B. W. Gnedenko, A. J. Chintschin

Elementare Einführung in die Wahrscheinlichkeitsrechnung

Übersetzung: Karl-Heinz Rupp
1955 Deutscher Verlag der Wissenschaften
MSB: Nr. 10
Abschrift und LaTex-Satz: 2021 https://mathematikalpha.de

Vorwort zur ersten Auflage

Die Beherrschung der theoretischen Grundlagen einer mathematischen Disziplin ermöglicht immer eine bewusstere und aktivere Anwendung ihrer Ergebnisse in der Praxis. Mit der Wahrscheinlichkeitsrechnung ist es genauso.
Mit den praktischen Anwendungen dieser Wissenschaft hat eine große Zahl leitender Funktionäre (und mitunter auch einfacher Mitarbeiter) der Armee, der Industrie, der Landwirtschaft, der Wirtschaftswissenschaft usw. zu tun, deren mathematische Bildung natürlicherweise begrenzt ist.

Unser Büchlein soll nun in möglichst verständlicher Form diese Werktätigen mit den Grundbegriffen der Theorie und den Methoden der Wahrscheinlichkeitsrechnung bekannt machen. Es ist allen Absolventen einer Ober-(Zehnjahres-)Schule durchaus zugänglich.
Absolventen einer allgemeinbildenden Schule können höchstens an ganz vereinzelten Stellen Schwierigkeiten haben. Das Büchlein ist in allen seinen Teilen auf der Basis konkreter praktischer Beispiele aufgebaut. Bei der Auswahl dieser Beispiele gingen wir jedoch weniger von der praktischen Aktualität als von ihrem anschaulichen Wert für das Verständnis der entsprechenden theoretischen Leitsätze aus.

Moskau, den 7. Januar 1945.

Vorwort zur zweiten Auflage

Diese Auflage unterscheidet sich kaum von der ersten Schlussbemerkungen sowie einige erläuternde Bemerkungen im Text wurden hinzugefügt.

Kiew-Moskau, den 7. November 1949.

Vorwort zur dritten Auflage

Die dritte Auflage unterscheidet sich kaum von der zweiten. Es wurden nur einige unbedeutende Verbesserungen aufgenommen.

17. April 1952

Inhaltsverzeichnis

Erster Teil. Wahrscheinlichkeiten

1 Wahrscheinlichkeiten von Ereignissen

1.1 Der Begriff der Wahrscheinlichkeit

Sagt man von irgendeinem Schützen, er erziele bei feststehenden Schussbedingungen 92 % der möglichen Treffer, so bedeutet das, dass von hundert unter bestimmten Bedingungen (ein und dasselbe Ziel in derselben Entfernung, mit demselben Gewehr usw.) abgegebenen Schüssen im Mittel etwa 92 treffen (und nur 8 vorbei- gehen).
Es werden freilich nicht von jeden hundert Schüssen 92 treffen: Manchmal werden es 91 oder 90 und manchmal auch 93 oder 94 sein. Die Trefferzahl kann zuweilen auch merklich kleiner bzw. bedeutend größer als 92 sein.
Im Mittel jedoch wird bei oftmaliger Wiederholung des Schießens unter denselben Bedingungen die angegebene Prozentzahl unverändert bleiben, solange nicht im Laufe der Zeit irgendeine wesentliche Änderung eintritt (so könnte unser Schütze sein Können steigern und es im Mittel auf 95 und mehr Treffer bringen).

Die Erfahrung zeigt nun, dass so ein Schütze meistens eine Trefferanzahl bei jeweils hundert Schüssen erreichen wird, die in der Nähe von 92 liegt. Es werden allerdings auch solche Serien auftreten, in denen z. B. die Anzahl der Treffer kleiner als 88 oder größer als 96 ist. Das wird jedoch nur verhältnismäßig selten geschehen. Die Zahl 92 %, die als Kennziffer für die Meisterschaft unseres Schützen dient, wird in der Regel sehr stabil sein, d. h., die Anzahl der Treffer der (unter denselben Bedingungen) abgegebenen Schüsse wird für den gleichen Schützen fast immer dieselbe sein. Sie wird nur in seltenen Ausnahmefällen beträchtlich von ihrem Mittelwert abweichen.

Wir betrachten noch ein anderes Beispiel. In einem Betrieb wurde festgestellt, dass unter gegebenen Produktionsbedingungen im Mittel 1,6 % der gefertigten Gegenstände nicht normgerecht waren und daher Ausschuss darstellten.
Das bedeutet also, dass in einem Posten von sagen wir 1000 Erzeugnissen, die noch nicht durch die Gütekontrolle gingen, etwa 16 unbrauchbar sind. Manchmal wird die Anzahl der unbrauchbaren Erzeugnisse freilich etwas größer und manchmal auch etwas kleiner sein.
Im Mittel wird diese Anzahl jedoch in der Nähe von 16 liegen, und in den meisten Posten von je 1000 Erzeugnissen wird sie nahe bei 16 sein.

Es ist selbstverständlich, dass wir auch hier die Produktionsbedingungen (Organisation des technologischen Prozesses, Einrichtung, Rohstoffe, Qualifikation der Arbeiter usw.) als unveränderlich voraussetzen.

Es ist klar, dass diese Beispiele noch beliebig erweitert werden könnten. In allen diesen Beispielen sehen wir aber, dass bei gleichartigen Massenerscheinungen (mehrmaliges Schießen, Massenproduktion von Erzeugnissen u. dgl. m.) dieser oder jener Form die Anzahl der für uns wichtigen Ereignisse (Treffer, unbrauchbare Erzeugnisse usw.) unter gegebenen Bedingungen fast immer dieselbe ist.
Nur in ganz seltenen Fällen wird sie irgendwie bedeutend von einem gewissen mittleren Wert abweichen.

Man kann daher sagen, dass diese mittlere Ziffer ein Charakteristikum der gegebenen Massenerscheinung (bei gegebenen, streng feststehenden Bedingungen) ist. Der Prozentsatz an

Treffern beschreibt uns die Meisterschaft des Schützen, und der Prozentsatz an Ausschuss gibt uns ein Bild von der Güte der Produktion.

Es versteht sich von selbst, das eine solche Kennziffer von bedeutendem Wert in den verschiedensten Gebieten, wie z. B. im Kriegswesen, in der Technik, in der Wirtschaft, in der Physik, in der Chemie usw., ist.
Mit Hilfe dieser Kennziffer können wir nicht nur Massenerscheinungen, die schon stattgefunden haben, abschätzen, sondern auch den Ausgang dieser oder jener zukünftigen Massenerscheinung voraussagen.

Trifft ein Schütze unter bestimmten Bedingungen bei 100 Schüssen im Mittel 92mal das Ziel, so sagt man, dass unter diesen Bedingungen für diesen Schützen die Treffwahrscheinlichkeit gleich 92 % (oder 0,92 oder 92/100) ist.
Befinden sich unter je 1000 Fertigerzeugnissen eines Betriebes, die unter bestimmten Bedingungen hergestellt werden, im Mittel 16 unbrauchbare, so sagen wir, die Wahrscheinlichkeit der Ausschussfabrikation betrage für diesen Betrieb 0,016 oder 1,6

Was wollen wir nun aber im allgemeinen als Wahrscheinlichkeit eines Ereignisses innerhalb einer gegebenen Massenerscheinung bezeichnen?
Jetzt ist es schon nicht mehr schwierig, hierauf zu antworten. Eine Massenerscheinung besteht immer aus der Wiederholung oder dem gleichzeitigen Auftreten einer großen Anzahl untereinander ähnlicher Einzelerscheinungen (das Schießen aus Einzelschüssen, die Massenproduktion aus der Herstellung einzelner Gegenstände usw.).

Uns interessiert nun ein bestimmtes Resultat einer Einzelerscheinung (das Treffen eines Einzelschusses, Unbrauchbarkeit eines Einzelerzeugnisses usw.) und vor allem die Anzahl der Resultate in der betreffenden Massenerscheinung (wieviel Schüsse das Ziel treffen, wieviel Erzeugnisse Ausschuss sind usw.).

Wir nennen nun die Prozentzahl (oder allgemein den Teil) der "günstigen" [1] Resultate die Wahrscheinlichkeit dieses für uns wichtigen Resultates in der gegebenen Massenerscheinung,
Hierbei muss man immer im Auge behalten, dass die Frage nach der Wahrscheinlichkeit eines Ereignisses (Resultates) nur unter genau bestimmten Bedingungen, unter denen unsere Massenerscheinung abläuft, einen Sinn hat. Jede wesentliche Änderung dieser Bedingungen hat in der Regel eine Änderung der uns interessierenden Wahrscheinlichkeit zur Folge.

Wird in einer Massenoperation das Ereignis A (z. B. Treffen des Ziels) im Mittel a-mal bei b Einzeloperationen (Schüssen) beobachtet, so ist die Wahrscheinlichkeit des Ereignisses A unter fest gegebenen Voraussetzungen $\frac{a}{b}$ (oder $\frac{100a}{b}\%$).
Wir können daher sagen, dass wir unter der Wahrscheinlichkeit eines "günstigen" Ausganges einer Einzeloperation den Quotienten aus der im Mittel beobachteten Anzahl der "günstigen" Ausgänge und der Anzahl aller Einzeloperationen, aus denen sich die Massenoperation zusammensetzt, verstehen.

Selbstverständlich kann, wenn die Wahrscheinlichkeit irgendeines Ereignisses $\frac{a}{b}$ ist, in jeder gegebenen Serie von b Einzeloperationen dieses Ereignis sowohl öfter als auch weniger als a mal auftreten. Nur im Mittel tritt es a mal ein, und in der Mehrzahl solcher Serien von b Operationen wird sich die Anzahl der Fälle, in denen das Ereignis A eintritt, der Zahl a nähern,

[1] Im zweiten Beispiel müsste man eher "ungünstig" sagen. In der Wahrscheinlichkeitsrechnung ist es jedoch üblich, jene Resultate als "günstig" zu bezeichnen, die zur Verwirklichung des uns in dem Problem interessierenden Ereignisses führen.

insbesondere, wenn b eine große Zahl ist.

Beispiel 1. Im Laufe des ersten Quartals wurden in einer gewissen Stadt geboren:

im Januar 145 Knaben und 135 Mädchen
im Februar 142 Knaben und 136 Mädchen
im März 152 Knaben und 140 Mädchen

Wie groß ist die Wahrscheinlichkeit der Geburt eines Knaben?
Der Anteil der Knabengeburten war:

im Januar: $\frac{145}{280} \approx 0,518 = 51,8\%$
im Februar: $\frac{142}{278} \approx 0,511 = 51,1\%$
im März: $\frac{152}{292} \approx 0,520 = 52,0\%$

Wir sehen, dass das arithmetische Mittel der Anteile der einzelnen Monate nahe bei $0,516 = 51,6\%$ liegt.
Die gesuchte Wahrscheinlichkeit ist unter den gegebenen Bedingungen also etwa 0,516 oder 51,6 %. Diese Ziffer ist aus der Demographie, der Wissenschaft, die die Bevölkerungsbewegung studiert, gut bekannt. Es zeigt sich, dass in verschiedenen Perioden der Anteil der Knabengeburten unter gewöhnlichen Bedingungen nur unbedeutend von dieser Zahl abweicht.

Beispiel 2. Zu Anfang des vergangenen Jahrhunderts wurde eine bemerkenswerte Erscheinung entdeckt, die (nach dem Entdecker, dem englischen Botaniker Brown) den Namen Brownsche Molekularbewegung erhielt. Bei dieser Erscheinung wird beobachtet, dass sich kleinste Stoffteilchen, die in einer Flüssigkeit verteilt sind[2], in chaotischer Bewegung befinden. Diese Bewegung läuft ohne unmittelbar erkennbare Ursache ab.

Man konnte sich die Ursache dieser sozusagen spontanen Bewegung solange nicht erklären, bis die kinetische Gastheorie eine einfache und erschöpfende Erklärung gab: Die Bewegung der in auftreten. Nur im Mittel tritt es a mal ein, und in der Mehrzahl solcher Serien von b Operationen wird sich die Anzahl der Fälle, in denen das Ereignis A eintritt, der Zahl a nähern, insbesondere, wenn b eine große Zahl ist.

Beispiel 1. Im Laufe des ersten Quartals wurden in einer gewissen Stadt geboren:

im Januar 145 Knaben und 135 Mädchen
im Februar 142 Knaben und 136 Mädchen
im März 152 Knaben und 140 Mädchen

Wie groß ist die Wahrscheinlichkeit der Geburt eines Knaben?

Der Anteil der Knabengeburten war:

im Januar: $\frac{145}{280} \approx 0,518 = 51,8\%$,
im Februar: $\frac{142}{278} \approx 0,511 = 51,1\%$,
im März: $\frac{152}{292} \approx 0,520 = 52,0\%$.

Wir sehen, dass das arithmetische Mittel der Anteile der einzelnen Monate nahe bei 0,516 $=$ 51,6% liegt. Die gesuchte Wahrscheinlichkeit ist unter den gegebenen Bedingungen also etwa 0,516 oder 51,6%.
Diese Ziffer ist aus der Demographie, der Wissenschaft, die die Bevölkerungsbewegung studiert, gut bekannt. Es zeigt sich, dass in verschiedenen Perioden der Anteil der Knabengeburten

[2]Das heißt, die sich in einem Zustand indifferenten Gleichgewicht befinden.

unter gewöhnlichen Bedingungen nur unbedeutend von dieser Zahl abweicht.

Beispiel 2. Zu Anfang des vergangenen Jahrhunderts wurde eine bemerkenswerte Erscheinung entdeckt, die (nach dem Entdecker, dem englischen Botaniker Brown) den Namen Brownsche Molekularbewegung erhielt. Bei dieser Erscheinung wird beobachtet, dass sich kleinste Stoffteilchen, die in einer Flüssigkeit verteilt sind[3], in chaotischer Bewegung befinden. Diese Bewegung läuft ohne unmittelbar erkennbare Ursache ab.

Man konnte sich die Ursache dieser sozusagen spontanen Bewegung solange nicht erklären, bis die kinetische Gastheorie eine einfache und erschöpfende Erklärung gab: Die Bewegung der in einer Flüssigkeit verteilten Teilchen ist das sichtbare Ergebnis von Stößen der Flüssigkeitsmoleküle gegen diese Teilchen.
Durch die kinetische Gastheorie hat man nun die Möglichkeit, die Wahrscheinlichkeiten dafür zu berechnen, dass ein Teilchen des verteilten Stoffes gar nicht oder nur einmal oder zweimal oder dreimal usw. angestoßen wird. Um die Resultate der Theorie zu überprüfen, wurde eine ganze Reihe von Experimenten durchgeführt.

Wir zitieren die Ergebnisse von 518 Beobachtungen des schwedischen Physikers Svedberg. Zu seinen Versuchen benutzte er kleinste im Wasser verteilte Goldteilchen. Es ergaben sich bei der Beobachtung eines Raumteiles während gewisser Zeitmomente die folgenden Ergebnisse: 112 mal wurde kein Teilchen, 168 mal wurde 1 Teilchen, 130 mal wurden 2 Teilchen, 69 mal wurden 3 Teilchen, 32 mal wurden 4 Teilchen, fünfmal wurden 5 Teilchen, einmal 6 Teilchen und einmal 7 Teilchen beobachtet.
Der Anteil dieser oder jener Anzahl beobachteter Teilchen ist daher gleich:

$$\begin{array}{ll}
\text{0 Teilchen: } \frac{112}{518} \approx 0,216; & \text{4 Teilchen: } \frac{32}{518} \approx 0,062; \\
\text{1 Teilchen: } \frac{168}{518} \approx 0,325; & \text{5 Teilchen: } \frac{5}{518} \approx 0,010; \\
\text{2 Teilchen: } \frac{130}{518} \approx 0,251; & \text{6 Teilchen: } \frac{1}{518} \approx 0,002; \\
\text{3 Teilchen: } \frac{69}{518} \approx 0,133; & \text{7 Teilchen: } \frac{1}{518} \approx 0,002.
\end{array}$$

Die so erhaltenen Beobachtungsergebnisse stimmten mit den theoretisch vorhergesagten Wahrscheinlichkeiten sehr gut überein.

1.2 Unmögliche und sichere Ereignisse

Die Wahrscheinlichkeit eines Ereignisses ist offenbar immer eine positive Zahl oder gleich Null. Sie kann aber auch nicht größer als Eins sein, weil in dem Bruch, der die Wahrscheinlichkeit bestimmt, der Zähler nie größer als der Nenner sein kann (die Anzahl der "günstigen" Operationen kann ja nie größer als die Anzahl aller vorgenommenen Operationen sein).

Wir wollen die Wahrscheinlichkeit eines Ereignisses A mit $P(A)$ bezeichnen. Für jedes Ereignis ist also

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

Je größer $P(A)$ ist, desto öfter tritt das Ereignis ein. Je größer z.B. bei einem Schützen die Treffwahrscheinlichkeit ist, desto öfter werden bei ihm günstige Schüsse vorkommen, desto höher wird seine Meisterschaft sein.

Ist die Wahrscheinlichkeit eines Ereignisses sehr klein, so wird es auch sehr selten eintreten. Im Falle $P(A) = 0$ wird das Ereignis A entweder niemals oder so selten eintreten, dass man

[3] Das heißt, die sich in einem Zustand indifferenten Gleichgewicht befinden.

es praktisch als unmöglich ansehen kann. Liegt $P(A)$ dagegen dicht bei Eins, so ist in dem Bruch, der diese Wahrscheinlichkeit angibt, der Zähler nur wenig kleiner als der Nenner, d. h., die weitaus meisten der Operationen sind "günstige Fälle". Ein solches Ereignis tritt also in der Mehrzahl aller Fälle ein.
Wenn $P(A) = 1$ ist, tritt das Ereignis A immer oder fast immer ein, so dass man es praktisch, wie man sagt, als "sicher" ansehen kann, d. h., dass man sicher mit seinem Eintreten rechnen kann.

Ist $P(A) = \frac{1}{2}$, so wird das Ereignis A etwa in der Hälfte aller Fälle eintreten. Das bedeutet also, dass es etwa genauso viel "günstige" wie "ungünstige" Operationen gibt.
Im Falle $P(A) > \frac{1}{2}$ haben wir mehr "günstige" als "ungünstige" Operationen, und für $P(A) < \frac{1}{2}$ ergibt sich genau das umgekehrte Bild.

Wie klein muss nun die Wahrscheinlichkeit eines Ereignisses sein, damit man es praktisch als unmöglich annehmen kann?
Im allgemeinen kann man auf diese Frage keine Antwort geben, weil alles von der Wichtigkeit des Ereignisses, von dem die Rede ist, abhängt. Eine kleine Zahl ist z.B. 0,01. Wenn wir etwa eine Serie von Geschossen haben, von der die Wahrscheinlichkeit, dass ein Geschoss beim Aufschlag nicht explodiert, gleich 0,01 ist, so bedeutet das, dass circa 1 % der Schüsse keine Wirkung zeigt.
Hiermit kann man sich durchaus zufrieden geben. Hat man jedoch einen Fallschirm vor sich, bei dem die Wahrscheinlichkeit, dass er sich beim Absprung nicht öffnet, 0,01 ist, so darf man sich damit keineswegs zufrieden geben. Das würde nämlich bedeuten, dass bei jeweils einem von hundert Absprüngen das kostbare Leben eines Fallschirmspringers vernichtet wird.

Diese Beispiele zeigen, dass wir erst in jeder Aufgabe vorher auf Grund praktischer Überlegungen feststellen müssen, wie klein die Wahrscheinlichkeit eines Ereignisses sein muss, damit man ohne Schaden mit seinem Auftreten nicht zu rechnen braucht.

1.3 Eine Aufgabe

Aufgabe. Ein Schütze schafft 80 % Treffer und ein anderer (unter denselben Schussbedingungen) 70 %. Wie groß wird nun die Treffwahrscheinlichkeit, wenn beide Schützen gleichzeitig auf das Ziel schießen? Das Ziel gilt als getroffen, wenn wenigstens eine der beiden Kugeln ein Treffer war.

Erstes Lösungsverfahren. Wir nehmen an, dass 100 Doppelschüsse abgegeben werden. Bei etwa 80 von ihnen wird das Ziel vom ersten Schützen getroffen werden. Es bleiben also rund 20 Schüsse, die bei diesem Schützen vorbeigehen.
Da der zweite Schütze von 100 Schüssen im Mittel 70 mal trifft, also von 10 Schüssen etwa 7 mal, können wir erwarten, dass unter jenen 20 Fehlschüssen des ersten etwa 14 Treffer des zweiten Schützen zu finden sind. Somit wird also das Ziel bei 100 Schüssen etwa $80 + 14 = 94$ mal getroffen werden.
Bei gleichzeitigem Schießen unserer beiden Schützen wird die Treffwahrscheinlichkeit daher gleich 94 % oder 0,94.

Zweites Lösungsverfahren. Wir nehmen wiederum an, dass 100 Doppelschüsse abgegeben werden. Oben wurde schon festgestellt, dass der erste Schütze etwa 20 mal am Ziel vorbeischießt. Da der zweite Schütze von 100 Schüssen etwa 30 und somit von 10 etwa 3 vorbeischießt, kann man erwarten, dass unter den 20 Fehlschüssen des ersten rund 6 Schüsse vorkommen werden,

in denen auch der zweite fehlt.
Bei jedem dieser sechs Schüsse wird also das Ziel ungetroffen bleiben, während bei jedem der übrigen 94 Schüsse wenigstens einer der beiden Schützen ein günstiges Ergebnis erzielt und somit das Ziel trifft. Wir kommen somit wiederum zu dem Schluss, dass das Ziel beim Doppelschuss in etwa 94 von 100 Fällen getroffen werden wird, d. h., dass die Treffwahrscheinlichkeit 94 % oder 0,94 wird.

Die eben von uns betrachtete Aufgabe war sehr einfach. Sie liefert uns aber trotzdem eine sehr wichtige Schlussfolgerung:
Diese Fälle, in denen es möglich ist, aus den Wahrscheinlichkeiten einzelner Ereignisse die Wahrscheinlichkeiten anderer komplizierterer Ereignisse zu berechnen, sind für die Wahrscheinlichkeitsrechnung von erstrangiger Bedeutung.
Sie kommen in der Tat sogar sehr oft vor und nicht nur im Kriegswesen, sondern auch in jeder Wissenschaft sowie bei jeder praktischen Tätigkeit, die sich mit Massenerscheinungen befasst.

Es wäre nun allerdings sehr unbequem, müsste man für jede vorkommende Aufgabe der beschriebenen Art ein besonderes Lösungsverfahren suchen. Die Wissenschaft strebt immer danach, allgemeine Regeln aufzustellen, deren Kenntnis die mechanische oder fast mechanische Lösung einzelner ähnlicher Aufgaben ermöglicht.

Die Wissenschaft, die sich mit der Aufstellung solcher Regeln für Massenerscheinungen befasst, heißt Wahrscheinlichkeitsrechnung. In diesem Buche sollen die ersten Begriffe dieser Wissenschaft dargelegt werden.

Die Wahrscheinlichkeitsrechnung ist eine der Disziplinen der Mathematik, die Ähnlichkeiten mit der Arithmetik, der Analysis bzw. Geometrie aufzuweisen hat. Ihr Weg ist daher der Weg exakter Überlegungen, und als Hilfsmittel dienen ihr Formeln, Tabellen, Zeichnungen usw.

2 Das Additionstheorem für Wahrscheinlichkeiten

2.4 Die Ableitung des Additionstheorems für Wahrscheinlichkeiten

Die einfachste und wichtigste Regel, die bei der Berechnung von Wahrscheinlichkeiten benutzt wird, ist das Additionstheorem.
Mit ihm wollen wir uns jetzt beschäftigen.

Für einen Schuss aus einer bestimmten Entfernung auf eine Zielseheibe, wie sie in Abb. 1 dargestellt ist, ergeben sich für verschiedene Schützen verschiedene Wahrscheinlichkeiten, in eines der Gebiete 1, 2, 3, 4, 5, 6 zu treffen.

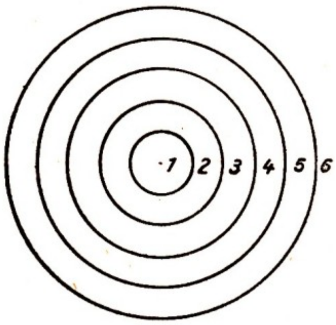


Abb. 1

Es möge einmal für einen Schützen die Wahrscheinlichkeit, in das Gebiet 1 zu treffen, gleich 0,24, und für das Gebiet 2 gleich 0,17 sein. Wie wir schon wissen, bedeutet das, dass von 100

Schüssen, die von diesem Schützen abgegeben werden, (im Mittel) 24 Kugeln in das Gebiet 1 und 17 Kugeln in das Gebiet 2 gelangen werden.

Es möge nun bei einem Schießwettbewerb ein Treffer im Gebiet 1 mit dem Prädikat "ausgezeichnet" und ein Treffer im Gebiet 2 mit dem Prädikat "gut" bedacht werden. Wie groß ist die Wahrscheinlichkeit dafür, dass ein Schuss unseres Schützen das Prädikat gut oder das Prädikat ausgezeichnet erhält?

Diese Frage ist leicht zu beantworten. Von je 100 von dem Schützen abgegebenen Kugeln fallen etwa 24 in das Gebiet 1 und etwa 17 in das Gebiet 2. Somit werden von je 100 Kugeln etwa $24 + 17 = 41$ entweder in das Gebiet 1 oder in das Gebiet 2 fallen. Die gesuchte Wahrscheinlichkeit ist daher gleich $0{,}24 + 0{,}17 = 0{,}41$. Die Wahrscheinlichkeit dafür, dass ein Schuss entweder ausgezeichnet oder gut wird, ist folglich gleich der Summe der Wahrscheinlichkeiten der ausgezeichneten und der guten Schüsse.

Wir betrachten noch ein anderes Beispiel. Ein Fahrgast wartet auf die Straßenbahn Nr. 26 oder Nr. 16 an einer Haltestelle, an der vier Linien vorbeikommen: die Nummern 16, 22, 26, 31.
Wir nehmen an, dass im Mittel alle Linien gleich oft verkehren. Gesucht ist die Wahrscheinlichkeit dafür, dass die erste an der Haltestelle haltende Straßenbahn eine der von dem Fahrgast benötigten Linien ist.

Es ist klar, dass die Wahrscheinlichkeit dafür, dass als erste Straßenbahn die Linie Nr. 16 vorbeikommt, gleich $\frac{1}{4}$ ist. Dieselbe Wahrscheinlichkeit ergibt sich auch für die Straßenbahnlinie Nr. 26. Die gesuchte Wahrscheinlichkeit wird damit offenbar gleich $\frac{1}{2}$, weil

$$\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

ist. Wir können also sagen, dass die Wahrscheinlichkeit dafür, dass als erste Straßenbahn die Linie Nr. 16 oder Nr. 26 ankommt, gleich der Summe der Wahrscheinlichkeiten des Erscheinens der Linie Nr. 16 und der Linie Nr. 26 ist.

Wir wollen jetzt die allgemeine Überlegung durchführen.
Bei der Durchführung einer gewissen Massenoperation sei festgestellt worden, dass in jeder Serie von b Einzeloperationen im Mittel

a_1 mal ein gewisses Resultat A_1,
a_2 mal ein gewisses Resultat A_2,
a_3 mal ein gewisses Resultat A_3,

usw. beobachtet wurde. Oder anders ausgedrückt:

die Wahrscheinlichkeit des Ereignisses A_1 ist gleich $\frac{a_1}{b}$,
die Wahrscheinlichkeit des Ereignisses A_2 ist gleich $\frac{a_2}{b}$,
die Wahrscheinlichkeit des Ereignisses A_3 ist gleich $\frac{a_3}{b}$, usw.

Wie groß ist nun die Wahrscheinlichkeit, dass in einer gewissen Einzeloperation irgendeines (ganz gleich, welches) der Ereignisse $A_1, A_2, A_3, ...$ beobachtet wird?

Man kann das uns interessierende Ereignis "A_1 oder A_2 oder A_3 oder ..."[4] nennen. Dieses Ereignis wird in einer Serie von b Einzeloperationen $a_1 + a_2 + a_3 + ...$-mal beobachtet. Somit

[4] Hier und in allen anderen ähnlichen Fällen bedeuten die Punkte so viel wie "und so weiter".

ist die gesuchte Wahrscheinlichkeit gleich

$$\frac{a_1 + a_2 + a_3 + ...}{b} = \frac{a_1}{b} + \frac{a_2}{b} + \frac{a_3}{b} + ...$$

Diesen Sachverhalt kann man auch durch die folgende Formel ausdrücken:

$$P(A_1 \text{ oder } A_2 \text{ oder } A_3 \text{ oder } ...) == P(A_1) + P(A_2 + P(A_3) + ...$$

Wie in unseren Beispielen setzen wir auch hier in den allgemeinen Überlegungen stets voraus, dass jeweils zwei beliebige der betrachteten Ereignisse (z. B. A_1 und A_2) unverträglich miteinander sind, d. h. in einer einzigen Operation nicht gleichzeitig vorkommen können. Im Beispiel der ankommenden Straßenbahn kann nicht gleichzeitig eine benötigte und eine nichtbenötigte Linie erscheinen. Sie fährt entweder zum Fahrtziel des Wartenden oder nicht.

Die Voraussetzung, dass die einzelnen interessierenden Ereignisse unverträglich sind, ist äußerst wichtig. Berücksichtigt man sie nicht, so wird das aufgestellte Additionstheorem falsch, und seine Anwendung führt zu schwerwiegenden Fehlern. Betrachten wir z. B. noch einmal die Aufgabe, die wir am Schluss des letzten Paragraphen lösten.
Dort ermittelten wir die Wahrscheinlichkeit dafür, dass beim Doppelschuss wenigstens einer der beiden Schützen das Ziel trifft, wobei die Treffwahrscheinlichkeit für den ersten Schützen gleich 0,8 und für den zweiten gleich 0,7 war.

Wenn wir nun zur Lösung dieser Aufgabe das Additionstheorem verwenden würden, fänden wir unmittelbar, dass die gesuchte Wahrscheinlichkeit gleich $0{,}8 + 0{,}7 = 1{,}5$ wäre.
Das Ergebnis ist aber ganz offensichtlich Unsinn, weil wir schon wissen, dass die Wahrscheinlichkeit eines Ereignisses nicht größer als Eins sein kann.

Zu dieser falschen und sinnlosen Antwort kamen wir eben dadurch, dass wir das Additionstheorem auf einen Fall anwandten, auf den es gerade nicht angewandt werden darf. Jene zwei Ereignisse, von denen in dieser Aufgabe die Rede war, sind durchaus miteinander verträglich. Es ist nämlich ohne weiteres möglich, und es muss sogar vorkommen, dass das Ziel bei ein und demselben Doppelschuss von beiden Schützen getroffen wird.

Ein sehr wesentlicher Teil der Fehler, die anfangs bei der Berechnung von Wahrscheinlichkeiten gemacht wurden, beruhte gerade auf solchen falschen Anwendungen des Additionstheorems. Man muss sich daher unbedingt vor diesem Fehler schützen, indem man vor jeder Anwendung des Additionstheorems sehr sorgfältig nachprüft, ob auch wirklich von den Ereignissen, auf die wir es anwenden wollen, je zwei miteinander unverträglich sind.

Wir sind jetzt in der Lage, das Additionstheorem allgemein zu formulieren:

Additionstheorem. Die Wahrscheinlichkeit für das Eintreffen irgendeines (ganz gleich welches) der Ereignisse A_1, A_2, A_3, ..., A_n in einer gewissen Operation ist gleich der Summe der Wahrscheinlichkeiten dieser Ereignisse, wenn jeweils zwei von ihnen miteinander unverträglich sind.

2.5 Das vollständige System von Ereignissen

Bei der dritten Staatsanleihe zur Wiederherstellung und Entwicklung der Volkswirtschaft werden während der zwanzigjährigen Laufzeit ein Drittel der Obligationen gewinnen und die restlichen zwei Drittel gezogen und zum Nominalwert getilgt. Jede Obligation dieser Anleihe hat

mit anderen Worten offenbar eine Wahrscheinlichkeit, zu gewinnen, die $\frac{1}{3}$ ist, und die Wahrscheinlichkeit $\frac{2}{3}$ gezogen zu werden.
Das Gewinnen und das Gezogenwerden sind entgegengesetzte Ereignisse, d. h. zwei Ereignisse, von denen unbedingt eins und nur eins für jede der Obligationen eintreten kann. Die Summe ihrer Wahrscheinlichkeiten ist

$$\frac{1}{3}+\frac{2}{3}=1$$

Das ist aber keine zufällige Tatsache. Wenn A_1 und A_2 zwei entgegengesetzte Ereignisse sind, und wenn in einer Serie von b Operationen das Ereignis A_1 genau a_1 mal und das Ereignis A_2 genau a_2 mal beobachtet wird, dann ist offenbar $a_1 + a_2 = b$. Also

$$P(A_1)=\frac{a_1}{b} \qquad , \qquad P(A_2)=\frac{a_2}{b}$$

so dass

$$P(A_1)+P(A_2)=\frac{a_1}{b}+\frac{a_2}{b}=\frac{a_1+a_2}{b}=1$$

Dasselbe Resultat kann man auch aus dem Additionstheorem erhalten: Da entgegengesetzte Ereignisse auch unverträglich sind, gilt

Das Ereignis "A_1 oder A_2" ist aber das sichere Ereignis; denn aus der Definition der entgegengesetzten Ereignisse folgt, dass es unbedingt eintreten muss. Daher wird die Wahrscheinlichkeit dieses Ereignisses gleich Eins, und wir erhalten wiederum

$$P(A_1)+P(A_2)=1$$

Die Summe der Wahrscheinlichkeiten zweier entgegengesetzter Ereignisse ist gleich Eins.

Diese Regel gestattet noch eine äußerst wichtige Verallgemeinerung, die man mit Hilfe desselben Verfahrens beweisen kann. Wir mögen n (beliebige Anzahl) Ereignisse A_1, A_2, ..., A_n haben. Diese Ereignisse mögen so beschaffen sein, dass in jeder einzelnen Operation bestimmt eine und nur eins dieser Ereignisse eintritt.

Wir vereinbaren, eine solche Gruppe von Ereignissen ein vollständiges System zu nennen. Insbesondere bildet jedes Paar entgegengesetzter Ereignisse ein solches vollständiges System.

Die Summe der Wahrscheinlichkeiten von Ereignissen, die ein vollständiges System bilden, ist gleich Eins.

In der Tat sind auf Grund der Definition eines vollständigen Systems jeweils zwei beliebige Ereignisse dieses Systems miteinander unverträglich. Also folgt aus dem Additionstheorem:

$$P(A_1)+P(A_2)+...+P(A_n)=P(A_1 \text{ oder } A_2 \text{ oder ... oder } A_n)$$

Auf der rechten Seite dieser Gleichung steht die Wahrscheinlichkeit des sicheren Ereignisses; sie ist daher gleich Eins. Also gilt für ein vollständiges System

$$P(A_1)+P(A_2)+...+P(A_n)=1$$

was zu beweisen war.

Beispiel 1. Von je 100 Schüssen auf die in der Abb. 1 dargestellte Scheibe erzielt ein Schütze im Mittel

44 Treffer im Gebiet 1, 30 Treffer im Gebiet 2, 15 Treffer im Gebiet 3,

6 Treffer im Gebiet 4, 4 Treffer im Gebiet 5, 1 Treffer im Gebiet 6.

$(44 + 30 + 15 + 6 + 4 + 1 = 100)$. Diese sechs Schussresultate bilden offensichtlich ein vollständiges System. Ihre Wahrscheinlichkeiten sind dann

$$0,44; 0,30; 0,15; 0,06; 0,04; 0,01;$$

und es ist

$$0,44 + 0,30 + 0,15 + 0,06 + 0,04 + 0,01 = 1$$

Die Kugeln, die in das Gebiet 6 fallen, d. h. die Scheibe nicht treffen, brauchen gar nicht gezählt zu werden. Das hindert uns jedoch nicht, die Wahrscheinlichkeit der Treffer in diesem Gebiet zu finden, weil wir zu diesem Zweck nur die Summe der Wahrscheinlichkeiten der Treffer in den anderen Gebieten von Eins subtrahieren müssen.

Beispiel 2. Aus einer Statistik geht hervor, dass in einer gewissen Weberei von hundert Störungen am Webstuhl, die zusätzliche Arbeiten der Weberin erfordern, im Mittel

22 entstehen durch Kettengarnriss,
31 entstehen durch Einschussgarnriss,
27 entstehen durch Schiffchenwechsel,
3 entstehen durch Treiberbruch

und die restlichen Störungen des Webstuhls aus anderen Ursachen entstehen.
Wir sehen, dass wir außer den anderen Ursachen vier bestimmte Gründe für die Störungen des Webstuhls haben; deren Wahrscheinlichkeiten sind: 0,22; 0,31; 0,27; 0,03.

Die Summe dieser Wahrscheinlichkeiten ist gleich 0,83. Zusammen mit den anderen Ursachen bilden die aufgezählten Gründe der Störungen des Webstuhls ein vollständiges System von Ereignissen. Die Wahrscheinlichkeit der Störung des Stuhls aus anderen Ursachen ist daher gleich $1 - 0,83 = 0,17$.

2.6 Beispiele

Auf unseren Satz über ein vollständiges System von Ereignissen gründet man oft mit Erfolg sogenannte Wahrscheinlichkeiten "a priori", d. h. vor der Erfahrung berechnete Wahrscheinlichkeiten.
Es möge z.B. eine kleine rechteckige Fläche (Abb. 2) einem Artilleriebeschuss unterzogen werden.

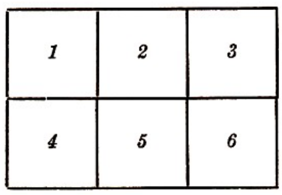


Abb. 2

Diese Fläche sei in sechs gleiche Quadrate aufgeteilt, die in der Zeichnung durchnummeriert wurden.
Meistens befinden sich die uns interessierenden Flächen unter gleichen Beschussbedingungen. Daher ist es auch nicht verwunderlich, dass wir annehmen, keines dieser sechs Quadrate werde öfter als ein anderes getroffen.

Wir nehmen daher ganz naturgemäß an, dass im Mittel jedes der Quadrate gleich oft getroffen wird, d. h., dass die Treffwahrscheinlichkeiten $p_1, p_2, p_3, p_4, p_5, p_6$ dieser Quadrate gleich sind. Nimmt man nun noch an, dass bei jedem Schuss unbedingt ein Quadrat getroffen wird, dann folgt hieraus, dass jede der Zahlen p gleich $\frac{1}{6}$ ist. Sie müssen nämlich untereinander gleich sein, und ihre Summe muss nach dem von uns bewiesenen Satz Eins ergeben.

Diese Aussage, die auf einer Reihe von Annahmen fußt, verlangt natürlich eine Bestätigung durch die praktische Erfahrung. Durch ähnliche Fälle sind wir jedoch an einen positiven Ausgang des Experimentes gewöhnt, und wir verlassen uns auch hier aus Zweckmäßigkeitsgründen auf unsere theoretischen Annahmen und ihre Kontrolle durch die Erfahrung.

Man sagt gewöhnlich in solchen Fällen, dass die gegebene Operation aus n verschiedenen gleichwahrscheinlichen Ereignissen (so wird in unserem Beispiel ein Schuss auf die in der Abb. 2 dargestellte Fläche als Ergebnis einen Treffer in einem der sechs Quadrate liefern) besteht. Die Wahrscheinlichkeit eines jedes einzelnen dieser Ereignisse ist in diesem Falle gleich $\frac{1}{n}$.

