpfe werden betrachtet", zum anderen unter der Bedingung $B_2 \equiv$ „nur die brauchbaren Strümpfe werden betrachtet". Die Größe $P\,(A)$, das heißt die Wahrscheinlichkeit des Ereignisses A ist offenbar abhängig („bedingt") von der Bedingung B_1 bzw. B_2. Es gibt also in diesem Zusammenhang keine Größe $P\,(A)$ schlechthin, sondern nur in Abhängigkeit von einer der Bedingungen. Diesen Sachverhalt deuten wir durch die Schreibweise $P_{B_1}\,(A)$ und $P_{B_2}\,(A)$ an. Wir nennen diese Wahrscheinlichkeiten *bedingt*. Zwischen den hier angegebenen bedingten Wahrscheinlichkeiten besteht offenbar die Relation

$$P_{B_1}\,(A) \leqq P_{B_2}\,(A)\ ;$$

das Gleichheitszeichen gilt genau dann, wenn der Betrieb keinen Ausschuß produziert.

Im Grunde genommen ist jede Wahrscheinlichkeit bedingt, da ein Ereignis stets unter gewissen Bedingungen stattfindet. Wir wollen die Wahrscheinlichkeit eines Ereignisses nichtbedingt oder *unbedingt* nennen, wenn das Ereignis unter einem breiten und beständigen Komplex von Bedingungen auftritt, für die wir uns nicht besonders interessieren.

Um zu weiteren Erkenntnissen zu gelangen, verweilen wir noch bei dem letzten Beispiel. Wir betrachten die Ereignisse $K_i \equiv$ „ein Strumpf gehört zur i-ten Güteklasse" ($i = 1, 2, 3, 4$) und $K_j{}^* \equiv$ „ein brauchbarer Strumpf gehört zur j-ten Wahl" ($j = 1, 2, 3$); dabei soll es im Betrieb die 4 Qualitätsstufen (Güteklassen) „I. Wahl", „II. Wahl", III. Wahl",, und „Ausschuß" geben. Gelten zum Beispiel für die K_i ($i = 1,2,3,4$) die relativen Häufigkeiten

$$h_1 = 0{,}70;\ h_2 = 0{,}15;\ h_3 = 0{,}10;\ h_4 = 0{,}05,$$

dann betragen die relativen Häufigkeiten für die Ereignisse $K_j{}^*$ ($j = 1, 2, 3$)

$$h_1^* = \frac{0{,}70}{0{,}95}\ ;\quad h_2^* = \frac{0{,}15}{0{,}95}\ ;\quad h_3^* = \frac{0{,}10}{0{,}95}.$$

Offenbar besteht $h_j < h_j^*$ ($j = 1, 2, 3$). Dagegen ist aber

$$\frac{h_i^*}{h_j^*} = \frac{h_i}{h_j} \quad (i, j = 1, 2, 3), \tag{*}$$

was unmittelbar einzusehen ist. Das heißt: Beim Übergang vom System der Ereignisse K_i ($i = 1, 2, 3, 4$) zu dem System der Ereignisse K_j^* ($j = 1, 2, 3$) ändern sich zwar die relativen Häufigkeiten der in beiden Systemen vorhandenen Ereignisse; der Quotient der relativen Häufigkeiten jeweils zweier dieser Ereignisse bleibt aber bei diesem Übergang unverändert.

Wir gehen nun vermöge der statistischen Definition der Wahrscheinlichkeit von den relativen Häufigkeiten zu den entsprechenden Wahrscheinlichkeiten über; dann ergibt sich wegen der Tatsache, daß $K_i = K_i^*$ ($i = 1, 2, 3$) unter der Bedingung $B =$ „der Strumpf ist verwendungsfähig" ist, aus (*) sofort

$$\frac{P_B(K_i)}{P_B(K_j)} = \frac{P(K_i)}{P(K_j)} \quad (i, j = 1, 2, 3)$$

Die voranstehenden Überlegungen verallgemeinern wir nun. Gegegen sei ein vollständiges System von Ereignissen $A_1, A_2, \ldots, A_m, \ldots, A_n$. Wir verengen dieses System, das heißt, wir verändern die Bedingungen so, daß gewisse Ereignisse dieses Systems unter den neuen Bedingungen unmöglich werden. Durch Hinzunahme der Bedingung B zum ursprünglichen Komplex mögen die Ereignisse $A_{m+1}, \ldots, A_n$ des obigen Systems unmöglich werden. Demzufolge gilt

$$P_B(A_{m+1}) = \ldots = P_B(A_n) = 0\,.$$

Unter den neuen Bedingungen erhalten wir dann das vollständige System $A_1, \ldots, A_m$ mit $P_B(A_1) + \ldots + P_B(A_m) = 1$ und

$$\boxed{\frac{P_B(A_i)}{P_B(A_j)} = \frac{P(A_i)}{P(A_j)} \quad (i, j = 1, 2, \ldots, m)}\,. \tag{1.13}$$

Wir betrachten hierzu ein Beispiel. Bei der Massenerscheinung „Würfeln mit einem normierten Würfel" haben wir bekanntlich das vollständige System $A_1, A_2, \ldots, A_6$ mit $P(A_k) = \frac{1}{6}$ ($k = 1, 2, \ldots 6$). Unter der Bedingung $B \equiv$ „gerade Augenzahl bei einem Wurf" ergibt sich das neue vollständige System A_2, A_4, A_6 mit den bedingten Wahrscheinlichkeiten $P_B(A_l) = \frac{1}{3}$ ($l = 2, 4, 6$). Die Beziehung (1.13) hat hier die Form

$$\frac{P_B(A_i)}{P_B(A_j)} = \frac{P(A_i)}{P(A_j)} \quad (i, j, = 2, 4, 6)\,,$$

die offenbar richtig ist.

Wir leiten nun die Formel für die bedingte Wahrscheinlichkeit ab. Zuvor behandeln wir folgendes Beispiel: In zwei Werken werden Glühlampen hergestellt; im ersten Werk gehören 80 Prozent, im zweiten nur 75 Prozent der Lampen zur Güteklasse 1.

Die im Handel erhältlichen Glühlampen entstammen zu 60 Prozent dem zweiten und zu 40 Prozent dem ersten Werk. Wie groß ist die bedingte Wahrscheinlichkeit dafür, eine Lampe des ersten Werkes zu kaufen, falls diese zur Güteklasse 1 gehört?

Es handelt sich bei dieser Aufgabe um die Ereignisse $A \equiv$ „Die Lampe stammt aus dem ersten Werk" und $B \equiv$ „die Lampe gehört zur Güteklasse 1" sowie um die bedingte Wahrscheinlichkeit $P_B(A)$, die wir wie folgt gewinnen: Von 100 Lampen sind im Mittel 40 aus dem ersten und 60 aus dem zweiten Werk. Unter diesen 40 bzw. 60 Glühlampen sind durchschnittlich 32 bzw. 45 von der Güteklasse 1. Folglich sind unter 100 Lampen im Mittel $32 + 45 = 77$ von der Güteklasse 1 und davon 32 aus dem ersten Werk. Nach der statistischen Definition erhalten wir dann das Ergebnis

$$P_B(A) = \frac{32}{77} \approx 0{,}416.$$

Etwa 41,6 Prozent der zur Güteklasse 1 gehörigen Glühlampen stammen also aus dem ersten Werk.

Wir leiten nun eine allgemeine Formel ab. Zu diesem Zweck betrachten wir zwei verschiedene Ereignisse A und B (mit $P(B) \neq 0$) und konstruieren das System AB, $\bar{A}B$, $\bar{B}$ das vollständig ist. Denn bei jedem Versuch tritt entweder B oder $\bar{B}$ auf, weiter tritt B entweder mit A oder mit $\bar{A}$ auf. Wir verengen nun dieses System durch den Ausschluß des Ereignisses $\bar{B}$, d. h. wir stellen die zusätzliche Bedingung, daß B eintritt. Dann ergibt sich das verengte System AB, $\bar{A}B$ mit den (durch B) bedingten Wahrscheinlichkeiten $P_B(AB)$ und $P_B(\bar{A}B)$. Wegen der Relation (1.13) gilt

$$\frac{P(\bar{A}B)}{P(AB)} = \frac{P_B(\bar{A}B)}{P_B(AB)}$$

oder

$$\frac{P(\bar{A}B) + P(AB)}{P(AB)} = \frac{P_B(\bar{A}B) + P_B(AB)}{P_B(AB)}$$

Nun ist aber $\bar{A}B + AB = B$, so daß

$$\frac{P(B)}{P(AB)} = \frac{P_B(B)}{P_B(AB)}$$

bleibt. Gemäß der Definition für die bedingte Wahrscheinlichkeit eines Ereignisses ergeben sich weiter die Vereinfachungen $P_B(B) = 1$ und $P_B(AB) = P_B(A)$. Damit haben wir schon die *Formel für die bedingte Wahrscheinlichkeit* oder *Satz 1.6:* Die bedingte Wahrscheinlichkeit $P_B(A)$ berechnet sich nach der Formel

$$\boxed{P_B(A) = \frac{P(AB)}{P(B)}, \quad (P(B) \neq 0)} \tag{1.14}$$

aus den unbedingten Wahrscheinlichkeiten $P(AB)$ und $P(B)$.

Wir haben schon in vorstehenden Beispielen bedingte Wahrscheinlichkeiten berechnet, die wir nun nach der Regel (1.14) bestimmen. Im letzten

Beispiel (mit den Lampen) erhalten wir wegen $P(AB) = 0{,}32$ und $P(B) = 0{,}77$ sofort das Ergebnis $P_B(A) = \frac{32}{77} = 0{,}416$. Im vorangegangenen Würfelbeispiel gelten wegen $P(A_2 B) = P(A_4 B) = P(A_6 B) = \frac{1}{6}$ und $P(B) = \frac{1}{2}$ die uns schon geläufigen Werte $P_B(A_2) = P_B(A_4) = P_B(A_6) = \frac{1}{3}$.

1.3.2. Der Multiplikationssatz

Die Formel (1.14) dient in der Praxis meist dazu, die Wahrscheinlichkeit $P(AB)$ aus den (aus der jeweiligen Aufgabenstellung her bekannten) Werten für die Wahrscheinlichkeiten $P(B)$ und $P_B(A)$ zu berechnen. Multiplizieren wir (1.14) mit $P(B) \neq 0$ und vertauschen wir in der so erhaltenen Beziehung die Rollen von A und B, dann ergibt sich wegen $AB = BA$ insgesamt der Multiplikationssatz für die Wahrscheinlichkeiten der gemeinsam auftretenden Ereignisse A und B zu

$$\boxed{P(AB) = P(A) \cdot P_A(B) = P(B) \cdot P_B(A)} \tag{1.15}$$

oder

Satz 1.7: Die Wahrscheinlichkeit für das gemeinsame Auftreten zweier Ereignisse ist gleich der Wahrscheinlichkeit des einen Ereignisses multipliziert mit der (bedingten) Wahrscheinlichkeit des anderen Ereignisses unter der Voraussetzung, daß das erste Ereignis bereits stattgefunden hat.

Die Formel (1.15) gilt im Vergleich zu (1.14) auch für $P(B) = 0$ und (A und B vertauscht) $P(A) = 0$.

Erläutern wir nun den Multiplikationssatz an zwei Beispielen. Zunächst betrachten wir einen Betrieb, in dem 96 Prozent der hergestellten Erzeugnisse absatzfähig sind. Von jeweils 100 absatzfähigen Erzeugnissen gehören im Mittel 75 zur Sorte 1. Wie groß ist die Wahrscheinlichkeit dafür, daß ein im Betrieb hergestelltes Erzeugnis zur Sorte 1 gehört? Wir haben hier die Ereignisse $A \equiv$ „das Erzeugnis ist absatzfähig“ und $B \equiv$ „das Erzeugnis gehört zur Sorte 1“ mit $P(A) = 0{,}96$ und $P_A(B) = 0{,}75$. Aus dem ersten Teil der Gleichung (1.15) erhalten wir das Ergebnis für $P(AB)$ zu

$$P(AB) = 0{,}96 \cdot 0{,}75 = 0{,}72\,.$$

Nun die zweite Aufgabe. Ein Fünftel aller Kunden, die einen Industrieladen aufsuchen, begibt sich in die Schuhabteilung. Im Durchschnitt kaufen 55 Prozent dieser Kunden in dieser Abteilung Damenschuhe. Wie groß ist die Wahrscheinlichkeit dafür, daß ein Kunde, der den Laden betritt, in der Schuhabteilung Damenschuhe kauft? Analog zur ersten Aufgabe liegen die Ereignisse $A \equiv$ „ein Kunde besucht die Schuhabteilung“ und $B \equiv$ „ein Kunde kauft Damenschuhe“ mit $P(A) = 0{,}20$ und $P_A(B) = 0{,}55$ vor. Die gesuchte Größe $P(AB)$ wird vermöge (1.15)

$$P(AB) = 0{,}20 \cdot 0{,}55 = 0{,}11\,.$$

Wir spezifizieren nun den Satz 1.7. Zu diesem Zweck führen wir einen neuen Begriff ein. Wir betrachten zwei Ereignisse A und B, deren Wahrscheinlichkeiten $\neq 0$ sind, und nennen B *unabhängig von* A, wenn B in keiner Weise durch das Ereignis A beeinflußt wird, wenn also $P_A(B) = P(B)$ gilt. Ist nun B unabhängig von A, so offenbar auch A unabhängig von B; denn aus den rechtsseitigen Gleichungen von (1.15) folgt wegen $P_A(B) = P(B) \neq 0$ sofort $P(A) = P_B(A)$ und damit die Unabhängigkeit des Ereignisses A von B.

Sind nun A und B voneinander unabhängig, d. h. bestehen die Relationen $P(A) = P_B(A)$ und $P(B) = P_A(B)$, so ergibt sich aus (1.15) der *Multiplikationssatz* für die Wahrscheinlichkeiten der gemeinsam auftretenden und voneinander unabhängigen Ereignisse A und B zu

$$\boxed{P(AB) = P(A) \cdot P(B)} \tag{1.16}$$

Satz 1.8: Die Wahrscheinlichkeit für das gemeinsame Auftreten zweier voneinander unabhängiger Ereignisse ist gleich dem Produkt der Wahrscheinlichkeiten der einzelnen Ereignisse.

Induktiv läßt sich das vorstehende Ergebnis verallgemeinern zum

Satz 1.9: Sind $A_1, \ldots, A_n$ paarweise voneinander unabhängige Ereignisse, dann gilt für das gemeinsame Auftreten dieser n Ereignisse[1]

$$\boxed{P(A_1 A_2 \ldots A_n) = \prod_{j=1}^{n} P(A_j)} \,. \tag{1.17}$$

Wir geben nun zwei Beispiele. Das erste entnehmen wir der Arbeit [3]. Ein Arbeiter bedient drei Webstühle, die unabhängig voneinander arbeiten. Die Wahrscheinlichkeit dafür, daß ein Webstuhl im Laufe einer Stunde die Aufmerksamkeit eines Arbeiters nicht erfordert, ist bekannt und zwar 0,9 für den ersten, 0,8 für den zweiten und 0,85 für den dritten. Wie groß ist die Wahrscheinlichkeit dafür, daß im Laufe einer Stunde keiner der drei Webstühle die Wartung durch den Arbeiter beansprucht? Ist A_i das Ereignis „der i-te Webstuhl arbeitet im Laufe einer Stunde ungestört" ($i = 1, 2, 3$), dann lauten die zugehörigen Wahrscheinlichkeiten $P(A_1) = 0{,}9$, $P(A_2) = 0{,}8$ und $P(A_3) = 0{,}85$. Die gesuchte Größe $P(A_1 A_2 A_3)$ ergibt sich nach (1.17) für $n = 3$ zu

$$P(A_1 A_2 A_3) = 0{,}9 \cdot 0{,}8 \cdot 0{,}85 = 0{,}612.$$

Nun zum zweiten Beispiel. Statistische Untersuchungen innerhalb eines Betriebes haben ergeben, daß während einer bestimmten Zeitspanne die Elektromotoren zum Antrieb der Werkzeugmaschinen nur mit einer Wahrscheinlichkeit von 40 Prozent in Betrieb sind. Wir betrachten zwei dieser Motoren und stellen folgende Fragen (immer bezogen auf die erwähnte Zeitspanne).

[1] Ist $n = \infty$, dann muß die rechte Seite von (1.17) konvergent sein!

a) Wie groß ist die Wahrscheinlichkeit p_1 dafür, daß beide Motoren gleichzeitig arbeiten?

b) Wie groß ist die Wahrscheinlichkeit p_2 dafür, daß mindestens einer der beiden Motoren in Betrieb ist?

c) Wie groß ist die Wahrscheinlichkeit p_3 dafür, daß beide Motoren still stehen?

d) Wie groß ist die Wahrscheinlichkeit p_4 dafür, daß nur einer der beiden Motoren arbeitet?

Wir führen die Bezeichnungen $A \equiv$ „der erste Motor arbeitet" und $B \equiv$ „der zweite Motor arbeitet" ein. Dann gilt zunächst $P(A) = P(B) = 0{,}4$. Die beiden Motoren sollen unabhängig voneinander in Betrieb sein. Die Ereignisse A und B sind also voneinander unabhängig. Weiter benutzen wir folgenden Sachverhalt, den wir im Anschluß an dieses Beispiel beweisen: Sind A und B unabhängig voneinander, dann auch die Paare A und $\bar{B}$, $\bar{A}$ und B sowie $\bar{A}$ und $\bar{B}$. Nun sind wir in der Lage, die Größen p_i ($i = 1, 2, 3, 4$) der Reihe nach zu berechnen; dabei wenden wir den Satz 1.8 an. Zunächst ist $p_1 = P(AB) = 0{,}4 \cdot 0{,}4 = 0{,}16$.

Betrachten wir das Ereignis $C \equiv$ „genau ein Motor arbeitet", dann gilt offenbar $p_2 = P(C + AB)$ oder wegen der Unverträglichkeit von C und AB nach dem Satz 1.1 (Additionssatz) $p_2 = P(C) + P(AB)$. $P(AB)$ ist bereits zu 0,16 ermittelt, bleibt also noch $P(C)$ zu berechnen. Nun können wir aber schreiben $C = A\bar{B} + \bar{A}B$ oder wieder nach Satz 1.1 $P(C) = P(A\bar{B}) + P(\bar{A}B)$. Wenden wir auf beide Glieder der rechten Seite der letzten Gleichung Satz 1.8 mit $P(A) = P(B) = 0{,}4$ und $P(\bar{A}) = P(\bar{B}) = 0{,}6$ an, dann ergibt sich $P(C) = 0{,}48$ und damit schließlich $p_2 = 0{,}48 + 0{,}16 = 0{,}64$. Die dritte Größe p_3 erhalten wir sofort zu $p_3 = P(\bar{A}\bar{B}) = 0{,}6 \cdot 0{,}6 = 0{,}36$ unter Beachtung von Satz 1.8. Die letzte Wahrscheinlichkeit p_4 wurde bereits zu $p_4 = 0{,}48$ ermittelt, denn es besteht die Beziehung $p_4 = P(C)$. Die Frage b) hätten wir auch so beantworten können: Offenbar ist $p_2 = 1 - P(\bar{A}\bar{B}) = 0{,}64$; denn es arbeitet genau dann mindestens ein Motor, wenn nicht beide stillstehen.

Nachträglich beweisen wir den

Satz 1.10: Sind A und B voneinander unabhängig, so auch die Paare A und $\bar{B}$, $\bar{A}$ und B sowie $\bar{A}$ und $\bar{B}$.

Beweis: Wegen der Voraussetzung können wir

$$P(B) = P_A(B) = 1 - P_A(\bar{B})$$

schreiben; andererseits ist $P(B) = 1 - P(\bar{B})$. Folglich gilt $P(\bar{B}) = P_A(\bar{B})$. Damit ist der erste Teil des Satzes bewiesen. Der zweite Teil wird ganz entsprechend behandelt. Wir wenden uns dem letzten Fall zu. Auf Grund des schon bewiesenen Teils des vorliegenden Satzes ist $P(A) = P_B(A) = P_{\bar{B}}(A)$ Daher folgt aus $P(\bar{A}) = 1 - P(A) = 1 - P_{\bar{B}}(A) = 1 - (1 - P_{\bar{B}}(\bar{A})) = P_{\bar{B}}(\bar{A})$ kurzerhand die Behauptung $P(\bar{A}) = P_{\bar{B}}(\bar{A})$.

Wir geben nun noch eine Verallgemeinerung der letzten Aussage an.

Satz 1.11: Sind die Ereignisse $A_1, A_2, \ldots, A_n$ paarweise voneinander unabhängig, so auch die Ereignisse $\bar{A}_{i_1}, \bar{A}_{i2}, \ldots, \bar{A}_{i_m}, A_{i_{m+1}}, \ldots, A_{i_n}$; dabei ist $i_1, i_2, \ldots, i_n$ eine beliebige Permutation der Zahlen 1, 2, ..., und $0 \leqq m \leqq n$.

1.3.3. *Folgerungen aus dem Multiplikationssatz*

Wir betrachten die paarweise voneinander unabhängigen Ereignisse $A_1, \ldots, A_n$. Dann sind nach Satz 1.11 auch die Ereignisse $\bar{A}_1, \ldots, \bar{A}_n$ paarweise voneinander unabhängig. Das Ereignis $A = A_1 + \ldots + A_n$ („mindestens eines der Ereignisse $A_1, \ldots, A_n$") besitzt offenbar das entgegengesetzte Ereignis $\bar{A} = \bar{A}_1 \bar{A}_2 \ldots \bar{A}_n$ („keines der Ereignisse $A_1, \ldots, A_n$"). Mithin gilt

$$P(A_1 + \ldots + A_n) = 1 - P(\bar{A}_1 \ldots \bar{A}_n)$$

Wenden wir auf die rechts stehende Wahrscheinlichkeit $P(\bar{A}_1 \bar{A}_2 \ldots \bar{A}_n)$ den Satz 1.9 an, so erhalten wir den *Additionssatz* für die Wahrscheinlichkeiten voneinander unabhängiger Ereignisse zu

$$\boxed{P(A_1 + A_2 + \ldots + A_n) = 1 - \prod_{j=1}^{n} (1 - P(A_j))} \qquad (1.18)$$

oder den

Satz 1.12: Die Wahrscheinlichkeit für das Auftreten irgendeines Ereignisses einer Menge paarweise voneinander unabhängiger Ereignisse ergibt sich, indem die Zahl 1 um das Produkt der Wahrscheinlichkeiten aller Komplementärereignisse der betrachteten Menge vermindert wird.[1]

Für $n = 2$ spezialisiert sich (1.18) zu

$$P(A + B) = 1 - [1 - P(A)] \cdot [1 - P(B)]$$

oder zu $$P(A + B) = P(A) + P(B) - P(AB) \qquad (1.19)$$

Sind die Ereignisse A und B unverträglich, d. h. gilt $P(AB) = 0$, dann ergibt sich aus (1.19) die uns schon bekannte Formel (1.8).

Wir betrachten nun zur Erläuterung der vorstehenden Aussagen zwei Beispiele. Das erste Beispiel handelt von zwei unabhängig voneinander arbeitenden Aggregaten G_1 und G_2 eines Betriebes. Durch umfangreiche Untersuchungen ist festgestellt worden, daß im Laufe einer Woche die Wahrscheinlichkeit für den Ausfall von G_1 0,3 und von G_2 0,2 beträgt. Wie groß ist nun die Wahrscheinlichkeit dafür, daß mindestens ein Aggregat im Laufe einer Woche ausfällt? Wir setzen $A_i \equiv$ „das Aggregat G_i fällt in der Woche aus" ($i = 1{,}2$) und haben $P(A_1) = 0{,}3$, $P(A_2) = 0{,}2$. Die Ereignisse A_1 und A_2 sind voneinander unabhängig, so daß nach (1.17) für $n = 2$ folgt $P(A_1 A_2) = 0{,}06$. Wenden wir nun die Formel (1.18) für $n = 2$ (oder (1.19)!) an, so

[1] Umfaßt die in Rede stehende Menge unendlich viele Elemente, dann muß das unendliche Produkt in (1.18) konvergent sein.

erhalten wir das gesuchte Ergebnis zu $P(A_1 + A_2) = 0{,}3 + 0{,}2 - 0{,}06 = 0{,}44$. In 44 Prozent aller Fälle erhält mindestens ein Aggregat im Laufe einer Woche einen Defekt.

Das zweite Beispiel handelt von einem normierten Würfel. p_1 sei die Wahrscheinlichkeit dafür, mit einem Würfel bei 4 Würfen eine „6" zu würfeln; p_2 sei dagegen die Wahrscheinlichkeit dafür, mit zwei Würfeln bei 24 Würfen mindestens eine „doppelte 6" zu erzielen. Welche der Wahrscheinlichkeiten p_i $(i = 1, 2)$ ist die größere?

Wir führen die Bezeichnungen $E_i \equiv$ „beim i-ten Wurf mit einem Würfel eine „6" $(i = 1, 2, 3, 4)$ und $Z_j \equiv$ „beim j-ten Wurf mit 2 Würfeln eine doppelte „6" $(j = 1, 2, \ldots, 24)$ ein. Klar ist $P(E_i) = \frac{1}{6}$ $(i = 1, 2, 3, 4)$. Daher erhalten wir gemäß (1.18) für $n = 4$

$$p_1 = P(E_1 + E_2 + E_3 + E_4) = 1 - \left(1 - \frac{1}{6}\right)^4 \approx 0{,}516\,.$$

Ganz analog berechnen wir p_2. Nach dem Multiplikationssatz (1.16) ergibt sich zunächst $P(Z_j) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$ $(j = 1, 2 \ldots 24)$. Wenden wir wieder (1.18) für $n = 24$ an, so folgt schon

$$p_2 = P(Z_1 + \ldots + Z_{24}) = 1 - \left(1 - \frac{1}{36}\right)^{24} \approx 0{,}491\,.$$

Die erste der ermittelten Wahrscheinlichkeiten ist also die größere.

Wir ziehen nun eine weitere Folgerung aus dem Multiplikationssatz.

Gegeben sei zu diesem Zweck ein vollständiges System von Ereignissen $A_1, A_2, \ldots, A_n$ und ein beliebiges Ereignis B. Wir fragen nach der Wahrscheinlichkeit für das Auftreten des Ereignisses B schlechthin, d. h. unabhängig davon, mit welchem der A_i $(i = 1, 2, \ldots, n)$. Diese Wahrscheinlichkeit nennen wir die *totale* oder die *vollständige Wahrscheinlichkeit* für das Ereignis B. Wir leiten jetzt eine Formel für diese Größe $P(B)$ ab. Zunächst gilt auf Grund des Satzes 1.1

$$P(B) = P(A_1 B + \ldots + A_n B) = \sum_{i=1}^{n} P(A_i B)$$

und weiter nach dem Satz 1.7

$$\boxed{P(B) = \sum_{i=1}^{n} P(A_i)\, P_{A_i}(B)}\,. \tag{1.20}$$

Satz 1.13: Die vollständige oder totale Wahrscheinlichkeit für das Ereignis B bezüglich des vollständigen Systems $A_1, \ldots, A_n$ berechnet sich nach der Formel (1.20).

Wir erläutern den Sachverhalt wieder an zwei Beispielen. Zunächst ein Beispiel aus der Landwirtschaft: Für die Aussaat von Weizen wird ein Gemenge von vier Sorten S_i $(i = 1, 2, 3, 4)$ mit den Anteilen 96 Prozent, 1 Prozent, 2 Prozent und 1 Prozent verwendet. Die Wahrscheinlichkeit dafür, daß aus einem Korn der Sorte S_i $(i = 1, 2, 3, 4)$ eine Ähre mit mindestens 50 Körnern wird, sei der Reihe nach $p_1 = 0{,}5$, $p_2 = 0{,}15$, $p_3 = 0{,}2$ und $p_4 = 0{,}05$. Wie groß ist die Wahrscheinlichkeit dafür, daß aus einem beliebigen Korn des Gemenges eine Ähre mit mindestens 50 Körnern wird? Wir setzen $A_i \equiv$ „ein Korn gehört zur Sorte S_i" $(i = 1, 2, 3, 4)$ und $B \equiv$ „das Korn wird zu einer Ähre mit mindestens 50 Körnern". Die Ereignisse $A_1, \ldots, A_4$ bilden ein vollständiges System mit den einzelnen Wahrscheinlichkeiten $P(A_1) = 0{,}96$, $P(A_2) = 0{,}01$, $P(A_3) = 0{,}02$ und $P(A_4) = 0{,}01$. Weiter sind uns die bedingten Wahrscheinlichkeiten $P_{A_1}(B) = 0{,}5$, $P_{A_2}(B) = 0{,}15$, $P_{A_3}(B) = 0{,}20$ und $P_{A_4}(B) = 0{,}05$ bekannt. Die Regel (1.20) liefert dann das Ergebnis zu $P(B) = 0{,}486$. Zu dem Ereignis B gehört damit die totale Wahrscheinlichkeit von 48,6 Prozent (vgl. [3]).

Das zweite Beispiel stammt aus dem Handel. In der Abteilung Damenkonfektion eines Kaufhauses wird nach langjährigen Erfahrungen folgendes Sortiment geführt:

Größenklasse	Anteil am Gesamtsortiment	Änderungswahrscheinlichkeiten
groß (g)	0,26	0,18
mittel (m)	0,44	0,10
klein (k)	0,30	0,15

Des weiteren ist bekannt, mit welcher Wahrscheinlichkeit nach dem Verkauf Änderungen am Kleidungsstück notwendig sind. Die betreffenden Werte sind in der letzten Spalte des obigen Schemas angegeben. Von Interesse ist die Wahrscheinlichkeit dafür, daß beim Verkauf eines beliebigen Kleidungsstückes eine Änderung erforderlich ist! Wir betrachten die Ereignisse $A_i \equiv$ „Kleidungsstück gehört zur i-ten Größenklasse" $(i = 1, 2, 3)$ und $B \equiv$ „am Kleidungsstück sind Änderungen nötig". Die Ereignisse A_i $(i = 1, 2, 3)$ bilden ein vollständiges System von Ereignissen mit den Wahrscheinlichkeiten $P(A_1) = 0{,}26$, $P(A_2) = 0{,}44$ und $P(A_3) = 0{,}30$. Ferner sind die bedingten Wahrscheinlichkeiten $P_{A_1}(B) = 0{,}18$, $P_{A_2}(B) = 0{,}10$ und $P_{A_3}(B) = 0{,}15$ bekannt. Die vollständige Wahrscheinlichkeit für B ergibt sich dann aus (1.20) zu

$$P(B) = 0{,}26 \cdot 0{,}18 + 0{,}44 \cdot 0{,}10 + 0{,}30 \cdot 0{,}15 \approx 0{,}136\,.$$

Wir ziehen nun die letzte Folgerung aus dem Multiplikationssatz: Gegeben sei wieder ein vollständiges System von Ereignissen $A_1, A_2, \ldots, A_n$ und ein beliebiges Ereignis B. Aus der Gleichung (1.15) folgt

$$P(B) \cdot P_B(A_i) = P(A_i)\, P_{A_i}(B) \; (i = 1, 2, \ldots, n)\,.$$

Ersetzen wir nun im letzten Ausdruck $P(B)$ durch die rechte Seite von (1.20), so erhalten wir das Theorem von BAYES:

$$\boxed{P_B(A_i) = \frac{P(A_i)\, P_{A_i}(B)}{\sum_{j=1}^{n} P(A_j)\, P_{A_j}(B)} \quad (i = 1, 2 \ldots, n)}\,, \tag{1.21}$$

das in der Praxis häufig Anwendung findet. Damit erhalten wir den *Satz 1.14:* Die bedingte Wahrscheinlichkeit eines zu einem vollständigen System gehörigen Ereignisses in bezug auf ein beliebiges Ereignis wird nach der BAYESschen Regel (1.21) ermittelt.

Belegen wir den vorstehenden Satz mit einem Beispiel aus der Industrie: In einem Betrieb seien 96 Prozent aller Erzeugnisse normgerecht. Bei der Gütekontrolle wird ein Kontrollsystem benutzt, das 98 Prozent der normgerechten Erzeugnisse als normgerecht und 95 Prozent nicht normgerechten Erzeugnisse als nicht normgerecht erkennt. Um eine größere Genauigkeit zu erzielen, werden alle Erzeugnisse nach dem obigen Verfahren doppelt geprüft. Wie groß ist die Wahrscheinlichkeit dafür, daß ein bei der doppelten Kontrolle als normgerecht ausgewiesenes Erzeugnis in der Tat der Norm genügt? Die Ereignisse $A_1 \equiv$ „Erzeugnis ist normgerecht" und $A_2 \equiv$ „Erzeugnis genügt nicht der Norm" bilden ein vollständiges System mit den Wahrscheinlichkeiten $P(A_1) = 0{,}96$ und $P(A_2) = 0{,}04$. Weiter betrachten wir die Ereignisse $C_1 \equiv$ „normgerechtes Erzeugnis wird als normgerecht angesehen" und $C_2 \equiv$ „nicht normgerechtes Erzeugnis wird als normgerecht ausgewiesen" mit $P(C_1) = 0{,}98$ und $P(C_2) = 1 - 0{,}95 = 0{,}05$. Von besonderem Interesse ist das Ereignis $B \equiv$ „Erzeugnis besteht die doppelte Kontrolle". Gesucht ist die bedingte Wahrscheinlichkeit $P_B(A_1)$. Zunächst berechnen wir auf Grund der Relation (1.16) die bedingten Wahrscheinlichkeiten $P_{A_1}(B) = 0{,}98^2$ und $P_{A_2}(B) = 0{,}05^2$, denn die beiden auszuführenden Kontrollen sind unabhängig voneinander. Für die gesuchte Wahrscheinlichkeit $P_B(A_1)$ folgt dann aus (1.21)

$$P_B(A_1) = \frac{0{,}96 \cdot 0{,}98^2}{0{,}96 \cdot 0{,}98^2 + 0{,}04 \cdot 0{,}05^2} \approx 0{,}9999\,.$$

Wird also das oben charakterisierte Kontrollverfahren angewendet, dann ist damit zu rechnen, daß von 10000 für normgerecht befundenen Erzeugnissen lediglich ein Erzeugnis der Norm nicht entspricht.

1.4. Vermischte Aufgaben nebst Lösungen

1.4.1. Aufgaben

1. Wie groß ist die Wahrscheinlichkeit dafür, mit vier normierten Würfeln bei einem Wurf 10 Augen zu erzielen?

2. Wie groß ist die Wahrscheinlichkeit dafür, mit zwei Würfeln bei einem Wurf mindestens 7 Augen zu würfeln?
3. In einem Ort gibt es drei Industrieläden, die normgerechte Glühlampen mit der relativen Häufigkeit 0,90, 0,80 bzw. 0,85 zum Kauf anbieten. Die Käufer suchen diese Läden mit der Wahrscheinlichkeit 0,40, 0,35 bzw. 0,25 beim Kauf einer Glühlampe auf. Wie groß ist die Wahrscheinlichkeit dafür, daß ein Käufer bei einem Kauf im angegebenen Ort eine normgerechte Glühlampe erhält?
4. Von zwei unabhängigen Ereignissen ist bekannt, daß das erste mit der Wahrscheinlichkeit 0,4 und das zweite mit der Wahrscheinlichkeit 0,7 auftritt.
   a) Mit welcher Wahrscheinlichkeit p_1 tritt mindestens eines der Ereignisse ein?
   b) Mit welcher Wahrscheinlichkeit p_2 bzw. p_3 tritt nur das erste bzw. nur das zweite Ereignis ein?
   c) Mit welcher Wahrscheinlichkeit p_4 treten beide gemeinsam auf?
   d) Mit welcher Wahrscheinlichkeit p_5 ereignet sich keines der beiden Ereignisse?
5. Wie groß ist in der letzten Aufgabe des Abschnittes 1.3.3. die Wahrscheinlichkeit dafür, daß ein beliebiges Erzeugnis die doppelte Kontrolle besteht?

1.4.2. Lösung der Aufgaben

1. Zur Berechnung der Wahrscheinlichkeit des Ereignisses $A =$ „10 Augen bei einem Wurf mit 4 Würfeln" gehen wir wie auf Seite 20 vor. Ein Wurf mit 4 Würfeln hat $n = 126$ gleich mögliche Ausgänge (Anzahl der Kombinationen von 6 Elementen zur 4. Klasse mit Wiederholung $= \binom{6+4-1}{4} = \binom{9}{4} = 126$). Lediglich $n = 8$ Ausgänge sind für das Ereignis A günstig, nämlich

$$\begin{aligned} A_1 &= \text{„}6+2+1+1\text{“}, & A_5 &= \text{„}4+2+2+2\text{“}, \\ A_2 &= \text{„}5+3+1+1\text{“}, & A_6 &= \text{„}4+4+1+1\text{“}, \\ A_3 &= \text{„}5+2+2+1\text{“}, & A_7 &= \text{„}3+3+3+1\text{“}, \\ A_4 &= \text{„}4+3+2+1\text{“}, & A_8 &= \text{„}3+3+2+2\text{“}, \end{aligned}$$

so daß

$$A = \sum_{i=1}^{8} A_i$$

besteht. Gemäß der klassischen Definition (1.12) haben wir damit das Resultat

$$P(A) = \frac{m}{n} = \frac{8}{126} \approx 0{,}0635\,.$$

2. Ist A das Ereignis „mindestens 7 Augen bei einem Wurf mit zwei Würfeln", dann gilt $\bar{A}$ = „höchstens 6 Augen bei einem Wurf mit zwei Würfeln". Wir berechnen $P(\bar{A})$ analog zu der vorstehenden Aufgabe. Ein Wurf mit 2 Würfeln hat $n = 21$ gleichmögliche Ausgänge (Anzahl der Kombinationen von 6 Elementen zur 2. Klasse mit Wiederholung $= \binom{6+2-1}{2} = \binom{7}{2} = 21$). Nur $m = 9$ dieser Ereignisse sind für $\bar{A}$ günstig, nämlich

$$A_1 = „1 + 1", \quad A_4 = „2 + 2", \quad A_7 = „1 + 5"$$
$$A_2 = „1 + 2", \quad A_5 = „1 + 4", \quad A_8 = „2 + 4",$$
$$A_3 = „1 + 3", \quad A_6 = „2 + 3", \quad A_9 = „3 + 3",$$

so daß

$$\bar{A} = \sum_{i=1}^{9} A_i$$

gilt. Aus (1.12) folgt dann

$$P(\bar{A}) = \frac{m}{n} = \frac{9}{21} \approx 0{,}429 \, .$$

Die gesuchte Wahrscheinlichkeit ist damit zu

$$P(A) \approx 1 - 0{,}429 = 0{,}571$$

gefunden.

3. Wir konstatieren zunächst die Ereignisse A_i = „der Kunde sucht den i-ten Laden auf" ($i = 1, 2, 3$) mit den Wahrscheinlichkeiten $P(A_1) = 0{,}40$, $P(A_2) = 0{,}35$ und $P(A_3) = 0{,}25$. Diese drei Ereignisse bilden offenbar ein vollständiges System. Uns interessiert das Ereignis B „der Kunde erhält eine normgerechte Glühlampe". Bekannt sind die bedingten Wahrscheinlichkeiten

$$P_{A_1}(B) = 0{,}90, \quad P_{A_2}(B) = 0{,}80, \quad P_{A_3}(B) = 0{,}85 \, .$$

Durch Anwendung des Satzes 1.13 erhalten wir das Ergebnis

$$P(B) = 0{,}40 \cdot 0{,}90 + 0{,}35 \cdot 0{,}80 + 0{,}25 \cdot 0{,}85 \approx 0{,}853 \, .$$

4. Wir bezeichnen die in Rede stehenden Ereignisse mit A und B.

Dann gilt zunächst $P(A) = 0{,}4$ und $P(B) = 0{,}7$.

a) Für p_1 erhalten wir gemäß (1.19) und Satz 1.8

$$p_1 = P(A + B) = 0{,}4 + 0{,}7 - 0{,}28 = 0{,}82$$

b$_1$) Bezüglich p_2 ergibt sich aus Satz 1.8

$$p_2 = P(A\bar{B}) = P(A)\, P(\bar{B}) = 0{,}4 \cdot 0{,}3 = 0{,}12 \, .$$

b$_2$) Entsprechend wird

$$p_3 = P(\bar{A}B) = P(\bar{A})\, P(B) = 0{,}6 \cdot 0{,}7 = 0{,}42$$

c) Wir wenden wiederum Satz 1.8 an und erhalten

$$p_4 = p(AB) = p(A) \cdot p(B) = 0{,}4 \cdot 0{,}7 = 0{,}28$$

d) Nochmalige Anwendung von Satz 1.8 liefert

$$p_5 = P(\overline{AB}) = P(\overline{A}) \cdot P(\overline{B}) = 0{,}6 \cdot 0{,}3 = 0{,}18$$

5. Wir notieren zunächst die Ereignisse A_1 = „ein normgerechtes Erzeugnis besteht die doppelte Kontrolle“ und A_2 = „ein nicht normgerechtes Erzeugnis besteht die doppelte Kontrolle“ mit den Wahrscheinlichkeiten $p(A) = 0{,}96$ und $p(A_2) = 0{,}04$, die ein vollständiges System darstellen. Von Interesse ist das Ereignis B = „ein beliebiges Erzeugnis besteht die doppelte Kontrolle“. Bekannt sind die bedingten Wahrscheinlichkeiten $P_{A1}(B) = 0{,}98^2$ und $P_{A2}(B) = 0{,}05^2$. Der Satz 1.13 liefert nun das Ergebnis

$$P(B) = 0{,}96 \cdot 0{,}98^2 + 0{,}04 \cdot 0{,}05^2 \approx 0{,}922\,.$$

2. Zufallsgrößen und deren Verteilungen

2.1. Begriff der Zufallsgröße

Wir betrachten Massenerscheinungen, deren Einzelerscheinungen die Eigenschaft haben, daß sie durch eine gewisse Zahlengröße charakterisiert werden können. Diese Größe kann verschiedenartig sein. Beim Würfeln ist sie die Augenzahl, bei der Überprüfung von Maschinenteilen, die auf Drehmaschinen bearbeitet werden, deren Durchmesser. Diese Größe spiegelt zwar in irgendeiner Weise das Resultat des betreffenden Versuchs, das Ergebnis der betreffenden Beobachtung wider, ändert sich jedoch von Versuch zu Versuch, von Beobachtung zu Beobachtung auf Grund vielfältiger zufälliger Einflüsse; sie kann alle möglichen Werte annehmen. Eine solche Größe nennen wir *Zufallsgröße* oder *-variable* der zugehörigen Massenerscheinung. Um Aussagen über zufällige Größen herleiten zu können, müssen wir ihre *Verteilung*, d. h. die möglichen Werte der Zufallsvariablen sowie die einzelnen Wahrscheinlichkeiten kennen, mit der diese möglichen Werte angenommen werden.

Die erwähnten Beispiele lassen schon zwei verschiedenartige Typen von Zufallsgrößen (und damit auch von Verteilungen) erkennen: die diskreten und die stetigen. Wir nennen eine Zufallsgröße *diskret*, wenn sie nur endlich oder abzählbar unendlich viele Werte der reellen Zahlengeraden annehmen kann. Die Zufallsvariable beim Würfeln ist diskret. Es kommen nur die ganzen Zahlen 1, 2, . . ., 6 in Betracht. Demgegenüber heißt eine Zufallsgröße *kontinuierlich* oder stetig, wenn sie jeden Wert eines bestimmten Intervalls der reellen Zahlengeraden annehmen kann. Der Durchmesser von Drehteilen ist eine stetige zufällige Variable. Handelt es sich etwa um Bolzen mit der durchschnittlichen Stärke von 10 mm, so sind alle Werte beispielsweise zwischen 9,8 und 10,2 mm für den Durchmesser denkbar.

2.2. Diskrete Verteilungen

2.2.1. Verteilungstabelle und Verteilungsfunktion

Wir betrachten eine diskrete Zufallsgröße x mit den möglichen Werten x_1, $x_2, \ldots, x_n$; ohne Beschränkung der Allgemeinheit können wir $x_1 < x_2 < \ldots < x_n$ annehmen. Die Zufallsvariable x möge den Wert x_i $(i = 1, 2., \ldots, n)$ mit der Wahrscheinlichkeit p_i $(i = 1, 2, \ldots, n)$ annehmen; es soll also gelten

$$p_i = P\,(x = x_i)\ (i = 1, 2, \ldots, n)\,.$$

Das Verzeichnis

x	x_1	x_2	x_3	$\ldots$	x_n
p	p_1	p_2	p_3	$\cdots$	p_n

(2.1)

heißt die *Verteilungstabelle der diskreten Zufallsgröße x*. Da $x_1, \ldots, x_n$ alle möglichen Werte von x umfassen, müssen die Ereignisse „$x = x_i$“ ($i = 1, 2, \ldots, n$) ein vollständiges System bilden, es muß also $p_1 + p_2 + \ldots + p_n = 1$ sein. Die Verteilungstabelle (2.1) gibt praktisch an, wie die Zufallsgröße x auf die möglichen Werte „verteilt“ ist.

Wir geben nun zwei Beispiele. Bei der Massenerscheinung Würfeln mit einem normierten Würfel betrachten wir die diskrete Zufallsgröße $x =$ Augenzahl nach einem Wurf. Zu dieser Variablen gehört offenbar die Verteilungstabelle

x	1	2	3	4	5	6
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

In einem Betrieb werden die hergestellten Erzeugnisse in der Hinsicht auf ihre Qualität untersucht, daß zu jedem Erzeugnis die Anzahl seiner Fehler angegeben wird. Diese Anzahl x der Erzeugnisfehler ist offenbar eine diskrete Zufallsvariable. Ist etwa bekannt, daß 90 Prozent der Erzeugnisse keinen, 8 Prozent einen und 2 Prozent zwei Fehler aufweisen, dann ergibt sich für diese Zufallsgröße x die Verteilungstabelle

x	0	1	2
p	0,9	0,08	0,02

Die Verteilungstabelle (2.1) kann auch graphisch dargestellt werden. Tragen wir auf der x-Achse die Werte $x_1, x_2, \ldots, x_n$ der Zufallsvariablen x ab, auf der y-Achse die entsprechenden Wahrscheinlichkeiten $p_1, p_2, \ldots, p_n$ und verbinden wir benachbarte der so erhaltenen Punkte (x, p), dann ergibt sich ein Streckenzug oder *Wahrscheinlichkeitsdiagramm*. In den angeführten Beispielen ($x =$ Augenzahl, $x =$ Anzahl der Fehler) erhalten wir die Wahrscheinlichkeitsdiagramme

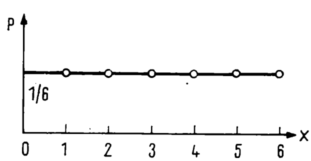

Abb. 1 Diagramm der Augenzahl

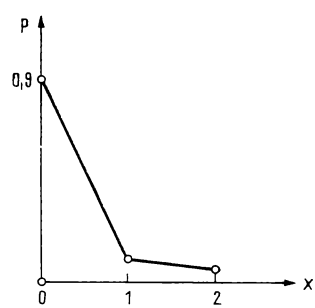

Abb. 2 Diagramm der Fehleranzahl

Wir nennen denjenigen Wert der Zufallsgröße x den *wahrscheinlichsten Wert*, zu dem die größte Wahrscheinlichkeit gehört, dem also der höchste Punkt des Wahrscheinlichkeitsdiagramms entspricht. Dieser Wert braucht nicht eindeutig zu sein.

In unserem ersten Beispiel gibt es keinen wahrscheinlichsten Wert der Zufallsvariablen $x =$ Augenzahl; alle sechs Werte sind hier gleichwahrscheinlich. Im zweiten Beispiel besitzt die Zufallsgröße $x =$ Anzahl der Fehler den wahrscheinlichsten Wert $x = 0$.

Wir betrachten wieder eine allgemeine diskrete Zufallsgröße x mit der Verteilungstabelle (2.1). Wir nennen die Funktion

$$\boxed{F(\xi) = P(x < \xi)} \tag{2.2}$$

die *Verteilungsfunktion der diskreten Zufallsgröße* x; $F(\xi)$ gibt die Wahrscheinlichkeit dafür an, daß die Zufallsvariable x einen kleineren Wert als ξ annimmt. Es gilt nun der

Satz 2.1: Die Verteilungsfunktion (2.2) der diskreten Zufallsgröße x mit der Verteilungstabelle (2.1) hat die Gestalt

$$\boxed{F(\xi) = \begin{cases} 0, & \text{wenn } -\infty < \xi \leqq x_1; \\ \sum\limits_{i=1}^{k} p_i, & \text{wenn } x_k < \xi \leqq x_{k+1} \quad (k = 1, 2, \ldots, n-1); \\ 1, & \text{wenn } x_n < \xi < \infty \end{cases}} \tag{2.3}$$

Beweis: $F(\xi) = 0$ für $\xi \leq x_1$ ist klar, da x_1 der kleinste Wert der Zufallsgröße x ist und demzufolge „$x < x_1$" ein unmögliches Ereignis darstellt. Im zweiten Fall ist definitionsgemäß $F(\xi) = P(„x = x_1“, + \ldots +, „x = x_k“)$ oder nach dem Additionssatz (1.9) weiter $F(\xi) = p_1 + \ldots + p_k$. Schließlich ist noch $F(\xi) = 1$ für $\xi > x_n$, da x_n der größte zufällige Wert ist und damit das Ereignis „$x \leqq x_n$" sicher ist.

Das Bild der Verteilungsfunktion $F(\xi)$ heißt die *Verteilungskurve der diskreten Zufallsgröße* x. Im allgemeinen Fall (2.1) hat die Verteilungskurve folgende Gestalt

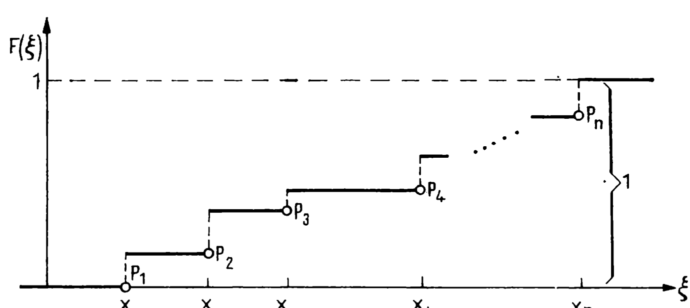

Abb. 3 Verteilungskurve einer diskreten Zufallsgröße

Die Verteilungskurve einer diskreten Zufallsgröße ist also eine nicht fallende Treppenfunktion, d. h. eine stückweise konstante Funktion, die an den Stellen x_i ($i = 1, \ldots, n$) endliche Sprünge macht und stets zwischen den Ordinatenwerten 0 und 1 bleibt. Die mit ◯ hervorgehobenen Punkte in der graphischen Darstellung der Verteilungsfunktion stellen die Werte der Verteilungsfunktion an den Stellen $x_1, \ldots, x_n$ dar. In unseren Beispielen erhalten wir die Verteilungskurven

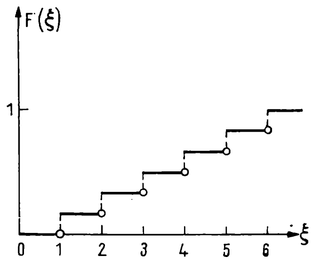

Abb. 4 Verteilungskurve zu Abb. 1

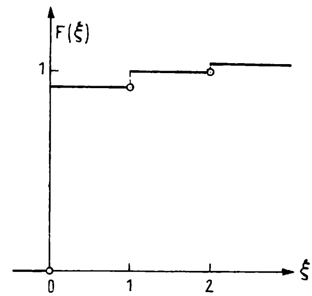

Abb. 5 Verteilungskurve zu Abb. 2

Anhand der ersten Darstellung können wir unmittelbar feststellen, mit welcher Wahrscheinlichkeit das Ereignis „$3 \leqq x < 5$" („mindestens 3, aber weniger als 5 Augen zu würfeln") eintritt. Wir haben lediglich die Differenz $F(5) - F(3) = \frac{5}{6} - \frac{3}{6} = \frac{1}{3}$ zu bilden. Das Ergebnis ist offenbar richtig, da $P(„3 \leqq x < 5") = P(„x = 3" + „x = 4") = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$ gilt.

Ganz allgemein besteht folgender Zusammenhang:

Satz 2.2: Die Wahrscheinlichkeit dafür, daß eine diskrete Zufallsvariable x mit der Verteilungsfunktion $F(\xi)$ einen Wert annimmt, der mindestens a und kleiner als b ($a < b$) ist, ergibt sich zu $F(b) - F(a)$, d. h.

$$\boxed{P(a \leqq x < b) = F(b) - F(a)}\,. \tag{2.4}$$

Beweis: Das uns interessierende Ereignis $E \equiv „a \leqq x < b"$ ist zunächst das Produkt der voneinander unabhängigen Ereignisse $A \equiv „a \leqq x"$ und $B \equiv „x < b"$, also $E = A \cdot B$. Nach der Formel (1.19) erhalten wir demzufolge

$$P(E) = P(AB) = P(A) + P(B) - P(A + b) \tag{*}$$

Wegen (2.2) ist nun $P(B) = F(b)$ und $P(A) = 1 - P(\bar{A}) = 1 - F(a)$. Weiter ist das Ereignis $A + B$ sicher, da jeder mögliche Wert der Zufallsgröße x nicht kleiner als a oder kleiner als b ist, mithin folgt $P(A + B) = 1$. Insgesamt ergibt sich dann aus (*)

$$P(E) = 1 - F(a) + F(b) - 1 = F(b) - F(a)\,.$$

Damit ist der Satz bewiesen.

Auf dem bewiesenen Satz beruht die Bedeutung der Verteilungsfunktion $F(\xi)$. Ist $F(\xi)$ bekannt, dann lassen sich die Wahrscheinlichkeiten der Art $P(a \leqq x < b)$ berechnen. Ist nun die Funktion $F(\xi)$ tabelliert, dann ergibt sich die gesuchte Wahrscheinlichkeit durch die einfache Differenzbildung $F(b) - F(a)$. Liegt dagegen $F(\xi)$ in Form der Verteilungskurve vor, dann kann die gesuchte Wahrscheinlichkeit unmittelbar abgesteckt werden.

2.2.2. *Der Erwartungswert einer diskreten Zufallsgröße*

Wir betrachten wieder eine allgemeine Zufallsgröße x mit der Verteilungstafel (2.1). Wir denken uns N Versuche ausgeführt, bei denen die Variable x genau N_i mal den Wert x_i $(i = 1, 2, \ldots, n)$ annimmt; offenbar ist $N_1 + N_2 + \ldots + N_n = N$. Für große Werte von N stellen die Quotienten $\frac{N_i}{N}$ $(i = 1, 2, \ldots, n)$ gemäß der statistischen Definition (1.3) die Wahrscheinlichkeiten p_i $(i = 1, 2, \ldots, n)$ dar. Bilden wir nun die Summe aller N beobachteten Werte, dann ergibt sich wegen $N_i = N_{p_i}$ $(i = 1, 2, \ldots, n)$ der Ausdruck

$$\sum_{i=1}^{n} N_i x_i = N \sum_{i=1}^{n} x_i p_i .$$

Als *Erwartungswert* oder *Mittelwert* $E(x)$ der diskreten Zufallsgröße x mit der Verteilungstabelle (2.1) führen wir die Größe[1]

$$\boxed{E(x) = \sum_{i=1}^{n} x_i p_i} \tag{2.5}$$

ein, die dem durchschnittlichen Wert der Ergebnisse in einer großen Versuchsserie entspricht.

Der Erwartungswert der Augenzahl x bei einem Wurf mit einem normierten Würfel hat gemäß (2.5) den Wert

$$E(x) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3{,}5 .$$

Bei einer großen Anzahl von Würfen ergeben sich demzufolge pro Wurf durchschnittlich 3,5 Augen.

Wir leiten nun eine Reihe von Eigenschaften her, die eine diskrete Zufallsgröße mit der Verteilungstabelle (2.1) besitzt. Als Vorbereitung dazu benötigen wir den

Satz 2.3: Besitzt die diskrete Zufallsgröße x die Verteilungstabelle (2.1) und ist $f(x)$ eine eindeutige Funktion in x, dann besteht für $z = f(x)$ die Verteilungstabelle

z	z_1	$z_2 \ldots z_n$
p	p_1	$p_2 \cdots p_n$

mit $z_i = f(x_i)$ $(i = 1, 2, \ldots, n)$.

[1] Ist $n = \infty$, dann muß die rechte Seite in (2.5) konvergent sein.

Beweis: Sind $x_1, \ldots, x_n$ die möglichen Werte der Variablen x, so besitzt die Zufallsgröße z die möglichen Werte $z_1, z_2, \ldots, z_n$ mit $z_i = f(x_i)$ ($i = 1, 2, \ldots, n$). Die Ereignisse „$x = x_i$" und „$z = z_i$" sind offenbar gleichwahrscheinlich; damit ist der Satz bewiesen.

Satz 2.4: Der Erwartungswert einer diskreten Zufallsgröße hat die Eigenschaft

$$\boxed{E(\alpha x + \beta) = \alpha E(x) + \beta}\,, \tag{2.6}$$

wobei α und β beliebige Konstante sind.

Beweis: Wir setzen $z = f(x) = \alpha x + \beta$ und erhalten aus Satz 2.3 und Definition (2.5)

$$E(z) = \sum_{i=1}^{n} z_i p_i = \sum_{i=1}^{n} \alpha x_i p_i + \sum_{i=1}^{n} \beta p_i = \alpha \sum_{i=1}^{n} x_i p_i + \beta \sum_{i=1}^{n} p_i$$

$$= \alpha E(x) + \beta\,.$$

Dabei ist die Vollständigkeitsrelation $p_1 + p_2 + \ldots + p_n = 1$ benutzt worden.

Satz 2.5: Der Erwartungswert des Quadrates einer diskreten Zufallsgröße ist größer als das Quadrat des Erwartungswertes dieser Zufallsgröße

$$\boxed{E(x^2) > E^2(x)}\,. \tag{2.7}$$

Beweis: Nach Satz 2.3 und Gleichung (2.5) gilt zunächst

$$E(x) = \sum_{i=1}^{n} x_i p_i, \quad E(x^2) = \sum_{i=1}^{n} x_i^2 p_i.$$

Wir bilden

$$2 E^2(x) = 2 E(x) \cdot E(x) = 2 E(x) \cdot \sum_{i=1}^{n} x_i p_i = \sum_{i=1}^{n} 2 E(x) x_i p_i$$

und weiter

$$E^2(x) = E^2(x) \cdot \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} E^2(x) p_i$$

und erhalten damit

$$E(x^2) - E^2(x) = E(x^2) - 2 E^2(x) + E^2(x) = \sum_{i=1}^{n} (x_i - E(x))^2 p_i; \tag{2.8}$$

die letzte Summe ist aber positiv, da sie aus positiven Summanden besteht. Damit folgt die Behauptung

$$E(x^2) - E^2(x) > 0\,.$$

Satz 2.6: Der Erwartungswert für die Summe zweier diskreter Zufallsgrößen ist gleich der Summe der Erwartungswerte der einzelnen Zufallsgrößen, es gilt also die Beziehung

$$\boxed{E(x + y) = E(x) + E(y)}\,. \tag{2.9}$$

Beweis: Neben der Zufallsgröße x mit der Tabelle (2.1) betrachten wir noch eine diskrete Variable y mit der Verteilungstabelle

$$\begin{array}{c|cccc} y & y_1 & y_2 & \cdots & y_m \\ \hline q & q_1 & q_2 & \cdots & q_m \end{array}\,.$$

Neben den Ereignissen $X_i \equiv$ „$x = x_i$“ ($i = 1, 2, \ldots, n$) und $y_j \equiv$ „$y = y_j$“ ($j = 1, 2, \ldots, m$) interessieren uns die zusammengesetzten Ereignisse $S_{ij} \equiv X_i + Y_j$ ($i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$) mit den Wahrscheinlichkeiten $P(S_{ij}) = s_{ij}$ ($i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$). Die zuletzt erwähnten Wahrscheinlichkeiten erfüllen wegen

$$\left.\begin{array}{l} X_i\,Y_1 + \ldots + X_i\,Y_m = X_i \;(i = 1, 2, \ldots, n) \\ X_1\,Y_j + \ldots + X_n\,Y_j = Y_j \;(j = 1, 2, \ldots, m) \end{array}\right\} \tag{2.10}$$

die Relationen

$$\sum_{i=1}^{n} s_{ij} = q_j \;(j = 1, 2, \ldots, m), \quad \sum_{j=1}^{m} s_{ij} = p_i \;(i = 1, 2, \ldots, n)\,. \tag{2.11}$$

Bilden wir nun

$$E(x + y) = \sum_{i=1}^{n} \sum_{j=1}^{m} (x_i + y_j) s_{ij},$$

so ergibt sich unter Beachtung von (2.11) gerade die Behauptung (2.9). Der Satz 2.6 läßt sich verallgemeinern zum

Satz 2.7: Der Erwartungswert einer Summe von beliebig vielen diskreten Zufallsvariablen ist gleich der Summe der Erwartungswerte der einzelnen Zufallsgrößen:

$$\boxed{E(x + y + z + \ldots) = E(x) + E(y) + E(z) + \ldots}\,. \tag{2.12}$$

Wir leiten nun noch eine Eigenschaft des Erwartungswertes diskreter Zufallsgrößen ab.

Satz 2.8: Der Erwartungswert des Produktes zweier voneinander unabhängiger diskreter Zufallsvariablen ist gleich dem Produkt der Erwartungswerte der einzelnen Zufallsgrößen:

$$\boxed{E(x \cdot y) = E(x) \cdot E(y)}\,. \tag{2.13}$$

Beweis: Wir knüpfen an die Bezeichnungen beim Beweis des Satzes 2.6 an und betrachten die Ereignisse $T_{ij} = X_i\, Y_j$ $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$ mit den Wahrscheinlichkeiten $t_{ij} = p_i q_j$ $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$ (wegen (1.16)!). Gemäß (2.5) erhalten wir nun das Ergebnis

$$E\,(xy) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i\, y_j\, t_{ij} = \left(\sum_{i=1}^{n} x_j\, p_i\right) \left(\sum_{j=1}^{m} y_i\, q_j\right) = E\,(x) \cdot E\,(y)\,.$$

Auch diese Aussage läßt sich erweitern zum

Satz 2.9: Der Erwartungswert eines Produktes beliebig vieler paarweise voneinander unabhängiger diskreter Zufallsvariablen ist gleich dem Produkt der Erwartungswerte der einzelnen Zufallsgrößen:

$$E\,(x \cdot y \cdot z \cdot \ldots) = E\,(x) \cdot E\,(y) \cdot E\,(z) \ldots \tag{2.14}$$

Nun betrachten wir einige Beispiele. Zunächst denken wir uns einen ‚falschen' Würfel mit der Verteilungstabelle

x	1	2	3	4	5	6
p	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{12}$

für die Augenzahl x bei einem Wurf. Für den diesbezüglichen Mittelwert erhalten wir $E\,(x) = \frac{1}{12}\,(1 + 6 + 3 + 16 + 10 + 6) = \frac{42}{12} = 3{,}5$. Wir haben hier also denselben Erwartungswert wie im Normalfall. Dennoch unterscheiden sich beide Verteilungen offensichtlich; die Verteilung im Normalfall hat keinen wahrscheinlichsten Wert, die hier behandelte Verteilung besitzt als einzigen wahrscheinlichsten Wert $x = 4$. Schon diese Überlegungen demonstrieren, daß der Erwartungswert einer diskreten Zufallsvariablen die zugehörige Verteilung nicht ausreichend beschreibt. Gegeben seien nun zwei normierte Würfel, und wir betrachten als diskrete Zufallsvariable die Summe bzw. das Produkt der Augenzahlen der einzelnen Würfel nach einem Wurf. Bezeichnen wir mit x bzw. y die Augenzahl des einen bzw. des anderen Würfels nach einem Wurf, dann interessiert uns die Zufallsgröße $x + y$ bzw. $x \cdot y$. Für die erste Zufallsvariable ergibt sich nach (2.9) der Erwartungswert

$$E\,(x + y) = \frac{21}{6} + \frac{21}{6} = 7\,.$$

Im Mittel werden also bei jedem Wurf mit zwei Würfen 7 Augen erzielt. Beachten wir schließlich noch die Tatsache, daß x und y unabhängig voneinander sind, dann folgt aus (2.13) für den Erwartungswert von $x \cdot y$ der Wert

$$E\,(x \cdot y) = \frac{21}{6} \cdot \frac{21}{6} = 12{,}25\,.$$

Im Mittel beläuft sich also das Produkt der Augenzahlen der beiden Würfel nach einem Wurf auf 12,25.

Schließlich betrachten wir noch zwei voneinander unabhängige diskrete Zufallsgrößen x bzw. y mit den Verteilungstabellen

x	5	2	4
p	0,6	0,1	0,3

y	7	9
q	0,8	0,2

Wie groß sind die Erwartungswerte $E(x)$, $E(y)$, $E(x+y)$ und $E(xy)$? Wir erhalten zunächst nach (2.5)

$$E(x) = 5 \cdot 0{,}6 + 2 \cdot 0{,}1 + 4 \cdot 0{,}3 = 4{,}4,$$
$$E(y) = 7 \cdot 0{,}8 + 9 \cdot 0{,}2 = 7{,}4.$$

Für die Zufallsgröße $x + y$ erhalten wir gemäß (2.9)

$$E(x+y) = E(x) + E(y) = 4{,}4 + 7{,}4 = 11{,}8.$$

Beachten wir noch die Unabhängigkeit der Variablen x und y voneinander, so ergibt sich aus (2.13) ohne viel Rechnung

$$E(xy) = E(x) \cdot E(y) = 4{,}4 \cdot 7{,}4 = 32{,}56.$$

2.2.3. *Die Streuung einer diskreten Zufallsgröße*

Wir haben bereits an einem Beispiel gezeigt, daß der Erwartungswert einer diskreten Zufallsvariablen nicht ausreicht, die betreffende Verteilung vollständig zu charakterisieren. Wir wollen diesen Sachverhalt nun verdeutlichen. Zu diesem Zweck betrachten wir zwei diskrete Zufallsgrößen x und y, denen die Verteilungstabellen

x	8	10	12
p	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

y	2	10	15
q	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

entsprechen. Für die zugehörigen Erwartungswerte gilt

$$E(x) = E(y) = \frac{31}{3}\,.$$

Beide diskrete Verteilungen haben also denselben Erwartungswert. Wir erkennen aber unmittelbar, daß die erste Verteilung „besser" ist als die zweite. Die Verteilung der Variablen x „streut" nämlich nicht so sehr wie die Verteilung der Zufallsgröße y; die x-Werte liegen nur in dem Intervall $[8, 12]$, die Werte der Variablen y dagegen in dem viel größeren Bereich $[2, 15]$.

Wir suchen nun ein Maß für die eben erwähnte „Streuung" einer diskreten Zufallsvariablen. Zu diesem Zweck sei eine diskrete Zufallsgröße x allgemein gegeben. Wir könnten versuchen, als Maß hierfür den Erwartungswert der Zufallsgröße $z = x - E(x)$, also der Abweichung der Variablen x von deren Erwartungswert zu nehmen. Das ist aber nicht möglich, da dieser Wert nach Satz 2.4 stets verschwindet, d. h. für alle diskreten Verteilungen gleich ist.

Dagegen ist es prinzipiell möglich, als Maß für die „Streuung" einer Zufallsgröße x den Mittelwert der Zufallsvariablen $z = |x - E(x)|$ anzusehen. Wir verzichten jedoch darauf, da das Operieren mit absoluten Beträgen umständlich ist.

Wir betrachten eine diskrete Zufallsgröße x mit der Verteilungstabelle (2.1). Wir fassen nun die Zufallsvariable $z = [x - E(x)]^2$ ins Auge, die die Quadrate der möglichen Abweichungen $x_i - E(x)$ $(i = 1, 2, \ldots, n)$ der Zufallsgröße x von deren Erwartungswert $E(x)$ mit den Wahrscheinlichkeiten p_i $(i = 1, 2, \ldots, n)$ annimmt. Gemäß (2.5) ist dann

$$\boxed{\sigma^2(x) = E([x - E(x)]^2) = \sum_{i=1}^{n} [x_i - E(x)]^2 p_i} \tag{2.15}$$

der Erwartungswert für die diskrete Zufallsgröße $z = [x - E(x)]^2$. Die Größe σ^2 heißt die *Streuung der diskreten Zufallsvariablen* x. Auf Grund der Definition (2.15) ist σ^2 in der Tat ein Maß für die oben charakterisierte Streuung einer diskreten Verteilung. Berechnen wir die Streuungen der beiden letzten konkreten Verteilungen, so erhalten wir im ersten Fall

$$\sigma^2(x) = 1{,}89$$

und im zweiten

$$\sigma^2(y) = 18{,}89.$$

Wir sehen also, daß die erste Verteilung eine kleinere Streuung als die zweite besitzt.

Aus (2.15) geht hervor, daß $\sigma^2(x)$ eine nichtnegative Zahl ist. Anstelle σ^2 wird in der Praxis häufig die positive Wurzel aus der Streuung, d. h. die Größe

$$\boxed{\sigma(x) = \sqrt{\sum_{i=1}^{n} [x_i - E(x)]^2 p_i}} \tag{2.16}$$

benutzt, die die Namen *mittlere quadratische Abweichung*, *Standardabweichung* und *Streuungsmaß* der betreffenden Verteilung trägt. Die Größe $\sigma(x)$ hat gegenüber $\sigma^2(x)$ den Vorteil, daß $\sigma(x)$ dieselbe Dimension wie die Zufallsgröße x besitzt.

In unseren beiden Beispielen ergeben sich gemäß (2.16) die Streuungsmaße $\sigma(x) = 1{,}37$ und $\sigma(y) = 4{,}34$.

Wir leiten nunmehr einige Eigenschaften der Streuung einer diskreten Zufallsvariablen her. Zunächst beweisen wir den

Satz 2.10: Für die Streuung einer diskreten Zufallsgröße x gilt die Formel

$$\boxed{\sigma^2(x) = E(x^2) - E^2(x)}\,. \tag{2.17}$$

Beweis: Die Formel (2.17) ergibt sich unmittelbar aus (2.16) in Verbindung mit der Relation (2.8). Sie wird häufig in der Praxis benutzt.

Wir berechnen nun nach der Regel (2.17) die Streuung $\sigma^2(x)$ der diskreten Zufallsgröße x = Augenzahl bei einem Wurf mit einem normierten Würfel. Wir haben schon im Abschnitt 2.22 die Größe $E(x) = \frac{21}{6}$ ermittelt. Wir benötigen zur Anwendung der Formel (2.17) noch $E(x^2)$. Aus der Verteilungstabelle

x^2	1	4	9	16	25	36
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

für die Größe x^2 ergibt sich

$$E(x^2) = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{6}.$$

Nunmehr liefert die Regel (2.17)

$$\sigma^2(x) = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{35}{12}.$$

Wir geben nun zwei weitere Sätze bezüglich der Streuung.

Satz 2.11: Die Streuung einer diskreten Zufallsgröße x hat die Eigenschaft

$$\boxed{\sigma^2(\alpha x) = \alpha^2\sigma^2(x)} \tag{2.18}$$

mit α als beliebiger Konstante.

Beweis: Wir setzen $z = \alpha x = f(x)$. Dann folgt aus den Sätzen 2.3 und 2.6 sowie aus (2.15)

$$\sigma^2(z) = \sum_{i=1}^{n} [z_i - E(z)]^2 p_i = \alpha^2 \sum_{i=1}^{n} [x_i - E(x)]^2 p_i = \alpha^2\sigma^2(x).$$

Satz 2.12: Die Streuung für die Summe zweier voneinander unabhängiger diskreter Zufallsvariablen ist gleich der Summe der Streuung der einzelnen Zufallsgrößen:

$$\boxed{\sigma^2(x + y) = \sigma^2(x) + \sigma^2(y)}. \tag{2.19}$$

Beweis: Wir betrachten die Zufallsgröße $z = x + y$. Wenden wir Satz 2.10 auf z an, so erhalten wir

$$\sigma^2(z) = E(z^2) - E^2(z) = E(x^2 + 2xy + y^2) - [E(x + y)]^2.$$

Beachten wir die Sätze 2.7 und 2.8 sowie nochmals Satz 2.10, so wird

$$\sigma^2(z) = E(x^2) - E^2(x) + E(y^2) - E^2(y) = \sigma^2(x) + \sigma^2(y)$$

Den vorstehenden Satz können wir noch allgemeiner aussprechen:

Satz 2.13: Die Streuung für die Summe beliebig vieler paarweise voneinander unabhängiger diskreter Zufallsgrößen ist gleich der Summe der Streuungen der einzelnen Zufallsvariablen:

$$\boxed{\sigma^2(x + y + z + \ldots) = \sigma^2(x) + \sigma^2(y) + \sigma^2(z) + \ldots}. \tag{2.20}$$

Bevor wir einen weiteren Satz herleiten, betrachten wir ein Beispiel[1] [3]. Ein gewisser Mechanismus soll aus 9 Einzelteilen montiert werden, die alle auf einer Achse fest aneinander angebracht wurden. Die Länge $x^{(i)}$ ($i = 1, 2, \ldots, 9$) jedes Einzelteils ist eine Zufallsvariable. Diese Zufallsgrößen sind offenbar unabhängig voneinander. Die entsprechenden Erwartungswerte bzw. Streuungsmaße seien $E\,(x^{(i)}) = 10$ cm bzw. $\sigma\,(x^{(i)}) = 0{,}2$ cm ($i = 1, 2, \ldots, 9$). Wie groß sind die Parameter $E\,(x)$ und $\sigma\,(x)$ der Länge x der Kette dieser 9 aneinander montierten Einzelteile? Zunächst wird nach Satz 2.7

$$E\,(x) = \sum_{i=1}^{9} E\,(x^{(i)}) = 90 \text{ cm}\,.$$

Aus Satz 2.13 folgt weiter

$$\sigma\,(x) = \sqrt{\sum_{i=1}^{9} \sigma^2\,(x^{(i)})} = \sqrt{9 \cdot 0{,}04} = 0{,}6\,.$$

Ist also die Streuung der Einzelteillänge 2 Prozent des zugehörigen Erwartungswertes, dann beträgt die Streuung der Kettenlänge nur noch $^2/_3$ Prozent des zugehörigen Erwartungswertes. Die an diesem Beispiel beobachtete Verringerung der zum Erwartungswert relativen Streuung bei der Addition von zufälligen Größen spielt bei der Montage von Präzisionsinstrumenten eine wichtige Rolle.

Wir beweisen nun den

Satz 2.14: Für jede diskrete Zufallsvariable x gilt die *TSCHEBYSCHEFFsche Ungleichung* (Streuungsungleichung)

$$\boxed{P\,(\,|\,x - E\,(x)\,|\, < \gamma) \geqq 1 - \frac{\sigma^2\,(x)}{\gamma^2}}\,; \tag{2.21}$$

dabei erfüllt die Konstante γ die Bedingung $\gamma \geqq \sigma\,(x)$.

Beweis: Wir betrachten die Zufallsgröße x mit der Verteilungstabelle (2.1) und schätzen die Wahrscheinlichkeit für das Ereignis $A \equiv$ „$|x - E\,(x)| \geqq \gamma$" nach oben ab.

Zu diesem Zweck lassen wir in dem Ausdruck (2.15) diejenigen Glieder fort, für die $[x_i - E\,(x)]^2 < \gamma^2$ oder $|x_i - E\,(x)| < \gamma$ besteht. Bezeichnen wir mit $\sum^*$ die Summe über die verbleibenden Glieder in (2.15), die also dem Betrag nach $\geqq \gamma^2$ sind, so folgt

$$\sigma^2\,(x) \geqq \textstyle\sum^* [x_i - E\,(x)]^2 p_i \geqq \gamma^2 \sum^* p_i\,.$$

Wegen $\sum^* p_i = P\,(A)$ haben wir damit die Abschätzung

$$P\,(A) \leqq \frac{\sigma^2\,(x)}{\gamma^2}\,.$$

[1] Die in diesem Beispiel vorkommenden Zufallsgrößen sind zwar nicht diskret, sondern kontinuierlich. Das tut aber nichts zur Sache, da die Streuung einer kontinuierlichen Zufallsvariable dieselben Eigenschaften aufweist, die wir bisher bei einer diskreten Zufallsveränderlichen kennengelernt haben (Vgl. hierzu Satz 2.24).

Gehen wir nun zum Ereignis $\bar{A} \equiv$ „$|x - E(x)| < \gamma$" über, dann ergibt sich die behauptete Ungleichung (2.21). Die Einschränkung $\gamma \geqq \sigma(x)$ ist erforderlich, damit die rechte Seite der Ungleichung (2.21) nicht negativ wird.

Wir behandeln nun zwei Beispiele: In einer Zündholzfabrik werden die Zündholzschachteln automatisch abgepackt. Die Anzahl der in einer Schachtel enthaltenen Hölzer ist offenbar eine diskrete Zufallsvariable, die die Verteilungstabelle

x	57	58	59	60	61	62	63
p	0,05	0,10	0,15	0,40	0,15	0,10	0,05

besitzt.

Der Erwartungswert oder Mittelwert der Anzahl der Hölzer pro Schachtel beträgt gemäß der Formel (2.5) $E(x) = 60$ (das ist auch aus Symmetriegründen klar!). Für die Streuung folgt nach (2.15) $\sigma^2(x) = 2$. Wie groß ist nun die Wahrscheinlichkeit p_0 dafür, daß die tatsächliche Anzahl x der Hölzer in einer Schachtel um weniger als 3 vom Mittelwert $E(x) = 60$ abweicht? Wenden wir die TSCHEBYSCHEFFsche Ungleichung (2.21) an, so erhalten wir (mit $\gamma = 3$)

$$p_0 = P(|x - E(x)| < 3) \geqq 1 - \frac{2}{9} = \frac{7}{9}.$$

Die gesuchte Wahrscheinlichkeit beträgt also mindestens $\frac{7}{9} \approx 0{,}78$. Der obigen Verteilungstabelle entnehmen wir den genauen Wert

$$p_0 = 0{,}10 + 0{,}15 + 0{,}40 + 0{,}15 + 0{,}10 = 0{,}90;$$

denn es gilt

$$P(|x - E(x)| < 3) = P(x = 58) + P(x = 59) + P(x = 60) + \\ + P(x = 61) + P(x = 62).$$

Die Streuungsungleichung liefert also eine grobe Abschätzung für die Größe p_0.

In einem Betrieb, der Kugeln (z. B. für Kugellager) herstellt, werden die gefertigten Kugeln maschinell abgepackt. Die Anzahl x der Kugeln pro Packung ist eine diskrete Zufallsgröße. Statistische Untersuchungen ergaben die Parameter $E(x) = 10000$ und $\sigma^2(x) = 100$; die Verteilungstabelle der Zufallsvariablen ist nicht bekannt. Gesucht ist nun die Wahrscheinlichkeit dafür, daß die wahre Kugelanzahl in einer Packung um weniger als 50 vom Sollwert = Mittelwert abweicht? Auf Grund der Beziehung (2.21) erhalten wir sogleich die Antwort auf diese Frage zu

$$P[|x - E(x)| < 50] \geqq 1 - \frac{100}{2500} = 0{,}96.$$

Wir können demzufolge damit rechnen, daß nur in 4 Prozent aller Packungen die Anzahl der Kugeln um mehr als 50 vom Sollwert abweicht.

Die im Satz 2.14 hergeleitete Ungleichung (2.21) wird in der Praxis häufig benutzt. Sie ermöglicht die näherungsweise Berechnung von Wahrscheinlichkeiten dafür, daß die diskrete Zufallsgröße einem um ihren Mittelwert symmetrisch liegenden Intervall angehört oder nicht. Die TSCHEBYSCHEFFsche Ungleichung wird insbesondere dann angewendet, wenn die Verteilungstabelle der Zufallsgröße nicht bekannt ist, aber die Parameter E und σ vorliegen (vgl. letztes Beispiel!).

Wir schließen nun die allgemeinen Bemerkungen über diskrete Zufallsgrößen ab. Prinzipiell können wir uns diskrete Verteilungen der verschiedensten Formen vorstellen.

Im nun folgenden Paragraphen lernen wir die binomische Verteilung kennen, die für praktische Belange von Bedeutung ist und als Prototyp einer diskreten Verteilung angesehen werden kann. Im Paragraphen 2.4 verallgemeinern wir diese Verteilung zu der Poissonschen Verteilung, die in der Theorie der Massenbedingung (vgl. Kapitel 4) eine fundamentale Rolle spielt.

2.3. Die binomische Verteilung

Wir beginnen unsere Betrachtungen mit zwei praktischen Fragen. In einem Betrieb sind 96 Prozent der hergestellten Erzeugnisse normgerecht. Wie groß ist die Wahrscheinlichkeit p_1 dafür, daß in einer Packung von 10 Erzeugnissen alle normgerecht sind? — In einer Abteilung eines Industriebetriebes ist der Wasserverbrauch an einem Tage zu 75 Prozent normal. Wie groß ist die Wahrscheinlichkeit p_2 dafür, daß innerhalb einer Woche (6 Tage) der Wasserverbrauch nur an drei Tagen normal ist?

Beide Fragen berühren einen Sachverhalt, den wir nun ganz allgemein an dem *Urnenschema von BERNOULLI* charakterisieren: Eine Urne mit schwarzen und weißen Kugeln ist gegeben. Mit p bzw. $q = 1 - p$ bezeichnen wir die Wahrscheinlichkeit dafür, daß eine aufs Geratewohl gezogene Kugel weiß bzw. schwarz ist. Es werden n zufällige *Ziehungen* vorgenommen, d. h., es werden der Reihe nach aufs Geratewohl n Kugeln ausgewählt. Die Anzahl x der dabei gezogenen weißen Kugeln ist offenbar eine diskrete Zufallsvariable mit den möglichen Werten $0, 1, 2, \ldots, n - 1, n$. Wie groß ist nun die Wahrscheinlichkeit $P_n(x)$ dafür, daß wir bei n zufälligen Ziehungen genau x weiße (und damit $n - x$ schwarze) Kugeln erhalten?

Um die aufgeworfene Frage zu beantworten, setzten wir die Unabhängigkeit der einzelnen Ziehungen voraus. Diese Voraussetzung ist offenbar erfüllt, wenn die Anzahl der schwarzen und weißen Kugeln in der Urne sehr groß ist. Aber auch bei geringer Kugelanzahl kann diese Annahme eingehalten werden, wenn Ziehungen mit Zurücklegung erfolgen, d. h., wenn nach jeder Ziehung die ausgewählte Kugel in die Urne zurückgelegt wird. Sprechen wir im folgenden von Ziehungen, so sind damit stets unabhängige zufällige Ziehungen gemeint. Wir berechnen nun die Wahrscheinlichkeit $P_n(x)$ für das Ereignis $A \equiv$ „x weiße und $n - x$ schwarze Kugeln bei n Ziehungen". Aus der Kombinatorik ist bekannt, daß x weiße und $n - x$ schwarze Kugeln genau $m = \binom{n}{x}$ Permutationen $A_1, A_2, \ldots, A_m$ zulassen, so daß

$$A = A_1 + A_2 + \ldots + A_m; \; m = \binom{n}{x} \tag{2.22}$$

gilt. Die Ereignisse $A_1, \ldots, A_m$ sind unverträglich und gleichwahrscheinlich; nach Satz 1.9 (unabhängige Ziehungen) erhalten wir für alle A_i ($i = 1, 2, \ldots, m$) die gleiche Wahrscheinlichkeit zu

$$P(A_1) = P(A_2) = \ldots = P(A_m) = p^x q^{n-x}. \tag{2.23}$$

Aus (2.22) und (2.23) folgt gemäß Satz 1.1

$$P_n(x) = P(A) = \sum_{i=1}^{m} p^x q^{n-x} = \binom{n}{x} p^x q^{n-x}$$

oder der

Satz 2.15: Die Wahrscheinlichkeit $P_n(x)$ dafür, daß sich bei n Ziehungen x weiße Kugeln ergeben, berechnet sich nach der *BERNOULLIschen Formel*

$$\boxed{P_n(x) = \binom{n}{x} p^x q^{n-x} \quad (x = 0, 1, \ldots, n)}\,. \tag{2.24}$$

Wir sind nun in der Lage, die eingangs gestellten Fragen zu beantworten. Im ersten Fall ist $n = 10$, $p = 0{,}96$ ($q = 0{,}04$) und $x = 10$; daher folgt aus (2.24)

$$p_1 = P_{10}(10) = \binom{10}{10} 0{,}96^{11} \cdot 0{,}04^0 = 0{,}665\,.$$

Im zweiten Beispiel ergibt sich wegen $n = 6$, $p = 0{,}75$ ($q = 0{,}25$) und $x = 3$ ganz analog gemäß (2.24)

$$p_2 = P_6(3) = \binom{6}{3} \cdot \left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right)^3 = 0{,}132\,.$$

In Verbindung mit der BERNOULLIschen Formel (2.24) definieren wir nun als *binomische* oder *BERNOULLIsche Verteilung* diejenige Verteilung, die zu der diskreten Zufallsgröße x mit der Verteilungstabelle

Wert	0	1	2	n
Wahrscheinlichkeit	$P_n(0)$	$P_n(1)$	$P_n(2)$	$P_n(n)$

(2.25)

gehört. Die binomische Verteilung hängt also von den Parametern n und p mit $n > 0$ ganz und $0 \leqq p \leqq 1$ ab.

Wir zeigen, daß die Tabelle (2.25) die Vollständigkeitsrelation

$$\sum_{m=0}^{n} P_n(x) = 1 \tag{2.26}$$

erfüllt. Beachten wir den binomischen Satz

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

so erhalten wir die Behauptung (2.26)

$$\sum_{x=0}^{n} P_n(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (p + q)^n = 1.$$

Als *binomische* oder *BERNOULLIsche Verteilungsfunktion* führen wir die Funktion

$$\boxed{B_n(\xi) = P(x < \xi) = \sum_{x<\xi} P_n(x)} \tag{2.27}$$

ein, die — wie im allgemeinen Fall einer diskreten Verteilung (vgl. Abschnitt 3.21) — die Wahrscheinlichkeit dafür angibt, daß die Zufallsgröße x einen kleineren Wert als ξ annimmt. $B_n(\xi)$ bedeutet demzufolge die Wahrscheinlichkeit dafür, daß bei n Ziehungen weniger als ξ Kugeln weiß sind. Wir betrachten nun als Beispiel die BERNOULLIsche Verteilung mit den Parametern $n = 6$ und $p = 0{,}75$. Die Verteilungstabelle (2.25) lautet dann

Wert	0	1	2	3	4	5	6
Wahrscheinlichkeit	$\frac{1}{4096}$	$\frac{18}{4096}$	$\frac{135}{4096}$	$\frac{540}{4096}$	$\frac{1215}{4096}$	$\frac{1458}{4096}$	$\frac{729}{4096}$

Das zugehörige Wahrscheinlichkeitsdiagramm hat die Form

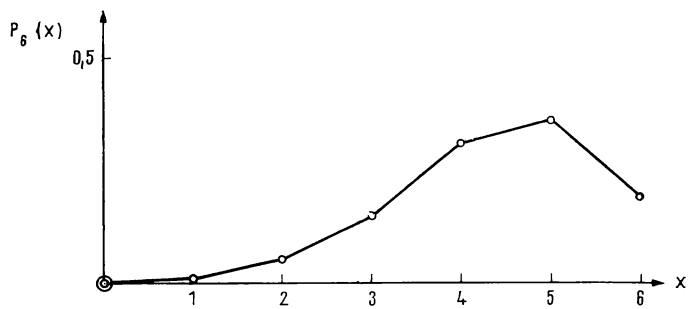

Abb. 6 Wahrscheinlichkeitsdiagramm einer BERNOULLIschen Verteilung

Wir erkennen, daß die vorliegende Verteilung den einzigen wahrscheinlichsten Wert $x_0 = 5$ besitzt. Die zugehörige Verteilungsfunktion $B_6(\xi)$ weist die graphische Darstellung (Verteilungskurve) auf:

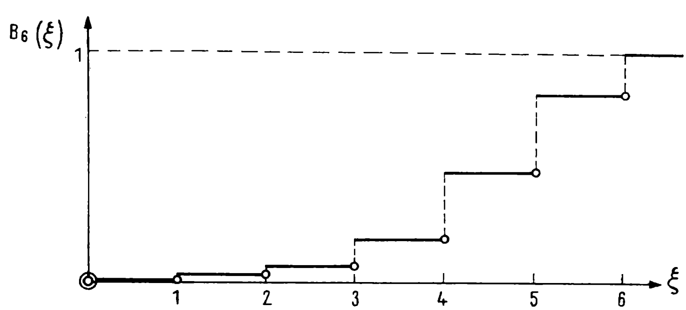

Abb. 7 Verteilungskurve einer BERNOULLIschen Verteilung

Wir geben nun wieder ein Zahlenbeispiel: In einem Betrieb gehören 60 Prozent der normgerechten Erzeugnisse zur Sorte 1. Es werden nun Packungen von je 100 normgerechten Erzeugnissen (durch zufällige Zusammenstellung) hergestellt. Wie groß ist die Wahrscheinlichkeit p_1 dafür, daß eine Packung keine normgerechten Erzeugnisse enthält, die zur Sorte 1 gehören? Wie groß ist die Wahrscheinlichkeit p_2 dafür, daß in einer Packung weniger als die Hälfte zur Sorte 1 gehört? Die erste Größe ermitteln wir gemäß Formel (2.24) zu

$$p_1 = P_{100}(0) = \binom{100}{0} 0{,}6^{\circ} \cdot 0{,}4^{100} \approx 0 \,.$$

Es ist praktisch ausgeschlossen, daß unter 100 Erzeugnissen kein Erzeugnis der Sorte 1 ist. Für die zweite Größe p_2 erhalten wir den Ausdruck

$$p_2 = B'_{100}(50) = \sum_{x=0}^{49} P_{100}(x) \approx 0{,}021 \,, \tag{2.28}$$

dessen Berechnung schon sehr mühevoll ist (immerhin sind 50 Summanden zu berechnen, die aus Binomialkoeffizienten und Potenzen bestehen). Im Zusammenhang mit der Normalverteilung (vgl. Paragraph 2.6) werden wir eine Näherungsformel kennenlernen, mit deren Hilfe Ausdrücke der Gestalt (2.27) äußerst einfach zu berechnen sind. Das Ergebnis in (2.28) bedeutet also, daß in etwa 1/50 aller Packungen weniger als die Hälfte der Erzeugnisse zur Sorte 1 gehören.

Wir betrachten nun wieder eine allgemeine BERNOULLIsche Verteilung und beweisen

Satz 2.16: Ist $np + p$ keine ganze Zahl, dann ist der einzige wahrscheinlichste Wert der zugehörigen binomischen Verteilung die ganze Zahl zwischen $np + p - 1$ und $np + p$. Ist $np + p$ eine ganze Zahl, dann besitzt die zugehörige binomische Verteilung die beiden gleichwahrscheinlichsten Werte $np + p - 1$ und $np + p$.

Beweis: Ist x_0 der wahrscheinlichste Wert der binomischen Verteilung, d. h., hat das zugehörige Wahrscheinlichkeitsdiagramm bei x_0 ein Maximum, dann muß offenbar

$$P_n(x_0) \geqq P_n(x_0 - 1), \; P_n(x_0) \geqq P_n(x_0 + 1)$$

oder wegen (2.24)

$$(n - x_0 + 1)\, p \geqq x_0\, q, \quad (x_0 + 1)\, q \geqq (n - x_0)\, p$$

gelten. Lösen wir die beiden letzten Ungleichungen nach x_0 auf und beachten wir $p + q = 1$, dann ergibt sich die Beziehung

$$np + p - 1 \leqq x_0 \leqq np + p \,.$$

Ist nun $np + p$ eine ganze Zahl, dann sind $np + p - 1$ und $np + p$ die beiden gleichwahrscheinlichsten Werte der Verteilung. Ist $np + p$ keine ganze Zahl, dann liegt aber in dem Intervall $np + p - 1 \leqq x_0 \leqq np + p$ der Länge 1 eine ganze Zahl, die den einzigen wahrscheinlichsten Wert ausmacht.

Wir bestimmen nun die wahrscheinlichsten Werte der uns schon bekannten binomischen Verteilungen. Bei der ersten Verteilung liegt wegen $n = 10$, $p = 0{,}96$ und $np + p = 9{,}6 + 0{,}96 = 10{,}56$ der einzige wahrscheinlichste Wert $x_0 = 10$ vor. Im zweiten Beispiel haben wir $n = 6$, $p = 0{,}75$ und $np + p = 4{,}5 + 0{,}75 = 5{,}25$; daher gibt es hier den einzigen wahrscheinlichsten Wert $x_0 = 5$. Auch im letzten Fall existiert nur ein wahrscheinlichster Wert, denn wegen $n = 100$, $p = 0{,}6$ ist $np + p = 60 + 0{,}6 = 60{,}6$ nicht ganz; es gilt also $x_0 = 60$. Die Verteilung mit den Parametern $n = 14$ und $p = \frac{1}{3}$ besitzt dagegen die beiden gleichwahrscheinlichsten Werte $x'_0 = np + p - 1 = \frac{14}{3} + \frac{1}{3} - 1 = 4$ und $x''_0 = np + p = 5$.

Wir beweisen nun den wichtigen

Satz 2.17: Die BERNOULLIsche Verteilung besitzt den Erwartungswert

$$\boxed{E(x) = np} \tag{2.29}$$

und die Streuung

$$\boxed{\sigma^2 = npq} \tag{2.30}$$

Beweis: Für den Erwartungswert der diskreten Zufallsvariable x mit der Verteilungstabelle (2.25) erhalten wir zunächst definitionsgemäß

$$E(x) = \sum_{x=0}^{n} x\, P_n(x)$$

oder

$$E(x) = \sum_{x=1}^{n} \frac{n!}{(x-1)!\,(n-x)!}\, p^x q^{n-x} =$$

$$np \cdot \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,[(n-1)-(x-1)]!}\, p^{x-1} \cdot q^{(n-1)-(x-1)}$$

Führen wir die Abkürzungen $n - 1 = m$ und $x - 1 = y$ ein, dann wird

$$E(x) = np \cdot \sum_{y=0}^{m} \binom{m}{y} p^y q^{m-y} = np\,,$$

da die vorstehende Summe wegen (2.26) den Wert 1 hat. Damit ist der erste Teil der Behauptung bewiesen. Für die Streuung der diskreten Zufallsgröße x mit der Verteilungstabelle (2.25) erhalten wir gemäß Definition

$$\sigma^2 = \sum_{x=0}^{n} [x - E(x)]^2\, P_n(x) = \sum_{x=0}^{n} (x - np)^2\, P_n(x)$$

oder ausführlicher

$$\sigma^2 = \sum_{x=0}^{n} x^2\, P_n(x) - 2np \sum_{x=0}^{n} x\, P_n(x) + n^2p^2 \sum_{x=0}^{n} P_n(x)\,.$$

Beachten wir im vorstehenden Ausdruck die Beziehungen (2.26) und (2.29), so ergibt sich

$$\sigma^2 = \sum_{x=0}^{n} x^2 P_n(x) - n^2p^2 . \tag{2.31}$$

Um die Summe

$$S_1 = \sum_{x=0}^{n} x^2 P_n(x) \tag{2.32}$$

zu berechnen, ermitteln wir zunächst den Ausdruck

$$S_2 = \sum_{x=0}^{n} x(x-1) P_n(x),$$

der mit S_1 durch die Relation

$$S_1 = S_2 + np \tag{2.33}$$

verknüpft ist. Süг S_2 erhalten wir den Term

$$S_2 = \sum_{x=2}^{n} \frac{x(x-1)\,n!}{x!\,(n-x)!} p^x q^{n-x}$$

$$= n(n-1)\,p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!\,[(n-2)-(x-2)]!} p^{x-2} \cdot q^{(n-2)-(x-2)} ,$$

der nach Einführung der Abkürzungen $n-2=r$ und $x-2=z$ die einfache Gestalt

$$S_2 = n(n-1)\,p^2 \sum_{z=0}^{r} \binom{r}{z} p^z q^{r-z}$$

annimmt. Wegen der Vollständigkeitsrelation (2.26) hat die Summe in der letzten Gleichung den Wert 1. Mithin bleibt

$$S_2 = n(n-1) \cdot p^2$$

oder gemäß (2.33), (2.32), und (2.31)

$$\sigma^2 = n(n-1)\,p^2 + np - n^2p^2 = np\,(1-p) = npq .$$

Damit ist der Satz vollständig bewiesen.

In den von uns bisher behandelten Beispielen ergeben sich nun folgende Mittelwerte und Streuungen: Der Erwartungswert für die Anzahl der normgerechten Erzeugnisse unter 10 Erzeugnissen beträgt $E(x) = 0{,}96 \cdot 10 = 9{,}6$. Der Erwartungswert für die Anzahl der Tage mit normalem Wasserverbrauch im Laufe einer Woche beläuft sich auf

$$E(x) = 0{,}75 \cdot 6 = 4{,}5 .$$

Die mittlere Anzahl der unter 100 Erzeugnissen zur Sorte 1 gehörigen Elemente hat den Wert $E(x) = 0{,}6 \cdot 100 = 60$. Die einzelnen Streuungen betragen der Reihe nach

$$\sigma^2 = 0{,}384, \quad \sigma^2 = 1{,}125 \text{ und } \sigma^2 = 24 .$$

Setzen wir in die allgemeine TSCHEBYSCHEFFsche Ungleichung (2.21) die Parameterwerte (2.29) und (2.30) für die BERNOULLIsche Verteilung ein, so ergibt sich die BERNOULLI*sche Streuungsungleichung* zu

$$\boxed{P(|x - n \cdot p| < \gamma) \geqq 1 - \frac{npq}{\gamma^2}} \; ; \tag{2.34}$$

dabei ist γ eine beliebige Konstante mit $\gamma^2 \geqq npq$.

Wie im allgemeinen Fall so liefert auch die Ungleichung (2.34) sehr grobe Abschätzungen für die tatsächlichen Wahrscheinlichkeiten. In der Praxis werden meistens genauere Näherungsverfahren zur Ermittlung von Wahrscheinlichkeiten $P(|x - np| < \gamma)$ bei der binomischen Verteilung benutzt. Wir werden sie bei der Normalverteilung kennenlernen.

Bevor wir eine wichtige Schlußfolgerung aus der Beziehung (2.34) ziehen, geben wir noch zwei Beispiele: In einem Betrieb sind 95 Prozent aller Erzeugnisse fehlerfrei. Wieviel normgerechte Erzeugnisse werden in 200000 Stück erwartet? Wie groß ist die Wahrscheinlichkeit dafür, daß die tatsächliche Zahl der fehlerfreien Erzeugnisse zwischen 188000 und 192000 liegt? Den gesuchten Erwartungswert erhalten wir vermöge (2.29) zu

$$E(x) = np = 200000 \cdot 0{,}95 = 190000 \,.$$

Die betreffende Streuung hat nach (2.30) den Wert

$$\sigma^2 = npq = 190000 \cdot 0{,}05 = 9500 \,.$$

Aus der Ungleichung (2.34) ergibt sich dann weiter die gesuchte Wahrscheinlichkeit zu

$$P(188000 < x < 192000) = P(|x - E(x)| < 2000) =$$
$$\geqq 1 - \frac{9500}{(2000)^2} \approx 0{,}998 \,.$$

Es werden also 190000 fehlerfreie Erzeugnisse erwartet. Mit einer Wahrscheinlichkeit von mindestens 99,8 Prozent liegt die tatsächliche Zahl oer fehlerfreien Erzeugnisse weiterhin zwischen den Grenzen 188000 und 192000.

Das zweite Beispiel entnehmen wir der Arbeit [3]. Ein Viertel der Beschäftigten eines Industriezweiges hat Oberschulbildung. Für eine Untersuchung werden nach dem Zufallsprinzip 200000 Beschäftigte ausgesucht. Wie groß ist der Erwartungswert für die Anzahl der Beschäftigten mit Oberschulbildung unter diesen 200000? Wie groß ist die Wahrscheinlichkeit dafür, daß die wirkliche Anzahl um weniger als 1,6 Prozent vom Erwartungswert abweicht? Der Mittelwert $E(x)$ beläuft sich zunächst auf $E(x) = np = 200000 \cdot 0{,}25 = 50000$; für die Streuung ergibt sich

$$\sigma^2 = npq = 50000 \cdot 0{,}75 = 37500 \,.$$

Aus der Streuungsungleichung (2.34) folgt dann

$$P(|x = E(x)| < 0{,}016 \cdot E(x) \geqq 1 - \frac{37500}{(800)^2} \approx 0{,}941 \,.$$

Im Mittel werden also 5000 Beschäftigte Oberschulbildung besitzen. Mit einer Wahrscheinlichkeit von 94,1 Prozent weicht die gesuchte Zahl der Beschäftigtən um weniger als 1,6 Prozent vom Erwartungswert, d. h. um weniger als 800 von dem Wert 50000 ab.

Wir leiten nun aus der Streuungsungleichung (2.34) den *Satz von* BERNOULLI her, der eine fundamentale Rolle in der Wahrscheinlichkeitsrechnung spielt.

Satz 2,18: Wählen wir die Zahl n der Ziehungen hinreichend groß, dann ist mit einer nahe bei 1 liegenden Wahrscheinlichkeit zu erwarten, daß die relative Häufigkeit $\frac{x}{n}$ der gezogenen weißen Kugeln von der Wahrscheinlichkeit p nur unbedeutend abweicht.

Beweis: Setzen wir in der Ungleichung (2.34) $\gamma = n\varepsilon$, dann erhalten wir

$$P\left(\left|\frac{x}{n} - p\right| < \varepsilon\right) \geqq 1 - \frac{p(1-p)}{n\varepsilon^2}. \tag{2.35}$$

Die Größe ε sei nun eine beliebig kleine positive Zahl. Die Anzahl der Ziehungen n wählen wir derart groß, daß auch $n\varepsilon^2$ über alle Grenzen wächst. Damit haben wir aber schon die Behauptung

$$P\left(\frac{x}{n} - p > \varepsilon\right) = 0$$

für hinreichend kleine $\varepsilon > 0$ und genügend große n.

Die Bedeutung des BERNOULLIschen Satzes besteht in folgendem: Wir betrachten eine Serie von Versuchen, die alle unabhängig voneinander ablaufen, und beobachten dabei xmal das Ereignis A (und $[n - x]$-mal das Ereignis $\bar{A}$.) Dann ergibt sich für das Auftreten des Ereignisses A innerhalb dieser Serie die relative Häufigkeit

$$h_n(A) = \frac{x}{n}. \tag{2.36}$$

Wenn wir nun genügend viele Versuche, also eine sehr umfangreiche Versuchsserie durchführen, dann strebt die relative Häufigkeit (2.36) nach dem vorstehenden Satz gegen die Wahrscheinlichkeit des Ereignisses A, d. h., es gilt dann die bei der statistischen Definition der Wahrscheinlichkeit angegebene Beziehung

$$\boxed{\lim_{n \to \infty} h_n(A) = P(A)} \tag{2.37}$$

2.4. Die POISSONsche Verteilung

Wir führen im folgenden den auf POISSON zurückgehenden Grenzübergang durch, der darin besteht, daß die BERNOULLIsche Formel (2.24) für große n

betrachtet wird, wobei allerdings die Größe $a = np$ als angenähert konstant angesehen wird. Zu diesem Zweck formen wir zunächst (2.24) zu

$$P_n(x) = \frac{a^x}{x!}\left(1 + \frac{-a}{n}\right)^n K(n, x) \tag{2.38}$$

mit

$$K(n, x) = \frac{\left(1 - \frac{a}{n}\right)^{-x} n!}{n^x \cdot (n - x)!} = \left(1 - \frac{a}{n}\right)^{-x} \prod_{j=0}^{x-1}\left(1 - \frac{j}{n}\right)$$

um. Beachten wir

$$\lim_{n \to \infty} K(n, x) = 1$$

und

$$\lim_{n \to \infty} \left(1 + \frac{\delta}{n}\right)^n = e^\delta ,$$

so ergibt sich aus (2.38) die Beziehung

$$\boxed{P_n(x) = \xrightarrow[a = np \text{ konstant}]{n \longrightarrow \infty} P^*(x) = \frac{a^x e^{-a}}{x!}} \tag{2.39}$$

Wir definieren nun als Poissonsche *Verteilung* diejenige Verteilung, zu der die diskrete Zufallsvariable x mit der Verteilungstabelle

Wert	0	1	2 ...	k ...
Wahrscheinlichkeit	$P^*(0)$	$P^*(1)$	$P^*(2) \ldots$	$P^*(k) \ldots$

(2.40)

gehört. Als Poissonsche *Verteilungsfunktion* bezeichnen wir entsprechend die Funktion

$$\boxed{F^*(\xi) = P(x < \xi) = \sum_{x < \xi} P^*(x)} \tag{2.41}$$

Wir leiten nun einige Eigenschaften der Poissonschen Verteilung ab. Zunächst zeigen wir, daß die Tabelle (2.40) in der Tat die Vollständigkeitsrelation

$$\sum_{x=0}^{\infty} P_n^*(x) = 1 \tag{2.42}$$

erfüllt. Beachten wir

$$\sum_{x=0}^{\infty} \frac{\delta^x}{x!} = e^\delta , \tag{2.43}$$

so erhalten wir in Verbindung mit (2.39)

$$\sum_{x=0}^{\infty} P^*(x) = e^{-a} \sum_{x=0}^{\infty} \frac{a^x}{x!} = e^{-a} e^a = 1 .$$

Weiter beweisen wir den

Satz 2.19: Für die Parameter der POISSONschen Verteilung gilt die Beziehung

$$\boxed{E(x) = \sigma^2 = a = np} \tag{2.44}$$

Beweis: Zunächst ist gemäß der Definition (2.5) des Erwartungswertes und der Gleichung (2.43)

$$E(x) = \sum_{x=0}^{\infty} \frac{x a^x e^{-a}}{x!} = a e^{-a} \sum_{x=1}^{\infty} \frac{a^{x-1}}{(x-1)!} = a e^{-a} e^{a} = a\,. \tag{2.45}$$

Für die Streuung gilt definitionsgemäß

$$\sigma^2 = \sum_{x=0}^{\infty} [x - E(x)]^2 \frac{a^x e^{-a}}{x!}$$

oder ausführlicher mit $E(x) = a$

$$\sigma^2 = \sum_{x=0}^{\infty} x^2 \frac{a^x e^{-a}}{x!} - 2a \sum_{x=0}^{\infty} x \frac{a^x e^{-a}}{x!} + a^2 \sum_{x=0}^{\infty} \frac{a^x e^{-a}}{x!}\,.$$

Berücksichtigen wir (2.45) und (2.42), dann bleibt

$$\sigma^2 = \sum_{x=0}^{\infty} x^2 \frac{e^{-a} a^x}{x!} - a^2 \equiv S - a^2\,. \tag{2.46}$$

Für den Ausdruck S schreiben wir

$$S = \sum_{x=0}^{\infty} x(x-1) \frac{a^x e^{-a}}{x!} + \sum_{x=0}^{\infty} \frac{x a^x e^{-a}}{x!}$$

Bei Beachtung der Beziehung (2.45) vereinfacht sich der vorstehende Term zu

$$S = a^2 \sum_{x=2}^{n} \frac{a^{x-2} e^{-a}}{(x-2)!} + a\,.$$

Setzen wir schließlich noch $x - 2 = u$, dann ist unmittelbar ersichtlich, daß der unter dem Summenzeichen der letzten Gleichung stehende Ausdruck gemäß (2.42) den Wert 1 hat. Damit haben wir $S = a^2 + a$ und wegen (2.46) die Behauptung $\sigma^2 = a$.

Wir haben oben festgestellt, daß die Wahrscheinlichkeit $P_n(x)$ für große n und kleine p durch die Wahrscheinlichkeit $P^*(x)$ angenähert wird. Um ein Bild von dieser Näherung zu erhalten, vergleichen wir im folgenden einige Werte von $P_n(x)$ und $P^*(x)$ für $n = 500$ und $p = \frac{1}{365}$ miteinander:

x	$P^*(x)$	$P_n(x)$	$\Delta(x) = P^*(x) - P_n(x)$
0	0,2541	0,2537	0,0004
1	0,3484	0,3484	0
2	0,2385	0,2388	— 0,0003
3	0,1089	0,1089	0
4	0,0372	0,0372	0
5	0,0102	0,0101	0,0001
6	0,0023	0,0023	0

Die $P^*(x)$ stellen also eine brauchbare Näherung für die Größen $P_n(x)$ dar, sobald n groß und p klein ist. In der Praxis wird daher stets $P^*(x)$ anstelle von $P_n(x)$ angewandt, wenn die erwähnten Bedingungen erfüllt sind, da die Berechnung von $P^*(x)$ gemäß (2.39) unter Verwendung von Tafeln bedeutend einfacher ist als die Ermittlung der Größe $P_n(x)$ nach der Formel (2.24) mit den schwer zu bewältigenden Binomialkoeffizienten.

Wir geben zum Abschluß noch drei praktische Beispiele zur POISSONschen Verteilung.

In einem Werk sind 1,5 Prozent der gefertigten Schrauben fehlerhaft. Wie groß ist der Erwartungswert für die Anzahl der fehlerhaften Schrauben in einem Posten von 100 Stück? Wie groß ist die Wahrscheinlichkeit dafür, daß in einer 100-Stück-Packung keine fehlerhaften Schrauben sind? Für den Erwartungswert erhalten wir

$$E(x) = np = 100 \cdot 0{,}015 = 1{,}5.$$

Es sind also 1—2 Schrauben im Mittel in einem Posten von 100 Stück fehlerhaft. Die gesuchte Wahrscheinlichkeit ist genau $P_{100}(o)$ oder vermöge (2.24)

$$P_{100}(0) = (1 - 0{,}015)^{100} \approx 0{,}221\,.$$

Nähern wir $P_{100}(o)$ durch $P^*(o)$ gemäß (2.39) an, so bekommen wir

$$P_{100}(0) \approx P^*(0) = \frac{(1{,}5)^0\, e^{-1{,}5}}{0!} \approx 0{,}223\,.$$

Der Näherungsfehler beträgt also lediglich 0,002.

In einem Blechband von 100 m Länge treten durchschnittlich 5 Fehler auf. Das Band wird in kleine Bleche von 3 m Länge zerschnitten. Wieviel fehlerhafte Blechstücke sind zu erwarten? Zunächst ist die Wahrscheinlichkeit dafür, daß in einem Meter Blech ein Fehler auftritt, gleich $\frac{1}{20}$. In n Metern Blech treten dann x Fehler mit der Wahrscheinlichkeit

$$P_n(x) \approx P^*(x) = \frac{\left(\frac{n}{20}\right)^x e^{-\frac{n}{20}}}{x!} \tag{*}$$

auf. Für die gesuchte Wahrscheinlichkeit $P_3(o)$ dafür, daß 3 Meter Blech fehlerfrei sind, ergibt sich daher aus (*) die Näherung

$$P^*(0) = \frac{e^{-0{,}15}}{0!} \approx 0{,}86$$

Zerschneiden wir das Band in Stücke zu 3 Metern, dann sind

$$\frac{100}{3} \cdot 0{,}86 \approx 29$$

fehlerhafte Bleche zu erwarten.

In einem Geschäft mögen in der Zeiteinheit (z. B. einer Stunde) im Mittel λ Kunden rein zufällig eintreten. Der Erwartungswert für die Anzahl der während einer beliebigen Zeitspanne T eintreffenden Kunden ist dann λT. Nach dem POISSONschen Gesetz (2.39) liefert dann

$$E(\lambda;\, x;\, T) \equiv \frac{(\lambda T)^x}{x!}\, e^{-\lambda T} \tag{2.47}$$

die Wahrscheinlichkeit dafür, daß innerhalb des Zeitraumes T genau x Kunden zufällig eintreffen. Die Formel (2.47) spielt in der Theorie der Massenbedienung eine wesentliche Rolle (vgl. hierzu Kapitel 4).

2.5. Kontinuierliche Verteilungen

2.5.1. *Dichtefunktion und Verteilungsfunktion*

Wir betrachten im folgenden eine *stetige* oder *kontinuierliche* Zufallsgröße, d. h. eine zufällige Variable, die in jedem Intervall der reellen Zahlengeraden mit einer bestimmten Wahrscheinlichkeit liegt. Wir nennen nun die stetige Funktion $w(x)$ die *kontinuierliche Dichtefunktion* der kontinuierlichen Zufallsgröße x, wenn folgende Bedingungen erfüllt sind:

a) Für alle x der reellen Zahlengeraden gilt

$$\boxed{w(x) \geqq 0} \tag{2.48}$$

b) Für beliebige reelle Zahlen a und b mit $a < b$ besteht

$$\boxed{\int_a^b w(x)\,dx = P(a \leqq x < b)}\,. \tag{2.49}$$

Die erste geforderte Bedingung besagt, daß die Funktion $w(x)$ nirgends auf der reellen Achse negativ ist. Die zweite Forderung (2.49) bedeutet, daß der Inhalt des Flächenstücks zwischen der Kurve $y = w(x)$, den Geraden $x = a$ und $x = b$ sowie dem Intervall $a \leqq x < b$ gleich der Wahrscheinlichkeit dafür ist, daß die Zufallsgröße x diesem Intervall angehört.

Auf Grund der Stetigkeit der Funktion $y = w(x)$ gelten offenbar neben (2.49) noch die Beziehungen

$$\boxed{\int_a^b w(x)\,dx = P(a < x \leqq b) = P(a \leqq x \leqq b) = P(a < x \leqq b)}\,. \tag{2.50}$$

Wir beweisen nun den

Satz 2.20: Für jede kontinuierliche Dichtefunktion $w(x)$ besteht die Vollständigkeitsrelation

$$\boxed{\int_{-\infty}^{+\infty} w(x)\,dx = 1}\,. \tag{2.51}$$

Beweis: Aus (2.50) folgt

$$\int_{-\infty}^{+\infty} w(x)\,dx = P(-\infty < x < \infty)\,.$$

Das Ereignis $A \equiv {}'' -\infty < x < \infty''$ ist offenbar gleich dem Ereignis „die Variable x nimmt einen beliebigen reellen Wert an“. Daher gilt $A = S$ und damit $P(A) = 1$. Der Satz ist hiermit bewiesen.