Die Wichtigkeit solcher "apriorischer" Berechnungen besteht gerade darin, dass sie in vielen Fällen die Vorausbestimmung der Wahrscheinlichkeit eines Ereignisses unter Bedingungen gestatten, unter denen die Durchführung der Massenerscheinungen entweder gar nicht möglich oder außerordentlich schwierig ist.

Beispiel 1. Auf den Obligationen unserer Staatsanleihe werden die Seriennummern gewöhnlich durch fünfstellige Zahlen dargestellt.
Wir wollen die Wahrscheinlichkeit dafür bestimmen, dass die letzte Ziffer einer auf gut Glück genommenen gewinnenden Serie gleich 7 ist (wie z.B. in der Serie Nr. 59607).

Gemäß der Definition der Wahrscheinlichkeit müssen wir zu diesem Zweck die lange Reihe der Gewinnlisten hernehmen und nachzählen, wie viele der gewinnenden Serien eine auf 7 endigende Nummer haben.
Der Quotient dieser Anzahl zur Gesamtzahl der gewinnenden Serien wird dann die gesuchte Wahrscheinlichkeit sein. Wir haben jedoch allen Grund, anzunehmen, dass jede der zehn Ziffern $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ genauso gut an letzter Stelle der Nummer einer gewinnenden Serie stehen kann wie irgendeine andere.
Daher sprechen wir ohne jedes Zögern die Vermutung aus, dass die gesuchte Wahrscheinlichkeit gleich 0,1 ist.

Sie können leicht die Richtigkeit dieser theoretischen Vorhersage nachprüfen; führen Sie alle notwendigen Zahlungen in einer Gewinnliste durch und überzeugen Sie sich, dass wirklich jede der zehn Ziffern in etwa 1/10 aller Fälle an letzter Stelle steht.

Beispiel 2. Eine Telephonleitung, die zwei im Abstand von 2 km befindliche Punkte A und B miteinander verbindet, sei an einer unbekannten Stelle zerrissen.
Wie groß ist nun die Wahrscheinlichkeit, dass sich die Schadenstelle nicht weiter als 450 m vom Punkte A entfernt befindet?

Die ganze Leitung zerlegen wir in Gedanken in einzelne Meter. Wegen der tatsächlichen Homogenität aller dieser Teilstücke können wir annehmen, dass die Wahrscheinlichkeit der Beschädigung für jeden Meter dieselbe ist. Ähnlich dem vorhergehenden ergibt sich hieraus leicht die gesuchte Wahrscheinlichkeit zu

$$\frac{450}{2000} = 0,225$$

3 Bedingte Wahrscheinlichkeiten und die Multiplikationsregel

3.7 Der Begriff der bedingten Wahrscheinlichkeit

In zwei Werken mögen Glühlampen hergestellt werden. Das erste der beiden Werke liefere 70 % und das zweite 30 % der gesamten erforderlichen Produktion. Im Mittel sind von je 100 Lampen des ersten Werkes 83 und von 100 Lampen des zweiten Werkes nur 63 normgerecht.[5]

Wie man leicht aus diesen Angaben ausrechnet, werden im Mittel von je 100 Glühlampen, die die Verbraucher kaufen, 77 normgerecht[6] sein.
Somit wird die Wahrscheinlichkeit, eine Normallampe zu kaufen, gleich 0,77 sein. Wir nehmen jetzt einmal an, dass wir erfahren haben, dass die Glühlampen, die sich in einem Warenhaus befinden, alle im ersten Werk hergestellt wurden. Dann wird sich folglich auch die Wahrscheinlichkeit, eine normgerechte Lampe zu kaufen, ändern. Sie wird nun

$$\frac{83}{100} = 0,83$$

Das betrachtete Beispiel zeigt, dass die Gesamtbedingungen, unter denen die Operation (in unserem Beispiel der Lampenkauf) vor sich geht, durch eine neue wesentliche Bedingung (in unserem Beispiel die Kenntnis des Herstellungsortes der Lampen) ergänzt werden.

Diese kann die Wahrscheinlichkeit des Ausganges einer Einzeloperation merklich ändern. Das ist auch klar: Erforderte doch die Definition des Wahrscheinlichkeitsbegriffes eine genaue Bestimmung der Bedingungen, unter denen die vorgegebene Massenoperation abläuft.

Fügt man nun zu dieser Gesamtheit eine neue Bedingung hinzu, so wird sich diese Gesamtheit im allgemeinen wesentlich verändern. Also wurde unsere Massenoperation schon unter neuen Bedingungen durchgeführt. Faktisch ist das aber schon eine völlig neue Operation, und die Wahrscheinlichkeit eines Ereignisses wird nicht mehr dieselbe bleiben wie unter den ursprünglichen Bedingungen.

Wir haben also zwei verschiedene Wahrscheinlichkeiten für ein und dasselbe Ereignis, das Kaufen einer Normallampe. Diese Wahrscheinlichkeiten wurden aber unter verschiedenen Bedingungen berechnet:
Solange wir nicht die zusätzliche Bedingung berücksichtigen (den Herstellungsort der Glühlampe), haben wir die unbedingte Wahrscheinlichkeit des Kaufens einer Normallampe zu 0,77. Jedoch bei Berücksichtigung der zusätzlichen Bedingung (dass die Glühlampe im ersten Werk hergestellt wurde) erhalten wir die bedingte Wahrscheinlichkeit 0,83, die von der vorhergehenden verschieden ist.

Wenn wir mit A das Ereignis des Kaufens einer Normallampe und mit B das Ereignis, dass sie im ersten Werk hergestellt wurde, bezeichnen, dann werden gewöhnlich mit $P(A)$ die unbedingte Wahrscheinlichkeit des Ereignisses A und mit $P(A|B)$ die Wahrscheinlichkeit desselben Ereignisses unter der Voraussetzung des Ereignisses B (d. h., dass die Lampe aus dem ersten Werk kommt) bezeichnen. Wir haben somit

$$P(A) = 0,77 \qquad , \qquad P(A|B) = 0,83$$

[5] Wir werden eine Glühlampe normgerecht oder normal nennen, wenn sie eine Brenndauer von wenigstens 1200 Brennstunden hat. Andernfalls werden wir diese Lampe als Ausschuss bezeichnen.
[6] In der Tat ist $0,83 \cdot 70 + 0,63 \cdot 30 = 77$.

Da man von der Wahrscheinlichkeit irgendeines Ereignisses nur unter gewissen genau bestimmten Voraussetzungen sprechen kann, ist streng genommen jede Wahrscheinlichkeit eine bedingte Wahrscheinlichkeit.

Eine unbedingte Wahrscheinlichkeit kann im eigentlichen Sinne dieses Wortes gar nicht existieren. In der Mehrzahl der konkreten. Aufgaben haben wir es jedoch damit zu tun, dass allen zu untersuchenden Operationen in einer Aufgabe ein ganz bestimmtes System K von Bedingungen zugrunde liegt, das für alle Operationen als erfüllbar vorausgesetzt wird. Werden nun bei der Berechnung irgendeiner Wahrscheinlichkeit keine anderen Bedingungen berücksichtigt, so nennen wir eine solche Wahrscheinlichkeit eine unbedingte.

Wir nennen aber eine Wahrscheinlichkeit bedingt, wenn sie unter der Voraussetzung berechnet wurde, dass außer dem für alle Operationen gemeinsamen System K von Bedingungen noch irgendwelche anderen genau festgelegten ergänzenden Bedingungen erfüllt werden.

So setzten wir in unserem Beispiel etwa voraus, dass die Glühlampenproduktion unter gewissen bestimmten Bedingungen vor sich geht, die für alle im Handel befindlichen Lampen dieselben sind. Diese Voraussetzung ist unumgänglich und so selbstverständlich, dass wir es bei der Formulierung der Aufgabe gar nicht für notwendig erachteten, sie zu erwähnen.
Würden wir einer Lampe keinerlei zusätzliche Bedingungen auferlegen, so wäre die Wahrscheinlichkeit dieses oder jenes Resultates einer Lampenprüfung eine unbedingte. Verlangen wir aber noch irgendwelche zusätzlichen Forderungen, so wird die unter diesen Forderungen berechnete Wahrscheinlichkeit schon bedingt.

Beispiel 1. In der Aufgabe, die von uns am Anfang dieses Paragraphen beschrieben wurde, ist die Wahrscheinlichkeit einer Lampenlieferung aus dem zweiten Werk offenbar gleich 0,3. Es wurde festgestellt, dass die Glühlampen von normaler Qualität waren. Wie groß ist nun die Wahrscheinlichkeit dafür, dass diese Lampen im zweiten Werk angefertigt worden sind?

Von jeweils 1000 im Handel befindlichen Glühlampen sind im Mittel 770 normgerecht, wobei 581 von ihnen aus dem ersten und. 189 aus dem zweiten Werk stammen[7].
Nach der gemachten Beobachtung ist die Wahrscheinlichkeit, eine normgerechte Lampe des zweiten Werkes zu kaufen, daher

$$\frac{189}{770} \approx 0,245$$

Das ist die bedingte Wahrscheinlichkeit des Kaufes einer Lampe des zweiten Werkes, die unter der Voraussetzung berechnet wurde, dass die betreffende Lampe normgerecht ist. In unseren früheren Bezeichnungen können wir auch schreiben:

$$P(\overline{B}) = 0,3 \qquad ; \qquad P(\overline{B}|A) \approx 0,245$$

(das Ereignis $\overline{B}$ bedeutet das Nichteintreten des Ereignisses B).

Beispiel 2. In einem Rayon durchgeführte langjährige Beobachtungen ergaben, dass von 100000 Kindern, die das zehnte Lebensjahr erreichten, im Mittel 82277 das vierzigste und 37977 das siebzigste Lebensjahr erlebten.
Wie groß ist dann die Wahrscheinlichkeit, das siebzigste Lebensjahr zu erreichen, für einen Menschen, der bereits vierzig Jahre ist?

[7] Das lässt sich auf folgende Weise berechnen: Von jeweils 1000 Glühlampen wurden im Mittel 700 im ersten Werk hergestellt, und von je 100 Lampen des ersten Werkes sind im Mittel 83 normgerecht. Somit werden von 700 Glühlampen des ersten Werkes im Mittel $7 \cdot 83 = 581$ normale Qualität haben. Die restlichen normgerechten 189 Glühlampen sind also im zweiten Werk hergestellt worden.

Da von 82277 Vierzigjährigen im Mittel 37977 das siebzigste Jahr erleben, ist die Wahrscheinlichkeit für einen Vierzigjährigen, siebzig Jahre alt zu werden, gleich

$$\frac{37977}{82277} \approx 0,46$$

Bezeichnet man mit A bzw. B die folgenden Ereignisse: erstens - ein Zehnjähriger erlebt das siebzigste Jahr, und zweitens - er erlebt das vierzigste Lebensjahr, so ergibt sich offenbar:

$$P(A) = 0,3797 \approx 0,38 \qquad ; \qquad P(A|B) \approx 0,46$$

3.8 Die Ableitung der Multiplikationsregel für Wahrscheinlichkeiten

Wir halten uns an das erste Beispiel des vorigen Paragraphen.
Von jeweils 1000 Glühlampen wurden im Mittel 300 im zweiten Werk hergestellt, und von diesen 300 Lampen waren im Mittel 189 normgerecht. Hieraus ergibt sich nun für die Fertigung einer Lampe im zweiten Werk (Ereignis B) die Wahrscheinlichkeit

$$P(\overline{B}) = \frac{300}{1000} = 0,3$$

Die Wahrscheinlichkeit ihrer Normalqualität wird damit unter der Voraussetzung der Fertigung im zweiten Werk gleich

$$P(A|\overline{B}) = \frac{189}{300} = 0,63$$

Da von jeweils 1000 Lampen 189 normgerecht waren und gleichzeitig im zweiten Werk hergestellt wurden, wird die Wahrscheinlichkeit des gemeinsamen Eintretens der Ereignisse A und $\overline{B}$ gleich

$$P(A \text{ und } \overline{B}) = \frac{189}{1000} = \frac{300}{1000} \cdot \frac{189}{300} = P(\overline{B}) \cdot P(A|\overline{B})$$

Diese "Multiplikationsregel" kann auch gleich allgemein formuliert werden. Möge das Resultat B in jeder Serie von n Operationen im Mittel m mal eintreten, und in jeder von m derartigen Operationen, in denen das Ereignis B schon eingetreten ist, möge l mal das Ereignis A auftreten. Dann werden in jeder Serie von n Operationen die Ereignisse A und B im Mittel l mal gleichzeitig eintreten. Also

$$P(B) = \frac{m}{n} \qquad , \qquad P(A|B) = \frac{l}{m}$$

$$P(A \text{ und } B) = \frac{l}{n} = \frac{m}{n} \cdot \frac{l}{m} = P(B) \cdot P(A|B) \tag{1}$$

Multiplikationsregel. Die Wahrscheinlichkeit des gleichzeitigen Eintretens zweier Ereignisse ist gleich dem Produkt der Wahrscheinlichkeit des ersten Ereignisses mit der bedingten Wahrscheinlichkeit des zweiten, welche unter der Voraussetzung berechnet wird, dass das erste Ereignis bereits eingetreten ist.

Wir können selbstverständlich jedes der zwei gegebenen Ereignisse als erstes ansehen, denn man kann die Formel (1) aus den gleichen Gründen auch in der Form

$$P(A \text{ und } B) = P(A) \cdot P(B|A) \tag{1'}$$

schreiben. Hieraus erhalten wir noch die wichtige Beziehung:

$$P(A) \cdot P(B|A) = P(B) \cdot P(A|B) \tag{2}$$

In unserem Beispiel war

$$P(A \text{ und } \overline{B}) = \frac{189}{1000}, \qquad P(A) = \frac{77}{100}, \qquad P(\overline{B}|A) = \frac{189}{700}$$

in Übereinstimmung mit Formel (1').

Beispiel. In einem Betrieb erwiesen sich 96 % der hergestellten Erzeugnisse als brauchbar (Ereignis A). Von jeweils 100 brauchbaren Erzeugnissen waren im Mittel 75 Güteklasse 1 (Ereignis B). Es ist die Wahrscheinlichkeit dafür zu bestimmen, dass ein in diesem Betrieb hergestelltes Erzeugnis zur Güteklasse 1 gehört.

Wir müssen also die Wahrscheinlichkeit $P(A \text{ und } B)$ berechnen; denn damit das Erzeugnis erstklassig ist, muss es erstens brauchbar (Ereignis A) und zweitens aus der Güteklasse 1 (Ereignis B) sein.
Aus den Voraussetzungen ergibt sich

$$P(A) = 0,96 \qquad ; \qquad P(B|A) = 0,75$$

Daher wird auf Grund der Formel (1')

$$P(A \text{ und } B) = 0,96 \cdot 0,75 = 0,72$$

3.9 Unabhängige Ereignisse

Bei Festigkeitsversuchen an zwei auf verschiedenen Maschinen hergestellten Garndocken zeigte es sich, dass ein Muster einer gewissen Länge der ersten Garndocke eine bestimmte Normbelastung mit der Wahrscheinlichkeit 0,84 und der zweiten Docke mit der Wahrscheinlichkeit 0,78 aushielt.[8]
Wie groß ist nun die Wahrscheinlichkeit dafür, dass zwei Garnmuster, die aus zwei verschiedenen Docken genommen wurden, imstande sind, die Normbelastung auszuhalten ?

Mit A bezeichnen wir das Ereignis, dass das aus der ersten Garndocke genommene Muster die Normbelastung aushält. Dann sei B das analoge Ereignis für ein Muster aus der zweiten Garndocke. Da $P(A \text{ und } B)$ gesucht wird, wenden wir die Multiplikationsregel an:

$$P(A \text{ und } B) = P(A) \cdot P(B[A)$$

Offenbar ist hier $P(A) = 0,84$. Was ist aber $P(B|A)$?
Gemäß der allgemeinen Definition der bedingten Wahrscheinlichkeit ist das die Wahrscheinlichkeit dafür, dass das Garnmuster aus der zweiten Docks die Normbelastung aushält, nachdem das Muster aus der ersten Docke diese Belastung schon überstanden hatte.

Die Wahrscheinlichkeit des Ereignisses B hängt aber nicht vom Eintreten oder Nichteintreten des Ereignisses A ab; denn wir können ja die Versuche gleichzeitig ausführen und die Garnmuster aus vollkommen verschiedenen an verschiedenen Maschinen gefertigten Docken auswählen.
Das bedeutet also praktisch, dass die Anzahl der günstigen Versuche an den Proben aus der zweiten Docke völlig unabhängig davon ist, welche Festigkeit ein beliebiges Muster aus der ersten Docke hat, d. h.

$$P(B|A) = P(B) = 0,78$$

[8] Wenn die Normbelastung sagen wir 400 Gramm ist, so bedeutet das folgendes: Von 100 Mustern aus der ersten Docke halten im Mittel 84 eine solche Belastung aus, 16 überstehen diese Probe nicht und zerreißen.

Hieraus folgt nun

$$P(A \text{ und } B) = P(A) \cdot P(B) = 0,84 \cdot 0,78 = 0,655$$

Die Besonderheit, die dieses Beispiel von allen vorhergehenden unterscheidet, besteht, wie wir sehen, darin, dass sich die Wahrscheinlichkeit des Ereignisses B nicht ändert, auch wenn wir zu den allgemeinen Bedingungen noch eine das Ereignis A betreffende Forderung hinzufügen. Mit anderen Worten, die bedingte Wahrscheinlichkeit $P(B|A)$ ist also gleich der unbedingten Wahrscheinlichkeit $P(B)$.
In einem solchen Fall werden wir kurz sagen, dass das Ereignis B vom Ereignis A unabhängig ist.

Es ist leicht nachzuprüfen, dass, wenn B von A unabhängig ist, auch A von B unabhängig sein muss. In der Tat folgt aus $P(B) = P(B|A)$ wegen der Formel (2) auch $P(A|B) = P(A)$, was gerade die Unabhängigkeit des Ereignisses A vom Ereignis B bedeutet.
Die Unabhängigkeit zweier Ereignisse ist also eine symmetrische Eigenschaft. Hieraus ersehen wir, dass die Multiplikationsregel für unabhängige Ereignisse eine besonders einfache Form annimmt:

$$P(A \text{ und } B) = P(A) \cdot P(B) \tag{3}$$

Man hat hier genauso, wie man vor jeder Anwendung des Additionstheorems die Unverträglichkeit je zweier gegebener Ereignisse feststellen muss, auch vor jeder Anwendung der Regel (3) sich davon zu überzeugen, dass die Ereignisse A und B voneinander unabhängig sind.
Die Nichtbeachtung dieses Hinweises führt ebenfalls zu einer großen Anzahl von Fehlern. Wenn die Ereignisse A und B abhängig sind, ist die Formel (3) nicht richtig und muss durch die Formeln (1) oder (1') ersetzt werden.

Die Regel (3) kann man auch leicht auf allgemeinere Fälle erweitern, so dass man auch die Wahrscheinlichkeit des Eintretens von drei oder mehr unabhängigen Ereignissen berechnen kann.
Mögen wir es z. B. mit drei unabhängigen Ereignissen A, B und C zu tun haben (das bedeutet, dass die Wahrscheinlichkeit jedes dieser Ereignisse davon unabhängig ist, ob die beiden anderen Ereignisse eintreten oder nicht). Aus der Unabhängigkeit der Ereignisse A, B und C folgt nach Regel (3)

$$P(A \text{ und } B \text{ und } C) = P(A \text{ und } B) \cdot P(C)$$

Setzen wir nun hier wieder an Stelle von $P(A \text{ und } B)$ den entsprechenden Ausdruck für die Wahrscheinlichkeit aus Formel (3) ein, ergibt sich

$$P(A \text{ und } B \text{ und } C) = P(A) \cdot P(B) \cdot P(C) \tag{4}$$

Damit ist aber auch schon klar, dass diese Regel auch gilt, wenn die betrachtete Gruppe eine beliebige Anzahl von Ereignissen enthält, nur müssen die Ereignisse voneinander unabhängig sein (d. h. die Wahrscheinlichkeit eines jeden einzelnen muss vom Eintreten oder Nichteintreten der übrigen Ereignisse unabhängig sein).

Die Wahrscheinlichkeit des gleichzeitigen Eintretens einer beliebigen Anzahl unabhängiger Ereignisse ist gleich dem Produkt der Wahrscheinlichkeiten dieser Ereignisse.

Beispiel 1. Ein Arbeiter bediene drei Webstühle. Die Wahrscheinlichkeit dafür, dass ein Webstuhl im Laufe einer Stunde die Aufmerksamkeit des Arbeiters nicht erfordert, sei für den ersten Webstuhl gleich 0,9, für den zweiten 0,8 und für den dritten 0,85.

Es wird nun die Wahrscheinlichkeit dafür gesucht, dass im Laufe einer Stunde keiner der Webstühle die Aufmerksamkeit des Arbeiters in Anspruch nimmt.

Nimmt man an, die Webstühle arbeiteten unabhängig voneinander, so findet man aus Formel (4) die Wahrscheinlichkeit

$$0,9 \cdot 0,8 \cdot 0,85 = 0,612$$

Beispiel 2. Unter den Voraussetzungen des Beispiels 1 ist die Wahrscheinlichkeit dafür zu bestimmen, dass wenigstens einer der drei Webstühle nicht die Aufmerksamkeit des Arbeiters im Laufe einer Stunde beansprucht.

Wir haben es hier in der Aufgabe mit einer Wahrscheinlichkeit der Form $P(A$ oder B oder $C)$ zu tun. Unser erster Gedanke wird natürlich sein, das Additionstheorem anzuwenden. Wir überzeugen uns jedoch sofort davon, dass diese Regel im vorliegenden Fall gar nicht anwendbar ist, weil jeweils zwei der drei zu betrachtenden Ereignisse miteinander vereinbar sind (denn es liegt kein Grund vor, warum nicht zwei Webstühle innerhalb einer Stunde ungestört durcharbeiten sollten).
Ja auch unabhängig von dieser Überlegung sieht man sofort, dass die Summe der drei Wahrscheinlichkeiten Eins bedeutend übersteigt, was natürlich für keine Wahrscheinlichkeit eintreten kann.

Um die gestellte Aufgabe zu lösen, erwähnen wir noch, dass die Wahrscheinlichkeit dafür, dass ein Webstuhl die Aufmerksamkeit des Arbeiters erfordert, für den ersten 0,1, für den zweiten 0,2 und. für den dritten 0,15 ist. Da nun diese drei Ereignisse voneinander unabhängig sind, ergibt sich für die Verwirklichung aller drei Ereignisse die Wahrscheinlichkeit nach Formel (4) zu

$$0,1 \cdot 0,2 \cdot 0,15 = 0,0003$$

Aber die Ereignisse "alle drei Webstühle erfordern die Aufmerksamkeit" und "wenigstens einer arbeitet ruhig" sind doch offensichtlich ein Paar entgegengesetzter Ereignisse. Daher ist die Summe ihrer Wahrscheinlichkeiten gleich Eins, und die gesuchte Wahrscheinlichkeit wird somit $1 - 0,0003 = 0,9997$.
Da aber die Wahrscheinlichkeit dieses Ereignisses so nahe bei Eins liegt, kann man es in der Praxis als sicher annehmen. Das bedeutet also, dass im Laufe einer Stunde fast stets einer der drei Webstühle ohne jede Störung laufen wird.

Beispiel 3. Unter bestimmten Bedingungen ist die Wahrscheinlichkeit, ein feindliches Flugzeug mit einem Schuss aus einem Gewehr abzuschießen, gleich 0,004. Es ist die Wahrscheinlichkeit der Vernichtung eines feindlichen Flugzeuges bei gleichzeitigem Schießen aus 250 Gewehren zu bestimmen.

Bei einem einzelnen Schuss ist die Wahrscheinlichkeit, dass das Flugzeug nicht abstürzt, gleich $1-0,004 = 0,996$. Die Wahrscheinlichkeit, dass es auch nicht bei allen 250 Schüssen abstürzt, wird nach der Multiplikationsregel für unabhängige Ereignisse gleich dem Produkt der 250 Faktoren, von denen jeder gleich 0,996 ist.

Die Wahrscheinlichkeit ist also $(0,996)^{250}$. Für den Fall, dass wenigstens einer der 250 Schüsse den Absturz des Flugzeugs herbeiführt, ergibt sich die Wahrscheinlichkeit

$$1 - (0,996)^{250}$$

Die logarithmische Berechnung, die wir hier aber nicht durchführen werden, lehrt, dass diese Zahl annähernd gleich $\frac{5}{8}$ ist. Obwohl nun die Wahrscheinlichkeit, ein feindliches Flugzeug mit einem Schuss zum Absturz zu bringen, winzig klein ist (0,004), ist die Wahrscheinlichkeit eines positiven Ergebnisses bei gleichzeitigem Beschuss aus einer größeren Anzahl Gewehre doch schon ganz beträchtlich.

Die in den letzten beiden Beispielen benutzte Überlegung lässt sich ohne weiteres verallgemeinern und führt dann auf eine wichtige allgemeine Regel. In beiden Fällen sprachen wir von der Wahrscheinlichkeit $P(A_1$ oder A_2 oder ... oder $A_n)$ des Eintretens von wenigstens einem der voneinander unabhängigen Ereignisse $A_1, A_2, ..., A_n$. Bezeichnen wir mit $\overline{A}_k$ das Nichteintreten von A_k. Dann sind die Ereignisse $\overline{A}_k$ und A_k entgegengesetzte, und es gilt

$$P(A_k) + P(\overline{A}_k) = 1$$

Andererseits sind aber auch offensichtlich die Ereignisse $\overline{A}_1$, $\overline{A}_2$, ..., $\overline{A}_n$ voneinander unabhängig, so dass

$$\begin{aligned} P(\overline{A}_1 \text{ und } \overline{A}_2 \text{ und ... und } \overline{A}_n) &= P(\overline{A}_1) \cdot P(\overline{A}_2) \cdot ... \cdot P(\overline{A}_n) \\ &= [1 - P(A_1)] \cdot [1 - P(A_2)] \cdot ... \cdot [1 - P(A_n)] \end{aligned}$$

Offenbar sind die Ereignisse (A_1 oder A_2 oder A_3 oder ... oder A_n) und ($\overline{A}_1$ und $\overline{A}_2$ und ... und $\overline{A}_n$) ebenfalls zueinander entgegengesetzt (entweder tritt wenigstens eins der Ereignisse A_k oder es treten alle Ereignisse $\overline{A}_k$ ein). Daher wird

$$\begin{aligned} P(A_1 \text{ oder } A_2 \text{ oder ... oder } A_n) &= 1 - P(\overline{A}_1 \text{ und } \overline{A}_2 \text{ und ... und } \overline{A}_n) \\ &= 1 - [1 - P(A_1)] \cdot [1 - P(A_2)] \cdot ... \cdot [1 - P(A_n)] \qquad (5) \end{aligned}$$

Mit Hilfe dieser wichtigen Formel kann man die Wahrscheinlichkeit des Eintretens von wenigstens einem der Ereignisse $A_1, A_2, ..., A_n$ bei vorgegebenen Einzelwahrscheinlichkeiten dann und nur dann berechnen, wenn die einzelnen Ereignisse voneinander unabhängig sind.
In dieser allgemeinen Formel ist auch der Spezialfall enthalten, dass alle Ereignisse A_k die gleiche Wahrscheinlichkeit p haben (wie es beispielsweise im Beispiel 3 der Fall war):

$$P(A_1 \text{ oder } A_2 \text{ oder ... oder } A_n) = 1 - (1-p)^n \qquad (6)$$

Beispiel 4. Beim Schießen auf ein Luftziel unter gewissen Schussbedingungen wird ein Schuss als wirksam anerkannt, wenn der Punkt, in dem das Geschoss explodiert, vom Mittelpunkt des Ziels nicht mehr als 10 m in jeder der drei Richtungen (Länge, Breite, Höhe) entfernt ist. Anders ausgedrückt, das Geschoss muss, um wirksam zu sein, im Innern des Würfels explodieren, dessen Mittelpunkt mit dem Mittelpunkt des Ziels zusammenfällt und dessen Kantenlänge 20 m ist.

Die Wahrscheinlichkeit einer Abweichung von mehr als 10 m sei
in der Länge $p_1 = 0,08$, in der Breite $p_2 = 0,12$, in der Höhe $p_3 = 0,1$.
Gesucht ist die Wahrscheinlichkeit P eines unwirksamen Schusses.

Für die Unwirksamkeit eines Schusses ist es also notwendig, dass der Explosionsort des Geschosses in wenigstens einer der drei Richtungen um mehr als 10 m vom Mittelpunkt des Ziels abweicht. Da diese drei Ereignisse gewöhnlich als voneinander unabhängig angenommen werden können (weil sie von verschiedenen Ursachen herrühren), ist es möglich, zur Lösung der Aufgabe die Formel (5) anzuwenden. Wir erhalten:

$$P = 1 - (1-p_1) \cdot (1-p_2) \cdot (1-p_3) \approx 0,27$$

Folglich können wir annehmen, dass von jeweils 100 Schüssen etwa 73 wirksam sein werden.

4 Folgerungen aus dem Additionstheorem und der Multiplikationsregel

4.10 Die Ableitung einiger Ungleichungen

Wir wenden uns von neuem dem Beispiel mit den Glühlampen aus dem vorigen Kapitel zu. Für die Ereignisse führen wir die folgenden Bezeichnungen ein:

A, wenn die Lampe normgerecht ist,
$\overline{A}$, wenn die Lampe Ausschuss ist,
B, wenn die Lampe aus dem ersten Werk stammt,
$\overline{B}$, wenn die Lampe aus dem zweiten Werk stammt.

Die Ereignisse A und $\overline{A}$ bilden offenbar ein Paar entgegengesetzter Ereignisse. Genau dasselbe ist aber auch bei B und $\overline{B}$ der Fall.
Wenn eine Lampe normgerecht ist (A), so ist sie entweder im ersten Betrieb (A und B) oder im zweiten (A und $\overline{B}$) hergestellt worden. Da aber die letzten beiden Ereignisse offenbar unverträglich sind, gilt nach dem Additionstheorem

$$P(A) = P(A \text{ und } B) + P(A \text{ und } \overline{B}) \tag{1}$$

Auf demselben Wege finden wir auch

$$P(B) = P(A \text{ und } B) + P(\overline{A} \text{ und } B) \tag{2}$$

Betrachten wir jetzt das Ereignis (A oder B). Man hat offensichtlich zu seiner Realisierung die folgenden drei Möglichkeiten:

1. A und B, 2. A und $\overline{B}$, 3. $\overline{A}$ und B.

Von diesen drei Möglichkeiten sind aber jeweils zwei miteinander unverträglich. Daher ist nach dem Additionstheorem

$$P(A \text{ oder } B) = P(A \text{ und } B) + P(A \text{ und } \overline{B}) + P(\overline{A} \text{ und } B) \tag{3}$$

Wir addieren die Gleichungen (1) und (2) und subtrahieren von dieser Summe die Gleichung (3). So finden wir leicht

$$P(A) + P(B) - P(A \text{ oder } B) = P(A \text{ und } B)$$

woraus

$$P(A \text{ oder } B) = P(A) + P(B) - P(A \text{ und } B) \tag{4}$$

folgt. Wir kamen so zu einem sehr wichtigen Resultat.
Obwohl wir die Überlegung für ein spezielles Beispiel durchführten, war sie doch so allgemein, dass die Schlussfolgerung als gültig für ein beliebiges Paar von Ereignissen A und B angenommen werden kann.
Vorher hatten wir den Ausdruck für die Wahrscheinlichkeit von $P(A$ oder $B)$ nur unter ganz speziellen Voraussetzungen bezüglich des Zusammenhanges zwischen den Ereignissen A und B erhalten (wir setzten zuerst ihre Unverträglichkeit und später ihre Unabhängigkeit voraus).

Die jetzt abgeleitete Formel (4) gilt jedoch ohne jegliche zusätzliche Forderungen für jedes Paar von Ereignissen A, B. Wir dürfen aber trotzdem einen wesentlichen Unterschied zwischen der

Formel (4) und unseren früheren Formeln nicht vergessen. In den früheren Formeln wurde die Wahrscheinlichkeit $P(A$ oder $B)$ stets durch die Wahrscheinlichkeiten $P(A)$ und $P(B)$ ausgedrückt.
Es genügte also die Kenntnis der Wahrscheinlichkeiten $P(A)$ und $P(B)$, um die Wahrscheinlichkeit des Ereignisses $(A$ oder $B)$ eindeutig zu bestimmen. In der Formel (4) sieht die Sache anders aus:

Zur Berechnung von $P(A$ oder $B)$ ist außer $P(A)$ und $P(B)$ noch die Kenntnis der Größe $P(A$ und $B)$ notwendig, d.h. die Wahrscheinlichkeit des gleichzeitigen Eintretens der Ereignisse A und B. Diese Wahrscheinlichkeit aber im allgemeinen Fall zu bestimmen, d. h. bei beliebigem Zusammenhang zwischen den Ereignissen A und B, ist gewöhnlich auch nicht leichter als gleich $P(A$ oder $B)$ zu finden. Daher wird diese Formel für praktische Rechnungen nur sehr selten zu gebrauchen sein. Ihre theoretische Bedeutung ist dafür aber um so größer.

Wir überzeugen uns vor allem, dass man unsere früheren Formeln leicht aus der Formel (4) als Spezialfall ableiten kann. Wenn die Ereignisse A und B unverträglich sind, ist das Ereignis $(A$ und $B)$ unmöglich, $P(A$ und $B) = 0$, und die Formel (4) führt auf die Beziehung

$$P(A \text{ oder } B) = P(A) + P(B)$$

d. h. zum Additionstheorem. Sind die Ereignisse A und B unabhängig, so folgt wegen Formel (3) auf Seite 28

$$P(A \text{ und } B) = P(A) \cdot P(B)$$

und die Formel (4) liefert

$$P(A \text{ oder } B) = P(A) + P(B) - P(A) \cdot P(B) = 1 - [1 - P(A)] \cdot [1 - P(B)]$$

d. h. die Formel (5) auf Seite 31 (für den Fall $n = 2$).
Wir ziehen ferner aus Formel (4) noch eine wichtige Folgerung. Da in allen Fällen $P(A$ und $B) \geq 0$ ist, folgt in jedem Fall aus Formel (4)

$$P(A \text{ oder } B) \leq P(A) + P(B) \tag{5}$$

Diese Ungleichung lässt sich leicht auf eine beliebige Anzahl Ereignisse erweitern. So wird z. B. im Falle dreier Ereignisse wegen (5)

$$P(A \text{ oder } B \text{ oder } C) \leq P(A \text{ oder } B) + P(C) \leq P(A) + P(B) + P(C)$$

Auf demselben Wege kann man nun von drei zu vier Ereignissen übergeben usw. Wir erhalten also das allgemeine Resultat:

Die Wahrscheinlichkeit des Eintretens von wenigstens einem von mehreren Ereignissen ist nie größer als die Summe der Wahrscheinlichkeiten dieser Ereignisse.

Hierbei gilt das Gleichheitszeichen nur in dem Fall, dass die gegebenen Ereignisse paarweise unverträglich sind.

4.11 Die Formel für die totale Wahrscheinlichkeit

Wir wenden uns noch einmal dem Beispiel mit den Glühlampen (Seite 17) zu und benutzen im folgenden für die verschiedenen Beobachtungsergebnisse die eingeführten Bezeichnungen.

Die Wahrscheinlichkeit der Normalqualität einer Glühlampe unter der Voraussetzung, dass sie im zweiten Werk hergestellt wurde, ist, wie wir schon früher gesehen haben, gleich

$$P(A|\overline{B}) = \frac{189}{300} = 0,63$$

Die Wahrscheinlichkeit dieses Ereignisses unter der Voraussetzung der Herstellung dieser Lampe im ersten Werk war

$$P(A|B) = \frac{581}{700} = 0,83$$

Wir nehmen an, diese zwei Zahlen seien uns bekannt, und ebenso mögen auch die Wahrscheinlichkeiten für die Herstellung einer Lampe im ersten Werk $P(B) = 0,7$ und im zweiten Werk $P(\overline{B}) = 0,3$ bekannt sein.

Es wird die unbedingte Wahrscheinlichkeit $P(A)$ gesucht, d. h. die Wahrscheinlichkeit der Normalqualität einer einzelnen Lampe ohne jede Voraussetzung über den Herstellungsort.

Um diese Aufgabe zu lösen, schließen wir wie folgt:
Mit E wollen wir ein doppeltes Ereignis bezeichnen. Es soll darin bestehen, dass die Lampe 1. aus dem ersten Werk stammt und 2. normgerecht ist. Außerdem möge F das analoge Ereignis für das zweite Werk sein. Da jede Normallampe entweder im ersten oder im zweiten Werk gefertigt sein muss, ist das Ereignis A mit dem Ereignis "E oder F" gleichwertig. Die Ereignisse E und F sind aber unverträglich. Also ist nach dem Additionstheorem

$$P(A) = P(E) + P(F) \tag{6}$$

Andererseits muss aber, damit das Ereignis E eintritt, 1. die Lampe aus dem ersten Werk (B) und 2. normgerecht (A) sein. Daher ist E mit dem Ereignis "B und A" gleichwertig, woraus nach der Multiplikationsregel

$$P(E) = P(B) \cdot P(A|B)$$

folgt. Auf dem gleichen Wege finden wir

$$P(E) = P(\overline{B}) \cdot P(A|\overline{B})$$

und durch Einsetzen dieser Ausdrücke in die Gleichung (6) folgt

$$P(A) = P(B) \cdot P(A|B) + P(\overline{B}) \cdot P(A|\overline{B})$$

Diese Formel löst gerade die von uns gestellte Aufgabe. Man setzt die vorgegebenen Zahlen ein und erhält $P(A) = 0,77$.

Beispiel. Für die Aussaat wurde Saatweizen der Sorte I mit einer gewissen Menge Zusatz anderer Sorten II, III und IV vor- bereitet. Wir nehmen eins dieser Körner.
Das Ereignis, welches darin besteht, dass dieses Korn Sorte I ist, bezeichnen wir mit A_1, dass es Sorte II mit A_2, dass es Sorte III mit A_3 und schließlich Sorte IV ist, mit A_4.

Es seien die Wahrscheinlichkeiten dafür bekannt, dass ein aufs Geratewohl genommenes Korn von dieser oder jener Sorte ist. Wir wählen die folgenden Wahrscheinlichkeiten:

$$P(A:1) = 0,96; \quad P(A_2) = 0,01; \quad P(A_3) = 0,02; \quad P(A_4) = 0,01$$

(Die Summe dieser vier Zahlen ist gleich Eins, wie das für ein vollständiges System von Ereignissen der Fall sein muss.)

Dass aus einem Korn eine Ähre wächst, die nicht weniger als 50 Körner enthält, geschehe mit den Wahrscheinlichkeiten

1. $0,50$ für ein Korn der Sorte I,
2. $0,15$ für ein Korn der Sorte II,
3. $0,20$ für ein Korn der Sorte III,
4. $0,05$ für ein Korn der Sorte IV.

Es wird nun die unbedingte Wahrscheinlichkeit für den Fall gesucht, dass eine Ähre nicht weniger als 50 Körner enthält.
K sei das Ereignis, welches darin besteht, dass die Ähre nicht weniger als 50 Körner enthält. Dann ergeben sich aus den Voraussetzungen der Aufgabe

$$P(K|A_1) = 0,50, \quad P(K|A_2) = 0,15, \quad P(K|A_3) = 0,20, \quad P(K|A_4) = 0,05.$$

Unsere Aufgabe ist die Bestimmung von $P(K)$.
Wir führen nun für die Tatsache, dass ein Korn von der Sorte I ist und die Ähre, die aus ihm hervorgeht, nicht weniger als 50 Körner enthält, die Bezeichnung E_1 ein. E_1 ist also dem Ereignis (A_1 und K) gleichwertig. Auf analoge Weise wählen wir die Bezeichnungen E_2 für das Ereignis (A_2 und K), E_3 für das Ereignis (A_3 und K), E_4 für das Ereignis (A_4 und K).