Geometrisch bedeutet die Relation (2.51), daß die Fläche zwischen der Kurve der Funktion $y = w(x)$ und der reellen Zahlengeraden den Wert 1 hat (siehe die folgende Abbildung!).

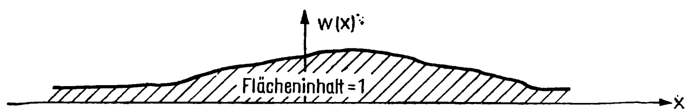

Abb. 8 Dichtefunktion einer stetigen Zufallsgröße

Die Dichtefunktion $w(x)$ einer stetigen Zufallsgröße entspricht offenbar der Verteilungstabelle (2.1) einer diskreten Zufallsvariablen. Insbesondere entspricht der Vollständigkeitsrelation $p_1 + p_2 + \ldots + p_n = 1$ im diskreten Fall die Vollständigkeitsrelation (2.51) im kontinuierlichen Fall. Kennen wir die Dichtefunktion $y = w(x)$, so sind wir in der Lage (wie bei Kenntnis der Verteilungstabelle (2.1) im diskreten Fall!) alle Wesenszüge der betreffenden kontinuierlichen Verteilung anzugeben.

Vielfach ist nicht die Dichtefunktion $w(x)$ einer kontinuierlichen Zufallsgröße x, sondern nur eine zu $w(x)$ proportionale Funktion $w^*(x)$ bekannt. Dann gilt aber zunächst

$$w(x) = c \cdot w^*(x)\,, \tag{2.52}$$

wobei c ein Proportionalitätsfaktor ist. Beachten wir weiter die für $w(x)$ gültige Beziehung (2.51), so erhalten wir für c den Wert

$$c = \frac{1}{\int_{-\infty}^{+\infty} w^*(x)\,dx}$$

und damit gemäß (2.52)

$$\boxed{w(x) = \frac{w^*(x)}{\int_{-\infty}^{+\infty} w^*(x)\,dx}}\,. \tag{2.53}$$

Damit haben wir die gesuchte Dichtefunktion $w(x)$ eindeutig bestimmt.

Wir erläutern den vorstehenden Sachverhalt an einem Beispiel: Gegeben sei die Funktion

$$w^* (x) = \frac{1}{1 + x^2},$$

die einer Dichtefunktion für die Zufallsgröße x proportional sein soll. Wir setzen $w (x) = \frac{c}{1 + x^2}$ mit c als konstantem Faktor und erhalten aus (2.51)

$$c = \frac{1}{\int\limits_{-\infty}^{+\infty} \frac{dx}{1 + x^2}} \tag{*}$$

Um das im Nenner des vorstehenden Ausdrucks stehende Integral zu berechnen, beachten wir, daß $w^* (x) = \frac{1}{1 + x^2}$ eine gerade Funktion ist, d. h. $w^* (-\alpha) = w^* (\alpha)$ für jedes reelle α erfüllt ist (und damit die Kurve von $y = w^* (x)$ spiegelbildlich zur y-Achse liegt!). Wir erhalten dann

$$\int\limits_{-\infty}^{+\infty} \frac{d\,x}{1 + x^2} = 2 \int\limits_{0}^{\infty} \frac{d\,x}{1 + x^2} = 2 \operatorname{arc\,tg} x \Big|_0^\infty = \pi .$$

Berücksichtigen wir das letzte Ergebnis in (*), so haben wir unsere Dichtefunktion $w (x)$ eindeutig zu

$$w (x) = \frac{1}{\pi (1 + x^2)} \tag{2.54}$$

ermittelt.

Wir verweilen noch einen Augenblick bei der Dichtefunktion (2.54), um den Verlauf der zugehörigen Kurve kennenzulernen. Auf Grund der oben erwähnten Eigenschaft von $w^* (x)$ ist auch das Bild von $w (x)$ symmetrisch zur y-Achse. Weiter erkennen wir die einzigen Nullstellen von $w (x)$ in den uneigentlichen Punkten $x = \pm \infty$. Bei $x = 0$ hat die Kurve ferner ein Maximum (das einzige Extremum!), da

$$w' (x) = \frac{-2\,x}{\pi (1 + x^2)^2} \tag{2.55}$$

außer an der Stelle $x = 0$ nirgends im Endlichen verschwindet und

$$w'' (x) = \frac{2}{\pi} \frac{3\,x^2 - 1}{(1 + x^2)^3} \tag{2.56}$$

im Ursprung negativ ist. Der zugehörige Maximalwert der Dichtefunktion ergibt sich nach (2.54) zu $\frac{1}{\pi}$. Die Beziehung (2.55) sagt weiter aus, daß die Kurve von $-\infty$ bis 0 monoton steigt, im Punkt $\left(0, \frac{1}{\pi}\right)$ kulminiert und von 0 bis ∞ monoton fällt. Schließlich entnehmen wir noch der Gleichung (2.56), daß die einzigen Wendepunkte der in Rede stehenden Kurve bei den Abzsissen $x = \pm \frac{1}{\sqrt{3}}$ liegen; die zugehörigen (auf Grund der Symmetrie) gleichen Wendepunktsordinaten besitzen den Wert $\frac{3}{4\pi}$. Nun sind wir in der Lage, die Gestalt der Kurve der Dichtefunktion (2.54) anzugeben.

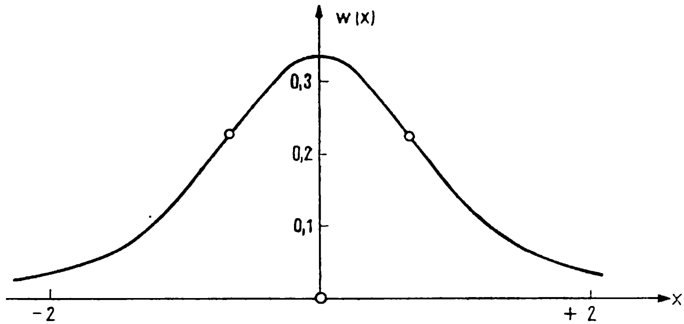

Abb. 9: Verlauf der Dichtefunktion $w(x) = \frac{1}{\pi(1+x^2)}$

Wir kehren nun zu der allgemeinen kontinuierlichen Verteilung mit der Dichtefunktion $w(x)$ zurück. Als *kontinuierliche Verteilungsfunktion* führen wir die Funktion

$$\boxed{W(\xi) = P(x < \xi)} \tag{2.57}$$

ein und beweisen den

Satz 2.21: Für die Verteilungsfunktion $W(\xi)$ der kontinuierlichen Zufallsvariablen x mit der Dichtefunktion $w(x)$ gilt die Darstellung

$$\boxed{W(\xi) = \int_{-\infty}^{\xi} w(x)\, dx}\,. \tag{2.58}$$

Beweis: Der Beweis der vorstehenden Aussage ergibt sich aus den Beziehungen (2.57) und (2.50):

$$W(\xi) = P(x < \xi) = P(-\infty < x < \xi) = \int_{-\infty}^{\xi} w(x)\, dx\,.$$

Aus Satz 2.21 folgt der

Satz 2.22: Die Verteilungsfunktion $W(\xi)$ einer kontinuierlichen Zufallsgröße ist eine monoton wachsende und stetige Funktion, die für $\xi = -\infty$ verschwindet und für $\xi = \infty$ den Wert 1 annimmt.

Beweis: Differenzieren wir das in (2.58) vorkommende Integral nach der oberen Grenze ξ, dann erhalten wir nach den Regeln der Analysis

$$\frac{d\,W(\xi)}{d\xi} = W'(\xi) = w(\xi). \tag{2.59}$$

Wegen der Stetigkeit von $w(\xi)$ ist $W(\xi)$ stetig differenzierbar und damit erst recht stetig in ξ. Aus (2.59) und (2.48) folgt $W'(\xi) \geqq 0$, d. h. die Tatsache, daß $W(\xi)$ monoton steigt. Aus der Beziehung (2.58) erhalten wir weiter die Relation

$$W(-\infty) = 0. \tag{2.60}$$

denn aus (2.58) resultiert

$$W(-\infty) = \int_{-\infty}^{-\infty} w(x)\,dx = 0.$$

(Ein Integral zwischen gleichen Grenzen verschwindet!). Schließlich liefert

$$W(+\infty) = 1 \tag{2.61}$$

wegen (2.58) die schon bewiesene Vollständigkeitsrelation (2.51). Damit ist der Satz bewiesen.

Die *Verteilungskurve* einer kontinuierlichen Zufallsvariablen, d. h. die Kurve der zugehörigen Verteilungsfunktion $W(\xi)$ zeigt auf Grund des vorstehenden Satzes 2.22 folgenden Verlauf:

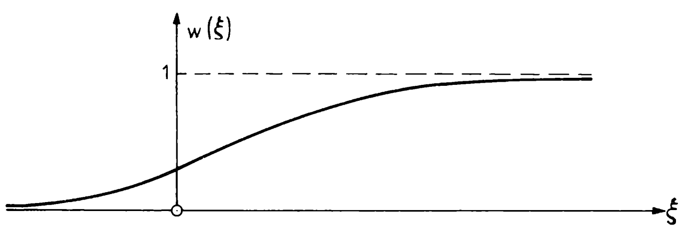

Abb. 10 Verteilungskurve einer stetigen Zufallsgröße

Die uns schon vertraute Dichtefunktion (2.54) besitzt gemäß (2.58) die Verteilungsfunktion

$$W(\xi) = \frac{1}{\pi} \int_{-\infty}^{\xi} \frac{dx}{1+x^2} = \frac{1}{\pi} \operatorname{arc\,tg} x \Big|_{-\infty}^{\xi} = \frac{1}{\pi}\left(\operatorname{arc\,tg} \xi + \frac{\pi}{2}\right)$$

oder kurz

$$W(\xi) = \frac{1}{\pi} \operatorname{arc\,tg} \xi + \frac{1}{2}. \tag{2.62}$$

Die entsprechende Verteilungskurve hat die Gestalt

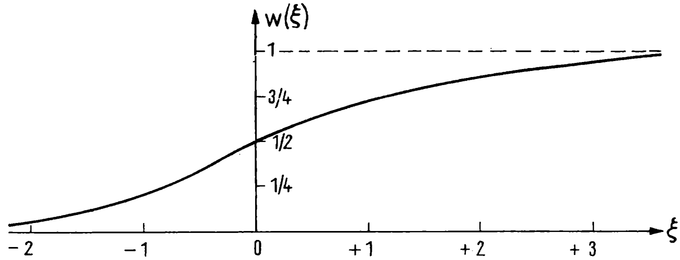

Abb. 11 Verteilungskurve zu Abb. 9

Wir beweisen nun den

Satz 2.23: Ist $W(\xi)$ die Verteilungsfunktion einer kontinuierlichen Zufallsgröße x, dann besteht die Beziehung

$$\boxed{P(a \leqq x < b) = W(b) - W(a)}\,. \tag{2.63}$$

Beweis: Die Verteilungsfunktion $W(x)$ ist wegen (2.59) eine unbestimmte Lösung des Integrals in (2.49). Daher folgt aus (2.49)

$$P(a \leqq x < b) = \int_a^b w(x)\,dx = W(x)\Big|_a^b = W(b) - W(a)\,.$$

Damit ist schon alles gezeigt. Da weiter nach dem Satz 2.22 die Verteilungsfunktion $W(\xi)$ bezüglich ihres Arguments stetig ist, gelten neben (2.63) auch noch die Relationen

$$\boxed{P(a < x < b) = P(a \leqq x \leqq b) = P(a < x \leqq b) = W(b) - W(a)} \tag{2.64}$$

Wir wenden uns nun wieder der stetigen Zufallsvariablen x mit der Dichte (2.54) und der Verteilungsfunktion (2.62) zu. Wir fragen zunächst nach der Wahrscheinlichkeit p_1 dafür, daß x dem Intervall $\left(-\frac{1}{\sqrt{3}} \leqq x \leqq \frac{1}{\sqrt{3}}\right)$ angehört, d. h. daß x einen beliebigen Punkt der reellen Achse zwischen den Wendepunktabszissen annimmt. Aus (2.64) folgt vorerst

$$P_1 = P\left(-\frac{1}{\sqrt{3}} \leqq x \leqq \frac{1}{\sqrt{3}}\right) = W\left(\frac{1}{\sqrt{3}}\right) - W\left(-\frac{1}{\sqrt{3}}\right).$$

Beachten wir (2.62), so wird weiter

$$p_1 = \frac{1}{\pi}\left[\operatorname{arc\,tg}\frac{1}{\sqrt{3}} - \operatorname{arc\,tg}\left(-\frac{1}{\sqrt{3}}\right)\right] = \frac{2}{\pi}\operatorname{arc\,tg}\frac{1}{\sqrt{3}}\,.$$

Der tg-Tafel entnehmen wir den Wert $\operatorname{arc\,tg}\frac{1}{\sqrt{3}} = \operatorname{arc\,tg} 30^0 = \frac{\pi}{6}\,.$

Damit haben wir das Ergebnis $p_1 = \frac{1}{3}$. Die Wahrscheinlichkeit dafür, daß x dem Intervall $-\frac{1}{\sqrt{3}} \leqq x \leqq \frac{1}{\sqrt{3}}$ angehört, ist demzufolge halb so groß wie die Wahrscheinlichkeit dafür, daß x diesem Intervall nicht angehört. Mit anderen Worten nimmt die über diesem Intervall liegende Fläche unterhalb der Kurve der Dichtefunktion ein Drittel der Gesamtfläche ein.

Ganz entsprechend berechnen wir die Wahrscheinlichkeit $p_2 = P(|x| \leqq 1)$ zu $p_2 = P(-1 \leqq x \leqq 1) = W(1) - W(-1) = \frac{2}{\pi} \operatorname{arc\,tg} 1 = \frac{1}{2}$.

Die Fläche zwischen der Kurve der Dichtefunktion, der x-Achse und den Geraden $x = \pm 1$ ist demzufolge die Hälfte der Gesamtfläche. Führen wir die Beziehungen $A \equiv$" $-1 \leqq x \leqq 1$" und $B \equiv$" $-\infty < x < 1$" + "$1 < x < +\infty$" ein, dann sind die Ereignisse A und B offenbar gleichwahrscheinlich mit der Wahrscheinlichkeit $\frac{1}{2}$.

Schließlich fragen wir in diesem Zusammenhang nach der Zahl α, für die

$$P(x \leqq \alpha) = \frac{3}{\pi} \approx 0{,}96 \qquad (*)$$

erfüllt ist. Mit anderen Worten: Wie groß muß die Konstante α gewählt werden, damit über dem Intervall $-\alpha \leqq x \leqq +\alpha$ bereits 96 Prozent der Fläche zwischen der Kurve der Dichtefunktion und der x-Achse liegen? Aus (2.62), (2.64) und (*) erhalten wir

$$P(-\alpha \leqq x \leqq \alpha) = W(\alpha) - W(-\alpha) = \frac{2}{\pi} \operatorname{arc\,tg} \alpha = \frac{3}{\pi}$$

oder kurz

$$\operatorname{arc\,tg} \alpha = \frac{\pi}{2} \cdot \frac{3}{\pi} = \frac{3}{2}.$$

Der tg-Tafel entnehmen wir den Wert $\alpha \approx 14{,}1$. Außerhalb des Intervalls $-14{,}1 \leqq x \leqq 14{,}1$ tritt die kontinuierliche Zufallsgröße x also nur mit einer Wahrscheinlichkeit von annähernd 4 Prozent auf.

Verdeutlichen wir die letzte Diskussion noch an einer Abbildung:

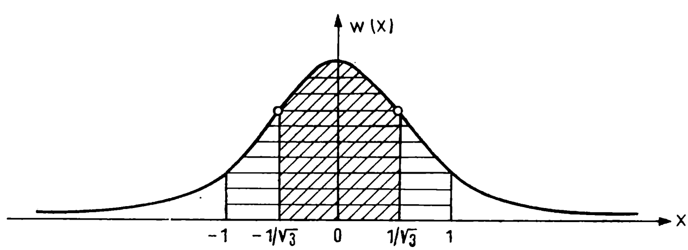

Abb. 12: Verlauf der Dichtefunktion $w(x) = \frac{1}{\pi(1+x^2)}$

Die schräg gestrichelte Fläche hat den Inhalt $\frac{1}{3}$, die horizontal gestrichelte dagegen $\frac{1}{2}$.

2.5.2. *Erwartungswert und Streuung kontinuierlicher Verteilungen*

Als *Erwartungswert* $E(x)$ bzw. als *Streuung* $\sigma^2(x)$ der stetigen Zufallsvariablen x mit der Dichtefunktion $w(x)$ führen wir ohne nähere Begründung die Integrale[1]

$$\boxed{E(x) = \int_{-\infty}^{+\infty} x\, w(x)\, dx} \qquad \boxed{\sigma^2 = \int_{-\infty}^{+\infty} [x - E(x)]^2\, w(x)\, dx} \tag{2.65}$$

ein, die als Verallgemeinerung der Summen (2.5) und (2.15) im diskreten Fall angesehen werden können.

Ohne Beweis (vgl. hierzu [4]) erwähnen wir den

Satz 2.24: Alle in den Sätzen 2.4—2.13 hergeleiteten Eigenschaften des Erwartungswertes $E(x)$ und der Streuung $\sigma^2(x)$ gelten auch im Fall einer stetigen Zufallsgröße x.

Wir berechnen nachstehend für zweistetige Verteilungen die Parameter $E(x)$ und $\sigma^2(x)$. Zunächst erörtern wir die Verteilung (2.54), die wegen (2.65_1) den Erwartungswert

$$E(x) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x dx}{1 + x^2} \tag{2.66}$$

besitzt. Um das vorstehende Integral zu bestimmen, untersuchen wir vorerst das Integral

$$I(\alpha) = \int_{-\alpha}^{+\alpha} \frac{x}{1 + x^2}\, dx\,.$$

Nach den Regeln der Analysis wird

$$I(\alpha) = \frac{1}{2} \int_{-\alpha}^{+\alpha} \frac{(1 + x^2)'}{1 + x^2}\, dx = \frac{1}{2} \ln(1 + x^2) \Big|_{-\alpha}^{+\alpha} = 0\,.$$

Beachten wir $I(\alpha) = 0$ in (2.66), so erhalten wir das Ergebnis

$$E(x) = \frac{1}{\pi} \lim_{\alpha \to \infty} I(\alpha) = \frac{1}{\pi} \cdot 0 = 0\,.$$

Der Erwartungswert der Verteilung (2.54) ist also null. Dieses Resultat ergibt sich auch unmittelbar aus der Symmetrie der Kurve der Dichtefunktion (2.54) Das Ergebnis folgt sogar sofort aus der Tatsache, daß die Funktion $\frac{x}{1 + x^2}$ eine ungerade Funktion ist.

[1] Selbstverständlich wird die Konvergenz der Integrale vorausgesetzt.

Wir interessieren uns nun für die Streuung der kontinuierlichen Verteilung (2.54). Gemäß (2.65_2) und $E(x) = 0$ ist zunächst

$$\sigma^2 = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x^2}{1+x^2}\, dx\,. \tag{2.67}$$

Wir behandeln wieder vorerst das Integral

$$I^*(\alpha) = \int_{-\alpha}^{+\alpha} \frac{x^2}{1+x^2}\, dx\,.$$

Wegen $\frac{x^2}{1+x^2} = 1 - \frac{1}{1+x^2}$ erhalten wir

$$I^*(x) = \int_{-\alpha}^{+\alpha} \left(1 - \frac{1}{1+x^2}\right) dx = x - \operatorname{arc\,tg} x \Big|_{-\alpha}^{+\alpha} = 2\,(\alpha - \operatorname{arc\,tg} \alpha)\,,$$

so daß

$$\lim_{\alpha \to \infty} I^*(\alpha) = \infty$$

wird. Demzufolge übersteigt die Streuung

$$\sigma^2(x) = \frac{1}{\pi} \lim_{\alpha \to \infty} I^*(\alpha)$$

jeden endlichen Wert.
Wir beschäftigen uns nun mit der kontinuierlichen Zufallsgröße x, die die Dichtefunktion

$$w(x) = \begin{cases} \frac{1}{2} e^{x}, & \text{wenn } x \leqq 0\,; \\ \frac{1}{2} e^{-x}, & \text{wenn } x \geqq 0 \end{cases} \tag{2.68}$$

aufweist. Wir bestätigen die Gültigkeit der Vollständigkeitsrelation (2.51) für die eben eingeführte Funktion $w(x)$. In der Tat ist

$$\int_{-\infty}^{+\infty} w(x)\,dx = \frac{1}{2}\left[\int_{-\infty} e^{x}\,dx + \int_{0}^{\infty} e^{-x} dx\right] = \frac{1}{2}\left[e^{x}\Big|_{-\infty}^{0} - e^{-x}\Big|_{0}^{\infty}\right] = 1\,.$$

Die Kurve der Dichtefunktion (2.68) zeigt den Verlauf.

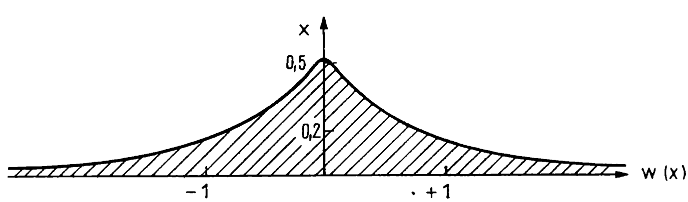

Abb. 13: Verlauf der Dichtefunktion (2 · 68)

Sie ist speziell symmetrisch zur y-Achse. Aus dieser Symmetrie folgt, daß der gemäß (2.65_1) gebildete Erwartungswert der Verteilung (2.68) verschwindet:

$$E(x) = \frac{1}{2}\left[\int\limits_{-\infty}^{0} x\, e^{x}\, d\, x + \int\limits_{0}^{\infty} x\, e^{-x}\, d\, x\right] = 0\,; \tag{2.69}$$

dieses Ergebnis ergibt sich auch schnell auf rechnerischem Wege. Die Formel (2.65_2) liefert für die Streuung der Verteilung (2.68) wegen (2.69) den Wert

$$\sigma^2 = \frac{1}{2}\left[\int\limits_{-\infty}^{0} x^2\, e^{x}\, d\, x + \int\limits_{0}^{\infty} x^2\, e^{-x}\, d\, x\right], \tag{2.70}$$

den wir nun berechnen. Beachten wir die durch mehrfache partielle Integration nachweisbare Beziehung

$$\int f(x)\, e^{\alpha x}\, d\, x = \frac{e^{\alpha x}}{\alpha}\left[f(x) - \frac{f'(x)}{\alpha} + \frac{f''(x)}{\alpha^2} \mp \ldots\right], \tag{2.71}$$

wobei $f(x)$ ein Polynom beliebiger Ordnung ist und demzufolge die Reihe in der eckigen Klammer abbricht, dann erhalten wir für das erste Integral $[f(x) = x^2, \alpha = 1)]$ in (2.70) das Ergebnis

$$\int\limits_{-\infty}^{0} x^2\, e^{x}\, d\, x = e^{x}\,(x^2 - 2\, x + 2)\,\Big|_{-\infty}^{0} = 2\,. \tag{*}$$

Für das zweite Integral $(f(x) = x^2, \alpha = -1)$ in (2.70) ergibt sich analog

$$\int\limits_{0}^{\infty} x^2\, e^{-x}\, d\, x = -\, e^{-x}\,(x^2 + 2\, x + 2)\,\Big|_{0}^{\infty} = 2\,. \tag{**}$$

Berücksichtigen wir schließlich (*) und (**) in (2.70), dann liegt das Resultat mit

$$\sigma^2 = \frac{1}{2}\,(2 + 2) = 2$$

vor.

2.5.3. *Der Zusammenhang zwischen diskreten und kontinuierlichen Verteilungen*

Wir machen im folgenden einige Bemerkungen zum Zusammenhang zwischen diskreten und stetigen Verteilungen. Zu diesem Zweck betrachten wir zunächst eine Reihe von Meßwerten $m_1, m_2, \ldots, m_n$ einer kontinuierlichen Zufallsgröße, die z. B. die Längen von n kontrollierten Erzeugnissen in 10^{-2} cm bedeuten. Wir ordnen nun die bei der Messung der Erzeugnisse angefertigte „Urliste" von Meßwerten, um die *absoluten Häufigkeiten* h_i $(i = 1, 2, \ldots, n)$

der Meßwerte zu ermitteln. Gehen wir zu den *relativen Häufigkeiten* $\bar{h}_i$ der Meßwerte m_i $(i = 1, 2, \ldots, n)$ über, so stellt das Schema

Meßwert	m_1	m_2	$m_3 \ldots m_n$
relative Häufigkeit	$\bar{h}_1$	$\bar{h}_2$	$\bar{h}_3 \ldots \bar{h}_n$

eine „*empirische Verteilung*" diskreter Natur dar. Da nun aber z. B. der Meßwert $x_K = 1{,}42$ alle möglichen Erzeugnislängen x mit $1.415 \leqq x < 1.425$ repräsentiert, handelt es sich bei dieser Verteilung genauer um eine Verteilung von Klassen der Breite 0,01 cm für die Erzeugnislänge x. Tragen wir über der jeweiligen Klassenmitte, d. h. über dem die Klasse repräsentierenden Meßwert m_i die entsprechende relative Häufigkeit $\bar{h}_i$ $(i = 1, 2, \ldots, n)$ auf und verbinden wir benachbarte der so entstehenden Punkte miteinander, so erhalten wir ein *Häufigkeitsdiagramm* für die einzelnen Klassen, die wir K_1, $K_2, \ldots, K_n$ nennen.

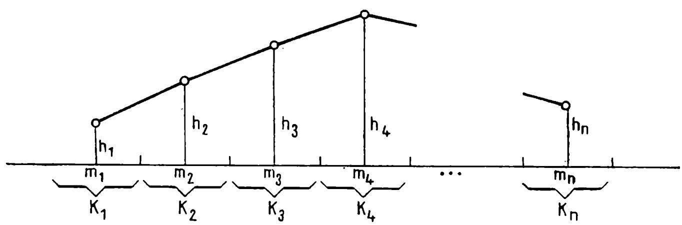

Abb. 14: Diagramm relativer Häufigkeiten von Klassen

Die Summe der Strecken muß offenbar die Einheitsstrecke ergeben. Dieses Diagramm approximiert die Dichtefunktion $w(x)$ der stetigen Zufallsgröße x (z. B. Erzeugnislänge) im Bereich der Klassen $K_1, K_2, \ldots, K_n$. Verwenden wir ein genaueres Meßinstrument, d. h. messen wir die Längen x auf 10^{-3} oder 10^{-4} cm genau, dann wird die Klasseneinteilung „feiner" und die Approximation der unbekannten Dichtefunktion durch das Klassenhäufigkeitsdiagramm „besser". Eine weitergehende Erörterung der hier aufgeworfenen Problematik erfolgt im Kapitel 4.

Diese Ausführungen lassen erkennen, daß jede stetige Verteilung durch eine diskreter Verteilung, jede Dichtefunktion $w(x)$ durch eine Verteilungstabelle der Form (2.1) angenähert werden kann. Unterteilen wir nämlich die reelle Zahlengerade in nicht notwendig gleiche Intervalle (Klassen) K_1, $K_2, K_3 \ldots$ und ordnen wir der Klasse K_i die Wahrscheinlichkeit $p_i = P$ („die stetige Zufallsgröße x tritt in der Klasse K_i auf") $(i = 1, 2, \ldots)$ zu, so haben wir eine diskrete Verteilung der Klassen K_i $(i = 1, 2, \ldots)$ mit der Verteilungstabelle

Klasse	K_1	K_2	$K_3 \ldots K_n$
Wahrscheinlichkeit	p_1	p_2	$p_3 \cdots p_n$

vor uns.

Schließlich erwähnen wir noch, daß eine diskrete Zufallsvariable x mit der Verteilungstabelle (2.1) als eine kontinuierliche Zufallsgröße mit der allerdings unstetigen Dichtefunktion

$$w(x) = \begin{cases} p_i, \text{ wenn } x = x_i \ (i = 1, 2, \ldots, n); \\ 0, \text{ sonst}. \end{cases}$$

aufgefaßt werden kann.

2.6. Die Normalverteilung

2.6.1. Die allgemeine Normalverteilung

Nach den Darlegungen des Paragraphen 2.5 können wir uns ohne große Schwierigkeiten die verschiedenen Formen stetiger Verteilungen verdeutlichen. Die Praxis hat nun gezeigt, daß die meisten Verteilungen praktischer Zufallsgrößen kontinuierlicher Natur mit einer geringen Zahl von Verteilungstypen übereinstimmen (vgl. Abschnitt 2.63). Ein solcher Typ ist die *allgemeine Normalverteilung* oder *GAUSSsche Verteilung* einer kontinuierlichen Zufallsvariablen x, die die *normale Dichtefunktion*

$$\boxed{w(x) = \frac{s}{\sqrt{\pi}} e^{-s^2 (x - t)^2} \quad (s > 0)} \tag{2.72}$$

besitzt.
Die Funktion (2.72) enthält zwei unbestimmte Konstante t und s so daß (2.72) genauer eine zweiparametrige Schar von Dichtefunktionen darstellt. Mithin gibt es eine zweiparametrige Schar von Normalverteilungen.

Wir zeigen zunächst einmal, daß die Funktion (2.72) die Vollständigkeitsrelation (2.51) erfüllt. Führen wir in dem Integral

$$J = \frac{s}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-s^2 (x - t)^2} dx$$

die neue Veränderliche $z = s(x - t)$ mit $dz = s dx$ ein, so erhalten wir

$$J = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-z^2} dz .$$

Beachten wir nun die Relation

$$\boxed{\int_{-\infty}^{+\infty} e^{-z^2} dz = \sqrt{\pi}} , \tag{2.73}$$

auf die wir nicht weiter eingehen (vgl. hierzu [5]), so ergibt sich die Behauptung $J = 1$.

Wir untersuchen nun den Verlauf der Kurve der allgemeinen Dichtefunktion (2.72). Zunächst liegt die Kurve symmetrisch zur Geraden $x = t$, da $w(t + x) = w(t - x)$ gilt. Weiter besitzt die Kurve die einzigen Nullstellen bei $x = \pm\infty$. Für $x = t$ liegt ein Maximum mit der Ordinate $\frac{s}{\sqrt{\pi}}$ vor, da

$$w'(x) = \frac{-2 s^3 (x - t) e^{-s^2 (x-t)^2}}{\sqrt{\pi}} \tag{2.74}$$

außer an der Stelle $x = t$ nirgends im Endlichen verschwindet und

$$w''(x) = \frac{-2 s^3 [1 - 2 s^2 (x - t)^2] e^{-s^2 (x-t)^2}}{\sqrt{\pi}} \tag{2.75}$$

für $x = t$ negativ ist. Aus (2.74) folgt, daß die Kurve von $-\infty$ bis t monoton steigt, im Punkte $\left(t; \frac{s}{\sqrt{\pi}}\right)$ kulminiert und von t bis $+\infty$ monoton fällt. Die zwei notwendigen Wendepunkte liegen gemäß (2.75) bei $t \pm \frac{1}{s\sqrt{2}}$ mit den auf Grund der Symmetrie gleichen) Ordinaten $\frac{s}{\sqrt{e\pi}} \approx 0{,}6 \cdot \frac{s}{\sqrt{\pi}}$.

Wir zeichnen nachstehend drei Kurven, die zu den Funktionen (2.72) mit den Parameterwerten

1. $t = s = 1$
2. $t = s = \frac{1}{2}$
3. $t = 0,\ s = \frac{1}{\sqrt{2}}$ gehören.

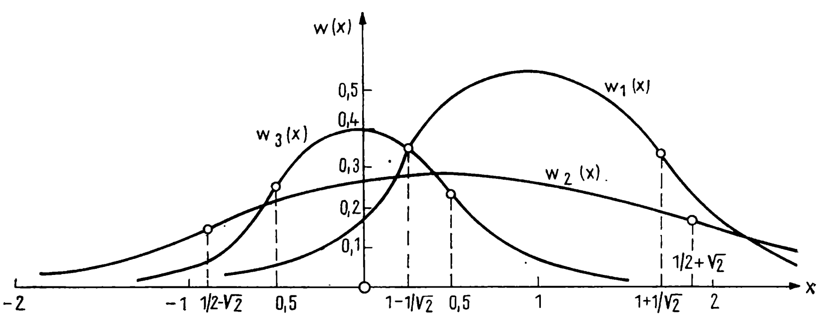

Abb. 15: Verlauf normaler Dichtefunktionen

Wir betrachten nun wieder eine allgemeine normal verteilte Zufallsgröße x, d. h. eine kontinuierliche Zufallsvariable x, die die Dichtefunktion (2.72) besitzt. Die Funktion

$$\boxed{W(\xi) = P(x < \xi)}, \tag{2.76}$$

gibt die Wahrscheinlichkeit dafür an, daß die Zufallsgröße x einen kleineren Wert als ξ annimmt. Wir nennen sie die *normale Verteilungsfunktion.* Auf Grund des Satzes 2.21 und in Verbindung mit (2.72) gestattet (2.76) die Darstellung

$$\boxed{W(\xi) = \frac{s}{\sqrt{\pi}} \int_{-\infty}^{\xi} e^{-s^2 (x-t)^2} dx} . \tag{2.77}$$

Die zugehörige *normale Verteilungskurve,* d. h. die Kurve der Funktion (2.77) zeigt den folgenden Verlauf

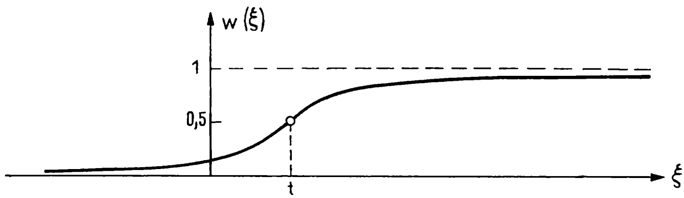

Abb. 16: Normale Verteilungskurve

In den folgenden Sätzen bestimmen wir den Mittelwert $E(x)$ und die Streuung $\sigma^2(x)$ der allgemeinen Normalverteilung (2.72).

Satz 2.25: Die Normalverteilung (2.72) hat den Erwartungswert

$$\boxed{E(x) = t} . \tag{2.78}$$

Beweis: Aus Symmetriegründen des Verlaufs der Dichtefunktionskurve ist der behauptete Sachverhalt klar. Wir weisen ihn nun rechnerisch nach: Zunächst ergibt sich aus (2.65_1) und (2.72)

$$E(x) = \frac{s}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x e^{-s^2 (x-t)^2} dx$$

oder

$$E(x) = \frac{1}{s\sqrt{\pi}} \int_{-\infty}^{+\infty} z e^{-z^2} dz + \frac{t}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-z^2} dz$$

Beachten wir die Tatsache, daß ze^{-z^2} eine ungerade Funktion ist, und berücksichtigen wir die Relation (2.73), so bekommen wir kurzerhand $E(x) = t$. Wir nennen das Integral

$$\boxed{M_n = \frac{s}{\sqrt{\pi}} \int_{-\infty}^{+\infty} [x - E(x)]^n e^{-s^2(x-t)^2} dx \; (n = 0, 1, \ldots)} \tag{2.79}$$

das *n-te Moment* oder das *Moment n-ter Ordnung* der Normalverteilung (2.72) und beweisen den

Satz 2.26: Für die Momente M_n $(n = 0, 1, 2, \ldots)$ der Normalverteilung (2.72) besteht die Rekursionsformel

$$\boxed{M_{n+2} = \frac{n+1}{2s^2} M_n} \,. \tag{2.80}$$

Insbesondere sind alle Momente ungerader Ordnung null.

Beweis: Wenden wir auf die Definitionsgleichung (2.79) für M_n die Regel der partiellen Integration an, so erhalten wir

$$M_n = \frac{s}{\sqrt{\pi}} \frac{(x-t)^{n+1}}{n+1} e^{-s^2(x-t)^2} \Bigg|_{-\infty}^{+\infty} + \frac{2s^3}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{(x-t)^{n+2}}{n+1} \cdot e^{-s^2(x-t)^2} \, dx$$

oder vereinfacht (das erste Glied fällt weg!)

$$M_n = \frac{2s^2}{n+1} \cdot \frac{s}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (x-t)^{n+2} \, e^{-s^2(x-t)^2} \, dx \,.$$

Beachten wir (2.79) für den Index $n + 2$, so wird klar, daß der vorstehende Ausdruck die Beziehung (2.80) darstellt; dabei ist n eine beliebige ganze nichtnegative Zahl. Um den zweiten Teil des Satzes zu beweisen, brauchen wir nur $M_1 = 0$ zu zeigen; denn ist $M_1 = 0$, so folgt aus (2.80) sofort $M_3 = M_5 = \ldots = 0$. In der Tat ist wegen $E(x) = t$

$$M_1 = \frac{s}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x e^{-s^2(x-t)^2} \, dx - \frac{s \cdot t}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-s^2(x-t)^2} \, dx = 0 \,.$$

Das erste Integral hat nach Satz 2.25 den Wert t, das zweite Integral ist bis auf den Faktor t gerade die Vollständigkeitsrelation (2.51) in Verbindung mit (2.72).

Aus dem letzten Satz folgt der

Satz 2.27: Die Normalverteilung (2.72) hat die Streuung

$$\boxed{\sigma^2(x) = \frac{1}{2s^2}} \,. \tag{2.81}$$

Beweis: Das nullte Moment (2.79) ist gerade die Vollständigkeitsrelation, folglich gilt $M_0 = 1$. Weiter stellt das zweite Moment (2.79) die gesuchte Streuung $\sigma^2(x)$ dar; mithin folgt $M_2 = \sigma^2(x)$. Setzen wir nun in (2.80) $n = 0$, so ergibt sich die Behauptung

$$M_2 = \sigma^2(x) = \frac{1}{2s^2} \cdot M_0 = \frac{1}{2s^2} \,.$$

Aus der Beziehung (2.81) folgt unmittelbar, daß die Streuung $\sigma^2(x)$ der Normalverteilung mit wachsendem Parameter s abnimmt. Je größer also der Parameter s ist, um so kleiner ist die Streuung $\sigma^2(x)$, d.h. um so „besser“

ist die Normalverteilung. Aus diesem Grunde trägt die positive Konstante s die Bezeichnung *Präzisionsmaß* der Normalverteilung.

Satz 2.28: Hat eine normalverteilte Zufallsvariable x den Erwartungswert $E(x)$ und das Streuungsmaß $\sigma(x)$, dann lautet die zugehörige Verteilungsdichte

$$\boxed{w(x) = \frac{1}{\sigma(x)\sqrt{2\pi}}\, e^{-\frac{[x-E(x)]^2}{2\sigma^2(x)}}}\,. \tag{2.82}$$

Beweis: Ersetzen wir in (2.72) den Parameter t vermöge (2.78) durch $E(x)$ und die Größe s durch den sich aus (2.81) für s ergebenden Wert $\frac{1}{\sigma(x)\sqrt{2}}$, so resultiert gerade die behauptete Darstellung (2.82). Die Darstellung (2.82) macht insbesondere deutlich, daß jede Normalverteilung durch ihren Erwartungswert $E(x)$ und ihre Streuung $\sigma^2(x)$ eindeutig bestimmt ist. Besitzt beispielsweise die normalverteilte Zufallsgröße x den Mittelwert $E(x) = 5$ und das Streuungsmaß $\sigma(x) = 2$, so folgt aus (2.82), daß die zugehörige Dichtefunktion die Gestalt

$$w(x) = \frac{1}{2\sqrt{2\pi}}\, e^{-\frac{(x-5)^2}{8}}$$

aufweist.

Beachten wir die Beziehung (2.82), so folgt für die Verteilungsfunktion einer Normalverteilung die Darstellung

$$\boxed{W(\xi) = \frac{1}{\sigma(x)\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-\frac{[x-E(x)]^2}{2\sigma^2(x)}}\, dx}\,. \tag{2.83}$$

Im eben betrachteten Beispiel lautet mithin die Verteilungsfunktion

$$W(\xi) = \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-\frac{(x-5)^2}{8}}\, dx\,.$$

2.6.2. *Die standardisierte Normalverteilung*

Die *standardisierte Normalverteilung* ist diejenige Normalverteilung, die den Erwartungswert $E(x) = 0$ und das Streuungsmaß $\sigma(x) = 1$ besitzt. Gemäß (2.82) hat die standardisiert normalverteilte Zufallsveränderliche x die Dichtefunktion

$$\boxed{\varphi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}}\,. \tag{2.84}$$

Die Funktion $\varphi(x)$ geht also aus der allgemeinen normalen Wahrscheinlichkeitsdichte (2.72) hervor, wenn $t = o$ und $s = \frac{1}{\sqrt{2}}$ gesetzt wird. Da $\varphi(x)$ eine gerade Funktion in x ist, d. h. $\varphi(-x) = \varphi(x)$ für alle reellen x gilt, liegt das Bild der Dichte (2.84) symmetrisch zur y-Achse. Weiter folgt aus der Diskussion des Kurvenverlaufs der allgemeinen Dichtefunktion (2.72), daß $\varphi(x)$ bei $x = 0$ das Maximum $\varphi(o) = \frac{1}{\sqrt{2\pi}} \approx 0{,}4$ (das einzige Extremum überhaupt) und in $\left(\pm 1, \frac{1}{\sqrt{2e\pi}} \approx 0{,}24\right)$ die einzigen Wendepunkte aufweist. Das Bild der Kurve der standardisierten Normalverteilungsdichte (2.84) ist bereits in der vorletzten Skizze gezeichnet worden; es ist dort mit der Bezeichnung $w_3(x)$ versehen.

Berücksichtigen wir in (2.83) die Parameterwerte der standardisierten Normalverteilung, so nimmt die betreffende Verteilungsfunktion die folgende Gestalt an:

$$\boxed{\Phi(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\xi} e^{-\frac{x^2}{2}}\, dx} \; ; \qquad (2.85)$$

Das zugehörige Kurvenbild hat das Aussehen

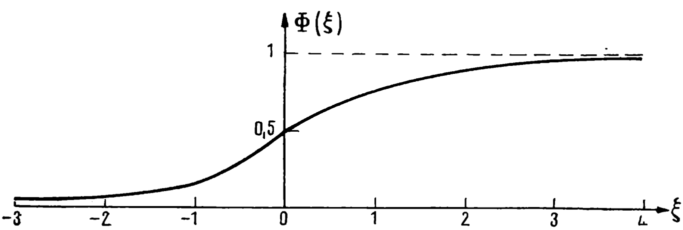

Abb. 17: Die standardisierte normale Verteilungskurve

Ist also x die in standardisiertem Sinne normalverteilte Zufallsgröße, so gilt in Verbindung mit (2.85)

$$\boxed{P(a \leqq x < b) = \Phi(b) - \Phi(a)} \, . \qquad (2.86)$$

Wir beweisen nunmehr den

Satz 2.29: Ist x eine beliebige normalverteilte Zufallsvariable mit dem Erwartungswert $E(x)$ der Dichte $w(x)$ und dem Streuungsmaß $\sigma(x)$, dann ist

$$\boxed{y = \frac{x - E(x)}{\sigma(x)}} \qquad (2.87)$$

die standardisiert normalverteilte Zufallsgröße y mit der Dichte

$$\boxed{\varphi(y) = \sigma(x) \cdot w(x) = \sigma(x)\, w\,[\sigma(x)\, y + E(x)]} \tag{2.88}$$

und der Verteilungsfunktion

$$\boxed{\Phi(\eta) = W(\xi) = W\big(\sigma(x)\eta + E(x)\big)} \tag{2.89}$$

Beweis: Wegen (2.87) hat y zunächst den Erwartungswert null:

$$E(y) = \frac{1}{\sigma(x)}\, E[x - E(x)] = \frac{1}{\sigma(x)}\, [E(x) - E(x)] = 0$$

und das Streuungsmaß eins:

$$\sigma(y) = \frac{1}{\sigma(x)}\, \sigma[x - E(x)] = \frac{1}{\sigma(x)} \cdot \sigma(x) = 1.$$

Weiter folgt aus (2.84), (2.87) und (2.82)

$$\varphi(y) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{y^2}{2}} = \sigma(x)\, \frac{1}{\sigma(x)\sqrt{2\pi}}\, e^{-\frac{[x - E(x)]^2}{2\sigma^2(x)}} = \sigma(x)\, w(x)$$

oder die Behauptung (2.88). Wenden wir schließlich die Substitution (2.87) auf das Integral in

$$\Phi(\eta) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\eta} e^{-\frac{y^2}{2}}\, dy$$

an, so wird wegen $dy = \dfrac{dx}{\sigma(x)}$

$$\Phi(\eta) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\eta} e^{-\frac{y^2}{2}}\, dy = \frac{1}{\sigma(x)\sqrt{2\pi}} \int\limits_{-\infty}^{\xi} e^{-\frac{[x - E(x)]^2}{2\sigma^2(x)}}\, dx = W(\xi)$$

mit $\xi = \eta\sigma(x) + E(x)$; das bedeutet aber schon die Behauptung (2.89).

Gemäß dem vorstehenden Satz läßt sich also jede normalverteilte Zufallsgröße durch die lineare Transformation (2.87) auf die standardisiert normalverteilte Zufallsvariable zurückführen. Die Umkehrung dieses Sachverhalts — die wegen der Linearität der Transformation (2.87) möglich ist — bringt der folgende Satz zum Ausdruck.

Satz 2.30: Ist y die Zufallsvariable der standardisierten Normalverteilung und sind die Parameter $\bar{x}$ und $r > 0$ gegeben, dann repräsentiert die Zufallsgröße

$$\boxed{x = ry + \bar{x}} \tag{2.90}$$

eine Normalverteilung mit dem Erwartungswert $\bar{x}$, dem Streuungsmaß r, der Dichtefunktion

$$\boxed{w(x) = \frac{1}{r}\,\varphi(y) = \frac{1}{r}\,\varphi\left(\frac{x-\bar{x}}{r}\right)} \tag{2.91}$$

und der Verteilungsfunktion

$$\boxed{W(\xi) = \Phi(\eta) = \Phi\left(\frac{\xi-\bar{x}}{r}\right)} \tag{2.92}$$

Beweis: Wegen (2.90) hat x zunächst den Erwartungswert $\bar{x}$:

$$E(x) = E(ry + \bar{x}) = rE(y) + E(\bar{x}) = \bar{x}$$

und das Streuungsmaß r:

$$\sigma(x) = \sigma(ry + \bar{x}) = r \cdot \sigma(y) = r.$$

Weiter folgt aus (2.82), (2.90) und (2.84)

$$w(x) = \frac{1}{r\sqrt{2\pi}}\,e^{-\frac{(x-\bar{x})^2}{2r^2}} = \frac{1}{r}\cdot\frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2}{2}} = \frac{1}{r}\,\varphi(y)$$

oder die Behauptung (2.91). Wenden wir schließlich die Substitution (2.90) auf das Integral in (2.83) an, so ergibt sich wegen $dx = r\,dy$

$$W(\xi) = \frac{1}{r\sqrt{2\pi}}\int_{-\infty}^{\xi} e^{-\frac{(x-\bar{x})^2}{2r^2}}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\eta} e^{-\frac{y^2}{2}}\,dy = \Phi(\eta)$$

mit $\eta = \frac{1}{r}(\xi - \bar{x})$; das ist aber gerade die Aussage (2.92). Aus den Sätzen 2.29 und 2.30 geht hervor, daß jede Normalverteilung in die standardisierte Normalverteilung und umgekehrt die standardisierte Normalverteilung in jede Normalverteilung transformiert werden kann. Aus diesem Sachverhalt ergibt sich unmittelbar, daß jede Normalverteilung in eine beliebige andere Normalverteilung übergeführt werden kann. Selbstverständlich kann diese Überführung mittelbar über die standardisierte Normalverteilung vorgenommen werden. Im folgenden Satz geben wir die diesbezüglichen direkten Transformationsformeln an.

Satz 2.31: Sind x_i $(i = 1, 2)$ normalverteilte Zufallsvariable mit den Erwartungswerten $\bar{x}_i$, den Streuungsmaßen σ_i, den Dichtefunktionen $w_i(x_i)$ und den Verteilungsfunktionen $W_i(\xi_i)$, dann gelten die Zusammenhänge

$$\boxed{\sigma_1 \cdot (x_2 - \bar{x}_2) = \sigma_2 \cdot (x_1 - \bar{x}_1)} \tag{2.93}$$

und

$$\boxed{\sigma_1\,w_1(x_1) = \sigma_2\,w_2(x_2)} \tag{2.94}$$

sowie

$$\boxed{W_1(\xi_1) = W_2(\xi_2)} \tag{2.95}$$

mit

$$\boxed{\sigma_1 \cdot (\xi_2 - \bar{x}_2) = \sigma_2 \cdot (\xi_1 - \bar{x}_1)}\,. \tag{2.95'}$$

Beweis: Bezeichnet y die standardisiert normalverteilte Zufallsgröße mit der Dichte $\varphi(y)$ und der Verteilungsfunktion $\Phi(\eta)$, dann resultieren aus dem Satz 2.29 die Beziehungen

$$\left.\begin{aligned} &y = \frac{x_i - \bar{x}_i}{\sigma_i} \\ &\text{und} \\ &\varphi(y) = \sigma_i\, w_i(\sigma_i\, y + \bar{x}_i) = \sigma_i\, w_i(x_i) \\ &\text{sowie} \\ &\Phi(\eta) = W_i(\sigma_i\, \eta + \bar{x}_i) = W_i(\xi)\,. \end{aligned}\right\} \quad (i = 1, 2) \qquad (*)$$

Eliminieren wir aus diesen 3 Gleichungspaaren die Größen y, $\varphi(y)$ und $\Phi(\eta)$, dann erhalten wir die Behauptung des Satzes 2.31.

Betrachten wir an dieser Stelle ein Beispiel: Gegeben sind die normalverteilten Zufallsgrößen x_1 und x_2 mit dem Erwartungswert $\bar{x}_1 = 3$ bzw. $\bar{x}_2 = 5$ und dem Streuungsmaß $\sigma_1 = 10$ bzw. $\sigma_2 = 20$. Wie hängen die Zufallsvariablen x_i, ihre Dichtefunktionen $w_i(x_i)$ und ihre Verteilungsfunktionen $W_i(\xi_i)$ miteinander zusammen? Aus (2.93) folgt zunächst $10(x_2 - 5) = 20(x_1 - 3)$ oder kurz

$$x_2 = 2\,x_1 - 1\,.$$

Wegen (2.94) wird weiter

$$w_1(x_1) = 2\,w_2(x_2)\,.$$

Schließlich ergibt sich aus (2.95) nebst (2.95')

$$W_1(\xi_1) = W_2(\xi_2)$$

mit $\xi_2 = 2\xi_1 - 1\,.$

Kehren wir nun zu den allgemeinen Überlegungen zurück. Wir haben das Ergebnis abgeleitet, daß sich jede Normalverteilung auf die standardisierte Normalverteilung zurückführen läßt. Diese weist unter allen Normalverteilungen die einfachste Struktur auf. Aus diesem Grunde läßt sich jede Fragestellung zu einer beliebigen Normalverteilung in eine Frage zur standardisierten Normalverteilung umwandeln. Aus den allgemeinen Bemerkungen über kontinuierliche Verteilungen geht hervor, daß bei der Beantwortung von Fragen und Erörterung von Problemen stetiger Zufallsgrößen laufend mit der Dichtefunktion und der Verteilungsfunktion gearbeitet wird. Folglich ist unschwer einzusehen, daß die Funktionen $\varphi(x)$ und $\Phi(\xi)$ eine fundamentale Rolle in der Theorie der Normalverteilungen spielen. Um mit diesen Funktionen, die in (2.84) und (2.85) ausgeschrieben sind, praktisch umgehen zu können, sind $\varphi(x)$ und $\Phi(\xi)$ tabelliert worden; in den meisten Büchern über Wahrscheinlichkeitsrechnung und Statistik sind solche Tabellen im Anhang zu finden.
Mitunter wird anstelle von $\Phi(\xi)$ die Funktion

$$\boxed{\Theta(\xi) = \int_{-\xi}^{+\xi} \varphi(x)\, d\,x = \Phi(\xi) - \Phi(-\xi)} \qquad (2.96)$$

benutzt, die wegen (2.85) die Wahrscheinlichkeit dafür liefert, daß die standardisiert normalverteilte Zufallsgröße x dem Intervall $(-\xi, +\xi)$ angehört, d. h. der Bedingung $|x| < \xi$ genügt

$$\boxed{\Theta(\xi) = P(|x| < \xi)}\,. \tag{2.97}$$

Wir beweisen nun den

Satz 2.32: Zwischen $\Phi(\xi)$ und $\Theta(\xi)$ besteht der Zusammenhang

$$\boxed{\Theta(\xi) = 2\,\Phi(\xi) - 1}\,. \tag{2.98}$$

Beweis: Da $\varphi(x)$ eine gerade Funktion in x ist, gilt zunächst

$$\Theta(\xi) = 2\int_0^\xi \varphi(x)\,dx\,.$$

Formen wir das letzte Integral um und beachten wir (2.85), so wird

$$\Theta(\xi) = 2\left[\int_{-\infty}^{\xi} \varphi(x)\,dx - \int_{-\infty}^{0} \varphi(x)\,dx\right] = 2\,(\Phi(\xi) - \Phi(0))\,.$$

Auf Grund der Symmetrie der Kurve von $\varphi(x)$ bezüglich der y-Achse folgt $\Phi(o) = P(-\infty < x < o) = \frac{1}{2}$ und damit die Behauptung (2.98).

Vielfach ist die Funktion $\Theta(\xi)$ sogar tabelliert. Durch Anwendung der aus (2.98) resultierenden Beziehung

$$\boxed{\Phi(\xi) = \frac{1}{2}\left[\Theta(\xi) + 1\right]} \tag{2.99}$$

ergibt sich aus einer Θ-Tabelle ohne große Rechnung der entsprechende Wert für Φ.

Wir geben im Anhang eine Tabelle für die Verteilungsfunktion $\Phi(\xi)$ der standardisierten Normalverteilung. Dabei beschränken wir uns auf nichtnegative Argumente ξ. Um der erwähnten Tabelle auch Φ-Werte negativer Argumente entnehmen zu können, brauchen wir nur zu beachten

Satz 2.33: Die Funktion $\Phi(\xi)$ besitzt die Eigenschaft

$$\boxed{\Phi(-\xi) = 1 - \Phi(\xi)}\,. \tag{2.100}$$

Beweis: Aus (2.85) folgt

$$\Phi(-\xi) = \int_{-\infty}^{-\xi} \varphi(x)\,dx\,.$$

Das vorstehende Integral gestattet nun die Umformung

$$\int_{-\infty}^{-\xi} \varphi(x)\, d\,x = \int_{-\infty}^{+\infty} \varphi(x)\, d\,x - \int_{-\infty}^{\xi} \varphi(x)\, d\,x = 1 - \Phi(\xi)\,,$$

wenn wir wieder die Symmetrie der Dichtefunktion $\varphi(x)$ berücksichtigen. Damit ist schon alles gezeigt.

Wir betrachten nun einige Beispiele, bei deren Behandlung wir uns der im Anhang befindlichen Φ-Tabelle bedienen.

1. Wie groß ist die Wahrscheinlichkeit dafür, daß die standardisiert normalverteilte Zufallsvariable x im Intervall $(-1, +1)$ auftritt? Diese Frage ist offenbar der folgenden gleichwertig: Wie groß ist der Inhalt des Flächenstückes zwischen der x-Achse, der Kurve $y = \varphi(x)$ und deren Wendepunktsordinaten? Der Flächeninhalt sei F_1.

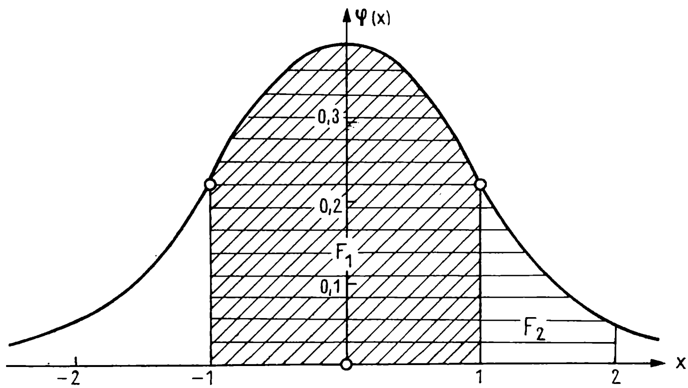

Abb. 18: Die standardisierte normale Dichtefunktion

Wir erhalten gemäß (2.96) und (2.97)

$$F_1 = P(-1 < x < 1) = P(|x| < 1) = \Theta(1)\,.$$

Durch Anwendung der Beziehung (2.98) wird dann weiter

$$F_1 = \Theta(1) = 2\Phi(1) - 1\,.$$

Der Φ-Tabelle entnehmen wir

$$\Phi(1) = 0{,}8413\,.$$

Mithin beträgt die gesuchte Wahrscheinlichkeit (oder der gesuchte Flächeninhalt)

$$F_1 = 2 \cdot 0{,}8413 - 1 = 0{,}6826\,.$$

2. Mit welcher Wahrscheinlichkeit tritt die in Rede stehende Zufallsveränderliche x im Intervall $(-1, 2)$ auf? Diese gesuchte Wahrscheinlichkeit ist offenbar gleich dem Inhalt des Flächenstückes zwischen der x-Achse, der Kurve $y = \varphi(x)$ und den Geraden $x = -1$ und $x = 2$. Dieser Inhalt sei F_2. Aus (2.64) und (2.86) folgt

$$F_2 = P(-1 < x < 2) = \Phi(2) - \Phi(-1)$$

oder auf Grund (2.100)

$$F_2 = \Phi(2) + \Phi(1) - 1\,.$$

Unter Zuhilfenahme der Φ-Tabelle ergibt die gesuchte Wahrscheinlichkeit (oder der gesuchte Flächeninhalt) den Wert

$$F_2 = 0{,}9772 + 0{,}8413 - 1 = 0{,}8185.$$

3. In welchem symmetrisch um den Ursprung liegenden Intervall ist die gegebene Zufallsgröße x mit einer Wahrscheinlichkeit von 50 Prozent anzutreffen? Bezeichnen wir das gesuchte Intervall mit $(-l, +l)$, so muß zunächst wegen (2.96)

$$\Theta(l) = P(-l < x < +l) = \frac{1}{2}$$

sein. Beachten wir (2.99), so wird aus der vorstehenden Gleichung

$$\Phi(l) = \frac{1}{2}\left(\frac{1}{2} + 1\right) = \frac{3}{4}\,.$$

Gemäß der Φ-Tabelle gehört zu dem Φ-Wert 0,75 das Argument $\xi = 0{,}674$ (linear interpoliert!), so daß

$$l = 0{,}674$$

das gesuchte Resultat ist. Berücksichtigen wir noch die Symmetrie der Dichtefunktion $\varphi(x)$, so bedeutet dieses Ergebnis, daß die mit A, B, C und D bezeichneten Flächen der folgenden Skizze gleichen Inhalt besitzen.

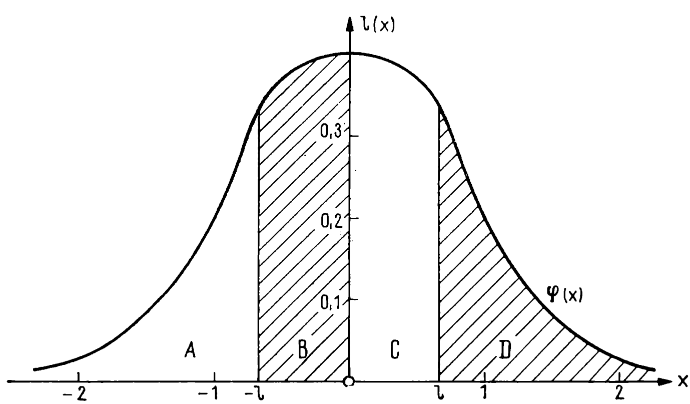

Abb. 19: Die standardisierte normale Dichtefunktion

4. Wie groß muß die Konstante α_0 sein, damit die betrachtete Zufallsvariable x mit einer Wahrscheinlichkeit von 60 Prozent rechts von α_0 vorkommt? Es soll also

$$P(\alpha_0 < x < \infty) = \Phi(\infty) - \Phi(\alpha_0) = 0{,}6$$

oder infolge Satz 2.22

$$\Phi(\alpha_0) = 0{,}4$$

gelten. Aus der Φ-Tafel geht hervor, daß α_0 negativ sein muß. Daher verwenden wir (2.100) in der Form

$$\Phi(\alpha_0) = 1 - \Phi(-\alpha_0)$$

und bekommen

$$\Phi(-\alpha_0) = 0{,}6\,.$$

Interpolieren wir wiederum in der Φ-Tabelle, so wird $-\alpha_0 = 0{,}253$ und damit das gewünschte Resultat

$$\alpha_0 = -0{,}253\,.$$

Wir wenden uns nun wieder den allgemeinen Betrachtungen zu. Zuerst beweisen wir den wichtigen

Satz 2.34: Hat die normalverteilte Zufallsgröße x den Erwartungswert $E(x)$ und die Streuung $\sigma^2(x)$, dann besteht der Zusammenhang

$$\boxed{P(\alpha \leqq x < \beta) = \Phi\left(\frac{\beta - E(x)}{\sigma(x)}\right) - \Phi\left(\frac{\alpha - E(x)}{\sigma(x)}\right)}\,. \tag{2.101}$$

Beweis: Ist $W(\xi)$ die Verteilungsfunktion der Zufallsvariablen x, dann ist auf Grund (2.63)

$$P(\alpha \leqq x < \beta) = W(\beta) - W(\alpha)\,.$$

Wenden wir nun auf diese letzte Gleichung den Satz 2.29, insbesondere die Beziehung (2.89) an, dann folgt gerade die Behauptung (2.101).

Da jede Normalverteilungsdichte stetig ist, sind neben (2.101) auch noch die Beziehungen

$$\boxed{\begin{aligned} P(\alpha \leqq x \leqq \beta) &= P(\alpha < x \leqq \beta) = P(\alpha < x < \beta) \\ &= \Phi\left(\frac{\beta - E(x)}{\sigma(x)}\right) - \Phi\left(\frac{\alpha - E(x)}{\sigma(x)}\right) \end{aligned}} \tag{2.102}$$

gültig.

Weiter besteht der

Satz 2.35: Hat die normalverteilte Zufallsgröße x den Erwartungswert $E(x)$ und das Streuungsmaß $\sigma(x)$, dann gilt die Formel

$$\boxed{P[x - E(x) < \varrho] = \Theta\left(\frac{\varrho}{\sigma(x)}\right)}\,. \tag{2.103}$$

Beweis: Wir betrachten die Zufallsveränderliche

$$y = \frac{x - E(x)}{\sigma(x)},$$

die nach Satz 2.29 der standardisierten Normalverteilung gehorcht. Dann wird wegen $\sigma(x) > 0$

$$|x - E(x)| = |\sigma(x) \cdot y| = \sigma(x) \cdot |y|$$

und folglich für die linke Seite von (2.103)

$$P[|x - E(x)| < \varrho] = P[\sigma(x)|y| < \varrho]$$

oder wiederum wegen $\sigma(x) > 0$

$$P[|x - E(x)| < \varrho] = P\left[|y| < \frac{\varrho}{\sigma(x)}\right].$$

Gemäß Formel (2.97), in der die standardisiert normalverteilte Zufallsveränderliche mit x bezeichnet ist, ergibt sich dann die Behauptung:

$$P\left[|y| < \frac{\varrho}{\sigma(x)}\right] = \Theta\left[\frac{\varrho}{\sigma(x)}\right].$$

Die abgeleitete Formel (2.103) liefert also die Wahrscheinlichkeit dafür, daß die betreffende Zufallsvariable x in einem symmetrisch um ihren Erwartungswert $E(x)$ liegenden Bereich vorkommt. Diese Wahrscheinlichkeit hängt offenbar nicht von der Größe des Erwartungswertes $E(x)$ ab, sondern wird lediglich durch die Länge dieses Intervalls und das Streuungsmaß $\sigma(x)$ bestimmt. Diesen Sachverhalt bringt auch die Beziehung (2.103) zum Ausdruck.

Wir geben zur Illustration der Sätze 2.34 und 2.35 nunmehr einige Beispiele. Dabei interpolieren wir linear in der Φ-Tafel.

1. Gegeben sei eine normalverteilte Zufallsgröße x mit dem Mittelwert $E(x) = 1$ und dem Streuungsmaß $\sigma(x) = 3$. Gesucht werden die Wahrscheinlichkeiten

   a) p_a dafür, daß x dem Intervall $-1 \leqq x \leqq 3$ angehört;

   b) p_b dafür, daß x im Intervall $-2 \leqq x \leqq 1$ auftritt;

   c) p_c dafür, daß x die Zahl $x_0 = 4$ übertrifft und

   d) p_d dafür, daß x zwischen $x_1 = 5$ und $x_2 = 6$ vorkommt.

   Wir ermitteln die gesuchten Wahrscheinlichkeiten der Reihe nach. Für p_a ergibt sich gemäß (2.102)

$$p_a = \Phi\left(\frac{3-1}{3}\right) - \Phi\left(\frac{-1-1}{3}\right) = \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{2}{3}\right)$$

   oder als Folge von (2.96) und (2.98)

$$p_a = \Phi\left(\frac{2}{3}\right) = 2\,\Phi\left(\frac{2}{3}\right) - 1.$$

Unter Zuhilfenahme der Φ-Tafel wird

$$p_a = 2 \cdot 0{,}7475 - 1 = 0{,}4950 \,.$$

Für p_b erhalten wir zunächst

$$p_b = \Phi\left(\frac{1-1}{3}\right) - \Phi\left(\frac{-2-1}{3}\right) = \Phi\,(0) - \Phi\,(-1)$$

oder wegen (2.100)

$$p_b = \Phi\,(o) + \Phi\,(1) - 1 \,.$$

Schlagen wir die Φ-Werte in der Tafel nach, so finden wir

$$p_b = 0{,}5 + 0{,}8413 - 1 = 0{,}3413 \,.$$

Bei der Ermittlung von p_c gehen wir von

$$\begin{aligned} p_c &= P\,(4 < x < \infty) = \Phi\left(\frac{\infty - 1}{3}\right) - \Phi\left(\frac{4-1}{3}\right) \\ &= \Phi\,(\infty) - \Phi\,(1) \,. \end{aligned}$$

aus; Φ (1) haben wir oben schon aufgesucht. Beachten wir außerdem noch $\Phi\,(\infty) = 1$, dann lautet das Ergebnis

$$p_c = 1 - 0{,}8413 = 0{,}1587 \,.$$

Schließlich berechnet sich p_d zu

$$p_d = \Phi\left(\frac{6-1}{3}\right) - \Phi\left(\frac{5-1}{3}\right) = \Phi\left(\frac{5}{3}\right) - \Phi\left(\frac{4}{3}\right)$$

oder anhand der Φ-Tabelle

$$p_d = 0{,}9522 - 0{,}9088 = 0{,}0434 \,.$$

2. Welche Normalverteilungen haben die Eigenschaft, daß ihre Zufallsvariablen x mit einer Wahrscheinlichkeit von 95 Prozent um weniger als eine Einheit von den betreffenden Erwartungswerten $E\,(x)$ abweichen? Für die gesuchten Verteilungen soll mithin

$$P\,[\,|x - E\,(x)| < 1] = 0{,}95$$

oder gemäß Satz 2.35

$$\Phi\left[\frac{1}{\sigma\,(x)}\right] = 0{,}95$$

gelten. Bringen wir die Formel (2.99) in Anwendung, so wird

$$\Phi\left[\frac{1}{\sigma\,(x)}\right] = \frac{0{,}95 + 1}{2} = 0{,}975 \,.$$

In der Φ-Tafel steht: Φ (1.96) = 0,975. Daher folgt

$$\frac{1}{\sigma\,(x)} = 1{,}96$$

oder kurzerhand

$$\sigma\,(x) \approx 0{,}51 \,.$$

Die Antwort auf die gestellte Frage lautet damit: Alle Normalverteilungen mit dem Streuungsmaß $\sigma(x) \approx 0{,}51$ haben die Besonderheit, daß ihre Zufallsgrößen mit einer Wahrscheinlichkeit von 95 Prozent um weniger als eine Einheit von den betreffenden Erwartungswerten $E(x)$ abweichen.

3. Gegeben ist eine normalverteilte Zufallsveränderliche mit dem Mittelwert $E(x) = -2$ und der Streuung $\sigma^2(x) = 100$. Gesucht sind die Wahrscheinlichkeit p_0 dafür, daß x dem Intervall (0,5) angehört, und ein solches symmetrisch um $x_0 = -2$ liegendes Intervall $(-2 - l, -2 + l)$, in dem die Zufallsvariable x mit einer Wahrscheinlichkeit von 99,9 Prozent anzutreffen ist.

   Für p_0 erhalten wir
$$p_0 = \Phi\left(\frac{5+2}{10}\right) - \Phi\left(\frac{0+2}{10}\right) = \Phi(0{,}7) - \Phi(0{,}2)$$
   oder bei Zuhilfenahme der Φ-Tafel
$$p_0 = 0{,}7580 - 0{,}5793 = 0{,}1787.$$
   Bei der Bestimmung der Größe 1 beachten wir den Satz 2.35, demzufolge
$$P[|x - E(x)| < l] = \Theta\left(\frac{l}{\sigma(x)}\right) = \Theta\left(\frac{l}{10}\right) = 0{,}999$$
   besteht. Die Beziehung (2.99) führt uns weiter zu
$$\Phi\left(\frac{l}{10}\right) = \frac{1}{2}(0{,}999 + 1) = 0{,}9995.$$
   In der Φ-Tafel gehört zu $\Phi = 0{,}9995$ das Argument $\xi = 3{,}30$. Mithin gilt
$$\frac{l}{10} = 3{,}30$$
   oder das Resultat
$$l = 33\,.$$
   In dem Intervall $(-35, 31)$ tritt also die in Rede stehende Zufallsgröße x mit einer Wahrscheinlichkeit von 0,999 auf. Außerhalb dieses Intervalls ist die Zufallsgröße daher nur mit einer Wahrscheinlichkeit von 1 Promille anzutreffen.

4. Gegeben ist wiederum eine normalverteilte zufällige Größe x; in diesem Fall seien $E(x) = 3$ und $\sigma(x) = 4$ die zugehörigen Parameter. Wie groß muß nun die Zahl z sein, damit die Variable x in dem Intervall $(z, 4)$ wenigstens mit einer Wahrscheinlichkeit von 25 Prozent auftritt?
   Wir gehen bei der Ermittlung von z von der gestellten Bedingung
$$P(z < x < 4) = \Phi\left(\frac{4-3}{4}\right) - \Phi\left(\frac{z-3}{4}\right) =$$
$$= \Phi\left(\frac{1}{4}\right) - \Phi\left(\frac{z-3}{4}\right) \geqq 0{,}25$$
   aus. Wegen $\Phi(0{,}25) = 0{,}5987$ ergibt sich weiter
$$\Phi\left(\frac{z-3}{4}\right) \leqq 0{,}5987 - 0{,}25 = 0{,}3487.$$

Um auf einen Φ-Wert $\geqq 0{,}5$ zu gelangen und damit die Φ-Tafel im Anhang anwenden zu können, beachten wir die Relation (2.100), die

$$\Phi\left(\frac{3-z}{4}\right) \geqq 1 - \Phi\left(\frac{z-3}{4}\right) \geqq 0{,}6513$$

zur Folge hat. Auf Grund der Monotonie der Φ-Funktion und der Tatsache $\Phi\,(0{,}389) = 0{,}6513$ erhalten wir

$$\frac{3-z}{4} \leqq 0{,}389$$

oder schließlich das Ergebnis

$$z \leqq 1{,}444\,.$$

Die gesuchte Größe z darf also nicht größer als die Zahl 1,444 sein.

5. Wie groß muß die Zahl L sein, damit die in der vorstehenden Aufgabe behandelte Zufallsgröße x in dem Intervall $(3-L,\ 3+L)$ mit einer geringeren Wahrscheinlichkeit als 15 Prozent vorkommt? Es soll also

$$P\,[\,|x-E(x)|<L] = \Theta\left(\frac{L}{4}\right) < 0{,}15$$

oder in Verbindung mit (2.99)

$$\Phi\left(\frac{L}{4}\right) < \frac{1}{2}\,(0{,}15+1) = 0{,}575$$

sein. Aus $\Phi\,(0{,}189) = 0{,}575$ und der Monotonie von $\Phi\,(\xi)$ resultiert

$$\frac{L}{4} < 0{,}189$$

oder

$$L < 0{,}756\,.$$

Die gesuchte Zahl L muß also kleiner als 0,756 sein, damit der geforderte Sachverhalt gültig ist.

2.6.3. *Grenzwertsätze*

Die sogenannten *Grenzwertsätze* beschäftigen sich mit gewissen Grenzübergängen von Verteilungen, die zu Normalverteilungen führen. Wir beschreiben im folgenden zwei solcher Grenzwertsätze, die die praktische Bedeutung der Normalverteilung unterstreichen. Bezüglich der entsprechenden Beweise verweisen wir auf die einschlägige Literatur, z. B. [4]. Zunächst erwähnen wir den *Zentralen Grenzwertsatz*, den

Satz 2.36: Eine Zufallsgröße x ist normalverteilt, wenn sie die Summe einer sehr großen Zahl von Zufallsvariablen $x_1, x_2, \ldots, x_n$ ist, die erstens voneinander unabhängig sind und zweitens im Verhältnis zur Summe unbedeutend sind, d. h. nur einen geringfügigen Beitrag zur Summe liefern.

Die praktische Bedeutung des vorstehenden Satzes beruht darauf, daß die Verteilungen der Zufallsgrößen x_i ($i = 1, 2, \ldots, n$) ganz beliebig sein können. Die einzelnen Verteilungen brauchen auch gar nicht explizit bekannt zu sein. Entscheidend ist eben nur die summarische Wirkung der Zufallsgrößen.

Wir behandeln nun zwei Beispiele, die den Sachverhalt des Satzes 2.36 noch verdeutlichen:

1. Wir betrachten die Messung einer physikalischen Größe mittels eines Meßinstrumentes. Eine beliebige Messung liefert lediglich einen Näherungswert für die zu messende Größe, da eine Vielzahl zufälliger Faktoren (z. B. Temperaturänderung, Schwingungen des Meßinstruments, Unvollkommenheit des Auges des Beobachters) unabhängig voneinander das Meßergebnis beeinflussen. Jeder dieser Faktoren bewirkt einen unbedeutenden Teilfehler des Meßresultats. Da jedoch die Anzahl dieser Faktoren sehr groß ist, bewirken alle diese Einflußfaktoren zusammen einen bemerkenswerten Gesamtfehler des Meßergebnisses. Der betrachtete Gesamtfehler ist also die Summe einer Vielzahl voneinander unabhängiger und geringfügiger Teilfehler. Gemäß dem Zentralen Grenzwertsatz ist dann zu erwarten, daß der Gesamtfehler eine Verteilung aufweist, die der Normalverteilung sehr nahe kommt.

2. Wir betrachten die Herstellung von Bolzen auf einer automatischen Werkzeugmaschine. Dabei nehmen wir an, daß die Qualität eines Bolzens durch die Größe seines Durchmessers bestimmt ist. Als Zufallsgröße x fassen wir die Abweichung der Durchmesser der hergestellten Bolzen von dem Normwert für den Bolzendurchmesser ins Auge, auf den die automatische Anlage eingestellt ist. Offenbar gibt es eine große Anzahl von Faktoren, die auf die Tätigkeit des Automaten einen Einfluß ausüben (z. B. die Qualität des zu verarbeitenden Materials, das Zusammenspiel der einzelnen Aggregate der Maschine, die Abnutzung des Drehmeißels, die Temperaturänderung). Jeder dieser Faktoren, die wir als unabhängig voneinander wirkend ansehen, bedingt eine geringfügige Abweichung des Bolzendurchmessers von der vorgegebenen Norm. Wirkt jedoch der Komplex aller dieser Einflußfaktoren, dann ergibt sich schon eine beachtenswerte Abweichung für den Durchmesser des Bolzens von der Normalgröße. Nach dem Satz 2.36 besitzt nun die Zufallsvariable x eine angenäherte Normalverteilung.
   Es ist selbstverständlich, daß die im letzten Beispiel gezogenen Schlußfolgerungen nicht nur für die Herstellung von Bolzen, sondern für jede Massenproduktion von Erzeugnissen gelten, sobald der technologische Prozeß nicht verändert wird, solange also die einmal zugrunde gelegten technologischen Bedingungen konstant bleiben.

Fassen wir unsere Überlegungen bezüglich des von Tschebyscheff, Ljapunow und Markow stammenden Zentralen Grenzwertsatzes zusammen, dann können wir folgendes feststellen: Alle zufälligen Merkmale genügen einer Normalverteilung, wenn sie folgende Charakterisierung zulassen: Das zufällige Merkmal x steht unter dem Einfluß einer Vielzahl voneinander unabhängig wirkender zufälliger Ursachen, von denen jede nur einen äußerst geringen Einfluß auf das Merkmal x, d. h. auf die Gesamtwirkung aller Ursachen ausübt. Von den ursächlichen Faktoren, die in der Praxis nicht einmal alle aufgezählt und genannt werden können, wird sonst nichts vorausgesetzt; insbesondere dürfen diese Faktoren beliebig verteilt sein.

Wir beschreiben nun den *Grenzwertsatz von* LAPLACE-MOIVRE, den

Satz 2.37: Die binomische oder BERNOULLIsche Verteilung (2.25) strebt gegen die Normalverteilung mit dem Erwartungswert $E(x) = np$ und der Streuung $\sigma^2(x) = np(1 - p)$, sofern mit n auch np sehr groß wird.

Im Abschnitt 2.3 haben wir uns recht ausführlich mit der binomischen Verteilung beschäftigt. Erhöhen wir nun die Anzahl n der Ziehungen im BERNOULLIschen Schema beträchtlich und liegt die Wahrscheinlichkeit p (für das Ziehen einer weißen Kugel) nicht in der Nähe der Zahl null, so daß dann auch np beträchtlich erhöht wird, dann ist gemäß dem vorstehenden Satz die Zufallsgröße x (Anzahl der insgesamt gezogenen weißen Kugeln) annähernd normalverteilt. Die im Abschnitt 2.3 skizzierten Wahrscheinlichkeitsdiagramme nähern dann also die Kurve für die Dichtefunktion der Normalverteilung mit den Parametern $E(x) = np$ und $\sigma^2(x) = np(1 - p)$ an.

Bei hinreichend großem n und nicht zu kleinem p kann demzufolge die BERNOULLIsche Verteilungsfunktion $B_n(\xi)$ in (2.27) durch die normale Verteilungsfunktion $W(\xi)$ in (2.83) mit den Parametern $E(x) = np$ und $o^2(x) = np(1 - p)$ ersetzt werden. Beachten wir nun noch den Satz 2.29, so ergibt sich aus dem angegebenen Grenzwertsatz von LAPLACE-MOIVRE der

Satz 2.38: Wird mit n auch np sehr groß, dann gilt:

$$\boxed{\lim_{\substack{n \to \infty \\ np \to \infty}} B_n(\xi) = \Phi\left(\frac{\xi - np}{\sqrt{np(1-p)}}\right)}\,. \tag{2.104}$$

Die Bedeutung des in Rede stehenden Grenzwertsatzes bringt schon der letzte Satz zum Ausdruck. Sind die diesbezüglichen Voraussetzungen erfüllt, so kann die schwierig zu berechnende Funktion $B_n(\xi)$ — vgl. die Struktur der Formel (2.27) im Zusammenhang mit (2.24) — näherungsweise durch die tabellarisch gegebene Funktion Φ ersetzt werden. Je größer die Zahl n ist und je näher die Größe p bei der Zahl eins liegt, um so genauer ist diese Ersetzung von B_n durch Φ.

Zwei Beispiele mögen diesen Sachverhalt noch klarer hervortreten lassen.

1. Wir betrachten die im Abschnitt 2.3 behandelte Aufgabe, bei deren Lösung die Gleichung (2.28) auftrat: In einem Betrieb sind 60 Prozent der normgerechten Erzeugnisse von der Sorte 1. Es werden Packungen zu je 100 Erzeugnissen (durch beliebige Zusammenstellung der normgerechten) hergestellt. Wie groß ist die Wahrscheinlichkeit p_2 dafür, daß in einer Packung weniger als die Hälfte, d. h. weniger als 50 der Gütesorte 1 angehören? Für die gesuchte Wahrscheinlichkeit ergibt sich

$$p_2 = \sum_{x=0}^{49} P_{100}(x) = B_{100}(50)\,.$$

Beachten wir die BERNOULLIsche Formel (2.24), so erkennen wir, daß zur Berechnung des vorstehenden Ausdrucks die Berechnung von 50 Summanden erforderlich ist, die jeweils aus Produkten von Binomialkoeffizienten und Potenzen bestehen. Durch Anwendung des Satzes 2.38 — hier ist $n = 100$ und $p = 0,6$ — erhalten wir dagegen

$$p_2 = B_{100}(50) \approx \Phi\left(\frac{50 - 60}{\sqrt{24}}\right) = \Phi(-2,04)$$

oder vermöge (2.100)

$$p_2 \approx 1 - \Phi(2,04)$$

Um diesen Ausdruck zu berechnen, werfen wir einen Blick in die Φ-Tafel. Wegen $\Phi(2,04) = 0,9793$ lautet das Ergebnis

$$p_2 = 0,0207 .$$

Dieses Ergebnis stimmt mit dem Resultat in (2.28) überein: In etwa 1/50 aller Packungen gehören weniger als die Hälfte der Erzeugnisse zur Qualitätsstufe 1.

2. Wir betrachten eine Urne mit schwarzen und weißen Kugeln. Es sei $p = 0,8$ die Wahrscheinlichkeit dafür, daß bei einer zufälligen Ziehung eine weiße Kugel erscheint. Wie groß ist nun die Wahrscheinlichkeit p_0 dafür, daß die tatsächliche Zahl x der gezogenen weißen Kugeln bei 1000 unabhängigen Ziehungen mindestens 780 und höchstens 820 beträgt? Für p_0 besteht zunächst die Gleichung

$$p_0 = \sum_{x=780}^{820} P_{1000}(x) = \sum_{x=0}^{820} P_{1000}(x) - \sum_{x=0}^{779} P_{1000}(x)$$
$$= B_{1000}(820) - B_{1000}(779) .$$

Um die mühsame Berechnung der vorstehenden Summe von 41 Summanden zu umgehen, wenden wir den Grenzwertsatz von LAPLACE-MOIVRE an, gemäß dem

$$B_{1000}(\xi) = \Phi\left(\frac{\xi - 800}{\sqrt{160}}\right)$$

gilt und erhalten

$$p_0 \approx \Phi\left(\frac{20}{\sqrt{160}}\right) - \Phi\left(\frac{-21}{\sqrt{160}}\right) = \Phi(1,58) - \Phi(-1,66)$$

oder vermöge (2.100)

$$p_0 \approx \Phi(1,58) + \Phi(1,66) - 1 .$$

Unter Zuhilfenahme der Φ-Tafel bekommen wir die gesuchte Wahrscheinlichkeit schließlich zu

$$p_0 \approx 0,9429 + 0,9515 - 1 = 0,8944 .$$

Wir berechnen nun noch p_0 nach der BERNOULLIschen Streuungsungleichung (2.34) zu

$$q_0 = P[|x - E(x)| < 21] \geqq 1 - \frac{160}{21^2} = 1 - 0{,}3628$$
$$= 0{,}6372\,.$$

Wir sehen an dieser Stelle nochmals, daß die BERNOULLIsche Ungleichung eine sehr grobe Abschätzung für die gesuchte Wahrscheinlichkeit liefert. Die beiden betrachteten Beispiele demonstrieren in der Tat, daß die Anwendung des Grenzwertsatzes von LAPLACE-MOIVRE bei der Behandlung der binomischen Verteilung bedeutende rechnerische Vorteile in sich birgt und äußerst leicht zu handhaben ist. Insgesamt haben wir damit zwei Grenzübergänge bezüglich der BERNOULLIschen Verteilung kennengelernt, die sich bei $n \to \infty$ ergeben. Wächst bei diesem Grenzübergang auch np über alle Grenzen, dann strebt die Verteilung — wie wir soeben festgestellt haben — gegen eine Normalverteilung. Bleibt dagegen $np = a$ annähernd konstant, so ergibt sich im Grenzfall die POISSONsche Verteilung, die wir im Abschnitt 2.4 — vgl. (2.39) — untersucht haben.

2.6.4. Anwendungsbeispiele

1. In einem Sägewerk werden Leisten zugeschnitten. Die Leistenlänge ist eine normalverteilte Zufallsgröße x mit dem Erwartungswert (Mittelwert) $E(x) = 200$ cm und dem Streuungsmaß $\sigma(x) = 3$ cm. Wieviel Prozent der zugeschnittenen Leisten sind länger als 199 cm und kürzer als 202 cm? Wie groß ist die Wahrscheinlichkeit dafür, daß die tatsächliche Länge um weniger als 2,5 cm vom Mittelwert abweicht?
   Bezeichnen wir die gefragten Wahrscheinlichkeiten mit p_1 und p_2, so erhalten wir gemäß (2.102)

$$1 - p_1 = \Phi\left(\frac{202 - 200}{3}\right) - \Phi\left(\frac{199 - 200}{3}\right) = \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right)$$

oder vermöge (2.100)

$$1 - p_1 = \Phi\left(\frac{2}{3}\right) + \Phi\left(\frac{1}{3}\right) - 1 = 0{,}7475 + 0{,}6306 - 1 = 0{,}3781$$

und infolge (2.103)

$$p_2 = \Theta\left(\frac{2{,}5}{3}\right) = \Theta\,(0{,}833)$$

oder wegen (2.98)

$$p_2 = 2\,\Phi\,(0{,}833) - 1 = 2 \cdot 0{,}7976 - 1 = 0{,}5952\,.$$

37,81 Prozent der zugeschnittenen Leisten sind damit länger als 199 cm und kürzer als 202 cm. Mit einer Wahrscheinlichkeit von 59,52 Prozent weicht die Leistenlänge um weniger als 2,5 cm vom Mittelwert 200 cm ab.

2. Von einer Großbäckerei werden die Backwarengeschäfte mit Zwiebackpackungen beliefert, deren Mindestgewicht 170 Gramm betragen soll. Eine umfangreiche Untersuchung ergab für das normalverteilte Gewicht x der Packungen den Erwartungswert $E(x) = 176$ Gramm und das Streuungsmaß $\sigma(x) = 3{,}3$ Gramm. Mit welcher Wahrscheinlichkeit wird das Mindestgewicht der Packungen unterschritten? Um wieviel Gramm muß das ursprünglich auf einem Gewicht von 180 Gramm beruhende Kalkulationsgewicht für die Rohstoffe einer Zwiebackpackung erhöht werden, damit praktisch eine Unterschreitung des Mindestgewichts nicht mehr vorkommt? Die gesuchte Wahrscheinlichkeit berechnet sich zu

$$P(-\infty < x < 170)\,\Phi = \left(\frac{170-176}{3{,}3}\right) - \Phi\left(\frac{-\infty-176}{3{,}3}\right) =$$
$$= \Phi(-1{,}82) - \Phi(-\infty)$$

oder wegen (2.100) zu

$$\Phi(\infty) - \Phi(1{,}82) = 1 - 0{,}9656 = 0{,}0344\,.$$

Damit unterschreiten 3,44 Prozent oder 1/29 aller Zwiebackpackungen das Mindestgewicht von 170 Gramm. Das ursprünglich auf 180 angesetzte Kalkulationsgewicht liegt um 4 Gramm über dem Mittelwert und ist auf Grund des erhaltenen Resultats zu niedrig festgelegt. Erhöhen wir das Kalkulationsgewicht auf 184 Gramm, dann wird $E(x) = 180$ Gramm und damit die gewünschte Unterschreitungswahrscheinlichkeit

$$P(-\infty < x < 170) = \Phi\left(\frac{170-180}{3{,}3}\right) - \Phi\left(\frac{-\infty-180}{3{,}3}\right) =$$
$$\Phi(-3{,}03) - \Phi(-\infty)$$

oder gemäß (2.100)

$$\Phi(\infty) - \Phi(3{,}03) = 1 - 0{,}9987 = 0{,}0013\,.$$

Nunmehr sind lediglich 0,13 Prozent oder 1/769 der Packungen mit unzulässigem Gewicht. Würden wir nochmals das Kalkulationsgewicht um 4 Gramm erhöhen, dann würde sich wegen $E(x) = 184$ die Unterschreitungswahrscheinlichkeit zu

$$P(-\infty < x < 170) = \Phi\left(\frac{170-184}{3{,}3}\right) - \Phi\left(\frac{-\infty-184}{3{,}3}\right) =$$
$$= \Phi(-4{,}24) - \Phi(-\infty)$$

oder vermöge (2.100) zu

$$= \Phi(\infty) - \Phi(4{,}24) = 1 - 0{,}99999 = 0{,}00001$$

ergeben. Jetzt sind nur noch 1/1000 Prozent Zwiepackpackungen vorhanden, deren Gewicht unterhalb des Mindesgewichts liegt.

3. Auf einer Maschine werden Kugeln für Kugellager hergestellt, deren normalverteilter Durchmesser 8 mm beträgt. Durch Regulierung der Maschine ist es möglich, die Streuung der Durchmesser der herzustellenden Kugeln einzustellen, wobei allerdings zu beachten ist, daß mit einer Ver-

kleinerung bzw. Vergrößerung der Streuung die Produktionskosten für die Kugeln zunehmen bzw. absinken (sonst würde man von vornherein die minimalste Streuung für die Kugeldurchmesser einstellen). Auf welchen Wert muß nun das Streuungsmaß σ gebracht werden, damit 95 Prozent der Kugeln den Abnahmebedingungen gerecht werden, die eine Toleranz von weniger als 0,02 mm Abweichung vom Mittelwert für den Durchmesser der Kugeln vorschreiben? Wir gehen von (2.103) aus:

$$P\left[|x - E(x)| < 0{,}02\right] = \Theta\left(\frac{0{,}02}{\sigma(x)}\right) = 0{,}95$$

und bekommen auf Grund (2.99)

$$\Phi\left(\frac{0{,}02}{\sigma(x)}\right) = \frac{1}{2}\,(0{,}95 + 1) = 0{,}975\,.$$

In der Φ-Tafel finden wir den Zusammenhang $\Phi\,(1{,}96) = 0{,}975$ und damit das Ergebnis zu

$$\frac{0{,}02}{\sigma(x)} = 1{,}95$$

oder kurz

$$\sigma(x) = 0{,}01\,.$$

Wird also das Streuungsmaß auf 0,01 mm eingestellt, dann entsprechen nur 5 Prozent der Kugeln den Abnahmebedingungen nicht. Wieviel Prozent der hergestellten Kugeln sind nicht abnahmefähig, wenn das Streuungsmaß auf das Doppelte eingestellt wird? Wir finden

$$P\left[|x - E(x)| < 0{,}02\right] = \Theta\left(\frac{0{,}02}{0{,}02}\right) = \Theta\,(1)$$

und weiter vermöge (2.98)

$$= 2\,\Phi\,(1) - 1 = 2 \cdot 0{,}8413 - 1 = 0{,}6826\,.$$

Erhöhen wir also das Streuungsmaß auf 0,02 mm, dann werden immerhin schon 31,74 Prozent der gefertigten Kugeln nicht mehr abgenommen.

4. Bei verschiedenen praktischen Problemen — z. B. in der Stichprobentheorie — wird damit gerechnet, daß eine normalverteilte Zufallsvariable x mit dem Streuungsmaß $\sigma\,(x)$ um weniger als $3\,\sigma\,(x)$ von ihrem Erwartungswert $E\,(x)$ abweicht. Anders ausgedrückt: Es wird die Beziehung

$$P\left(|x - E(x)| < 3\,\sigma\right) = 1$$

angenommen. Welche Gründe gibt es für diese Faustregel? Für die linke Seite der letzten Gleichung erhalten wir nach (2.103)

$$P\left[|x - E(x)| < 3\,\sigma\right] = \Theta\left(\frac{3\sigma(x)}{\sigma(x)}\right) = \Theta\,(3)$$

oder als Folge von (2.98)

$$= 2\,\Phi\,(3) - 1 = 2 \cdot 0{,}9986 - 1 = 0{,}9972\,.$$

In 99,72 Prozent aller Fälle ist der behauptete Sachverhalt mithin gültig.

5. Von einem Ort O_1 werden zu einem anderen Ort O_2 Telefonleitungen gelegt. In O_1 gibt es $n = 5000$ Teilnehmer, die pro Stunde durchschnittlich eine Minute $\left(p = \frac{1}{60}\right)$ mit O_2 zu telefonieren beabsichtigen. Wieviel Leitungen müssen nun von O_1 nach O_2 gelegt werden, damit im Mittel höchstens 5 Prozent aller Gespräche auf besetzte Leitungen treffen? Wir haben hier eine BERNOULLIsche Verteilung mit $n = 5000$, $p = \frac{1}{60}$ und $q = \frac{59}{60}$ vor uns[1], die den Erwartungswert

$$E(x) = np \approx 83{,}33$$

und das Streuungsmaß

$$\sigma(x) = \sqrt{npq} \approx 9{,}05$$

besitzt. Bezeichnen wir die gesuchte Anzahl der zu legenden Telefonleitungen mit l, dann soll

$$B_{5000}(l+1) \geqq 0{,}95$$

bestehen. ($B_{5000}(l+1)$ gibt die Wahrscheinlichkeit dafür an, daß höchstens l Gespräche gleichzeitig geführt werden!).
Wir können uns vorstellen, wie mühsam die Ermittlung von l aus der letzten Ungleichung wäre, wenn der Grenzwertsatz von LAPLACE-MOIVRE nicht verwendet würde. Gemäß (2.104) erhalten wir

$$B_{5000}(l+1) = \Phi\left(\frac{l+1-83{,}3}{9{,}05}\right) \geqq 0{,}95$$

Wegen $\Phi(1{,}645) = 0{,}95$ und der Monotonie der Φ-Funktion ergibt sich

$$\frac{l+1-83{,}3}{9{,}05} \geqq 1{,}645$$

oder kurz

$$l \approx 97{,}2\,.$$

Es müssen demzufolge mindestens 98 Leitungen gelegt werden, damit lediglich mit einer Wahrscheinlichkeit von höchstens 5 Prozent alle Leitungen von O_1 nach O_2 besetzt sind.

6. Von einem Ort O_1 führt täglich ein Personenzug zu einem Ort O_2. In O_1 gibt es 1500 Personen, die im Mittel zweimal im Monat diesen Zug benutzen und jeweils — soweit vorhanden — dann einen Sitzplatz einnehmen. Wie groß ist die Wahrscheinlichkeit dafür, daß ein Reisender in einem Zug mit 120 Sitzplätzen keinen Sitzplatz vorfindet? Mit wieviel Plätzen muß der Zug von O_1 aus eingesetzt werden, damit höchstens

[1] Daß hier in der Tat eine Binomialverteilung vorliegt, erkennen wir sofort, wenn wir uns des Urnenschemas von BERNOULLI (vgl. S. 57) vergegenwärtigen; der weißen Kugel dort entspricht hier ein telefonierender Fernsprechteilnehmer.

1 Prozent der Reisenden keinen Sitzplatz erhalten? Wie groß ist schließlich die Wahrscheinlichkeit dafür, daß ein Zug mit 110 Plätzen nach Verlassen von O_1 noch mindestens 20 freie Sitzplätze aufweist?
Wir haben hier wiederum eine binomische Verteilung vor uns. Die Zufallsgröße ist die Zahl der den Zug benutzenden Reisenden aus O_1; die zugehörigen Parameter sind $n = 1500$, $p = \frac{1}{15}$ und $q = \frac{14}{15}$. Die hier vorliegende Verteilung hat daher den Erwartungswert

$$E(x) = np = 100$$

und das Streuungsmaß

$$\sigma(x) = \sqrt{npq} \approx 9{,}66\,.$$

Für die erste Wahrscheinlichkeit gilt zunächst

$$P(x > 120) = 1 - P(x < 121) = 1 - B_{1500}(121)$$

oder vermöge des Grenzwertsatzes

$$P(x > 120) = 1 - \Phi\left(\frac{121 - 100}{9{,}66}\right) = 1 - \Phi(2{,}17)\,.$$

Unter Zuhilfenahme der Φ-Tafel lautet das Ergebnis

$$P(x < 120) = 1 - 0{,}9850 = 0{,}0150\,.$$

Die gesuchte Wahrscheinlichkeit beträgt somit 1,5 Prozent. Wir ermitteln nun die gefragte Anzahl der Plätze, die wir mit k bezeichnen, indem wir von

$$P(x > k) = 1 - P(x < k + 1) = 1 - B_{1500}(k + 1) \leqq 0{,}01.$$

ausgehen. Wenden wir auf die letzte Gleichung den Satz 2.38 an, so wird

$$1 - \Phi\left(\frac{k + 1 - 100}{9{,}66}\right) \leqq 0{,}01$$

oder

$$\Phi\left(\frac{k + 1 - 100}{9{,}66}\right) \geqq 0{,}99\,.$$

Aus $\Phi(2{,}323) = 0{,}99$ und der Monotonie von $\Phi(\xi)$ resultiert schließlich

$$\frac{k + 1 - 100}{9{,}66} \geqq 2{,}323$$

oder kurzerhand

$$k \approx 123{,}4\,.$$

Werden mithin mindestens 123 Plätze eingesetzt, dann bekommt nur höchstens jeder Hundertste der Reisenden keinen Sitzplatz.

Wir gehen nun an die Berechnung der letzten gesuchten Wahrscheinlichkeit, für die

$$P\,(x < 90) = B_{1560}\,(90) = \Phi\left(\frac{90 - 100}{9{,}66}\right) = \Phi\,(-\,1{,}035)$$

oder vermöge (2.100) und Φ-Tafel

$$P\,(x < 90) = 1 - \Phi\,(1{,}035) = 1 - 0{,}8497 = 0{,}1503$$

gilt. Ein mit 110 Plätzen versehener Zug wird also in 15 Prozent aller Fälle mindestens 20 freie Plätze nach Verlassen von O_1 aufweisen.

7. Auf einer automatischen Werkzeugmaschine werden Walzen hergestellt, deren Durchmesser x als eine normalverteilte Zufallsgröße mit dem Erwartungswert $E\,(x) = 30$ cm und dem Streuungsmaß $\sigma\,(x) = 0{,}1$ cm erscheint. Wieviel Prozent der gefertigten Erzeugnisse sind stärker als 30,1 cm bzw. schwächer als 29,8 cm? Wie groß ist die Wahrscheinlichkeit dafür, daß die Abweichung des Durchmessers vom Sollwert = Mittelwert mindestens 1,5 mm beträgt? Am Tag werden 7000 Walzen hergestellt. Eine Walze ist weiterhin normgerecht, wenn ihr Durchmesser von dem Sollwert um weniger als 2 mm abweicht? Wieviel normgerechte Walzen erzeugt der Automat durchschnittlich im Laufe eines Monats (= 30 Tage)? Für die erste gesuchte Wahrscheinlichkeit ergibt sich

$$P\,(x > 30{,}1) = 1 - P\,(x \leq 30{,}1) = 1 - \Phi\left(\frac{30{,}1 - 30}{0{,}1}\right) = 1 - \Phi\,(1)$$

oder

$$P\,(x > 30{,}1) = 1 - 0{,}8413 = 0{,}1587\,.$$

Entsprechend wird im zweiten Fall

$$P\,(x < 29{,}8) = \Phi\left(\frac{29{,}8 - 30}{0{,}1}\right) = \Phi\,(-\,2) = 1 - \Phi\,(2) = $$
$$= 1 - 0{,}9772 = 0{,}0228\,.$$

Die dritte Wahrscheinlichkeit finden wir zu

$$P\,[|\,x - E\,(x)|\, \geqq 0{,}15] = 1 - P\,[|\,x - E\,(x)| < 0{,}15] =$$
$$= 1 - \Theta\left(\frac{0{,}15}{0{,}1}\right) = 1 - \Theta\,(1{,}5)$$

oder gemäß (2.98) zu

$$P[x - E\,(x)] \geqq 0{,}15 = 1 - [2\Phi(1{,}5) - 1] = 1 - (2\cdot 0{,}9332 - 1) = 0{,}1336\,.$$

Die Wahrscheinlichkeit dafür, daß eine Walze der Norm genügt, beträgt

$$P\,[|\,x - E\,(x)| < 0{,}2] = \Theta\left(\frac{0{,}2}{0{,}1}\right) = \Theta\,(2)$$

oder wieder gemäß (2.98)

$$P\,[|x - E\,(x)| < 0{,}2] = 2\,\Phi\,(2) - 1 = 2\cdot 0{,}9772 - 1 = 0{,}9544\,.$$

Im Monat werden $30 \cdot 7000 = 210000$ Erzeugnisse hergestellt. Demzufolge beläuft sich die mittlere Anzahl der gefertigten normgerechten Walzen pro Monat auf $210000 \cdot 0{,}9544 = 200424$ Stück.

2.7. Vermischte Aufgaben nebst Lösungen

2.7.1. Aufgaben

1. Bestimme den Erwartungswert $E(x)$, die Streuung $\sigma^2(x)$ und das Streuungsmaß $\sigma(x)$ der diskreten Zufallsgröße x mit der Verteilungstabelle

x	7	8	9	10	11	12
p	$\frac{1}{9}$	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{9}$.

Berechne ferner die Wahrscheinlichkeit dafür, daß x vom Erwartungswert $E(x)$ um weniger als 2 abweicht, und zwar
a) nach der gegebenen Verteilungstabelle,
b) nach der TSCHEBYSCHEFFschen Ungleichung!

2. Berechne die wahrscheinlichsten Werte der BERNOULLIschen Verteilung mit den Parametern

$$a)\ p = \frac{2}{3},\ n = 17 \text{ und}$$

$$b)\ p = \frac{1}{4},\ n = 28!$$

3. In einem Betrieb sind 0,5 Prozent der hergestellten Erzeugnisse fehlerhaft. Wie groß ist die Wahrscheinlichkeit dafür, daß in einem Posten von 1000 Erzeugnissen alle fehlerfrei sind?

4. In einem Geschäft treffen im Durchschnitt in einer halben Stunde 20 Kunden ein. Wie groß ist die Wahrscheinlichkeit dafür, daß während der Dauer von 2 Stunden

   a) 60,

   b) 100 und

   c) 75 Kunden eintreffen?

5. Wie groß muß die Konstante C sein, damit $w(x) = C \cdot w^*(x)$ mit

$$w^*(x) \begin{cases} e^x & \text{für } x \leqq 0 \\ e^{-x} & \text{für } x \geqq 0 \end{cases}$$

eine Dichtefunktion für die stetige Zufallsgröße x wird?

6. Welcher Zahlenwert muß der Konstanten a erteilt werden, damit

$$w(x) = \frac{a}{e^x + e^{-x}}$$

eine Dichtefunktion für die kontinuierliche Zufallsvariable x wird? Zeichne diese Dichtefunktion und die zugehörige Verteilungsfunktion $W(\xi)$! Ferner berechne man die Wahrscheinlichkeit $P(-1 \leqq x \leqq 2)$.

7. Man bestimme Mittelwert $E(x)$ und Streuungsmaß $\sigma(x)$ der stetigen Zufallsgröße x mit der Dichtefunktion

$$w(x) = \begin{cases} 0, \text{ für } x < 0; \\ \frac{1}{2} \sin x, \text{ für } 0 \leqq x \leqq \pi; \\ 0, \text{ für } x > \pi. \end{cases}$$

8. Wie lautet die Dichtefunktion $w(x)$ einer normalverteilten Zufallsgröße x mit dem Erwartungswert $E(x) = 6$ und dem Streuungsmaß $\sigma(x) = 2$? Man gebe weiter die Wahrscheinlichkeit $P(0 < x < 5)$ an!

9. In einem pharmazeutischen Betrieb werden Medikamente in Ampullen abgefüllt. Der Mittelwert des Ampulleninhalts beträgt 10 ccm, das Streuungsmaß 0,1 ccm. Wie groß ist die Wahrscheinlichkeit p_1 dafür, daß der Ampulleninhalt um weniger als 0,25 ccm vom Sollwert (= Mittelwert) abweicht? Eine Apotheke erhält eine Lieferung von 10000 Ampullen. Wieviel Ampullen darunter haben erwartungsgemäß

   a) weniger als 9,8 ccm und

   b) mehr als 10,1 ccm Inhalt?

10. In der DDR werden in einem Zeitraum von etwa 4 Monaten $n = 10^5$ Kinder geboren. Die Wahrscheinlichkeit für eine Knabengeburt beträgt $p = 0{,}514$. Wie groß sind Erwartungswert $E(x)$ und Streuungsmaß $\sigma(x)$ für die Anzahl x der Knabengeburten unter diesen 10^5 Geburten? Man berechne die Wahrscheinlichkeit dafür, daß die tatsächliche Anzahl der Knabengeburten zwischen den Werten $E(x) + 300$ und $E(x) - 300$ liegt und zwar

    a) nach der BERNOULLIschen Streuungsungleichung,

    b) nach dem Grenzwertsatz von LAPLACE-MOIVRE.

2.7.2. Lösungen

1. Zunächst berechnen wir aus (2.5) den Mittelwert

$$E(x) = 7 \cdot \frac{1}{9} + 8 \cdot \frac{1}{6} + 9 \cdot \frac{1}{9} + 10 \cdot \frac{1}{3} + 11 \cdot \frac{1}{6} + 12 \cdot \frac{1}{9}$$

$$\text{oder } E(x) = \frac{173}{18} \approx 9{,}61\,.$$

Die Streuung erhalten wir gemäß (2.15) zu

$$\sigma^2(x) = \left(-\frac{47}{18}\right)^2 \frac{1}{9} + \left(-\frac{29}{18}\right)^2 \frac{1}{6} + \left(-\frac{11}{18}\right)^2 \cdot \frac{1}{9} + \left(\frac{7}{18}\right)^2 \frac{1}{3} + \left(\frac{25}{18}\right)^2 \cdot \frac{1}{6} + \left(\frac{43}{18}\right)^2 \cdot \frac{1}{9}$$

oder

$$\sigma^2(x) = \frac{725}{324} \approx 2{,}24\,.$$

Damit beträgt das Streuungsmaß

$$\sigma(x) = \frac{\sqrt{725}}{18} \approx 1{,}50\,.$$

a) Für die zu ermittelnde Wahrscheinlichkeit gilt

$$P\left(\left|x - \frac{173}{18}\right| < 2\right) = P(''x = 8'' + ''x = 9'' + ''x = 10'' + ''x = 11'')$$

oder aufgrund des Satzes 1.1 und der gegebenen Verteilungstabelle

$$P\left(\left|x - \frac{173}{18}\right| < 2\right) = \frac{1}{6} + \frac{1}{9} + \frac{1}{3} + \frac{1}{6} = \frac{7}{9} \approx 0{,}78\,.$$

b) Unter Zuhilfenahme der TSCHEBYSCHEFFschen Streuungsungleichung (2.21) erhalten wir die folgende grobe Abschätzung für die soeben berechnete Wahrscheinlichkeit

$$P\left(\left|x - \frac{173}{18}\right| < 2\right) \geqq 1 - \frac{725}{324{,}4} \approx 0{,}44\,.$$

2. a) Wir bilden zunächst die Größe

$$np + p = 17 \cdot \frac{2}{3} + \frac{2}{3} = 12\,.$$

Da diese Größe ganz ist, folgt aus Satz 2.16: die vorgegebene binomische Verteilung besitzt die beiden gleichwahrscheinlichsten Werte $x_0' = 12 - 1 = 11$ und $x_0'' = 12$.

b) Wir gehen wieder von der Größe

$$np + p = 28 \cdot \frac{1}{4} + \frac{1}{4} = \frac{29}{4}$$

aus, die hier nicht ganz ist. Als Folge des Satzes 2.16 weist daher die vorgelegte binomische Verteilung den einzigen wahrscheinlichsten Wert $x_0 = 7$ auf.

3. Wir haben eine BERNOULLIsche Verteilung mit den Parametern $p = 0{,}005$ und $n = 1000$ vor uns. Die gesuchte Wahrscheinlichkeit berechnet sich mühsam gemäß der Formel (2.24) zu

$$P_{1000}(o) = \binom{1000}{0} 0{,}005^0 (1 - 0{,}005)^{1000}$$

oder

$$P_{1000}(o) = 1000 \cdot 0{,}995^{1000} \approx 0{,}0063\,.$$

Einfacher ist folgender Weg: Da n sehr groß ist, p dagegen sehr klein, nähern wir die vorliegende Binomialverteilung durch die zugehörige POISSONsche Verteilung mit dem Parameter $a = np = 5$ an. Dann gilt

$$P_{1000}(o) \approx P^*_{1000}(o)$$

oder wegen (2.39)

$$P_{1000}(0) \approx \frac{a^o \cdot e^{-a}}{o!} = e^{-5} \approx 0{,}0067\,.$$

4. Wir bedienen uns der Beziehung (2.47) und wählen als Zeiteinheit die halbe Stunde.

   a) In diesem Fall liegen die Werte $\lambda = 20$, $T = 4$ und $x = 60$ vor, so daß wird

   $$E(20; 60; 4) = \frac{80^{60} \cdot e^{-80}}{60!} \approx 0{,}0033\,.$$

   b) Entsprechend wird bei $\lambda = 20$, $T = 4$ und $x = 100$

   $$E(20; 100; 4) = \frac{80^{100} \cdot e^{-80}}{100!} \approx 0{,}0033\,.$$

   c) Schließlich wird bei $\lambda = 20$, $T = 4$ und $x = 75$

   $$E(20; 75; 4) = \frac{80^{75}\, e^{-80}}{75!} \approx 0{,}0391\,.$$

5. Die gegebene Verteilungsdichte muß die Vollständigkeitsrelation (2.51) erfüllen, d. h.

$$1 = c \int_{-\infty}^{+\infty} w^*(x)\,dx = c \left\{ \int_{-\infty}^{o} w^*(x)\,dx + \int_{o}^{\infty} w^*(x)\,dx \right\}$$

oder wegen der speziellen Struktur von $w^*(x)$

$$1 = c \left\{ \int_{-\infty}^{o} e^x\,dx + \int_{o}^{\infty} e^{-x}\,dx \right\}.$$

Beachten wir die unbestimmten Integrale

$$\int e^x\,dx = e^x\,, \int e^{-x}\,dx = -\,e^{-x}\,,$$

so erhalten wir

$$1 = c\left\{ \Big[\, e^x \,\Big]_{-\infty}^{o} + \Big[-e^{-x} \Big]_{0}^{\infty} \right\} = c\left\{ (1 - 0) + (0 + 1) \right\} = 2\,c$$

oder

$$c = \frac{1}{2}\,.$$

6. Wir wenden wieder die Vollständigkeitsrelation (2.51) an, die

$$1 = a \int_{-\infty}^{\infty} \frac{dx}{e^x + e^{-x}}$$

zur Folge hat. Weil der Integrand gerade ist, das heißt, die Eigenschaft $w(x) = w(-x)$ besitzt, können wir für den letzten Ausdruck auch

$$1 = 2a \int_o^\infty \frac{dx}{e^x + e^{-x}} \tag{*}$$

schreiben. Um das unbestimmte Integral

$$I = \int \frac{dx}{e^x + e^{-x}}$$

zu berechnen, erweitern wir den Integranden mit e^x, so daß

$$I = \int \frac{e^x\,dx}{(e^x)^2 + 1}$$

entsteht. Nun führen wird die Substitution $t = e^x$ mit $dt = e^x\,dx$ ein. Dann ergibt sich das bekannte Integral

$$I = \int \frac{dt}{1 + t^2} = \operatorname{arctg} t.$$

Machen wir die Transformation $t = e^x$ rückgängig, dann erhalten wir schließlich

$$I = \operatorname{arctg} t = \operatorname{arctg} e^x.$$

Dann resultiert aus (*)

$$1 = 2a \left[\operatorname{arctg} e^x\right]_o^\infty$$

oder $1 = 2a \left(\operatorname{arctg} \infty - \operatorname{arctg} 1\right) = 2a \left(\frac{\pi}{2} - \frac{\pi}{4}\right) = \frac{a\pi}{2}.$

Damit ist

$$a = \frac{2}{\pi}$$

das gesuchte Ergebnis.

Die Dichtefunktion der vorliegenden Verteilung lautet mithin

$$w(x) = \frac{2}{\pi} \cdot \frac{1}{e^x + e^{-x}}.$$

Gemäß (2.58) leiten wir nun die zugehörige Verteilungsfunktion $W(\xi)$ ab. Wir haben oben bereits

$$\int \frac{dx}{e^x + e^{-x}} = \operatorname{arctg} e^x$$

ermittelt, so daß wird

$$W(\xi) = \frac{2}{\pi} \int_{-\infty}^{\xi} \frac{dx}{e^x + e^{-x}} = \frac{2}{\pi} \left(\operatorname{arctg} e^\xi - \operatorname{arctg} o\right)$$

oder $W(\xi) = \frac{2}{\pi} \operatorname{arctg} e^\xi$

Die Funktionen $w(x)$ und $W(\xi)$ stellen wir im folgenden graphisch dar.

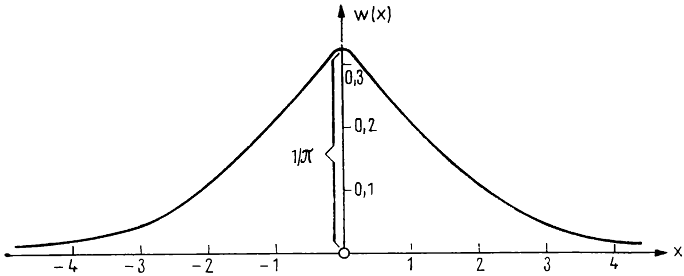

Abb. 20: Graphische Darstellung von $w(x) = \frac{2}{\pi} \frac{1}{e^x + e^{-x}}$

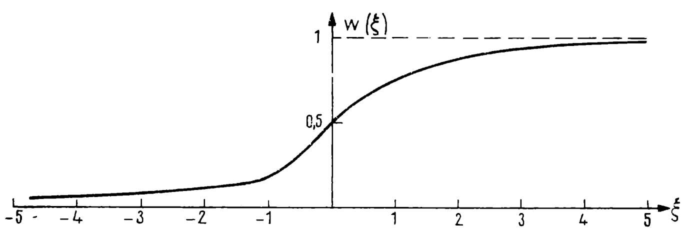

Abb. 21: Graphische Darstellung von $W(\xi) = \frac{2}{\pi} \operatorname{arc\,tg} e^{\xi}$

7. Wir benutzen die Beziehungen (2.65). Dabei unterteilen wir das Integrationsintervall $(-\infty, +\infty)$ in die Teilbereiche $(-\infty, o)$, (o, π) und $(\pi, +\infty)$, da die gegebene Dichtefunktion in diesen unterschiedlich erklärt ist.