Damit das Ereignis K eintritt, ist offensichtlich notwendig, dass wenigstens eins der Ereignisse E_1, E_2, E_3, E_4 realisiert wird. Da aber je zwei beliebige dieser Ereignisse miteinander unverträglich sind, ergibt sich aus dem Additionstheorem die Beziehung

$$P(K) = P(E_1) + P(E_2) + P(E_3) + P(E_4) \tag{7}$$

Andererseits folgt aus der Multiplikationsregel

$$\begin{aligned}
P(E_1) &= P(A_1 \text{ und } K) = P(A_1) \cdot P(K|A_1),\\
P(E_2) &= P(A_2 \text{ und } K) = P(A_2) \cdot P(K|A_2),\\
P(E_3) &= P(A_3 \text{ und } K) = P(A_3) \cdot P(K|A_3),\\
P(E_4) &= P(A_4 \text{ und } K) = P(A_4) \cdot P(K|A_4)
\end{aligned}$$

Wir setzen diese Ausdrücke in die Gleichung (7) ein und erhalten

$$P(K) = P(A_1) \cdot P(K|A_1) + P(A_2) \cdot P(K|A_2) + P(A_3) \cdot P(K|A_3) + P(A_4) \cdot P(K|A_4)$$

was offensichtlich gerade die Lösung unserer Aufgabe ist.
Wir setzen die vorgegebenen Zahlen ein und finden $P(K) = 0,486$.

Die beiden hier eingehend betrachteten Beispiele führen uns auf eine wichtige allgemeine Formel. Wir können sie jetzt ohne jede Schwierigkeit formulieren und beweisen.
Die Ereignisse $A_1, A_2, ..., A_n$ mögen ein vollständiges System von Ereignissen in einer gegebenen Operation bilden. (Wir erinnern: Das bedeutet, dass je zwei beliebige Ereignisse unverträglich sind und dass irgendeines von ihnen unbedingt eintreten muss.)
Dann gilt für ein beliebiges mögliches Ereignis K dieser Operation die Beziehung

$$P(K) = P(A_1) \cdot P(K|A_1) + P(A_2) \cdot P(K|A_2) + ... + P(A_n) \cdot P(K|A_n) \tag{8}$$

Diese Regel (8) nennt man nun gewöhnlich die "Formel der totalen Wahrscheinlichkeit". Ihr Beweis wird genau so geführt, wie wir es in den beiden betrachteten Beispielen taten. Erstens erfordert das Auftreten des Ereignisses K das Eintreten von einem der Ereignisse "A_i und K", so dass nach dem Additionstheorem

$$P(K) = \sum_{i=1}^{n} P(A_i \text{ und } K) \tag{9}$$

wird; zweitens folgt aus der Multiplikationsregel

$$P(A_i \text{ und } K) = P(A_i) \cdot P(K|A_i)$$

Diese Ausdrücke setzen wir in die Gleichung (9) ein und kommen so zur Formel (8).

4.12 Die Formel von Bayes

Die Formeln des vorhergehenden Paragraphen ermöglichen uns, ein wichtiges Resultat abzuleiten, das vielfach angewandt wird.
Wir beginnen mit der formalen Ableitung dieses Ergebnisses. Die Erläuterung des realen Sinne der Formel, die sich schließlich ergibt, wollen wir bis zur Betrachtung von Beispielen verschieben.

Die Ereignisse $A_1, A_2, ..., A_n$ mögen wiederum ein vollständiges System von Ereignissen einer gewissen Operation bilden. Wenn K dann ein willkürliches Ereignis dieser Operation ist, gilt nach der Multiplikationsregel

$$P(A_i \text{ und } K) = P(A_i) \cdot P(K|A_i) = P(K) \cdot P(A_i|K), \qquad (1 \leq i \leq n)$$

woraus

$$P(A_i|K) = \frac{P(A_i) \cdot P(K|A_i)}{P(K)}, \qquad (1 \leq i \leq n)$$

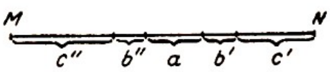
folgt, oder, indem wir in den Nenner des erhaltenen Bruches für $P(K)$ den Ausdruck der totalen Wahrscheinlichkeit (8) des vorhergehenden Paragraphen einsetzen,

$$P(A_i|K) = \frac{P(A_i) \cdot P(K|A_i)}{\sum_{r_1}^{n} P(A_r) \cdot P(K|A_r)}, \qquad (1 \leq i \leq n) \tag{10}$$

Das ist die Formel von Bayes, die in der Praxis bei der Berechnung von Wahrscheinlichkeiten vielfach Anwendung findet. Am häufigsten wird sie in Situationen angewandt, die sich durch folgendes Beispiel illustrieren lassen.

Möge ein Schuss auf ein auf einer geradlinigen Strecke M N (Abb. 3) liegendes Ziel abgegeben werden.

Abb. 3



Diese Strecke zerlegen wir in Gedanken in fünf nicht zu große Bereiche a, b', b'', c', c''. Die genaue Lage des Ziels sei uns unbekannt. Wir kennen nur die Wahrscheinlichkeiten dafür, dass sich das Ziel in diesem oder jenem unserer fünf Bereiche befindet.
Diese Wahrscheinlichkeiten mögen gleich

$$P(a) = 0,48; P(b') = P(b'') = 0,21; P(c') = P(c'') = 0,05$$

sein, wobei mit a, b', b'', c', c'' die folgenden Ereignisse bezeichnet seien: Das Ziel befindet sich im Bereich a, b', b'', c', c'' (die Summe dieser Zahlen ist gleich Eins).
Da die größte Wahrscheinlichkeit dem Bereich a entspricht, werden wir natürlich darauf unseren Schuss abfeuern. Wegen der unvermeidlichen Fehler beim Schießen jedoch kann das Ziel auch dann getroffen werden, wenn es sich nicht im Intervall a, sondern in irgendeinem anderen Bereich befindet.

Die Wahrscheinlichkeiten, das Ziel zu treffen (Ereignis K), seien die folgenden:

$P(K|a) = 0,56$, wenn das Ziel im Intervall a liegt,
$P(K|b') = 0,18$, wenn das Ziel im Intervall b' liegt,
$P(K|b'') = 0,16$, wenn das Ziel im Intervall b'' liegt,
$P(K|c') = 0,06$, wenn das Ziel im Intervall c' liegt,
$P(K|c'') = 0,02$, wenn das Ziel im Intervall c'' liegt.

Angenommen, ein Schuss sei abgegeben und das Ziel getroffen worden (das ist das Ereignis K). Auf Grund dieses Ereignisses werden die Wahrscheinlichkeiten für verschiedene Lagen des Ziels, wie wir sie oben hatten (d. h. die Zahlen $P(a)$, $P(b')$, ...), leicht zu hoch geschätzt. Die qualitative Seite dieser Schätzung ist ohne jede Rechnung klar.

Wir schießen auf den Abschnitt a und treffen das Ziel, also wird sich die Wahrscheinlichkeit $P(a)$ vergrößern. Wir wollen aber die durch den von uns abgegebenen Schuss hervorgerufene Schätzung quantitativ genau berechnen, d. h., wir wollen die genauen Ausdrücke für die Wahrscheinlichkeiten $P(a|K)$, $P(b'|K)$, ... der verschiedenen möglichen Lagen des Ziels unter der Voraussetzung berechnen, dass das Ziel von dem abgegebenen Schuss getroffen wurde.
Die Formel von Bayes (10) gibt uns sofort eine Antwort auf diese Frage.
So ist

$$P(a|K) = \frac{P(a) \cdot P(K|a)}{P(a)P(K|a) + P(b')P(K|b') + P(b'')P(K|b'') + P(c')P(K|c') + P(c'')P(K|c'')} \approx 0,8$$

Wie wir sehen, ist $P(a|K)$ tatsächlich größer als $P(a)$.

Auf analoge Art ergeben sich auch leicht die Wahrscheinlichkeiten $P(b'|K)$, ... für die anderen möglichen Lagen des Ziels.
Zur Ausführung der Berechnungen möchten wir noch erwähnen, dass sich die einzelnen Ausdrücke, die für diese Wahrscheinlichkeiten durch die Formel von Bayes geliefert werden, nur durch ihre Zähler voneinander unterscheiden. Ihre Nenner sind alle einander gleich, und zwar gleich $P(K) \approx 0,34$.

Für ähnliche Fälle kann man das allgemeine Schema auch wie folgt beschreiben. Die Bedingungen einer Operation enthalten ein gewisses unbekanntes Element, bezüglich dessen man n verschiedene "Hypothesen" $A_1, A_2, ..., A_n$ machen kann. Diese n "Hypothesen" müssen ein vollständiges System von Ereignissen bilden.
Aus irgendwelchen Ursachen seien uns die Wahrscheinlichkeiten $P(A_i)$ dieser Hypothesen bekannt. Bekannt sei uns ferner, dass die Hypothese A_i einem gewissen Ereignis K (beispielsweise dem Treffen des Ziels) die Wahrscheinlichkeit $P(K|A_i)$ $(1 \leq i \leq n)$ "erteilt" [hier ist $P(K|A_i)$ die Wahrscheinlichkeit des Ereignisses K, die unter der Bedingung berechnet wird, dass die Hypothese A_i begründet ist].

Wenn nun bei der Ausführung des Versuchs das Ereignis K eintritt, so ruft das eine zu hohe Schätzung der Wahrscheinlichkeiten von Hypothesen A_i hervor. Die Aufgabe besteht nun darin, die neuen Wahrscheinlichkeiten $P(A_i|K)$ dieser Hypothesen zu finden. Die Lösung dieses

Problems kann nun aber mittels der Formel von Bayes gegeben werden.

In der artilleristischen Praxis führt man das sogenannte Einschießen durch. Es hat den Zweck, unsere Kenntnisse von den Schussbedingungen zu verbessern. Dabei kann als unbekanntes Element, das genau bestimmt werden soll, nicht nur die Lage des Ziels, sondern auch jedes andere die Schussbedingungen in ihrer Wirksamkeit beeinflussende Element dienen (insbesondere eine Eigentümlichkeit der verwendeten Waffe).

Sehr oft ist es nicht so, dass nur ein Probeschuss abgegeben wird, sondern eine ganze Serie von Probeschüssen. Es wird dann die Frage nach der Berechnung der neuen Wahrscheinlichkeiten auf Grund der erhaltenen Schussresultate gestellt. In allen solchen Fällen kann man mit Erfolg und leicht die vorgelegte Aufgabe mittels der Formel von Bayes lösen.

Zur Abkürzung setzen wir in dem von uns betrachteten allgemeinen Schema

$$P(A_i) = P_i, \qquad P(K|A_i) = p_i, \qquad (1 \leq i \leq n)$$

so dass die Formel von Bayes eine einfache Gestalt annimmt:

$$P(A_i|K) = \frac{P_i p_i}{\sum\limits_{r=1}^{n} P_r p_r}$$

Wir nehmen an, dass s Probeschüsse abgegeben werden sind. Das Ereignis K sei hierbei m-mal eingetreten und $(s-m)$-mal nicht eingetreten.
Mit K^* bezeichnen wir das erhaltene Ergebnis einer Serie von s Schüssen. Wir können hierbei annehmen, dass die Resultate der einzelnen Schüsse voneinander unabhängige Ereignisse sind. Ist die Hypothese A_i begründet, so ist die Wahrscheinlichkeit des Ereignisses K gleich p_i, und somit die Wahrscheinlichkeit des entgegengesetzten Ereignisses (d.h., dass K nicht eintritt) gleich $1-p$.

Daher wird die Wahrscheinlichkeit des zusammengesetzten Ereignisses K^* nach der Multiplikationsregel für unabhängige Ereignisse gleich dem Produkt von 8 Faktoren, von denen m gleich p_i, und $s-m$ gleich $1-p_i$ sind. Anders ausgedrückt heißt das:

$$P(K^*|A_i) = \binom{s}{m} p_i^m (1-p_i)^{s-m}, \qquad (1 \leq i \leq n)$$

und die Formel von Bayes gibt:

$$P(A_i|K) = \frac{P_i p_i^m (1-p_i)^{s-m}}{\sum\limits_{r=1}^{n} P_r p_r^m (1-p_r)^{s-m}}, \qquad (1 \leq i \leq n) \tag{11}$$

Damit ist die gestellte Aufgabe gelöst.

Natürlich gibt es solche Aufgaben nicht nur in der Praxis des Artilleristen, sondern in allen anderen Gebieten des menschlichen Lebens.

Beispiel 1. In der von uns am Anfang dieses Paragraphen betrachteten Aufgabe ist die Wahrscheinlichkeit dafür zu finden, dass das Ziel sich im Abschnitt a befand, wenn zwei aufeinanderfolgende Schüsse auf diesen Abschnitt Treffer ergaben.
K^* soll einen Doppeltreffer im Ziel bedeuten. Gemäß Formel (11) gilt

$$P(a|K^*) = \frac{P(a) \cdot [P(K|a)]^2}{P(a) \cdot [P(K|a)]^2 + P(b') \cdot [P(K|b')]^2 + ...}$$

Wir überlassen dem Leser diese einfache Rechnung und erwähnen nur, dass als Ergebnis eines zweifachen Treffers sich die Wahrscheinlichkeit dafür, dass sich das Ziel im Abschnitt befand, noch weiter vergrößert hat.

Beispiel 2. Die Wahrscheinlichkeit dafür, dass in einem gewissen Werk ein Erzeugnis der Norm genügt, sei 0,96. Verlangt wird ein vereinfachtes Kontrollsystem[9], welches für die der Norm genügenden Erzeugnisse ein positives Resultat mit einer Wahrscheinlichkeit von 0,98 ergibt. Für Erzeugnisse, die der Norm nicht genügen, möge sie ein positives Resultat nur mit einer Wahrscheinlichkeit von 0,05 liefern.
Wie groß ist nun die Wahrscheinlichkeit dafür, dass ein bei der vereinfachten Kontrolle zweimal als normgerecht befundenes Erzeugnis der Norm genügt?

Hier besteht das vollständige System von Hypothesen aus zwei entgegengesetzten Ereignissen: 1. Das Erzeugnis genügt der Norm; 2. das Erzeugnis genügt der Norm nicht.
Die Wahrscheinlichkeiten dieser Hypothesen seien vor dem Versuch gleich $P_1 = 0,96$ bzw. $P_2 = 0,04$. Bei der ersten Hypothese ist die Wahrscheinlichkeit dafür, dass das Erzeugnis die Prüfung übersteht, gleich $p_1 = 0,98$ und bei der zweiten entsprechend $p_2 = 0,05$.
Nach einer zweimaligen Probe ist die Wahrscheinlichkeit der ersten Hypothese auf Grund der Formel (11) gleich

$$\frac{P_1 p_1^2}{P_1 p_1^2 + P_2 p_2^2} = \frac{0,96 \cdot 0,98^2}{0,96 \cdot 0,98^2 + 0,04 \cdot 0,05^2} \approx 0,9999$$

Wir sehen also: wenn das Erzeugnis die in den Voraussetzungen der Aufgabe erwähnte Probe erfüllt, so können wir nur in einem von zehntausend Fällen einen Fehler begehen, indem wir das Erzeugnis als normgerecht annehmen. Ein solches Ergebnis wird aber den Anforderungen der Praxis vollkommen gerecht.

Beispiel 3. Die Untersuchung eines Kranken möge den Verdacht auf eine von drei Krankheiten nach sich ziehen: A_1, A_2, A_3.
Ihre Wahrscheinlichkeiten unter gegebenen Voraussetzungen seien

$$P_1 = \frac{1}{2}, \qquad P_2 = \frac{1}{6}, \qquad P_3 = \frac{1}{3}$$

Zur genaueren Bestimmung der Diagnose wird eine Analyse durchgeführt, die ein positives Resultat im Falle der Krankheit A_1 mit der Wahrscheinlichkeit $0,1$, im Falle der Krankheit A_2 mit der Wahrscheinlichkeit $0,2$ und im Falle der Krankheit A_3 mit der Wahrscheinlichkeit $0,9$ ergibt.
Die Analyse wurde fünfmal durchgeführt und lieferte vier positive und ein negatives Resultat. Wie groß ist die Wahrscheinlichkeit jeder Krankheit nach der Analyse ?

Im Falle der Krankheit A_1 ist die Wahrscheinlichkeit des erwähnten Analysenresultats $p_1 = \binom{5}{4} \cdot 0,1^4 \cdot 0,9$.
Für die zweite Hypothese wird die Wahrscheinlichkeit gleich $p_2 = \binom{5}{4} \cdot 0,2^4 \cdot 0,8$ und für die dritte $p_3 = \binom{5}{4} \cdot 0,9^4 \cdot 0,1$.

[9]Die Notwendigkeit einer vereinfachten Kontrolle tritt in der Praxis sehr oft auf. So z. B. bei der Prüfung der hergestellten Glühlampen. Wollte man z. B. jede Glühlampe einer Brenndauerprüfung von, sagen wir, 1200 Stunden unterziehen, so erhielte der Verbraucher nur ausgebrannte oder fast ausgebrannte Glühlampen zu kaufen. Es ist also diese Art der Kontrolle durch eine geeignetere Prüfung zu ersetzen. In der Praxis ersetzt man nun diese Kontrolle durch ein kurzzeitiges Einschalten jeder Glühlampe (bzw. lässt man von jeder unter gleichen Fabrikationsbedingungen hergestellten Groß-Serie eine kleine Anzahl bis zum Ausbrennen leuchten).

Aus der Formel von Bayes ergeben sich die folgenden Resultate: Die Wahrscheinlichkeit der Krankheit A_1 wird

$$\frac{P_1p_1}{P_1p_1+P_2p_2+P_3p_3} = \frac{\frac{1}{2}\cdot 0,1^4\cdot 0,9}{\frac{1}{2}\cdot 0,1^4\cdot 0,9+\frac{1}{6}\cdot 0,2^4\cdot 0,8+\frac{1}{3}\cdot 0,9^4\cdot 0,1} \approx 0,002$$

die Wahrscheinlichkeit der Krankheit A_2:

$$\frac{P_2p_2}{P_1p_1+P_2p_2+P_3p_3} = \frac{\frac{1}{6}\cdot 0,2^4\cdot 0,8}{\frac{1}{2}\cdot 0,1^4\cdot 0,9+\frac{1}{6}\cdot 0,2^4\cdot 0,8+\frac{1}{3}\cdot 0,9^4\cdot 0,1} \approx 0,010$$

und die Wahrscheinlichkeit der Krankheit A_3:

$$\frac{P_3p_3}{P_1p_1+P_2p_2+P_3p_3} = \frac{\frac{1}{3}\cdot 0,9^4\cdot 0,1}{\frac{1}{2}\cdot 0,1^4\cdot 0,9+\frac{1}{6}\cdot 0,2^4\cdot 0,8+\frac{1}{3}\cdot 0,9^4\cdot 0,1} \approx 0,988$$

Da diese drei Ereignisse A_1, A_2, A_3 auch auf Grund der Erfahrung ein vollständiges System von Ereignissen bilden, kann man zur Kontrolle der durchgeführten Rechnungen die drei erhaltenen Zahlen addieren und sich davon überzeugen, dass ihre Summe wiederum gleich Eins ist.

5 Das Bernoullische Schema

5.13 Beispiele

Beispiel 1. Baumwollfasern einer bestimmten Sorte haben im Mittel zu 75 % eine Länge, die kleiner als 45 mm ist, und 25 % der Baumwollfasern sind länger als (oder gleich) 45 mm. Zu bestimmen ist die Wahrscheinlichkeit dafür, dass von drei beliebig herausgegriffenen Baumwollfasern zwei kürzer und eine länger als 45 mm sind.

Es sei A das Ereignis der Auswahl einer Faser, die kürzer als 45 mm ist, und B dasjenige, dass die ausgewählte Faser länger als 45 mm ist. Dann ist selbstverständlich

$$P(A) = \frac{3}{4} \quad , \quad P(B) = \frac{1}{4}$$

Ferner verabreden wir, mit dem Schema AAB das folgende zusammengesetzte Ereignis zu bezeichnen: Die ersten zwei ausgewählten Baumwollfasern sind länger und die dritte ist kürzer als 45 mm. Damit dürfte aber auch klar sein, welche Bedeutung die Schemata BBA, ABA usw. haben.

Unsere Aufgabe besteht nun darin, die Wahrscheinlichkeit des Ereignisses C zu berechnen, welches darin besteht, dass von den drei ausgewählten Baumwollfädchen zwei kürzer und eines länger als 45 mm sind. Zu diesem Zwecke muss sich offensichtlich eines der folgenden Schemata verwirklichen:

$$AAB, \quad ABA, \quad BAA \tag{1}$$

Da jeweils zwei dieser drei Ereignisse miteinander unverträglich sind, folgt aus dem Additionstheorem

$$P(C) = P(AAB) + P(ABA) + P(BAA)$$

Alle drei Summanden der rechten Seite dieser Gleichung sind einander gleich, weil die Ereignisse der Faserauswahl als voneinander unabhängig angenommen werden können. Die Wahrscheinlichkeit jedes der Schemata (1) wird nach der Multiplikationsregel für die Wahrscheinlichkeiten unabhängiger Ereignisse durch ein Produkt von drei Faktoren gebildet, von denen zwei gleich $P(A) = \frac{3}{4}$ und einer gleich $P(B) = \frac{1}{4}$ sind. Also ist die Wahrscheinlichkeit eines jeden der drei Schemata (1) gleich

$$\left(\frac{3}{4}\right)^2 \cdot \frac{1}{4} = \frac{9}{64} \qquad \text{und somit} \qquad P(C) = 3 \cdot \frac{9}{64} = \frac{27}{64}$$

womit die gestellte Aufgabe gelöst ist.

Beispiel 2. Durch Beobachtungen, die über viele Jahrzehnte fortgesetzt Wurden, stellte man fest, dass von tausend Neugeborenen im Mittel 515 Knaben und 485 Mädchen sind. In einer gewissen Familie mögen sechs Kinder sein.
Wie groß ist die Wahrscheinlichkeit dafür, dass von diesen sechs Kindern nicht mehr als zwei Mädchen sind ?

Zur Verwirklichung dieses Ereignisses, dessen Wahrscheinlichkeit wir suchen, dürfen in dieser Familie nicht mehr als zwei Mädchen sein, d. h., es müssen entweder kein oder ein oder zwei Mädchen sein.
Die Wahrscheinlichkeiten dieser Teilereignisse bezeichnen wir mit P_0, P_1 bzw. P_2. Offenbar wird die gesuchte Wahrscheinlichkeit nach dem Additionstheorem

$$P = P_0 + P_1 + P_2 \tag{2}$$

Für jedes Kind ist die Wahrscheinlichkeit, dass es ein Knabe ist, gleich 0,515 und somit die Wahrscheinlichkeit dafür, dass es ein Mädchen ist, gleich 0,485.

Am einfachsten ist P_0 zu finden. Das ist die Wahrscheinlichkeit dafür, dass alle Kinder in der Familie Knaben sind. Da wir das Ereignis - die Geburt eines Kindes irgendeines Geschlechts - als unabhängig von der Geburt der anderen Kinder ansehen können, folgt aus der Multiplikationsregel, dass die Wahrscheinlichkeit dafür, dass alle sechs Kinder Knaben sind, gleich dem Produkt von sechs gleichen Faktoren 0,515 ist, d.h.

$$P_0 = (0,515)^6 \approx 0,018$$

Wir kommen jetzt zur Berechnung von P_1, d. h. der Wahrscheinlichkeit dafür, dass sich unter den sechs Kindern in der Familie fünf Knaben und ein Mädchen befinden. Zur Realisierung dieses Ereignisses hat man sechs verschiedene Möglichkeiten.
Sie ergeben sich aus der Reihenfolge der Geburten, d. h. als wievieltes Kind das Mädchen geboren wurde (als erstes, zweites usw.). Wir betrachten irgendeine der Abarten unseres Ereignisses, z. B. dass das Mädchen als viertes Kind geboren wurde.

Die Wahrscheinlichkeit dieser Abart ist nach der Multiplikationsregel gleich einem Produkt von sechs Faktoren, von denen fünf gleich 0,515 sind und der sechste (der an vierter Stelle steht) gleich 0,485 ist, d. h., diese Wahrscheinlichkeit ist gleich $(0,515)^5 \cdot 0,485$.
Genau dieselbe Wahrscheinlichkeit besitzt auch jede der möglichen fünf anderen Abarten des uns interessierenden Ereignisses. Daher wird die Wahrscheinlichkeit P_1 nach dem Additionstheorem gleich der Summe dieser sechs Zahlen, die gleich $(0,515)^5 \cdot 0,485$ sind, d. h.,

$$P_1 = 6 \cdot (0,515)^5 \cdot 0,485 \approx 0,105$$

Wir wenden uns jetzt der Berechnung von P_2 zu (der Wahrscheinlichkeit, dass zwei Kinder Mädchen und vier Knaben sind).

Ähnlich wie bei den vorherigen Überlegungen stellen wir sofort fest, dass das Ereignis eine Reihe von möglichen Varianten aufzuweisen hat (eine Variante ist beispielsweise: das zweite und fünfte Kind in der Reihenfolge der Geburten sind Mädchen und die übrigen Kinder Knaben). Die Wahrscheinlichkeit jeder Abart wird nach der Multiplikationsregel gleich $(0,515)^4 \cdot (0,485)^2$. Um nun P_2 zu erhalten, muss man nach dem Additionstheorem die Zahl $(0,515)^4 \cdot (0,485)^2$ mit der Anzahl aller möglichen Abarten des betrachteten Typs multiplizieren. Auf diese Weise ist die ganze Aufgabe auf die Bestimmung dieser letzten Zahl zurückgeführt.

Jede Abart wird dadurch charakterisiert, dass von sechs Kindern jeweils zwei Mädchen und vier Knaben sind. Die Anzahl der verschiedenen Abarten ist folglich gleich der Anzahl der verschiedenen Arten, auf die man von sechs Kindern zwei auswählen kann. Eine solche Auswahl ist auf soviel Arten möglich, wie man sechs Gegenstände zu je zweien ohne Wiederholungen miteinander kombinieren kann, d. h. $\binom{6}{2} = \frac{6\cdot 5}{1\cdot 2} = 15$.
Also wird

$$P2 = \binom{6}{2} \cdot (0,515)^4 \cdot (0,485)^2 = 15 \cdot (0,515)^4 \cdot (0,485)^2 \approx 0,247$$

Wir fassen die erhaltenen Resultate zusammen und finden:

$$P = P_0 + P_1 + P_2 \approx 0,018 + 0,105 + 0,247 = 0,370$$

In weniger als vier von zehn Fällen werden in solchen kinderreichen Familien nicht mehr als ein Drittel Mädchen und somit, nicht weniger als zwei Drittel Knaben geboren werden.

5.14 Die Bernoullische Formel

Im vorhergehenden Paragraphen machten wir uns durch eine Reihe von Beispielen mit dem Schema wiederholter Versuche bekannt.
In jedem Versuch konnte ein gewisses Ereignis A auftreten. Das Wort "Versuch" hat bei uns einen sehr vielfältigen Sinn.

Wenn wir auf ein gewisses Ziel einen Schuss abgeben, verstehen wir unter Versuch einen einzelnen Schuss. Beim Prüfen der Brenndauer einer Glühlampe bedeutet Versuch die Kontrolle einer einzelnen Glühlampe.
Wenn wir die Einteilung von Neugeborenen nach Geschlecht, Gewicht oder Statur (Größe) untersuchen, verstehen wir unter Versuch die Untersuchung eines einzigen Säuglings.
Im allgemeinen werden wir unter Versuch im folgenden die Erfüllung gewisser Bedingungen verstehen, bei, deren Vorhandensein ein gewisses uns interessierendes Ereignis eintreten kann.

Wir beginnen jetzt mit der Betrachtung eines der Hauptschemata der Wahrscheinlichkeitsrechnung. Diese Schemata haben, abgesehen von der Bedeutung für die Anwendung in den verschiedenartigsten Wissensgebieten, auch eine sehr große Bedeutung in der Wahrscheinlichkeitsrechnung als mathematische Wissenschaft.
Das Schema, welches wir hier betrachten wollen, besteht darin, dass eine Folge unabhängiger Ereignisse untersucht wird. In einer Folge unabhängiger Ereignisse ist die Wahrscheinlichkeit eines Ereignisses unabhängig davon, was die bisherigen Versuche für Ergebnisse hatten oder was die noch auszuführenden Versuche für Resultate im einzelnen liefern werden.

In jedem dieser Versuche kann ein Ereignis A mit einer Wahrscheinlichkeit p, die nicht von der Nummer des Versuchs abhängt, eintreten (oder nicht eintreten). Das beschriebene Schema erhielt den Namen Bernoullisches Schema, weil es zuerst von dem bekannten Schweizer

Gelehrten Jakob Bernoulli systematisch untersucht wurde. Jakob Bernoulli lebte gegen Ende des XVII. Jahrhunderts.

Mit dem Bernoullischen Schema hatten wir es schon in verschiedenen Beispielen zu tun. Um sich hiervon zu überzeugen, braucht, man sich nur an die Beispiele des letzten Paragraphen zu erinnern. Jetzt wollen wir aber das allgemeine Problem lösen; denn die von uns bis jetzt in diesem Kapitel betrachteten Beispiele waren alle nur Spezialfälle dieses Schemas.

Aufgabe. Unter bestimmten Voraussetzungen sei die Wahrscheinlichkeit für das Auftreten eines Ereignisses A in jedem Versuch gleich p. Gesucht ist die Wahrscheinlichkeit für das k-malige Auftreten und $(n-k)$-malige Nichtauftreten des Ereignisses A in einer Serie von n unabhängigen Versuchen.

Das Ereignis, dessen Wahrscheinlichkeit wir suchen, lässt sich auf verschiedene Weisen realisieren. Um eine bestimmte Realisierung zu erhalten, müssen wir aus der vorgegebenen Serie irgendwelche k Versuche herausgreifen und annehmen, dass sich bei diesen k Versuchen A ereignet und bei den restlichen $(n-k)$ Versuchen A nicht eintritt.
Also erfordert jede solche Realisierung das Eintreten von n bestimmten Ereignissen, und zwar k-maliges Eintreten und $(n-k)$-maliges Nichteintreten des Ereignisses A. Aus der Multiplikationsregel ergibt sich, dass die Wahrscheinlichkeit jeder bestimmten Realisierung gleich

$$p^k(1-p)^{n-k}$$

ist. Die Anzahl der verschiedenen Realisierungen ist gleich der Anzahl der verschiedenen Gruppen zu je k Versuchen, die man aus den n Versuchen zusammenstellen kann, d.h. gleich $\binom{n}{k}$. Durch Anwendung des Additionstheorems und der bekannten Formel für die Anzahl der Kombinationen ohne Wiederholungen,

$$\binom{n}{k} = \frac{n(n-1)...(n-(k-1))}{k(k-1)...2\cdot 1}$$

finden wir, dass die gesuchte Wahrscheinlichkeit für das k-malige Auftreten des Ereignisses A bei n unabhängigen Versuchen gleich

$$P_n(k) = \frac{n(n-1)...(n-(k-1))}{k(k-1)...2\cdot 1} p^k(1-p)^{n-k} \tag{3}$$

ist. Damit ist die gestellte Aufgabe auch schon gelöst.
Oft ist es noch günstiger, dem Ausdruck (2) eine andere Form zu geben. Wir multiplizieren dazu Zähler und Nennen mit dem Produkt $(n-k)(n-(k+1))...2\cdot 1$ und erhalten

$$\binom{n}{k} = \frac{n(n-2)...2\cdot 1}{k(k-1)...2\cdot 1\cdot (n-k)(n-(k+1))...2\cdot 1}$$

oder, wenn wir zur Abkürzung mit $m!$ das Produkt aller ganzen Zahlen von 1 bis m einschließlich bezeichnen,

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

Für $P_n(k)$ ergibt das:

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k(1-p)^{n-k} \tag{4}$$

Die Formeln (3) und (4) nennt man gewöhnlich Bernoullische Formeln.

Für große Werte von n und k ist die Berechnung von $P_n(k)$ auf Grund dieser Formeln sehr mühselig, weil $n!$, $k!$ und $(n-k)!$ sehr große und nur schwer berechenbare Zahlen sind. Für Berechnungen dieser Art werden daher sowohl spezielle Tafeln für die Fakultäten als auch gewisse Näherungsformeln sehr viel verwandt.

Beispiel. Die Wahrscheinlichkeit dafür, dass in einem gewissen Industriebetrieb der Wasserverbrauch normal ist (nicht größer als eine bestimmte Anzahl von Litern in 24 Stunden), sei gleich $\frac{3}{4}$.
Es ist die Wahrscheinlichkeit dafür zu finden, dass innerhalb der nächsten sechs Tage der Wasserverbrauch im Laufe von $0, 1, 2, 3, 4, 5, 6$ Tagen normal ist.

Mit $P_6(k)$ bezeichnen wir die Wahrscheinlichkeit eines normalen Wasserverbrauchs im Verlaufe von k von sechs Tagen. Aus Formel (3) ergibt sich (man muss $p = \frac{3}{4}$ setzen)

$$P_6(6) = \left(\frac{3}{4}\right)^6 = \frac{3^6}{4^6}$$

$$P_6(5) = 6 \cdot \left(\frac{3}{4}\right)^5 \frac{1}{4} = \frac{6 \cdot 3^5}{4^6}$$

$$P_6(4) = \binom{6}{4} \cdot \left(\frac{3}{4}\right)^4 \cdot \left(\frac{1}{4}\right)^2 = \binom{6}{2} \cdot \left(\frac{3}{4}\right)^4 \cdot \left(\frac{1}{4}\right)^2 = \frac{6 \cdot 5}{1 \cdot 2} \cdot \left(\frac{3}{4}\right)^4 \cdot \left(\frac{1}{4}\right)^2 = \frac{15 \cdot 3^4}{4^6}$$

$$P_6(3) = \binom{6}{3} \cdot \left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right)^3 = \frac{6 \cdot 5 \cdot 4}{1 \cdot 2 \cdot 3} \cdot \left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right)^3 = \frac{20 \cdot 3^3}{4^6}$$

$$P_6(2) = \binom{6}{2} \cdot \left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^4 = \frac{6 \cdot 5}{1 \cdot 2} \cdot \left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^4 = \frac{15 \cdot 3^2}{4^6}$$

$$P_6(1) = 6 \cdot \frac{3}{4} \cdot \left(\frac{1}{4}\right)^5 = \frac{6 \cdot 3}{4^6}$$

Offensichtlich ist $P_6(0)$ gleich $\frac{1}{4^6}$. (Das ist die Wahrscheinlichkeit dafür, dass an jedem der sechs Tage ein Überverbrauch vorhanden ist).
Die Brüche aller sechs Wahrscheinlichkeiten haben denselben Nenner, nämlich $4^6 = 4096$. Hierdurch lassen sich die Rechnungen um sehr vieles abkürzen. Die Berechnungen ergeben:

$$P_6(6) \approx 0,18; \quad P_6(5) \approx 0,36; \quad P_6(4) \approx 0,30; \quad P_6(3) \approx 0,13;$$
$$P_6(2) \approx 0,03; \quad P_6(1) \approx P_6(0) \approx 0$$

Wir sehen, dass man von den sechs Tagen am wahrscheinlichsten mit einem Überverbrauch an ein oder zwei Tagen rechnen muss.
Dagegen ist die Wahrscheinlichkeit dafür, dass an fünf oder sechs Tagen ein Überverbrauch vorhanden ist, praktisch gleich Null.

5.15 Die wahrscheinlichste Anzahl von Wiederholungen eines Ereignisses

Das gerade betrachtete Beispiel zeigt, dass die Wahrscheinlichkeit eines Überverbrauches an Wasser im Laufe von 10 Tagen mit wachsendem k zuerst zunimmt und dann nach dem Erreichen eines Höchstwertes wieder absinkt.
Diesen Tatbestand sieht man am deutlichsten, wenn man sich die Änderung der Wahrscheinlichkeiten $P_6(k)$ für wachsendes k in Form eines Diagramme geometrisch darstellt, wie dies Abbildung 4 zeigt.

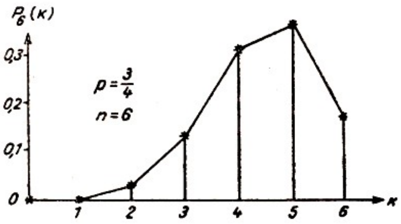


Abb. 4

Ein noch überzeugenderes Bild von der Änderung der Größe $P_n(k)$ für wachsendes k geben diese Diagramme, wenn die Zahl n bedeutend vergrößert wird. So hat beispielsweise für $n = 15$ und $p = \frac{1}{2}$ ein solches Diagramm die in der Abbildung 5 angegebene Form.

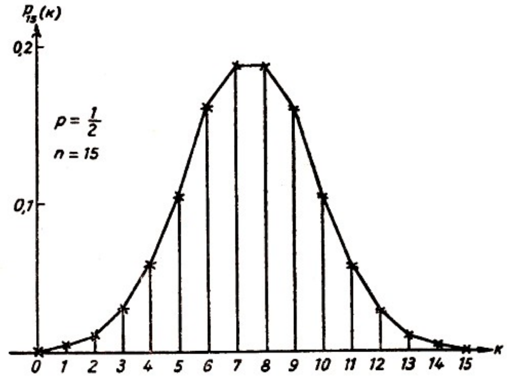


Abb. 5

In der Praxis wird manchmal gefragt, welche Anzahl von Wiederholungen eines Ereignisses die wahrscheinlichste ist, d. h., für welche Zahl k die Wahrscheinlichkeit $P_n(k)$ ein Größtwert wird (hierbei werden natürlich p und n als gegeben vorausgesetzt).

Die Bernoullische Formel gibt in allen Fällen eine einfache Lösung dieser Frage. Wir werden uns hiermit jetzt beschäftigen.
Zuerst wollen wir die Größe $\frac{P_n(k+1)}{P_n(k)}$ berechnen. Nach Formel (4) ist

$$P_n(k+1) = \frac{n!}{(k+1)!(n-k-1)!} p^{k+1}(1-p)^{n-k-1} \tag{5}$$

und aus den Formeln (3) und (5) ergibt sich

$$\frac{P_n(k+1)}{P_n(k)} = \frac{n!k!(n-k)!p^{k+1}(1-p)^{n-k-1}}{(k+1)!(n-k-1)!n!p^k(1-p)^{n-k}} = \frac{n-k}{k+1} \cdot \frac{p}{1-p}$$

Die Wahrscheinlichkeit $P_n(k+1)$ wird also größer, gleich oder kleiner als die Wahrscheinlichkeit $P_n(k)$, je nach dem, ob $\frac{P_n(k+1)}{P_n(k)}$ größer, gleich oder kleiner als Eins ist. Das führt aber, wie wir sahen, auf die Frage, welche der drei Beziehungen

$$\frac{n-k}{k+1} \cdot \frac{p}{1-p} > 1, \qquad \frac{n-k}{k+1} \cdot \frac{p}{1-p} = 1, \qquad \frac{n-k}{k+1} \cdot \frac{p}{1-p} < 1 \tag{6}$$

sich als richtig erweist. Um z. B. zu erklären, für welche k-Werte die Ungleichung $P_n(k+1) > P_n(k)$ erfüllt wird, muss man wissen, für welche k-Werte die Ungleichung

$$\frac{n-k}{k+1} \cdot \frac{p}{1-p} > 1 \qquad \text{oder} \qquad (n-k)p > (k+1)(1-p)$$

gilt. Hieraus ergibt sich aber:

$$np - (1-p) > k$$

Solange also das wachsende k unterhalb der Größe $np-(1-p)$ bleibt, gilt immer die Ungleichung $P_n(k+1) > P_n(k)$, d. h. nimmt mit wachsendem k auch die Wahrscheinlichkeit $P_n(k)$ zu.