   Wir erhalten für den Erwartungswert gemäß (2.65_1)

$$E(x) = \int_{-\infty}^{o} x \cdot o \, dx + \int_{o}^{\pi} x \cdot \frac{1}{2} \sin x \, dx + \int_{\pi}^{\infty} x \cdot o \, dx$$

   oder einfach

$$E(x) = \frac{1}{2} \int_{0}^{\pi} x \cdot \sin x \, dx \,.$$

   Durch partielle Integration finden wir das unbestimmte Integral

$$\int x \cdot \sin x \, dx = \sin x - x \cdot \cos x \,. \qquad (*)$$

(dieses Resultat läßt sich durch Differentiation auf beiden Seiten schnell bestätigen!)

Damit ergibt sich der Mittelwert zu

$$E(x) = \frac{1}{2}\Big[\sin x - x \cdot \cos x\Big]_0^\pi = \frac{1}{2}\Big(o + \pi + o + o\Big) = \frac{\pi}{2}.$$

Ganz entsprechend bekommen wir für die Streuung $\sigma^2(x)$ vermöge (2.65_2)

$$\sigma^2(x) = \frac{1}{2}\int_0^\pi \left(x - \frac{\pi}{2}\right)^2 \cdot \sin x\, dx$$

oder

$$\sigma^2(x) = \frac{1}{2}\int_0^\pi x^2 \sin x\, dx - \frac{\pi}{2}\int_0^\pi x \cdot \sin x\, dx + \frac{\pi^2}{8}\int_0^\pi \sin x\, dx.$$

Beachten wir das unbestimmte Integral

$$\int x^2\, sin\, x\, dx = - x^2 cos\, x + 2x\, sin\, x + 2\, cos\, x$$

sowie (*), so folgt

$$\sigma^2(x) = \frac{1}{2}\Big[(2 - x^2)\cos x + 2x \cdot \sin x\Big]_0^\pi - \frac{\pi}{2}\Big[\sin x - x\cos x\Big]_0^\pi + \frac{\pi^2}{8}\Big[- \cos x\Big]_0^\pi$$

oder

$$\sigma^2(x) = \frac{1}{2}(\pi^2 - 4) - \frac{\pi^2}{2} + \frac{\pi^2}{4} = \frac{\pi^2}{4} - 2.$$

Das Streuungsmaß $\sigma(x)$ hat damit den Wert

$$\sigma(x) = \sqrt{\frac{\pi^2}{4} - 2} \approx 0{,}684.$$

8. Wir setzen in dem allgemeinen Ausdruck (2.82) für eine normale Verteilungsdichte die Parameter $E(x) = 6$ und $\sigma(x) = 2$ ein.

Es resultiert dann

$$w(x) = \frac{1}{2\sqrt{2\pi}}\, e^{-\frac{(x-6)^2}{8}}$$

Wenden wir die Beziehung (2.102) an, dann erhalten wir für die gesuchte Wahrscheinlichkeit

$$P(o < x < 5) = \Phi\left(\frac{5-6}{2}\right) - \Phi\left(\frac{0-6}{2}\right)$$

oder

$$P(o < x < 5) = \Phi\left(-\frac{1}{2}\right) - \Phi\left(-3\right).$$

Berücksichtigen wir (2.100) und die Φ-Tafel, so erhalten wir das Resultat

$$P(o < x < 5) = \Phi(3) - \Phi\left(\frac{1}{2}\right) = 0{,}9986 - 0{,}6915 = 0{,}3071\,.$$

9. Gegeben sind die Parameter $E(x) = 10$ und $\sigma(x) = 0{,}1$ für die Verteilung der Zufallsgröße $x =$ Ampulleninhalt, die wir als normal ansehen. Von Interesse ist die Wahrscheinlichkeit

$$p_1 = P(|x - E(x)| < 0{,}25)\,,$$

die wir gemäß (2.103) zu

$$p_1 = \Theta\left(\frac{0{,}25}{0{,}1}\right) = \Theta\,(2{,}5)$$

ermitteln. Beachten wir (2.98) und die Φ-Tafel, so resultiert schließlich

$$p_1 = 2\,\Phi\,(2{,}5) - 1 = 2 \cdot 0{,}9938 - 1 = 0{,}9876\,.$$

a) Wir berechnen nun die Wahrscheinlichkeit $q_1 = P(x < 9{,}8)$ auf Grund (2.102) zu

$$q_1 = P(-\infty < x < 9{,}8) = \Phi\left(\frac{9{,}8-10}{0{,}1}\right) - \Phi\left(\frac{-\infty-10}{0{,}1}\right)$$

oder

$$q_1 = \Phi\,(-2) - \Phi\,(-\infty)\,.$$

Wegen $\Phi(-\infty) = 0$ und (2.100) ergibt sich unter Zuhilfenahme der Φ-Tafel

$$q_1 = 1 - \Phi\,(2) = 1 - 0{,}9772 = 0{,}0228\,.$$

Demzufolge fällt der Inhalt einer Ampulle mit einer Wahrscheinlichkeit von 2,28 Prozent kleiner als 9,8 ccm aus. Wieviel Ampullen (z) solcher Art gibt es erwartungsgemäß unter einer Lieferung von $n = 1000$ Stück? Diese Frage beantworten wir, indem wir unsere Ergebnisse über die Binomialverteilung heranziehen. Die gesuchte Ampullenanzahl z ist offenbar gleich dem Erwartungswert der Binomialverteilung mit den Parametern $p = q_1 = 0{,}0228$ und $n = 1000$. Sie ergibt sich aus (2.29) zu

$$z = 0{,}0228 \cdot 1000 = 228\,.$$

In der Lieferung werden mithin 228 Ampullen erwartet, deren Inhalt kleiner als 9,8 ccm ist.

b) Ganz entsprechend wie in a) berechnen wir

$$q_2 = P(x > 10{,}1) = 1 - P(x \leqq 10{,}1)$$

zu

$$q_2 = 1 - P(-\infty < x \leqq 10{,}1) = 1 - \left\{\Phi\left(\frac{10{,}1 - 10}{0{,}1}\right) - \Phi\left(\frac{-\infty - 10}{0{,}1}\right)\right\}$$

oder

$$q_2 = 1 - \{\Phi(1) - \Phi(-\infty)\}.$$

Beachten wir wieder $\Phi(-\infty) = 0$ und die Φ-Tafel, so bleibt

$$q_2 = 1 - \Phi(1) = 1 - 0{,}8413 = 0{,}1587.$$

Weiter folgt aus diesem Teilergebnis, — in Analogie zu den Ausführungen in a) — daß in einer Lieferung von 1000 Ampullen

$$z = 0{,}1587 \cdot 10000 = 1587$$

Ampullen erwartet werden, deren Inhalt größer als 10,1 ccm ausfällt.

10. Gegeben ist die Binomialverteilung mit den Parametern $p = 0{,}514$ ($q = 1 - p = 0{,}486$) und $n = 10^5$. Den gesuchten Mittelwert $E(x)$ berechnen wir vermöge (2.29) zu

$$E(x) = n \cdot p = 0{,}514 \cdot 10^5 = 51\,400.$$

Aus (2.30) resultiert das Streuungsmaß

$$\sigma(x) = \sqrt{npq} = \sqrt{2{,}498} \cdot 10^2 \approx 158$$

Wir ermitteln nun die gefragte Wahrscheinlichkeit

$$p_0 = P(|x - E(x)| < 300)$$

a) Nach der Streuungsungleichung (2.34) ergibt sich unter Beachtung des für $\sigma(x)$ bereits gefundenen Werts sowie $\gamma = 300$

$$p_0 \geqq 1 - \frac{2{,}498 \cdot 10^4}{9 \cdot 10^4} = 0{,}7225.$$

b) Wenden wir den Grenzwertsatz von Laplace-Moivre an, indem wir das Merkmal x als normalverteilt ansehen, so erhalten wir vermöge (2.103)

$$p_0 = \Theta\left(\frac{300}{\sqrt{2{,}498} \cdot 10^2}\right) = \Theta\,\frac{3}{1{,}58} = \Theta(1{,}9).$$

Benutzen wir (2.98) und die Φ-Tafel, dann ergibt sich schließlich das Resultat

$$p_0 = 2\,\Phi(1{,}9) - 1 = 2 \cdot 0{,}9713 - 1 = 0{,}9426.$$

3. Einführung in die Theorie der Stichproben

3.1. Grundgesamtheit und Stichprobe

Eine *Stichprobe* ist eine Menge von Elementen, die einer Gesamtheit von Elementen, der sogenannten *Grundgesamtheit* oder *generellen Gesamtheit* entstammen. Die Anzahl der Elemente der Stichprobe heißt *Umfang der Stichprobe*. Der *generelle Umfang* oder der *Umfang der Grundgesamtheit* ist die Anzahl der Elemente der Grundgesamtheit. Der Stichprobenumfang kann offenbar nicht größer als der generelle Umfang sein. Wählen wir zum Beispiel aus 10000 Erzeugnissen eines Betriebes 50 aus, so haben wir eine Stichprobe des Umfanges 50 vor uns; die Grundgesamtheit besitzt in diesem Falle den Umfang 10000.

In diesem Sinne ist jede Reihe von Beobachtungs- und Versuchsergebnissen eine Stichprobe aus der Gesamtheit der möglichen Beobachtungs- und Versuchsresultate. Der Umfang solcher Stichproben ist gleich der Anzahl der durchgeführten Beobachtungen und Versuche.

Die Stichproben spielen eine große Rolle in der Statistischen Qualitätskontrolle. Wegen des hohen Zeit- und Kostenaufwandes ist es nämlich oft untragbar, alle aus einem Fertigungsprozeß stammenden Erzeugnisse eines Betriebes auf ihre Qualität hin zu untersuchen. Falls das Erzeugnis bei der Kontrolle (Prüfung) zerstört (zum Beispiel Kunstfasern, Glühlampen) bzw. sein Gebrauchswert herabgesetzt wird, verbietet sich eine Prüfung aller hergestellten Erzeugnisse (das heißt der Grundgesamtheit) von selbst. In beiden Fällen ist die Beschränkung der Kontrolle auf einen Teil der Grundgesamtheit, das heißt auf eine Stichprobe, unumgänglich.

Wir führen nun einige Begriffsbildungen und Bezeichnungen ein, die für das weitere Verständnis der Stichprobentheorie bedeutsam sind. Zunächst nennen wir eine Stichprobe *zufällig*, wenn ihre einzelnen Elemente rein zufällig der Grundgesamtheit entnommen werden. In diesem Kapitel betrachten wir lediglich zufällige Stichproben.

Unter einer *großen Stichprobe* verstehen wir im folgenden stets eine zufällige Stichprobe, deren Umfang mindestens 100 beträgt. Im anderen Fall sprechen wir von einer *kleinen Stichprobe*.

Eine *Grundgesamtheit erster Art* ist eine solche, deren Elemente durch eine stetige oder diskrete Zufallsvariable x, das *generelle Merkmal*, charakterisiert werden können. Das Merkmal läßt sich zahlenmäßig ausdrücken. So bildet beispielsweise die Menge der in einem Sägewerk zugeschnittenen Leisten eine Grundgesamtheit erster Art, wenn die Leistenlänge x als Charakteristikum der Leisten angesehen wird. Den Erwartungswert bzw. die Streuung des generellen Merkmals x nennen wir das *generelle Mittel* bzw. die *generelle Streuung* und bezeichnen diese Größen mit $\bar{x}$ bzw. $\bar{\sigma}^2$. Bilden wir eine Stichprobe aus einer generellen Gesamtheit erster Art, so liefern die zur Stichprobe vereinigten Elemente der Grundgesamtheit ein *Stichprobenmittel* $\tilde{x}$ und eine *Stichprobenstreuung* $\tilde{\sigma}^2$ für das Merkmal x. Die Größen $\tilde{x}$ und $\tilde{\sigma}$ sind wegen des zufälligen Charakters der Stichprobe Zufallsgrößen.

Eine *Grundgesamtheit zweiter Art* ist eine solche, in der gewisse Elemente eine Sonderstellung einnehmen. Sie sind durch ein Merkmal zu charakterisieren, das sich nicht zahlenmäßig ausdrücken läßt. Betrachten wir zum Beispiel die im Monat hergestellten Glühlampen eines Betriebes, so bilden diese eine Grundgesamtheit zweiter Art, sobald wir uns für die qualitätsgerechten Erzeugnisse interessieren. Die einwandfreien Glühlampen nehmen in dieser generellen Gesamtheit die oben erwähnte Sonderstellung ein. Mit p bezeichnen wir die *generelle Wahrscheinlichkeit*, das heißt die Wahrscheinlichkeit dafür, daß ein zur Grundgesamtheit zweiter Art gehöriges Element eine Sonderstellung einnimmt. Bilden wir eine Stichprobe aus einer generellen Gesamtheit zweiter Art, so liefern die zur Stichprobe vereinigten Elemente der Grundgesamtheit eine *Stichprobenhäufigkeit* h für die eine Sonderstellung einnehmenden Elemente. Diese Größe h ist gleichfalls eine Zufallsvariable, sie ändert sich von Stichprobe zu Stichprobe.

Die Parameter $\bar{x}$, $\bar{\sigma}$ und p charakterisieren offenbar die betreffende generelle Gesamtheit. Sie sind in der Praxis meistens unbekannt und müssen daher ermittelt werden. Dies ist auf dem Wege über die Stichprobenparameter $\tilde{x}$, und h möglich, die mehr oder weniger von den generellen Parametern $\bar{x}$ und p abweichen. Diese Abweichungen hängen einmal von dem Umfang der Stichprobe ab. Zu einem größeren Stichprobenumfang gehören geringere Abweichungen.

Die Abweichungen der Stichprobenparameter $\tilde{x}$ und $\tilde{\sigma}$ von den Parametern $\tilde{x}$ und $\tilde{\sigma}$ einer Grundgesamtheit erster Art werden darüber hinaus wesentlich von deren generellem Streuungsmaß $\bar{\sigma}$ beeinflußt. Ist zum Beispiel $\bar{\sigma} = 0$, das heißt, besitzen alle Elemente der generellen Gesamtheit denselben Merkmalswert x, dann liefert schon ein einziges Element — also eine Stichprobe des Umfanges 1 — eine ausreichende Information über die Grundgesamtheit. In diesem Fall sind die betrachteten Abweichungen null. Bei zunehmender Streuung $\bar{\sigma}^2$ wachsen diese Abweichungen.

Die Abweichung der Stichprobenhäufigkeit h von der Wahrscheinlichkeit p einer Grundgesamtheit zweiter Art wird neben dem Stichprobenumfang von der generellen Wahrscheinlichkeit p beeinflußt. Ist zum Beispiel $p = 0$ oder $p = 1$, das heißt, nimmt kein oder jedes Element der generellen Gesamtheit eine Sonderstellung ein, dann zeichnet schon ein einziges Element — also eine Stichprobe des Umfanges 1 — ein vollständiges Bild von der Grundgesamtheit. In diesem Fall beträgt die Abweichung zwischen h und p null. Bewegt sich die Größe p — von 0 oder 1 ausgehend — in Richtung $p = \frac{1}{2}$, dann nimmt diese Abweichung zu.

Wir formulieren die zwei Hauptaufgaben der Stichprobentheorie:

a) Von den Parametern einer Stichprobe gegebenen Umfanges ist auf die entsprechenden Parameter der generellen Gesamtheit zu schließen. Ferner ist die Genauigkeit (Zuverlässigkeit) dieses Schlusses anzugeben.

b) Der Umfang einer vorzunehmenden Stichprobe ist zu ermitteln, der für eine vorgeschriebene Genauigkeit (Zuverlässigkeit) des Schlusses von den Parametern dieser Stichprobe auf die entsprechenden generellen Parameter erforderlich ist.

In den Abschnitten 3.2 und 3.3 behandeln wir die Aufgabe a) der Stichprobentheorie. Der Aufgabenstellung b) widmen wir uns im Abschnitt 3.4. Den Abschluß dieses Kapitels bildet ein Abschnitt über einige Verfahren, die bei der Erhebung einer Stichprobe angewendet werden.

3.2. Stichproben aus Grundgesamtheiten erster Art

3.2.1. Große Stichproben

Wir betrachten eine Grundgesamtheit erster Art mit sehr großem Umfang und den Parametern $\bar{x}$ und $\bar{\sigma}$. Dieser generellen Gesamtheit entnehmen wir rein zufällig n ($n > 100$) Elemente mit den Merkmalswerten $x_1, x_2, \ldots, x_n$, die wir zu einer großen Stichprobe des Umfangs n vereinigen.

Wir beweisen nun bezüglich des Mittelwerts

$$\boxed{\tilde{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}} \tag{3.1}$$

dieser Stichprobe folgenden

Satz 3.1: Das Stichprobenmittel $\tilde{x}$ ist eine normalverteilte Zufallsvariable mit dem Erwartungswert

$$\boxed{E(\tilde{x}) = \bar{x}} \tag{3.2}$$

und der Streuung

$$\boxed{u^2 = E[(\tilde{x} - \bar{x})^2] = \frac{\bar{\sigma}^2}{n}}\,. \tag{3.3}$$

Beweis: Zunächst sind die Größen x_i ($i = 1, 2, \ldots, n$) stetige Zufallsgrößen, sie ändern sich von Stichprobe zu Stichprobe. Damit ist dann $\tilde{x}$ gemäß (3.1) auch eine stetige Zufallsvariable. Da die zugrunde gelegte generelle Gesamtheit einen sehr großen Umfang aufweist, sind die für die Bildung einer Stichprobe erforderlichen „Ziehungen" aus der Grundgesamtheit unabhängig.

Damit sind dann die Zufallsgrößen x_i ($i = 1, 2, \ldots, n$) unabhängig voneinander. Weil nun praktisch jedes Element der Grundgesamtheit als Stichprobenelement mit dem Merkmalswert x_i auftreten kann, besitzen die Zufallsvariablen x_i ($i = 1, 2, \ldots, n$) den Erwartungswert $\bar{x}$ und das Streuungsmaß $\bar{\sigma}$. Also gilt

$$\boxed{E(x_i) = \bar{x}, \quad \sigma(x_i) = \bar{\sigma} \ (i = 1, 2, \ldots, n)} \tag{3.4}$$

Beachten wir die vorstehenden Beziehungen, so ergibt sich wegen der Unabhängigkeit der x_i ($i = 1, 2, \ldots, n$) und des Satzes *2.24* aus (3.1)

$$E(\tilde{x}) = \frac{1}{n} \sum_{i=1}^{n} E(x_i) = \frac{1}{n} \sum_{i=1}^{n} \bar{x} = \bar{x}$$

und

$$u^2 = \sigma^2(\tilde{x}) = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2(x_i) = \frac{1}{n^2} \sum_{i=1}^{n} \bar{\sigma}^2 = \frac{\bar{\sigma}}{n}\,.$$

Da nun $n > 100$ besteht, folgt schließlich aus dem Zentralen Grenzwertsatz 2.36 die Tatsache, daß $\tilde{x}$ dem Normalverteilungsgesetz gehorcht. Damit ist alles gezeigt.

Wir nennen die Größe u, das heißt das Streuungsmaß des Stichprobenmittels $\tilde{x}$, das *Genauigkeitsmaß der Stichprobe.* Die abgeleitete Formel (3.3) drückt den bereits im Abschnitt 3.1 angedeuteten Zusammenhang zwischen dem Genauigkeitsmaß u, dem Stichprobenumfang n und dem generellen Streuungsmaß $\bar{\sigma}$ aus: Das Genauigkeitsmaß einer Stichprobe nimmt mit wachsender Grundgesamtheitsstreuung zu und mit wachsendem Stichprobenumfang ab. Dieser Sachverhalt rechtfertigt die Bezeichnung „Genauigkeitsmaß" für die Größe u. Wenden wir den Satz 2.35 nebst der Relation (2.98) auf die normalverteilte Zufallsgröße $\tilde{x}$ an, so ergibt sich sofort der

Satz 3.2: Es gilt die Beziehung

$$\boxed{P(|\tilde{x} - \bar{x}| < \varrho) = 2\,\Phi\left(\frac{\varrho}{u}\right) - 1}\,. \tag{3.5}$$

Die vorstehende Formel versetzt uns in die Lage, die Abweichung des Stichprobenmittels $\tilde{x}$ von dem generellen Mittel $\bar{x}$ abzuschätzen. Setzen wir in (3.5) der Reihe nach $\varrho = u$, $\varrho = 2u$ und $\varrho = 3u$ ein, so erhalten wir unter Zuhilfenahme der Φ-Tafel im Anhang die Ausdrücke

$$\left.\begin{aligned} P(|\tilde{x} - \bar{x}| < u) &= 0{,}6826\,, \\ P(|\tilde{x} - \bar{x}| < 2u) &= 0{,}9544\,, \\ P(|\tilde{x} - \bar{x}| < 3u) &= 0{,}9972\,. \end{aligned}\right\} \tag{3.6}$$

Diese Gleichungen verdeutlichen die Aussage der Formel (3.5). Sie lassen klar erkennen, in welcher Weise das zufällige Stichprobenmittel $\tilde{x}$ um den konstanten Wert $\bar{x}$ des generellen Mittels schwankt. So sagt die letzte dieser drei Gleichungen insbesondere aus: Mit einer Wahrscheinlichkeit von 99,72 Prozent weicht das Stichprobenmittel $\tilde{x}$ um weniger als $3u$ vom Grundgesamtheitsmittel $\bar{x}$ ab. Haben wir also von einer vorgenommenen Stichprobe den Mittelwert $\bar{x}$ und das Genauigkeitsmaß u ermittelt, so können wir mit einer Wahrscheinlichkeit von 99,72 Prozent garantieren, daß der gesuchte Erwartungswert $\bar{x}$ des generellen Merkmals x weniger als $3u$ von unserem Wert $\tilde{x}$ abweicht.

Wir betrachten nun ein Beispiel, das die letzten Ausführungen unterstreicht. Gegeben ist eine sehr große Gesamtheit von Zuckerpackungen, deren Reingewicht ein Streuungsmaß von $\bar{\sigma} = 10$ Gramm aufweist. Gesucht ist der Erwartungswert $\bar{x}$ für das Reingewicht dieser Packungen. Zu diesem Zweck wird der Gesamtheit eine Stichprobe von 200 Packungen entnommen. Das Stichprobenmittel $\tilde{x}$ beläuft sich auf 500 Gramm. Wie groß ist das Streuungsmaß des Mittels $\tilde{x}$, d. h. wie groß ist das Genauigkeitsmaß dieser Stichprobe? Welches symmetrisch um das gefundene Stichprobenmittel liegende Intervall kann mit einer Wahrscheinlichkeit von 98 Prozent für das generelle Reingewicht der Zuckerpackungen garantiert werden? Wie lautet das Intervall, wenn nur eine Garantie von 90 Prozent gegeben wird?

Es liegen in diesem Beispiel die Werte $n = 200$, $\bar{\sigma} = 10$ Gramm und $\tilde{x} = 500$ Gramm vor. Aus (3.3) berechnen wir das Genauigkeitsmaß der Stichprobe zu

$$u = \sqrt{\frac{100}{200}} = \sqrt{\frac{1}{2}} \approx 0{,}71\,,$$

das somit annähernd 0,71 Gramm beträgt. Bezeichnen wir das erste gesuchte Intervall mit $|\tilde{x} - \bar{x}| < \varrho_1$, so folgt aus (3.5)

$$P\,(\,|\tilde{x} - \bar{x}\,| < \varrho_1) = 2\,\Phi\left(\frac{\varrho_1}{0{,}71}\right) - 1 = 0{,}98$$

oder

$$\Phi\left(\frac{\varrho_1}{0{,}71}\right) = 0{,}99\,.$$

Unter Zuhilfenahme der Φ-Tafel erhalten wir

$$\varrho_1 = 1{,}65.$$

Damit gilt die folgende Schätzung für den Wert des generellen Mittelwerts $\bar{x}$:

$$P\,(498{,}35 < \bar{x} < 501{,}65) = 0{,}98\,.$$

Ganz entsprechend leiten wir für das zweite gesuchte Intervall $|\tilde{x} - \bar{x}| < \varrho_2$ das Resultat

$$\varrho_2 = 1{,}18$$

her, das die Abschätzung

$$P\,(498{,}82 < x < 501{,}18) = 0{,}90$$

sicherstellt.

Das vorstehende Beispiel macht deutlich, wie wir die Aussagen über das generelle Mittel $\bar{x}$ gewinnen können. Als Schätzwert für $\bar{x}$ dient praktisch das Stichprobenmittel $\bar{x}$, dessen Erwartungswert nach Satz *3.1* mit $\bar{x}$ übereinstimmt. Wir betrachten nun die Streuung $\tilde{\sigma}^2$ der Stichprobe, das heißt die Streuung der Merkmalswerte $x_1, \ldots, x_n$ um den Mittelwert $\tilde{x}$ der Stichprobe und beweisen den

Satz 3.3: Die Stichprobenstreuung $\tilde{\sigma}^2$ ist eine stetige Zufallsgröße mit dem Erwartungswert

$$\boxed{E(\tilde{\sigma}^2) = \left(1 - \frac{1}{n}\right)\bar{\sigma}^2} \, . \tag{3.7}$$

Beweis: Da die Größen x_i ($i = 1, 2, \ldots, n$) stetige Zufallsvariable sind, ist $\tilde{\sigma}^2$ zunächst eine kontinuierliche Zufallsveränderliche. Bezeichnen wir mit $\widetilde{x^2}$ den Mittelwert der Quadrate der Zahlen $x_1, \ldots, x_n$, das heißt setzen wir

$$\widetilde{x^2} = \frac{1}{n}(x_1{}^2 + x_2{}^2 + \ldots + x_n{}^2),$$

so folgt aus Satz 2.24 und Gleichung (2.17)

$$\tilde{\sigma}^2 = \widetilde{x^2} - \tilde{x}^2 .$$

Bilden wir nun den Erwartungswert des letzten Ausdrucks, so wird

$$E(\tilde{\sigma}^2) = E(\widetilde{x^2}) - E(\tilde{x}^2) .$$

oder

$$E(\tilde{\sigma}^2) = E(\widetilde{x^2}) - E\left[\left\{(\tilde{x} - \bar{x}) + \bar{x}\right\}^2\right].$$

Quadrieren wir die geschweifte Klammer aus, so ergibt sich weiter

$$E(\tilde{\sigma}^2) = E(\widetilde{x^2}) - E\left[(\tilde{x} - \bar{x})^2\right] - 2\,\bar{x}\,E(\tilde{x} - \bar{x}) - \bar{x}^2$$

oder auf Grund der Eigenschaften der Zufallsgröße $\tilde{x}$ (die Beziehung $E(\widetilde{x^2}) = \overline{x^2}$ läßt sich analog wie die Formel (3.2) des Satzes 3.1 herleiten!).

$$E(\tilde{\sigma}^2) = \overline{x^2} - \bar{x}^2 - u^2 .$$

Nun ist aber — wieder als Folge von (2.17) —

$$\bar{\sigma}^2 = \overline{x^2} - \bar{x}^2 ,$$

so daß

$$E(\tilde{\sigma}^2) = \bar{\sigma}^2 - u^2$$

wird. Beachten wir (3.3), so ist die Behauptung (3.7) perfekt. Der Erwartungswert der Stichprobenstreuung $\tilde{\sigma}^2$ ist also nicht gleich der generellen Streuung $\bar{\sigma}^2$, sondern wegen $1 - \frac{1}{n} < 1$ stets kleiner als $\bar{\sigma}^2$.

Aus diesem Grunde werden wir nicht $\tilde{\sigma}^2$, sondern die Größe

$$\boxed{\sigma_0{}^2 = \frac{n}{n-1}\,\tilde{\sigma}^2} \tag{3.8}$$

als *Schätzwert für die generelle Streuung* $\bar{\sigma}^2$ ansehen. Es gilt nämlich der

Satz 3.4: Die in (3.8) eingeführte Größe $\bar{\sigma}_0{}^2$ ist eine kontinuierliche Zufallsvariable, deren Erwartungswert gleich der generellen Streuung $\bar{\sigma}^2$ ist:

$$\boxed{E\,(\sigma_0{}^2) = \bar{\sigma}^2} \,. \tag{3.9}$$

Beweis: Der erste Teil der Behauptung ist unmittelbar klar. Der zweite Teil resultiert aus (3.8) und (3.7) zu

$$E\,(\sigma_0{}^2) = E\left(\frac{n}{n-1}\,\tilde{\sigma}^2\right) = \frac{n}{n-1}\left(1 - \frac{1}{n}\right)\bar{\sigma}^2 = \bar{\sigma}^2 \,.$$

Den Schätzwert $\bar{\sigma}_0{}^2$ für die generelle Streuung $\bar{\sigma}^2$ können wir gemäß (3.8) für alle Stichproben bilden, die mindestens aus zwei Elementen bestehen, deren Umfang n also größer als 1 ist. Für $n = 1$ verliert die Formel (3.8) ihre Gültigkeit. Den Fall einer Stichprobe mit nur einem Element müssen wir daher ausschließen, sobald unsere Überlegungen auf die Größe $\sigma_0{}^2$ Bezug nehmen.

Wegen $\dfrac{n}{n-1} > 1$ ist der Schätzwert σ_0^2 für die Streuung $\bar{\sigma}^2$ der Grundgesamtheit immer größer als die Stichprobenstreuung $\tilde{\sigma}^2$. Bei großen Stichproben ist der Unterschied zwischen diesen beiden Zufallsgrößen äußerst gering. Haben wir jedoch eine kleine Stichprobe — zum Beispiel mit $n = 20$ — vor uns, so ist die Abweichung beider Größen voneinander doch erheblich.

Wir haben im Satz 3.1 festgestellt, daß das Genauigkeitsmaß u einer Stichprobe vom generellen Streuungsmaß $\bar{\sigma}$ abhängt. Diese Abhängigkeit findet ihren Ausdruck in der Beziehung (3.3). Da nun aber die generelle Streuung $\bar{\sigma}^2$ in der Praxis meistens unbekannt ist, müssen wir uns mit einem Näherungswert für $\bar{\sigma}^2$ begnügen. Als einen solchen Näherungswert für $\bar{\sigma}^2$ betrachten wir die eingeführte Größe $\sigma_0{}^2$, deren Erwartungswert mit $\bar{\sigma}^2$ übereinstimmt. Ersetzen wir nun $\bar{\sigma}^2$ in (3.3) durch ihren Schätzwert $\sigma_0{}^2$, so bekommen wir für das Genauigkeitsmaß u der Stichprobe einen Näherungswert u_0, den wir *Schätzwert für das Genauigkeitsmaß* nennen.

Satz 3.5: Für den Schätzwert u_0 des Genauigkeitsmaßes u einer Stichprobe gilt

$$\boxed{u_0 = \frac{\tilde{\sigma}}{\sqrt{n-1}}} \,. \tag{3.10}$$

Beweis: Aus (3.3) und (3.8) folgt sofort

$$u^2 = \frac{\bar{\sigma}^2}{n} \approx \frac{1}{n}\,\frac{n}{n-1}\,\tilde{\sigma}^2 = \frac{\tilde{\sigma}^2}{n-1} = u_0{}^2$$

oder die Behauptung (3.10).

Die Schätzung von u durch u_0 ist bei unbekannter genereller Streuung $\bar{\sigma}^2$ erforderlich, um überhaupt Aussagen über die Grundgesamtheit aus der vorgenommenen Stichprobe zu gewinnen. Diese Schätzung ist um so genauer, je größer der Umfang n der betreffenden Stichprobe ist. Bei kleineren Stichproben wird diese Annäherung von u durch u_0 dagegen grob sein.

Wir behandeln nun zwei Beispiele:

a) Aus einer sehr großen Gesamtheit von Erzeugnissen werden 170 Erzeugnisse ausgewählt und auf ihre Länge x (= generelles Merkmal) untersucht. Diese 170 Stichprobenelemente liefern den Mittelwert $\tilde{x} = 10$ cm und das Streuungsmaß $\tilde{\sigma} = 0{,}4$ cm. Welche Aussagen können wir bezüglich der mittleren Länge $\bar{x}$ der Gesamtheit aller Erzeugnisse treffen? Wir schätzen zunächst das Genauigkeitsmaß u dieser Stichprobe vermöge (3.10) zu

$$u \approx u_0 = \frac{0{,}4}{\sqrt{169}} \approx 0{,}031\,,$$

das heißt zu 0,31 mm. Für das gesuchte Mittel $\bar{x}$ ergibt sich dann aus (3.5) die Aussage

$$P(|x - 10| < \varrho) = 2\,\Phi\left(\frac{\varrho}{0{,}031}\right) - 1\,.$$

Setzen wir beispielsweise $\varrho = 3\,u_0 = 0{,}093$ cm, dann erhalten wir wegen (3.6)

$$P\,(9{,}907 < \bar{x} < 10{,}093) = 0{,}9972\,.$$

oder das Ergebnis: In 99,72 Prozent aller Fälle weicht der Erwartungswert $\bar{x}$ der Erzeugnislängen der Grundgesamtheit um weniger als 0,093 cm $\approx$ 1 mm vom Stichprobenmittel $\tilde{x} = 10$ cm ab.

b) In einer Großstadt wird das monatliche Einkommen von 145 Familien untersucht. Diese Untersuchung ergibt einen Durchschnitt $\tilde{x} = 750$ MDN und ein Streuungsmaß $\tilde{\sigma} = 90$ MDN. In welcher Höhe bewegt sich das mittlere monatliche Familieneinkommen x dieser Großstadt?

Wir bilden entsprechend (3.10) das geschätzte Genauigkeitsmaß dieser Stichprobe

$$u_0 = \frac{90}{\sqrt{144}} = \frac{90}{12} = 7{,}5\,,$$

das mithin 7,5 MDN beträgt. Sodann folgt aus (3.5) die Beziehung

$$P(|x - 750| < \varrho) = 2\,\Phi\left(\frac{\varrho}{7{,}5}\right) - 1\,,$$

die die gestellte Frage beantwortet. Ist insbesondere $\varrho = 2u_0 = 15$ MDN, so resultiert aus (3.6)

$$P\,(735 < \bar{x} < 765) = 0{,}9544\,.$$

Mit einer Wahrscheinlichkeit von 95,44 Prozent liegt demzufolge das generelle Familieneinkommen pro Monat zwischen 735 MDN und 765 MDN

Wir haben bislang vorausgesetzt, daß die zugrunde gelegte generelle Gesamtheit erster Art einen sehr großen Umfang besitzt, damit die zur Bildung einer Stichprobe erforderlichen „Ziehungen“ aus der Grundgesamtheit unabhängig voneinander sind. In Wirklichkeit liegt jedoch auch bei einer solchen Grundgesamtheit eine gewisse Abhängigkeit dieser Ziehungen vor.

Um diesen Sachverhalt allgemein zu erörtern, betrachten wir nunmehr eine Grundgesamtheit erster Art, die aus N Elementen besteht, die also den Umfang N aufweist. Wir denken uns diese N Elemente durchnumeriert. Erhalten wir nun als Ergebnis einer Ziehung aus der generellen Gesamtheit das Element mit der Nummer i, so kann dieses Element offenbar bei darauffolgenden Ziehungen nicht mehr als Resultat erscheinen. Eine Ziehung beeinflußt damit die darauffolgenden Ziehungen, die für die Bildung einer Stichprobe vorgenommen werden. Das bedeutet aber, daß die Ziehungen voneinander abhängen. Wird jedoch jedes durch eine Ziehung erhaltene Element nach der Erfassung seines Wertes in die generelle Gesamtheit zurückgegeben, aus der dann die nächste Ziehung erfolgt usw., so sind die Ziehungen offenbar unabhängig voneinander. Derartige „Ziehungen mit Zurücklegen“ sollen uns jedoch in diesem Zusammenhang nicht weiter interessieren.

Ist der generelle Umfang N sehr groß, dann ist die Abhängigkeit der Ziehungen aus der Grundgesamtheit äußerst gering; sie kann in diesem Fall praktisch vernachlässigt werden. Bei ökonomischen Untersuchungen ist N meistens sehr groß. Die meisten Betriebe haben beispielsweise einen solch großen Produktionsumfang, daß uns viele Daten zur Verfügung stehen. Trotzdem wenden wir uns im folgenden den Grundgesamtheiten erster Art geringen Umfanges zu, da derartige Gesamtheiten mitunter auftreten und mit den Methoden der Stichprobentheorie gleichfalls beurteilt werden müssen. Wir verdeutlichen nunmehr den Einfluß der Abhängigkeit der Ziehungen aus einer generellen Gesamtheit geringen Umfangs auf die bereits erzielten Resultate. Dieser Einfluß wird sich unter anderem darin äußern, daß an den Formeln (3.3) und (3.7) gewisse Korrekturen anzubringen sind, die von N abhängen und für $N = \infty$ verschwinden. Zunächst beweisen wir den

Satz 3.6: Der Mittelwert $\tilde{x}$ einer Stichprobe des Umfangs n aus einer Grundgesamtheit des Umfangs N ist eine diskrete Zufallsgröße mit dem Erwartungswert

$$\boxed{E(\tilde{x}) = \bar{x}}\,. \tag{3.2}$$

Beweis: Die zugrunde gelegte generelle Gesamtheit möge aus N_1 Elementen mit dem Merkmalswert $x_1, \ldots, N_k$ Elementen mit dem Merkmalswert x_k bestehen. Dann gelten znuächst die Beziehungen $N = N_1 + N_2 + \ldots N_k$ und

$$\boxed{\bar{x} = \frac{1}{N} \sum_{i=1}^{k} N_i x_i}\,. \tag{3.11}$$

Wir entnehmen nun dieser Grundgesamtheit eine Stichprobe des Umfangs n, die n_1 Elemente mit dem Merkmalswert $x_1, \ldots, n_k$ Elemente mit dem

Merkmalswert x_k aufweist. Dann bestehen offenbar die Gleichungen $n = n_1 + n_2 + \ldots + n_k$ und

$$\boxed{\tilde{x} = \frac{1}{n} \sum_{i=1}^{k} n_i \, x_i} \; . \tag{3.12}$$

Das Stichprobenmittel $\tilde{x}$ ist eine Zufallsgröße; der zahlenmäßige Wert für $\tilde{x}$ ändert sich von Stichprobe zu Stichprobe. Da insgesamt $\binom{N}{n}$ verschiedene Stichproben des Umfangs n der in Rede stehenden generellen Gesamtheit entnommen werden können, repräsentiert $\tilde{x}$ eine diskrete Zufallsvariable, die damit $\binom{N}{n}$ möglicher Werte fähig ist. Unter diesen $\binom{N}{n}$ Stichproben gibt es genau

$$\prod_{i=1}^{k} \binom{N_i}{n_i} = \binom{N_1}{n_1} \cdot \binom{N_2}{n_2} \cdot \ldots \cdot \binom{N_k}{n_k}$$

Stichproben des Umfangs n, bei denen n_i Elemente den Merkmalswert x_i ($i = 1, 2, \ldots, k$) besitzen. Die Wahrscheinlichkeit für das Auftreten einer Stichprobe der letztgenannten Art beträgt demzufolge (klassische Definition der Wahrscheinlichkeit eines Ereignisses)

$$\boxed{P(n_1, n_2, \ldots, n_k \, ; n) = \frac{\prod_{i=1}^{k} \binom{N_i}{n_i}}{\binom{N}{n}}} \; . \tag{3.13}$$

Jeder Zerlegung des Umfangs n ist eine Summe von k Summanden n_i mit $0 \leqq n_i \leqq N_i$ ($i = 1, 2, \ldots, k$) entspricht eine solche Wahrscheinlichkeit (3.13) Die Summe über alle diese Wahrscheinlichkeiten muß gemäß der Vollständigkeitsrelation gleich 1 sein. Wir drücken diesen Sachverhalt durch die Formel

$$\boxed{\sum_{n_1 + n_2 + \ldots + n_k = n} P(n_1, n_2, \ldots, n_k \, ; n) = 1} \tag{3.14}$$

aus.

Wir gehen nun an die Ermittlung von $E(\tilde{x})$. Da eine zu dem Mittelwert (3.12) Anlaß gebende Stichprobe des Umfangs n mit der Wahrscheinlichkeit (3.13) vorkommt, nimmt das Stichprobenmittel $\tilde{x}$ als diskrete Zufallsvariable den konkreten Wert (3.12) mit der Wahrscheinlichkeit (3.13) an. Wenden wir nun die Formel für den Erwartungswert einer diskreten Zufallsgröße an, so ergibt sich auf Grund der obigen Ausführungen

$$\boxed{E(\tilde{x}) = \sum_{i=1}^{k} \left\{ \frac{1}{n} \sum_{n_1+n_2+\ldots+n_k=n} n_i \, P(n_1, n_2, \ldots, n_k; n) \right\} x_i} . \tag{3.15}$$

Für den Ausdruck

$$\{\cdots\} = \frac{1}{n} \sum_{n_1+n_2+\ldots+n_k=n} n_i \frac{\binom{N_1}{n_1} \cdots \binom{N_i}{n_i} \cdots \binom{N_k}{n_k}}{\binom{N}{n}}$$

können wir wegen

$$\binom{a}{b} = \frac{a}{b} \binom{a-1}{b-1}$$

schreiben

$$\{\cdots\} = \frac{N_i}{N} \sum_{n_1+n_2+\ldots+n_k=n} \frac{\binom{N_1}{n_1} \cdots \binom{N_i-1}{n_i-1} \cdots \binom{N_k}{n_k}}{\binom{N-1}{n-1}} .$$

Setzen wir nun noch $N_j^* = N_j$, $n_j^* = n_j$ für $j \neq i$ und $N_i^* = N_i - 1$, $n_i^* = n_i - 1$ sowie $N^* = N - 1$ und $n^* = n - 1$, so erhalten wir unter Verwendung der Schreibweise (3.13) für den letzten Ausdruck

$$\{\cdots\} = \frac{N_i}{N} \sum_{n_1^*+n_2^*+\ldots+n_k^*=n^*} P(n_1^*, n_2^*, \ldots, n_k^*; n^*) .$$

Wegen (3.14) hat die vorstehende Summe den Wert 1. Damit gilt das wichtige Teilergebnis

$$\{\cdots\} = \frac{N_i}{N} .$$

Berücksichtigen wir nun dieses Ergebnis in der Gleichung (3.15), so ergibt sich die Relation

$$E(\tilde{x}) = \sum_{i=1}^{k} \frac{N_i x_i}{N} = \frac{1}{N} \sum_{i=1}^{k} N_i x_i ,$$

die infolge (3.11) gerade die Behauptung $E(\tilde{x}) = \bar{x}$ darstellt. Die im Satz 3.1 hergeleitete Beziehung (3.2) gilt also auch dann, wenn der Umfang der generellen Gesamtheit beschränkt ist. Der Merkmalswert x_i eines Stichprobenelements ist nach wie vor eine Zufallsgröße mit dem generellen Mittel als Erwartungswert

$$\boxed{E(x_i) = \bar{x}} . \tag{3.4}$$

Da nun aber die x_i nicht mehr unabhängig voneinander sind, müßte im vorliegenden Fall die Aussage (3.2) auf ganz anderem Wege — als im Beweis zu Satz 3.1 — hergeleitet werden.

Da wir in diesem Abschnitt lediglich große Stichproben untersuchen, beträgt der Umfang N der Grundgesamtheit mindestens 100 (denn eine Stichprobe kann nicht umfangreicher sein als die generelle Gesamtheit!). Daher wird die Abhängigkeit zwischen den zufälligen Größen x_i gering sein. Dieser Umstand berechtigt uns, das Stichprobenmittel $\tilde{x}$ in brauchbarer Annäherung als normalverteilt anzusehen. Ist allerdings das generelle Merkmal normalverteilt, dann ist $\tilde{x}$ stets eine Zufallsvariable mit Normalverteilung.

Wir fragen nun nach dem Streuungsmaß des Stichprobenmittels $\tilde{x}$, das heißt nach dem Genauigkeitsmaß u der vorgenommenen Stichprobe. Die diesbezügliche Antwort liefert uns der

Satz 3.7: Das Genauigkeitsmaß u einer Stichprobe des Umfangs n aus einer generellen Gesamtheit des Umfangs N beträgt

$$\boxed{u = \frac{\bar{\sigma}}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}} \; . \tag{3.16}$$

Beweis: Wir treffen zunächst einige Vorbereitungen zum eigentlichen Beweis. Entnehmen wir der generellen Gesamtheit des Umfangs N ein Element mit dem Merkmalswert x_i, dann liefern die verbleibenden $N - 1$ Elemente x_j der Grundgesamtheit den Mittelwert $\bar{x}_0$ mit

$$\boxed{E(x_j) = \bar{x}_0 = \frac{N\bar{x} - x_i}{N-1}} \; ; \tag{3.17}$$

denn $N\bar{x}$ ist die Summe aller N Merkmalswerte der generellen Gesamtheit. Wir halten nun die Größe x_i fest und berechnen den Erwartungswert der Zufallsvariablen $x_i x_j$. Wegen (3.17) wird

$$E(x_i x_j) = x_i \cdot E(x_j) = \frac{1}{N-1} \cdot (N\bar{x}\, x_i - x_i^2) \, . \tag{3.18}$$

Denken wir uns nun auch x_i veränderlich, so folgt aus (3.18)

$$E(x_i x_j) = \frac{1}{N-1} E(N\bar{x}\, x_i - x_i^2) = \frac{1}{N-1} \left[N\bar{x}\, E(x_i) - E(x_i^2)\right] . \tag{3.19}$$

Auf Grund der Relationen (3.4) und der Eigenschaften des Erwartungswertes einer Zufallsvariablen gilt

$$E(x_i^2) = \overline{x^2} = \bar{x}^2 + \overline{x^2} - \bar{x}^2 = \bar{x}^2 + \bar{\sigma}^2$$

(vgl. auch den Beweis zum Satz 3.3); damit vereinfacht sich (3.19) zu

$$E(x_i x_j) = \frac{1}{N-1} (N\bar{x}^2 - \bar{x}^2 - \bar{\sigma}^2) = \bar{x}^2 = \frac{\bar{\sigma}^2}{N-1} \tag{3.20}$$

Für den Erwartungswert

$$E[(x_i - \bar{x})(x_j - \bar{x})] = E(x_{ij}) = E(x_i x_j) - \bar{x} \cdot [E(x_j) + E(x_i)] + \bar{x}^2$$

erhalten wir nunmehr unter Beachtung von (3.20) und (3.4)

$$E(x_{ij}) = \bar{x}^2 - \frac{\bar{\sigma}^2}{N-1} - \bar{x}^2 - \bar{x}^2 + \bar{x}^2 = \frac{\bar{\sigma}^2}{1-N}. \qquad (3.21)$$

Wir gehen nun zum eigentlichen Beweis der Behauptung (3.16) über. Aus der Definition des Streuungsmaßes u des Stichprobenmittels $\tilde{x}$ und der Relation (3.1) folgt

$$u^2 = E[(\tilde{x} - \bar{x})^2] = \frac{1}{n^2} E\left[\sum_{i=1}^{n} (x_i - \bar{x})^2\right].$$

Quadrieren wir die Glieder der Summe des letzten Ausdrucks aus, so bleibt

$$u^2 = \frac{1}{n^2}\left\{\sum_{i=1}^{n} E\left[(x_i - \bar{x})^2\right] + 2 \sum_{i<j=2}^{n} E\left[(x_i - \bar{x})(x_j - x)\right]\right\}. \qquad (3.22)$$

Die n Summanden der ersten Summe haben gemäß (3.4) alle den Wert $\bar{\sigma}^2$. Die zweite Summe in (3.22) besteht aus $\binom{n}{2} = \frac{n(n-1)}{2}$ Gliedern, die wegen (3.21) alle den Wert $\frac{\bar{\sigma}^2}{1-N}$ aufweisen. Mithin bedeutet (3.22)

$$u^2 = \frac{1}{n^2}\left(n\,\bar{\sigma}^2 + 2\,\frac{n(n-1)\,\bar{\sigma}^2}{2(1-N)}\right) = \frac{\bar{\sigma}^2}{n}\,\frac{N-n}{N-1}$$

oder die Behauptung (3.16).

Im Vergleich zur Formel (3.3) tritt also im Falle einer beschränkten generellen Gesamtheit zu der Streuung u^2 des Stichprobenmittels $\tilde{x}$ der Korrekturfaktor $\frac{N-n}{N-1}$ hinzu, der für $N \to \infty$ gegen 1 strebt.

Wir betrachten nun ein Beispiel, bei dem wir $\tilde{x}$ als eine — angenähert — normalverteilte Zufallsveränderliche mit dem Erwartungswert (3.2) und dem Streuungsmaß (3.16) auffassen: Eine Gesamtheit von $N = 2001$ hergestellten Präzisionskugeln wird auf ihre Beschaffenheit untersucht. Zu diesem Zweck werden $n = 151$ Kugeln ausgewählt und deren Durchmesser bestimmt. Aus diesen 151 Werten ergibt sich das Stichprobenmittel zu $\tilde{x} = 10$ mm. Ferner ist den technologischen Unterlagen für den Produktionsprozeß (Herstellung der Präzisionskugeln!) zu entnehmen, daß die 2001 Kugeln der Grundgesamtheit ein Streuungsmaß von $\bar{\sigma} = 0{,}1$ mm für ihre Durchmesser aufweisen. Wie groß ist das Genauigkeitsmaß der vorgenommenen Stichprobe des Umfangs 151? Was kann über den Mittelwert $\bar{x}$ der 2001 Präzisionskugeln ausgesagt werden?

Zunächst resultiert aus (3.16) das gesuchte Genauigkeitsmaß zu

$$u = \frac{0{,}1}{\sqrt{151}} \sqrt{\frac{1850}{2000}} \approx 0{,}0078,$$

das heißt zu $u \approx 0{,}0078$ mm. Aus (3.5) folgt dann für das generelle Mittel $\bar{x}$ der Durchmesser der Präzisionskugeln

$$P(|\bar{x} - 10| < \varrho) = 2\Phi\left(\frac{\varrho}{0{,}0078}\right) - 1\,.$$

Setzen wir einmal wieder $\varrho = 3u = 0{,}023$ mm, so können wir gemäß vorstehender Formel und (3.6) sofort aussagen: Mit einer Wahrscheinlichkeit von 99,72 Prozent liegt der generelle Mittelwert $\bar{x}$ der 2001 Kugeldurchmesser zwischen 9,977 mm und 10,023 mm.

Wenden wir anstelle von (3.16) die Formel (3.3) für die Ermittlung des Genauigkeitsmaßes der obigen Stichprobe an, das heißt, vernachlässigen wir praktisch die Abhängigkeit der für eine Stichprobenerhebung erforderlichen Ziehungen, so erhalten wir einen um 0,0003 mm größeren Wert für das Genauigkeitsmaß:

$$u = \frac{0{,}01}{\sqrt{151}} = 0{,}0081\ mm\,.$$

Arbeiten wir mit diesem u-Wert, so erweisen sich die betreffenden Aussagen als „schlechter" im Vergleich zu unseren auf $u = 0{,}0078$ basierenden Ergebnissen des vorausgegangenen Beispiels.

Das vorstehende Beispiel macht deutlich, wie wir Aussagen über das generelle Mittel $\bar{x}$ gewinnen können, wenn der Umfang der Grundgesamtheit beschränkt ist. Als Schätzwert für $\bar{x}$ dient praktisch das Stichprobenmittel $\tilde{x}$, dessen Erwartungswert nach Satz 3.6 mit $\bar{x}$ übereinstimmt.

Wir betrachten nun die Streuung $\tilde{\sigma}^2$ der Stichprobe des Umfangs n, das heißt die Streuung der Merkmalswerte $x_1, x_2, \ldots, x_n$ der zur Stichprobe vereinigten n Elemente; dabei wird vorausgesetzt, daß die generelle Gesamtheit aus N Elementen besteht. Es gilt der

Satz 3.8: Die Streuung $\tilde{\sigma}^2$ einer Stichprobe des Umfangs n aus einer generellen Gesamtheit des Umfangs N ist eine diskrete Zufallsgröße mit dem Erwartungswert

$$\boxed{E(\tilde{\sigma}^2) = \left(1 - \frac{1}{n}\right)\frac{N}{N-1}\,\bar{\sigma}^2}\,. \tag{3.23}$$

Beweis: Da nur $\binom{N}{n}$ verschiedene Stichproben des Umfangs n aus einer generellen Gesamtheit des Umfangs N möglich sind, ist $\tilde{\sigma}^2$ eine Zufallsvariable, die nur endlich vieler Werte fähig ist. Damit ist $\tilde{\sigma}^2$ eine diskrete Zufallsvariable; die zu den $\binom{N}{n}$ möglichen $\tilde{\sigma}^2$-Werten gehörigen Wahrscheinlichkeiten stehen in (3.13). Der erste Teil des Satzes ist hiermit bewiesen.

Um den zweiten Teil als richtig nachzuweisen, gehen wir von der bereits im Beweis zum Satz 3.3 benutzten Beziehung

$$\tilde{\sigma}^2 = \widetilde{x^2} - \tilde{x}^2$$

aus, die sich aus der allgemeinen Relation (2.17) ergibt und gelangen ganz entsprechend zu dem Ausdruck:

$$E(\tilde{\sigma}^2) = \bar{\sigma}^2 - u^2$$

Ersetzen wir nun u durch den Ausdruck (3.16), so ist unsere Behauptung

$$E(\tilde{\sigma}^2) = \bar{\sigma}^2\left(1 - \frac{N-1}{n(N-1)}\right) = \frac{N}{n}\,\frac{n-1}{N-1}\,\bar{\sigma}^2$$

perfekt.

Im Vergleich zur Formel (3.7) tritt also im Falle einer beschränkten Grundgesamtheit zum Erwartungswert der Stichprobenstreuung $\tilde{\sigma}^2$ der Korrekturfaktor $\left(\frac{N}{N-1}\right)$ hinzu, der für $N \to \infty$ gegen 1 strebt.*

Als Schätzwert für den generellen Mittelwert $\bar{x}$ nehmen wir bekanntlich das Stichprobenmittel, auch wenn die Grundgesamtheit einen beschränkten Umfang besitzt. Für die generelle Streuung $\bar{\sigma}^2$ nehmen wir nicht die Stichprobenstreuung $\tilde{\sigma}^2$, sondern die Größe

$$\boxed{\sigma_0^2 = \frac{n}{N}\,\frac{N-1}{n-1}\,\tilde{\sigma}^2} \tag{3.24}$$

als *Schätzwert*. Es besteht nämlich der

Satz 3.9: Die in (3.24) eingeführte Größe σ_0^2 ist eine diskrete Zufallsgröße, deren Erwartungswert gleich der Streuung der generellen Gesamtheit des Umfangs N ist:

$$\boxed{E(\sigma_0^2) = \bar{\sigma}^2}\,. \tag{3.25}$$

Beweis: Der erste Teil der Behauptung folgt unmittelbar aus Satz 3.8. Der zweite Teil resultiert aus (3.24) und (3.23) zu

$$E(\sigma_0^2) = \frac{n}{N}\,\frac{N-1}{n-1}\,E(\tilde{\sigma}^2) = \frac{n}{N}\,\frac{N-1}{n-1}\cdot\frac{N}{n}\,\frac{n-1}{N-1}\,\bar{\sigma}^2 = \bar{\sigma}^2$$

Die bezüglich des Schätzwertes (3.8) im Falle einer unbeschränkten Grundgesamtheit gemachten Voraussetzungen gelten auch bezüglich (3.24). Bei beschränktem Umfang der Grundgesamtheit tritt demzufolge zum Schätzwert σ_0^2 für die generelle Streuung $\bar{\sigma}^2$ der Korrekturfaktor $\left(\frac{N}{N-1}\right)$ hinzu, der für $N \to \infty$ gegen 1 strebt. Da dieser Faktor ständig kleiner als 1 ist, verkleinert sich der Schätzwert σ_0^2 für die Streuung $\bar{\sigma}^2$ der generellen Gesamtheit, wenn dieselbe einen beschränkten Umfang besitzt.

* Da der Korrekturfaktor $\frac{N}{N-1}$ größer ist als 1, vergrößert sich also der Erwartungswert der Stichprobenstreuung im Falle einer beschränkten Grundgesamtheit.