Prüfen wir unsere Theorie gleich an dem Schema nach, welches in der Abb. 5 als Diagramm dargestellt ist ($p = \frac{1}{2}$, $n = 15$, $np-(1-p) = 7$). Dort muss also für $k < 7$, d.h. für alle k von 0 bis 6 einschließlich stets $P_n(k+1) > P_(k)$ sein.
Diese Behauptung bestätigt das Diagramm.

Auf völlig analogem Wege findet man, indem man von den zwei anderen Beziehungen (6) ausgeht, dass

$$P_n(k+1) = P_n(k), \quad \text{wenn} \quad k = np-(1-p) \quad \text{und}$$
$$P_n(k+1) < P_n(k), \quad \text{wenn} \quad k > np-(1-p)$$

Sobald also die wachsende Zahl k die Schranke $np-(1-p)$ überschreitet, nimmt die Wahrscheinlichkeit $P-n(k)$ ab, und zwar bis auf $P_n(n)$.

Durch diese Aussage werden wir vor allem davon überzeugt, dass das Verhalten der Größe $P_n(k)$ bei wachsendem k (erst zunehmen und dann abnehmen) in dem oben betrachteten Beispiel ein allgemeines, in allen Fällen geltendes Gesetz ist. Aber nicht nur das. Diese Aussage ermöglicht uns auch, unverzüglich die andere von uns gestellte Aufgabe zu lösen, nämlich den wahrscheinlichsten Wert der Zahl k zu bestimmen. Wir wollen den wahrscheinlichsten Wert mit k_0 bezeichnen. Dann muss

$$P_n(k_0+1) \leq P_n(k_0)$$

sein und auf Grund unserer Überlegungen

$$k_0 \geq np-(1-p)$$

Andererseits soll aber auch

$$P_n(k_0-1) \leq P_n(k_0)$$

sein; also muss auch die Ungleichung

$$k_0 - 1 \leq np-(1-p) \qquad \text{oder} \qquad k_0 \leq np-(1-p)+1 = np+p$$

gelten. Also muss der wahrscheinlichste Wert k_0 der Zahl k einer zweifachen Ungleichung genügen:

$$np-(1-p) \leq k_0 \leq np+p \tag{7}$$

Das Intervall von $np-(1-p)$ bis $np+p$, in welchem die Zahl k_0 liegt, hat, wie eine einfache Subtraktion zeigt, die Länge 1.
Wenn daher ein Intervallende, sagen wir die Zahl $np-(1-p)$, keine ganze Zahl ist, dann wird es zwischen diesen Enden eine und nur eine ganze Zahl geben. In diesem Falle ist also k_0 eindeutig bestimmt. Diesen Fall können wir als den normalen ansehen; denn da $p < 1$ ist, wird die Größe $np-(1-p)$ nur in Ausnahmefällen eine ganze Zahl werden.
Tritt dieser besondere Umstand ein, so liefert uns Ungleichung (7) für k_0 zwei Werte: $np-(1-p)$ und $np+p$, die sich gerade um den Wert 1 voneinander unterscheiden. Dann sind diese beiden Werte die wahrscheinlichsten.

Ihre Wahrscheinlichkeiten sind einander gleich und größer als die Wahrscheinlichkeiten für alle anderen k. Mit solch einem Spezialfall haben wir es gerade in unserem Diagramm der Abb. 5 zu tun.

Hier ist $n = 15$, $p = \frac{1}{2}$ und somit $np - (1 - p) = 7$ und $np + p = 8$. Es ist also am wahrscheinlichsten, dass das Ereignis 7-mal oder 8-mal eintritt. Die Wahrscheinlichkeiten dafür sind einander gleich, jede von ihnen beträgt etwa 0,196 (das kann man auch alles unmittelbar aus dem Diagramm ablesen).

Beispiel 1. Auf Grund langjähriger Beobachtungen wurde in einer gewissen Gegend folgendes festgestellt: Dafür, dass es am 1. Juli regnet, ist die Wahrscheinlichkeit gleich 4/17. Wie groß ist nun die wahrscheinlichste Anzahl der verregneten 1. Julis in den nächsten 50 Jahren?

Hier ist $n = 50$, $p = 4/17$ und somit

$$np - (1 - p) = 50 \cdot \frac{4}{17} - \frac{13}{17} = 11$$

Diese Zahl ist ganz. Also haben wir es wieder mit einem Ausnahmefall zu tun. Als wahrscheinlichste Anzahl für die Regentage treten also 11 und 12 auf.

Beispiel 2. Das Geschütz Nr. 1 möge in einem gewissen Zeitintervall 60 Schüsse abgeben können. Die Treffwahrscheinlichkeit für jeden Einzelschuss sei 0,7. Das Geschütz Nr. 2 möge in dem gleichen Zeitintervall 50 Schüsse mit einer Treffwahrscheinlichkeit von 0,8 abgeben können. Welches dieser Geschütze hat die größere wahrscheinlichste Trefferzahl?

Für das Geschütz Nr. 1 gilt: $n = 60, \quad p = 0,7, \quad np - (1 - p) = 41,7, \quad k = 42$
Für das Geschütz Nr. 2 gilt: $n = 50, \quad p = 0,8, \quad np - (1 - p) = 39,8, \quad k = 40$

Wir sehen, dass die wahrscheinlichste Trefferzahl für das Geschütz Nr. 1 etwas größer ist als für das Geschütz Nr. 2.

In der Praxis kommt es verhältnismäßig oft vor, dass die Anzahl n sehr groß ist (Massenschüsse, Massenproduktion von Erzeugnissen usw.). In einem solchen Falle wird auch das Produkt np eine sehr große Zahl (wenn nicht gerade die Wahrscheinlichkeit p außerordentlich klein ist).
Da aber in den Ausdrücken $np - (1 - p)$ und $np + p$, zwischen denen die Anzahl liegt, mit der das Ereignis mit der größten Wahrscheinlichkeit vorkommt, die zweiten Glieder kleiner als 1 sind, werden beide Zahlen und somit auch die zwischen ihnen eingeschlossene wahrscheinlichste Anzahl ziemlich dicht bei np liegen.

Gibt man z. B. 1000 Schüsse mit einer Treffwahrscheinlichkeit von 0,74 ab, so kann man als wahrscheinlichsten Wert für die Anzahl der Treffer $1000 \cdot 0,74 = 740$ annehmen.

Dieser Aussage kann man noch eine genauere Form geben. Bei n Versuchen sei k_0 die wahrscheinlichste Anzahl des Eintretens des Ereignisses. Dann ist $\frac{k_0}{n}$ der wahrscheinlichste "Anteil" für einen günstigen Ausgang dieser n Versuche. Aus der Ungleichung (7) ergibt sich:

$$p - \frac{1-p}{n} \leq \frac{k_0}{n} \leq p + \frac{p}{n}$$

Wir stellen uns vor, dass wir die Wahrscheinlichkeit p für das Eintreten des Ereignisses in einem Einzelversuch unverändert lassen, aber die Anzahl der Versuche n immer weiter vergrößern (hierbei wird sich selbstverständlich auch der wahrscheinlichste Wert k_0 vergrößern).

Die Brüche $\frac{1-p}{n}$ und $\frac{p}{n}$, die auf der linken bzw. rechten Seite der letzten Ungleichung stehen, werden dann immer kleiner und kleiner werden. Für sehr große n kann man auf diese Weise diese Brüche vernachlässigen, d. h., man kann annehmen, dass sowohl die linke als auch die rechte Seite der Ungleichung und somit auch der zwischen ihnen eingeschlossene Bruch $\frac{k_0}{n}$

etwa gleich p werden. Damit haben wir aber die Aussage:

Der wahrscheinlichste Anteil der günstigen Ergebnisse bei einer großen Anzahl von Versuchen ist praktisch gleich der Wahrscheinlichkeit für das Auftreten dieses Ereignisses bei einem Einzelversuch.

Wenn also unter gewissen Schussbedingungen die Treffwahrscheinlichkeit für einen Einzelschuss gleich 0,84 ist, dann kann man bei einer großen Anzahl von Schüssen mit großer Wahrscheinlichkeit erwarten, dass etwa 84 % der Schüsse Treffer sein werden.

Das bedeutet nun aber wiederum nicht, dass die Wahrscheinlichkeit, genau 84 % Treffer zu erzielen, groß wird. Im Gegenteil, diese "größte Wahrscheinlichkeit" selbst wird bei einer großen Schusszahl sogar sehr klein sein. (In dem Schema der Abb. 5 war die größte Wahrscheinlichkeit 0,196, wobei es sich nur um 15 Versuche handelte. Bei einer großen Versuchsanzahl wird sie aber noch bedeutend kleiner.)
Diese Wahrscheinlichkeit ist nur die "größte" im relativen Sinne: Die Wahrscheinlichkeit, 84 % Treffer zu erhalten, ist größer als die Wahrscheinlichkeit, nur 83 % Treffer zu erzielen.

Es ist andererseits aber leicht verständlich, dass bei langen Schussserien die Wahrscheinlichkeit einer speziellen Zahl von Treffern nicht von besonderem Interesse sein wird. Wenn z. B. 200 Minen gelegt werden, so wird es wenig zweckmäßig sein, die Wahrscheinlichkeit dafür zu berechnen, dass 137 von ihnen treffen, weil es praktisch gleichgültig ist, ob die Trefferzahl 137 oder 136 oder 138 oder vielleicht sogar 140 ist.
Praktisches Interesse besitzt dagegen z. B. die Frage nach der Wahrscheinlichkeit, ob mehr als 100 der 200 ausgelegten Minen treffen, oder ob die Trefferzahl zwischen 100 und 125 liegt, oder ob sie kleiner als 50 ist usw. Wie hat man nun eine solche Art Wahrscheinlichkeit zu berechnen?

Wir wollen z. B. die Wahrscheinlichkeit dafür bestimmen, dass die Trefferzahl zwischen 100 und 120 (120 einschließlich) liegt. Oder genauer: Es wird die Wahrscheinlichkeit der Ungleichung

$$100 < k \leq 120$$

gesucht, wobei k die Anzahl der Treffer ist. Um diese Ungleichung zu befriedigen, muss k gleich einer der zwanzig Zahlen 101, 102, ..., 120 sein. Auf Grund des Additionstheorems ist diese Wahrscheinlichkeit gleich

$$P\{100 < k \leq 120\} = P_{200}(101) + P_{200}(102) + ... + P_{200}(120)$$

Zur unmittelbaren Berechnung dieser Summe müssten wir vorher zwanzig Einzelwahrscheinlichkeiten vom Typ $P_n(k)$ nach Formel (3) berechnen. Dieser Methode stehen aber bei größerer Summandenzahl unüberwindliche Schwierigkeiten entgegen.

In der Praxis wird man eine Summe der angegebenen Gestalt niemals durch direkte Rechnungen bestimmen. Hierzu existieren bequeme Näherungsformeln und Tabellen. Diese Formeln und Tabellen werden auf Grund komplizierter Verfahren der mathematischen Analysis, auf die wir hier jedoch nicht weiter eingehen werden, berechnet und aufgestellt. Man kann jedoch auch durch einfache Überlegungen einige Kenntnisse über Wahrscheinlichkeiten der Gestalt $P(100 < k \leq 120)$ erhalten.
Diese elementaren Kenntnisse führen in vielen Fällen schon zur erschöpfenden Lösung des Problems. Im folgenden Kapitel werden wir uns etwas eingehender damit beschäftigen.

6 Der Bernoullische Satz

6.16 Der Inhalt des Bernoullischen Satzes

Wir betrachten noch einmal das Diagramm in der Abb. 5.
Dort sind die Wahrscheinlichkeiten $P_{15}(k)$ für das k-malige Eintreten des betrachteten Ereignisses für die verschiedenen Werte von k als vertikale Linien dargestellt.
Die Wahrscheinlichkeit, die auf irgendeinen Bereich von k-Werten kommt, d. h. die Wahrscheinlichkeit dafür, dass das uns interessierende Ereignis x-mal auftritt, wobei x irgendeine der Zahlen dieses Bereiches ist, wird nach der Additionsregel gleich der Summe der Wahrscheinlichkeiten aller Zahlen dieses Bereiches, d. h. gleich der Summe aller vertikalen Linien, die über diesem Bereich aufgetragen sind.

Die Zeichnung zeigt anschaulich, dass diese Summe für verschiedene Bereiche gleicher Länge nicht gleich zu bleiben braucht. Zum Beispiel haben die Bereiche $2 \leq k < 5$ und $7 \leq k < 10$ die gleiche Länge. Die Wahrscheinlichkeit jedes Bereiches erhalten wir als Summe der Längen der drei vertikalen Linien.
Man sieht aber sofort, dass die Summe für den zweiten Bereich wesentlich größer ist als für den ersten. Wir wissen schon, dass die Diagramme der Wahrscheinlichkeiten $P_n(k)$ für alle n qualitativ dieselbe Gestalt wie das Diagramm in Abb. 5 haben, dass nämlich die Größe $P_n(k)$ zuerst mit wachsendem k zunimmt, um dann nach Erreichen eines Größtwertes abzunehmen.

Verständlicherweise wird daher von zwei Bereichen gleicher Länge stets derjenige eine größere Wahrscheinlichkeit haben, der näher beim wahrscheinlichsten Wert k_0 liegt. Insbesondere wird einem Bereich, der symmetrisch um k_0 angeordnet ist, stets eine größere Wahrscheinlichkeit zukommen als irgendeinem anderen Bereich derselben Länge.

Es zeigt sich nun, dass man in dieser Hinsicht noch bedeutend mehr aussagen kann. Insgesamt gibt es bei n Versuchen $n+1$ Möglichkeiten für die Zahl k, d. h. für die Anzahl des Auftretens des Ereignisses.

Wir wählen einen Bereich, in dem k_0 in der Mitte liegt und der nur einen kleinen Teil, z. B. ein Hundertstel der möglichen Werte für k enthält. Unter diesen Verhältnissen ergibt sich für eine sehr große Gesamtzahl n von Versuchen, dass diesem Bereich eine erdrückende Wahrscheinlichkeit zukommt, während alle anderen statt dessen genommenen Werte der Zahl k nur eine winzige Wahrscheinlichkeit besitzen.
Obwohl also der von uns ausgewählte Bereich im Verhältnis zu n winzig klein ist (in der Zeichnung nimmt er nur den einhundertsten Teil der Länge des Diagramms ein), wird die Summe der über ihm angeordneten vertikalen Linien doch bedeutend größer als die Summe aller übrigen vertikalen Linien zusammen.

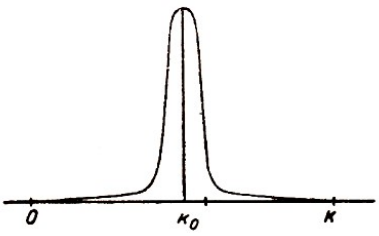


Abb. 6

Der Grund dafür ist, dass die Linien im zentralen Teil des Diagramms um viele mal größer sind als die mehr am Rande aufgetragenen.

Für große n hat das Diagramm der Größe $P_n(k)$ etwa eine Gestalt, wie sie in Abb. 6 dargestellt ist.

Praktisch bedeutet das offenbar das folgende: Wenn wir eine sehr große Anzahl von Versuchen durchführen; dann können wir mit einer Wahrscheinlichkeit, die sehr nahe bei Eine liegt, erwarten, dass die Anzahl k des Auftretens des Ereignisses A sehr nahe beim wahrscheinlichsten Wert k_0 liegt und sich davon nur um einen unbedeutenden Bruchteil der Anzahl der durchgeführten Versuche unterscheidet.

Diese Aussage, die unter dem Namen Bernoullischer Satz bekannt ist und zu Anfang des achzehnten Jahrhunderts entdeckt wurde, ist eines der wichtigsten Gesetze der Wahrscheinlichkeitsrechnung. Bis zur Mitte des letzten Jahrhunderts erforderten noch alle Beweise dieses Satzes komplizierte mathematische Hilfsmittel, und erst dem großen russischen Mathematiker P. L. Tschebyscheff gelang als erstem eine sehr einfache und kurze Begründung dieses Gesetzes. Diesen bemerkenswerten Beweis Tschebyscheffs werden wir hier auch bringen.

6.17 Der Beweis des Bernoullischen Satzes

Wir wissen schon, dass sich für eine große Anzahl n von Versuchen die wahrscheinlichste Zahl k_0 des Eintretens des Ereignisses A kaum von der Größe np unterscheidet, wobei p, wie immer, die Wahrscheinlichkeit des Ereignisses A bei einem einzelnen Versuch bedeutet.
Es genügt also vollkommen, wenn wir beweisen, dass die Anzahl k des Auftretens des Ereignisses A bei einer großen Anzahl von Versuchen sich mit sehr großer Wahrscheinlichkeit nur wenig von np unterscheidet - nicht mehr als um einen beliebig kleinen Teil der Zahl n (nicht mehr als z. B. um $0,01n$ oder 0,001 % oder allgemein nicht mehr als um εn, wobei ε eine beliebig kleine positive Zahl ist).

Mit anderen Worten: Wir müssen also zeigen, dass die Wahrscheinlichkeit

$$P\{|kn - p| > \varepsilon n\} \tag{1}$$

für hinreichend großes n beliebig klein gemacht werden kann.

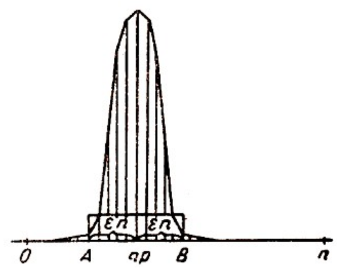


Abb. 7

Um uns von dieser Tatsache zu überzeugen, bemerken wir, dass die Wahrscheinlichkeit (1) auf Grund des Additionstheorems gleich der Summe der Wahrscheinlichkeiten $P_n(k)$ für alle jene k-Werte ist, die von np um mehr als εn entfernt sind.
In unserem Diagramm (Abb. 7) ist diese Summe gleich der Summe der Längen aller vertikalen Linien, die außerhalb des Bereiches AB liegen, sowohl links als auch rechts von ihm. Da die Gesamtsumme aller vertikalen Linien (als Summe der Wahrscheinlichkeiten eines vollständigen Systems von Ereignissen) gleich Eins ist, bedeutet das, dass fast die ganze Summe über dem Intervall AB aufgetragen ist und nur ein verschwindend kleiner Anteil auf dem Gebiet, welches außerhalb dieses Intervalls liegt.

Somit wird
$$P\{|k-np|>\varepsilon n\} = \sum_{|k-np|>\varepsilon n} P_n(k) \tag{2}$$

Wir wenden uns jetzt der Tschebyscheffschen Überlegung zu.
Da jedes Glied der hingeschriebenen Summe

$$\left|\frac{k-np}{\varepsilon n}\right| > 1 \qquad \text{und somit auch} \qquad \left(\frac{k-np}{\varepsilon n}\right)^2 > 1$$

ist, können wir diese Summe nur vergrößern, wenn jedes der Glieder $P_n(k)$ durch den Ausdruck

$$\left(\frac{k-np}{\varepsilon n}\right)^2 P_n(k)$$

ersetzt wird. Daher ist

$$P\{|k-np|>\varepsilon n\} < \sum_{|k-np|>\varepsilon n} \left(\frac{k-np}{\varepsilon n}\right)^2 = \frac{1}{\varepsilon^2 n^2} \sum_{|k-np|>\varepsilon n} (k-np)^2 P_n(k)$$

Ferner vergrößern wir diese Summe offenbar noch mehr, wenn wir zu den schon vorhandenen Summanden noch neue hinzufügen.
Wir lassen die Zahl k nicht nur den Bereich bis $np-\varepsilon n$ und von $np+\varepsilon n$ ab durchlaufen, sondern die ganze Reihe der für sie möglichen Werte, d. h. die ganze Reihe der Zahlen von 0 bis n einschließlich. Es gilt damit auch erst recht

$$P\{|k-np|>\varepsilon n\} < \frac{1}{\varepsilon^2 n^2} \sum_{k=0}^{n} (k-np)^2 P_n(k) \tag{3}$$

Die letzte Summe unterscheidet sich insbesondere dadurch von allen vorhergehenden, dass man sie genau berechnen kann. Das Tschebyscheffsche Beweisverfahren besteht gerade darin, die schwierig abzuschätzende Summe (2) durch die Summe (3) zu ersetzen, die eine genaue Berechnung gestattet.

Wir wollen diese Berechnung jetzt durchführen. Ihre Länge besagt gar nichts, sie zeigt nur die Schwierigkeit der Beweisanordnung, mit der aber jeder fertig wird, der einige Tatsachen aus der Algebra kennt. Die bemerkenswerte Idee von Tschebyscheff wurde bereits ausgenutzt. Sie besteht gerade in dem Übergang von der Gleichung (2) zur Ungleichung (3).
Leicht zu finden ist vor allem:

$$\sum_{k=0}^{n} (k-np)^2 P_n(k) = \sum_{k=0}^{n} k^2 P_n(k) - 2np \sum_{k=0}^{n} k P_n(k) + n^2p^2 \sum_{k=0}^{n} P_n(k) \tag{4}$$

Von den drei Summanden der rechten Seite ist der letzte gleich n^2p^2, weil $\sum\limits_{k=0}^{n} P_n(k)$ die Summe der Wahrscheinlichkeiten eines vollständigen Systems von Ereignissen ist. Somit haben wir also nur noch die Summen

$$\sum_{k=0}^{n} kP_n(k) \qquad \text{und} \qquad \sum_{k=0}^{n} k^2 P_n(k)$$

zu berechnen. Hierbei verschwinden in beiden Summen die Glieder für $k=0$, so dass man mit der Summation bei $i=1$ beginnen kann.

1. Zur Berechnung der beiden Summen drücken wir $P_n(k)$ mit Hilfe der Formel (4) aus Kapital 5 aus. Wir finden:

$$\sum_{k=1}^{n} kP_n(k) = \sum_{k=1}^{n} \frac{kn!}{k!(n-k)!} p^k (1-p)^{n-k}$$

Da offenbar $n! = n(n-1)!$ und $k! = k(k-1)!$ ist, erhalten wir

$$\sum_{k=1}^{n} kP_n(k) = np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

oder, indem wir in der Stimme der rechten Seite $k-1 = l$ setzen,

$$\sum_{k=1}^{n} kP_n(k) = np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!(n-1-l)!} p^l (1-p)^{n-1-l} = np \sum_{l=0}^{n-1} P_{n-1}(l)$$

Hierzu ist nur zu bemerken, dass l von 0 bis $n-1$ variiert, wenn k sich von 1 bis n ändert. Die letzte Summe

$$\sum_{l=0}^{n-1} P_{n-1}(l)$$

ist selbstverständlich wieder gleich Eins, weil sie die Summe der Wahrscheinlichkeiten eines vollständigen Systems von Ereignissen ist. Es handelt sich um alle Möglichkeiten für das Auftreten des Ereignisses l bei $n-1$ Versuchen. Für

$$\sum_{k=0}^{n} kP_n(k)$$

erhalten wir auf diese Weise den außerordentlich einfachen Ausdruck

$$\sum_{k=0}^{n} kP_n(k) = np \tag{5}$$

2. Zur Berechnung der zweiten Summe bestimmen wir zunächst die Größe

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

Da das Glied für $k = 1$ offensichtlich gleich Null wird, kann man die Summation bei $k = 2$ beginnen. Berücksichtigt man $n! = n(n-1)(n-2)!$ und $k! = k(k-1)(k-2)!$, so schließt man leicht, indem man analog dem vorhergehenden $k-2 = m$ setzt:

$$\begin{aligned}
\sum_{k=1}^{n} k(k-1)P_n(k) &= \sum_{k=2}^{n} k(k-1)P_n(k) = \sum_{k=2}^{n} \frac{k(k-1)n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= n(n-1)p^2 \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)![(n-2)-(k-2)]!} p^{k-2} (1-p)^{(n-2)-(k-2)} \\
&= n(n-1)p^2 \sum_{m=0}^{n-2} \frac{(n-2)!}{m!(n-2-m)!} p^m (1-p)^{n-2-m} \\
&= n(n-1)p^2 \sum_{m=0}^{n-2} P_{n-2}(m) = n(n-1)p^2 \qquad (6)
\end{aligned}$$

weil die letzte Summe wiederum als Summe der Wahrscheinlichkeiten eines gewissen vollständigen Systems von Ereignissen bei $n-2$ Versuchen gleich Eins ist.

Schließlich erhalten wir aus den Formeln (5) und (6):

$$\sum_{k=1}^{n} k^2 P_n(k) = \sum_{k=1}^{n} k(k-1)P_n(k) + \sum_{k=1}^{n} kP_n(k) = n(n-1)p^2 + np = n^2p^2 + np(1-p) \tag{7}$$

Damit haben wir die beiden benötigten Summen berechnet. Wir setzen die Ergebnisse (5) und (7) in die Beziehung (4) ein und erhalten somit schließlich:

$$\sum_{k=0}^{n} (k-np)^2 P_n(k) = n^2p^2 + np(1-p) - 2np \cdot np + n^2p^2 = np(1-p)$$

Diesen außerordentlich einfachen Ausdruck setzt man nun in die Ungleichung (3) ein und erhält

$$P\{|k-np| > \varepsilon n\} < \frac{np(1-p)}{\varepsilon^2 n^2} = \frac{p(1-p)}{\varepsilon^2 n} \tag{8}$$

Diese Ungleichung beweist aber alles Notwendige. In der Tat konnten wir die Zahl ε beliebig wählen, also auch beliebig klein.
Nachdem wir sie aber einmal ausgewählt haben, halten wir sie fest. Die Anzahl n der Versuche kann auf Grund unserer Behauptung aber beliebig groß sein. Dann kann aber der Bruch $\frac{p(1-p)}{\varepsilon^2 n}$ beliebig klein gemacht werden, weil der Nenner dieses Bruches mit wachsendem n immer größer wird und der Zähler sich nicht ändert.

Sei z. B. $p = 0,75$ und somit $1 - p = 0,25$ und

$$p(1-p) = 0,1875 < 0,2$$

Wir wählen $\varepsilon = 0,01$. Dann ergibt die Ungleichung (8):

$$P\left\{\left|k - \frac{3}{4}n\right| > \frac{1}{100}n\right\} < \frac{0,2}{0,0001 \cdot n} = \frac{2000}{n}$$

Wenn zum Beispiel $n = 200000$ ist, so folgt

$$P\{|k - 150000| > 2000\} < 0,01$$

Praktisch bedeutet das beispielsweise folgendes: In einem gewissen Betrieb mögen für einen feststehenden technologischen Prozess im Mittel 75 % der Erzeugnisse eine gewisse Eigenschaft besitzen (z. B. zur 1. Sorte zu gehören), dann werden von 200000 Erzeugnissen 148000 bis 152000 Erzeugnisse dieselbe Eigenschaft mit einer Wahrscheinlichkeit von mehr als 0,99 besitzen (d. h. praktisch fast sicher).

Es ist notwendig, zu diesen Ausführungen noch zwei Bemerkungen zu machen:

1. Die Ungleichung (8) gibt für die Wahrscheinlichkeit $P\{|k-np] > \varepsilon n\}$ nur eine sehr grobe Abschätzung.
Tatsächlich ist diese Wahrscheinlichkeit besonders für große n bedeutend kleiner. In der Praxis verwendet man daher auch eine genauere Abschätzung, deren Begründung jedoch wesentlich komplizierter ist.

2. Die durch die Ungleichung (8) gegebene Abschätzung wird bedeutend genauer, wenn die Wahrscheinlichkeit p sehr klein oder umgekehrt sehr dicht bei Eins gelegen ist.

Sei in dem gerade angeführten Beispiel die Wahrscheinlichkeit dafür, dass ein Erzeugnis eine gewisse Eigenschaft besitzt, gleich $p = 0,95$. Dann wird $1 - p = 0,05$ und $p(1-p) < 0,05$. So erhalten wir für $\varepsilon = 0,005$ und $n = 200000$

$$\frac{p(1-p)}{\varepsilon^2 n} < \frac{0,05 \cdot 1000000}{25 \cdot 200000} = 0,01$$

wie vorhin. Jetzt ist aber εn nicht gleich 2000, sondern nur gleich 1000. Hieraus (weil $np = 190000$ ist) schließen wir, dass die Anzahl der Erzeugnisse, die die betrachtete Eigenschaft besitzen, bei der Gesamtzahl von 200000 Erzeugnissen mit praktischer Sicherheit zwischen 189 000 und 191000 liegt.

Für $p = 0,95$ garantiert also praktisch die Ungleichung (8), dass das Intervall für die Erzeugnisse mit der uns interessierenden Eigenschaft die halbe Länge wie für $p = 0,75$ hat; denn es war ja

$$P\{|k - 190000] > 1000\} < 0,01$$

Aufgabe. Es sei bekannt, dass ein Viertel der Arbeiter eines gewissen Produktionszweiges Mittelschulbildung hat. Für eine gewisse Untersuchung wurden aufs Geratewohl 200000 Arbeiter ausgewählt. Es ist zu bestimmen

1. die wahrscheinlichste Anzahl der Arbeiter mit Mittelschulbildung von den 200000 ausgewählten, und
2. die Wahrscheinlichkeit dafür, dass die tatsächliche Anzahl solcher Arbeiter von der wahrscheinlichsten Anzahl um nicht mehr als 1,6 % abweicht.

Zur Lösung dieser Aufgabe gehen wir von der Tatsache aus, dass für jeden der aufs Geratewohl ausgewählten Arbeiter die Wahrscheinlichkeit, Mittelschulbildung zu haben, gleich ein Viertel ist (hierin liegt der Sinn des Begriffes "aufs Geratewohl"). Auf diese Weise wird in unserer Aufgabe

$$n = 200000; \quad p = \frac{1}{4}; \quad k_0 = np = 50000; \quad p(1-p) = \frac{3}{16}$$

Es wird also die Wahrscheinlichkeit dafür gesucht, dass $|k-np| < 0,016np$ oder $|k-np] < 800$ ist, wobei k die Anzahl der Arbeiter mit Mittelschulbildung ist.
Wir wählen ε derart, dass $\varepsilon n = 800$ wird. Hieraus erhalten wir: $\varepsilon = \frac{800}{n} = 0,004$. Aus Formel (8) ergibt sich dann

$$P\{|k - 50000| > 800\} < \frac{3}{16 \cdot 0,000016 \cdot 200000} \approx 0,06$$

Hieraus folgt nun

$$P\{|k - 50000| < 800\} > 0,94$$

Antwort: Der gesuchte wahrscheinlichste Wert ist gleich 50000, und die gesuchte Wahrscheinlichkeit ist größer als 0,94.
In Wirklichkeit liegt die gesuchte Wahrscheinlichkeit noch wesentlich näher bei Eins.

Zweiter Teil. Zufallsgrößen

7 Zufallsgröße und Verteilungsgesetz

7.18 Der Begriff der Zufallsgröße

In allen bisherigen Ausführungen begegneten wir solchen Größen, deren numerischer Wert nicht ein für allemal bestimmt werden konnte. Er änderte sich unter der Einwirkung irgendwelcher zufälliger Einflüsse.
So wird beispielsweise die Anzahl der Knaben unter einhundert Neugeborenen nicht für jedes Hundert ein und dieselbe sein. Oder die Länge einer Baumwollfaser einer bestimmten Sorte schwankt nicht nur für verschiedene Anbaugebiete sehr stark, sondern auch für einen Strauch und sogar für eine Samenkapsel.

Zur Erläuterung bringen wir noch einige Beispiele solcher Größen.

1. Schießt man mit dem gleichen Geschütz unter unveränderlichen Schussbedingungen auf ein und dasselbe Ziel, so beobachtet man nichtsdestoweniger, dass die Geschosse an verschiedenen Stellen einschlagen. Diese Erscheinung nennt man die "Streuung" der Geschosse. Die Entfernung des Aufschlagpunktes vom Abschussort des Geschosses ist eine Größe, die sich bei verschiedenen Schüssen in Abhängigkeit von nicht durch vorherige Berechnung beeinflussbaren zufälligen Umständen ergibt.

2. Die Geschwindigkeit eines Gasmoleküls ist nicht unveränderlich und ändert sich bei jedem Zusammenstoß mit einem anderen Molekül. Mit Rücksicht darauf, dass ja jedes Molekül mit jedem anderen Gasmolekül entweder zusammenstoßen oder nicht zusammenstoßen kann, besitzt die Geschwindigkeitsänderung einen rein zufälligen Charakter.

3. Die Anzahl der Meteoriten, die im Laufe eines Jahres in die Lufthülle der Erde eintreten und die Erdoberfläche erreichen, ist nicht konstant. Sie ist sogar bedeutenden Schwankungen unterworfen, die von einer ganzen Reihe von Ursachen zufälligen Charakters abhängen.

4. Das Gewicht von Weizenkörnern, die in einem gewissen Gebiet gezüchtet werden, ist keine bestimmte Größe, sondern ändert sich von Korn zu Korn. Da es unmöglich ist, den Einfluss aller Fakten (Qualität des Bodens, auf welcher Ähre das gegebene Korn gewachsen ist, Wasserhaushalt, Lichteinflüsse usw.), die das Wachstum des Korns beeinflussen, zu berücksichtigen, sind dies alles Größen, die sich in Abhängigkeit vom Zufall ändern.

Ungeachtet der Verschiedenartigkeit der angeführten Beispiele ergeben sie doch von dem jetzt interessierenden Standpunkt alle das gleiche Bild. In jedem dieser Beispiele hatten wir es mit Größen zu tun, die in irgendeiner Weise das Resultat der vorgenommenen Operation charakterisierten (z. B. die Zählung der Meteoriten, die Änderung der Fadenlänge).
Jede dieser Größen kann bei den verschiedenen Operationen, die natürlich unter gleichartigen Bedingungen auszuführen sind, die verschiedensten Werte annehmen.

Diese verschiedenen Werte kommen dadurch zustande, dass uns zufällige Unterschiede im Ablauf dieser Operationen bei der Kontrolle entgehen. Diese Art von Größen nennt man in der Wahrscheinlichkeitsrechnung zufällige Größen oder Zufallsgrößen. Die angeführten Beispiele reichen schon vollkommen aus, um uns von der Wichtigkeit des Studiums der Zufallsgrößen gerade für die Anwendung der Wahrscheinlichkeitsrechnung auf die verschiedensten Wissensgebiete und auf die Praxis zu überzeugen.

Eine vorgegebene Zufallsgröße zu kennen, bedeutet selbstverständlich nicht, ihren numerischen Wert zu kennen. Wissen wir z. B., dass ein Geschoss 926 m vom Abschussort entfernt aufschlägt, so ist diese Entfernung schon keine Zufallsgröße mehr, weil sie ja einen bestimmten Zahlenwert annimmt.
Was müssen wir aber nun von einer Zufallsgröße wissen, damit wir eine möglichst vollständige Kenntnis von ihr gerade als Zufallsgröße haben?

Zu diesem Zweck müssen wir offenbar alle Zahlenwerte kennen, die sie annehmen kann. Sei beispielsweise beim Schießen mit einem Geschütz unter gewissen Bedingungen die kleinste beobachtete Reichweite des Geschützes 904 m und die größte 982 m. Die Entfernung des Einschlagortes vom Abschussort des Geschosses kann also alle Werte annehmen, die zwischen diesen Grenzen liegen. Im Beispiel 3 kann offensichtlich als Anzahl der Meteoriten, die im Laufe eines Jahres die Erdoberfläche erreichen, irgendeine nichtnegative Zahl, d. h. 0, 1, 2, 3 usw. auftreten.

Die Kenntnis des Verzeichnisses der möglichen Werte einer Zufallsgröße reicht allerdings noch nicht aus, um als Material für eine praktisch notwendige Abschätzung zu dienen. Im zweiten Beispiel betrachteten wir ein Gas bei zwei verschiedenen Temperaturen. Die Zahlenwerte der Molekülgeschwindigkeiten sind bei beiden Zuständen die gleichen. Folglich gibt uns die Liste der möglichen Werte nicht die Möglichkeit, irgendeine vergleichende Aussage über diese Temperaturen zu machen.

Der Temperaturunterschied weist aber trotzdem auf einen sehr wesentlichen Unterschied im Zustand des Gases hin, auf einen Unterschied, von dem uns die Liste der möglichen Geschwindigkeitswerte der Moleküle keinerlei Vorstellung gibt. Wenn wir die Temperatur eines Gases abschätzen wollen und nur eine Liste der möglichen Geschwindigkeitswerte der Moleküle vorgegeben ist, werden wir natürlich fragen: Wie oft wird diese oder jene Geschwindigkeit beobachtet?

Mit anderen Worten, wir werden natürlich versuchen, die Wahrscheinlichkeiten der verschiedenen möglichen Werte der uns interessierenden Zufallsgröße kennenzulernen.

7.19 Der Begriff des Verteilungsgesetzes

Wir werden uns zunächst mit einem ganz einfachen Beispiel beschäftigen. Es werde ein Schießen auf eine Zielscheibe, wie sie in Abb. 8 dargestellt ist, durchgeführt.

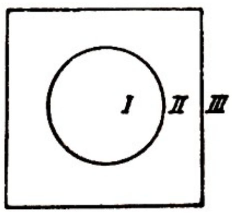


Abb. 8

Für einen Treffer im Gebiet I erhält der Schütze drei Punkte, im Gebiet II zwei und im Gebiet III einen Punkt.[10]

Als Zufallsgröße wollen wir die Anzahl der Punkte betrachten, die für einen einzelnen Schuss

[10] Der Leser könnte einwenden, dass man für einen Treffer im Gebiet III, d.h. für einen Fehlschuss, keinen Punkt ausgeben sollte. Wenn jedoch verabredet wurde, jeden Schuss mit mindestens einem Punkt zu bewerten, so muss auch ein schlechter Schuss einen Punkt bekommen. Die Zweckmäßigkeit wird auch aus dem folgenden klar werden.

ausgegeben wird. Sämtliche möglichen Werte sind hier die Zahlen 1, 2 und 3. Mit p_1, p_2 und p_3 bezeichnen wir die Wahrscheinlichkeiten dieser drei Werte.
Zum Beispiel bedeutet p_3 dann die Treffwahrscheinlichkeit für das Gebiet I der Zielscheibe. Während die möglichen Werte unserer Zufallsgrößen für alle Schützen dieselben sind, können die Wahrscheinlichkeiten p_1, p_2 und p_3 sich für verschiedene Schützen sehr wesentlich voneinander unterscheiden.

Durch diesen Unterschied wird gerade der Unterschied in der Qualität des Schießens bestimmt. So kann z. B. für einen sehr guten Schützen $p_3 = 0,8$; $p_2 = 0,2$; $p_1 = 0$, für einen mittleren Schützen $p_3 = 0,3$; $p_2 = 0,5$; $p_1 = 0,2$ und für einen ganz schlechten Schützen $p_3 = 0,1$; $p_2 = 0,3$; $p_1 = 0,6$ sein.

Werden bei einem Schießen 12 Schüsse abgegeben, so kann als mögliche Anzahl der Treffer in jedem der Gebiete I, II und III jede ganze Zahl von 0 bis 12 einschließlich auftreten. Aber diese Tatsache allein gibt uns noch nicht die Möglichkeit, etwas über die Güte des Schießens auszusagen.
Wir erhalten dagegen eine erschöpfende Auskunft, wenn uns außer den möglichen Werten auch die Wahrscheinlichkeiten dieser Werte gegeben werden, d. h. die Zahlen, die angeben, wie oft von einer Serie von 12 Schüssen diese oder jene Trefferzahl in dem einen oder anderen Gebiet erzielt wird.

Selbstverständlich wird es in irgendwelchen anderen Fällen genauso sein. Kennen wir die Wahrscheinlichkeiten der verschiedenen möglichen Werte der Zufallsgröße, so werden wir damit auch wissen, wie oft man mit dem Auftreten der mehr oder weniger vorteilhaften Werte zu rechnen hat. Das ist aber offenbar auch hinreichend für ein Urteil über die Wirksamkeit oder die gute Qualität derjenigen Operationen, mit welcher die gegebene Zufallsgröße zusammenhängt.
Die Praxis zeigt, dass die Kenntnis der Wahrscheinlichkeiten aller möglichen Werte der untersuchten Zufallsgröße tatsächlich zur Lösung jeder Frage ausreicht, die mit der Abschätzung dieser Zufallsgröße als Gradmesser der Güte der entsprechenden Operation zusammenhängt.

Somit kommen wir zu der Aussage, dass zur vollständigen Charakterisierung irgendeiner Zufallsgröße notwendig und hinreichend ist, folgendes zu kennen:

1. Ein Verzeichnis aller möglichen Werte dieser Größe,
2. die Wahrscheinlichkeit für jeden dieser Werte.

Hieraus ist ersichtlich, dass es zweckmäßig ist, eine Zufallsgröße durch eine Tabelle aus zwei Zeilen zu charakterisieren: die obere Zeile enthält in irgendeiner Reihenfolge alle möglichen Werte der Zufallsgröße und die untere ihre Wahrscheinlichkeiten, so dass unter jedem der möglichen Werte die zugehörige Wahrscheinlichkeit steht.
So kann beispielsweise für das oben betrachtete Beispiel die Anzahl der Punkte, die für einen Schuss des besten Schützen vergeben werden, als Zufallsgröße durch die Tabelle

1	2	3
0	0,2	0,8

(I)

dargestellt werden.
Im allgemeinen Fall wird also eine Zufallsgröße, deren mögliche Werte $x_1, x_2, ..., x_n$ und deren entsprechende Wahrscheinlichkeiten $p_1, p_2, ..., p_n$ sind, durch die Tabelle

x_1	x_2	...	x_n
p_1	p_2	...	p_n

gegeben.
Eine solche Tabelle vorzugeben, d.h. alle möglichen Werte der Zufallsgröße mit ihren zugehörigen Wahrscheinlichkeiten vorzugeben bedeutet, wie man sagt, das Verteilungsgesetz dieser Zufallsgröße vorzugeben.

Die Kenntnis des Verteilungsgesetzes einer gegebenen Zufallsgröße gestattet nun die Lösung aller irgendwie damit zusammenhängenden Fragen.

Aufgabe. Die Anzahl der Punkte, die für einen einzelnen Schuss des Schützen ausgegeben werden, möge das Verteilungsgesetz (I) haben. Die Punktzahl für einen anderen Schützen habe das Verteilungsgesetz

1	2	3
0,2	0,5	0,3

(II)

Es ist das Verteilungsgesetz für die Summe der an beide Schützen verteilten Punkte zu bestimmen.
Die Summe, von der hier die Rede ist, ist offensichtlich eine Zufallsgröße. Unsere Aufgabe ist es nun, ihre Tabelle aufzustellen. Zu diesem Zweck müssen wir alle Möglichkeiten beim gleichzeitigen Schießen der beiden Schützen betrachten. Wir schreiben diese Resultate in die folgende Tabelle, wobei die Wahrscheinlichkeiten jedes Resultate nach der Multiplikationsregel für unabhängige Ereignisse ermittelt werden; dabei sind x und y die Punktzahlen, die für den ersten bzw. für den zweiten Schützen ausgegeben werden.

Nr. des Resultats	x	y	$x+y$	Wahrscheinlichkeit
1)	1	1	2	$0 \cdot 0,2 = 0$
2)	1	2	3	$0 \cdot 0,6 = 0$
3)	1	3	4	$0 \cdot 0,3 = 0$
4)	2	1	3	$0,2 \cdot 0,2 = 0,04$
5)	2	2	4	$0,2 \cdot 0,5 = 0,1$
6)	2	3	5	$0,2 \cdot 0,3 = 0,06$
7)	3	1	4	$0,8 \cdot 0,2 = 0,16$
8)	3	2	5	$0,8 \cdot 0,5 = 0,4$
9)	3	3	6	$0,8 \cdot 0,3 = 0,24$

Diese Tabelle zeigt, dass die uns interessierende Summe $x + y$ nur die Werte 3, 4, 5 und 6 annehmen kann. Der Wert 2 ist für sie unmöglich, weil die Wahrscheinlichkeit gleich Null ist.[11]
Im Falle der Resultate 2 und 4 haben wir $x + y = 3$. Damit die Summe $x + y$ den Wert 3 annimmt, muss notwendigerweise das Resultat 2 oder das Resultat 4 eintreten. Nach dem Additionstheorem ist die Wahrscheinlichkeit dafür gleich der Summe der Wahrscheinlichkeiten dieser Resultate, d. h. gleich $0 + 0,04 = 0,04$.
Für die Gleichung $x + y = 4$ muss wenigstens eine der Resultate 3 , 5 oder 7 eintreten. Wir erhalten daher als Wahrscheinlichkeit dieser Gleichung (wiederum nach dem Additionstheorem) $0 + 0,1 + 0,16 = 0,26$. Auf analoge Weise finden wir auch die Wahrscheinlichkeit dafür, dass die Summe $x + y$ den Wert 5 annimmt: $0,06 + 0,4 = 0,46$.
Die Wahrscheinlichkeit des Wertes 6, die nur im Falle des Resultates 9 eintritt, ist 0,24. Also haben wir für die Zufallsgröße $x + y$ die folgende Tabelle der möglichen Werte:

3	4	5	6
0,04	0,26	0,46	0,24

(III)

[11] Man kann diese Zahl selbstverständlich auch als möglichen Wert der Größe $x + y$ annehmen, der dann die Wahrscheinlichkeit Null hat; ähnlich machten wir dies für den Wert 1 in Tabelle (I).

Durch die Angabe der Tabelle (III) ist die gestellte Aufgabe vollständig gelöst.

Die Summe aller vier Wahrscheinlichkeiten der Tabelle (III) ist gleich Eins. Diese Eigenschaft muss natürlich jedes Verteilungsgesetz besitzen, weil ja von der Summe der Wahrscheinlichkeiten aller möglichen Werte der Zufallsgröße die Rede ist, d. h. von der Summe der Wahrscheinlichkeiten eines gewissen vollständigen Systems von Ereignissen. Mit Hilfe dieser Eigenschaft der Verteilungsgesetze kann man bequem die Richtigkeit der durchgeführten Berechnungen kontrollieren.

8 Mittelwerte

8.20 Die Bestimmung des Mittelwertes einer Zufallsgröße

Jene zwei Schützen, von denen wir gerade sprachen, können, wenn sie gleichzeitig schießen, in Abhängigkeit von zufälligen Umständen entweder 3 oder 4 oder 5 oder 6 Punkte erreichen. Die Wahrscheinlichkeiten dieser vier möglichen Resultate sind der obigen Tabelle (III) zu entnehmen. Auf die Frage: "Wieviel Punkte erhalten unsere beiden Schützen bei einem (Doppel-) Schuss?" können wir keine Antwort geben, weil verschiedene Schüsse verschiedene Resultate ergeben. Wir werden uns bei der Beurteilung des Schießens unseres Paares nicht für das Ergebnis eines einzelnen Schusses (dieses Ergebnis kann zufällig sein), sondern für den Mittelwert einer ganzen Schussserie interessieren.

Wieviel Punkte erzielt unser Schützenpaar im Mittel bei einem Schuss ?
Diese Frage ist schon vollkommen vernünftig gestellt, und auf sie können wir auch eine klare Antwort geben.
Wir werden wie folgt schließen. Wenn unser Schützenpaar einhundert Doppelschüsse abgibt, so ergibt die Tabelle (III), dass

etwa 4 dieser Schüsse 3 Punkte ergeben
etwa 26 dieser Schüsse 4 Punkte ergeben
etwa 46 dieser Schüsse 5 Punkte ergeben
etwa 24 dieser Schüsse 6 Punkte ergeben

Im Mittel erhält unser Schützenpaar also für je einhundert Doppelschüsse eine Gesamtpunktzahl, die gleich der Summe

$$3 \cdot 4 + 4 \cdot 26 + 5 \cdot 46 + 6 \cdot 24 = 490$$

ist. Dividieren wir diese Zahl durch 100, dann ergibt sich, dass für einen Schuss im Mittel 4,9 Punkte verteilt werden. Das ist aber auch schon die Antwort auf die von uns gestellte Frage.

Wir könnten auch, statt die fertige Summe (490) durch 100 zu teilen (wie wir das gerade gemacht haben), noch vor der Addition jeden Summanden durch 100 dividieren. Dann liefert uns die Summe direkt die mittlere Punktzahl für einen Schuss.

Am einfachsten ist es allerdings, wenn man die Division so durchführt, dass man in jedem Summanden jeweils den zweiten Faktor durch 100 teilt; denn diese Faktoren ergaben sich durch Multiplikation der aus der Tabelle (III) entnommenen Wahrscheinlichkeiten mit 100. Man braucht also nur die ursprünglichen Wahrscheinlichkeiten zu nehmen. Für die mittlere Wahrscheinlichkeit, die auf einen Schuss kommt, erhalten wir also den Ausdruck:

$$3 \cdot 0,04 + 4 \cdot 0,26 + 5 \cdot 0,46 + 6 \cdot 0,24 = 4,9$$

Die auf der linken Seite stehende Summe ist, wie man unmittelbar sieht, aus der gegebenen Tabelle (III) nach einer sehr einfachen Regel gebildet worden:
Jeder der in der oberen Zeile der Tabelle vorkommende mögliche Wert wurde mit der unter ihm in der Tabelle stehenden Wahrscheinlichkeit multipliziert, und sämtliche so erhaltenen Produkte wurden zueinander addiert.

Wir führen jetzt dieselben Überlegungen ganz allgemein durch. Angenommen, eine gewisse Zufallsgröße sei durch die Tabelle

x_1	x_2	...	x_k
p_1	p_2	...	p_k

gegeben. Wir erinnern an folgendes:
Wenn die Wahrscheinlichkeit eines Wertes x_1 der Größe x gleich p_1 ist, so bedeutet das, dass dieser Wert x_1 in einer Serie von n Operationen etwa n_1-mal beobachtet wird, wobei $\frac{n_1}{n} = p_1$, woraus $n_1 = np_1$ folgt.
Analog wird der Wert x_2 etwa $n_2 = np_2$-mal auftreten usw. Eine Serie von n Operationen wird also im Mittel enthalten:

$n_1 = np_1$ Operationen, in denen $x = x_1$,
$n_2 = np_2$ Operationen, in denen $x = x_2$,
...
$n_k = np_k$ Operationen, in denen $x = x_k$.

Die Summe der Werte der Größe x in allen n durchgeführten Operationen wird daher etwa

$$x_1 n_1 + x_2 n_2 + ... + x_k n_k = n(x_1 p_1 + x_2 p_2 + ... + x_k p_k)$$

Für den Mittelwert $\overline{x}$ einer Zufallsgröße, die einer Einzeloperation entspricht und die aus der oben stehenden Summe durch Division durch die Anzahl n der Operationen in einer gegebenen Serie erhalten wird, ergibt sich daher

$$\overline{x} = x_1 p_1 + x_2 p_2 + ... + x_k p_k$$

Wir kommen somit zu der folgenden wichtigen Regel:
Um den Mittelwert einer Zufallsgröße zu erhalten, muss man jeden ihrer möglichen Werte mit der ihm entsprechenden Wahrscheinlichkeit multiplizieren und alle so erhaltenen Produkte zueinander addieren.

Welche Vorteile hat nun eigentlich die Kenntnis des Mittelwertes einer Zufallsgröße ? Um auf diese Frage überzeugender antworten zu können, wollen wir erst noch einige Beispiele betrachten.

Beisiel 1. Wir wenden uns noch einmal unseren zwei Schützen zu. Die an sie ausgegebenen Punktzahlen sind Zufallsgrößen, deren Verteilungsgesetze durch die Tabelle (I) für den ersten Schützen und durch die Tabelle (II) für den zweiten Schützen gegeben sind.
Ein aufmerksamer Blick auf diese Tabellen zeigt uns schon, dass der erste Schütze besser als der zweite ist. In der Tat ist die Wahrscheinlichkeit des besten Resultates (3 Punkte) bei ihm bedeutend größer als beim zweiten Schützen, während die Wahrscheinlichkeit des schlechtesten Resultates umgekehrt beim zweiten Schützen größer ist als beim ersten. Ein solcher Vergleich befriedigt uns jedoch nicht. Er trägt rein qualitativen Charakter.

Wir sehen hierbei noch kein Maß, keine Zahl, mit deren Größe wir direkt die Güte des Schießens

irgendeines Schützen abschätzen könnten, so wie man die Temperatur beispielsweise direkt zur Abschätzung der Erwärmung eines physikalischen Körpers benutzt.
Ohne ein solches abschätzendes Maß zu haben, kann stets der Fall eintreten, dass die unmittelbare Betrachtung uns keinerlei Antwort gibt oder dass sich diese Antwort als anfechtbar erweisen kann. Hätten wir es z. B. statt mit den Tabellen (I) und (II) mit den Tabellen

1	2	3
0,4	0,1	0,5

(I')

1	2	3
0,1	0,6	0,3

(II')

für den ersten Schützen für den zweiten Schützen zu tun, so wäre es äußerst schwierig, auf den ersten Blick zu entscheiden, welcher nun eigentlich der bessere Schütze ist. Tatsächlich ist das beste Resultat (3 Punkte) beim ersten Schützen wahrscheinlicher als beim zweiten. Aber gleichzeitig ist auch das schlechteste Resultat beim ersten Schützen wahrscheinlicher als beim zweiten, dagegen ist das Resultat 2 Punkte beim zweiten Schützen wahrscheinlicher.

Gemäß der oben erwähnten Regel stellen wir jetzt für jeden Schützen die Mittelwerte ihrer Punktzahlen auf:
1. Für den ersten Schützen: $1 \cdot 0,4 + 2 \cdot 0,1 + 3 \cdot 0,5 = 2,1$;
2. für den zweiten Schützen: $1 \cdot 0,1 + 2 \cdot 0,6 + 3 \cdot 0,3 = 2,2$.

Wir sehen also, dass der zweite Schütze eine etwas bessere mittlere Punktzahl aufzuweisen hat als der erste. Das heißt also praktisch, dass der zweite Schütze bei mehrmaligem Schießen im allgemeinen ein etwas besseres Resultat erzielen wird als der erste.
Jetzt können wir auch mit Sicherheit sagen, dass der zweite Schütze besser schießt als der erste. Durch den Mittelwert der Anzahl der ausgegebenen Punkte haben wir somit ein Maß, womit wir leicht und ohne irgendwelche Zweifel offen zu lassen, die Schießfertigkeit verschiedener Schützen vergleichen können.

Beispiel 2. Zur genauen Anpassung eines gewissen Einzelteils beim Zusammenbau eines Präzisionsgerätes werden in Abhängigkeit vom Zufall 1, 2, 3, 4 oder 5 Proben erforderlich sein.
Die zur Erreichung eines befriedigenden Arbeitens des Gerätes notwendige Anzahl an der Proben ist somit eine Zufallsgröße mit den möglichen Werten 1, 2, 3, 4 und 5. Die Wahrscheinlichkeiten dieser Werte mögen durch die Tabelle

1	2	3	4	5
0,07	0,16	0,555	0,21	0,1

gegeben sein. Wir könnten nun vor die Aufgabe gestellt sein, einen Monteur mit einer solchen Anzahl Einzelteile zu versehen, wie für den Zusammenbau von 20 Geräten notwendig sind.[12]

Um diese Anzahl ungefähr abzuschätzen, können wir die vorgegebene Tabelle nicht unmittelbar benutzen. Sie lehrt uns nur, dass die verschiedenen Fälle verschieden oft auftreten. Wenn wir den Mittelwert $\overline{x}$ der Anzahl der Proben x, die für ein Gerät notwendig sind, ermittelt haben und diesen Mittelwert mit 20 multiplizieren, so erhalten wir offenbar den ungefähren Wert der gesuchten Anzahl. Es ergibt sich so:

$$\overline{x} = 1 \cdot 0,07 + 2 \cdot 0,16 + 3 \cdot 0,55 + 4 \cdot 0,21 + 5 \cdot 0,1 = 3,38;$$

$$20\overline{x} = 3,38 \cdot 20 = 67,6 \approx 68$$

[12] Wir werden hierbei voraussetzen, dass diejenigen Einzelteile, die sich bei der Montage eines Gerätes als unbrauchbar erwiesen, nicht mehr weiter benutzt werden.

Damit der Monteur nun noch einen kleinen Vorrat für die Fälle hat, dass der tatsächliche Verbrauch an Einzelteilen den erwarteten übertrifft, wird man ihm praktischerweise 70 bis 75 Einzelteile geben.

In den betrachteten Beispielen verlangte die Praxis für eine gewisse Zufallsgröße eine gewisse orientierungsweise Abschätzung.
Ein einziger Blick auf die Tabelle liefert uns diese Abschätzung allerdings noch nicht. Die Tabelle sagt uns nur, dass die Zufallsgröße die und die Werte mit den und den Wahrscheinlichkeiten annehmen kann. Der nach dieser Tabelle berechnete Mittelwert der Zufallsgröße gibt uns aber schon eine solche Abschätzung, weil es gerade jener Wert ist, den unsere Zufallsgröße bei mehr oder weniger oftmaliger Wiederholung der Operation im Mittel annehmen wird.
Wir sehen also, dass der Mittelwert vom praktischen Standpunkt her eine Zufallsgröße besonders gut charakterisiert, wenn wir es mit Massenoperationen oder mehrmaligen Wiederholungen einer Operation zu tun haben.

Aufgabe 1. Es werde eine Reihe von Versuchen durchgeführt.
Ein gewisses Ereignis A möge bei diesen Versuchen stets mit der gleichen Wahrscheinlichkeit p auftreten, wobei die Ergebnisse der einzelnen Versuche voneinander unabhängig sein sollen. Man bestimme den Mittelwert für die Anzahl des Auftretens des Ereignisses A in einer Serie von n Versuchen.

Die Anzahl des Auftretens des Ereignisses A in einer Serie von n Versuchen ist eine Zufallsgröße mit den möglichen Werten $0, 1, ..., n$, wobei die Wahrscheinlichkeit des Wertes k gleich

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

ist. Daher wird der gesuchte Mittelwert gleich

$$\sum_{k=0}^{n} k P_n(k)$$

Diese Summe berechneten wir beim Beweise des Bernoullischen Satzes und sahen, dass sie gleich np ist. Seinerzeit überzeugten wir uns auch davon, dass die wahrscheinlichste Anzahl von Wiederholungen des Ereignisses A bei n Versuchen im Fall großer n nahe bei np liegt.
Jetzt sehen wir, dass die mittlere Anzahl des Auftretens des Ereignisses A für beliebiges n genau gleich np ist.
In dem vorliegenden Falle fällt also der wahrscheinlichste Wert einer Zufallsgröße mit ihrem Mittelwert zusammen. Diese Übereinstimmung gilt jedoch nicht für beliebige Zufallsgrößen.

Vor diesem Trugschluss muss nachdrücklich gewarnt werden. Im allgemeinen kann der wahrscheinlichste Wert einer Zufallsgröße sogar sehr stark von ihrem Mittelwert abweichen. So ist z. B. für eine Zufallsgröße mit dem Verteilungsgesetz

0	5	10
0,7	0,1	0,2

der wahrscheinlichste Wert 0 und der Mittelwert 2,5.

Aufgabe 2. Es wird eine Reihe von Schüssen mit einer Treffwahrscheinlichkeit von 0,8 abgefeuert. Das Schießen wird nur bis zum ersten Treffer fortgesetzt. Andererseits dürfen aber nicht mehr als vier Schüsse abgegeben werden. Es ist nun der Mittelwert für die Anzahl der abgefeuerten Schüsse zu bestimmen.

Die Anzahl der Schüsse, die unter den Schussbedingungen abgegeben werden können, muss gleich 1, 2, 3 oder 4 sein. Wir müssen die Wahrscheinlichkeiten dieser vier Werte berechnen. Damit nur ein Schuss notwendig war, muss der erste Schuss ein Treffer gewesen sein, also wird die Wahrscheinlichkeit

$$p_1 = 0,8$$

Damit zwei Schüsse erforderlich waren, muss der erste fehlgegangen und der zweite ein Treffer gewesen sein. Daher ist die Wahrscheinlichkeit nach der Multiplikationsregel für unabhängige Ereignisse gleich

$$p_2 = (1-0,8) \cdot 0,8 = 0,16$$

Damit nun drei Schüsse notwendig waren, müssen die ersten beiden Schüsse vorbeigegangen sein und der dritte getroffen haben. Daher wird

$$p_3 = (1-0,8)^2 \cdot 0,8 = 0,032$$

Vier Schüsse waren schließlich erforderlich, wenn die ersten drei Fehlschüsse gewesen sind (unabhängig davon, wie der vierte Schuss ausfällt). Daher wird

$$p_4 = (1-0,8)^3 = 0,008$$

5 20. Die Bestimmung des Mittelwertes einer Zufallsgröße 79 Die Anzahl der abgegebenen Schüsse wird also als Zufallsgröße durch das Verteilungsgesetz

1	2	3	4
0,8	0,16	0,032	0,008

bestimmt. Der Mittelwert dieser Zufallsgröße wird daher

$$1 \cdot 0,8 + 2 \cdot 0,16 + 3 \cdot 0,032 + 4 \cdot 0,008 = 1,248$$

Wenn beispielsweise ein Geschütz 100 solcher Schussserien abzugeben hat, dann werden dabei etwa $1,248 \cdot 1000 = 125$ Geschosse verbraucht.

Aufgabe 3. Eine Fläche habe die Form eines Quadrates, dessen Seiten nach vorliegenden Luftbildmessungen gleich 350 m lang sind. Die Güte der Luftaufnahme wird dadurch bestimmt, dass ein Fehler von[13]

0 m die Wahrscheinlichkeit 0,42 hat,
± 10 m die Wahrscheinlichkeit 0,16 hat,
± 20 m die Wahrscheinlichkeit 0,08 hat,
± 30 m die Wahrscheinlichkeit 0,05 hat,

Es ist der Mittelwert des Flächeninhalts dieser Fläche zu bestimmen.

In Abhängigkeit von den Zufälligkeiten einer Luftbildvermessung stellen die Seitenlängen der Fläche eine Zufallsgröße dar. Das Verteilungsgesetz dieser Zufallsgröße wird gegeben durch die Tabelle

320	330	340	350	360	370	380
0,05	0,08	0,16	0,42	0,16	0,08	0,05

[13] Das ist so zu verstehen, dass ein Fehler von +10 m die gleiche Wahrscheinlichkeit 0,16 hat wie eine Fehler von -10 m. In dem gleichen Sinne sind auch die weiteren Wahrscheinlichkeiten zu verstehen.

Hieraus könnten wir nun sofort den Mittelwert dieser Größe bestimmen. Im vorliegenden Falle besteht dazu aber gar keine Notwendigkeit.
Wir brauchen unsere Rechenregel gar nicht anzuwenden. Da die gleichen Fehler in beiden Richtungen gleichwahrscheinlich sind, ergibt sich schon aus Symmetrie, dass der Mittelwert der Länge einer Quadratseite gleich dem beobachteten Wert, d. h. gleich 350 m ist. Oder noch etwas ausführlicher: Der Ausdruck für den Mittelwert besteht aus den Gliedern:

$$\begin{aligned}
&(340+300)\cdot 0,10 = [(350-10)+(350+10)]\cdot 0,16 = 2\cdot 350\cdot 0,16;\\
&(330+370)\cdot 0,08 = 2\cdot 350\cdot 0,08,\\
&(320+380)\cdot 0,05 = 2\cdot 350\cdot 0,05.
\end{aligned}$$

Er ist daher gleich $350\cdot(0,42+2\cdot 0,10+2\cdot 0,08+2\cdot 0,05)=350$.
Man könnte nun denken, dass auf Grund derselben Symmetrieüberlegungen der Mittelwert des Flächeninhaltes ebenfalls gleich 350^2 m^2 $= 122500$ m^2 werden muss. Das wäre so, wenn der Mittelwert des Quadrates einer Zufallsgröße gleich dem Quadrat ihres Mittelwertes wäre. Das ist jedoch nicht der Fall.
In unserem Beispiel kann der Flächeninhalt des Quadrates die Werte

$$320^2, 330^2, 340^2, 350^2, 360^2, 370^2, 380^2$$

annehmen. Welcher Wert gilt nun in Wirklichkeit?
Das hängt davon ab, welcher der in der Tabelle dargestellten Fälle vorliegt. Somit sind die Wahrscheinlichkeiten dieser sieben Fälle die gleichen wie die Wahrscheinlichkeiten in der Tabelle. Oder kürzer gesagt: Das Verteilungsgesetz des Flächeninhaltes des Quadrates wird durch die Tabelle

320^2	330^2	340^2	350^2	360^2	370^2	380^2
0,05	0,08	0,16	0,42	0,16	0,08	0,05

bestimmt. Folglich ist sein Mittelwert gleich

$$320^2\cdot 0,05+330^2\cdot 0,08+340^2\cdot 0,16+350^2\cdot 0,42+360^2\cdot 0,16+370^2\cdot 0,08+380^2\cdot 0,05$$

Auch hierbei können wir wieder zur Abkürzung der Rechnung die Symmetrie ausnutzen. Es lohnt sich zu untersuchen, wie das gemacht wird, weil sich ähnliche Möglichkeiten zur Vereinfachung des öfteren ergeben. Wir können die angegebenen Ausdrücke in die folgende Form umschreiben:

$$\begin{aligned}
&350^2\cdot 0,42+(340^2+360^2)\cdot 0,16+(330^2+370^2)\cdot 0,08+(320^2+380^2)\cdot 0,05\\
&\quad = 350^2\cdot 0,42+[(350-10)^2+(350+10)^2]\cdot 0,16+[(350-20)^2+(350+20)^2]\cdot 0,08\\
&\quad + [(350-30)^2+(350+30)^2]\cdot 0,05 = 350^2[0,42+2\cdot 0,16+2\cdot 0,08+2\cdot 0,05]\\
&\quad + 2\cdot 10^2\cdot 0,16+2\cdot 20^2\cdot 0,08+2\cdot 30^2\cdot 0,05 = 350^2+2\cdot(16+32+45) = 122686.
\end{aligned}$$

Bei diesem Verfahren können alle auszuführenden Rechnungen im Kopf erledigt werden.

Wir sehen somit, dass der Mittelwert des Flächeninhaltes eines Quadrates etwas größer ist (in diesem Falle ist der Unterschied praktisch unbedeutend) als das Quadrat des Mittelwertes einer Seite (d. h. als $350^2 = 122\,500$).
Es ist leicht zu beweisen, dass dieser Tatsache eine allgemeine Regel zugrunde liegt:

Der Mittelwert der Quadrate beliebiger Zufallsgrößen ist stets größer als die Quadrate ihrer Mittelwerte. In der Tat: Sei x eine Zufallsgröße mit völlig beliebigem Verteilungsgesetz

x_1	x_2	...	x_k
p_1	p_2	...	p_k

Dann hat die Zufallsgröße x^2 das Verteilungsgesetz

x_1^2	x_2^2	...	x_k^2
p_1	p_2	...	p_k

Die Mittelwerte dieser zwei Zufallsgrößen sind dann

$$\overline{x} = x_1p_1 + x_2p_2 + ... + x_kp_k \qquad \text{bzw.} \qquad \overline{x^2} = x_1^2p_1 + x_2^2p_2 + ... + x_k^2p_k$$

Nun ist offenbar

$$\overline{x^2} - (\overline{x})^2 = \overline{x^2} - 2(\overline{x})^2 + (\overline{x})^2$$

Da nun aber $p_1 + ... + p_k = 1$ ist, können wir die drei Glieder der rechten Seite auch in der Form

$$\begin{aligned}
\overline{x^2} &= \sum_{i=1}^{k} x_i^2 p_i \\
2\overline{x}^2 &= 2(\overline{x})(\overline{x}) = 2(\overline{x})\sum_{i=1}^{k} x_i p_i = \sum_{i=1}^{k} 2\overline{x} x_i p_i \\
(\overline{x})^2 &= (\overline{x})^2 \sum_{i=1}^{k} p_i = \sum_{i=1}^{k} (\overline{x})^2 p_i
\end{aligned}$$

darstellen. Daher wird

$$\overline{x^2} - (\overline{x})^2 = \sum_{i=1}^{k} [x_i^2 + 2\overline{x}x_i + (\overline{x})^2] p_i = \sum_{i=1}^{k} (x_i - \overline{x})^2 p_i$$

Da nun alle Glieder der Summe, die auf der rechten Seite steht, nicht negativ sind; ist somit

$$\overline{x^2} - (\overline{x})^2 > 0$$

was zu beweisen war.

9 Mittelwerte von Summen und Produkten

9.21 Satz über den Mittelwert einer Summe

Man ist in der Praxis sehr oft gezwungen, den Mittelwert einer Summe von zwei (und manchmal auch einer größeren Anzahl) Zufallsgrößen, deren Mittelwerte bekannt sind, zu berechnen. Zwei Betriebe mögen beispielsweise die gleichen Produkte herstellen, wobei bekannt sei, dass der erste Betrieb im Mittel täglich 120 Erzeugnisse und der zweite 180 ausstößt.
Ob wir mit Hilfe dieser zwei Angaben den Mittelwert der Anzahl der Erzeugnisse feststellen können, der täglich von beiden Betrieben zusammen zu erwarten ist? Oder sind diese Angaben unzulänglich, und wir müssten außer den Mittelwerten noch etwas von den zwei betrachteten Zufallsgrößen wissen (z. B. vollständig ihre Verteilungsgesetze kennen)?

Es ist äußerst Wichtig, dass man zur Berechnung des Mittelwertes einer Summe nur immer die Mittelwerte der Summanden zu kennen braucht. Der Mittelwert der Summe drückt sich auf die allereinfachste Weise durch die Mittelwerte der Summanden aus: Der Mittelwert einer Summe ist stets gleich der Summe der Mittelwerte der Summanden.

Wenn also x und y zwei vollkommen willkürliche Zufallsgrößen sind, dann ist

$$\overline{x+y} = \overline{x} + \overline{y}$$

In dem oben angeführten Beispiel ist x die Anzahl der Erzeugnisse des ersten Betriebes und y die Anzahl der Erzeugnisse des zweiten: $x = 120$, $y = 180$, und somit

$$\overline{x+y} = \overline{x} + \overline{y} = 300$$

Um die behauptete Regel in allgemeiner Form beweisen zu können, nehmen wir an, dass die Zufallsgrößen x und y den Verteilungsgesetzen

$$\begin{array}{|c|c|c|c|} x_1 & x_2 & ... & x_k \\ \hline p_1 & p_2 & ... & p_k \end{array} \quad \text{(I)} \qquad \text{und} \qquad \begin{array}{|c|c|c|c|} y_1 & y_2 & ... & y_k \\ \hline q_1 & q_2 & ... & q_k \end{array} \quad \text{(II)}$$

unterworfen sind. Dann werden als mögliche Werte der Größe $x + y$ alle möglichen Summen der Form $x_i + y_i$ auftreten, wo $1 \leq i \leq k$ und $1 \leq j \leq l$ ist. Die Wahrscheinlichkeit des Wertes $x_i + y_j$ die wir mit p_{ij} bezeichnen wollen, ist uns unbekannt.

Dies ist die Wahrscheinlichkeit des doppelten Ereignisses $x = x_i, y = y_j$ d. h. die Wahrscheinlichkeit dafür, dass die Größe x den Wert x_i und die Größe y den Wert y_1 annimmt. Wenn wir annehmen könnten, dass diese zwei Ereignisse unabhängig sind, dann ergäbe sich nach der Multiplikationsregel

$$p_{ij} = p_i q_j \tag{1}$$

Wir werden jedoch keinesfalls die Unabhängigkeit dieser Ereignisse voraussetzen.
Somit wird also die Gleichung (1) im allgemeinen nicht gelten, und wir müssen damit rechnen, dass wir im allgemeinen aus der Kenntnis der Tabellen (I) und (II) noch nichts über die Größen p_{ij} aussagen können.

Nach der allgemeinen Regel ist der Mittelwert der Größe $x+y$ gleich der Summe der Produkte aller möglichen Werte dieser Größe mit den entsprechenden Wahrscheinlichkeiten, d.h.

$$\overline{x+y} = \sum_{i=1}^{k} \sum_{j=1}^{l} (x_i + y_j) p_{ij} = \sum_{i=1}^{k} x_i \left(\sum_{j=1}^{l} p_{ij} \right) + \sum_{j=1}^{l} y_I \left(\sum_{i=1}^{k} p_{ij} \right) \tag{2}$$

Wir wollen jetzt erst einmal genau die Summe $\sum\limits_{j=1}^{l} p_{ij}$ betrachten. Das ist die Summe der Wahrscheinlichkeiten aller möglichen Ereignisse der Gestalt $(x = x_i, y = y_j)$, wobei die Zahl i in allen Gliedern der Summe dieselbe ist, und die Zahl j die Reihe der für sie möglichen Werte (von 1 bis l einschließlich) durchläuft.
Da aber die Ereignisse $y = y_j$ für verschiedene j offenbar miteinander unverträglich sind, ist nach dem Additionstheorem die Summe $\sum\limits_{j=1}^{l} p_{ij}$ gerade die Wahrscheinlichkeit für das Eintreten irgendeines der l Ereignisse $(x = x_i, y = y_j)$ $(j = 1, 2, ..., l)$.

Statt zu sagen "es ist irgendeine der Ereignisse $(x = x_i, y = y_j)$ $(1 \leq j \leq l)$ eingetreten", kann man sich auch einfach so ausdrücken: "es ist das Ereignis $(x = x_i)$ eingetreten".
In der Tat:

1. Wenn eins der Ereignisse $(x = x_i, y = y_j)$ (j ganz beliebig) eingetreten ist, dann ist offensichtlich auch das Ereignis $x = x_i$ eingetreten.
2. Wenn dagegen das Ereignis $x = x_i$, eingetreten ist, dann muss auch, weil y unbedingt einen seiner möglichen Werte $y_1, y_2, ..., y_l$ annimmt, irgendeins der Ereignisse $(x = x_i, y = y_j)$ $(1 \leq j \leq l)$ eingetreten sein.
Also ist $\sum\limits_{j=1}^{l} p_{ij}$, die die Wahrscheinlichkeit für das Eintreten irgendeine der Ereignisse $(x = x_i, y = y_j)$ $(1 \leq j \leq l)$ ist, einfach die Wahrscheinlichkeit des Ereignisses $x = x_o$, d. h.

$$\sum_{j=1}^{l} p_{ij} = p_i$$

Auf völlig analoge Weise überzeugen wir uns schließlich von der Tatsache

$$\sum_{i=1}^{k} p_{ij} = q_j$$

Setzt man diese Ausdrücke nun aber in die Gleichung (2) ein, so ergibt sich

$$\overline{x+y} = \sum_{i=1}^{k} x_i p_i + \sum_{j=1}^{l} y_j q_j = \overline{x} + \overline{y}$$

was zu beweisen war.

Der Satz, der eben von uns für den Fall zweier Summanden bewiesen wurde, lässt sich auch ohne weiteres auf drei und mehr Summanden erweitern. In der Tat können wir auf Grund des eben Bewiesenen schreiben:

$$\overline{x+y+z} = \overline{x+y} + \overline{z} = \overline{x} + \overline{y} + \overline{z} \qquad \text{usw.}$$

Beispiel. Es werden n Schüsse mit den Treffwahrscheinlichkeiten $p_1, p_2, ..., p_n$ abgefeuert. Es wird der Mittelwert der gesamten Trefferzahl gesucht.

Bei einem Einzelschuss ist die Trefferzahl eine Zufallsgröße, die nur die beiden Werte 1 (Treffer) und 0 (Fehlschuss) annehmen kann. Die Wahrscheinlichkeiten dieser Werte für den ersten Schuss sind p_1 bzw. $1 - p_1$. Also wird der Mittelwert der Trefferzahl des ersten Schusses gleich

$$1 \cdot p_1 + 0 \cdot (1 - p_1) = p_1$$

Genau so schließt man, dass sie für den zweiten Schuss gleich p_2 ist, usw. Die Gesamttrefferzahl ist die Summe der Trefferzahlen in den einzelnen Schüssen. Ihr Mittelwert ist daher, wegen der gerade von uns aufgestellten Regel für die Addition von Mittelwerten, gleich der Summe der Mittelwerte der Trefferzahlen, die den Einzelschüssen entsprechen, d. h. gleich

$$p_1 + p_2 + ... + p_n$$

Damit ist aber die gestellte Aufgabe gelöst.

Wenn, insbesondere alle Schüsse die gleiche Treffwahrscheinlichkeit $(p_1 = p_2 = ... = p_n = p)$ besitzen, dann wird der Mittelwert der Gesamtzahl der Treffer gleich np. Dieses Resultat hatten wir schon erhalten. Besonders interessant ist es, die schwierigen Rechnungen, die dort zu diesem Zweck notwendig waren, mit der einfachen, keine Rechnungen erfordernden Überlegung zu vergleichen, die hier zu demselben Ergebnis führte. Wir erzielten hier jedoch nicht nur eine größere Einfachheit, sondern auch eine größere Allgemeinheit.

Bei unserer ersten Ableitung setzten wir die Unabhängigkeit der Resultate der einzelnen Schüsse voraus, und diese Methode lieferte uns eine Aussage, die nur unter dieser Voraussetzung anwendbar war. Jetzt können wir aber ohne diese Voraussetzung arbeiten, weil die Additionsregel für die Mittelwerte, mit der wir in unserer neuen Ableitung operieren, für beliebige Zufallsgrößen ohne jede Einschränkung gilt.
Obwohl die Einzelschüsse voneinander abhängig sein können, ist der Mittelwert der Gesamtzahl der Treffer immer gleich np, wenn nur die Treffwahrscheinlichkeiten für alle Schüsse untereinander gleich sind.

9.22 Satz über den Mittelwert eines Produktes

Dieselbe Frage, die wir für Summen von Zufallsgrößen beantworteten, muss auch oft für Produkte von Zufallsgrößen gelöst werden.
Es seien x und y wiederum zwei Zufallsgrößen, die den in den Tabellen (I) bzw. (II) angegebenen Verteilungsgesetzen unterliegen sollen. Dann ist das Produkt xy eine Zufallsgröße, deren mögliche Werte durch alle Produkte der Form x_iy_j $(1 \leq i \leq k,\ 1 \leq j \leq l)$ bestimmt werden.

Die Wahrscheinlichkeit des Wertes x_iy_j sei gleich p_{ij}. Die Aufgabe besteht nun darin, eine solche Regel anzugeben, die es gestatten würde, den Mittelwert $\overline{xy}$ der Größe xy in allen Fällen durch die Mittelwerte der Faktoren auszudrücken. Die Lösung dieser Aufgabe erweist sich jedoch im allgemeinen Fall als unlösbar.
Durch die Kenntnis der Mittelwerte $\overline{x}$ und $\overline{y}$ wird die Größe $\overline{xy}$ im allgemeinen nicht eindeutig bestimmt (d. h. für ein und dieselben $\overline{x}$ und $\overline{y}$ sind verschiedene Werte der Größe $\overline{xy}$ möglich). Infolgedessen kann eine allgemeine Formel, die $\overline{xy}$ durch $\overline{x}$ und $\overline{y}$ ausdrückt, nicht existieren.