Ersetzen wir in dem Ausdruck (3.16) die generelle Streuung $\bar{\sigma}^2$ durch ihren Schätzwert (3.24), so bekommen wir einen *Schätzwert* u_0 für das Genauigkeitsmaß u der Stichprobe. Genauere Auskunft hierüber gibt der

Satz 3.10: Der Schätzwert für das Genauigkeitsmaß u einer Stichprobe des Umfangs n aus einer Grundgesamtheit des Umfangs N beträgt

$$\boxed{u_0 = \sqrt{\frac{N-n}{N(n-1)}}\,\tilde{\sigma}}\,. \tag{3.26}$$

Beweis: Ersetzen wir $\bar{\sigma}$ in (3.16) durch den Schätzwert σ_0 in (3.24), so resultiert in der Tat

$$u_0 = \frac{1}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\cdot\sqrt{\frac{n}{N}\,\frac{N-1}{n-1}}\,\tilde{\sigma} = \sqrt{\frac{N-n}{N(n-1)}}\,\tilde{\sigma}.$$

Im Vergleich zur Formel (3.10) tritt damit im Falle einer beschränkten Grundgesamtheit zum Schätzwert u_0 für das Genauigkeitsmaß der Stichprobe der Korrekturfaktor $\frac{N-n}{N}$ hinzu, der für $N \to \infty$ gegen 1 strebt. Da dieser Faktor stets kleiner als 1 ist, verkleinert sich mithin der Schätzwert u_0 im Falle einer beschränkten generellen Gesamtheit.

Wir betrachten nun ein Beispiel: Ein Posten von $N = 1000$ Schraubenbolzen ist auf seine Qualität (Durchmesser der Bolzen) zu beurteilen. Zu diesem Zweck werden $n = 100$ Bolzen ausgewählt, die einen Mittelwert von $\tilde{x} = 30$ mm und ein Streuungsmaß von $\tilde{\sigma} = 1$ mm aufweisen. Wie groß ist das geschätzte Genauigkeitsmaß dieser Stichprobe des Umfangs 100? Welche Aussagen können über den generellen Mittelwert $\bar{x}$, d. h. über die Durchmesser der 1000 zu beurteilenden Bolzen getroffen werden?

Wegen $N = 1000$ und $n = 100$ sowie $\tilde{\sigma} = 1$ beläuft sich der Schätzwert u_0 für das gesuchte Genauigkeitsmaß u dieser Stichprobe auf

$$u_0 = \sqrt{\frac{900}{99\,000}} = 0{,}095\,,$$

das heißt auf 0,095 mm. Setzen wir voraus, daß $\tilde{x}$ normalverteilt ist — das ist ganz bestimmt der Fall, wenn das generelle Merkmal, nämlich der Bolzendurchmesser, normalverteilt ist — dann können wir sofort sagen

$$P(|\bar{x} - 30| < \varrho) = 2\,\Phi\left(\frac{\varrho}{0{,}095}\right) - 1 \tag{*}$$

(vgl. (3.5)!) oder gemäß (3.6) beispielsweise

$$\left.\begin{aligned} &P(29{,}715 < \bar{x} < 30{,}285) = 0{,}9972\,,\\ &P(29{,}810 < \bar{x} < 30{,}190) = 0{,}9544\,,\\ &P(29{,}905 < \bar{x} < 30{,}095) = 0{,}6826\,. \end{aligned}\right\} \quad (**)$$

Berücksichtigen wir den Umfang $N = 1000$ in diesem Beispiel nicht, würden wir also stillschweigend annehmen, daß die generelle Gesamtheit unbeschränkten Umfang besitzt, dann erhielten wir gemäß (3.10) den Schätzwert

$$u_0 = \frac{1}{\sqrt{99}} = 0{,}1005\,,$$

das heißt, $u_0 = 0{,}1005$ mm für das Genauigkeitsmaß der vorgenommenen Stichprobe des Umfangs $n = 100$. Dann würde anstelle der Gleichung (*) folgende Relation gelten

$$P\,(\,|\,x - 30\,| < \varrho) = 2\,\Phi\left(\frac{\varrho}{0{,}1005}\right) - 1\,.$$

Sie vermittelt im Vergleich zu (**) die schlechteren Aussagen:

$$P\,(29{,}6985 < \bar{x} < 30{,}3015) = 0{,}9972\,,$$
$$P\,(29{,}7990 < \bar{x} < 30{,}2010) = 0{,}9544\,,$$
$$P\,(29{,}8995 < \bar{x} < 30{,}1005) = 0{,}6826\,.$$

Wir berechnen nun abschließend den Schätzwert σ_0 für das generelle Streuungsmaß $\bar{\sigma}$ in dem vorliegenden Beispiel. Aus (3.24) resultiert unmittelbar

$$\sigma_0 = \sqrt{\frac{100 \cdot 999}{1\,000 \cdot 99}} = \sqrt{\frac{111}{110}} \approx 1{,}0091\,.$$

Die geschätzte generelle Streuung beträgt damit 1,0091 mm.

Wir leiten nunmehr eine Schlußfolgerung aus dem Satz 3.7 ab. Zu diesem Zweck führen wir die Begriffe des absoluten und relativen Umfangs einer Stichprobe ein. Unter dem *absoluten Umfang* einer Stichprobe verstehen wir die absolute Anzahl der Stichprobenelemente, die wir stets mit n bezeichnet haben. Unser bisheriger Umfangsbegriff fällt also mit dem Begriff des absoluten Umfangs zusammen. Der *relative Umfang* einer Stichprobe ist das Verhältnis zwischen dem absoluten Umfang der Stichprobe und dem Umfang der generellen Gesamtheit, also gleich dem Verhältnis $\frac{n}{N}$. Ist die Grundgesamtheit beschränkt, dann ist der relative Umfang einer Stichprobe eine Zahl, die größer als null und höchstens gleich 1 ist. Bei unbeschränkter Grundgesamtheit ist der relative Umfang jeder Stichprobe null. In diesem Zusammenhang gilt der

Satz 3.11: Das Genauigkeitsmaß u einer Stichprobe aus einer generellen Gesamtheit großen Umfangs hängt stärker von dem absoluten als von dem relativen Umfang der Stichprobe ab.

Beweis: Ist N eine große Zahl, dann können wir die im Nenner des Wurzelausdrucks (3.16) vorkommende Zahl $N - 1$ durch N ersetzen, ohne dabei einen nennenswerten Fehler zu begehen. Es bleibt dann von (3.16) die Relation

$$\boxed{u^2 = \frac{\bar{\sigma}^2}{n}\left(1 - \frac{n}{N}\right)} \tag{3.27}$$

übrig, aus der die Behauptung erhellt.

Wir verdeutlichen den Sachverhalt des vorstehenden Satzes an einem Beispiel: Auf dem Gebiet der DDR gibt es N_R Landwirtschaftliche Produktionsgenossenschaften, von denen N_B dem Bezirk Rostock angehören. Zwecks Untersuchung eines bestimmten Merkmals der Genossenschaften, z. B. der Höhe der Arbeitseinheit, werden n_B Genossenschaften des Bezirkes Rostock zufällig herausgegriffen. Diese Bezirksstichprobe hat dann den absoluten Umfang n_B und den relativen Umfang $\frac{n_B}{N_B}$. Wir nehmen nun an, daß in den anderen Bezirken der DDR gleichfalls der Prozentsatz $\frac{n_B}{N_B}$ an Landwirtschaftlichen Produktionsgenossenschaften zufällig ausgewählt wird, um Aussagen über die Höhe der Arbeitseinheit in den Bezirken abzuleiten. Alle diese Bezirksstichproben besitzen den gleichen relativen Umfang $\frac{n_B}{N_B}$, aber verschiedene absolute Umfänge. Wir denken uns nun alle diese Bezirkstichproben zu einer Stichprobe im Republikmaßstab vereinigt, die n_R Elemente umfassen möge. Diese Republikstichprobe hat denselben relativen Umfang $\frac{n_R}{N_R} = \frac{n_B}{N_B}$ wie die Bezirkstichproben, ihr absoluter Umfang ist offenbar größer als der Umfang jeder Bezirkstichprobe. Wir stellen nun die Frage: Gibt die Republikstichprobe eine zuverlässigere Auskunft über die Arbeitseinheit der Genossenschaften im Republikmaßstab als die im Bezirk Rostock vorgenommene Stichprobe im Bezirkmaßstab? Um diese Frage zu beantworten, setzen wir der Einfachheit halber voraus, daß das Streuungsmaß $\bar{\sigma}$ der Zufallsvariablen „Arbeitseinheit" in allen Bezirken der DDR ein und denselben Wert besitzt. Bezeichnen wir ferner das Genauigkeitsmaß der „kleinen" Stichprobe mit u_B und der „großen" Stichprobe mit u_R, dann folgt aus (3.27)

$$u^2_B = \frac{\bar{\sigma}^2}{n_B}\left(1 - \frac{n_B}{N_B}\right), \quad u^2_R = \frac{\bar{\sigma}^2}{n_R}\left(1 - \frac{n_R}{N_R}\right).$$

Wegen $\frac{n_B}{N_B} = \frac{n_R}{N_R}$ und $n_B < n_R$ besteht die Abschätzung

$$u^2_B = \frac{\bar{\sigma}^2}{n_B}\left(1 - \frac{n_R}{N_R}\right) > \frac{\bar{\sigma}^2}{n_R}\left(1 - \frac{n_R}{N_R}\right) = u^2_R .$$

Es gilt demnach $u_B > u_R$. Der Wert u_B der „kleinen" Stichprobe ist also größer als der Wert u_R der „großen" Stichprobe. Das hat zur Folge, daß das Stichprobenmittel im ersten Falle stärker um das generelle Mittel schwankt als im zweiten Fall. Da das Stichprobenmittel als Schätzwert für das generelle

Mittel — hier für die mittlere Höhe der Arbeitseinheit — dient, bedeutet dieser Umstand eine positive Beantwortung der oben gestellten Frage.

Um in diesem Zusammenhang ein konkretes Zahlenbeispiel vor Augen zu haben, betrachten wir die Werte

$$N_R = 10000 \qquad n_R = 1000$$
$$N_B = 1000 \qquad n_B = 100$$

und erhalten aus (32.7)

$$u^2{}_B = \frac{\bar{\sigma}^2}{100}\left(1 - \frac{1}{10}\right) = \frac{9\bar{\sigma}^2}{1000}$$

und

$$u^2{}_R = \frac{\bar{\sigma}^2}{1000}\left(1 - \frac{1}{10}\right) = \frac{9\bar{\sigma}^2}{10000} .$$

In diesem Zahlenbeispiel besteht speziell der Zusammenhang

$$u_B = \sqrt{10}\, u_R ,$$

3.2.2. Kleine Stichproben

Wir beschäftigen uns in diesem Abschnitt mit kleinen Stichproben, das heißt mit Stichproben, die weniger als 100 Elemente umfassen. Derartige Stichproben treten in der Praxis insbesondere dort auf, wo die Erhebung einer Stichprobe mit großem Zeit- und Kostenaufwand verbunden ist.

Wir bezeichnen abermals mit $\tilde{x}$ das Stichprobenmittel, mit $\bar{x}$ das generelle Mittel, mit $\tilde{\sigma}$ das Streuungsmaß der Stichprobe und mit $\bar{\sigma}$ das generelle Streuungsmaß. Das Stichprobenmittel

$$\tilde{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

ist wieder eine Zufallsgröße. Setzen wir eine unbeschränkte Grundgesamtheit mit stetigem Merkmal voraus, dann ist x eine stetige Zufallsvariable. Die Merkmale x_i ($i = 1, 2, \ldots, n$) der zur Stichprobe vereinigten n Elemente sind dann wieder unabhängig voneinander.

Das Stichprobenmittel $\tilde{x}$ braucht jedoch — im Gegensatz zu einer großen Stichprobe — nicht normalverteilt zu sein. Da wir nämlich kleine Stichproben vor uns haben, ist der Zentrale Grenzwertsatz auf das Stichprobenmittel $\tilde{x}$ nicht mehr anwendbar. Gehorcht allerdings das generelle Merkmal einer Normalverteilung — das setzen wir im weiteren voraus —, dann ist natürlich auch $\tilde{x}$ normalverteilt.

Satz 3.12: Ist das generelle Merkmal normalverteilt, dann repräsentiert das Stichprobenmittel $\tilde{x}$ eine normalverteilte Zufallsvariable mit dem Erwartungswert

$$\boxed{E(\tilde{x}) = \bar{x}} \tag{3.2}$$

und dem Streuungsmaß (Genauigkeitsmaß)

$$\boxed{u = \frac{\bar{\sigma}}{\sqrt{n}}}\,. \tag{3.3}$$

Auf den Nachweis dieser Aussage verzichten wir, da er im Prinzip wie der Beweis des Satzes 3.1 verläuft.
Wir betrachten nun die im standardisierten Sinne normalverteilte Zufallsgröße

$$\boxed{z = \frac{\tilde{x} - \bar{x}}{u}}\,. \tag{3.28}$$

Da das Genauigkeitsmaß u vermöge (3.3) mit dem meist unbekannten generellen Streuungsmaß $\bar{\sigma}$ zusammenhängt, sind wir auch bei kleinen Stichproben gezwungen, mit einem Schätzwert für u auszukommen. Der bereits hergeleitete Schätzwert

$$\boxed{u_0 = \frac{\tilde{\sigma}}{\sqrt{n-1}}} \tag{3.10}$$

nähert u jedoch sehr grob an, da der Umfang der kleinen Stichproben gering ist (vgl. die betreffenden Ausführungen über u_0 im Abschnitt 3.21). Demzufolge approximiert die gleichfalls stetige Zufallsvariable

$$\boxed{z_0 = \frac{\tilde{x} - \bar{x}}{u_0}} \tag{3.29}$$

die im standardisierten Sinne normalverteilte Größe (3.28) sehr grob. Es ist daher nicht angebracht, die Zufallsvariable z_0 als angenähert normalverteilt zu betrachten. Über die tatsächliche Verteilung der Zufallsveränderlichen z_0 gibt uns Auskunft der

Satz 3.13: Die Zufallsgröße z_0 in (3.29) besitzt die STUDENTsche Dichtefunktion

$$\boxed{s(z_0; n) = B_n \left(1 + \frac{z_0^2}{n-1}\right)^{-\frac{n}{2}}} \tag{3.30}$$

mit

$$\boxed{B_n = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\;\Gamma\left(\frac{n-1}{2}\right)}} \tag{3.31}$$

als einer von n abhängigen Konstanten.[1] Für $n \to \infty$ strebt die STUDENTsche Dichtefunktion (3.30) gegen die Dichtefunktion der standardisierten Normalverteilung

$$\boxed{\lim_{n\to\infty} S(z_0; n) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z_0^2}{2}}} . \tag{3.32}$$

Beweis: Auf den Nachweis der ersten Aussage verzichten wir, da er in diesem Rahmen zu weit führt. Den zweiten Teil der Behauptung machen wir plausibel. Bezüglich eines exakten Beweises dieses Satzes verweisen wir auf die einschlägige Literatur, z. B. [4]. Wir verdeutlichen in diesem Zusammenhang lediglich, daß (3.30) für $n \to \infty$ gegen die standardisierte Normalverteilungsdichte strebt. Zunächst können wir

$$\lim_{n\to\infty}\left(1 + \frac{z_0^2}{n-1}\right)^{-\frac{n}{2}} = \frac{1}{\sqrt{\lim\limits_{n\to\infty}\left(1 + \frac{z_0^2}{n}\right)^n}}$$

schreiben. Unter Verwendung der schon bei der POISSONschen Verteilung angewandten Relation

$$\lim_{m\to\infty}\left(1 + \frac{a}{m}\right)^m = e^a$$

folgt dann

$$\lim_{n\to\infty}\left(1 + \frac{z_0^2}{n-1}\right)^{-\frac{n}{2}} = e^{-\frac{z_0^2}{2}} .$$

Die in (3.30) vorkommende Konstante B_n strebt für $n \to \infty$ gegen eine gewisse Zahl B, so daß insgesamt gilt

$$\lim_{n\to\infty} [S(z_0; n)] = B \cdot e^{-\frac{z_0^2}{2}} .$$

Wenden wir schließlich auf die rechte Seite der letzten Gleichung die Vollständigkeitsrelation für stetige Dichtefunktionen an, dann ergibt sich aufgrund der im Kapitel 2 angestellten Betrachtungen $B = \frac{1}{\sqrt{2\pi}}$. Damit ist die Beziehung (3.32) bewiesen.

Ist also der Umfang n der Stichprobe bekannt, dann haben wir die Verteilungsdichte für die Zufallsgröße z_0 vermöge (3.30) in der Hand. Damit lassen sich dann Aussagen über das Verhalten dieser Zufallsvariablen herleiten, die wegen

$$\tilde{x} = z_0 \cdot u_0 + \bar{x}$$

[1] Γ ist die *Euler*sche oder *Gammafunktion*. Für natürliche Argumente m gilt $\Gamma(m) = (m-1)!$. Vgl. [5].

eine Charakterisierung des Stichprobenmittels $\tilde{x}$ zulassen. Mit dieser Problematik befaßt sich der

Satz 3.14: Es besteht die Beziehung

$$\boxed{P(|\tilde{x} - \bar{x}| < \varrho) = \Psi(\xi_0; n)} \tag{3.33}$$

mit

$$\boxed{\Psi(\xi_0; n) = \int_{-\xi_0}^{+\xi_0} S(t; n)\, dt} \tag{3.34}$$

und

$$\boxed{\xi_0 = \frac{\varrho \sqrt{n-1}}{\tilde{\sigma}}}\,. \tag{3.35}$$

Beweis: Zunächst folgt aus (3.29) und (3.10)

$$P(|\tilde{x} - \bar{x}| < \varrho) = P(|u_0 z_0| < \varrho) = P(|z_0| < \xi_0)$$

mit dem in (3.35) stehenden Wert für ξ_0. Da nun aber z_0 die in (3.30) aufgeschriebene Dichtefunktion besitzt, folgt aus den allgemeinen Darlegungen über stetige Verteilungen

$$P(|z_0| < \xi_0) = P(-\xi_0 < z < \xi_0) = \int_{-\xi_0}^{\xi_0} S(t; n)\, dt\,.$$

Das ist aber wegen (3.34) gerade die Behauptung (3.33).

Im Anhang geben wir eine Tabelle für die in (3.34) eingeführte Ψ-Funktion, um den vorstehenden Satz praktisch anwenden zu können (vgl. die folgenden Beispiele!). Diese Tabelle besitzt einen doppelten Eingang, einen für den Stichprobenumfang n und einen für die Größe ξ_0. Haben wir also ein ganz bestimmtes Paar von Zahlen für n und ξ_0 im Auge, dann finden wir den betreffenden Ψ-Wert in der erwähnten Tafel in der Spalte n und in der Zeile ξ_0.

Die Funktion $\Psi(\xi_0; n)$ spielt offenbar im Rahmen der Studentschen Verteilung diejenige Rolle, die der Funktion $\Theta(\xi)$ bei der standardisierten Normalverteilung zukommt. Aufgrund des Satzes 3.13 gilt sogar

$$\boxed{\lim_{n \to \infty} \Psi(\xi_0; n) = \Theta(\xi_0)}\,. \tag{3.36}$$

Die in der letzten Spalte der Ψ-Tafel stehenden Werte sind daher genau die entsprechenden Werte $\Theta(\xi_0)$.

Wir behandeln nun zwei Beispiele.

1. In einem Kunstfaserwerk werden an 6 Fäden Zerreißproben ausgeführt. Die einzelnen Fäden reißen bei den in Kilopond gemessenen Belastungen

$$\begin{aligned} x_1 &= 1{,}90, & x_4 &= 2{,}50, \\ x_2 &= 2{,}30, & x_5 &= 2{,}00, \\ x_3 &= 1{,}50, & x_6 &= 2{,}40. \end{aligned}$$

Damit besitzt die vorgenommene Stichprobe des Umfangs $n = 6$ den Mittelwert

$$\tilde{x} = \frac{1}{6}(1{,}90 + 2{,}30 + 1{,}50 + 2{,}50 + 2{,}00 + 2{,}40) = 2{,}1$$

und die Streuung

$$\tilde{\sigma}^2 = \frac{1}{6}(0{,}2^2 + 0{,}2^2 + 0{,}6^2 + 0{,}4^2 + 0{,}1^2 + 0{,}3^2) = 0{,}117$$

sowie das Streuungsmaß

$$\tilde{\sigma} = 0{,}342\,.$$

Die Verteilung der zum Zerreißen eines Fadens erforderlichen Belastung x fassen wir als normal auf; diese Voraussetzung dürfte in der Praxis erfüllt sein. Wir fragen nun nach der Wahrscheinlichkeit dafür, daß das gesuchte generelle Mittel $\bar{x}$ (das heißt die durchschnittliche Zerreißbelastung) vom Stichprobenmittelwert $\tilde{x} = 2{,}1$ kp um weniger als 0,1 kp abweicht. Wegen $\varrho = 0{,}1$ folgt aus (3.35)

$$\xi_0 = \frac{\sqrt{5} \cdot 0{,}1}{0{,}342} = 0{,}654$$

und damit aus (3.33)

$$P\,(|x - 2{,}1| < 0{,}1) = \Psi\,(0{,}654;\,6)\,.$$

Interpolieren wir linear in der zu $n = 6$ gehörigen Zeile der Ψ-Tafel, so erhalten wir das Resultat

$$P\,(|x - 2{,}1| < 0{,}1) = 0{,}457\,.$$

Die gesuchte Wahrscheinlichkeit ist also kleiner als 50 Prozent.

2. Ein Backwarengeschäft erhält eine große Lieferung von Kekspackungen, deren Sollgewicht 200 Gramm betragen soll. Eine Untersuchung von 10 Packungen liefert die in Gramm angegebenen Werte

$$\begin{aligned} x_1 &= 195 & x_6 &= 194 \\ x_2 &= 202 & x_7 &= 199 \\ x_3 &= 201 & x_8 &= 200 \\ x_4 &= 204 & x_9 &= 203 \\ x_5 &= 198 & x_{10} &= 201, \end{aligned}$$

die den Mittelwert

$$\tilde{x} = 199{,}7\,,$$

die Streuung

$$\tilde{\sigma}^2 = 9{,}61$$

und das Streuungsmaß

$$\tilde{\sigma} = 3{,}1$$

bewirken. Mit welcher Wahrscheinlichkeit weicht das mittlere Gewicht $\bar{x}$ der Kekspackungen vom Stichprobenmittel $\tilde{x} = 199{,}7$ Gramm um weniger als 2 Gramm ab? Wegen $\varrho = 2$ resultiert aus (3.35)

$$\xi_0 = \frac{2 \cdot 3}{3{,}1} = 1{,}935$$

und damit aus (3.33)

$$P\,(|\bar{x} - 199{,}7| < 2) = \Psi\,(1{,}935;\,10)\,.$$

Interpolieren wir wiederum linear in der zu $n = 10$ gehörigen Zeile der Ψ-Tafel, so ergibt sich das gesuchte Resultat

$$P\,(|\bar{x} - 199{,}7| < 2) = 0{,}914\,.$$

Das gesuchte generelle Mittel $\bar{x}$ weicht also lediglich mit einer Wahrscheinlichkeit von 8,6 Prozent um mindestens 2 Gramm vom Stichprobenwert $\tilde{x} = 199{,}7$ Gramm ab.

Wir leiten nun noch eine Folgerung aus dem letzten Satz ab. Zu diesem Zweck bezeichnen wir mit $\Lambda\,(n;\,p_0)$ denjenigen Wert von ξ_0, der $\Psi\,(\xi_0,\,n) = p_0$ erfüllt, so daß die Identität

$$\boxed{\Psi\,[\Lambda\,(n;\,p_0)\,;\,n] = p_0} \tag{3.37}$$

besteht. $\Lambda\,(n;\,p_0)$ stellt praktisch die nach ξ_0 aufgelöste Relation $\Psi\,(\xi_0;\,n) = p_0$ dar (diese Auflösung ist übrigens eindeutig, wie die Ψ-Tafel im Anhang zeigt). Mit dieser Funktion $\Lambda\,(n;\,p_0)$ gilt nun der

Satz 3.15: Es besteht der Zusammenhang

$$\boxed{P\,(|\tilde{x} - \bar{x}| < \varrho) = p_0} \tag{3.38}$$

mit

$$\boxed{\varrho = \frac{\tilde{\sigma}\,\Lambda\,(n\,;\,p_0)}{\sqrt{n-1}}}\,. \tag{3.39}$$

Beweis: Wir gehen von der Beziehung (3.33) aus. Aus $\Psi\,(\xi_0;\,n) = p_0$ folgt gemäß obiger Vereinbarung (3.37) $\xi_0 = \Lambda\,(n;\,p_0)$, so daß (3.35) gerade (3.39) zur Folge hat. Damit ist der Satz schon bewiesen.

Die Aussage des vorstehenden Satzes besteht — anders ausgedrückt — in folgendem: Geben wir uns eine bestimmte Wahrscheinlichkeit p_0 vor, dann können wir vermöge (3.38) und (3.39) ohne große Mühe dasjenige symmetrisch um den Stichprobenmittelwert $\tilde{x}$ liegende Intervall $(\tilde{x} - \varrho, \tilde{x} + \varrho)$ angeben, in dem das gesuchte generelle Mittel mit der Wahrscheinlichkeit p_0 anzutreffen ist. Bei der Lösung solcher in der Praxis häufig auftretender Aufgaben spielt die Funktion $\Lambda\,(n;\,p_0)$ eine wesentliche Rolle. Daher geben wir im Anhang eine Tabelle für diese Funktion mit den vier — für praktische Überlegungen und Anwendungen wichtigsten — p_0-Werten $p_0 = 0{,}90$, $p_0 = 0{,}95$, $p_0 = 0{,}99$ und $p_0 = 0{,}999$. Diese Tabelle besitzt wieder einen doppelten Eingang, einen für den Stichprobenumfang n und einen für die vorgegebene Wahrscheinlichkeit p_0. Haben wir also ein ganz bestimmtes Paar von Zahlen für n und p_0 im Auge, dann finden wir den betreffenden Λ-Wert in der erwähnten Tafel in der Spalte p_0 und in der Zeile n.

Zur Verdeutlichung des soeben erörterten Sachverhalts betrachten wir nun zwei Beispiele, die sich an die beiden letzten Beispiele anlehnen.

1. In welchem Intervall $(\tilde{x} - \varrho, \tilde{x} + \varrho)$ liegt die generelle mittlere Zerreißbelastung $\bar{x}$ für Kunstfasern mit einer Wahrscheinlichkeit von $p_0 = 95$ Prozent? Wegen $n = 6$, $\tilde{\sigma} = 0{,}342$ und $\Lambda\,(6;\,0{,}95) = 2{,}57$ folgt aus (3.39)
$$\varrho = \frac{0{,}342 \cdot 2{,}57}{\sqrt{5}} = 0{,}393\,.$$
Das gesuchte Intervall lautet mithin (1,707; 2,493).

2. Welchen Wert muß die Zahl ϱ annehmen, damit das generelle mittlere Gewicht der Kekspackungen dem Intervall $|\,\tilde{x} - \bar{x}\,| < \varrho$ in 99,9 Prozent aller Fälle angehört? Beachten wir die Parameter $n = 10$, $\tilde{\sigma} = 2{,}92$ und $\Lambda\,(10;\,0{,}999) = 4{,}78$, so ergibt sich das Resultat
$$\sigma = \frac{3{,}1 \cdot 4{,}78}{3} = 4{,}939\,.$$

Bislang haben wir in diesem Abschnitt angenommen, daß die generelle Gesamtheit einen unbeschränkten Umfang aufweist. Umfaßt die Grundgesamtheit indes nur N Elemente, so verlaufen die Überlegungen prinzipiell analog. Es ist lediglich zu beachten, daß nunmehr (3.16) anstelle von (3.10) als geschätztes Genauigkeitsmaß der kleinen Stichprobe zu verwenden ist. Daher besteht der

Satz 3.16: Wird einer Grundgesamtheit mit normalverteiltem Merkmal des Umfangs N eine kleine Stichprobe des Umfangs n entnommen, so bleiben die Aussagen der Sätze 3.14 und 3.15 gültig, sofern die Relationen (3.35) und (3.39) durch

$$\boxed{\xi_0 = \frac{\varrho\,\sqrt{N\,(n-1)}}{\sigma\,\sqrt{N-n}}} \tag{3.40}$$

bzw.

$$\boxed{\varrho = \frac{\tilde{\sigma}\,\sqrt{N-n}\;\Lambda\,(n\,;\,p_0)}{\sqrt{N\,(n-1)}}} \tag{3.41}$$

ersetzt werden.

Wir geben abschließend zum letzten Satz ein Beispiel: In einem Betrieb werden Glühlampen hergestellt, deren Brenndauer x eine normalverteilte Zufallsgröße ist. Um einen Posten von $N = 900$ Glühlampen zu beurteilen, werden $n = 20$ Lampen ausgewählt und auf ihre Brenndauer x untersucht. Diese kleine Stichprobe des Umfangs $n = 20$ liefert den Mittelwert $\tilde{x} = 5000$ Stunden und das Streuungsmaß $\tilde{\sigma} = 1000$ Stunden. Welche Aussagen gelten daraufhin bezüglich des generellen Mittels $\bar{x}$, das heißt bezüglich der mittleren Lebensdauer der Glühlampen des betreffenden Postens? Wie groß ist die Wahrscheinlichkeit dafür, daß $\bar{x}$ zwischen 4546 Stunden und 5454 Stunden liegt? Wie groß muß die Zahl ϱ sein, damit $\bar{x}$ weniger als ϱ vom Stichprobenmittel $\tilde{x} = 5000$ Stunden mit einer Wahrscheinlichkeit von 99 Prozent abweicht?

Zunächst resultiert aus (3.40) wegen $N = 900$, $n = 20$, $\varrho = 454$ und $\tilde{\sigma} = 1000$

$$\xi_0 = \frac{454 \sqrt{900 \cdot 19}}{1000 \sqrt{880}} = 2 .$$

Dann folgt aus den Sätzen 3.17 und 3.15

$$P(|\bar{x} - \tilde{x}| < 454) = \Psi(2; 20) = 0{,}94$$

Die gesuchte Wahrscheinlichkeit beträgt mithin 94 Prozent. Für die Zahl ϱ erhalten wir aufgrund der Sätze 3.16 und 3.15, insbesondere der Relation (3.41)

$$\varrho = \frac{1000 \sqrt{880} \cdot 2{,}86}{\sqrt{900 \cdot 19}} = 649 .$$

3.3. Stichproben aus Grundgesamtheiten zweiter Art

Wir setzen zunächst voraus, daß die betrachtete Grundgesamtheit zweiter Art einen sehr großen Umfang besitzt. Darüber hinaus nehmen wir an, daß die generelle Wahrscheinlichkeit p, das heißt die Wahrscheinlichkeit dafür, daß ein beliebiges Element der Grundgesamtheit eine ausgezeichnete Stellung einnimmt, größer als $\frac{1}{2}$ ist. Diese Annahme bedeutet keine Einschränkung der Allgemeinheit. Denn ist $p < \frac{1}{2}$, dann sehen wir eben die nichtausgezeichneten Elemente der generellen Gesamtheit als ausgezeichnet an, zu denen die generelle Wahrscheinlichkeit $q = 1 - p > \frac{1}{2}$ gehört.

Dieser Grundgesamtheit entnehmen wir nun eine zufällige Stichprobe des Umfangs n. Unter den n ausgewählten Elementen seien m ausgezeichnet. Die Zahl m heißt die *absolute Häufigkeit* der ausgezeichneten Elemente der Stichprobe. Sie ist eine diskrete Zufallsgröße mit den möglichen Werten $m = 0, \ldots, m = n$. Es gilt der

Satz 3.17: Die absolute Häufigkeit m der ausgezeichneten Elemente einer Stichprobe des Umfangs n ist mit den Parametern

$$\boxed{E(m) = n \cdot p} \tag{3.43}$$

und

$$\boxed{\sigma(m) = \sqrt{np(1-p)}} \tag{8.44}$$

binomisch verteilt. Liegt eine große Stichprobe vor, so genügt m der Normalverteilung mit den Parametern (3.43) und (3.44).

Beweis: Der erste Teil der Aussage erhellt aus den Ausführungen über die BERNOULLIsche Verteilung im 2. Kapitel.

Der zweite Teil stellt eine Folgerung aus dem Grenzwertsatz von LAPLACE-MOIVRE (und der bezüglich p getroffenen Annahme) dar. Für die *relative Häufigkeit*

$$\boxed{h = \frac{m}{n}} \tag{3.45}$$

der ausgezeichneten Elemente der Stichprobe ergibt sich damit der

Satz 3.18: Die relative Häufigkeit h der ausgezeichneten Elemente einer Stichprobe des Umfangs n ist eine Zufallsgröße mit dem Erwartungswert

$$\boxed{E(h) = p} \tag{3.46}$$

und dem Streuungsmaß (Genauigkeitsmaß der Stichprobe)

$$\boxed{v = \sqrt{E[(h-p)^2]} = \sqrt{p\,\frac{(1-p)}{n}}}\,, \tag{3.47}$$

die für große Stichproben mit den Parametern (3.46) und (3.47) normalverteilt ist.

Beweis: Wir müssen nur (3.46) und (3.47) nachweisen, alle andere Aussagen resultieren aus Satz 3.17. Nun folgt aus (3.43) und (3.45)

$$E(h) = E\left(\frac{m}{n}\right) = \frac{1}{n} \cdot E(m) = \frac{1}{n} \cdot np = p$$

und aus (3.44) und (3.45)

$$v^2 = \sigma^2(h) = \sigma^2\left(\frac{m}{n}\right) = \frac{1}{n^2}\sigma^2(m) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n}.$$

Damit gilt

Satz 3.19: Für große Stichproben besteht die Beziehung

$$\boxed{P(|h-p| < \varrho) = 2\,\Phi\left(\frac{\varrho}{v}\right) - 1}\,. \tag{3.48}$$

Wir nähern die generelle Wahrscheinlichkeit p durch die relative Häufigkeit h der Stichprobe an; vgl. die Betrachtungen im Abschnitt 3.1. Die Beziehung (3.48) gestattet nun die Abschätzung dieser Approximation. Setzen wir der Reihe nach $\varrho = v$, $\varrho = 2v$ und $\varrho = 3v$, so erhalten wir die Abschätzungen

$$\left.\begin{array}{l} P(|h - p| < v) \;\; = 0{,}6826 \\ P(|h - p| < 2v) = 0{,}9544 \\ P(|h - p| < 3v) = 0{,}9972 \end{array}\right\},$$

die beispielsweise zeigen, in welcher Weise die relative Häufigkeit h um die konstante generelle Wahrscheinlichkeit p schwankt.

Wir erörtern nun ein Beispiel, daß die letzten Überlegungen verdeutlicht. In einem Betrieb sind $p = 64$ Prozent der Erzeugnisse von der Sorte I. Es wird nun eine Stichprobe von $n = 100$ dieser Erzeugnisse vorgenommen, die eine gewisse relative Häufigkeit h der zur Sorte I gehörigen Stichprobenelemente liefert. Wie groß ist das Genauigkeitsmaß dieser Stichprobe? In welchem symmetrisch um $p = 0{,}64$ liegendem Intervall ist die Häufigkeit h mit einer Wahrscheinlichkeit von 99 Prozent zu erwarten? Mit welcher Wahrscheinlichkeit gehört h dem Bereich (0,60; 0,68) an?
Wir erhalten gemäß (3.47) das Genauigkeitsmaß

$$v = \sqrt{\frac{0{,}64 \cdot 0{,}36}{100}} = 0{,}048\,.$$

Aus (3.48) resultiert weiter

$$2\Phi\left(\frac{\varrho}{0{,}048}\right) - 1 = 0{,}99$$

oder

$$\Phi\left(\frac{\varrho}{0{,}048}\right) = 0{,}995\,.$$

Wegen $\Phi\,(2{,}567) = 0{,}995$ wird schließlich

$$\varrho = 0{,}124.$$

In dem Intervall (0,516; 0,764) ist die Größe h demzufolge mit einer Wahrscheinlichkeit von 99 Prozent zu erwarten. Wir beantworten nun die letzte Frage. Wegen $\varrho = 0{,}04$ ergibt sich

$$P(|h - p| < 0{,}04) = 2\,\Phi\left(\frac{0{,}04}{0{,}048}\right) - 1$$

oder

$$P(|h - p| < 0{,}04) = 2\Phi\,(0{,}8333) - 1 = 2 \cdot 0{,}7976 - 1 = 0{,}5952\,.$$

In diesem Beispiel haben wir von der generellen Wahrscheinlichkeit p auf die Stichprobenhäufigkeit h geschlossen. In der Praxis geht es jedoch darum, von der Größe h auf die unbekannte Wahrscheinlichkeit p zu schließen. In diesem Fall sind wir genötigt, für das Streuungsmaß v einen *Schätzwert* v_0 zu verwenden, da v von dem unbekannten generellen Parameter p abhängt. Im folgenden leiten wir einen Schätzwert für v ab.

Wir denken uns die Grundgesamtheit zweiter Art durch eine Zufallsgröße x charakterisiert, die den Wert 1 annimmt, falls das betrachtete Element ein ausgezeichnetes ist, also beispielsweise normgerecht ist. Im anderen Fall wird x der Wert 0 gegeben. Durch diese Festsetzung können wir die zugrunde gelegte generelle Gesamtheit zweiter Art als eine Grundgesamtheit erster Art mit dem diskreten Merkmal x auffassen. Damit gewinnen wir Anschluß an die im vorangehenden Abschnitt angestellten Untersuchungen.

Satz 3.20: Das eingeführte generelle Merkmal x hat den Mittelwert

$$\boxed{\bar{x} = p} \tag{3.49}$$

und das Streuungsmaß

$$\boxed{\bar{\sigma} = \sqrt{p(1-p)}}\,. \tag{3.50}$$

Eine Stichprobe des Umfangs n besitzt den Mittelwert

$$\boxed{\tilde{x} = h} \tag{3.51}$$

und das Streuungsmaß

$$\boxed{\tilde{\sigma} = \sqrt{h(1-h)}}\,. \tag{3.52}$$

Für große Stichproben ist $\tilde{x}$ eine normalverteilte Zufallsgröße mit dem Erwartungswert

$$\boxed{E(\tilde{x}) = \bar{\bar{x}} = p} \tag{3.53}$$

und dem Streuungsmaß (Genauigkeitsmaß der Stichprobe!)

$$\boxed{u = \frac{\bar{\sigma}}{\sqrt{n}} = \sqrt{\frac{p(1-p)}{n}} = v}\,. \tag{3.54}$$

Die Stichprobenstreuung $\tilde{\sigma}^2$ ist eine Zufallsgröße mit dem Erwartungswert

$$\boxed{E(\tilde{\sigma}^2) = \left(1 - \frac{1}{n}\right)\bar{\sigma}^2 = \left(1 - \frac{1}{n}\right) p(1-p)}\,. \tag{3.55}$$

Beweis: Zunächst sind die Relationen (3.49) und (3.51) einleuchtend. Um (3.50) nachzuweisen, betrachten wir ein beliebiges Element der Grundgesamtheit, dessen Merkmal x der Verteilungstabelle

x-Wert	1	0
Wahrscheinlichkeit	p	$1-p$

genügt und demzufolge gemäß den Ausführungen über BERNOULLIsche Verteilungen die Streuung $p(1-p)$ aufweist. Da diese Überlegung für jedes Element der generellen Gesamtheit gilt, ist die Aussage (3.50) richtig. Ganz

entsprechend läßt sich der Ausdruck (3.52) gewinnen. Die Aussagen (3.53) bis (3.55) resultieren schließlich aus den Sätzen 3.1—3.3 und 3.18.
Aus den Sätzen 3.4—3.5 und 3.20 folgt weiter der

Satz 3.21: Der Schätzwert v_0 des Genauigkeitsmaßes einer Stichprobe des Umfangs n und der relativen Häufigkeit h berechnet sich zu

$$\boxed{v_0 = \sqrt{\frac{h(1-h)}{n-1}}}\,. \tag{3.56}$$

Von diesem hergeleiteten Schätzwert für v in (3.47) machen wir Gebrauch, wenn die generelle Wahrscheinlichkeit p unbekannt ist. Folgende zwei Beispiele erläutern diesen Sachverhalt:

1. Von 325 geprüften Erzeugnissen einer Grundgesamtheit erweisen sich 300 als normgerecht. Welche Aussagen bestehen in bezug auf die Wahrscheinlichkeit p dafür, daß ein beliebiges Erzeugnis der ins Auge gefaßten generellen Gesamtheit normgerecht ist? Wir haben $n = 325$ und $m = 300$ und daher nach (3.45)

$$h = \frac{300}{325} = \frac{12}{13} = 0.923\,.$$

   Für das Streuungsmaß der relativen Häufigkeit h, das heißt für das Genauigkeitsmaß v der vorgenommenen Stichprobe, ergibt sich gemäß (3.56) der Schätzwert

$$v_0 = \sqrt{\frac{12}{13^2 \cdot 324}} = \frac{\sqrt{3}}{117} \approx 0{,}015\,.$$

   Damit folgt dann aus (3.48) die gesuchte Aussage

$$P(\,|\,0{,}923 - p\,| < \varrho) = 2\,\Phi\left(\frac{\varrho}{0{,}015}\right) - 1\,.$$

   Wählen wir zum Beispiel $\varrho = 3v_0 = 0{,}045$, so resultiert nach (3.48)

$$P(\,|\,0{,}923 - p\,| < 0{,}045) = 0{,}9972\,.$$

   In 99,72 Prozent aller Fälle ist also die gefragte Wahrscheinlichkeit p der generellen Gesamtheit, das heißt der Anteil der normgerechten Erzeugnisse, in dem Intervall (0,878; 0,968) zu erwarten.

2. In einer Großstadt sind unter 145 Familien 60 Familien, zu denen mindestens 2 Kinder gehören. Wie läßt sich die Wahrscheinlichkeit p dafür schätzen, daß eine beliebige Familie dieser Stadt mindestens zwei Kinder besitzt? Wegen $n = 145$ und $m = 60$ wird

$$h = \frac{60}{145} = \frac{12}{29} = 0{,}414$$

   und demzufolge

$$v = \sqrt{\frac{12 \cdot 17}{29^2 \cdot 144}} = \frac{\sqrt{51}}{174} \approx 0{,}041\,.$$

Mithin folgt aus (3.48) beispielsweise

$$P(|p - 0{,}414| < 0{,}05) = 2\,\Phi\left(\frac{0{,}05}{0{,}041}\right) - 1 = 0{,}7776\,.$$

Das bedeutet: In 77,76 Prozent aller Fälle weicht die gesuchte Wahrscheinlichkeit p um weniger als 0,05 von der gefundenen Stichprobenhäufigkeit $h = 0{,}414$ ab.

Bislang haben wir vorausgesetzt, daß die Grundgesamtheit zweiter Art einen unbeschränkten Umfang besitzt. Ist das nicht der Fall, dann lassen sich analog zu den Ausführungen im Abschnitt 3.22 die bereits erhaltenen Ergebnisse entsprechend abändern. Insgesamt gilt der

Satz 3.22: Besitzt die generelle Gesamtheit zweiter Art den Umfang N, dann bleiben die Aussagen der Sätze 3.20—3.21 richtig, sofern die Relationen (3.47) und (3.54)—(3.56) der Reihe nach durch die Beziehungen

$$\boxed{v = \sqrt{\frac{p\,(1-p)\,(N-n)}{n\,(N-1)}}} \tag{3.57}$$

$$\boxed{u = \frac{\tilde{\sigma}}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = v\,,} \tag{3.58}$$

$$\boxed{E\,(\tilde{\sigma}^2) = \left(1 - \frac{1}{n}\right)\frac{N}{N-1}\cdot p\,(1-p)} \tag{3.59}$$

$$\boxed{v_0 = \sqrt{\frac{(N-n)\,h\,(1-h)}{N\quad(n-1)}}} \tag{3.60}$$

ersetzt werden.

Wir ersparen uns die Herleitung der vorstehenden Formeln, die für $N \to \infty$ wieder in (3.47), bzw. (3.54)—(3.56) übergehen; sie erfolgt wie im Abschnitt 3.22.

Ein Beispiel möge den Sachverhalt des letzten Satzes verdeutlichen: Eine Lieferung von $N = 800$ Bohrern soll daraufhin untersucht werden, mit welcher Wahrscheinlichkeit p ein beliebiger Bohrer zur Güteklasse 1 gehört. In einer Stichprobe von $n = 160$ Erzeugnisse sind $m = 120$ von der Klasse 1. Was kann über die Qualität der Lieferung ausgesagt werden? Wegen

$$h = \frac{120}{160} = \frac{3}{4} = 0{,}75$$

wird gemäß (3.60)

$$v_0 = \sqrt{\frac{3}{4^2}\cdot\frac{640}{159\cdot 800}} = \frac{1}{2}\cdot\frac{1}{\sqrt{265}} \approx 0{,}031$$

und damit

$$P(|p - 0{,}75| < \varrho) = 2\,\Phi\left(\frac{\varrho}{0{,}031}\right) - 1$$

als Folge von (3.48). Setzen wir etwa $\varrho = 0{,}02$, dann erhalten wir

$$P\left(|p - 0{,}75| < 0{,}02\right) = 2\,\Phi\left(\frac{0{,}02}{0{,}031}\right) - 1 = 0{,}4812$$

und damit das Ergebnis: Mit einer Wahrscheinlichkeit von 48,12 Prozent weicht die gesuchte generelle Wahrscheinlichkeit p von der gefundenen Stichprobenhäufigkeit $h = 0{,}75$ um weniger als 0,02 ab. Mit anderen Worten: Der Anteil der zur Qualitätsklasse 1 gehörenden Bohrer liegt mit einer Wahrscheinlichkeit von 0,4812 zwischen 73 und 77 Prozent.

Wir geben noch einen Hinweis zu kleinen Stichproben aus Grundgesamtheiten zweiter Art. In diesem Fall findet der Grenzwertsatz von LAPLACE-MOIVRE keine Anwendung, so daß die relative Häufigkeit h der Stichprobe — das heißt der Schätzwert für die generelle Wahrscheinlichkeit p — nicht mehr als normalverteilt angesehen werden kann. Dann verliert beispielsweise die Formel (3.48) ihre Berechtigung. Auf diese Problematik gehen wir jedoch in diesem Rahmen nicht weiter ein.

3.4. Die Ermittlung des Umfangs einer Stichprobe

3.4.1. Die Ermittlung bei Grundgesamtheiten erster Art

Wir betrachten zunächst wieder eine generelle Gesamtheit mit sehr großem Umfang und beweisen den

Satz 3.23: Für den Umfang n einer Stichprobe muß die Beziehung

$$\boxed{n = \frac{\xi_\varrho^2 \bar{\sigma}^2}{\varrho^2}} \tag{3.61}$$

bestehen, damit mit einer Wahrscheinlichkeit p_ϱ das generelle Mittel $\bar{x}$ von dem errechneten Stichprobenmittelwert $\tilde{x}$ um weniger als ϱ abweicht. Dabei berechnet sich ξ_ϱ aus der Gleichung

$$\boxed{\Phi(\xi_\varrho) = \frac{1 + p_\varrho}{2}} \tag{3.62}$$

Ist das generelle Merkmal nicht normalverteilt und fällt die nach (3.61) ermittelte Größe n kleiner als 100 aus, dann stellt (3.61) lediglich eine Näherung für den tatsächlich erforderlichen Umfang der kleinen Stichprobe dar.

Beweis: Wir gehen von der Voraussetzung $P(|\tilde{x} - \bar{x}| < \varrho) = p_\varrho$ aus, die wegen (3.3) und (3.5) die Relation

$$\Phi\left(\frac{\varrho\sqrt{n}}{\bar{\sigma}}\right) = \frac{1 + p_\varrho}{2}$$

zur Folge hat. Bezeichnen wir mit ξ_ϱ das Argument der Funktion $\Phi(\xi)$, für das (3.62) gilt, dann erhalten wir gerade die Formel (3.61). Der zweite Teil der Behauptung resultiert aus dem Satz 3.13.

Wir geben zur Erläuterung ein Beispiel: In einem Sägewerk werden Leisten mit einer Sollänge geschnitten. Das generelle Streuungsmaß für die Leistenlänge ist als $\bar{\sigma} = 1$ cm bekannt. Wieviel Leisten müssen nun auf ihre Länge untersucht werden, damit die mittlere Länge $\bar{x}$ der hergestellten Leisten mit einer Wahrscheinlichkeit von $p_\varrho = 95$ Prozent um weniger als $\varrho = 0{,}1$ cm von dem errechneten Stichprobenmittel $\tilde{x}$ abweicht? Aus (3.62) wird $\Phi(\xi_\varrho) = 0{,}975$ oder unter Zuhilfenahme der Φ-Tafel $\xi_\varrho = 1{,}96$ ermittelt. Gemäß (3.61) resultiert dann das Ergebnis

$$n = \frac{1{,}96^2 \cdot 1^2}{0{,}1^2} = 19{,}6^2 \approx 384\,.$$

Damit müssen 384 Leisten untersucht werden, um den gestellten Forderungen Genüge zu leisten.

In der Praxis ist meistens neben $\bar{x}$ auch $\bar{\sigma}$ unbekannt. Das erschwert die Bestimmung des Umfangs n einer vorzunehmenden Stichprobe, da ein Näherungswert für $\bar{\sigma}$ im allgemeinen nicht angegeben werden kann. (Die Stichprobe, die über $\bar{\sigma}$ stets gewisse Informationen liefert, wird ja erst nach der Bestimmung ihres Umfangs erhoben!)

Ist jedoch das generelle Merkmal x normalverteilt, dann können wir einen solchen *Schätzwert* $\bar{\sigma}_0$ *für* $\bar{\sigma}$ auf einfachem Wege herleiten

Satz 3.24: Bei normalverteiltem generellem Merkmal x besitzt das generelle Streuungsmaß den Schätzwert

$$\boxed{\bar{\sigma}_0 = \frac{x_g - x_k}{6}}\,; \tag{3.63}$$

dabei sind x_g der größte und x_k der kleinste Wert des generellen Merkmals.

Beweis: Wir gehen von der Beziehung

$$P(|x - \bar{x}| < 3\bar{\sigma}) = 0{,}9972 \approx 1$$

aus, die

$$x_k \approx \bar{x} - 3\bar{\sigma}, \quad x_g \approx \bar{x} + 3\bar{\sigma}$$

zur Folge hat, wenn x_k und x_g die im Satz 3.24 angegebene Bedeutung besitzen. Eliminieren wir aus den beiden letzten Näherungsrelationen den Parameter $\bar{x}$, dann folgt die Behauptung

$$\bar{\sigma} \approx \frac{x_g - x_k}{6} = \bar{\sigma}_0\,.$$

Ersetzen wir das generelle Streuungsmaß $\bar{\sigma}$ in (3.61) durch seinen Schätzwert $\bar{\sigma}_0$ in (3.63), so erhalten wir den

Satz 3.25: Für den Umfang n einer Stichprobe aus einer normalverteilten Grundgesamtheit muß die Relation

$$\boxed{n \approx \frac{(x_g - x_k)^2 \, \xi_\varrho^2}{36 \, \varrho^2}} \tag{3.64}$$

bestehen, damit die Bedingung $P\,(|\tilde{x} - \bar{x}| < \varrho) = p_\varrho$ erfüllt ist. Dabei haben ξ_ϱ, x_g und x_k die in den Sätzen 3.23 und 3.24 angegebene Bedeutung.

Wir geben nun zur Veranschaulichung der letzten Überlegungen ein Beispiel: Wieviel Personenkraftwagen gleichen Typs müssen auf ihren Benzinverbrauch pro 100 km getestet werden, damit der dabei errechnete Mittelwert mit einer Wahrscheinlichkeit $p_\varrho = 94$ Prozent um weniger als $\varrho = 0{,}1$ $^l/_{100}$ km von dem generellen Durchschnitt abweicht? Der niedrigste Benzinverbrauch wird mit $x_k = 8$ $^l/_{100}$ km und der höchste mit $x_g = 12$ $^l/_{100}$ km angenommen. Nach (3.62) erhalten wir

$$\Phi\,(\xi_\varrho) = \frac{1 + 0{,}94}{2} = 0{,}97\,.$$

Laut Φ-Tabelle gilt

$$\xi_\varrho = 1{,}88\,.$$

Nach (3.64) können wir nunmehr berechnen

$$n \approx \frac{(8 - 12)^2 \cdot 1{,}88^2}{36 \cdot 0{,}1^2}$$

oder

$$n \approx \frac{16 \cdot 3{,}53}{0{,}36} = 157\,.$$

Es brauchen also nur 157 Personenkraftwagen getestet zu werden, um die vorgeschriebene Genauigkeit einzuhalten. Wir beweisen nun den

Satz 3.26: Besitzt die generelle Gesamtheit den Umfang N, dann bleiben die Aussagen der Sätze 3.23 und 3.26 gültig, sofern die Formeln (3.61) und (3.64) durch

$$\boxed{n = \frac{\xi_\varrho^2 \, \bar{\sigma}^2 \, N}{\varrho^2 \, (N - 1) + \xi_\varrho^2 \, \bar{\sigma}^2}} \tag{3.65}$$

bzw.

$$\boxed{n \approx \frac{\xi_\varrho^2 \, (x_g - x_k)^2 \cdot N}{36 \, \varrho^2 \, (N - 1) + \xi_\varrho^2 \, (x_g - x_k)^2}} \tag{3.66}$$

ersetzt werden.

Beweis: Aus $P(|\tilde{x} - \bar{x}| < \varrho) = p_\varrho$ folgt unter Beachtung von (3.5) und (3.16) der Ausdruck

$$\Phi\left(\frac{\varrho\sqrt{n(N-1)}}{\bar{\sigma}\cdot\sqrt{N-n}}\right) = \frac{1+p_\varrho}{2},$$

der in Verbindung mit (3.62) gerade die Behauptung (3.65) liefert. Die Aussage (3.66) geht schließlich wegen (3.63) aus (3.65) hervor.

Wir wollen auch zu diesem Problem ein Beispiel besprechen. Dazu gehen wir von der letzten Aufgabe aus. Allerdings gelten andere Bedingungen. Es wird danach gefragt, wieviel Personenkraftwagen zu testen sind, damit der dabei errechnete Mittelwert mit einer Wahrscheinlichkeit $p_\varrho = 98$ Prozent um weniger als $\varrho = 0{,}1\ l/100$ km vom generellen Durchschnitt abweicht. Die Auswahl kann nur aus $N = 600$ Kraftwagen getroffen werden, da erst diese Menge des neuen Typs produziert worden ist. Der niedrigste Benzinverbrauch wird mit $x_k = 8\ l/100$ km und der höchste mit $x_g = 12\ l/100$ km angenommen. Wir berechnen ξ_ϱ aus

$$\Phi(\xi_\varrho) = \frac{1+0{,}98}{2} = 0{,}99.$$

Nach der Φ-Tafel gilt $\xi_\varrho \approx 2{,}32$. Damit stehen uns alle Größen zur Verfügung, um sie in (3.66) einzusetzen:

$$n \approx \frac{2{,}32^2\,(8-12)^2\cdot 600}{36\cdot 0{,}1^2\cdot(600-1) + 2{,}32^2\,(8-12)^2}$$

oder

$$n \approx \frac{5{,}38\cdot 16\cdot 600}{0{,}36\cdot 599 + 5{,}38\cdot 16} \approx 171$$

Demzufolge müssen 171 Wagen getestet werden.

3.4.2. *Die Ermittlung bei Grundgesamtheiten zweiter Art*

Wir untersuchen zunächst wieder eine generelle Gesamtheit zweiter Art mit sehr großem Umfang. In Analogie zum Satz 3.23 gilt:

Satz 3.27: Für den Umfang n einer Stichprobe muß die Beziehung

$$\boxed{n = \frac{\xi_\varrho^2\, p(1-p)}{\varrho^2}} \tag{3.67}$$

bestehen, damit $P(|h - p| < \varrho) = p_\varrho$ erfüllt ist. Dabei berechnet sich ξ_ϱ aus der Gleichung (3.62). Gilt für die nach (3.67) ermittelte Größe n die Ungleichung $n < 100$, dann stellt sie lediglich einen groben Näherungswert für den tatsächlich erforderlichen Umfang der zugehörigen kleinen Stichprobe dar.

Beweis: Wir gehen von der Voraussetzung $P(|h-p|<\varrho)=p_\varrho$ aus, die wegen (3.48) und (3.47) die Relation

$$\Phi\left(\frac{\varrho\sqrt{n}}{\sqrt{p(1-p)}}\right)=\frac{1+p_0}{2}$$

nach sich zieht. Beachten wir schließlich (3.62), dann folgt die Behauptung (3.67). Die restliche Aussage des vorstehenden Satzes resultiert aus den Überlegungen am Ende des Abschnitts 3.3.

Wir geben zu diesem Problemkreis ein Beispiel: In einem Betrieb sind 97 Prozent der Erzeugnisse normgerecht, die verpackt werden sollen. Wie groß muß das Fassungsvermögen der Packungen sein, damit zu 80 Prozent garantiert ist, daß der Anteil der in diesen Packungen vorhandenen normgerechten Erzeugnisse um weniger als $\varrho=1$ Prozent von der generellen Wahrscheinlichkeit $p=0{,}97$ abweicht? Aus $\Phi(\xi_\varrho)=\frac{1}{2}(1+0{,}80)=0{,}90$ folgt $\xi_\varrho=1{,}282$ und damit gemäß (3.67)

$$n=\frac{(1{,}282)^2\,0{,}97\cdot 0{,}03}{(0{,}01)^2}\approx 478.$$

In der Praxis ist die generelle Wahrscheinlichkeit p meistens unbekannt. In diesem Fall müssen wir die generelle Streuung $\bar{\sigma}^2=p\,(1-p)$ durch einen Schätzwert ersetzen. Dabei kann nur von Vorteil sein, wenn der betreffende Schätzwert $\bar{\sigma}_0^{\,2}$ größer als $\bar{\sigma}^2$ ausfällt, denn in diesem Fall sind die tatsächlichen Verhältnisse in der generellen Gesamtheit nicht „schlechter" als die der „Ersatzgrundgesamtheit" mit der Streuung $\bar{\sigma}_0$.

Satz 3.28: Für die generelle Streuung $\bar{\sigma}^2=p\,(1-p)$ einer Grundgesamtheit zweiter Art gilt stets die Abschätzung

$$\boxed{\bar{\sigma}^2\leqq\bar{\sigma}_0^{\,2}=\frac{1}{4}}\tag{3.68}$$

Beweis: Die generelle Streuung $\bar{\sigma}^2$ ist eine Funktion $S(p)=p\,(1-p)$ der generellen Wahrscheinlichkeit p. Diese Funktion ist wegen $0\leqq p\leqq 1$ nirgends negativ. Ihr einziges Extremum ist ein Maximum an der Stelle $p=\frac{1}{2}$; denn $S'(p)=1-2p$ verschwindet und $S''(p)=-2$ ist negativ an dieser Stelle $p=\frac{1}{2}$. Der zu $p=\frac{1}{2}$ gehörige Maximalwert der Funktion $S(p)$ hat den Wert $\frac{1}{4}$. Mithin gilt (3.68).

Nun konstatieren wir den

Satz 3.29: Der Umfang n einer Stichprobe mit

$$\boxed{n\approx\left(\frac{\xi_\varrho}{2\varrho}\right)^2}\tag{3.69}$$

erfüllt sicher die Bedingung $P(|h-p|<\varrho)=p_\varrho$. Die Größe ξ_ϱ genügt dabei der Gleichung (3.62).

Beweis: Ersetzen wir $\bar{\sigma}^2$ in (3.67) durch $\bar{\sigma}_o{}^2 \doteq \frac{1}{4}$ in (3.68), so folgt

$$n \leqq \left(\frac{\xi_\varrho}{2\varrho}\right)^2 = n_0 .$$

Der Umfang n muß also höchstens n_0 sein. Wählen wir $n = n_0$, dann ergibt sich aus den bisherigen Überlegungen über Stichproben, daß die geforderte Bedingung $P(|h - p| < \varrho) = p_\varrho$ mit Sicherheit erfüllt ist.

Beispiel: Wieviel Erzeugnisse einer Gesamtheit müssen auf ihre Qualität hin begutachtet werden, damit die relative Häufigkeit der normgerechten unter diesen mit einer Wahrscheinlichkeit von 90 Prozent um weniger als 0,05 von der gesuchten generellen Wahrscheinlichkeit p für die Normgerechtheit eines Erzeugnisses abweicht? Wir berechnen aus $\Phi(\xi_\varrho) = \frac{1 + 0{,}90}{2}$ $= 0{,}95$ den Wert $\xi_\varrho = 1{,}645$ und erhalten dann aus (3.69)

$$n = \frac{1{,}645^2}{4 \cdot 0{,}0023} = 270$$

Wir betrachten nun noch den Fall, daß die generelle Gesamtheit lediglich N Elemente umfaßt. Diesbezüglich ist maßgebend der

Satz 3.30: Besitzt die generelle Gesamtheit den Umfang N, dann bleiben die Aussagen der Sätze 3.27 und 3.29 erhalten, sofern die Formeln (3.67) und (3.69) der Reihe nach durch die Ausdrücke

$$\boxed{n = \frac{\xi_\varrho^2 N (1 - p) p}{\varrho^2 (N - 1) + \xi_\varrho^2 p (1 - p)}} \tag{3.70}$$

bzw.

$$\boxed{n \approx \frac{\xi_\varrho^2 N}{4\varrho^2 (N - 1) + \xi_\varrho^2}} \tag{3.71}$$

ersetzt werden.

Beweis: Aus $P(|h - p| < \varrho) = p_\varrho$ folgt wegen (3.48) und (3.57)

$$2\,\Phi\left(\frac{\varrho}{v}\right) - 1 = 2\,\Phi\left(\frac{\varrho\sqrt{n(N-1)}}{\sqrt{p(1-p)(N-n)}}\right) - 1 = p_\varrho$$

oder

$$\Phi\left(\frac{\varrho\sqrt{n(N-1)}}{\sqrt{p(1-p)(N-n)}}\right) = \frac{p_\varrho + 1}{2} .$$

Beachten wir (3.62), so ergibt sich aus

$$\frac{\varrho\sqrt{n(N-1)}}{\sqrt{p(1-p)(N-n)}} = \xi_\varrho$$

durch beiderseitiges Quadrieren und anschließende Auflösung nach n die Beziehung (3.70). Der erste Teil des Satzes ist damit bewiesen.

Ersetzen wir nun $p(1-p) = \bar{\sigma}^2$ in (3.70) durch den im Satz 3.28 abgeleiteten Schätzwert $\bar{\sigma}_0{}^2 = \frac{1}{4}$, dann erhalten wir gerade die Formel (3.71). Damit ist auch der zweite Teil des Satzes bewiesen.

Zum Abschluß erörtern wir folgendes Beispiel: Gegeben ist ein Posten von $N = 900$ Erzeugnissen. Wieviel Erzeugnisse dieses Postens müssen kontrolliert werden, damit die sich ergebende relative Häufigkeit der zur Sorte 1 gehörigen Erzeugnisse mit einer Wahrscheinlichkeit von 70 Prozent um weniger als 5 Prozent von der generellen Wahrscheinlichkeit für die Zugehörigkeit zur Sorte 1 abweicht? Aus $\Phi(\xi_\varrho) = 0{,}85$ folgt $\xi_\varrho = 1{,}037$ und damit gemäß (3.71)

$$n \approx \frac{(1{,}037)^2 \cdot 900}{(0{,}1)^2\, 899 + (1{,}037)^2} \approx 97\,.$$

3.5. Einige Auswahlverfahren

In diesem Abschnitt führen wir einige Verfahren an, nach denen die Erhebung von Stichproben vorgenommen werden kann. Das *herkömmliche Auswahlverfahren* besteht darin, daß der Grundgesamtheit rein zufällig der Reihe nach n Elemente entnommen werden, die insgesamt die erhobene Stichprobe ausmachen. Unsere bisherigen Überlegungen bezogen sich ausschließlich auf Stichproben, die in der herkömmlichen Weise gebildet werden.

In der Praxis wird eine Stichprobe häufig nach dem folgenden Verfahren erhoben: Die Grundgesamtheit besteht aus r *Teilgrundgesamtheiten*, die keine gemeinsamen Elemente aufweisen. Jeder Teilgrundgesamtheit wird auf dem Wege der herkömmlichen Auswahl eine *Teilstichprobe* entnommen. Alle diese Teilstrichpoben ergeben zusammen eine einzige Stichprobe aus der betreffenden generellen Gesamtheit. So zerfallen zum Beispiel die in der DDR beobachteten Niederschlagsmengen eines Monats in solche Ergebnisse, die den einzelnen Bezirken bzw. Kreisen der Republik entstammen.

Wir betrachten nun allgemein eine generelle Gesamtheit erster Art mit dem Umfang N, dem Mittelwert $\bar{x}$ und der Streuung $\bar{\sigma}^2$. Diese Gesamtheit besteht aus r elementefremden Teilgrundgesamtheiten mit dem Umfang N_j, dem Mittelwert $\bar{x}_j$ und der Streuung $\bar{\sigma}_j{}^2$ $(j = 1, 2, \ldots, r)$. Dann gilt $N = N_1 + \ldots + N_r$ und

Satz 3.31: Der Mittelwert der Grundgesamtheit ist gleich dem gewogenen Mittelwert der Mittelwerte der Teilgrundgesamtheiten mit den Umfängen der Teilgrundgesamtheiten als entsprechenden Gewichten:

$$\boxed{\bar{x} = \frac{1}{N} \sum_{j=1}^{r} \bar{x}_j N_j}\,. \tag{3.72}$$

Beweis: Der Ausdruck $\bar{x}_j N_j$ $(j = 1, 2, \ldots, r)$ stellt die Summe der Merkmalswerte der N_j Elemente der j-ten Teilgrundgesamtheit dar. Demzufolge ist $\sum_{j=1}^{r} \bar{x}_j N_j$ gerade die Summe der Merkmale der N Elemente der generellen Gesamtheit. Dividieren wird diese Summe durch den Umfang N, dann ergibt sich gerade die Behauptung (3.72)

Wir denken uns nun der j-ten Teilgrundgesamtheit eine Stichprobe des Umfangs n_j mit dem Mittelwert $\tilde{x}_j$, dem Genauigkeitsmaß u_j und der Streuung $\tilde{\sigma}_j^2$ $(j = 1, 2, \ldots, r)$ entnommen. Diese r Teilstichproben vereinigen wir zu einer Stichprobe des Umfangs $n = n_1 + \ldots + n_r$ mit dem Mittelwert $\tilde{x}$, dem Genauigkeitsmaß u und der Streuung $\tilde{\sigma}^2$.

Satz 3.32: Das Stichprobenmittel $\tilde{x}$ ist gleich dem gewogenen Mittelwert der Mittelwerte $\tilde{x}_j$ der Teilstichproben mit den Umfängen N_j $(j = 1, 2, \ldots, r)$ der Teilgrundgesamtheiten als entsprechenden Gewichten

$$\boxed{\tilde{x} = \frac{1}{N} \sum_{j=1}^{r} \tilde{x}_j N_j} \, . \tag{3.73}$$

Das Mittel $\tilde{x}$ ist eine Zufallsgröße mit dem Erwartungswert

$$\boxed{E(\tilde{x}) = \bar{x}} \tag{3.74}$$

und dem Streuungsmaß (Genauigkeitsmaß der Stichprobe)

$$\boxed{u = \frac{1}{N} \sqrt{\sum_{j=1}^{r} \frac{\bar{\sigma}_j^2}{n_j} N_j^2 \frac{N_j - n_j}{N_j - 1}}} \, . \tag{3.75}$$

Beweis: Zunächst ist $\tilde{x}$ eine Zufallsgröße, die mit einer Wahrscheinlichkeit von $\frac{N_j}{N}$ durch $\tilde{x}_j$ $(j = 1, 2, \ldots, r)$ determiniert ist, da mit einer Wahrscheinlichkeit von $\frac{N_j}{N}$ ein der Grundgesamtheit entnommenes Element der j-ten Teilgrundgesamtheit $(j = 1, 2, \ldots, r)$ entstammt. Mithin gilt (3.73). Die Behauptung (3.74) folgt aus (3.72) zu

$$E(\tilde{x}) = E\left(\frac{1}{N} \sum_{j=1}^{r} \bar{x}_j N_j\right) = \frac{1}{N} \sum_{j=2}^{r} E(\bar{x}_j) N_j = \frac{1}{N} \sum_{j=1}^{r} \bar{x}_j N_j = \bar{x} \, ;$$

denn es gilt $E(x_j) = \bar{x}_j$ $(j = 1, 2, \ldots, r)$ nach Satz 3.6. Wegen der sich aus (3.73) und (3.72) ergebenden Relation

$$\tilde{x} - \bar{x} = \frac{1}{N} \sum_{j=1}^{r} (\tilde{x}_j - \bar{x}_j) N_j$$

können wir schreiben

$$u^2 = E\left[(\tilde{x} - \bar{x})^2\right] = \frac{1}{N^2} E\left\{\left[\sum_{j=1}^{r} (x_j - \bar{x}_j) N_j\right]^2\right\}$$

oder

$$u^2 = \frac{1}{N^2}\left[\sum_{j=1}^{r} N_j^2 E\left\{(\tilde{x}_j - \bar{x}_j)^2\right\} + 2 \sum_{i<j=2}^{r} N_j N_j E\left\{(\tilde{x}_i - \bar{x}_i)(\tilde{x}_j - \bar{x}_j)\right\}\right].$$

Beachten wir nun (vgl. Satz 3.7)

$$E\left[(\tilde{x}_j - \bar{x}_j)^2\right] = u_j^2 = \frac{\bar{\sigma}_j^2}{n_j} \frac{N_j - n_j}{N_j - 1}$$

sowie die wegen der Unabhängigkeit der einzelnen Stichproben gültigen Beziehung

$$E\left[(\tilde{x}_i - \bar{x}_i)(\tilde{x}_j - \bar{x}_j)\right] = E(\tilde{x}_i - \bar{x}_i) \cdot E(\tilde{x}_j - \bar{x}_j) = 0 \cdot 0 = 0, \ (i \neq j)$$

dann bleibt

$$u^2 = \frac{1}{N^2} \sum_{j=1}^{r} \frac{\bar{\sigma}_j^2}{n_j} N_j^2 \frac{N_j - n_j}{N_j - 1}.$$

Damit ist alles gezeigt.

Wir behandeln nun ein Beispiel:

In einem Bezirk der DDR gibt es 10 Kreise mit den Einwohnerzahlen N_j ($j = 1, 2, \ldots, 10$), die für den monatlichen Pro-Kopf-Verbrauch an Zucker (in kg) die Streuungsmaße $\bar{\sigma}_j$ ($j = 1, 2, \ldots, 10$) aufweisen. Im j-ten Kreis werden n_j Personen über ihren Zuckerverbrauch im Monat befragt ($j = 1, 2, \ldots, 10$). Diese n_j Personen verbrauchen durchschnittlich pro Kopf und Monat die Zuckermenge $\tilde{x}_j$ ($j = 1, 2, \ldots, 10$). Für den Mittelwert $\tilde{x}$ des durchschnittlichen Zuckerverbrauchs pro Kopf und Monat der $n = n_1 + \ldots + n_{10}$ befragten Personen gilt dann gemäß (3.73)

$$\tilde{x} = \frac{\sum_{j=1}^{10} \tilde{x}_j N_j}{N_1 + N_2 + \cdots + N_{10}}.$$

Das Genauigkeitsmaß dieser Stichprobe des Umfangs n hat wegen (3.75) den Wert

$$u = \frac{1}{N^2} \sqrt{\sum_{j=1}^{10} \frac{\bar{\sigma}_j^2}{n_j} N_j^2 \frac{N_j - n_j}{N_j - 1}}.$$

Betrachten wir das Zahlenbeispiel

j	N_j	n_j	$\bar{\sigma}_j^2$	$\tilde{x}_j$
1	$5 \cdot 10^4$	120	0,03	0,35
4	$4 \cdot 10^4$	150	0,05	0,30
3	$3{,}5 \cdot 10^4$	200	0,04	0,37
4	$4{,}5 \cdot 10^4$	180	0,04	0,36
5	$3 \cdot 10^4$	125	0,03	0,34
6	$2{,}5 \cdot 10^4$	130	0,01	0,38
7	$2 \cdot 10^4$	250	0,02	0,35
8	$1{,}5 \cdot 10^4$	110	0,03	0,36
9	$3 \cdot 10^4$	180	0,01	0,34
10	$1 \cdot 10^4$	175	0,04	0,35

dann finden wir die konkreten Werte

$$\tilde{x} = 0{,}348$$

und $$u = 0{,}005.$$

Mit diesen Werten können wir dann über den unbekannten Mittelwert $\bar{x}$ des Pro-Kopf-Verbrauchs an Zucker in dem betrachteten Bezirk aussagen

$$P(|\tilde{x} - \bar{x}| < \varrho) = 2\,\Phi\left(\frac{\varrho}{0{,}005}\right) - 1$$

oder gemäß (3.6)

$$P(0{,}343 < \bar{x} < 0{,}353) = 0{,}6826$$
$$P(0{,}338 < \bar{x} < 0{,}358) = 0{,}9544$$
$$P(0{,}333 < \bar{x} < 0{,}363) = 0{,}9972.$$

Es ist demnach mit einer Wahrscheinlichkeit von 99,72 Prozent zu rechnen, daß der gesuchte generelle Mittelwert $\bar{x}$ zwischen 343 Gramm und 353 Gramm liegt.

Wir betrachten nun eine generelle Gesamtheit zweiter Art mit dem Umfang N und der generellen Wahrscheinlichkeit p. Diese Gesamtheit besteht aus r elementefremden Teilgrundgesamtheiten mit dem Umfang N_j und der generellen Wahrscheinlichkeit p_j $(j = 1, 2, \ldots, r)$. Wir denken uns weiter der j-ten Teilgrundgesamtheit eine Stichprobe des Umfangs n_j mit der relativen Häufigkeit h_j und dem Genauigkeitsmaß v_j $(j = 1, 2, \ldots, r)$ entnommen. Diese r Teilstichproben fassen wir als eine Stichprobe des Umfangs $n = n_1 + n_2 + \ldots + n_r$ auf, die die relative Häufigkeit h und das Genauigkeitsmaß v aufweist. In voller Analogie zu denS ätzen 3.31 und 3.32 besteht der

Satz 3.33: Die generelle Wahrscheinlichkeit der Grundgesamtheit p ist gleich dem gewogenen Mittel der generellen Wahrscheinlichkeiten der Teilgrundgesamtheiten mit den Umfängen der Teilgrundgesamtheiten als entsprechenden Gewichten

$$\boxed{p = \frac{1}{N} \sum_{j=1}^{r} p_j N_j}\,. \qquad (3.76)$$

Die relative Stichprobenhäufigkeit ist gleich dem gewogenen Mittel der relativen Häufigkeiten der Teilstichproben mit den Umfängen der Teilgrundgesamtheiten als entsprechenden Gewichten

$$\boxed{h = \frac{1}{N} \sum_{j=1}^{r} h_j N_j} \,. \tag{3.77}$$

Die Stichprobenhäufigkeit h ist eine Zufallsvariable mit dem Erwartungswert

$$\boxed{E(h) = p} \tag{3.78}$$

und dem Streuungsmaß (Genauigkeitsmaß der Stichprobe)

$$\boxed{v = \frac{1}{N} \sqrt{\sum_{j=1}^{r} \frac{p_j (1 - p_j)}{n_j} \, \frac{N_j - n_j}{N_j - 1} \, N_j^2}} \,. \tag{3.79}$$

Wir erwähnen schließlich noch ein letztes Verfahren, das in gewissem Sinne im Gegensatz zu der eben besprochenen Auswahlmethode steht. Dieses Verfahren besteht darin, daß der Grundgesamtheit rein zufällig geschlossene Elementegruppen — sogenannte *Serien* — entnommen werden. In diesem Falle liegt eine Stichprobe von Serien aus der betreffenden generellen Gesamtheit vor. Dieses Serienauswahlverfahren hat jedoch im Vergleich zu der eingangs behandelten Auswahlmethode keine besonderen Aspekte. Fassen wir nämlich die einzelnen Serien als Elemente (im weiteren Sinne) auf, so stellt das Serienauswahlverfahren gerade die herkömmliche Auswahlmethode dar.

4. Einführung in die Bedienungstheorie

4.1. Der Gegenstand der Bedienungstheorie

Die Bedienungstheorie ist eine mathematische Disziplin, die sich mit der Tätigkeit eines sogenannten *Systems der Massenbedienung* beschäftigt und sich dabei wahrscheinlichkeitstheoretischer Methoden bedient. Als Beispiel für solche Systeme der Massenbedienung (oder kurz *Bedienungssysteme*) können zunächst dienen: Warenhäuser, Telefonstationen, Dienstleistungseinrichtungen, Reparaturstationen, Auskunftbüros und Fahrkartenschalter. Weitere Beispiele aus der Praxis werden wir im weiteren noch kennenlernen.

Jedes Bedienungssystem besteht aus einer gewissen Anzahl von bedienenden Einheiten, die wir *Bedienungsstationen* oder *-aggregate* nennen. So fungieren als bedienende Einheiten die Verkäuferinnen und Verkäufer der Warenhäuser, die Verbindungslinien im Fernsprechverkehr, die Kassierer an den Fahrkartenschaltern usw. Die bedienenden Einheiten können also Personen und/oder Sacheinheiten (Automaten) sein. Darüber hinaus ist ersichtlich, daß ein Bedienungssystem eine oder mehrere Bedienungsstationen besitzt.

Die Tätigkeit eines beliebigen Bedienungssystems besteht in der Befriedigung (Erfüllung) von Forderungen, die in das System eintretende Einheiten erheben. Wir nennen im folgenden Einheiten mit Forderungen kurz *Forderungen.* Die Tätigkeit eines Warenhauses beinhaltet die Erfüllung der Wünsche der eintreffenden Kunden; hier sind also die zu bedienenden Kunden die eintreffenden Forderungen. Die Tätigkeit eines Verpackungsautomaten hat die Verpackung von Waren zur Aufgabe; hier stellen also die zu verpackenden Waren die eintreffenden Forderungen dar.

Die Forderungen treffen nacheinander in gewissen zufälligen Zeitmomenten in das Bedienungssystem ein. Die Bedienung oder Abfertigung einer Forderung durch eine Bedienungsstation dauert eine gewisse Zeit. Nach Ablauf dieser Bedienungszeit, die im allgemeinen von Forderung zu Forderung verschieden ist, wird die Station frei und damit bereit für die Bedienung einer nächsten Forderung.

Die bisher erwähnten Beispiele für praktische Bedienungssysteme lassen erkennen, daß neben der Qualität der Bedienung die Organisation der Bedienung eine große Bedeutung besitzt. Diese (organisatorische) Seite eines Bedienungsprozesses kann durch verschiedene Kennziffern ausgedrückt werden, von denen wir als Beispiele anführen: die *Wartezeit der eingetroffenen Forderung* bis zum Bedienungsbeginn, die *Anzahl der tätigen Bedienungsstationen*, die *Länge der Warteschlange*, die *mittlere Leerlaufzeit* (d. h. den

Auslastungsgrad) der Bedienungsstationen und den *Prozentsatz der nicht bedienten Forderungen*, die das System unbedient verlassen, denn nicht immer kann eine Forderung befriedigt — bedient — werden; das ist beispielsweise der Fall, wenn in einer Reparaturstation die für die Reparatur eines defekten Gerätes (Forderung) notwendigen Einzelteile fehlen oder die Reparaturstation überfüllt ist!). Alle diese Kennziffern hängen von den verschiedensten Bedingungen ab, die im einzelnen gar nicht immer berücksichtigt werden können. So hängt zum Beispiel die Länge der Warteschlange an einem Fahrkartenschalter von den Erfahrungen des Kassierers ab, von der Art der Operationen, die er beim Ausfüllen der Fahrkarten auszuführen hat, oder der Anzahl der Züge, für die an diesem Schalter gleichzeitig Karten verkauft werden usw. Aber alle diese Kennziffern charakterisieren den Bedienungsprozeß. Sie geben ein Bild davon, in welchem Maße das Bedienungssystem der Gesamtheit der Forderungen gerecht wird.

Der Gegenstand der Bedienungstheorie besteht in der Aufstellung von Beziehungen zwischen den ein Bedienungssystem charakterisierenden Parametern, das heißt zum Beispiel zwischen dem Charakter des Stromes der eintreffenden Forderungen, der Produktivität der einzelnen Bedienungsstationen, der Anzahl der Bedienungsstationen und der Effektivität der Bedienung.

Die Kenntnis dieser Beziehungen gestattet eine bestimmte Regulierung des Bedienungsprozesses, indem ein oder mehrere Parameter entsprechend verändert werden. Besitzen zum Beispiel in einem Bedienungssystem die Bedienungsstationen eine hohe Leerlaufzeit (beispielsweise 40 Prozent der gesamten Tätigkeitszeit), sind die Stationen also nur gering ausgelastet, dann läßt sich durch entsprechende Veränderungen der Anzahl der tätigen Bedienungsstationen eine bessere Auslastung der Stationen (etwa zu 90 Prozent) erreichen. Auf diesem Wege lassen sich dann optimale Zustände im Bedienungssystem herstellen.

Die Bedienungstheorie ist noch eine relativ neue mathematische Disziplin. Sie entstand in den zwanziger Jahren dieses Jahrhunderts bei der Behandlung von Problemen, die sich bei der Ausnutzung von Fernsprechanlagen ergaben. Seit dieser Zeit hat sich die Bedienungstheorie selbst wesentlich vervollkommnet und ihr Anwendungsgebiet hat sich ständig erweitert. Viele Aufgaben der Automatisierung der Produktion und der Organisation des Transports lassen eine Behandlung mittels bedienungstheoretischer Methoden zu. Des weiteren wird die Bedienungstheorie mit Erfolg bei der Projektierung und Entwicklung technischer Einrichtungen (Aggregate) angewandt, für die solche Kennziffern wie mittlere Zeit der störungsfreien Arbeit, die erforderliche Menge an Ersatzteilen und die mittlere Zeit des Leerstehens der zu konstruierenden Anlagen ausschlaggebend sind.

Im folgenden geben wir eine kurze Einführung in das Wesen und in die Methoden der Bedienungstheorie. Dabei müssen wir aus Platzmangel manche Beweise und bemerkenswerte Zusammenhänge beiseite lassen. Der interessierte Leser findet das Fehlende in der Literatur [9]—[12].

4.2. Der Eingangsstrom und die Bedienungsdauer

Wir beschäftigen uns nunmehr mit den für das Funktionieren eines Bedienungssystems so wichtigen Prozessen Ankunft und Bedienung der Forderungen. Zunächst wollen wir ein Bedienungssystem *regulär* nennen, wenn die einzelnen Forderungen in genau bestimmten Zeitpunkten eintreffen und eine genau bestimmte Bedienungszeit benötigen. Der denkbar einfachste Fall eines solchen Systems liegt vor, wenn die *Zwischenankunftszeit* (also die Zeit zwischen zwei aufeinanderfolgend eintreffenden Forderungen) und die *Bedienungsdauer* konstant sind. Die Ermittlung der die Güte der Organisation eines solchen Bedienungssystems beeinflussenden Parameter bereitet keine Schwierigkeiten. Trifft zum Beispiel in einem Warenhaus alle 5 Minuten eine Forderung ein, die für die Bedienung stets 10 Minuten beansprucht, so brauchen nur zwei Bedienungsstationen geöffnet zu werden, um Wartezeiten der Forderung bzw. die Nichtauslastung der Bedienungsstationen zu vermeiden.

Vergegenwärtigen wir uns nochmals die bisher aufgezählten Beispiele aus der Praxis für Bedienungssysteme, dann erkennen wir, daß diese nicht regulär sind. Die Zeitmomente des Eintreffens der einzelnen Forderungen sind nicht genau bestimmt, sondern zufällig. Diese Feststellung ist gleichbedeutend damit, daß die Zwischenankunftszeit eine Zufallsgröße ist. Veranschaulichen wir uns den *Eingangsstrom* der Forderungen derart, daß wir auf einer reellen Zahlengeraden die Ankunftsmomente durch Punkte kennzeichnen,

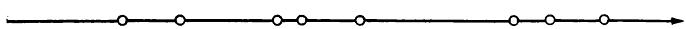

Abb. 22 Bild eines Eingangsstromes von Forderungen

dann weisen die in der Praxis vorherrschenden Eingangsströme sogenannte Verdichtungen und Verdünnungen auf (vgl. Abb. 22). Die Verdichtungen können zur Bildung von *Warteschlangen*, von *Stauungen* vor den Bedienungsstationen führen, die Verdünnungen zum unproduktiven *Leerstehen* der Bedienungsapparate.

Die Bedienungszeit, die in praktischen Bedienungssystemen für die einzelnen Forderungen benötigt wird, ist auch eine Zufallsvariable, sie ändert sich von Forderung zu Forderung zufällig. Diesen Sachverhalt belegen die bereits angeführten Beispiele.

Die für die Tätigkeit eines Bedienungssystems ausschlaggebenden Prozesse Ankunft und Bedienung der Forderungen sind also Zufallsprozesse, die durch die Zufallsgrößen Zwischenankunftszeit bzw. Bedienungszeit charakterisiert werden. Sind die Verteilungen dieser stetigen Zufallsvariablen gegeben, dann sind die entsprechenden Prozesse Ankunft und Bedienung der Forderungen eindeutig beschrieben. Von den vielfältigsten Typen dieser Zufallsprozesse werden wir im folgenden den für die Praxis wichtigsten Typ kennenlernen.

Wir beschreiben nunmehr den Eingangsstrom der Forderungen etwas ausführlicher. Zu diesem Zweck nennen wir einen Eingangsstrom *stationär*, wenn die Wahrscheinlichkeit dafür, daß in einem Zeitraum diese oder jene

Anzahl von Forderungen eintrifft, nur von der Länge dieses Zeitraumes und nicht von dem Beginn dieser Zeitspanne abhängt.

Viele Eingangsströme realer Bedienungssysteme besitzen diese Eigenschaft. Betrachten wir zum Beispiel das Bedienungssystem, das in der Wartung mehrerer Webautomaten durch einen Facharbeiter besteht. Der Arbeiter ist in diesem System die einzige Bedienungsstation; die Forderungen stellen die defekten Webautomaten dar. Die Bedienung der Forderungen ist hier mit der Reparatur der defekten Automaten identisch. Der Eingangsstrom der Forderungen in diesem System ist offenbar stationär.

Als stationär kann auch der Strom der Gespräche angesehen werden, die im Selbstwählverkehr geführt werden. Zwar ändert sich die Gesprächshäufigkeit im Verlaufe eines Tages in einem Selbstwählfernsprechamt beträchtlich, jedoch kann der Strom der Ferngespräche für einzelne Tagesabschnitte (morgens, mittags, nachts usw.) mit guter Näherung als stationär angenommen werden.

Eine ähnliche Situation herrscht in dem Bedienungssystem, das ein Dienstleistungsgeschäft verkörpert. Nehmen wir an, daß dieses Geschäft Kleidungsstücke zur chemischen Reinigung entgegennimmt. Hier ist der Eingangsstrom der Forderungen (Reinigungsaufträge) innerhalb der einzelnen Jahreszeiten (Sommer, Winter usw.) gleichfalls stationär.

Wir denken uns nun zwei Zeiträume T_1 und T_2 gegeben, von denen T_1 vor T_2 liegen möge, und bezeichnen mit $A\ (T_i)$ $(i = 1, 2)$ die Anzahl der in dem Zeitraum T_i $(i = 1, 2)$ eintreffenden Forderungen. Die eingeführten Größen $A\ (T_i)$ $(i = 1, 2)$ sind augenscheinlich diskrete Zufallsveränderliche. Wir nennen nun einen Eingangsstrom *ohne Nachwirkung*, wenn $A\ (T_2)$ von $A\ (T_1)$ unabhängig ist. In einem Forderungenstrom ohne Nachwirkung hängt demzufolge das Geschehen innerhalb eines Zeitraumes T nicht von der Vorgeschichte dieses Zeitintervalls ab. Zahlreiche reale Forderungenströme weisen diese Eigenschaft auf.

So ist zum Beispiel der Strom der Aufträge, der eine Reparaturstation erreicht, ohne Nachwirkung, weil in der Regel jeder Auftrag unabhängig davon aufgegeben wird, wann und wieviele Aufträge bis zu diesem Zeitpunkt bereits aufgegeben wurden.

Wir können auch den Strom der Forderungen als ohne Nachwirkung ansehen, der von defekten Maschinen (vgl. obiges Beispiel zu Webautomaten) ausgeht. Jedoch ist hierbei zu beachten, daß die Menge der Maschinen gewöhnlich beschränkt ist und der Strom der zu reparierenden Maschinen damit nur bedingt ohne Nachwirkung in Erscheinung tritt. Bei einer großen Anzahl von Maschinen kann dieser Strom als ohne Nachwirkung aufgefaßt werden; bei kleinerer Anzahl fehlt dem Forderungsstrom diese Besonderheit.

Wir betrachten nun eine Zeitspanne Δt und bezeichnen mit $P\ (\Delta t, n)$ die Wahrscheinlichkeit dafür, daß während der Zeitspanne Δt genau n Forderungen eintreffen. Ein Eingangsstrom heißt nun *ordinär*, wenn die Wahrscheinlichkeiten $P\ (\Delta t, k)$ $(k = 2, 3, \ldots)$ im Vergleich zu der Wahrscheinlichkeit $P\ (\Delta t, 1)$ klein sind. Diese Eigenschaft eines Forderungenstroms bedeutet mit anderen Worten, daß die Wahrscheinlichkeit dafür, daß in einem kleinen Zeitintervall mehr als eine Forderung eintrifft, sehr klein ist, daß also das Auftreten von zwei oder mehr Forderungen in einem kleinen Zeitintervall fast unwahrscheinlich ist.

Bei gewissen praktischen Eingangsströmen (z. B. im Bedienungssystem Tankstelle) ist diese Eigenschaft erkennbar; bei anderen Strömen ist diese Besonderheit intuitiv erkennbar, oder die Wirklichkeit wird mindestens mit ausreichender Annäherung als ordinär wiedergegeben. Selbstverständlich gibt es auch Bedienungssysteme, deren Forderungenstrom nicht ordinär ist. Auf solche Bedienungssysteme gehen wir jedoch in diesem Rahmen nicht ein. Vgl. hierzu [10].

Wir nennen nun einen Forderungenstrom eines Bedienungssystems POISSONsch, wenn er gleichzeitig stationär, ordinär und ohne Nachwirkung ist. POISSONsche Eingangsströme sind die in der Praxis vorherrschenden Eingangsströme. Demzufolge betrachten wir im weiteren Verlauf lediglich POISSONsche Forderungenströme.

Mit λ bezeichnen wir die sogenannte *Eingangsrate* des Forderungenstromes, das heißt die mittlere Anzahl der in der Zeiteinheit eintreffenden Forderungen. Diese Größe λ ist konstant, weil der POISSONsche Forderungenstrom stationär ist. Bedeutet weiter t_a die Zwischenankunftszeit des Eingangsstromes, dann gilt der

Satz 4.1: Die Zufallsgröße t_a ist im Intervall (o, ∞) mit der Dichte

$$\boxed{f_1(t_a) = \lambda e^{-\lambda t_a}} \tag{4.1}$$

stetig verteilt. Ihr Erwartungswert beträgt

$$\boxed{E(t_a) = \frac{1}{\lambda}}. \tag{4.2}$$

Beweis: Auf den Beweis des ersten Teils der Behauptung müssen wir verzichten; vgl. diesbezüglich [7]. Die Vollständigkeitsrelation für t_a ist wegen

$$\int_0^\infty \lambda e^{-\lambda t} dt = -e^{-\lambda t}\Big|_0^\infty = 1$$

erfüllt. Da weiter λ die mittlere Anzahl der in der Zeiteinheit eintreffenden Forderungen ausmacht, bedeutet $\frac{1}{\lambda}$ gerade diejenige Zeit, die im Durchschnitt zwischen zwei aufeinanderfolgenden Ankünften von Forderungen verstreicht. Damit ist aber (4.2) bewiesen.

Auch auf analytischem Wege ergibt sich die Beziehung (4.2) unmittelbar:

$$E(t_a) = \int_0^\infty t f_1(t)\,dt = e^{-\lambda t}\left(t + \frac{1}{\lambda}\right)\Big|_\infty^0 = \frac{1}{\lambda}.$$

Mit $E(\lambda; x, T)$ bezeichnen wir nun die Wahrscheinlichkeit dafür, daß während der Zeitspanne T genau x Forderungen eines POISSONschen Eingangsstromes mit der Eingangsrate λ eintreffen. Bezüglich dieser neuen Kennziffer des Forderungenstromes besteht der

Satz 4.2: Für die Wahrscheinlichkeit E (λ; x, T) gilt die Formel

$$\boxed{E\,(\lambda;\,x,\,T) = \frac{(\lambda T)^x\, e^{-\lambda T}}{x!}} \tag{4.3}$$

Wegen des Beweises dieser Formel verweisen wir wieder auf die Arbeit [7]. Wir erinnern bei dieser Gelegenheit daran, daß wir im Zusammenhang mit der *Poisson*schen Verteilung durch heuristische Überlegungen die Beziehung (4.3) bereits gewonnen haben; vgl. hierzu (2.47)!

Wie wir vorstehend gesehen haben, gestattet der POISSONsche Eingangsstrom eine recht einfache Beschreibung. Der Forderungenstrom wird vollständig durch die Eingangsrate λ bestimmt. Bei Kenntnis des konstanten Parameters λ können mithin alle Wesenszüge des Eingangsstromes praktisch erfaßt werden. Der wichtige Parameter λ wird bei Eingangsströmen in der Praxis auftretender Bedienungssysteme auf statistischem Wege ermittelt; auf diesen Sachverhalt können wir in diesem Rahmen leider auch nicht eingehen.

Wir betrachten nun ein Beispiel, um die eingeführten Begriffe und Formeln anzuwenden: In einer Buchhandlung werden Bestellungen für Bücher entgegengenommen, und zwar pro Viertelstunde durchschnittlich 2 Bestellungen.

Die Eingangsrate des hier bestehenden Forderungenstroms beträgt $\lambda = 2$, sofern die Viertelstunde als Zeiteinheit gewählt wird. Gemäß (4.2) verstreicht dann im Mittel zwischen zwei aufeinanderfolgenden Bestellungen die Zeit von$\left(\frac{1}{2}\right.$ Viertelstunde oder$\left.\right)$ $7^1/_2$ Minuten. Die Wahrscheinlichkeit dafür, daß in einer Stunde ($T = 4$) genau $x = 5$ Bestellungen aufgegeben werden, berechnet sich nach (4.3) zu

$$E\,(2;\,5,\,4) = \frac{8^5\, e^{-8}}{5!} = 0{,}093$$

oder 9,3 Prozent. Entsprechend ergibt sich die Wahrscheinlichkeit dafür, daß in einer halben Stunde ($T = 2$) überhaupt kein Kunde ($x = 0$) die Buchhandlung zwecks Buchbestellung betritt, zu

$$E\,(2;\,0,\,2) = \frac{4^0 \cdot e^{-4}}{0!} = 0{,}018$$

oder 1,8 Prozent. Es ist also äußerst selten, daß im Verlaufe einer halben Stunde keine Bestellung aufgegeben wird.

Wir untersuchen nun etwas eingehender die Bedienungsdauer der Forderungen, die in erster Linie ein Charakteristikum für das Funktionieren jeder Bedienungsstation eines Bedienungssystems ist. Sie gibt an, wieviel Zeit auf die Bedienung einer Forderung von der jeweiligen Bedienungsstation aufgewendet werden muß. Wir haben bereits oben konstatiert, daß die Bedienungsdauer eine Zufallsvariable ist. Die Ursachen für diesen Tatbestand sind insbesondere darin zu sehen, daß die eintreffenden Forderungen nicht völlig identisch sind und die Verschiedenheit der Forderungen zufälligen Einflüssen unterliegt. So haben zum Beispiel Fernsehempfänger und Rundfunkgeräte, die zur Reparatur in eine Elektro-Werkstatt gebracht werden, die unterschiedlichsten Defekte. Selbst wenn die Fehler der Aggregate identisch sind (zum Beispiel defekter Lautsprecher) kann die Reparaturdauer verschieden sein, sofern es sich beispielsweise um Geräte verschiedener Typen handelt. Es ist damit klar, daß die Bedienungsdauer im allgemeinen eine Zufallsgröße ist und demzufolge durch ein Verteilungsgesetz beschrieben werden kann.

Wir denken uns nun eine Bedienungsstation gegeben, die ein Strom von Forderungen erreicht. Da die bedienten Einheiten das Bedienungssystem verlassen, bildet sich ein *Ausgangsstrom* von bedienten (befriedigten) Einheiten. Aus den oben angestellten Untersuchungen geht hervor, daß dieser Strom ein Zufallsprozeß ist. Wir nehmen nun diesen Prozeß der Einfachheit halber gleichfalls als POISSONsch an, das heißt als stationär, ordinär und ohne Nachwirkung. Derartige Ausgangsströme treten vielfach in der Praxis auf bzw. approximieren mit brauchbarer Güte verschiedene reale Ausgangsströme. Darüber hinaus ist diese Annahme in den meisten theoretischen Darlegungen über praktische Bedienungssysteme getroffen worden. Demzufolge legen wir unseren weiteren Betrachtungen stets POISSONsche Ausgangsströme zugrunde.

Wir fassen nunmehr eine Bedienungsstation ins Auge, die ein Forderungenstrom erreicht. Mit μ bezeichnen wir die sogenannte Bedienungsrate dieser Bedienungsstation, das heißt die mittlere Anzahl der in der Zeiteinheit durch diese Station bedienten Forderungen. Diese Größe μ ist konstant, weil der Ausgangsstrom stationär ist. Bedeutet weiter t_b die Bedienungsdauer einer Forderung an der gegebenen Bedienungsstation, dann folgt analog zu Satz 4.1

Satz 4.3: Die Zufallsgröße t_b ist im Intervall $(0, \infty)$ mit der Dichte

$$\boxed{f_2(t_b) = \mu e^{-\mu t_b}}\,. \tag{4.4}$$

stetig verteilt. Ihr Erwartungswert beträgt

$$\boxed{E(t_b) = \frac{1}{\mu}}\,. \tag{4.5}$$

Wir beschäftigen uns nun mit einem Bedienungssystem, das n Bedienungsstationen enthält, zu denen die Bedienungsraten μ_i $(i = 1, 2, \ldots, n)$ gehören. Diese n Stationen mögen unabhängig voneinander arbeiten; diese Voraussetzung ist praktisch immer erfüllt. Es liegt nun nahe, dieses Bedienungssystem als ein System mit nur einer Bedienungsstation aufzufassen, das die Bedienungsrate μ besitzt. Es gilt dann

Satz 4.4: Ein Bedienungssystem, in dem n Stationen unabhängig voneinander mit den Bedienungsraten $\mu_1, \mu_2, \ldots, \mu_n$ tätig sind, besitzt insgesamt die Bedienungsrate

$$\boxed{\mu = \sum_{i=1}^{n} \mu_i} \,. \tag{4.6}$$

Beweis: Wir bezeichnen mit t_i $(i = 1, 2, \ldots, n)$ bzw. t die Bedienungsdauer einer Forderung durch die i-te Station $(i = 1, 2, \ldots, n)$ bzw. das ganze System. Die Parameter t_i $(i = 1, 2, \ldots, n)$ haben gemäß (4.4) die Verteilungsfunktion

$$\boxed{P(t_i < \tau) = 1 - e^{-\mu_i \tau} \quad (i = 1, 2, \ldots, n)} \,. \tag{4.7}$$

Wir berechnen nun die Verteilungsfunktion der Zufallsveränderlichen t. Zunächst besteht die Relation

$$P(t \leqq \tau) = P\left(\min_{i=1,2,\ldots,n} t_i \leqq \tau\right),$$

weil die Bedienung einer Forderung durch das Bedienungssystem beendet ist, sobald eine Bedienungsstation diese Forderung abgefertigt hat. Weiter folgt aufgrund des Multiplikationssatzes für unabhängige Wahrscheinlichkeiten (die Parameter t_i $(i = 1, 2, n)$ sind voneinander unabhängig!)

$$P\left(\min_{i=1,\ldots,n} t_i \geqq \tau\right) = P(t_1 \geqq \tau, \ldots, t_n \geqq \tau) = \prod_{i=1}^{n} P(t_i \geqq \tau)$$

oder infolge (4.7)

$$P(t < \tau) = 1 - \prod_{i=1}^{n} e^{-\mu_i \tau} = 1 - e^{-\mu \tau}$$

mit der in (4.6) stehenden Abkürzung für μ. Die Dichtefunktion für die Bedienungsdauer t des Systems beträgt mithin $f_2(t) = \mu e^{-\mu t}$. Damit ist der Satz bewiesen. Aus Satz 4.4 resultiert

Satz 4.5: Je größer die Anzahl der Bedienungsstationen eines Systems ist, umso kleiner wird die mittlere Dauer der Bedienung einer Forderung durch das Bedienungssystem.

Beweis: Aus (4.5) und (4.6) ergibt sich zwangsläufig die Behauptung

$$\boxed{E(t) = \frac{1}{\mu} = \frac{1}{\sum\limits_{i=1}^{n} \mu_i}} \,; \tag{4.8}$$

denn da die μ_i $(i = 1, 2, \ldots, n)$ nichtnegative Größen sind, verkleinert sich die linke Seite der letzten Gleichung (4.8) bei Vergrößerung der Zahl n der

Bedienungsstationen. Haben wir etwa ein Bedienungssystem mit $n = 3$ Stationen vor uns, die alle dieselbe Bedienungsrate $\mu_1 = \mu_2 = \mu_3 = 5$ aufweisen, dann beträgt die Bedienungsrate des Systems gemäß (4.6)

$$\mu = \mu_1 + \mu_2 + \mu_3 = 15$$

und die mittlere Bedienungsdauer

$$E\,(t_b) = \frac{1}{15} \text{ Stunde} = 4 \text{ Minuten},$$

falls als Zeiteinheit die Stunde (60 Minuten) zugrunde gelegt wird. Nimmt nun eine weitere Bedienungsstation mit gleicher Bedienungsrate $\mu_4 = 5$ ihre Tätigkeit im gegebenen Bedienungssystem auf, so steigt die Bedienungsrate des Systems auf

$$\mu = 20$$

an, während die mittlere Bedienungsdauer auf 3 Minuten zurückgeht. Wird andererseits im ursprünglichen System eine Bedienungsstation geschlossen, dann lauten die den Bedienungsprozeß charakterisierenden Parameter

$$\mu = 10$$

und

$$E\,(t_b) = \frac{1}{10} \text{ Stunde} = 6 \text{ Minuten}.$$

Wir bezeichnen nun mit $B\,(\mu;\, x,\, T)$ die Wahrscheinlichkeit dafür, daß während der Zeitspanne T genau x Forderungen in einem Bedienungssystem mit der Bedienungsrate μ abgefertigt werden. In Analogie zum Satz 4.2 gilt der

Satz 4.6: Für die Wahrscheinlichkeit $B\,(\mu;\, x,\, T)$ gilt die Formel

$$\boxed{B\,(\mu;\, x,\, T) = \frac{(\mu\, T)^x\, e^{-\mu T}}{x!}}\,. \tag{4.9}$$

Wir verdeutlichen diese Aussage an dem oben bereits behandelten Beispiel einer Bedienungsstation mit drei Stationen und der konstanten Bedienungsrate $\mu = 15$. Wir fragen nach der Wahrscheinlichkeit dafür, daß im Verlaufe einer halben Stunde $\left(T = \frac{1}{2}\right)$ nur $x = 10$ Forderungen bedient werden. Aus (4.9) folgt für die gesuchte Wahrscheinlichkeit der Wert

$$B\left(15;\ 10,\ \frac{1}{2}\right) = \frac{7{,}5^{10} \cdot e^{-7{,}5}}{10!} \doteq 0{,}085$$

oder 8,5 Prozent. Wird eine zusätzliche Station geöffnet bzw. eine der drei Bedienungsaggregate stillgelegt, dann ergeben sich für die gefragte Wahrscheinlichkeit die Werte

$$B\left(20;\ 10,\ \frac{1}{2}\right) = \frac{10^{10} \cdot e^{-10}}{10!} = 0{,}125$$

oder 12,5 Prozent bzw.

$$B\left(10;\ 10.\ \frac{1}{2}\right) = \frac{5^{10} \cdot e^{-5}}{10!} = 0{,}018$$

oder 1,8 Prozent.

Wie wir vorstehend festgestellt haben, gestattet der POISSONsche Ausgangsstrom bedienter Forderungen eine recht einfache Beschreibung. Dieser Strom wird vollständig durch die konstante Bedienungsrate μ des Systems charakterisiert. Was wir oben über die Ermittlung der Eingangsrate λ gesagt haben, trifft im gleichen Umfang für den Parameter μ eines praktischen Bedienungssystems zu.

Nachdem wir bisher ausführlich die Prozesse Eingang und Bedienung der Forderungen, die das Funktionieren eines jeden Bedienungssystems in allgemeinen Zügen bedingen, analysiert haben, wenden wir uns nunmehr dem Verhalten der eintreffenden Forderungen im Bedienungssystem zu. Entsprechend diesem Verhalten werden wir uns mit verschiedenen Grundtypen von Bedienungssystemen beschäftigen, die in der Praxis eine Rolle spielen.

4.3. Das Verlustsystem

Wir nennen ein Bedienungssystem ein sogenanntes *Verlustsystem*, wenn die eintreffenden Forderungen folgendes Verhalten zeigen: Ist mindestens eine Bedienungsstation frei, das heißt bereit zur Bedienung einer Forderung, dann läßt sich die eintreffende Forderung abfertigen. Sind allerdings alle Stationen besetzt, das heißt mit der Bedienung bereits früher angekommener Forderungen beschäftigt, dann verläßt die eingetroffene Forderung kurzerhand unbedient (unbefriedigt) das Bedienungssystem. Mit anderen Worten ist ein Verlustsystem ein System, das die eintreffenden Forderungen zurückweist, sobald diese auf besetzte Bedienungsstationen treffen. Es ist in diesem Zusammenhang klar, daß eine solche Bedienungssituation notwendigerweise eine endliche Anzahl der Bedienungsaggregate voraussetzt.

Ein Beispiel für ein Verlustsystem stellt ein Fernsprechamt dar, das nur eine beschränkte Anzahl von Fernsprechteilnehmern gleichzeitig vermitteln kann. Sind nun alle Verbindungslinien besetzt, das heißt wird in allen Verbindungsleitungen gesprochen, dann erhält jeder Teilnehmer, der sich zu diesem Zeitpunkt an das Fernamt wendet, eine Absage. Diese Situation ist uns allzu gut bekannt. Wählen wir zum Beispiel während der Hauptgeschäftszeit eine Telefonnummer, dann ertönt zuweilen schon nach dem Wählen der ersten Ziffer das Besetztzeichen in dem Telefonhörer. In diesem Fall erhalten wir eine Ablehnung und verlassen das Bedienungssystem, indem wir den Hörer auflegen.

Ein Verlustsystem kann (wie jedes Bedienungssystem) durch verschiedene Kennziffern charakterisiert werden. Diese Ziffern geben an, in welchem Maße das System organisiert und funktionsfähig ist. Eine solche Kennziffer ist zum Beispiel die *Verlustwahrscheinlichkeit*, das heißt die Wahrscheinlichkeit dafür, daß eine eintreffende Forderung alle Bedienungsstationen besetzt vorfindet, das System verläßt und damit dem Bedienungssystem verlorengeht. Betrachten wir bei dieser Gelegenheit nochmals das letzte Beispiel aus der Telefonie, dann stellen wir eine sehr hohe Verlustwahrscheinlichkeit für dieses Bedienungssystem während der Hauptgeschäftszeit fest. Eine weitere ein

Verlustsystem charakterisierende Kenngröße ist die mittlere Anzahl der tätigen, das heißt der mit der Bedienung von Forderungen beschäftigten Bedienungsstationen. Diese Größe drückt praktisch den Auslastungsgrad der Bedienungsaggregate aus. Die Verlustwahrscheinlichkeit gibt demgegenüber ein Bild von der mittleren Anzahl der unbedienten, dem System verlorengehenden Forderungen. Beide Kennziffern stehen in wechselseitigem Verhältnis zueinander: keine Verlustwahrscheinlichkeit bedingt einen niedrigen Auslastungsgrad und umgekehrt; hoher Auslastungsgrad hat eine große Verlustwahrscheinlichkeit zur Folge und umgekehrt. Auf diesen Zusammenhang kommen wir später ausführlich zu sprechen. Weitere Kenngrößen eines Verlustsystems werden uns im folgenden noch begegnen.

Wir gehen nun daran, Beziehungen zwischen den einzelnen Parametern herzuleiten, die ein Verlustsystem in seinem Funktionieren charakterisieren. Diese Beziehungen werden es uns ermöglichen, ein Verlustsystem optimal zu organisieren, das heißt das entsprechende Bedienungssystem auf optimale Parameter einzustellen. Zu diesem Zweck betrachten wir ein Verlustsystem mit n unabhängig voneinander arbeitenden Bedienungsaggregaten. Der Eingangsstrom ist Poissonsch mit der Eingangsrate λ. Die Bedienungsdauer einer Forderung durch die i-te Bedienungsstation ($i = 1, 2, \ldots, n$) ist mit der für alle Stationen gleichen Bedienungsrate μ — im Sinne von Satz 4.3 — exponentiell verteilt. Dann ist gemäß Satz 4.4 auch die Bedienungsdauer einer Forderung durch das gesamte System exponentiell verteilt.

Mit s bezeichnen wir die Anzahl der tätigen, das heißt mit der Bedienung von Forderungen beschäftigten Bedienungsstationen. Diese Größe s ist offenbar eine diskrete Zufallszahl mit den möglichen Werten $s = 0, 1, 2, \ldots, n$. Weiter seien Z_s das Ereignis (der Zustand) „s Bedienungsstationen sind tätig und n-s Bedienungsaggregate stehen leer" und $P_s(t)$ die Wahrscheinlichkeit dafür, daß der Zustand Z_s zur Zeit t vorherrscht; dabei ist jeweils $s = 0, 1, 2, \ldots, n$. Nun besteht folgender Zusammenhang:

Satz 4.7: Im stationären Verlustsystem gelten die Formeln:

$$\boxed{P_s(t) = p_s = \frac{\varrho^s}{s! \sum\limits_{k=0}^{n} \frac{\varrho^k}{k!}} \quad (s = 0, 1, \cdots, n)} \tag{4.10}$$

mit

$$\boxed{\varrho = \frac{\lambda}{\mu}}\,. \tag{4.10'}$$

Beweis: Der Kürze halber beweisen wir den vorstehenden Satz nur für den Fall $n = 1$, in dem die Formeln (4.10) die Gestalt

$$\boxed{P_s(t) = p_s = \frac{\varrho^s}{1+\varrho} \; (s = 0, 1)} \tag{4.11}$$

annehmen. Wegen des allgemeinen Beweises verweisen wir auf die Arbeit [6]. Wir betrachten einen Zeitmoment t sowie eine kleine Zeitspanne Δt und untersuchen im folgenden die Wahrscheinlichkeit $P_s\,(t + \Delta t)$ $(s = 0, 1)$, das heißt die Wahrscheinlichkeiten dafür, daß im Moment $t + \Delta t$ der Zustand Z_s $(s = 0, 1)$ besteht.

Der Zustand Z_0 zur Zeit $t + \Delta t$ kann sich wie folgt ergeben:

1. Zur Zeit t besteht der Zustand Z_0 (die einzige Station ist frei) und während der Zeitspanne Δt trifft keine Forderung ein.
2. Zur Zeit t herrscht der Zustand Z_1 (die einzige Station ist besetzt) und die betreffende Forderung verläßt im Laufe des Zeitraumes Δt das System.

Die Möglichkeit, daß zum Zeitpunkt t der Zustand Z_0 vorliegt und während der Zeitspanne Δt eine Forderung eintrifft und bedient wird, ist bei hinreichend klein gewählter Größe Δt ausgeschlossen. Ebenso ist es wegen der Ordinärität des Stromes der eintreffenden sowie der bedienten Forderungen nicht möglich, daß während der Zeitspanne Δt mehrere Forderungen eintreffen bzw. das System verlassen. Demzufolge kann der Zustand Z_0 zur Zeit $t + \Delta t$ nur auf einer der beiden Arten 1. oder 2. zustande kommen. Aus Satz 4.2 resultiert die Übergangswahrscheinlichkeit

$$p_{00} = E\,(\lambda;\, o, \Delta t) = e^{-\lambda \Delta t} \tag{4.12}$$

das heißt die Wahrscheinlichkeit für den Übergang $Z_0 \to Z_0$ während des Zeitraumes Δt (keine Forderung trifft ein!). Entsprechend folgt aus (4.9) die Übergangswahrscheinlichkeit

$$p_{10} = B\,(\mu;\, 1, \Delta t) = \mu \Delta t e^{-\mu \Delta t} \tag{4.13}$$

die Wahrscheinlichkeit für den Übergang $Z_1 \to Z_0$ während der Zeitspanne Δt (eine Forderung wird bedient). Wenden wir nun die Regel für die vollständige Wahrscheinlichkeit bezüglich des Ereignis Z_0 zur Zeit $t + \Delta t$ an, dann folgt bei Beachtung der Übergangswahrscheinlichkeiten (4.12) und (4.13)

$$\boxed{P_0\,(t + \Delta t) = P_0\,(t)\, e^{-\lambda \Delta t} + P_1\,(t)\, \mu\, \Delta t\, e^{-\mu \Delta t}}\,. \tag{4.14}$$

Denken wir uns für $e^{-\lambda \Delta t}$ bzw. $e^{-\mu \Delta t}$ die entsprechende Taylor-Reihe (bezüglich der Entwicklungsstelle $\Delta t = 0$) aufgeschrieben und in (4.14) eingesetzt, so ergibt sich die Gleichung

$$P_0\,(t + \Delta t) = (1 - \lambda \Delta t) P_0(t) + \mu \Delta t P_1(t) + o\,(\Delta t), \tag{4.15}$$

wobei alle Glieder mit kleinerer Ordnung als Δt zu dem Ausdruck $o\,(\Delta t)$ zusammengefaßt sind.

Der Zustand Z_1 zur Zeit $t + \Delta t$ kann wie folgt zustande kommen:

1. Zur Zeit t herrscht der Zustand Z_0 und während der Zeitspanne Δt tritt eine Forderung in das System ein.
2. Zur Zeit t besteht der Zustand Z_1 und während des Zeitraumes Δt wird die Bedienung der betreffenden Forderung nicht abgeschlossen.