Es gibt aber einen wichtigen Spezialfall, in welchem ein solcher Ausdruck angebbar ist. Der sich dann ergebende Zusammenhang trägt ebenfalls außerordentlich einfachen Charakter.
Wir werden zwei Zufallsgrößen x und y unabhängig nennen, wenn die Ereignisse $x = x_i$ und $y = y_i$ für beliebige i und j voneinander unabhängig sind, d. h., wenn die eine Zufallsgröße irgendeinen bestimmten Wert annimmt, so soll das keinerlei Einfluss auf das Verteilungsgesetz der anderen Zufallsgröße haben. Wenn die Zufallsgrößen x und y in dem gerade festgelegten Sinne unabhängig sind, so wird

$$p_{ij} = p_iq_j \qquad (i = 1, 2, ..., k; j = 1, 2, ..., l)$$

nach der Multiplikationsregel für unabhängige Ereignisse. Also gilt

$$\overline{xy} = \sum_{i=1}^{k}\sum_{j=1}^{l} x_iy_jp_{ij} = \sum_{i=1}^{k}\sum_{j=1}^{l} x_iy_jp_iq_j = \sum_{i=1}^{k} x_ip_i \sum_{j=1}^{l} y_jq_j = \overline{x} \cdot \overline{j}$$

Für unabhängige Zufallsgrößen ist der Mittelwert eines Produktes gleich dem Produkt der Mittelwerte der Faktoren.

Wie auch im Falle der Addition überträgt sich diese Regel, die von uns für ein Produkt von zwei Zufallsgrößen nachgewiesen wurde, automatisch auf ein Produkt einer beliebigen Anzahl von Zufallsgrößen. Hierbei ist nur die Unabhängigkeit dieser Faktoren notwendig, d. h. die Vorgabe irgendwelcher bestimmter Werte für einen Teil dieser Zufallsgrößen darf keinen Einfluss auf die Verteilungsgesetze der restlichen Zufallsgrößen haben.

Beispiel 1. Angenommen, es werde die Inhaltsbestimmung einer rechteckigen Fläche mittels Luftaufnahme gefordert. Die Messung der Seitenlängen dieses Rechtecks möge 72 m und 50 m ergeben haben. Das Verteilungsgesetz des Messfehlers möge auch unbekannt sein.
Es sei nur bekannt, dass Fehler gleicher Größe für eine Seite gleichwahrscheinlich sind.

Aus Symmetrieüberlegungen folgt dann ohne weiteres (und kann auch leicht bewiesen werden, siehe frühere Aufgabe 3), dass die Mittelwerte der Rechteckseiten mit den erhaltenen Messresultaten übereinstimmen.
Wenn man nun diese zwei Messresultate als voneinander unabhängige Zufallsgrößen annehmen kann, dann wird der Mittelwert des Flächeninhalts nach der gerade von uns bewiesenen Regel gleich dem Produkt der Mittelwerte der Seiten, d. h. gleich 72 m $\cdot$ 50 m $=$ 3600 m^2.

Manchmal können aber auch Gründe dafür vorliegen, dass man die Seitenmessungen als voneinander abhängig annehmen muss. Das ist beispielsweise der Fall, wenn beide Messungen mit Hilfe ein und derselben unzulänglich geprüften Vorrichtung ausgeführt werden. Wenn die Längenmessung ein Resultat liefert, welches die tatsächliche Länge bedeutend übersteigt, dann haben wir wirklich Grund vorauszusetzen, dass die Messgeräte im allgemeinen zu große Werte ergeben.
Infolgedessen erhöht sich auch die Wahrscheinlichkeit zu großer Werte für die Breitenmessung, so dass man diese beiden Messungen keinesfalls als unabhängig annehmen darf. In einem solchen Falle darf man dann auch nicht annehmen, dass der Mittelwert des Flächeninhalts gleich dem Produkt der Mittelwerte der Rechteckseiten ist. Zur Bestimmung dieses Mittelwertes sind unter diesen Umständen noch ergänzende Informationen erforderlich.

Beispiele 2. Durch einen Leiter, dessen Widerstand von zufälligen Umständen abhängt, fließe ein elektrischer Strom, dessen Stärke ebenfalls vom Zufall abhängig sei. Es sei fernerhin bekannt, dass der Mittelwert des Leiterwiderstandes 25 Ohm und die mittlere Stromstärke 6 Ampere betragen. Man berechne den Mittelwert der elektromotorischen Kraft E des durch den Widerstand fließenden elektrischen Stromes.

Nach dem Ohmschen Gesetz ist
$$E = RI$$
wobei R der Widerstand des Leiters und I die Stromstärke sind. Da nach unserer Annahme $\overline{R} = 25$, $\overline{I} = 6$ ist, ergibt sich unter der Voraussetzung, dass R und I voneinander unabhängige Größen sind:
$$\overline{E} = \overline{RI} = 25 \cdot 6 = 150 \text{ Volt}$$

10 Die Streuung und die mittlere Abweichung

10.23 Die Unzulänglichkeit des Mittelwertes zur Charakterisierung einer Zufallsgröße

Wir sahen schon mehrmals, dass der Mittelwert einer Zufallsgröße uns eine ungefähre, orientierende Vorstellung von dieser Zufallsgröße gibt. Für viele praktische Zwecke ist diese Vorstellung auch vollkommen ausreichend.

So genügte beispielsweise die Kenntnis der Mittelwerte der von zwei Schützen erreichten Punktzahlen, um ihre Schießkunst zu vergleichen. Um die Wirksamkeit zweier Verminungen eines gegebenen Feldes zu beurteilen, genügt es vollkommen, den Mittelwert des Schadens zu kennen, den diese zwei Systeme dem Gegner verursachen können usw.

In allen diesen Fällen erhalten wir wesentliche Aussagen, wenn wir unsere Zufallsgröße durch eine Zahl - den Mittelwert - beschreiben, statt sie durch ihr kompliziertes Verteilungsgesetz vorzugeben. Es sieht dann so aus, als ob wir es nicht mit einer zufälligen, sondern mit einer genau bekannten Größe mit einem vollständig bestimmten Wert zu tun hätten.

Weit öfter liegen die Dinge jedoch wesentlich anders, nämlich so, dass sich die für praktische Zwecke wichtigsten Züge einer Zufallsgröße keineswegs aus ihrem Mittelwert bestimmen lassen, sondern dass genauere Kenntnisse ihres Verteilungsgesetzes erforderlich sind. Einen typischen Fall dieser Art haben wir bei der Untersuchung der Flugweite eines Geschosses vor uns.
Es sei x die Entfernung des Aufschlagpunktes vom Abschussort des Geschosses. Bei bester Ausrichtung des Geschützes ist der Mittelwert $\overline{x}$ dieser Entfernung gleich der Entfernung des Geschützes vom Ziel.

Wir nehmen nun an, dass uns $\overline{x}$ bekannt und das Geschütz bestmöglich eingestellt worden sei. Wo werden die Geschosse einschlagen? Werden viele von ihnen das Ziel treffen? Schlagen sie in der Nähe des Ziels ein ?
Auf alle diese außerordentlich wichtigen Fragen können wir keine Antwort geben, wenn wir nur $\overline{x}$ kennen. Wir wissen nur, dass Weit- und Nahschüsse auftreten werden, und dass die Chancen für einen Nahschuss und einen Weitschuss etwa gleich groß sind, weil der Mittelwert der Flugweite mit der Zielentfernung übereinstimmt.

Die Hauptsache ist uns allerdings unbekannt: Ob die Mehrheit der Geschosse in die Nähe des Ziels fallen wird, so dass man mit einer sehr großen Trefferzahl rechnen kann, oder ob die Mehrheit der Geschosse in beiden Richtungen stark zerstreut in großem Abstand vom Ziel einschlägt, so dass man nur eine unbedeutende Trefferzahl zu erwarten hat. Beide Annahmen können offenbar für beliebig vorgegebene $\overline{x}$-Werte realisiert werden.

Zwei Geschütze, die ein und denselben Mittelwert $\overline{x}$ haben, können sich in dieser Beziehung vollkommen unterschiedlich verhalten. Das eine von ihnen kann eine bedeutend größere "Streuung" der Geschosse aufweisen als das andere. Das bedeutet, dass die Geschosse beim Schießen mit dem ersten Geschütz im Mittel bedeutend weiter nach beiden Richtungen vom Ziel abweichen werden als beim Schießen mit dem zweiten Geschütz. Das zweite Geschütz erweist sich also als besser, obwohl der Mittelwert $\overline{x}$ für beide Geschütze der gleiche ist.

Wir betrachten ein anderes Beispiel. Wir stellen uns vor, dass die Ergiebigkeit zweier Weizensorten geprüft werden soll. Der Ertrag eines Quadratmeters ist infolge zufälliger Umstände bedeutenden Schwankungen unterworfen (Niederschlagsmenge, Düngerverteilung, Sonnenbestrahlung u. a. m.) und ist daher eine Zufallsgröße.
Wir nehmen an, die mittlere Ernte sei für jede Sorte unter den gleichen Bedingungen ein und dieselbe (240 Gramm pro Quadratmeter). Kann man nun die Qualität einer Sorte nur nach dem Mittelwert des Ertrages beurteilen?

Offenbar doch nicht; denn jene Sorte ist doch von größerem wirtschaftlichen Interesse, deren Ergiebigkeit weniger von zufälligen meteorologischen oder irgendwelchen anderen zufälligen Erscheinungen abhängt. Anders ausgedrückt heißt das:
Von größerem Interesse ist die Sorte, deren Ergiebigkeit die kleinere "Streuung" aufzuweisen hat. Wir sehen also, dass bei der Untersuchung irgendeiner Weizensorte auf Ergiebigkeit die Amplitude ihrer möglichen Schwankungen keine geringere Bedeutung hat als die mittlere Ergiebigkeit.

10.24 Verschiedene Verfahren zur Messung der Streuung einer Zufallsgröße

Die angeführten sowie auch eine Reihe anderer analoger Beispiele zeigen überzeugend, dass zur Beschreibung der praktisch interessantesten Eigenschaften einer Zufallsgröße in vielen Fällen die Vorgabe ihres Mittelwertes vollkommen unzureichend ist. Diese praktisch interessierenden Züge bleiben dabei völlig unbekannt.

Um nun aber doch etwas darüber zu erfahren, muss man entweder das gesamte Verteilungsgesetz einer solchen Größe, das praktisch fast immer kompliziert und unbequem ist, kennen oder sich bemühen, außer dem Mittelwert einer solchen Größe noch eine oder zwei Zahlen ähnlicher Art einzuführen, so dass diese nicht zu große Gruppe von Zahlen in ihrer Gesamtheit eine praktisch hinreichende Charakterisierung jener Züge der zu untersuchenden Größe gibt, die wir als die wesentlichsten gekennzeichnet hatten.
Wir wollen nun untersuchen, wie sich diese letzte Möglichkeit verwirklichen lässt.

Wie die betrachteten Beispiele zeigen, erweist sich in vielen Fällen folgende Frage als die wichtigste: Wie groß sind im allgemeinen die Abweichungen der faktisch von der gegebenen Zufallsgröße angenommenen Werte von ihrem Mittelwert, d. h., wie stark gehen diese auseinander, streuen sie?
Gruppieren sie sich alle eng um den Mittelwert herum (und liegen sie somit auch dicht beieinander) oder weicht die Mehrzahl der Werte sehr stark vom Mittelwert ab (in diesem Falle werden sich auch gewisse der Werte sehr stark voneinander unterscheiden) ?

Das folgende grobe Schema gibt eine gute anschauliche Vorstellung von diesem Unterschied. Wir betrachten zwei Zufallsgrößen mit den folgenden Wahrscheinlichkeitsverteilungen:

-0,01	+0,01
0,5	0,5

(I)

-100	+100
0,5	0,5

(II)

Beide Zufallsgrößen, für die wir oben Tabellen angaben, haben den Mittelwert Null. Während aber die erste von ihnen sehr nahe bei Null (und daher auch sehr nahe beieinander) gelegene Werte annimmt, nimmt die zweite dagegen nur Werte an, die sehr wesentlich von Null (und damit auch voneinander) verschieden sind.

Für die erste Größe gibt uns die Kenntnis des Mittelwertes auch näherungsweise Anhaltspunkte für die tatsächlich möglichen Werte.
Für die zweite Größe ist der Mittelwert aber weit entfernt von den faktisch möglichen Werten und gibt für sie keinerlei Anhaltspunkte. Wir sagen, die möglichen Werte der Streuung seien im zweiten Falle unvergleichlich größer als im ersten Falle.

Wir stehen also jetzt vor der Aufgabe, eine Zahl zu finden, die uns auf vernünftige Weise ein Maß für die Streuung einer Zufallsgröße geben könnte. Wir suchen also eine Zahl, die uns angibt, wie groß die Abweichung dieser Zufallsgröße von ihrem Mittelwert im allgemeinen wird.
Die Abweichung $x - \overline{x}$ einer Zufallsgröße x von ihrem Mittelwert $\overline{x}$ ist selbst offenbar eine Zufallsgröße, ebenso der absolute Betrag $|x - \overline{x}|$ dieser Abweichung.
Das ist eine Größe, die ein Maß der Abweichung ohne Berücksichtigung des Vorzeichens gibt.
Gesucht wird also eine Zahl, die diese zufällige Abweichung $|x - \overline{x}|$ orientierungsweise charakterisieren könnte, die uns Auskunft über die etwaige Abweichung dieser Größe geben könnte.

Zur Lösung dieser Frage gibt es viele verschiedene Methoden, von denen in der Praxis die folgenden drei am gebräuchlichsten sind:

1. Die mittlere Abweichung.
Als orientierenden Näherungswert der Zufallsgröße $|x-\overline{x}|$ kann man natürlich ihren Mittelwert $\overline{|x-\overline{x}|}$ nehmen. Dieser Mittelwert des absoluten Betrages der Abweichung wird als mittlere Abweichung der Zufallsgröße x bezeichnet. Wenn eine Zufallsgröße an durch eine Tabelle

x_1	x_2	...	x_k
p_1	p_2	...	p_k

gegeben ist, dann hat die Tabelle der Größe $|x-\overline{x}|$ die Gestalt

$\lvert x_1-\overline{x}\rvert$	$\lvert x_2-\overline{x}\rvert$	...	$\lvert x_k-\overline{x}\rvert$
p_1	p_2	...	p_k

wobei $\overline{x}=\sum\limits_{k=1}^{n} x_i p_i$ ist. Für die mittlere Abweichung M_x der Größe x erhalten wir die Formel

$$M_x = \overline{|x-\overline{x}|} = \sum_{k=1}^{n} |x_i - \overline{x}| p_i$$

wobei selbstverständlich wiederum $\overline{x}=\sum\limits_{k=1}^{n} x_i p_i$ ist. Für die Zufallsgrößen, die durch die Tabellen (I) und (II) bestimmt werden, ist $\overline{x}=0$, und wir haben

$$M_{x_I} = 0,01 \qquad \text{bzw.} \qquad M_{x_{II}} = 100$$

Übrigens sind beide Beispiele trivial, weil der absolute Betrag der Abweichung in beiden Fällen nur einen einzigen Wert annehmen kann. Damit geht aber der Charakter einer Zufallsgröße verloren.

Wir berechnen noch die mittlere Abweichung der Zufallsgrößen, die mittels der Tabellen (I') und (II') bestimmt werden. Wir sahen dort, dass die Mittelwerte dieser Größen 2,1 bzw. 2,2 sind, d. h. sehr nahe beieinander liegen. Die mittlere Abweichung für die erste Größe ist gleich

$$|1-2,1| \cdot 0,4 + |2-2,1| \cdot 0,1 + |3-2,1| \cdot 0,5 = 0,9$$

und für die zweite gleich

$$|1-2,2| \cdot 0,1 + |2-2,2| \cdot 0,6 + |3-2,2| \cdot 0,3 = 0,48$$

Wir sehen also, dass die mittlere Abweichung für die zweite Größe nur etwa halb so groß ist wie für die erste. Praktisch bedeutet das offenbar, dass, obwohl beide Schützen im Mittel etwa gleiche Punktzahlen erreichen und in diesem Sinne als gleich geschickt angesehen werden können, der zweite Schütze Wesentlich gleichmäßiger schießt, dass seine Resultate bedeutend weniger streuen als die des ersten Schützen. Dieser schießt bei gleicher mittlerer Punktzahl viel ungleichmäßiger. Er erzielt oft sowohl sehr gute als auch sehr schlechte Ergebnisse.

2. Die mittlere quadratische Abweichung.
Die ungefähre Abweichung mit Hilfe der mittleren Abweichung zu messen, ist sehr natürlich, aber zugleich auch sehr unpraktisch. Rechnungen und Abschätzungen mit dem absoluten Betrag sind meistens kompliziert und manchmal sogar völlig undurchführbar. Aus diesem Grunde führt man in der Praxis für die Größe der Abweichung gewöhnlich ein anderes Maß ein.

Genauso wie die Abweichung $x-\overline{x}$ einer Zufallsgröße x von ihrem Mittelwert $\overline{x}$ ist auch das Quadrat $(x-\overline{x})^2$ dieser Abweichung eine Zufallsgröße, deren Tabelle in unseren alten Bezeichnungen die Gestalt

$(x_1-\overline{x})^2$	$(x_2-\overline{x})^2$	...	$(x_k-\overline{x})^2$
p_1	p_2	...	p_k

hat. Der Mittelwert dieses Quadrates ist daher gleich

$$\sum_{k=1}^{n}(x_i-\overline{x})^2 p_i$$

Diese Größe gibt uns eine Vorstellung davon, womit das Quadrat der Abweichung $x-\overline{x}$ ungefähr übereinstimmt. Ziehen wir nun aus dieser Größe die Quadratwurzel

$$Q_x=\sqrt{\sum_{k=1}^{n}(x_i-\overline{x})^2 p_i}$$

so erhalten wir eine Größe, die uns als Maß für die Größe der Abweichung dienen kann. Die Größe Q_x wird als mittlere quadratische Abweichung der Zufallsgröße x bezeichnet. Ihr Quadrat, d. h. die Größe Q_x^2 ist die Streuung.

Selbstverständlich hat dieses neue Maß für die Größe der Abweichung etwas künstlicheren Charakter als die oben eingeführte mittlere Abweichung. Hier gingen wir einen Umweg. Wir ermittelten zuerst einen orientierenden Näherungswert für das Quadrat der Abweichung, um danach durch Ziehen der Quadratwurzel wieder zur Abweichung selbst zurückzukehren.
Die Verwendung der mittleren quadratischen Abweichung wird aber dafür die Rechnungen, wie wir im folgenden Paragraphen sehen werden, bedeutend vereinfachen. Diese Tatsache veranlasst auch die Statistiken in der Praxis vorzugsweise die mittlere quadratische Abweichung zu benutzen.

Beispiel. Für die Zufallsgrößen, die durch die Tabellen (I') und (II') definiert sind, erhalten wir

$$Q^2_{x_{I'}}=(1-2,1)^2\cdot 0,4+(2-2,1)^2\cdot 0,1+(3-2,1)^2\cdot 0,5=0,89$$

bzw.

$$Q^2_{x_{II'}}=(1-2,2)^2\cdot 0,1+(2-2,2)^2\cdot 0,6+(3-2,2)^2\cdot 0,3=0,86$$

und somit

$$Q_{x_{I'}}=\sqrt{0,89}\approx 0,94 \qquad \text{und} \qquad Q_{x_{II'}}=0,6$$

Oben hatten wir für dieselben Größen als mittlere Abweichungen $M_{x_{I'}}=0,9$ und $M_{x_{II'}}=0,48$.
Wir sehen, dass die mittlere quadratische Abweichung für die erste Größe genau wie die mittlere Abweichung Wesentlich größer als für die zweite ist. Ob wir also die Streuung durch die mittlere oder durch die mittlere quadratische Abweichung messen, in beiden Fällen kommen wir zu denselben Aussagen: Die erste unserer beiden Größen streut mehr als die zweite.

In beiden Fällen war aber auch die mittlere quadratische Abweichung größer als die mittlere Abweichung. Man macht sich leicht klar, dass dies auch für beliebige Zufallsgrößen der Fall sein muss. In der Tat ist die Streuung Q_x^2 als Mittelwert des Quadrates der Zufallsgröße $|x-\overline{x}|$ nach der bewiesenen Regel nicht kleiner als das Quadrat des Mittelwertes M_x der Größe $|x-\overline{x}|$. Aus $Q_x^2\geq M_x^2$, folgt aber $Q_x\geq M_x$.

3. Die gemittelte (wahrscheinliche) Abweichung.
Oftmals, besonders im Militärwesen, ist noch ein anderes Verfahren zur Bestimmung des Maßes der Streuung erforderlich. Wir werden es an Hand artilleristischer Beispiele erläutern.

Angenommen, es werde ein Artillerieschießen vom Punkte O in Richtung OX durchgeführt (Abb. 9).

Abb. 9 

Die Entfernung x des Aufschlagsortes vom Abschussort des Geschosses ist eine Zufallsgröße, deren Mittelwert uns die Lage des "Treffzentrums" C ($OC = \overline{x}$) angibt. Um diesen Punkt herum, mehr oder weniger gestreut, werden die Aufschlagpunkte der einzelnen Geschosse liegen.

Die Abweichung $x - \overline{x}$ der untersuchten Zufallsgröße (Flugweite des Geschosses) von ihrem Mittelwert ist gleichzeitig die Abweichung des Aufschlagpunktes vom Treffzentrum C. Jede Abschätzung der Größe $|x-\overline{x}|$ ist daher auch gleichzeitig eine Abschätzung des Streuungsgrades. Die Streuung der Geschosse um dieses Zentrum C kann also als wichtiges Kennzeichen für die Güte des Schießens dienen.

Wir legen vom Punkte C aus nach links und rechts je eine kleine Strecke der Länge α. Dann werden in das Innere dieses so erhaltenen Intervalles der Länge 2α mit dem Mittelpunkt C nur wenige Geschosse fallen.
Oder anders ausgedrückt: die Wahrscheinlichkeit dafür, dass $|x - \overline{x}| < \alpha$, für kleines α, wird ebenfalls sehr klein.

Das so konstruierte Intervall werden wir jetzt verbreitern, indem wir die Zahl α (die doch willkürlich gewählt worden war) vergrößern. Je größer das so konstruierte Intervall wird, desto größer wird auch der Anteil der Geschosse werden, die in das Innere dieses Intervalles fallen, und desto größer wird auch die Wahrscheinlichkeit für ein einzelnes Geschoss, innerhalb des Intervalls aufzuschlagen. Wird α sehr groß, so werden praktisch alle Geschosse ins Innere dieses Intervalls fallen. Mit schrittweiser Vergrößerung der Zahl α wird somit die Wahrscheinlichkeit der Ungleichung

$$|x - \overline{x}| < \alpha$$

von Null bis Eins ansteigen. Für kleine α ist wahrscheinlicher, dass

$$|x - \overline{x}| > \alpha$$

gilt, d. h., dass ein Geschoss außerhalb des Intervalles aufschlägt.
Dann wird für große α wahrscheinlicher, dass $|x - \overline{x}| < \alpha|$ wird, d.h., dass ein Geschoss ins Innere des Intervalls fällt. Daher muss es irgendwo beim Übergang von kleinen zu größeren α-Werten ein α_0 geben, für welches ein Geschoss mit gleicher Wahrscheinlichkeit sowohl in das Innere als auch außerhalb des Intervalls der Länge $2\alpha_0$ mit dem Punkt C als Mittelpunkt fallen kann. Das heißt mit anderen Worten, dass die Ungleichungen

$$|x - \overline{x}| < \alpha_0 \qquad \text{und} \qquad |x - \overline{x}| > \alpha_0$$

gleichwahrscheinlich sind.
Somit muss die Wahrscheinlichkeit jeder Ungleichung gleich $\frac{1}{2}$ sein (wenn wir vereinbaren, die geringfügig kleine Wahrscheinlichkeit der genauen Gleichung $|x-\overline{x}| = \alpha_0$ zu vernachlässigen).

Für $\alpha < \alpha_0$ ist die zweite und für $\alpha > \alpha_0$ ist die erste Ungleichung wahrscheinlicher. Also gibt es eine eindeutig bestimmte Zahl α_0 derart, dass der absolute Betrag der Abweichung mit gleicher Wahrscheinlichkeit sowohl größer als auch kleiner als α_0 sein kann.

Wie groß nun α_0 ist, hängt von der Güte des schießenden Geschützes ab. Es ist leicht verständlich, dass die Größe α_0 als Maß für die Streuung der Geschosse dienen kann, genauso wie die mittlere oder die mittlere quadratische Abweichung.

In der Tat, wenn α_0 beispielsweise sehr klein ist, so bedeutet das, dass schon in eine sehr kleine Umgebung des Punktes C die Hälfte aller abgefeuerten Geschosse des Geschützes fallen. Das zeugt von einer verhältnismäßig unbedeutenden Streuung.
Wenn dagegen α_0 groß ist, dann müssen wir damit rechnen, obwohl wir den Punkt C mit einem großen Intervall umgeben haben, dass die Hälfte aller Geschosse außerhalb des Intervalls aufschlagen wird. Das zeigt offenbar, dass die Geschosse sehr stark um das Zentrum herum streuen.

Die Zahl α_0 wird gewöhnlich als gemittelte oder wahrscheinliche Abweichung der Zufallsgröße x bezeichnet. Die gemittelte oder wahrscheinliche Abweichung einer Zufallsgröße ist diejenige Zahl, die angibt, wann die Abweichung $x - \overline{x}$ mit gleicher Wahrscheinlichkeit dem absoluten Betrage nach sowohl größer als auch kleiner als diese Zahl sein kann. Obwohl die gemittelte Abweichung der Größe an, die wir mit E_x bezeichnen werden, für mathematische Berechnungen nicht bequemer als die mittlere Abweichung M_x und wesentlich unhandlicher als die mittlere quadratische Abweichung Q_x ist, benutzt man trotzdem bei der Artillerie zur Abschätzung aller Abweichungen die Größe E_x. Im folgenden werden wir kennenlernen, wieso dies in der Praxis keine besonderen Schwierigkeiten verursacht.

10.25 Sätze über die mittlere quadratische Abweichung

Wir überzeugen uns jetzt davon, dass die mittlere quadratische Abweichung tatsächlich besondere Eigenschaften besitzt, die sie vor jeder anderen Möglichkeit der Charakterisierung der Abweichung, der mittleren, der gemittelten (wahrscheinlichen) Abweichung u. a. m. auszeichnet. Wie wir uns später vergewissern werden, hat die folgende Aufgabe für die Anwendung grundlegende Bedeutung.

Seien $x_1, x_2, ..., x_n$ gewisse Zufallsgrößen mit den mittleren quadratischen Abweichungen $q_1, q_2, ..., q_n$. Wir setzen $x_1+x_2+...+x_n = X$ und fragen uns, wie man die mittlere quadratische Abweichung Q der Größe X finden kann, wenn $q_1, q_2, ..., q_n$ gegeben und die Zufallsgrößen x_i $(1 \leq i \leq n)$ als unabhängig vorausgesetzt werden.

Auf Grund des Satzes über Addition der Mittelwerte haben wir

$$\overline{X} = \overline{x}_1 + \overline{x}_2 + ... + \overline{x}_n$$

und folglich

$$X - \overline{X} = (x_1 - \overline{x}_1) + (x_2 - \overline{x}_2) + ... + (x_n - \overline{x}_n)$$

woraus

$$\overline{(X - \overline{X})^2} = \left[\sum_{i=1}^{n}(x_i - \overline{x}_i)\right]^2 = \sum_{i=1}^{n}(x_i - \overline{x}_i)^2 + \sum_{i=1}^{n}\sum_{k=1}^{n}(x_i - \overline{x}_k)(x_i - \overline{x}_k) \qquad (1)$$

folgt. Nun ist aber

$$\overline{(X - \overline{X})^2} = Q^2 \qquad , \qquad \overline{(x_i - \overline{x}_i)^2} = q_i^2 \quad (1 \leq i \leq n)$$

Wenden wir jetzt die Additionsregel für die Mittelwerte an, so ergibt sich

$$Q^2 = \sum_{i=1}^{n} q_i^2 + \sum_{i=1, i\neq k}^{n}\sum_{i=1}^{n}\overline{(x_i - \overline{x}_i)(x_k - \overline{x}_k)} \qquad (2)$$

Da aber die Größen x_i und x_k für $i \neq k$ nach Voraussetzung unabhängig sind, können wir die Multiplikationsregel für die Mittelwerte unabhängiger Zufallsgrößen verwenden:

$$\overline{(x_i - \overline{x}_i)(x_k - \overline{x}_k)} = \overline{(x_i - \overline{x}_i)} \cdot \overline{(x_k - \overline{x}_k)}$$

Hier sind beide Faktoren der rechten Seite gleich Null, weil z. B.

$$\overline{(x_i - \overline{x}_i)} = \overline{x_i} - \overline{x}_i = 0$$

ist. Also verschwindet in der letzten Summe der Gleichung (2) jedes einzelne Glied, und wir erhalten die Beziehung

$$Q^2 = \sum_{i=1}^{n} q_i^2$$

d.h., die Streuung einer Summe unabhängiger Zufallsgrößen ist gleich der Summe ihrer Streuungen.

Im Falle unabhängiger Zufallsgrößen schließt sich also, wie wir sehen, an die Additionsregel der Mittelwerte die sehr wichtige Additionsregel der Streuungen an. Für die mittlere quadratische Abweichung erhalten wir hieraus sofort

$$Q = \sqrt{\sum_{i=1}^{n} q_i^2} \qquad (3)$$

Diese Formel drückt einfach die mittlere quadratische Abweichung der Summe durch die mittleren quadratischen Abweichungen ihrer Summanden im Falle ihrer Unabhängigkeit aus. Das ist gerade einer der größten Vorzüge der mittleren quadratischen Abweichung im Vergleich zu der mittleren, der gemittelten (wahrscheinlichen) und anderen Abweichungen.

Beispiel 1. Für n Schüsse mit gleicher Treffwahrscheinlichkeit p ist der Mittelwert der Trefferzahl gleich np. Um orientierungsweise abschätzen zu können, wie groß die Abweichung der tatsächlichen Trefferzahl von diesem Mittelwert sein kann, suchen wir die mittlere quadratische Abweichung der Trefferzahl. Wir machen dies vor allem als Anwendungsbeispiel der Formel (3).

In der Tat kann man die Trefferzahl für n Schüsse als Summe der Trefferzahlen bei den einzelnen Schüssen auffassen. Da wir diese Zahlen als unabhängige Zufallsgrößen annehmen, können wir nach der Additionsregel für Streuungen zur Berechnung der mittleren quadratischen Abweichung der Gesamttrefferzahl die Formel (3) benutzen. Dabei sind $q_1, q_2, ..., q_n$ die mittleren quadratischen Abweichungen der Trefferzahlen bei den einzelnen Schüssen. Die Trefferzahl x_i beim i-ten Schuss wird durch die Tabelle

1	0
p	$1-p$

bestimmt. Daher wird $\overline{x}_i = p$ und

$$q_i^2 = \overline{x - \overline{x}_i}^2 = (1-p)^2 p + p^2(1-p) = p(1-p)$$

folglich ist

$$Q = \sqrt{\sum_{i=1}^{n} q_i^2} = \sqrt{np(1-p)}$$

Damit ist die gestellte Aufgabe gelöst. Wir bilden den Mittelwert np der Trefferzahl mit ihrer mittleren quadratischen Abweichung $\sqrt{np(1-p)}$.

Hieraus ergibt sich, dass für große n (d. h. für eine große Anzahl von Schüssen) letztere bedeutend kleiner als die erste ist und nur einen Bruchteil davon ausmacht. So ist für $n = 900$, $p = \frac{1}{2}$ der Mittelwert der Trefferzahl gleich 450 und die mittlere quadratische Abweichung gleich $\sqrt{900 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 15$. Also wird die tatsächliche Trefferzahl nur um etwa 3 bis 4 % von ihrem Mittelwert abweichen.

Beispiel 2. Wir stellen uns vor, dass ein gewisser Mechanismus aus n Einzelteilen montiert werden soll. Diese Einzelteile sollen alle auf einer Achse fest aneinander angebracht werden. Über alle Teile soll noch eine gewisse Hülle geschoben werden (siehe Abb. 10).

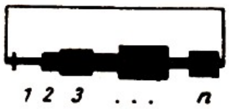


Abb. 10

Die Länge eines jeden Einzelteiles kann etwas von der entsprechenden Norm verschieden sein und ist daher eine Zufallsgröße. Wir setzen diese Zufallsgrößen als unabhängig voraus. Die mittleren 100 Kap. 10. Die Streuung und die mittlere Abweichung Längen der Einzelteile und die Streuung dieser Längen seien $a_1, a_2, ..., a_n$ bzw. $q_1, q_2, ..., a_n$. Der Mittelwert und die Streuung der Länge einer Kette aus n Einzelteilen sind dann

$$a = \sum_{i=1}^{n} a_k \qquad \text{und} \qquad q = \sum_{i=1}^{n} q_k^2$$

Wenn insbesondere $n = 9$, $a_1 = a_2 = ... = a_9 = 10$ cm und $q_1 = q_2 = ... = q_9 = 0,2$ cm sind, wird $a = 90$ cm und $q = \sqrt{9 \cdot (0,2)^2} = 0,6$ cm.
Wenn also die Länge eines Einzelteiles im Mittel von ihrem Mittelwert um 2 % abweicht, dann wird die ganze Kette aus diesen Einzelteilen im Mittel nur etwa um 2/3 % von ihrem Mittelwert abweichen.

Diese Tatsache - die Verringerung des relativen Fehlers bei der Addition von Zufallsgrößen - spielt eine außerordentliche Rolle bei der Montage von Präzisionsmechanismen. Würden sich die Abweichungen der Einzelteile von ihren normalen Maßen nicht gegenseitig kompensieren, so traten bei der Montage der Mechanismen ständig Fälle auf, in denen das umhüllende Einzelteil nicht die zu umhüllende Kette umfassen würde, oder es ergäbe sich umgekehrt ein außerordentlich großer Spielraum. In beiden Fällen erhielten wir offensichtlich Ausschuss.
Gegen diesen Ausschuss auf dem Wege einer Verkleinerung der "Toleranz" der Einzelteile zu kämpfen, d. h. die zulässigen Abweichungen der tatsächlichen Maße der Einzelteile herabzusetzen, wäre unzweckmäßig, weil eine relativ kleine Steigerung der Herstellungsgenauigkeit eine sehr große Erhöhung der Produktionskosten verursachen würde.[14]

Beispiel 3. Angenommen, an einer gewissen Größe werden unter unveränderlichen Bedingungen n Messungen vorgenommen. Als Resultat einer ganzen Reihe von Ursachen (Zustand der Vorrichtung, Beobachter, Schwankungen im Luftdruck, Staubgehalt der Luft u. a. m.) werden verschiedene Messungen im allgemeinen auch verschiedene Ergebnisse liefern, die auf zufälligen Messfehlern beruhen.
Wir werden die Ergebnisse der Messungen mit $x_1, x_2, ..., x_n$ bezeichnen, indem wir jedem x als Index die Nummer der Messung anhängen. Der Mittelwert für alle diese Messungen ist ein und

[14] Die Technik stellte in der letzten Zeit die Forderung auf, eine Theorie der Toleranzen zu entwickeln. Sie wird im wesentlichen auf Überlegungen und Folgerungen der Wahrscheinlichkeitsrechnung beruhen. Diese Theorie der Toleranzen wurde gleich von sowjetischen Wissenschaftlern aufgegriffen und wird nun von ihnen weiterentwickelt.

derselbe, nämlich $\overline{x}$. Die mittlere quadratische Abweichung q kann man ebenfalls ohne weiteres für alle Messungen als gleich voraussetzen, weil die Messungen ja unter unveränderlichen Bedingungen ausgeführt werden sollen. Schließlich nehmen wir die Größen $x_1, x_2, ..., x_n$ als unabhängig an.

Wir betrachten jetzt das arithmetische Mittel

$$\xi = \frac{x_1 + x_2 + ... + x_n}{n}$$

der n Messergebnisse. Das ist eine Zufallsgröße. Es seien ihr Mittelwert und ihre mittlere quadratische Abweichung zu bestimmen. Nach der Additionsregel ist

$$\overline{\xi} = \frac{1}{n}\overline{(x_1 + x_2 + ... + x_n} = \frac{1}{n}(\overline{x}_1 + \overline{x}_2 + ... + \overline{x}_n) = \frac{1}{n}(n\overline{x}) = \overline{x}$$

d. h. der Mittelwert ist, wie es offenbar auch sein muss, derselbe, wie für jede einzelne Messung. Ferner ist die mittlere quadratische Abweichung der Summe $x_1 + x_2 + ... + x_n$ nach der Additionsregel für die Streuung (3) gleich

$$Q = \sqrt{nq^2} = q\sqrt{n}$$

und die mittlere quadratische Abweichung der Größe ξ, die nur $\frac{1}{n}$ dieser Summe ist, wird gleich

$$\frac{Q}{n} = \frac{q}{\sqrt{n}}$$

Wir sind so zu dem sehr wichtigen Ergebnis gekommen: Das arithmetische Mittel von n unabhängigen und gleichverteilten Zufallsgrößen hat:

a) einen Mittelwert, der gleich dem Mittelwert der Größen ist, aus denen er gebildet wird,
b) eine mittlere quadratische Abweichung, die um $\sqrt{n}$ mal kleiner ist als jede der Größen, aus denen sie gebildet wird.

Wenn der Mittelwert der gemessenen Größe $x = 200$ m und die mittlere quadratische Abweichung $q = 5$ m sind, dann wird das arithmetische Mittel ξ von 100 Einzelmessungen einen Mittelwert von 200 m haben. Die mittlere quadratische Abweichung wird aber um $\sqrt{100} = 10$ mal kleiner als für eine Einzelmessung, d. h. gleich $\frac{q}{10} = 0,5$ m sein.
Wir haben also allen Grund zu erwarten, dass das arithmetische Mittel von 100 Einzelmessungen wesentlich dichter beim Mittelwert 200 m liegt als das Ergebnis irgendeiner Einzelmessung. Das arithmetische Mittel einer großen Anzahl unabhängiger Zufallsgrößen besitzt eine viele Male kleinere Streuung als jede dieser Einzelgrößen.

11 Das Gesetz der großen Zahlen

11.26 Die Tschebyscheffsche Ungleichung

Wir sprachen schon oft davon, dass die Kenntnis irgendeiner mittleren Abweichung (z. B. der mittleren quadratischen) einer Zufallsgröße eine ungefähre Vorstellung von den Abweichungen gibt, mit denen wir tatsächlich rechnen müssen. Diese Bemerkung selbst enthält jedoch noch keinerlei quantitative Abschätzungen und gibt keine Möglichkeit, angenähert zu berechnen, welche Wahrscheinlichkeit große Abweichungen haben. Dazu können wir aber auf Grund der folgenden einfachen Überlegung gelangen.

Sie wurde zuerst von Tschebyscheff angestellt. Wir gehen von dem Ausdruck für die Streuung einer Zufallsgröße x aus:

$$Q_x^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 p_i$$

Es sei α eine beliebige positive Zahl. Wenn wir aus dieser Summe alle Glieder weglassen, in denen $|x_i - \overline{x}| \leq \alpha$ ist, und nur jene stehenlassen, in denen $|x_i - \overline{x}| > \alpha$ ist, so kann sich diese Summe nur verkleinern:

$$Q_x^2 \geq \sum_{|x_i - \overline{x}| > \alpha} (x_i - \overline{x})^2 p_i$$

Diese Summe wird aber noch mehr verkleinert, wenn wir in jedem Glied den Faktor $(x_i - \overline{x})^2$ durch die kleinere Größe α^2 ersetzen:

$$Q_x^2 \geq \alpha^2 \sum_{|x_i - \overline{x}| > \alpha} p_i$$

Die jetzt auf der rechten Seite stehende Summe ist die Summe der Wahrscheinlichkeiten aller jener x_i der Zufallsgröße x, die von x nach irgendeiner Seite um mehr als α abweichen.
Nach der Additionsregel ist das gerade die Wahrscheinlichkeit dafür, dass die Größe zu irgendeinen dieser Werte annimmt. Mit anderen Worten, es ist die Wahrscheinlichkeit $P\{|x - \overline{x}| > \alpha\}$ dafür, dass die tatsächliche Abweichung größer als α ist. Also erhalten wir

$$P\{|x - \overline{x}| > \alpha\} \leq \frac{Q_x^2}{\alpha^2} \tag{1}$$

Hiermit können wir aber die Wahrscheinlichkeit einer Abweichung, die größer als eine beliebig vorgegebene Zahl α ist, abschätzen, wenn nur die mittlere quadratische Abweichung Q_x bekannt ist.

Oft erweist sich jedoch die Abschätzung, die man durch die "Tschebyscheffsche Ungleichung" (1) erhält, als zu grob. Manchmal kann sie aber verwendet werden und ist dann äußerst bequem. Ihre theoretische Bedeutung ist außerordentlich groß.

Am Ende des vorigen Paragraphen betrachteten wir ein Beispiel: Mittelwert der Messresultate 200 m, mittlere quadratische Abweichung 5 m. Unter diesen Voraussetzungen müsste die Wahrscheinlichkeit einer tatsächlichen Abweichung um mehr als 3 m ganz merklich sein (man könnte denken, dass sie größer als die Hälfte ist; ihren genauen Wert kann man selbstverständlich nur dann bestimmen, wenn das Verteilungsgesetz der Messresultate vollständig bekannt ist).

Wir sahen aber schon, dass für das arithmetische Mittel von 100 Messungen die mittlere quadratische Abweichung 0,5 m ist. Hieraus folgt wegen der Ungleichung (1)

$$P\{|\xi - 200| > 3\} \leq \frac{0,5^2}{3^2} = \frac{1}{36} \approx 0,03$$

Für das arithmetische Mittel von 100 Messungen eine Abweichung von mehr als 3 m zu erhalten, ist schon ziemlich unwahrscheinlich (sie ist tatsächlich noch kleiner, als die von uns ermittelte Grenze, so dass man praktisch mit der Unmöglichkeit einer solchen Abweichung rechnen kann).

Im Beispiel 1 auf den Seiten 67-68 erhielten wir als Trefferzahl bei 900 Schüssen den Mittelwert

450 und eine mittlere quadratische Abweichung 15. Die Wahrscheinlichkeit dafür, dass die tatsächliche Anzahl m der Treffer zwischen beispielsweise 400 und 500 liegt (d. h. $|m-450| \leq 50$), wird nach der Tschebyscheffschen Ungleichung:

$$P\{|m-450| \leq 50\} = 1 - P\{|m-450| > 50\} \geq 1 - \frac{15^2}{50^2} = 0,91$$

Diese Wahrscheinlichkeit ist aber in Wirklichkeit noch größer.

11.27 Das Gesetz der großen Zahlen

Es seien n unabhängige Zufallsgrößen $x_1, x_2, ..., x_n$ mit den gleichen mittleren quadratischen Abweichungen q gegeben. Das arithmetische Mittel dieser Zahlen,

$$\xi = \frac{x_1 + x_2 + ... + x_n}{n}$$

hat, wie wir schon sahen, den Mittelwert a und die mittlere quadratische Abweichung $\frac{q}{\sqrt{n}}$. Daher erhalten wir mittels der Tschebyscheffschen Ungleichung für beliebige positive α

$$P\{|\xi - a| > \alpha\} \leq \frac{q^2}{\alpha^2 n} \tag{2}$$

Wir mögen es beispielsweise mit dem arithmetischen Mittel von n Messungen einer gewissen Größe zu tun haben. Genau wie vorher sei $q = 5$ m und $a = 200$ m. Dann erhalten wir:

$$P\{|\xi - 200| > \alpha\} \leq \frac{25}{\alpha^2 n}$$

Wir können α sehr klein wählen, z. B. 0,5 m. Dann ist

$$P\{|\xi - 200| > 0,5\} \leq \frac{100}{n}$$

Wenn die Anzahl n der Messungen sehr groß wird, dann muss die rechte Seite unserer Ungleichung sehr klein werden. Für $n = 10000$ ist sie gleich 0,01, und für das arithmetische Mittel von 10000 Messungen erhalten wir

$$P\{|\xi - 200| > 0,5\} \leq 0,01$$

Wir wollen verabreden, solche wenig wahrscheinlichen Ereignisse zu vernachlässigen. Dann können wir sagen, dass das arithmetische Mittel von 10000 Messungen sicher nicht mehr als um 50 cm nach irgendeiner Seite von 200 m abweicht. Wenn wir eine noch größere Genauigkeit haben wollen, z. B. 10 cm, müssten wir $\alpha = 0,1$ m wählen und erhielten

$$P\{|\xi - 200| > 0,1\} \leq \frac{25}{0,01n} = \frac{2500}{n}$$

Um jetzt die rechte Seite kleiner als 0,01 zu machen, müssten wir nicht nur 10000 (das ist jetzt unzureichend), sondern 250000 Messungen ausführen. Wie klein α auch gewählt sein mag, stets kann man die rechte Seite der Ungleichung (2) beliebig klein machen, wir müssen n nur hinreichend groß wählen.
Bei hinreichend großem n können wir also annehmen, dass mit beliebig großer Annäherung an das sichere Ereignis die umgekehrte Ungleichung gilt:

$$|\xi - a| \leq \alpha$$

Wenn die Zufallsgrößen $x_1, x_2, ..., x_n$ unabhängig sind, und wenn sie alle den gleichen Mittelwert a und dieselbe mittlere quadratische Abweichung haben, dann wird die Größe

$$\xi = \frac{x_1 + x_2 + ... + x_n}{n}$$

für hinreichend großes n sich mit einer beliebig nahe bei Eins liegenden Wahrscheinlichkeit beliebig wenig von a unterscheiden.

Das ist der einfachste Spezialfall eines der wichtigsten Gesetze der Wahrscheinlichkeitsrechnung, des sogenannten Gesetzes der großen Zahlen. Es wurde Mitte des vergangenen Jahrhunderts von dem bekannten russischen Mathematiker Tschebyscheff entdeckt.[15]
Der tiefe Inhalt dieses bemerkenswerten Gesetzes besteht nun in folgendem:

Eine einzelne Zufallsgröße kann (wie wir wissen) sehr oft Werte annehmen, die weit von ihrem Mittelwert entfernt sind (eine große Streuung haben). Das arithmetische Mittel einer großen Anzahl von Zufallsgrößen verhält sich in dieser Beziehung völlig anders. Diese Größe streut sehr wenig. Sie nimmt mit sehr großer Wahrscheinlichkeit nur solche Werte an, die sehr nahe bei ihrem Mittelwert liegen.
Das liegt letztlich daran, dass die zufälligen Abweichungen nach dieser oder jener Seite sich gegenseitig aufheben. Infolgedessen wird auch die Summenabweichung in den meisten Fällen sehr klein werden.

Eine wichtige und in der Praxis oft anzutreffende Anwendung der Resultate des eben bewiesenen Tschebyscheffschen Satzes besteht darin, dass man aus einer relativ kleinen Probe (Auswahl) auf die Qualität einer größeren Menge eines homogenen Materials schließt.[16]

So kann man z. B. einen Ballen Baumwolle beurteilen, indem man an einer zufälligen Stelle ein kleines Bündel (Stapel) herauszieht und dieses beurteilt. Um eine große Sendung Saatkörner zu beurteilen, entnimmt man der Sendung an mehreren zufälligen Stellen ein Purka (Maß).[17]
An Hand dieser an verschiedenen Stellen der Sendung entnommenen Körner wird die Qualität der Sendung festgestellt.

Urteile über die Qualität einer Produktion, die auf Grund solcher Stichproben gefällt wurden, besitzen große Genauigkeit. Obwohl beispielsweise die Anzahl der Körner, die im Purka entnommen wurden, im Vergleich zur gesamten Körnermenge sehr klein ist, besteht gemäß dem Gesetz der großen Zahlen doch die Möglichkeit, das mittlere Gewicht eines Kornes sehr genau zu bestimmen.
Genauso ist es möglich, an Hand eines kleinen Stapels, der etwa einhundert Fäden enthält, die alles in allem nur etwa den zehnten Teil eines Gramms wiegen, die Qualität eines 20pudigen Baumwollballens zu beurteilen.[18]

11.28 Beweis des Gesetzes der großen Zahlen

Bis jetzt betrachteten wir nur solche Fälle, in denen alle Größen $x_1, x_2, ...$ den gleichen Mittelwert und dieselbe mittlere quadratische Abweichung haben. Das Gesetz der großen Zahlen gestattet jedoch noch viel allgemeinere Anwendungen.

[15] Es geht genau genommen auf Bernoulli zurück; die Bezeichnung wurde von Poisson eingeführt (d. Red.)
[16] Genau genommen wird hier schon das "starke" Gesetz der großen Zahlen benutzt (d. Red).
[17] Mit einem Purka werden etwa 100 bis 200 Gramm entnommen. Die ganze Sendung kann dagegen zehn oder vielleicht sogar einhundert Tonnen Saatkörper enthalten. Purka ist ein Schöpflöffel, um Stichproben zur Qualitätsprüfung zu entnehmen.
[18] 1 Pud = 16,38 kg (altrussisches Maß).

Wir werden jetzt Fälle betrachten, in denen die Mittelwerte der Größen $x_1, x_2, ..., x_n$ irgendwelche Zahlen sind (sie sollen mit $a_1, a_2, ...$ bzw. a_n bezeichnet werden), die im allgemeinen voneinander verschieden sind. Der Mittelwert der Größe

$$\xi = \frac{x_1 + x_2 + ... + x_n}{n}$$

wird dann offenbar

$$A = \frac{1}{n}(a_1 + a_2 + ... + a_n)$$

wobei nach der Tschebyscheffschen Ungleichung (1) für beliebiges positives α

$$P\{|\xi - A| > \alpha\} \leq \frac{Q_\xi^2}{\alpha^2} \tag{3}$$

ist. Wir sehen, dass sich alles auf die Abschätzung der Größe Q_ξ^2 zurückführen lässt. Diese Größe ist aber hier fast genauso einfach abzuschätzen, wie in dem vorher betrachteten Spezialfall. Q_ξ^2 ist die Streuung der Zufallsgröße ξ, die gleich der durch n dividierten Summe der n unabhängigen Zufallsgrößen x_i ist. (Die Voraussetzung der Unabhängigkeit behalten wir natürlich bei.)
Nach der Additionsregel für die Streuung erhalten wir daher

$$Q_\xi^2 = \frac{1}{n^2}(q_1^2 + q_2^2 + ...q_n^2)$$

wobei $q_1, q_2, ..., q_n$ die mittleren quadratischen Abweichungen der Größen $x_1, x_2, ..., x_n$ sind.
Es wird jetzt vorausgesetzt, dass diese mittleren quadratischen Abweichungen im allgemeinen voneinander verschieden sind. Wir verlangen jedoch, dass die mittleren quadratischen Abweichungen, wie viele Größen wir auch nehmen (d.h. wie groß die Anzahl n auch sei), stets kleiner als eine gewisse positive Zahl b sind.

In der Praxis ist diese Bedingung immer erfüllt, weil nur Größen fast gleichen Typs zusammengefasst werden, und ihr Streuungsgrad für verschiedene Größen sich nicht zu sehr ändert. Wir verlangen also, dass $q_i < b$ $(i = 1, 2, ...)$ ist. Dann ergibt aber die letzte Gleichung

$$Q_\xi^2 < \frac{1}{n^2} n b^2 = \frac{b^2}{n}$$

Infolgedessen können wir aus der Ungleichung (3) schließen, dass

$$P\{|\xi - A| > \alpha\} < \frac{b^2}{n\alpha^2}$$

ist. Wie klein α auch sein möge, nimmt man nur eine hinreichend große Anzahl n von Zufallsgrößen, so kann die rechte Seite dieser Ungleichung beliebig klein gemacht werden. Das beweist aber offenbar gerade das Gesetz der großen Zahlen in dem von uns jetzt betrachteten allgemeinen Fall.

Wenn die Größen $x_1, x_2, ...$ unabhängig sind und ihre mittleren quadratischen Abweichungen sämtlich kleiner als eine feste positive Zahl b sind, dann kann man das arithmetische Mittel

$$\xi = \frac{x_1 + x_2 + ... + x_n}{n}$$

für hinreichend großes n mit einer Wahrscheinlichkeit, die beliebig nahe bei Eins liegt, beliebig genau annähern.

Das ist gerade das Gesetz der großen Zahlen in der allgemeinen Form, wie es von Tschebyscheff formuliert worden war.
An dieser Stelle möchten wir auf einen wesentlichen Umstand aufmerksam machen. Wir nehmen an, dass eine Messung einer gewissen Größe a vorgenommen werden soll. Bei Wiederholung der Messungen unter gleichen Bedingungen erhält der Beobachter keine vollkommen übereinstimmenden numerischen Resultate $x_1, x_2, ..., x_n$. Als Näherungswert für a wird das arithmetische Mittel

$$a \sim \frac{1}{n}(x_1 + x_2 + ... + x_n)$$

genommen.
Kann man nun a beliebig genau ermitteln, wenn nur genügend viele Messungen ausgeführt werden?

Wenn die Messungen ohne systematische Fehler ausgeführt werden, d. h. wenn

$$\overline{x}_k = a \qquad (\text{für } k = 1, 2, ..., n)$$

ist und wenn die Werte selbst keine Unbestimmtheit besitzen, dann ist das der Fall. Wir können a also beliebig genau bestimmen, wenn die Werte am Messgerät tatsächlich abgelesen werden können.
Ist aber das Gerät so konstruiert, dass die Ablesegenauigkeit nicht größer als eine gewisse Größe δ ist (weil beispielsweise der Strichabstand auf der Skala, an der die Ablesung vorgenommen wird, gleich δ ist), so kann die berechnete Genauigkeit verständlicherweise nicht größer als $\pm\delta$ werden. Es ist auch klar, dass das arithmetische Mittel in diesem Falle dieselbe Unbestimmtheit δ besitzen wird, ebenso jedes der x_k.

Diese Bemerkung zeigt folgendes: Wenn ein Gerät die Messresultate mit einer gewissen Unbestimmtheit δ liefert, so ist die Hoffnung, den Wert a nach dem Gesetz der großen Zahlen mit einer größeren Genauigkeit zu bestimmen, eine Illusion und die Durchführung einer solchen Berechnung ein leerer arithmetischer Zeitvertreib.

12 Die Normalgesetze (Gausssche Verteilungen)

12.29 Die Aufgabenstellung

Wir haben gesehen, dass der Verlauf von sehr vielen Naturerscheinungen und Produktionsprozessen sowie gewissen militärisch-technischen Vorgängen von irgendwelchen Zufallsgrößen abhängt. Oft ist alles, was wir über diese Zufallsgrößen einer Erscheinung, eines Prozesses oder einer Operation in Erfahrung bringen können, ihr Verteilungsgesetz, d. h. die Liste ihrer möglichen Werte und der zugehörigen Wahrscheinlichkeiten.

Kann eine Größe unendlich viele Werte annehmen (Flugweite eines Geschosses, Größe der Messfehler usw.), so ist es vorteilhafter, nicht die Wahrscheinlichkeiten einzelner Werte, sondern die Wahrscheinlichkeiten ganzer Bereiche anzugeben (z.B. die Wahrscheinlichkeit dafür, dass das Geschoss in dem Bereich von 720 bis 740 m, von 740 bis 760 m usw. aufschlägt).
Diese Überlegung ändert jedoch nichts am Wesen der Sache. Um eine Zufallsgröße zu kennen, müssen wir uns eine möglichst genaue Darstellung ihres Verteilungsgesetzes verschaffen.

Würden wir ohne jede Überlegung und Mutmaßung allgemeinen Charakters an eine Zufallsgröße herangehen und ohne jede vorbereitende Voraussetzung versuchen, auf rein experimentellem

Wege alle Züge ihres Verteilungsgesetzes zu bestimmen, so standen wir vor einer außerordentlich schwierigen, praktisch unlösbaren Aufgabe.
In jedem neuen Fall wäre eine ungeheure Anzahl Experimente notwendig, um die wichtigsten Züge des neuen und unbekannten Verteilungsgesetzes zu bestimmen.

Daher gehen die Bemühungen der Wissenschaft dahin, wenn möglich für alle, wenigstens aber für eine große Klasse der in der Praxis vorkommenden Zufallsgrößen allgemeine Typen von Verteilungsgesetzen zu finden und sie auf vernünftige Weise zu charakterisieren.

Theoretisch waren einige solcher Typen schon frühzeitig angegeben und durch Versuche nachgeprüft worden. Es versteht sich, inwiefern es vorteilhaft ist, auf Grund von theoretischen Überlegungen und der ganzen vorherigen Erfahrung voraussagen zu können, von welchem Typ das Verteilungsgesetz einer neu auftretenden Zufallsgröße sein muss.
Wenn eine solche Vermutung sich als richtig erweist, so genügt gewöhnlich schon eine ganz geringe Zahl von Versuchen oder Beobachtungen, um alle notwendigen Züge des gesuchten Verteilungsgesetzes festzustellen.

Theoretische Untersuchungen zeigten, dass wir in sehr vielen praktischen Fällen mit hinreichender Begründung Verteilungsgesetze eines ganz bestimmten Typs erwarten dürfen. Es handelt sich dabei um die sog. Normalverteilung.
Von diesen Gesetzen werden wir ganz kurz in diesem Kapitel berichten. Auf Beweise und genaue Formulierungen werden wir wegen ihrer Kompliziertheit verzichten.

Unter den uns in der Praxis begegnenden Zufallsgrößen tragen sehr viele den Charakter von "zufälligen Fehlern" oder lassen sich wenigstens leicht auf solche "Fehler" zurückführen.
Wir untersuchen beispielsweise die Flugweite x eines Geschosses. Beim Schießen mit einem gewissen Geschütz müssen wir natürlich annehmen, dass eine gewisse normale mittlere Flugweite x_0 existiert, auf die die Zieleinrichtung des Geschützes eingestellt wird. Die Differenz $x - x_0$ ist dann der "Fehler" der Schussweite, und die Untersuchung der Zufallsgröße x ist unmittelbar auf die Untersuchung des "zufälligen Fehlers" $x - x_0$ zurückgeführt.

Ein solcher Fehler aber, der seine Größe von Schuss zu Schuss ändert, hängt in der Regel von sehr vielen unabhängig voneinander wirkenden Ursachen ab: den zufälligen Schwankungen im Gewicht und in der Form des Geschosses, den zufälligen Änderungen der atmosphärischen Bedingungen, die eine Änderung der Luftbewegung hervorrufen, den zufälligen Fehlern beim Richten (wenn das Richten vor jedem Schuss wiederholt werden muss oder auch vor jeder Gruppe von Schüssen).
Alle diese und außerdem noch viele andere Ursachen können Unterschiede in der Flugweite hervorrufen. Alle diese Fehler werden also voneinander unabhängige Zufallsgrößen sein, und ihr Anteil am Gesamtfehler ist für jede einzelne Zufallsgröße nur gering. Der Fehler $x - x_0$, den wir untersuchen wollen, wird letztlich einfach die summarische Wirkung aller dieser zufälligen Fehler sein, die von den einzelnen Ursachen herrühren.

In unserem Beispiel ist der uns interessierende Fehler also eine Summe einer großen Anzahl von unabhängigen Zufallsgrößen. Damit ist aber auch klar, dass sich die zufälligen Fehler, mit denen wir in der Praxis zu tun haben, fast immer auf ähnliche Art behandeln lassen.

Eine theoretische Überlegung, die wir hier nicht darlegen können, ergibt, dass das Verteilungsgesetz einer Zufallsgröße, welche die Summe einer sehr großen Anzahl unabhängiger Zufallsgrößen ist, ein Gesetz eines gewissen völlig bestimmten Typs ist, wenn nur jeder der

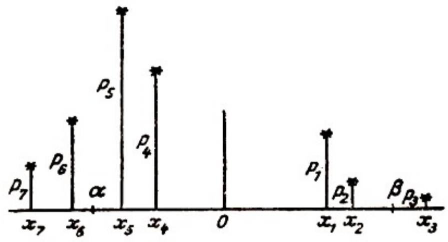
Summanden im Verhältnis zur ganzen Summe unbedeutend ist.[19]

Dieser Typus ist gerade der Typus der Normalgesetze.
Daher liegt die Vermutung nahe, dass ein ganz bedeutender Teil der in der Praxis vorkommenden Zufallsgrößen (alle Fehler, die sich aus einer großen Anzahl unabhängiger zufälliger Fehler zusammensetzen) angenähert nach einem Normalgesetz verteilt sind. Mit den Grundeigenschaften dieser Gesetze müssen wir uns jetzt vertraut machen.

12.30 Der Begriff der Verteilungskurve

Im Paragraph 15 hatten wir schon Verteilungsgesetze mit Hilfe eines Diagramms graphisch dargestellt. Dieses Verfahren ist insbesondere deshalb zweckmäßig, weil man mit einem Blick die wichtigsten Züge des zu untersuchenden Verteilungsgesetzes erfassen kann, ohne dass man auf eine Untersuchung der Tabellen angewiesen ist.

Abb. 11



Das Schema ist das folgende: Auf der horizontalen Geraden werden die verschiedenen möglichen Werte der gegebenen Zufallsgröße aufgetragen, indem man von einem gewissen Ursprung O ausgeht und die positiven Werte nach rechts und die negativen nach links aufträgt (siehe Abb. 11).
Über jedem dieser möglichen Werte trägt man die zugehörige Wahrscheinlichkeit auf. Den Maßstab wählt man in beiden Richtungen zweckmäßigerweise so, dass das ganze Diagramm eine bequem und leicht überschaubare Gestalt annimmt. Ein flüchtiger Blick auf Abb. 11 überzeugt uns davon, dass die dadurch charakterisierte Zufallsgröße den wahrscheinlichsten Wert x_5 (negativ) hat und dass mit zunehmender Entfernung von x_5 die Wahrscheinlichkeit der möglichen Werte dieser Größe (sogar ziemlich schnell) abnimmt.

Die Wahrscheinlichkeit dafür, dass die Größe einen in irgendeinem Intervall (α, β) eingeschlossenen Wert annimmt, ist nach dem Additionstheorem gleich der Summe der Wahrscheinlichkeiten aller möglichen Werte, die in diesem Intervall liegen.
Geometrisch betrachtet ist es gerade die Summe der Längen der senkrechten Striche, die über diesem Intervall aufgetragen sind.
In Abb. 11 ist

$$P\{\alpha < x < \beta\} = p_1 + p_2 + p_4 + p_5$$

Wenn, wie es tatsächlich in der Praxis oft vorkommt, die Anzahl der möglichen Werte sehr groß ist, nimmt man, um die Zeichnung in der Horizontalen nicht zu sehr auszudehnen, in horizontaler Richtung einen sehr kleinen Maßstab. Dadurch liegen die möglichen Werte außerordentlich dicht beieinander (Abb. 12).

[19]Siehe dazu auch die Schlussbemerkungen.

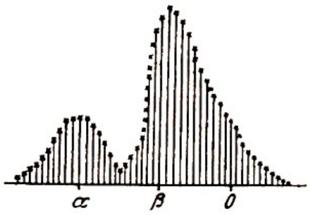


Abb. 12

Dann vereinigen sich die Spitzen der aufgetragenen senkrechten Striche für unser Auge zu einer zusammenhängenden Kurve, die man "Verteilungskurve"der vorgegebenen Zufallsgröße nennt.
Auch hier wird die Wahrscheinlichkeit der Ungleichung $\alpha < x < \beta$ graphisch als Summe der Längen der senkrechten Striche dargestellt, die über dem Intervall (α, β) aufgetragen sind.

Wir nehmen jetzt an, der Abstand zwischen zwei benachbarten Werten sei stets gleich Eins. Das ist z. B. der Fall, wenn sich die möglichen Werte durch eine Reihe aufeinanderfolgender ganzer Zahlen ausdrücken lassen. Das kann man praktisch stets erreichen, indem man eine hinreichend kleine Einheit wählt. Dann wird die Länge jedes senkrechten Striches zahlenmäßig gleich dem Flächeninhalt des Rechtecks, dessen Höhe dieser Strich ist. Die Grundlinie dieses Rechtecks ist gleich dem Abstand zweier benachbarter Striche, d. h. gleich Eins (Abb. 13).

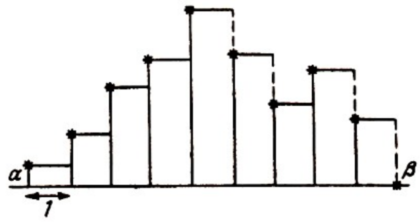


Abb. 13

Auf diese Weise wird die Wahrscheinlichkeit der Ungleichung $\alpha < x < \beta$ graphisch als Summe der Flächeninhalte der eingezeichneten Rechtecke, die über diesem Intervall liegen, dargestellt. Ordnet man die möglichen Werte sehr gedrängt an, etwa wie in Abb. 12, so wird die Summe der Flächeninhalte dieser Rechtecke sich praktisch nicht von dem Flächeninhalt der krummlinigen Figur unterscheiden, die von unten durch das Intervall (α, β), von oben durch die Verteilungskurve und an den Seiten durch die über den Punkten α und β (Abb. 14) aufgetragenen senkrechten Striche begrenzt wird.[20]

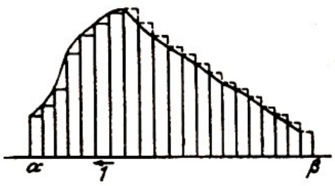


Abb. 14

Auf diese Weise kann in einem krummlinigen Diagramm der Form der Abb. 14 die Treffwahrscheinlichkeit der vorgegebenen Zufallsgröße in einem beliebigen Intervall einfach und bequem durch den Flächeninhalt ausgedrückt werden, der über diesem Intervall unterhalb der Verteilungskurve liegt. Besitzt das Verteilungsgesetz der gegebenen Zufallsgröße ein solches Kurvendiagramm, braucht man diese senkrechten Striche gar nicht zu ziehen.
Sie würden das Diagramm nur unübersichtlicher machen; ja selbst die Frage nach der Wahrscheinlichkeit eines einzelnen Wertes verliert hier ihre Aktualität.

[20] Wie früher nehmen wir hier als Einheit der Länge den Abstand zwischen zwei benachbarten möglichen Werten.

Gibt es sehr viele mögliche Werte (das trifft genau genommen für alle Kurvendiagramme zu), so wird die Wahrscheinlichkeit eines einzelnen Wertes winzig klein (praktisch gleich Null) und verliert damit jegliches Interesse. So ist es bei der Artillerie völlig uninteressant zu wissen, wie groß z.B. die Wahrscheinlichkeit dafür ist, dass der Aufschlagpunkt des Geschosses genau um 473 cm vom Zentrum abweicht. Von wesentlicher Bedeutung ist dagegen die Frage nach der Wahrscheinlichkeit dafür, dass die Abweichung des Geschosses in dem Intervall von 3 bis 5 m liegt.

Und so ist es auch in allen ähnlichen Fällen: Wenn eine Zufallsgröße sehr viele Werte annehmen kann, dann sind die Wahrscheinlichkeiten dieser einzelnen Werte völlig unwichtig, die Wahrscheinlichkeit eines ganzen Intervalle solcher Werte dagegen ist äußerst wichtig. Diese Wahrscheinlichkeiten ergeben sich aber anschaulich und unmittelbar, wie wir gerade sahen, durch Flächeninhalte im Kurvendiagramm.

12.31 Die Eigenschaften der normalen Verteilungskurven

Eine Größe, die einem Normalgesetz gehorcht, besitzt unendlich viele mögliche Werte. Daher kann man die Normalgesetze bequem graphisch als Kurvendiagramme darstellen.

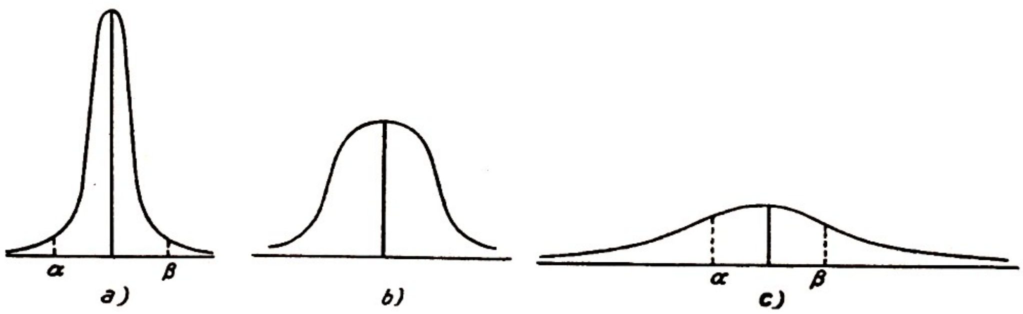


Abb. 15

In Abb. 15 sind einige Verteilungskurven von Normalgesetzen aufgezeichnet. Abgesehen von Unterschieden im einzelnen können wir einige deutliche Gemeinsamkeiten in diesen Zeichnungen entdecken:

1. Alle Kurven haben einen höchsten Punkt. Entfernt man sich von diesem Punkt nach rechts oder links, so nehmen sie monoton ab. Das bedeutet offenbar, dass mit zunehmender Entfernung vom wahrscheinlichsten Wert dieser Zufallsgröße die Wahrscheinlichkeiten ebenfalls monoton abnehmen.

2. Alle Kurven sind bezüglich der senkrechten Geraden, die durch den höchsten Punkt geht, symmetrisch. Offenbar bedeutet das, dass die Werte, die von dem wahrscheinlichsten Wert gleichweit entfernt sind, die gleichen Wahrscheinlichkeiten haben.

3. Alle Kurven haben eine glockenartige Gestalt. Unweit des höchsten Punktes sind sie nach oben gewölbt (konvex), in einem gewissen Abstand haben sie einen Wendepunkt und sind dann konvex nach unten. Dieser Abstand (wie auch die größte Höhe) sind für verschiedene Kurven verschieden.[21]

Wodurch unterscheiden sich die verschiedenen Normalkurven nun voneinander?

[21] Für diejenigen Leser, die mit den Elementen der höheren Mathematik vertraut sind, fügen wir hinzu, dass die Kurvengleichung, die das Normalgesetz analytisch darstellt, die folgende Form hat:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

Um eine klare Antwort auf diese Frage geben zu können, müssen wir uns daran erinnern, dass der unterhalb irgendeiner Verteilungskurve liegende Flächeninhalt gleich 1 ist, weil dieser Flächeninhalt gleich der Wahrscheinlichkeit dafür ist, dass die Zufallsgröße irgendeinen ihrer möglichen Werte annimmt, d. h. gleich der Wahrscheinlichkeit des sicheren Ereignisses ist.

Der Unterschied zwischen den einzelnen Verteilungskurven besteht daher nur darin, dass dieser summierte Flächeninhalt, der für alle Kurven ein und derselbe ist, auf verschiedene Weise auf die verschiedenen Bereiche verteilt ist. Für die Normalgesetze, wie sie die Kurven der Abb. 15 zeigen, können wir die Frage im Grunde jetzt anders stellen:
Wie groß ist der Anteil des summierten Flächeninhalts, der sich über einen unmittelbar an den wahrscheinlichsten Wert angrenzenden Bereich, beziehungsweise über einen Bereich erstreckt, der von dem wahrscheinlichsten Wert etwas weiter entfernt ist. Für das in Abb. 15a dargestellte Gesetz ist fast der gesamte Flächeninhalt in unmittelbarer Nähe des wahrscheinlichsten Wertes konzentriert.

Das heißt also, dass die Zufallsgröße mit erdrückender Wahrscheinlichkeit und somit in den meisten Fällen Werte annimmt, die sehr nahe beim wahrscheinlichsten Wert liegen. Da auf Grund der oben erwähnten Symmetrie der Normalgesetze der wahrscheinlichste Wert stets mit dem Mittelwert zusammenfällt, können wir sagen, dass eine dem Gesetz a) gehorchende Zufallsgröße wenig streut. Insbesondere ist damit ihre Streuung und ihre mittlere quadratische Abweichung sehr klein.

Dagegen ist der in Abb. 15 c) in unmittelbarer Nähe des wahrscheinlichsten Wertes konzentrierte Flächeninhalt nur ein kleiner Teil der Summenfläche. Den Unterschied zur Abb. 15 a) werden wir sogleich sehen, wenn wir nämlich die über Intervallen gleicher Länge befindlichen Flächeninhalte miteinander vergleichen.
Hier in der Abb. 15 c) ist es ganz wahrscheinlich, dass die Zufallsgröße einen Wert annimmt, der von dem wahrscheinlichen Wert merklich abweichen wird. Die Größe streut stark, ihre Streuung und die mittlere quadratische Abweichung sind groß.
Der Fall b) liegt offensichtlich zwischen den Fallen a) und c).

Um die Gesamtheit der Normalgesetze möglichst schnell kennenzulernen und anwenden zu können, werden wir zweckmäßigerweise von zwei Grundeigenschaften dieser Gesetze ausgehen. Diese Eigenschaften, die wir sofort ausführlich formulieren werden, können wir hier nicht beweisen. Zum Beweise wäre eine genaue Definition der Normalgesetze erforderlich, die aber wiederum vom Leser einige Kenntnisse der höheren Mathematik verlangen würde.

Eigenschaft 1. Wenn die Zufallsgröße x nach einem Normalgesetz verteilt ist, dann gilt

1. für beliebige Konstanten $c > 0$ und d ist die Zufallsgröße $cx + d$ ebenfalls nach einem Normalgesetz verteilt, und
2. es gibt für jedes Normalgesetz ein (eindeutig bestimmtes) Zahlenpaar $c > 0$ und d derart, dass die Größe $cx + d$ nach diesem Gesetz verteilt ist.

Wenn also eine Zufallsgröße x nach einem Normalgesetz verteilt ist, dann sind auch alle Verteilungsgesetze der Form $cx + d$ für alle möglichen Werte der Konstanten $c > 0$ und d wieder Normalgesetze.

wobei die Zahl $e = 2,71828...$ die Basis der natürlichen Logarithmen und $\pi = 3,14159...$ das Verhältnis der Länge der Kreisperipherie zum Durchmesser ist. Die Größen a und σ^2 sind der Mittelwert und die Streuung der Zufallsgröße. Die Kenntnis der analytischen Form des Normalgesetzes kann dem Leser das Verständnis des weiteren Stoffes bedeutend erleichtern. Die folgenden Darlegungen sind jedoch auch solchen Lesern zugänglich, die keine Kenntnis der höheren Mathematik besitzen.

Eigenschaft 2. Wenn die Zufallsgrößen x und y unabhängig und nach Normalgesetzen verteilt sind, dann ist auch ihre Summe $z = x + y$ nach einem gewissen Normalgesetz verteilt.

Aus diesen zwei hier ohne Beweis als gültig angesehenen Grundeigenschaften können wir schon eine ganze Reihe anderer Eigenschaften der Normalgesetze streng herleiten. Die abgeleiteten Eigenschaften sind für die Praxis besonders wichtig.

I. Zu je zwei Zahlen a und $q > 0$ gibt es ein eindeutig bestimmtes Normalgesetz mit dem Mittelwert a und der mittleren quadratischen Abweichung q.

In der Tat: Es sei x eine Zufallsgröße, die nach einem Normalgesetz mit dem Mittelwert $\overline{x}$ und der mittleren quadratischen Abweichung Q_x verteilt ist. Unsere Behauptung wird bewiesen sein, wenn es uns auf Grund der Eigenschaft 1 gelingt, ein eindeutig bestimmtes Zahlenpaar $c > 0$ und d anzugeben, derart, dass die Größe $cx + d$ den Mittelwert a und die mittlere quadratische Abweichung q hat. Hat die Tabelle der Größe x die Gestalt

x_1	x_2	...	x_n
p_1	p_2	...	p_n

(I)

dann wird der Größe $cx + d$ (wobei $c > 0$ und d bis jetzt noch beliebige Konstanten sind) die Tabelle

$cx_1 + d$	$cx_2 + d$	...	$cx_n + d$
p_1	p_2	...	p_n

entsprechen. Offensichtlich ist[22]

$$\sum_k x_k p_k = \overline{x} \qquad , \qquad \sum_k (x_k - \overline{x})^2 p_k = Q_x^2$$

Unsere Forderung führt so auf zwei Bedingungen:

$$\sum_k (cx_k + d)p_k = a \qquad , \qquad \sum_k (cx_k + d - a)^2 p_k = q^2$$

Die erste dieser Bedingungen ergibt:

$$c\sum_k x_k p_k + d\sum_k p_k = a \qquad \text{oder} \qquad c\overline{x} + d = a \tag{1}$$

und die zweite

$$\sum_k (cx_k + d - c\overline{x} - d)^2 p_k = c^2 \sum_k (x_k - \overline{x})^2 p_k = c^2 Q_x^2 = q^2$$

woraus (wegen $c > 0$) folgt:

$$c = \frac{q}{Q_x} \tag{2}$$

und somit aus (1)

$$d = a - c\overline{x} = a - \frac{q\overline{x}}{Q_x} \tag{3}$$

Somit können für gegebene a und q mit Hilfe der Formeln (2) und (3) stets solche Zahlen c und d bestimmt werden und dabei nur auf eine einzige Weise. Die Größe $cx + d$ unterliegt also dem Normalgesetz mit dem Mittelwert a und der mittleren quadratischen Abweichung q.

[22] Das Zeichen $\sum\limits_k$ ist eine Abkürzung. Es bedeutet $\sum\limits_{k=1}^{n}$.

Damit ist die Behauptung vollständig bewiesen.

Beschränkt man sich nicht nur auf Normalgesetze, sondern lässt alle möglichen Verteilungsgesetze zu, so gibt uns die Vorgabe des Mittelwertes und der Streuung oder der mittleren quadratischen Abweichung einer Zufallsgröße noch sehr wenig Auskunft über ihr Verteilungsgesetz.
Es gibt nämlich sehr viele (und dabei wesentlich voneinander verschiedene) Verteilungsgesetze, die ein und denselben Mittelwert und ein und dieselbe Streuung haben.

Im allgemeinen Falle charakterisiert uns die Vorgabe der Streuung und des Mittelwertes nur sehr angenähert das Verteilungsgesetz einer vorgegebenen Zufallsgröße.

Die Sache wird jedoch anders, wenn wir uns verabredungsgemäß nur auf die Betrachtung von Normalgesetzen beschränken. Wie wir uns gerade überzeugten, ist einerseits jede Voraussetzung über Mittelwert und Streuung einer gegebenen Zufallsgröße mit der Forderung vereinbar, dass sie nach einem Normalgesetz verteilt ist.
Bestehen Gründe zu der Vermutung, dass die gegebene Größe nach einem Normalgesetz verteilt ist, so bestimmt andererseits - und das ist die Hauptsache - die Vorgabe ihres Mittelwertes und der Streuung eindeutig ihr Verteilungsgesetz. Damit ist aber dann schon alles für diese Zufallsgröße vollkommen bekannt. Wenn wir insbesondere den Mittelwert und die Streuung einer solchen Größe kennen, können wir die Wahrscheinlichkeit dafür berechnen, dass sie einen Wert annimmt, der in diesem oder jenem willkürlich gewählten Intervall liegt.