Entsprechende Überlegungen wie oben zeigen, daß nur auf eine dieser Arten der Zustand Z_1 zur Zeit $t + \Delta t$ entstehen kann. Aus Satz 4.2 ergibt sich die Übergangswahrscheinlichkeit

$$p_{01} = E(\lambda; 1, \Delta t) = \lambda \Delta t e^{-\mu \Delta t}, \tag{4.16}$$

das heißt die Wahrscheinlichkeit für den Übergang $Z_0 \to Z_1$ während der Zeitspanne Δt (eine Forderung trifft ein). Die Übergangswahrscheinlichkeit

$$p_{11} = B(\mu; 0, \Delta t) = e^{-\mu \Delta t} \tag{4.17}$$

resultiert aus (4.9); sie stellt die Wahrscheinlichkeit für den Übergang $Z_1 \to Z_1$ während des Zeitraumes Δt dar (die Bedienung wird nicht abgeschlossen!). Nach der Regel für die vollständige Wahrscheinlichkeit des Ereignis Z_1 zur Zeit $t + \Delta t$ wird dann

$$\boxed{P_1(t + \Delta t) = P_0(t) \lambda \Delta t e^{-\lambda \Delta t} + P_1(t) e^{-\mu \Delta t}} \tag{4.18}$$

oder nach entsprechender Umformung wie oben (Gleichung (4.15))

$$P_1(t + \Delta t) = \lambda \Delta t\, P_0(t) + (1 - \mu \Delta t) P_1(t) + o(\Delta t). \tag{4.19}$$

Wir bringen nun die Gleichungen (4.15) und (4.19) auf folgende Gestalt

$$\left.\begin{aligned} \frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} &= -\lambda P_0(t) + \mu P_1(t) + \frac{o(\Delta t)}{\Delta t}, \\ \frac{P_1(t + \Delta t) - P_1(t)}{\Delta t} &= \lambda P_0(t) - \mu P_1(t) + \frac{o(\Delta t)}{\Delta t}. \end{aligned}\right\} \tag{4.20}$$

Nehmen wir nun den Grenzübergang $\Delta t \to 0$ vor, so ergibt sich wegen

$$\lim_{\Delta t \to 0} \frac{P_s(t + \Delta t) - P_s(t)}{\Delta t} = \frac{d P_s(t)}{dt} = P_s'(t) \quad (s = 0,1)$$

und

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$$

das lineare Differential — Differenzengleichungssystem

$$\boxed{\begin{aligned} P_0'(t) &= -\lambda P_0(t) + \mu P_1(t), \\ P_1'(t) &= \lambda P_0(t) - \mu P_1(t) \end{aligned}}\,. \tag{4.21}$$

Die Lösungen $P_s(t)$ $(s = 0, 1)$ dieses Systems sind die gesuchten Wahrscheinlichkeiten dafür, daß zur Zeit t der Zustand Z_s $(s = 0, 1)$ herrscht. Wir setzen nun voraus, daß das betrachtete Verlustsystem stationär ist. Sodann sind diese Wahrscheinlichkeiten von der Zeit t unabhängig: $P_s(t) = p_s$ $(s = 0, 1)$. Aus (4.21) folgt dann mit der in (4.10′) eingeführten Bezeichnung für ϱ

$$\boxed{p_1 = \varrho\, p_0}\,. \tag{4.22}$$

Da nun Z_0 und Z_1 ein vollständiges System von Ereignissen bilden, muß zwangsläufig

$$\boxed{p_1 + p_0 = 1} \tag{4.23}$$

bestehen. Aus (4.22) und (4.23) resultiert aber schon die Behauptung (4.11). Wir beweisen nun den

Satz 4.8: Ein stationäres Verlustsystem besitzt die *Ablehnungswahrscheinlichkeit*

$$\boxed{p_{ab} = \frac{\varrho^n}{n! \sum\limits_{k=0}^{n} \frac{\varrho^k}{k!}}} \tag{4.24}$$

und weist im Mittel

$$\boxed{E(s) = \frac{\sum\limits_{k=1}^{n} \frac{\varrho^k}{(k-1)!}}{\sum\limits_{k=0}^{n} \frac{\varrho^k}{k!}}} \tag{4.25}$$

tätige Bedienungsstationen auf.

Beweis: Eine ankommende Forderung erhält eine Ablehnung, wenn alle Bedienungsaggregate besetzt sind, wenn also der Zustand Z_n besteht. Damit gilt $p_{ab} = p_n$. In Verbindung mit (4.10) für $s = n$ ergibt sich dann gerade die Aussage (4.24). Um die Behauptung (4.25) nachzuweisen, gehen wir von der Verteilungstabelle für die diskrete Zufallsvariable s aus:

s	0	1	2	....	n
p	p_0	p_1	p_2	....	p_n

Aufgrund der über diskrete Verteilungen gemachten Ausführungen beträgt die mittlere Anzahl der tätigen Stationen, das heißt der Erwartungswert von s kurzerhand

$$E(s) = \sum_{k=0}^{n} k p_k = \sum_{k=1}^{n} k p_k .$$

Beachten wir (4.10) und die Beziehung $k\ (k-1)! = k!$, dann erhalten wir gerade die Behauptung (4.25). Damit ist der Satz vollständig bewiesen.

Die hergeleiteten Beziehungen (4.10) heißen Erlangsche *Formeln.* Sie sind unter der Annahme aufgestellt worden, daß die Bedienungsdauer der einzelnen Forderungen eine exponentiell verteilte Zufallsgröße darstellt (Vgl. (4.4)!). Kürzlich ist nachgewiesen worden, daß die Erlangschen Formeln auch dann gelten, wenn die Bedienungsdauer eine beliebige Verteilung mit dem Erwartungswert $\frac{1}{\mu}$ besitzt. Vgl. hierzu [8]. Mithin folgt

Satz 4.9: Die Aussagen der Sätze 4.7 und 4.8 gelten auch für Verlustsysteme mit beliebiger Verteilung für die Bedienungsdauer, deren Erwartungswert gleich der reziproken Bedienungsrate μ ist.

Wir geben nunmehr einige Beispiele zu den bisher gewonnenen Ergebnissen.

In einer Stadt wird ein Fernsprechamt projektiert, das in der Lage sein soll, n Gesprächsteilnehmern gleichzeitig die Gespräche zu vermitteln. Aufgrund statistischer Erhebungen ist ermittelt worden, daß der für dieses Fernsprechamt zu erwartende Forderungenstrom die Ankunftsrate $\lambda = 1$ Gespräch pro Minute besitzt. Des weiteren ist die Erfahrung gesammelt worden, daß die Gesprächsdauer der Teilnehmer im Durchschnitt 2 Minuten beträgt. Demzufolge besitzt jede Bedienungsstation die Bedienungsrate $\mu = \frac{1}{2}$ Gespräche pro Minute. Das gesamte Fernsprechamt fertigt sodann im Mittel pro Minute $\frac{n}{2}$ Gespräche ab. Für die Projektanten ist nun die Frage von Interesse: Wie groß muß die Zahl n sein, das heißt mit wieviel Verbindungskanälen muß das zukünftige Fernamt ausgestattet sein, damit höchstens 10 Prozent der Gesprächsteilnehmer auf besetzte Leitungen treffen? Gegeben ist also die Ablehnungswahrscheinlichkeit $p_{ab} \leqq 0,1$; gesucht ist die Anzahl n der Bedienungsstationen. Setzen wir in (4.24) die Werte $\lambda = 1$, $\mu = \frac{1}{2}$ und als Folge von (4.10′) $\varrho = 2$ ein, dann folgt aus der Bedienung $p_{ab} \leqq 0,1$ unmittelbar

$$0,1 \geqq \frac{2^n}{n! \sum\limits_{k=0}^{n} \frac{2^k}{k!}}$$

oder

$$S_l(n) \equiv n! \sum_{k=0}^{n} \frac{2^k}{k!} \geqq 10 \cdot 2^n \equiv S_r(n) \tag{4.26}$$

Aus dieser Ungleichung ist nun die Größe n zu ermitteln. Zunächst ist die rechte Seite für $n = 1$ größer als die linke Seite:

$$S_l(1) = 3 < 20 = S_r(1).$$

Als Antwort auf die oben gestellte Frage kommt damit $n = 1$ nicht in Betracht. Um nun den tatsächlichen Wert für n zu erhalten, berechnen wir die beiden Seiten $S_r(n)$ und $S_l(n)$ in (4.26) für $n = 2, \ldots, n_0$; dabei ist n_0 die kleinste natürliche Zahl, für die die rechte Seite die linke Seite nicht übertrifft. Das Ergebnis dieser Rechnungen stellen wir in der folgenden Tabelle zusammen:

n	$S_l(n)$	$S_r(n)$
1	3	20
2	10	40
3	38	80
4	168	160

Es ist demzufolge $n_0 = 4$. Mithin lautet die gesuchte Antwort

$$n \geqq 4.$$

Das im Projekt befindliche Fernsprechamt muß infolgedessen mindestens 4 Verbindungskanäle aufweisen, damit höchstens jeder zehnte Gesprächsteilnehmer auf besetzte Leitungen trifft.

Werden nun genau $n = 4$ Verbindungslinien installiert, so beträgt die Ablehnungswahrscheinlichkeit des zukünftigen Fernamtes gemäß (4.24)

$$p_{ab} = \frac{2^4}{4!\sum_{k=0}^{4} \frac{2^k}{k!}} = \frac{16}{168} \approx 0{,}095$$

oder 9,5 Prozent. Unter diesen Umständen treffen also von 1000 Teilnehmern im Mittel 95 auf durchwegs besetzte Leitungen.

Vermöge (4.10) geben wir zum Abschluß noch die vollständige Verteilungstabelle für die Zufallsgröße s = Anzahl der besetzten Verbindungslinien bei vier Kanälen an

s	0	1	2	3	4
p_s	0,143	0,286	0,286	0,190	0,095

Mit einer Wahrscheinlichkeit von 14,3 Prozent ist demzufolge das Kanalsystem der Fernsprechanlage unbesetzt (ohne Aufträge!); mit einer Wahrscheinlichkeit von 19 Prozent ist nur eine Verbindungslinie frei usw.

Wir führen nun als neue Kennziffer eines Verlustsystems den Quotienten $c = \frac{E(s)}{n}$ ein, der in Verbindung mit (4.25) die Gestalt

$$c = \frac{\sum_{k=1}^{n} \frac{\varrho^k}{(k-1)!}}{n \sum_{k=0}^{n} \frac{\varrho^k}{k!}} \tag{4.27}$$

besitzt. Dieser Parameter stellt offensichtlich den *Auslastungsgrad der Bedienungsstationen* eines Verlustsystems dar. Die Größe (4.27) zeigt also, welcher Anteil (in Prozent) der Kapazität des Verlustsystems im Mittel über einen längeren Zeitraum in Anspruch genommen wird. In dem letzten Beispiel, das von dem Fernsprechamt handelt, nimmt dieser Parameter den Wert $c = 45{,}2$ Prozent an.

Als abschließendes Beispiel für ein Verlustsystem betrachten wir ein Mitropa-Friseurgeschäft, in dem 3 Friseure arbeiten. Diesen Friseursalon suchen im Mittel 15 Reisende pro Stunde auf, um sich bedienen (frisieren, rasieren, . . .) zu lassen. Da die Reisenden wenig Zeit haben, verlassen sie das Friseurgeschäft, sobald sie warten müssen, das heißt, sobald alle 3 Friseure mit der Bedienung bereits früher eingetroffener Kunden beschäftigt sind. Jeder Friseur fertigt durchschnittlich 3 Reisende pro Stunde ab. Uns interessieren in diesem Zusammenhang folgende Fragen:

1. Wie groß ist die Wahrscheinlichkeit dafür, daß von den 3 Friseuren $s = 0$, 1, 2 bzw. 3 beschäftigt sind?
2. Wie groß ist der Auslastungsgrad des Friseurgeschäfts?
3. Wie lauten die Antworten auf die Fragen 1. und 2., wenn einer der drei Friseure wegen Krankheit ausfällt?

Ad 1) Aus (4.10) ergibt sich wegen $n = 3$, $\lambda = 15$, $\mu = 3$ und $\varrho = 5$ das Verteilungsgesetz für die Zufallsgröße s zu

s	0	1	2	3
p_s	0,025	0,127	0,318	0,530

Ad 2) Gemäß (4.27) liegt der Auslastungsgrad

$$c = \frac{185}{236} = 0{,}784$$

vor.

Ad 3) Nunmehr sind die Parameterwerte $n = 2$, $\lambda = 15$, $\mu = 3$ und $\varrho = 5$ maßgebend. Entsprechend oben erhalten wir

s	0	1	2
p_s	0,054	0,270	0,676

sowie $c = 81{,}1$ Prozent.

4.4. Das Wartesystem

Wir nennen ein Bedienungssystem ein sogenanntes Wartesystem, wenn die eintreffenden Forderungen folgendes Verhalten zeigen: Ist mindestens eine Bedienungsstation frei, dann läßt sich die eintreffende Forderung abfertigen. Sind allerdings alle Stationen besetzt, dann wartet die eintreffende Forderung solange, bis eine Bedienungsstation frei wird, die sie dann bedient. Mit anderen Worten ist ein Wartesystem ein Bedienungssystem, das eine eintreffende Forderung nur dann verläßt, wenn sie vollkommen bedient ist. Wir beschränken uns im folgenden auf den Fall, daß die wartenden Forderungen eine geordnete Linie bilden, die wir *Wartelinie* oder *Warteschlange* nennen. Diese Beschränkung bedeutet, daß eine Forderung, die bereits l_2 wartende Forderungen in der Schlange vorfindet, sich als $(l_2 + 1)$-te wartende Einheit in die Schlange „anstellt". Entsprechend der Reihenfolge bei der Ankunft werden die wartenden Forderungen durch die Bedienungsaggregate abgefertigt.

Ein Beispiel für ein Wartesystem stellt eine Reparaturstation dar. Die eintreffenden Forderungen sind Sacheinheiten (z. B. Maschinen oder andere Aggregate: Traktoren, Fernsehapparate, Küchenmaschinen, ...), die reparaturbedürftig sind; die Bedienung ist identisch mit der Reparatur durch einen Facharbeiter (zum Beispiel Fernsehmechaniker). Die Forderungen verweilen solange in der Reparaturstation, bis sie repariert sind. Die Forderungen warten meistens in einer geordneten Warteschlange.

In einem Geschäft können wir beispielsweise zwei Wartesysteme erkennen. Auf der einen Seite haben wir das System, das aus den Käufern (Forderungen) und den Verkäufern (Bedienungsstationen) besteht; die in der Schlange stehenden Personen warten darauf, von den Verkäufern mit Ware bedient zu werden. Auf der anderen Seite beobachten wir das System, das sich aus der Ware (Forderungen) und den Verkäufern (Bedienungsaggregaten) zusammensetzt; die auf Lager liegende Ware wartet darauf, durch die Verkäufer verkauft zu werden. In beiden Wartesystemen tritt also der Verkäufer als Bedienungsstation auf.

Die wartenden Forderungen müssen keine geordnete Wartelinie bilden, sie können auch einen ungeordneten Haufen repräsentieren (wie etwa im Wartesystem Ware — Verkäufer bei gewissen Waren und in bestimmten Geschäften: die Waren werden nicht in der Reihenfolge verkauft, in der sie angeliefert worden sind). Auf solche Wartesysteme gehen wir im folgenden nicht ein. Ferner braucht die Wartelinie nicht sichtbar zu sein, sie kann beispielsweise — gewissermaßen unsichtbar — nur auf dem Papier stehen (etwa in Form einer Bestelliste). Ein Wartesystem kann nun — wie jedes Bedienungssystem — durch verschiedene Kennziffern charakterisiert werden. Diese Ziffern geben ein Bild davon, in welchem Maße das Wartesystem organisiert und funktionsfähig ist. Eine solche Kennziffer ist zum Beispiel die *mittlere Schlangenlänge*, das heißt die mittlere Anzahl der in der Schlange auf Bedienung wartenden Forderungen. Betrachten wir die oben genannte Reparaturstation als Beispiel, so bedeutet dieser Parameter die mittlere Zahl der bereits zur Reparatur angenommenen, aber noch nicht instandgesetzten Aggregate. Eine weitere ein Wartesystem charakterisierende Größe ist die *mittlere Wartezeit* einer eintreffenden Forderung, das heißt die Zeit, die eine eintreffende Forderung im Mittel bis zu ihrer Bedienung warten muß. Diese Größe ist beispielsweise in dem Wartesystem Ware-Verkäufer von Interesse, sobald die Waren (zum Beispiel Lebensmittel und Obst) nur eine beschränkte Zeit auf Lager liegen dürfen, da anderenfalls ihr Wert (zum Beispiel Genußfähigkeit) herabsinkt. Die bereits aufgezählten Parameter charakterisieren das Wartesystem von der Seite der Forderungen her. Für die Einschätzung des Auslastungsgrades ist — wie im Fall eines Verlustsystems — auch hier die mittlere Anzahl der tätigen Bedienungsstationen bedeutsam. Weitere Kennziffern eines Wartesystems begegnen uns noch im folgenden.

Wir gehen nun an die Aufstellung von Relationen zwischen den einzelnen das Wartesystem kennzeichnenden Parametern. Diese Relationen gestatten die optimale Organisation eines Wartesystems, das heißt eine Einstellung des Systems auf optimale Parameter. Zu diesem Zweck betrachten wir n unabhängig voneinander arbeitende Bedienungsstationen; dabei ist n eine endliche Zahl. Bezüglich des Eingangsstroms und der Bedienungsdauer einer Forderung durch eine Station bzw. das gesamte System treffen wir die an entsprechender Stelle über Verlustsysteme gemachten Voraussetzungen.

Des weiteren beschränken wir uns im folgenden auf den für die Praxis wichtigen Fall $\lambda < n\mu$ oder gemäß (4.10′) $\varrho < n$. Es ist klar, daß die Ungleichung $\lambda > n\mu$ oder $\varrho > n$ zu einer tendentiell anwachsenden Schlangenlänge führt, denn in der Zeiteinheit treffen stets mehr Forderungen im Mittel ein als durchschnittlich bedient werden können. Aber auch der Fall $\varrho = n$ hat wegen des zufälligen Charakters des Eingangs- und Ausgangsstromes eine tendentiell anwachsende Schlangenlänge zur Folge. Dieser Sachverhalt wird später noch deutlicher hervortreten.

Wir bezeichnen mit l_1 die Anzahl der im Bedienungsprozeß stehenden Forderungen und mit l_2 die Schlangenlänge, das heißt die Anzahl der auf ihre Bedienung wartenden Forderungen. Offenbar sind l_1 und l_2 diskrete Zufallsvariable mit den möglichen Werten $l_1 = 0, 1, 2, \ldots, n$ bzw. $l_2 = 0, 1, 2, \ldots$ Weiter ist ersichtlich, daß die Ereignisse „$l_1 = n$" und „$l_2 = 0$" identisch sind:[1]

$$P\,(l_1 = n) = P\,(l_2 = 0).$$

Die Summe dieser beiden Zufallsgrößen ist wieder eine Zufallsveränderliche $l = l_1 + l_2$, die zuweilen *uneigentliche Schlangenlänge* im Gegensatz zur *eigentlichen Schlangenlänge* l_2 heißt. Es gilt

$$\boxed{l = \begin{cases} l_1, \text{ wenn keine Forderung wartet} \\ l_2 + n, \text{ wenn alle Stationen besetzt sind.} \end{cases}} \tag{4.28}$$

Die Zahl l gibt offenbar die Gesamtzahl der sich im Bedienungssystem aufhaltenden Forderungen an.

Mit Z_l bezeichnen wir weiter das Ereignis (den Zustand) „l Forderungen verweilen im Wartesystem" und mit $P_l\,(t)$ die Wahrscheinlichkeit dafür, daß der Zustand Z_l zu der Zeit t vorherrscht; dabei gilt stets $l = 0, 1, 2, 3, \ldots$ In diesem Zusammenhang konstatieren wir den

Satz 4.9: Im stationären Wartesystem bestehen die Formeln

$$\boxed{P_l\,(t) = p_l = \begin{cases} \dfrac{\varrho^l}{l!\,R} \text{ wenn } 0 \leqq l \leqq n\,; \\[2ex] \dfrac{\varrho^l}{n!\,n^{l-n}\cdot R}\,, \text{ wenn } n \leqq l < \infty \end{cases}} \tag{4.29}$$

mit $\varrho = \dfrac{\lambda}{\mu}$ — vgl. Gleichung (4.10′) — und

$$\boxed{R = \sum_{k=0}^{n-1} \frac{\varrho^k}{k\,!} + \frac{\mu\,\varrho^n}{(n-1)\,!\,(n\,\mu - \lambda)}}\,. \tag{4.29′}$$

[1] Befinden sich nämlich genau n Forderungen im System ($l_1 = n$), dann werden alle diese Forderungen bedient, da es ja n Bedienungskanäle gibt. Dann wartet also keine Forderung ($l_2 = 0$).

Diese Behauptung wird im Prinzip nach derselben Methode bewiesen, der wir uns beim Beweis des Satzes 4.7 für den Fall $n = 1$ bedienten. Aus diesem Grunde verzichten wir auf den Beweis des Satzes 4.9 und verweisen auf [7].

Wir wenden nun die Formel (4.29) auf ein Wartesystem an, das ein Auskunftsbüro darstellt. In der Stunde mögen durchschnittlich 20 Personen eintreffen. Das aus 3 Auskunftsschaltern bestehende Büro möge eine Bedienungsrate von 30 Personen aufweisen. Uns interessieren folgende Fragen:

1. Mit welcher Wahrscheinlichkeit stehen alle drei Schalter leer?
2. Mit welcher Wahrscheinlichkeit sind genau 6 Personen im Büro?
3. Mit welcher Wahrscheinlichkeit ist ein Schalter frei?

Wegen $n = 3, \lambda = 20, \mu = 10$ und $\varrho = \frac{\lambda}{\mu} = 2$ erhalten wir der Reihe nach aus (4.29) die Antworten

$$\textit{Ad } 1)\ p_0 = \frac{1}{R} = \frac{1}{9} \approx 0{,}111$$

$$\textit{Ad } 2)\ p_6 = \frac{32}{81 \cdot R} = \frac{32}{729} \approx 0{,}045$$

$$\textit{Ad } 3)\ p_2 = \frac{2}{R} = \frac{2}{9} \approx 0{,}222\,.$$

Wir beweisen nun den

Satz 4.10: Die Wahrscheinlichkeit dafür, daß in einem stationären Wartesystem alle Bedienungsstationen besetzt sind, beträgt

$$\boxed{P_b = \frac{\varrho^n}{R\,(n-1)!\,(n-\varrho)}}\,. \tag{4.30}$$

Beweis: Der Zustand „alle Stationen sind besetzt“ tritt ein, wenn mindestens n Forderungen im System verweilen. Das bedeutet gemäß (4.29′)

$$P_b = \sum_{l_2=0}^{\infty} P_{l_2} = \sum_{l=n}^{\infty} P_l = \frac{n^n}{R \cdot n!} \sum_{l=n}^{\infty} \left(\frac{\varrho}{n}\right)^l = \frac{n^n\,\varrho^n}{R \cdot n!\,n^n} \sum_{k=0}^{\infty} \left(\frac{\varrho}{n}\right)^k.$$

Wegen $\varrho = \frac{\lambda}{\mu} < n$ konvergiert die unendliche Reihe im letzten Ausdruck. Wir erhalten damit

$$P_b = \frac{\varrho^n}{R \cdot n!}\,\frac{1}{1-\frac{\varrho}{n}} = \frac{\varrho^n}{R\,(n-1)!\,(n-\varrho)}\,.$$

Im letzten Beispiel, das von dem Auskunftsbüro handelt, ergibt sich diese Wahrscheinlichkeit zu

$$P_b = \frac{4}{9} \approx 0{,}444$$

Satz 4.11: Die mittlere eigentliche Schlangenlänge in einem stationären Wartesystem beträgt

$$\boxed{L = E\,(l_2) = \frac{\varrho^{n+1}}{R\,(n-1)!\,(\varrho-n)^2}}\,. \tag{4.31}$$

Der Auslastungsgrad für die Bedienungsstationen eines stationären Wartesystems beläuft sich auf

$$\boxed{c = 1 - \frac{1}{Rn} \sum_{K=0}^{n-1} \frac{(n-K)\,\varrho^K}{K!}}. \tag{4.32}$$

Beweis: Aus (4.29) ergibt sich zunächst

$$L = E(l_2) = \sum_{l_2=0}^{\infty} l_1\, p_{n+l_1} = \frac{\varrho^n}{n!\,R} \sum_{k=0}^{\infty} k \cdot \left(\frac{\varrho}{n}\right)^k. \tag{*}$$

Die unendliche Reihe

$$S = \sum_{k=0}^{\infty} a_k \equiv \sum_{k=1}^{\infty} k \left(\frac{\varrho}{n}\right)^k \tag{**}$$

ist konvergent, weil der Quotient $\frac{a_{k+1}}{a_k}$ von zwei aufeinanderfolgenden Gliedern dieser Reihe für $k \to \infty$ kleiner als 1 ist. (Wir hatten $\varrho < n$ vorausgesetzt!) Wir formen die Summe (**) nunmehr um (diese Umformung ist wegen der bestehenden Konvergenz gestattet!)[1]

$$S = \lim_{m\to\infty} \sum_{j=1}^{m} \left[\sum_{j=i}^{m} \left(\frac{\varrho}{n}\right)^i\right] = \lim_{m\to\infty} \frac{\left(\frac{\varrho}{n}\right)^m - 1}{\frac{\varrho}{n} - 1} \sum_{j=1}^{m} \left(\frac{\varrho}{n}\right)^j$$

$$= \lim_{m\to\infty} \frac{\left[\left(\frac{\varrho}{n}\right)^m - 1\right]^2}{\frac{n}{\varrho}\left[\frac{\varrho}{n} - 1\right]^2},$$

indem wir zweimal die Summenformel für die endliche geometrische Reihe anwenden. Wegen $\varrho < n$ gilt $\lim_{m\to\infty} \left(\frac{\varrho}{n}\right)^m = 0$. Mithin resultiert aus (*) und (**) die Behauptung (4.31)

$$L = \frac{\varrho^n}{n!\cdot R} \cdot \frac{\varrho}{n} \frac{1}{\left(\frac{\varrho}{n} - 1\right)^2} = \frac{\varrho^{n+1}}{R\,(n-1)!\,(\varrho - n)^2}.$$

Mit f bezeichnen wir die Anzahl der freien Stationen eines Wartesystems.

[1] Wir denken uns dabei das allgemeine Glied $k\left(\frac{\varrho}{n}\right)^k$ in (**) durch die Summe der k-Größen derselben Form $\left(\frac{\varrho}{n}\right)^k$ ersetzt und die Glieder der so aus (**) hervorgehenden Reihe umgeordnet.

Die Größe f ist eine diskrete Zufallsvariable mit dem Verteilungsgesetz

f	0	1	2	$\ldots n-1$	n
p	p_b	p_{n-1}	p_{n-2}	$\cdots p_1$	p_0

(vgl. (4.29) und (4.30)!). Demzufolge beträgt ihr Mittelwert

$$E(f) = \sum_{k=0}^{n-1} (n-k)\, p_k .$$

Die mittlere Anzahl der tätigen Bedienungsaggregate besitzt den Wert $n - E(f)$. Dividieren wir diesen Ausdruck $n - E(f)$ durch n, so ergibt sich gerade der Auslastungskoeffizient c. Berücksichtigen wir in $E(f)$ noch die Gleichungen (4.29), dann resultiert unmittelbar die Aussage (4.32).
Wenden wir die hergeleiteten Formeln (4.31) und (4.32) auf das letzte Beispiel (Auskunftsbüro) an, dann erhalten wir folgende Werte: Die mittlere Schlangenlänge beträgt

$$L = \frac{8}{9} \approx 0{,}89$$

Forderungen. Der Auslastungsgrad der Bedienungsstationen erweist sich zu

$$c = 1 - \frac{1}{3} = \frac{2}{3}$$

oder zu 66,7 Prozent.

Wir hatten bereits zu Beginn der Untersuchungen über Wartesysteme festgestellt, daß der Fall $\varrho = n$ zu einer sehr großen Schlangenlänge führt. Diese Feststellung unterstreicht nun die Formel (4.31):

$$\lim_{\varrho \to n} L = \frac{n^{n+1}}{R\,(n-1)! \cdot 0} = \infty .$$

Wir haben vorstehend einige Kennziffern und Beziehungen zwischen diesen abgeleitet, die ein stationäres Wartesystem als Bedienungssystem charakterisieren. Wir beschäftigen uns im folgenden mit einem weiteren wichtigen Parameter, der *Wartezeit* der eintreffenden Forderungen. Dieser Parameter ist offenbar eine stetige Zufallsgröße, deren Verteilungsfunktion und Mittelwert uns besonders interessiert. Zunächst beweisen wir den

Satz 4.12: In einem stationären Wartesystem wartet eine eintreffende Forderung mit einer Wahrscheinlichkeit von

$$\boxed{W(\delta) = 1 - \frac{\varrho^n\, e^{-(n\mu-\lambda)\delta}}{R\,(n-1)!\,(n-\varrho)}} \tag{4.33}$$

weniger als $\delta > 0$ Zeiteinheiten.

Beweis: Eine eintreffende Forderung möge bei ihrer Ankunft l Forderungen im Bedienungssystem vorfinden. Die Wahrscheinlichkeit $P_l(\vartheta)$ dafür, daß diese Forderung mindestens δ Zeiteinheiten bis zum Bedienungsbeginn wartet, ist gleich der Wahrscheinlichkeit dafür, daß innerhalb der Zeitspanne (gerechnet von dem Moment der Ankunft) höchstens $l - n$ Forderungen abgefertigt werden. Diese letzte Wahrscheinlichkeit ergibt sich aus dem Summensatz für Wahrscheinlichkeiten und Satz 4.6 zu

$$P_l(\delta) = \sum_{m=0}^{l-n} \frac{(n\mu\delta)^m e^{-n\mu\delta}}{m!}, \tag{*}$$

denn während der Zeitspanne δ fertigt jede Bedienungsstation im Mittel n Forderungen ab. Nach der Regel für die vollständige Wahrscheinlichkeit stellt dann

$$W^*(\delta) = \sum_{l=0}^{\infty} P_l(\delta)\, p_l \tag{**}$$

die Wahrscheinlichkeit dafür dar, daß eine eintreffende Forderung mindestens δ Zeiteinheiten warten muß. Diese Wahrscheinlichkeit hängt mit der gesuchten Größe $W(\delta)$ in der Form $W(\delta) + W^*(\delta) = 1$ zusammen. Wir formen jetzt (**) um. Zunächst ist $P_l(\delta) = 0$, wenn $l < n$ ist; denn in diesem Fall wartet die eintreffende Forderung nicht, sie wird sofort bedient. Setzen wir weiter (*) in (**) ein, dann wird unter Berücksichtigung von (4.29')

$$W^*(\delta) = \sum_{l=n}^{\infty} P_l(\delta)\, p_l = \frac{n^n e^{-n\mu\delta}}{R \cdot n!} \sum_{l=n}^{\infty} \left(\frac{\varrho}{n}\right)^l \cdot \left(\sum_{m=0}^{l-n} \frac{(n\mu\delta)^m}{m!}\right). \tag{***}$$

Beachten wir noch die Identität[1]

$$\sum_{l=n}^{\infty} \left(\frac{\varrho}{n}\right)^l \left(\sum_{m=0}^{l-n} \frac{(n\mu\delta)^m}{m!}\right) = \sum_{k=n}^{\infty} \left(\frac{\varrho}{n}\right)^k \sum_{r=0}^{\infty} \frac{(\lambda\delta)^r}{r!},$$

dann erhalten wir wegen

$$\sum_{k=n}^{\infty} \left(\frac{\varrho}{n}\right)^k = \frac{\left(\frac{\varrho}{n}\right)^n}{1 - \frac{\varrho}{n}}$$

— unendliche geometrische Reihe mit dem Quotienten $\frac{\varrho}{n} < 1$ — und

$$\sum_{r=0}^{\infty} \frac{(\lambda\delta)^r}{r!} = e^{\lambda\delta}$$

Taylorsche Entwicklung der e-Funktion — aus (***)

$$W^*(\delta) = \frac{\varrho^n e^{-(n\mu-\lambda)\delta}}{R\,(n-1)!\,(n-\varrho)}.$$

Das ist aber wegen $W^*(\delta) + W(\delta) = 1$ schon die Behauptung (4.33).

[1] Diese Identität läßt sich leicht bestätigen, indem unter Beachtung von $\left(\frac{\lambda}{\mu}\right) = \varrho$ beide Seiten ausführlich als Potenzreihen in δ aufgeschrieben werden und ein Vergleich der entsprechenden δ-Koeffizienten (auf Identität) vorgenommen wird. Ist das geschehen, so ist die Identität beider Seiten durch den Identitätssatz über Potenzreihen sichergestellt.

Die hergeleitete Formel (4.33) stellt die Verteilungsfunktion der Wartezeit einer eintreffenden Forderung — gerechnet von der Ankunft bis zum Bedienungsbeginn — für positive Werte δ dar. Beachten wir die selbstverständliche Relation $W(\delta) = 0$ für $\delta \leqq 0$, dann ergibt sich insgesamt die Wartezeitverteilungsfunktion zu

$$\boxed{W(\delta) = \begin{cases} (4.33), \text{ wenn } \delta > 0 \\ 0, \text{ sonst.} \end{cases}} \tag{4.34}$$

Die Verteilungsfunktion $W(\delta)$ ist für $\delta = 0$ unstetig. Wir erkennen, daß sie beim Passieren des Ursprungs in Richtung wachsender δ-Werte einen endlichen Sprung von $W = 0$ auf $W = 1 - p_b$ (vgl. die Formel (4.30)!) macht. Die Sprunghöhe ist demzufolge gleich der Wahrscheinlichkeit dafür, daß nicht alle Bedienungsstationen besetzt sind, das heißt, daß mindestens ein Bedienungsaggregat frei ist. Wir beweisen nun den

Satz 4.13: In einem stationären Wartesystem besteht für eine eintreffende Forderung die mittlere Wartezeit

$$\boxed{E(t_w) = \frac{\varrho^n}{R\,(n-1)!\,(n-\varrho)^2\mu}}\,. \tag{4.35}$$

Beweis: Die Wartezeit t_w einer eintreffenden Forderung ist eine stetige Zufallsvariable mit der Verteilungsfunktion (4.34) nebst (4.33). Aufgrund der über stetige Verteilungen angestellten Untersuchungen können wir

$$E(t_w) = \int\limits_{-\infty}^{+\infty} t\,W'(t)\,dt \tag{*}$$

folgern; denn $W'(t_w) = \frac{d\,W(t_w)}{d\,t_w} = w(t_w)$ ist die Wahrscheinlichkeitsdichte für die Zufallsgröße t_w. Wegen (4.34_2) und (4.33) folgt aus (*)

$$E(t_w) = \frac{\varrho^n}{R\,(n-1)!\,(n-\varrho)} \cdot (n\mu - \lambda) \int\limits_0^{\infty} t\,e^{-(n\mu-\lambda)t}\,dt\,.$$

Wenden wir auf das vorstehende Integral die Formel

$$\alpha \int\limits_0^{\infty} t\,e^{-\alpha t}\,dt = -\,te^{-\alpha t}\Big|_0^{\infty} + \int\limits_0^{\infty} e^{-\alpha t}\,dt = \frac{e^{-\alpha t}}{\alpha}\Big|_{\infty}^{0} = \frac{1}{\alpha}$$

an, indem wir $\alpha = n\mu - \lambda$ setzen, so ergibt sich schon die Behauptung

$$E(t_w) = \frac{\varrho^n}{R\,(n-1)!\,(n-\varrho)} \cdot \frac{1}{n\mu-\lambda} = \frac{\varrho^n}{R\,(n-1)!\,(n-\varrho)^2\mu}\,.$$

Damit ist alles gezeigt.

Wegen (4.30) können wir die Formel (4.35) zu

$$\boxed{E(t_w) = \frac{P_b}{\mu\,(n-\varrho)}} \tag{4.36}$$

umschreiben. Diese Schreibweise läßt klar erkennen, daß die mittlere Wartezeit für eine eintreffende Forderung direkt proportional ist der Wahrscheinlichkeit dafür, daß alle Stationen besetzt sind, das heißt die betreffende Forderung warten muß. Desweiteren resultiert aus (4.36) unmittelbar die Tatsache, daß für $\varrho \to n$ die mittlere Wartezeit unendlich groß wird, vgl. hierzu die bereits oben getroffenen Bemerkungen. Die Formeln (4.33) und (4.35) wenden wir nun auf das letzte Beispiel an, das von dem Auskunftsbüro handelt. Die mittlere Wartezeit berechnet sich zu

$$E\,(t_w) = \frac{4}{90} \approx 0{,}044$$

Stunden oder ca. 2,7 Minuten. Die Wahrscheinlichkeit dafür, daß eine Forderung weniger als die durchschnittliche Wartezeit von rund 2,7 Minuten wartet, beläuft sich auf

$$W\left(\frac{4}{90}\right) = 1 - \frac{4}{9}\, e^{-\frac{4}{9}} \approx 0{,}714\,.$$

In 71,4 Prozent aller Fälle wartet demzufolge eine eintreffende Forderung weniger als die mittlere Wartezeit.

Als Abschluß unserer Betrachtungen über Wartesysteme behandeln wir folgendes Beispiel: Eine Werkstatt, in der $n = 5$ Mechaniker (gleicher Produktivität und unabhängig voneinander) arbeiten, erreicht ein aus defekten Aggregaten (z. B. Fernsehapparaten, Küchengeräten, Fahrzeugen usw.) bestehender Forderungenstrom. Statistische Erhebungen haben ergeben, daß zwischen der Anlieferung zweier aufeinanderfolgender Aufträge durchschnittlich eine Zeit von einer Stunde verstreicht und daß die Dauer der Reparatur (Bedienung) einer Forderung durch einen Mechaniker im Mittel vier Arbeitsstunden ausmacht. Wir stellen in diesem Zusammenhang folgende Fragen:

1. Wie lange hält sich ein Aggregat in der Werkstatt im Mittel auf?
2. Mit welcher Wahrscheinlichkeit verbleibt ein defektes Aggregat weniger als einen Tag in der Reparaturstation?
3. Wieviel Aggregate warten durchschnittlich auf ihre Instandsetzung?
4. Wie groß ist die Wahrscheinlichkeit dafür, daß in der Werkstatt drei defekte Aggregate auf Reparatur warten?
5. Wie groß ist die Wahrscheinlichkeit dafür, daß alle fünf Handwerker ohne Arbeit sind?
6. Wie hoch ist der Auslastungsgrad der Mechaniker?
7. Wie fallen die Antworten zu den Fragen 1. bis 6. aus, wenn nur drei Handwerker in der Reparaturstation beschäftigt sind?

Wir legen unseren weiteren Überlegungen als Zeiteinheit einen achtstündigen Arbeitstag zugrunde. Dann sind uns die Parameter $n = 5$, $\lambda = 8$ und $\mu = 2$ gegeben. Aus (4.10′) bzw. (4.29′) berechnen wir die Hilfsgrößen $\varrho = \frac{8}{2} = 4$ und

$$R = \sum_{k=0}^{4} \frac{4^k}{k\,!} + \frac{2 \cdot 4^5}{4\,! \cdot 2} = 77\,.$$

Die unerläßliche Voraussetzung $\varrho < n$ für ein Wartesystem ist damit erfüllt. Wir beginnen nun mit der Beantwortung der einzelnen Fragen

Ad 1: Aus (4.35) resultiert

$$E\,(t_w) = \frac{4^5}{77 \cdot 4\,! \cdot 2 \cdot 1^2} = \frac{64}{231} \approx 0{,}277\,.$$

Ein in Auftrag gegebenes Aggregat wartet mithin im Mittel 0,277 Arbeitstage oder annähernd zwei Stunden (genauer 133 Minuten!), ehe es in die Reparatur genommen wird. Da die Reparatur eines Gerätes durchschnittlich vier Stunden dauert, verbleibt demzufolge ein Aggregat im Mittel ca. 6 Stunden in der Werkstatt. Das heißt, es vergeht durchschnittlich eine Zeit von ungefähr 6 Stunden von der Aufgabe bis zur völligen Erledigung des Reparaturauftrages.

Ad 2: Wir wenden die Formel (4.33) mit $\delta = \frac{1}{2}$ an (da ein halber Tag durchschnittlich für die Instandsetzung benötigt wird!).

$$W\left(\frac{1}{2}\right) = 1 - \frac{4^5 \cdot e^{-1}}{77 \cdot 4! \cdot 1} = 1 - \frac{256 \cdot e^{-1}}{462} \approx 0{,}796\,.$$

Infolgedessen beträgt die Wahrscheinlichkeit dafür, daß eine Foderung weniger als einen Tag im System verbleibt, 79,6 Prozent. Lediglich in einem Fünftel aller Fälle hält sich ein defektes Aggregat mindestens einen Tag in der Werkstatt auf.

Ad 3: Wir benutzen die Formel (4.31)

$$L = \frac{4^6}{77 \cdot 4! \cdot 1} = \frac{1024}{462} \approx 2{,}2$$

Ad 4: Aus (4.29′) folgt mit $l = 8$

$$P_8 = \frac{4^8}{5! \cdot 5^3 \cdot 77} \approx 0{,}056$$

Mit einer Wahrscheinlichkeit von 5,6 Prozent warten demzufolge drei Aggregate auf Instandsetzung.

Ad 5: Wir erhalten aus (4.29) mit $l = 0$

$$p_0 = \frac{1}{77} \approx 0{,}013$$

Nur mit einer Wahrscheinlichkeit von 1,3 Prozent tritt infolgedessen der Zustand ein, daß alle Handwerker ohne Arbeit sind.

Ad 6: Aus (4.32) ergibt sich

$$c = 1 - \frac{1}{77 \cdot 5} \sum_{k=0}^{4} (5 - k)\,\frac{4^k}{k!} = 0{,}8\,.$$

Der Auslastungsgrad der Reparaturstation beträgt mithin 80 Prozent.

Ad 7: In diesem Falle ist die unbedingte Voraussetzung $\varrho < n$ nicht erfüllt: Die drei Mechaniker reparieren am Tage durchschnittlich 6 Aggregate, während 8 Aggregate am Tage im Mittel eintreffen. Wir haben somit ein Bedie-

nungssystem vor uns, das mit dem Strom der Forderungen nicht fertig wird; es entstehen unbeschränkte Wartezeiten und Schlangenlängen.
Mit diesem Beispiel beschließen wir unsere Ausführungen über die Bedienungstheorie. Neben den behandelten Warte- und Verlustsystemen gibt es noch zahlreiche andere Typen von Bedienungssystemen, die in der Praxis eine Rolle spielen und eine gewisse Zwischenstellung zu den beiden behandelten Typen einnehmen. Als zwei solcher Typen erwähnen wir nur kurz:

a) ein System, in dem die eintreffende Forderung nur dann wartet, wenn die Länge der bereits vorhandenen Schlange eine vorgegebene Grenze (beispielsweise die Lagerkapazität einer Reparaturstelle!) nicht übersteigt; andernfalls verläßt die Forderung das Bedienungssystem unbedient.
b) ein System, in dem die eintreffende Forderung nur eine gewisse Zeitspanne wartet; hat nach Ablauf dieser Zeitspanne die Bedienung der Forderung noch nicht begonnen, so verläßt die Forderung das Bedienungssystem unbedient.

Desweiteren ergeben sich Verallgemeinerungen, wenn die Eingangs- und Ausgangsströme instationär sind und beispielsweise die Eingangsrate nicht konstant (sondern abhängig von der Zeit und der Schlangenlänge) ist. Auf solche verallgemeinerten Bedienungssituationen können wir in diesem Rahmen nicht eingehen. Darüber hinaus müssen wir auch die Möglichkeiten unerwähnt lassen, die die moderne Rechentechnik (im Zusammenhang mit der Monte-Carlo-Methode) der Bedienungstheorie (Simulieren von realen Bedienungssituationen, insbesondere wirklicher Eingangsströme) bietet. Uns ist an dieser Stelle nur ein Hinweis auf die Literatur [9] bis [12] gestattet.

ANLAGENVERZEICHNIS

Nachstehend geben wir Tabellen für die Funktionen $\Phi(\xi)$, $\Psi(\xi_0; n)$ und $\Lambda(n; p_0)$, die in der Wahrscheinlichkeitsrechnung im allgemeinen und in der Stichprobentheorie im besonderen eine bedeutende Rolle spielen. Diese Tabellen sind auf der elektronischen Rechenanlage ZRA 1 des Rechenzentrums der Universität Rostock berechnet worden. Die letzte Stelle der in den Tabellen angegebenen Funktionswerte ist gerundet. Der absolute Fehler dieser Werte beläuft sich demzufolge dem Betrage nach auf höchstens einer halben Einheit der betreffenden letzten Dezimalstelle.

Tafel der Funktion $\Phi(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-\frac{x^2}{2}} dx$

ξ	$\Phi(\xi)$	ξ	$\Phi(\xi)$	ξ	$\Phi(\xi)$	ξ	$\Phi(\xi)$	ξ	$\Phi(\xi)$
0,00	0,5000	0,52	0,6985	1,04	0,8508	1,56	0,9406	2,16	0,9846
1	5040	53	7019	05	8531	57	9418	18	9854
2	5080	54	7054	06	8554	58	9429	20	9861
3	5120	55	7088	07	8577	59	9441	22	9868
4	5160	56	7123	08	8599	60	9452	24	9875
5	5199	57	7157	09	8621	61	9463	26	9881
6	5239	58	7190	10	8643	62	9474	28	9887
7	5279	59	7224	11	8665	63	9484	30	9893
8	5319	60	7257	12	8686	64	9495	32	9898
9	5359	61	7291	13	8708	65	9505	34	9904
10	5398	62	7324	14	8729	66	9515	36	9909
11	5438	63	7357	15	8749	67	9525	38	9913
12	5478	64	7389	16	8770	68	9535	40	9918
13	5517	65	7422	17	8790	69	9545	42	9922
14	5557	66	7454	18	8810	70	9554	44	9927
15	5596	67	7486	19	8830	71	9564	46	9931
16	5636	68	7517	20	8849	72	9572	48	9934
17	5675	69	7549	21	8869	73	9582	50	9938
18	5714	70	7580	22	8888	74	9591	52	9941
19	5753	71	7611	23	8907	75	9599	54	9945
20	5793	72	7642	24	8925	76	9608	56	9948
21	5832	73	7673	25	8944	77	9616	58	9951
22	5871	74	7703	26	8962	78	9625	60	9953
23	5910	75	7734	27	8980	79	9633	62	9956
24	5948	76	7764	28	8997	80	9641	64	9959
25	5987	77	7794	29	9015	81	9649	66	9961
26	6026	78	7823	30	9032	82	9656	68	9963
27	6064	79	7853	31	9049	83	9664	70	9965
28	6103	80	7881	32	9066	84	9671	72	9967
29	6141	81	7910	33	9082	85	9678	74	9969
30	6179	82	7939	34	9099	86	9686	76	9971
31	6217	83	7967	35	9115	87	9693	78	9973
32	6255	84	7995	36	9131	88	9699	80	9974
33	6293	85	8023	37	9147	89	9706	82	9976
34	6331	86	8051	38	9162	90	9713	84	9977
35	6368	87	8078	39	9177	91	9719	86	9979
36	6406	88	8106	40	9192	92	9726	88	9980
37	6443	89	8133	41	9207	93	9732	90	9981
38	6480	90	8159	42	9222	94	9738	92	9982
39	6517	91	8186	43	9236	95	9744	94	9984
40	6554	92	8212	44	9251	96	9750	96	9985
41	6591	93	8238	45	9265	97	9756	2,98	9986
42	6628	94	8264	46	9279	98	9761	3,00	9986
43	6664	95	8289	47	9292	1,99	9767	10	9990
44	6700	96	8315	48	9306	2,00	9772	20	9993
45	6736	97	8340	49	9319	02	9783	30	9995
46	6772	98	8365	50	9332	04	9793	40	9996
47	6808	0,99	8389	51	9345	06	9803	50	9997
48	6844	1,00	8413	52	9357	08	9812	60	9998
49	6879	01	8438	53	9370	10	9821	70	9998
50	6915	02	8461	54	9382	12	9830	80	9999
51	6950	03	8485	55	9394	14	9838	90	9999

Tafel der Funktion $\Psi\,(\xi_0\,;n) = \dfrac{2\,\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi\,(n-1)}\;\Gamma\left(\frac{n-1}{2}\right)} \displaystyle\int_0^{\xi_0} \left[1+\frac{x^2}{n-1}\right]^{-\frac{n}{2}} dx$

$\xi_0 \diagdown n$	2	3	4	5	6
0,1	0,063	0,070	0,073	0,075	0,076
2	126	140	146	149	151
3	185	207	216	221	224
4	242	272	284	290	294
5	295	333	348	356	361
6	344	390	409	419	425
7	388	443	465	477	485
8	429	492	517	531	540
9	466	536	565	581	590
1,0	500	577	609	626	636
1	530	614	648	666	678
2	557	647	683	703	716
3	582	676	715	736	749
4	605	703	743	765	779
5	625	727	769	791	806
6	644	749	792	815	829
7	661	768	812	835	850
8	677	786	830	853	868
9	691	802	846	869	884
2,0	704	816	860	883	897
1	717	829	873	896	910
2	728	841	884	907	920
3	738	851	894	916	930
4	748	861	904	925	938
5	757	870	912	933	945
6	766	878	919	939	951
7	774	885	926	945	957
8	781	892	932	950	961
9	788	898	937	955	966
3,0	795	904	942	959	969
1	801	909	946	963	972
2	807	914	950	966	975
3	812	919	954	969	978
4	817	923	957	972	980
5	822	927	960	974	982
6	827	930	963	977	984
7	831	933	965	978	985
8	836	937	967	980	987
9	840	939	969	982	988
4,0	843	942	971	983	989
1	847	945	973	984	990
2	851	947	975	986	991
3	854	949	976	987	992
4	857	951	978	988	992
5	860	953	979	988	993
6	863	955	980	989	993
7	866	957	981	990	994
8	869	959	982	991	994
9	871	960	983	991	995
5,0	874	962	984	992	995
1	876	963	985	992	996

Tafel der Funktion $\Psi(\xi_0; n) = \frac{2\,\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\,\Gamma\left(\frac{n-1}{2}\right)} \cdot \int\limits_0^{\xi_0} \left[1 + \frac{x^2}{n-1}\right]^{-\frac{n}{2}} dx$

$\xi_0 \diagdown n$	7	8	9	10	12	14	16
0,1	0,076	0,077	0,077	0,077	0,078	0,078	0,078
2	152	153	153	154	155	155	155
3	226	227	228	229	230	231	232
4	297	299	300	301	303	304	305
5	365	367	369	371	373	374	375
6	429	432	435	436	439	441	442
7	490	493	496	498	501	503	505
8	545	550	553	555	559	562	563
9	597	602	605	608	612	615	617
1,0	644	649	653	656	661	664	666
1	686	692	696	700	705	708	711
2	724	730	735	739	744	748	751
3	758	765	770	773	779	783	786
4	788	795	800	804	810	814	818
5	815	822	827	831	838	842	845
6	839	846	851	855	862	866	869
7	859	866	872	876	882	886	890
8	877	885	890	894	900	904	907
9	893	900	905	909	915	920	923
2,0	907	914	919	923	929	933	935
1	919	925	930	934	940	944	946
2	929	936	940	944	949	953	955
3	938	944	949	952	957	961	963
4	946	952	956	959	964	967	970
5	953	958	962	965	970	973	975
6	959	964	968	971	975	977	979
7	964	969	972	975	979	981	983
8	968	973	976	979	982	984	986
9	972	976	979	982	985	987	988
3,0	975	979	982	984	987	989	990
1	978	982	985	987	989	991	992
2	981	984	987	988	991	992	993
3	983	986	988	990	992	994	994
4	985	988	990	991	993	995	995
5	986	989	991	993	994	995	996
6	988	991	992	994	995	996	997
7	989	992	993	994	996	997	997
8	990	993	994	995	996	997	998
9	991	993	995	996	997	997	998
4,0	992	994	995	996	997	998	998
1	993	995	996	997	998	998	998

Tafel der Funktion $\Psi(\xi_0; n) = \frac{2\,\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\,\Gamma\left(\frac{n-1}{2}\right)} \cdot \int\limits_0^{\xi_0} \left[1 + \frac{x^2}{n-1}\right]^{-\frac{n}{2}} dx$

$\xi_0 \diagdown n$	18	20	25	30	40	∞
0,1	0,078	0,079	0,079	0,079	0,079	0,080
2	156	156	157	157	157	159
3	232	232	233	234	234	235
4	306	306	307	308	308	311
5	376	377	378	379	380	383
6	443	444	446	447	448	451
7	506	507	509	510	512	516
8	565	566	568	564	571	576
9	619	620	623	624	626	632
1,0	668	670	672	674	676	683
1	713	714	717	719	721	729
2	753	755	758	760	762	770
3	789	790	794	796	798	806
4	820	822	825	827	830	838
5	847	849	853	855	857	866
6	871	873	877	879	882	890
7	892	894	897	900	902	911
8	910	912	915	917	920	928
9	925	927	930	932	935	943
2,0	938	939	942	944	947	954
1	948	950	953	955	957	964
2	957	958	962	963	966	972
3	965	966	969	971	972	979
4	971	973	975	976	978	984
5	976	978	980	981	983	988
6	981	982	984	985	986	991
7	984	985	987	988	989	993
8	987	988	989	990	991	995
9	989	990	991	992	993	996
3,0	991	992	993	994	995	997
1	993	993	994	995	996	998
2	994	995	995	996	997	999
3	995	996	996	997	997	999
4	996	996	997	997	998	9995
5	997	997	997	998	998	9997
6	997	997	998	998	998	9998
7	998	998	998	998	999	9999
8	998	998	998	999	999	9999
9	998	998	999	999	999	9999
4,0	998	999	999	999	999	9999
1	999	999	999	999	999	99995

Tafel der Funktion $\xi_0 = \Lambda\,(n\,;\,p_0)$ mit $\Psi\,[\Lambda\,(n;\,p_0)\,;\,n] \equiv p_0$

$n \diagdown p_0$	0,90	0,95	0,99	0,999
2	6,37	12,88	68,43	
3	2,93	4,33	10,29	56,78
4	2,36	3,20	5,99	19,15
5	2,14	2,78	4,60	8,61
6	2,02	2,57	4,03	6,86
7	1,95	2,45	3,71	5,96
8	1,90	2,37	3,50	5,41
9	1,86	2,31	3,36	5,04
10	1,84	2,26	3,25	4,78
12	1,80	2,20	3,11	4,44
14	1,78	2,16	3,01	4,22
16	1,76	2,13	2,95	4,07
18	1,75	2,11	2,90	3,97
20	1,74	2,09	2,86	3,88
25	1,71	2,06	2,80	3,75
30	1,70	2,05	2,76	3,66
40	1,69	2,02	2,71	3,56
∞	1,65	1,96	2,58	3,30

[1] Kolmogoroff, A.N. Grundbegriffe der Wahrscheinlichkeitsrechnung, Ergebnisse der Mathematik und ihrer Grenzgebiete, Bd. 2, H. 3 Springer Verlag, Berlin 1933

[2] Gnedenko, B. W. Lehrbuch der Wahrscheinlichkeitsrechnung Akademie-Verlag, Berlin 1957

[3] Gnedenko, B. W., Chintschin, A. J. Elementare Einführung in die Wahrscheinlichkeitsrechnung VEB Deutscher Verlag der Wissenschaften, Berlin 1955

[4] Fisz, M. Wahrscheinlichkeitsrechnung und mathematische Statistik VEB Deutscher Verlag der Wissenschaften, Berlin 1958

[5] v. Mangoldt, H., Knopp, K. Einführung in die höhere Mathematik, Bd. 3 Hirzel-Verlag, Leipzig 1933

[6] Wentzel, E. S. Теория вероятностей (Wahrscheinlichkeitsrechnung) Государственное издательство Физико-математической литературы, Moskau 1962

[7] Rosenberg, W. J., Prochorow, A. J. Einführung in die Bedienungstheorie B. G. Teubner Verlagsgesellschaft, Leipzig 1964

[8] Sewastjanow, B. A. Ein ergodischer Satz für MARKOWsche Prozesse und seine Anwendung auf Verlustsysteme der Telefonie, in: Zeitschrift Теория вероятностей и её применения (Wahrscheinlichkeitsrechnung und ihre Anwendung) Nr. 2/1957, S. 106 bis 116

[9] Chintschin, A. J. Работы по математической теории массового обслуживания (Arbeiten zur mathematischen Theorie der Massenbedienung) Государственное издательство физико-математической литературы, Moskau 1963

[10] Buslenko, N. P., Schreider, J. A. Die Monte-Carlo-Methode und ihre Verwirklichung mit elektronischen Digitalrechnern B. G. Teubner Verlagsgesellschaft, Leipzig 1964

[11] Saaty, Th. L. Elements of queueing theory
(Elemente der Warteschlangentheorie)
Mc Graw-Hill, New York 1961

[12] Takacs, L. Introduction to the theory of queues
(Einführung in die Warteschlangentheorie)
Oxford Univ. Press 1962