II. Der Quotient aus der gemittelten (wahrscheinlichen) Abweichung und der mittleren quadratischen Abweichung ist für alle Normalgesetze ein und derselbe.

Es seien zwei beliebige Normalgesetze vorgegeben, und es sei x eine Zufallsgröße, die dem ersten dieser Gesetze unterliegt. Auf Grund der Eigenschaft 1 existieren konstante Zahlen $c > 0$ und d derart, dass die Größe $cx + d$ nach dem zweiten der gegebenen Gesetze verteilt ist.
Es seien Q_x und E_x die mittlere quadratische Abweichung und die gemittelte (wahrscheinliche) Abweichung der ersten Größe und q und e die entsprechenden Abweichungen der zweiten Größe. Nach Definition der wahrscheinlichen Abweichung gilt

$$P\{|(cx+d) - (c\overline{x}+d)| < e\} = \frac{1}{2} \qquad \text{oder} \qquad P\{c|x-\overline{x}| < e\} = \frac{1}{2}$$

oder schließlich

$$P\left\{|x-\overline{x}| < \frac{e}{c}\right\} = \frac{1}{2}$$

Hieraus folgt wiederum auf Grund der Definition der wahrscheinlichen Abweichung, dass $\frac{e}{c}$ die wahrscheinliche Abweichung der Größe x, d. h.

$$\frac{e}{c} = E_x \qquad \text{ist, woraus} \qquad \frac{e}{E_x} = c$$

folgt. Daher ergibt die Beziehung (2)

$$\frac{e}{E_x} = \frac{q}{Q_x} \qquad \text{woraus} \qquad \frac{e}{q} = \frac{E_x}{Q_x}$$

folgt, d. h., der Quotient aus der wahrscheinlichen und der mittleren quadratischen Abweichung ist für unsere zwei Gesetze der gleiche.
Da diese Gesetze auf Grund unserer Voraussetzung zwei beliebige Normalgesetze waren, ist unsere Behauptung bewiesen.

Der Quotient $\frac{e}{q}$ ist also eine absolute Konstante, die wir mit dem Buchstaben λ bezeichnen. Eine Berechnung dieses Wertes ergibt:

$$\lambda = \sqrt{\frac{2}{\pi}} \approx 0,674$$

Somit ist für jedes Normalgesetz

$$e = \sqrt{\frac{2}{\pi}} q$$

Wegen dieses einfachen Zusammenhanges zwischen den Zahlen e und q ist es für Zufallsgrößen, die nach einem Normalgesetz verteilt werden, praktisch gleichgültig, welche der beiden Zahlen wir zur Charakterisierung der Streuung benutzen werden.
Wir sahen oben, dass die mittleren quadratischen Abweichungen im allgemeinen (d. h. wenn wir uns nicht nur auf normal - d. h. nach einem Normalgesetz - verteilte Zufallsgrößen beschränken) eine ganze Reihe einfacher Eigenschaften besitzen, die andere Charakterisierungen nicht aufweisen können.

Diese Eigenschaften sind es wohl auch, die sowohl den Theoretiker als auch den Praktiker veranlassen, als Maß der Streuung die mittlere quadratische Abweichung zu wählen. Wir wiesen aber schon darauf hin, dass die Artilleristen nichtsdestoweniger fast immer die gemittelten Abweichungen verwenden. Jetzt ist uns auch klar, weshalb diese Tradition keinerlei Schaden mit sich bringen kann. Die Zufallsgrößen, mit denen die artilleristische Wissenschaft und die Praktiker zu tun haben, sind fast immer nach einem Normalgesetz verteilt, und für solche Zufallsgrößen ist es wegen der oben erwähnten Proportionalität völlig gleichgültig, welche der beiden Charakterisierungen praktisch verwendet wird.

III. Es seien x und y zwei normal verteilte Zufallsgrößen und $z = x + y$. Dann ist

$$E_z = \sqrt{E_x^2 + E_y^2}$$

wobei E_x, E_y bzw. E_z die wahrscheinlichen Abweichungen der Größen x, y und z sind.

Eine analoge Formel für die mittleren quadratischen Abweichungen gilt, wie wir aus Paragraph 25 wissen, auch für beliebige Verteilungsgesetze der Größen x und y. Sind dies schon Normalgesetze, so ist die Größe z auf Grund der Eigenschaft 2 ebenfalls nach einem Normalgesetz verteilt. Dann gilt aber wegen der gerade bewiesenen Eigenschaft II:

$$E_x = \lambda Q_x, \qquad E_y = \lambda Q_y, \qquad E_z = \lambda Q_z$$

und somit

$$E_z = \lambda\sqrt{Q_x^2 + Q_y^2} = \sqrt{(\lambda Q_x)^2 + (\lambda Q_y)^2} = \sqrt{E_x^2 + E_y^2}$$

Somit lässt sich eine der wichtigsten Eigenschaften der mittleren quadratischen Abweichungen im Falle der Normalgesetze unmittelbar auch auf die wahrscheinlichen (gemittelten) Abweichungen übertragen.

12.32 Lösung von Aufgaben

Als Hauptnormalgesetz wollen wir das Gesetz bezeichnen, welches den Mittelwert Null und die Streuung Eins hat. Unterliegt eine Zufallsgröße x dem Hauptnormalgesetz, so wollen wir der Kürze halber

$$P\{|x| < a\} = \Phi(a)$$

für beliebiges positives a schreiben. Also ist $\Phi(a)$ die Wahrscheinlichkeit dafür, dass die dem Hauptnormalgesetz unterliegende Größe zu dem absoluten Betrage nach die Zahl a nicht übersteigt.

Für die Größe $\Phi(a)$ gibt es sehr genaue Tabellen, die ihren Wert für verschiedene Werte der Zahl a enthalten. Eine solche Tabelle ist ein unentbehrliches Werkzeug für jeden, der mit der Berechnung von Wahrscheinlichkeiten zu tun hat.
Auszüge aus solchen Tabellen finden sich in jedem Buch, welches sich mit Wahrscheinlichkeitsrechnung befasst. Am Ende unseres Buches findet der Leser ebenfalls eine solche Tabelle.

An Hand der Tabelle der Funktionswerte von $\Phi(a)$ kann man leicht und mit großer Genauigkeit alle Berechnungen für beliebige normal verteilte Zufallsgrößen ausführen. Wir zeigen jetzt an Beispielen, wie dies gemacht wird.

Aufgabe I. Die Zufallsgröße x ist nach dem Normalgesetz mit dem Mittelwert $\overline{x}$ und der mittleren quadratischen Abweichung Q_x verteilt. Man bestimme die Wahrscheinlichkeit dafür, dass die Abweichung $x - \overline{x}$ dem absoluten Betrage nach die Zahl a nicht übersteigt.

Es sei z eine Zufallsgröße, die gemäß dem Hauptnormalgesetz verteilt ist. Wegen der Eigenschaft 1 kann man Zahlen $c > 0$ und d finden derart, dass die Größe $cz + d$ den Mittelwert $\overline{x}$ und die mittlere quadratische Abweichung Q_x hat, d. h. demselben Normalgesetz unterliegt wie die gegebene Größe x. Daher ist

$$P\{Y|x - \overline{x}| < a\} = P\{|(cz + d) - (c\overline{z} + d)| < a\} == P\{c|z - \overline{z}| < a\}$$

Nach Formel (2) ist hier $c = \frac{Q_x}{Q_z} = Q_x$, weil $Q_z = 1$ ist (für das Hauptnormalgesetz ist die Streuung gleich 1). Daher ergibt sich

$$P\{|x - \overline{x}| < a\} = P\{Q_x|z - \overline{z}| < a\} = P\left\{|z| < \frac{a}{Q_x}\right\} = \Phi\left(\frac{a}{Q_x}\right) \tag{4}$$

Damit ist die gestellte Aufgabe gelöst, weil wir die Größe $\Phi\left(\frac{a}{Q_x}\right)$ unmittelbar der Tabelle entnehmen können. Also ermöglicht unsere Tabelle, mit Hilfe der Formel (4) leicht die Wahrscheinlichkeit einer beliebigen Abweichung vom Mittelwert für eine nach einem beliebigen Normalgesetz verteilte Zufallsgröße zu berechnen.

Beispiel 1. Auf einer Maschine wird ein gewisses Einzelteil hergestellt. Es zeigt sich, dass die Länge x eine Zufallsgröße ist, die nach einem Normalgesetz verteilt ist und den Mittelwert 20 cm und die Streuung 0,2 cm hat.
Es ist die Wahrscheinlichkeit dafür zu finden, dass die Länge des Einzelteils zwischen 19,7 cm und 20,3 cm liegt, d. h. dass die Abweichung nach irgendeiner Seite 0,3 cm nicht übersteigt.

Nach Formel (4) und nach unserer Tabelle wird

$$P\{|x - 20| < 0,3\} = \Phi\left(\frac{0,3}{0,2}\right) = \Phi(1,5) = 0,866$$

Somit wird bei 87 % aller Erzeugnisse, die unter den gegebenen Bedingungen angefertigt werden, die Länge zwischen 19,7 cm und 20,3 cm liegen. Die restlichen 13 % haben eine größere Abweichung vom Mittelwert.

Beispiel 2. Man bestimme unter den Voraussetzungen des Beispiele 1, mit welcher Genauigkeit die Länge eines Erzeugnisses mit der Wahrscheinlichkeit 0,95 garantiert werden kann.

Die Aufgabe besteht offenbar im Aufsuchen einer solchen positiven Zahl a, für die

$$P\{|x-20| < a\} > 0,95$$

ist.
Die Rechnung des Beispiels 1 zeigt, dass $a = 0,3$ dafür zu klein ist, weil in diesem Falle die linke Seite der gerade hingeschriebenen Ungleichung kleiner als 0,87 wird. Da gemäß Gleichung (4)

$$P\{|x-20| < a\} = \Phi\left(\frac{a}{0,2}\right) = \Phi(5a)$$

ist, müssen wir in unserer Tabelle einen solchen Wert $5a$ suchen, für den $\Phi(5a) > 0,95$ ist. Diese Bedingung wird für $5a > 1,97$ erfüllt, und hieraus folgt $a > 0,394$.
Man kann also mit einer nicht unter 0,95 liegenden Wahrscheinlichkeit garantieren, dass die Längenabweichung 0,4 cm nicht übersteigt.

Beispiel 3. Bei einigen praktischen Problemen rechnet man damit, dass eine nach einem Normalgesetz verteilte Zufallsgröße x: nicht um mehr abweicht; als das Dreifache der mittleren quadratischen Abweichung Q_x ausmacht. Welche Gründe hat man für diese Behauptung?

Die Formel (4) und unsere Tabelle ergeben, dass

$$P\{|x-\overline{x}| < 3Q_x\} = \Phi(3) > 0,997 \qquad \text{und folglich} \qquad P\{|x-\overline{x}| > 3Q_x\} < 0,003$$

ist.
Praktisch bedeutet das, dass Abweichungen, die den absoluten Betrag von $3Q_x$ übersteigen, im Mittel in weniger als 3 von 1000 Fällen auftreten werden.
Kann man eine solch geringe Wahrscheinlichkeit vernachlässigen oder muss man sie berücksichtigen? Das hängt letzten Endes vom Inhalt der Aufgabe ab und kann nicht ein für allemal vorgeschrieben werden.

Wir wollen noch erwähnen, dass die Beziehung $P\{|x-\overline{x} < 3Q_x\} == \Phi(3)$ offenbar ein Spezialfall der Formel

$$P\{|x-\overline{x}| < aQ_x\} = \Phi(a) \tag{5}$$

ist, die sich aus Formel (4) ergibt und für jede nach einem Normalgesetz verteilte Zufallsgröße gilt.

Beispiel 4. Bei einem mittleren Gewicht eines Geschosses von 8,4 kg wurde festgestellt, dass die Abweichungen im Mittel bei drei von einhundert Geschossen den absoluten Betrag von 50 g übersteigen. Angenommen, das Geschossgewicht ist nach einem Normalgesetz verteilt. Wie groß ist die wahrscheinliche Abweichung?

Wir wissen, dass

$$P\{|x-8,4| > 0,05\} = 0,03$$

ist, wobei x das Gewicht eines willkürlich ausgewählten Geschosses ist. Hieraus folgt

$$0,97 = P\{|x-8,4| < 0,05\} = \Phi\left(\frac{0,05}{Q_x}\right)$$

Die Beziehung $\Phi(a) = 0,97$ ist nach unserer Tabelle für $a \approx 2,12$ erfüllt. Daher wird

$$\frac{0,05}{Q_x} \approx 2,12 \qquad \text{also} \qquad Q_x \approx \frac{0,05}{2,12}$$

Die wahrscheinlichste Abweichung ist, wie wir wissen, gleich $E_x = 0,674 Q_x \approx 0,0155$ kg $= 15,5$ g.

Beispiel 5. Beim Schießen mit einem Geschütz wird die Abweichung eines Geschosses vom Ziel durch drei voneinander unabhängige Ursachen hervorgerufen:
1. Fehler bei der Zielbestimmung,
2. Richtfehler und
3. Fehler von Ursachen, die sich von Schuss zu Schuss ändern (Geschossgewicht, atmosphärische Bedingungen u. a. m.).

Unter der Voraussetzung, dass alle drei Fehlermöglichkeiten nach einem Normalgesetz mit dem Mittelwert 0 verteilt sind, und dass die wahrscheinlichen Abweichungen 24 m, 8 m bzw. 12 m sind, bestimme man die Wahrscheinlichkeit dafür, dass die Gesamtabweichung vom Ziel nicht größer als 40 m ist.

Da die wahrscheinliche Abweichung des Gesamtfehlers x auf Grund der Eigenschaft III gleich

$$\sqrt{24^2 + 8^2 + 12^2} = 28 \text{ m}$$

ist, wird die mittlere quadratische Abweichung des Gesamtfehlers gleich $\frac{28}{0{,}674} \approx 41,5$ und somit

$$P\{|x| < 40\} = \Phi\left(\frac{40}{41,5}\right) \approx \Phi(0,964) = 0,665$$

Die Abweichungen, die nicht größer als 40 m sind, werden also etwa 2/3 aller Fälle ausmachen.

Aufgabe II. Die Zufallsgröße x sei nach dem Normalgesetz mit dem Mittelwert $\overline{x}$ und der mittleren quadratischen Abweichung Q_x verteilt. Es ist die Wahrscheinlichkeit dafür zu bestimmen, dass die Abweichung $x - \overline{x}$ dem absoluten Betrage nach zwischen den Zahlen a und b liegt $(0 < a < b)$.

Da nach dem Additionstheorem

$$P\{|x - \overline{x}| < b\} = P\{|x - \overline{x}| < a\} + P\{a < |x - \overline{x}| < b\}$$

ist, wird

$$P\{a < |x - \overline{x}| < b\} = P\{|x - \overline{x}| < b\} - P\{|x - \overline{x}| < a\} = \Phi\left(\frac{b}{Q_x}\right) - \Phi\left(\frac{a}{Q_x}\right) \qquad (6)$$

womit die gestellte Aufgabe auch schon gelöst ist.

Für die meisten praktischen Fragen sind solche Wertetabellen der Funktion $Phi(a)$, wie wir sie die ganze Zeit benutzten, jedoch unnötig schwere Rechenhilfsmittel. Gewöhnlich hat man nur die Treffwahrscheinlichkeit der Abweichung $x - \overline{x}$ in einem mehr oder weniger großen Intervall zu berechnen.
Daher wird man praktischerweise neben unserer "vollständigen" Tabelle auch eine gekürzte Tabelle verwenden. Eine solche Tabelle kann man sich leicht mit Hilfe der Formel (6) aus der vollständigen Tabelle selbst herstellen.

Wir bringen ein solches Konstruktionsbeispiel für eine Tabelle, die weit gröber als die Tabelle am Ende des Buches ist und trotzdem in vielen Fällen vollkommen ausreicht. Wir Zerlegen das ganze Änderungsintervall der Größe $|x - \overline{x}|$ in fünf Teile:

1. von Null bis $0,32Q_x$; 2. von $0,32Q_x$ bis $0,69Q_x$; 3. von $0,69Q_x$ bis $1,15Q_x$; 4. von $1,15Q_x$,

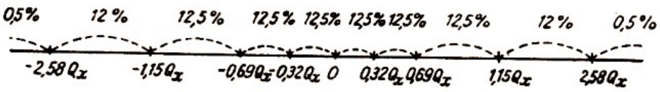
bis $2,58Q_x$, und 5. größer als $2,58Q_x$.

Unter Benutzung der Formel (4) finden wir:

$$\begin{aligned} P\{|x-\overline{x}| < 0,32Q_x\} &= \Phi(0,32) \approx 0,25;\\ P\{0,32Q_x < |x-\overline{x}] < 0,69Q_x\} &= \Phi(0,69) - \Phi(0,32) \approx 0,25;\\ P\{0,69Q_x < |x-\overline{x}] < 1,15Q_x\} &= \Phi(1,15) - \Phi(0,69) \approx 0,25;\\ P\{1,15Q_x < |x-\overline{x}] < 2,38Q_x\} &= \Phi(2,38) - \Phi(1,15) \approx 0,24;\\ P\{|x-\overline{x}] > 2,58Q_x\} &= 1 - \Phi(2,58) \approx 0,01 \end{aligned}$$

Das Ergebnis dieser Rechnung kann man etwa durch eine graphische Darstellung veranschaulichen (Abb. 16).

Abb. 16



Hier ist die ganze unendliche Gerade in 10 Teile, fünf positive und fünf negative, zerlegt. Über jedem Intervall steht, wieviel Prozent der tatsächlich beobachteten Abweichung im Mittel auf dieses Intervall kommen.
So müssen beispielsweise gemäß der oben durchgeführten Rechnung auf die Intervalle $(-1,15Q_x, -0,69Q_x)$ und $(0,69Q_x, 1,15Q_x)$ zusammengenommen etwa 25 % aller Abweichungen fallen.

Wegen der Symmetrie der Normalgesetze werden auf beide Teilintervalle angenähert gleichviel Geschosse fallen, so dass auf jedem von ihnen etwa 12,5 % der Gesamtabweichungen liegen.
Durch Benutzung dieses oder eines ähnlichen einfachen Schemas können wir die Abweichungen einer Zufallsgröße, die einem Normalgesetz mit beliebigem Mittelwert und beliebiger Streuung gehorcht, unmittelbar in ihren Grundzügen darstellen.

Wir wollen zum Schluss untersuchen, wie man die Treffwahrscheinlichkeit einer Zufallsgröße, die einem Normalgesetz unterliegt, in einem beliebig vorgegebenen Intervall berechnet.

Aufgabe III. Wenn wir wissen, dass die Zufallsgröße x nach einem Normalgesetz verteilt ist (Mittelwert $\overline{x}$, mittlere quadratische Abweichung Q_x), soll mit Hilfe der Tabelle die Wahrscheinlichkeit der Ungleichung $a < x < b$ berechnet werden, wobei a und b $(a < b)$ beliebige vorgegebene Zahlen sind.

Wir müssen drei Fälle in Abhängigkeit von der Anordnung der Zahlen a und b bezüglich $\overline{x}$ betrachten.

Erster Fall: $\overline{x} \leq a \leq b$.
Nach dem Additionstheorem ist

$$P\{\overline{x} < x < b\} = P\{\overline{x} < x < a\} + P\{a < x < b\}$$

woraus

$$\begin{aligned} P\{a < x < b\} &= P\{\overline{x} < x < b\} - P\{\overline{x} < x < a\}\\ &= P\{0 < x-\overline{x} < b-\overline{x}\} - P\{0 < x-\overline{x} < a-\overline{x}\} \end{aligned}$$

folgt. Für beliebiges $\alpha > 0$ ist wegen der Symmetrie der Normalgesetze

$$P\{0 < x - \overline{x} < \alpha\} = P\{-\alpha < x - \overline{x} < 0\} = \frac{1}{2}P\{-\alpha < x - \overline{x} < \alpha\}$$
$$= \frac{1}{2}P\{|x - \overline{x}| < \alpha\} = \frac{1}{2}\Phi\left(\frac{\alpha}{Q_x}\right) \quad (7)$$

Daher wird

$$P\{a < x < b\} = \frac{1}{2}\left[\Phi\left(\frac{b - \overline{x}}{Q_x}\right) - \Phi\left(\frac{a - \overline{x}}{Q_x}\right)\right]$$

Zweiter Fall: $a \leq \overline{x} \leq b$.

$$\begin{aligned} P\{a < x < b\} &= P\{a < x < \overline{x}\} + P\{\overline{x} < x < b\} \\ &= P\{a - \overline{x} < x - \overline{x} < 0\} + P\{0 < x - \overline{x} < b - \overline{x}\} \\ &= \frac{1}{2}\left[\Phi\left(\frac{\overline{x} - a}{Q_x}\right) - \Phi\left(\frac{b - \overline{x}}{Q_x}\right)\right] \end{aligned}$$

nach (7).

Dritter Fall: $a \leq b \leq \overline{x}$.

$$P\{a < x < \overline{x}\} = P\{a < x < b\} + P\{b < x < \overline{x}\}$$

woraus

$$\begin{aligned} P\{a < \overline{x} < b\} &= P\{a < x < \overline{x}\} + P\{b < x < \overline{x}\} \\ &= P\{a - \overline{x} < x - \overline{x} < 0\} - P\{b - \overline{x} < x - \overline{x} < 0\} \\ &= \frac{1}{2}\left[\Phi\left(\frac{\overline{x} - a}{Q_x}\right) - \Phi\left(\frac{\overline{x} - b}{Q_x}\right)\right] \end{aligned}$$

folgt.

Die Aufgabe ist in allen drei Fällen gelöst. Wir sehen also, dass für eine nach einem beliebigen Normalgesetz verteilte Zufallsgröße unsere Tabelle die Möglichkeit gibt, die Treffwahrscheinlichkeit dieser Größe für ein beliebiges Intervall zu finden. Dadurch wird auf ausführliche Weise ihr Verteilungsgesetz charakterisiert.

Um diese Ausführungen anschaulicher zu machen, führen wir die Rechnungen an dem folgenden Beispiel vor.

Beispiel. Vom Punkte O werde längs der Geraden OX ein Schießen durchgeführt. Die mittlere Flugweite eines Geschosses betrage 1200 m. Es wird vorausgesetzt, dass die Flugweite H nach einem Normalgesetz mit der mittleren quadratischen Abweichung 40 m verteilt ist.
Es ist zu bestimmen, wieviel Prozent der abgefeuerten Geschosse um 60 bis 80 m über das Ziel hinausfliegen.

Damit ein Schuss ein solcher Weitschuss ist, muss $1260 < H < 1280$ sein. Wir wenden die Schlussformel des ersten Falles der Aufgabe III an und erhalten:

$$P\{1260 < H < 1280\} = \frac{1}{2}\left[\Phi\left(\frac{1280 - 1200}{40}\right) - \Phi\left(\frac{1260 - 1200}{40}\right)\right] = \frac{1}{2}[\Phi(2) - \Phi(1,5)]$$

Aus der Tabelle entnimmt man: $\Phi(2) \approx 0,955$, $\Phi(l,5) \approx 0,866$, woraus

$$P\{1260 < H < 1280\} \approx 0,044$$

folgt. Also werden nicht mehr als 4 % der abgefeuerten Geschosse solche Weitschüsse sein.

13 Schlussbemerkungen

Die Wahrscheinlichkeitsrechnung entwickelte sich in den letzten Jahren zu einer der am schnellsten aufblühenden mathematischen Disziplinen. Neue theoretische Ergebnisse eröffneten große Möglichkeiten für die Anwendung der Methoden der Wahrscheinlichkeitsrechnung in den Naturwissenschaften. Ein intensives und ausführliches Studium der Naturerscheinungen förderte gleichzeitig die Wahrscheinlichkeitsrechnung bei der Entwicklung neuer Methoden und beim Auffinden neuer Gesetzmäßigkeiten.

Die Wahrscheinlichkeitsrechnung ist eine derjenigen mathematischen Disziplinen, die sich nicht vom Leben und von den Problemen anderer Wissenschaften abkapseln und mit der allgemeinen Entwicklung der Naturwissenschaften und der Technik Schritt halten.
Der Leser darf das Gesagte jedoch nicht falsch verstehen und denken, dass die Wahrscheinlichkeitsrechnung jetzt nur noch ein Anhängsel, ein Hilfsmittel zur Lösung praktischer Aufgaben sei.

Das ist durchaus nicht der Fall: Gerade in den letzten drei Jahrzehnten wurde die Wahrscheinlichkeitsrechnung mit ihren Problemen und Untersuchungsmethoden zu einer streng mathematischen Disziplin. Gleichzeitig zeigte sich, dass die wichtigsten und nächstliegenden Probleme der Wahrscheinlichkeitsrechnung als mathematische Disziplin der Lösung aktueller Probleme der Naturwissenschaft dienen.

Die mächtige Entwicklung der Wahrscheinlichkeitsrechnung ist eng mit den Traditionen und Erfolgen der russischen Wissenschaft verbunden. Während die Wahrscheinlichkeitsrechnung im vergangenen Jahrhundert in Westeuropa in eine Sackgasse geriet, fand in Russland der geniale Mathematiker P. L. Tschebyscheff einen neuen Weg zu ihrer Weiterentwicklung. Es handelt sich hier um die Untersuchung von Folgen unabhängiger Zufallsgrößen. Tschebyscheff selbst und auch seine Schüler A. A. Markoff und A. M. Ljapunow fanden auf diesem Wege einige fundamentale Ergebnisse (das Gesetz der großen Zahlen, den Satz von Ljapunow usw.).

Dem Leser ist das Gesetz der großen Zahlen schon bekannt.
Unsere nächste Aufgabe wird es nun sein, eine Vorstellung von der anderen wichtigen Aussage der Wahrscheinlichkeitsrechnung, dem Satz von Ljapunow (oder, wie er noch genannt wird, dem "Zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung") zu vermitteln.

Sein Wert besteht vor allen Dingen darin, dass eine große Anzahl Erscheinungen, deren Ausgang vom Zufall abhängig ist, sich in ihren Grundzügen nach folgendem Schema ergeben:
Die untersuchte Erscheinung steht unter dem Einfluss einer Vielzahl unabhängiger zufällig wirkender Ursachen. Jede dieser Ursachen hat auf die Gesamterscheinung nur einen verschwindend kleinen Einfluss. Ihre Auswirkungen drücken sich durch die Zufallsgrößen $\xi_1, \xi_2, ..., \xi_n$ aus, und die gesamte Wirkung auf die Erscheinung ist gleich der Summe $s_n = \xi_1 + \xi_2 + ... + \xi_n$. Da man die Auswirkung jeder einzelnen Ursache praktisch nicht ausrechnen kann (oder anders gesagt, die Verteilungsfunktionen der Größen ξ nicht angeben kann), ja sie nicht einmal einfach aufzählen kann, ist es verständlich, wie wichtig gerade die Ausarbeitung von Methoden ist, die die Untersuchung der summarischen Wirkung unabhängig von der Natur jedes einzelnen Summanden gestatten.

Die gewöhnlichen Untersuchungsmethoden sind zur Lösung dieser Problemstellung unbrauchbar. An ihre Stelle müssen andere Methoden treten, für welche die große Anzahl der Ursachen einer Erscheinung nicht nur kein Hindernis mehr bildet, die Lösung des Problems zu finden, sondern sie vielmehr erleichtert.

Diese Methoden müssen die unzureichende Kenntnis jeder einzelnen Ursache durch ihre große Anzahl kompensieren. Der zentrale Grenzwertsatz, der unter großen Schwierigkeiten in der Hauptsache von den Akademiemitgliedern P. L. Tschebyscheff (1821-1894), A. A. Markoff (1856-1922) und A. M. Ljapunow (1857-1918) aufgestellt wurde, sagt aus:

Wenn die wirkenden Ursachen $\xi_1, \xi_2, ..., \xi_n$ unabhängig sind, ihre Anzahlen sehr groß und die Auswirkung jeder einzelnen Ursache im Vergleich zur Summe ihrer Wirkungen nur unbedeutend ist, dann kann sich das Verteilungsgesetz der Summe s_n nur unbedeutend von einem Normalgesetz unterscheiden.

Wir bringen noch einige Beispiele, die sich in das gerade beschriebene Schema einordnen lassen.

Beim Schießen mit einem Geschütz sind Abweichungen zwischen dem Aufschlagpunkt des Geschosses und dem Visierpunkt unvermeidlich. Da die Streuung das Resultat einer riesigen Anzahl unabhängig wirkender Faktoren ist (z. B. Unregelmäßigkeiten beim Drehen des Geschosses, Unebenheiten der Geschossspitze, Schwankungen in der Dichte des Materials, aus dem der, Geschosskopf angefertigt wird, geringfügige Schwankungen in der Menge des Sprengstoffes bei verschiedenen Geschossen, kleine, für das Auge unmerkliche Fehler beim Richten des Geschützes, geringfügige Schwankungen des atmosphärischen Zustandes bei verschiedenen Schüssen und viele andere), von denen jeder die Flugbahn des Geschosses nur unbedeutend beeinflusst, folgt aus dem Satz von Ljapunow, dass die Streuung einem Normalgesetz unterliegen muss.

Dieser Umstand wird auch in der Ballistik berücksichtigt und. schon bei der Ausarbeitung der Schießregeln einkalkuliert.

Führt man irgendeine Beobachtung zur Messung irgendeiner physikalischen Konstanten durch, so werden im Ergebnis unserer Beobachtung unvermeidlich eine riesige Menge Faktoren auftreten, die man unmöglich im einzelnen berechnen kann, die aber irgendwelche Messfehler erzeugen.
Hierher gehören beispielsweise Mängel des Messgerätes, dessen Zeiger sich unter der Einwirkung verschiedenster Ursachen verändern können (es kann sich um atmosphärische, wärmetechnische, mechanische oder irgendwelche anderen Gründe handeln).
Ebenfalls hierher gehören auch die Unzulänglichkeiten beim Beobachter, die durch Besonderheiten seines Sehvermögens oder seines Gehörs hervorgerufen werden und sich auch in Abhängigkeit vom physischen und psychischen Zustand des Beobachters unmerklich ändern können.

Der tatsächliche Messfehler resultiert also aus einer Vielfalt der Größe nach geringfügiger, voneinander unabhängiger sogenannter elementarer, vom Zufall abhängiger Fehler. Nach dem Satz von Ljapunow können wir wieder annehmen, dass die Beobachtungsfehler nach einem Normalgesetz verteilt sind.

Die Anzahl der Beispiele kann noch beliebig vergrößert werden:
Die Lage und Geschwindigkeit der Gasmoleküle, die durch eine große Anzahl von Zusammenstößen mit anderen Molekülen bestimmt werden; die Menge eines durch eine Wand diffundierenden Stoffes; die Abweichungen der Einzelteile eines Mechanismus von vorgegebenen Ausmaßen bei der 'Massenherstellung solcher Mechanismen; die Verteilung des Wachstums der Tiere, der Pflanzen oder irgendwelcher Organismen usw.

Die Vervollkommnung der statistischen Physik und auch einige Zweige der Technik stellten

der Wahrscheinlichkeitsrechnung eine große Zahl völlig neuer Probleme, die ebenfalls nicht in den Rahmen des klassischen Schemas hineinpassten.
In der Physik und in der Technik interessierte die Untersuchung von Prozessen, d. h. von Erscheinungen, die sich mit der Zeit ändern. Die Wahrscheinlichkeitsrechnung hatte weder allgemeine Verfahren noch ausgearbeitete Spezialschemata zur Lösung der sich bei der Untersuchung dieser Erscheinungen ergebenden Probleme. Es entstand so das dringende Bedürfnis nach einer Ausarbeitung einer allgemeinen Theorie der zufälligen Prozesse, d. h. einer Theorie, die sich mit Zufallsgrößen beschäftigt, die von einem oder mehreren sich stetig ändernden Parametern abhängen.

Wir zählen hier einige Probleme auf, die auf die Betrachtung von Zufallsgrößen führen, die sich mit der Zeit ändern. Wir stellen uns vor, dass wir die Bewegung irgendeines Gasmoleküls oder irgendeines Flüssigkeitsmoleküls verfolgen. Dieses Molekül wird in einem zufälligen Zeitmoment mit einem anderen Molekül zusammenstoßen und dabei seine Bewegungsrichtung und seine Geschwindigkeit ändern.
Diese Änderung der Lage eines Moleküls unterliegt in jedem Moment der Wirkung des Zufalls. Um eine ganze Reihe physikalischer Vorgänge zu studieren, ist es gerade erforderlich, dass man die Wahrscheinlichkeit dafür berechnen kann, dass ein Molekül sich nach einer gewissen Zeit in dieser oder jener Lage befindet.

Bringt man beispielsweise zwei Gase oder zwei Flüssigkeiten miteinander in Berührung, so werden sich die Moleküle der beiden Flüssigkeiten miteinander vermischen. Wie schnell läuft dieser Diffusionsprozess ab? Nach welchen Gesetzen geht er vor sich ?
Wann hat sich eine praktisch homogene Mischung der Flüssigkeiten herausgebildet?
Auf alle diese Fragen kann die statistische Diffusionstheorie eine klare Antwort geben; auf Grund dieser Theorie sind wahrscheinlichkeitstheoretische Berechnungen, die Untersuchung der zufälligen Prozesse möglich.

Ein ähnliches Problem gibt es offenbar auch in der Chemie, wenn man den Prozess der chemischen Reaktion untersucht; welcher Teil der Moleküle wirkt bei der Reaktion mit; wie läuft die Reaktion mit der Zeit ab, wann ist die Reaktion praktisch als abgeschlossen zu betrachten?

Ein ganz wichtiger Kreis von Erscheinungen ergibt sich beim radioaktiven Zerfall. Bei dieser Erscheinung zerfallen Atome eines radioaktiven Stoffes; sie gehen in Atome eines anderen Elementes über. Jeder Atomzerfall erfolgt plötzlich, ähnlich einer Explosion unter Freiwerden einer gewissen Energiemenge.

Zahlreiche Beobachtungen zeigen nun, dass dieser Zerfall von Atomen in zufälligen Zeitmomenten und unabhängig voneinander vor sich geht (unter der Voraussetzung, dass die Masse des radioaktiven Stoffes nicht zu groß ist). Für die Untersuchung des radioaktiven Zerfalls ist es ganz wesentlich festzustellen, wie groß die Wahrscheinlichkeit dafür ist, dass in einem bestimmten Zeitintervall eine gewisse Anzahl von Atomen zerfällt. Dieses Beispiel ist ein typisches Beispiel der Theorie der zufälligen Prozesse.

Formal ist es, wenn man als Erklärung nur ein mathematisches Bild der Erscheinungen gibt; denn es gibt viele andere Erscheinungen, die genauso vor sich gehen. So z. B. die Frage nach der Belastung einer Telefonzentrale, d. h. nach der Zahl der Anrufe an die Telefonzentrale von Teilnehmern, oder die Brownsche Molekularbewegung, das Abreißen eines Fadens in einer Water-Maschine und viele andere.

Zu Beginn der dreißiger Jahre legten die sowjetischen Mathematiker A. N. Kolmogoroff und

A. J. Chintschin durch ihre fundamentalen Arbeiten den Grundstein zur allgemeinen Theorie der zufälligen Prozesse. Im ersten Jahrzehnt unseres Jahrhunderts, also etwas früher, untersuchte A. A. Markoff Folgen unabhängiger Zufallsgrößen, die den Namen Markoffsche Ketten erhielten.

Diese Theorie, zuerst nur als mathematische Disziplin entwickelt, erfuhr in den zwanziger Jahren in den Händen der Physiker eine Umwandlung in ein wirksames Hilfsmittel zur Erforschung der Natur. Seit jener Zeit lieferten viele berühmte Gelehrte S. N. Bernstein, W. I. Romanowski, A. N. Kolmogoroff, J. Hadamard, M. Frechet und andere) bedeutende Beiträge zur Theorie der Markoffschen Ketten.

In den zwanziger Jahren entdeckten A. N. Kolmogoroff, E. E. Sluzki, A. J. Chintschin und Paul Lévy einen engen Zusammenhang der Wahrscheinlichkeitsrechnung mit den mathematischen Disziplinen, welche die Mengen und den allgemeinen Funktionsbegriff (Mengenlehre und Theorie der Funktionen einer reellen Veränderlichen) untersuchen.
Die Entdeckung dieses Zusammenhangs erwies sich als außerordentlich fruchtbar. Auf diesem Wege gelangte man auch zu endgültigen Resultaten für klassische Probleme, die von Tschebyscheff aufgeworfen worden waren.

Schließlich sind noch die Arbeiten von S. N. Bernstein und A. N. Kolmogoroff zu erwähnen, in denen die Grundbegriffe der Wahrscheinlichkeitsrechnung geklärt werden.

Abschließend können wir sagen, dass russische Gelehrte eine entscheidende Rolle bei der Gesamtentwicklung der Wahrscheinlichkeitsrechnung spielten. Von wichtigen Ideen der Wahrscheinlichkeitsrechnung können wir mit Stolz behaupten, dass sie in Russland bzw. der Sowjetunion entweder entstanden oder mächtig weiterentwickelt wurden oder ihre glanzvolle Vollendung fanden.

14 Anhang

Wertetabelle der Größe $\Phi(a)$

a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$
0,00	0,000	0,40	0,311	0,80	0,576	1,20	0,770	1,60	0,890
0,01	0,008	0,41	0,318	0,81	0,582	1,21	0,774	1,61	0,893
0,02	0,016	0,42	0,326	0,82	0,588	1,22	0,778	1,62	0,895
0,03	0,024	0,43	0,333	0,83	0,593	1,23	0,781	1,63	0,897
0,04	0,032	0,44	0,340	0,84	0,599	1,24	0,785	1,64	0,899
0,05	0,040	0,45	0,34?	0,85	0,605	1,25	0,789	1,65	0,901
0,06	0,048	0,46	0,354	0,86	0,610	1,26	0,792	1,66	0,903
0,07	0,056	0,47	0,362	0,87	0,616	1,27	0,796	1,67	0,905
0,08	0,084	0,48	0,369	0,88	0,621	1,28	0,800	1,68	0,907
0,09	0,072	0,49	0,376	0,89	0,627	1,29	0,803	1,69	0,909
0,10	0,080	0,50	0,383	0,90	0,632	1,30	0,806	1,70	0,911
0,11	0,088	0,51	0,390	0,91	0,637	1,31	0,810	1,71	0,913
0,12	0,096	0,52	0,397	0,92	0,642	1,32	0,813	1,72	0,915
0,13	0,103	0,53	0,404	0,93	0,648	1,33	0,816	1,73	0,916
0,14	0,111	0,54	0,411	0,94	0,653	1,34	0,820	1,74	0,918
0,15	0,119	0,55	0,418	0,95	0,658	1,35	0,823	1,75	0,920
0,16	0,127	0,56	0,425	0,96	0,663	1,36	0,826	1,76	0,922
0,17	0,135	0,57	0,431	0,97	0,668	1,37	0,829	1,77	0,923
0,18	0,143	0,58	0,438	0,98	0,673	1,38	0,832	1,78	0,925
0,19	0,151	0,59	0,445	0,99	0,678	1,39	0,835	1,79	0,927
0,20	0,159	0,60	0,451	1,00	0,683	1,40	0,838	1,80	0,928
0,21	0,166	0,61	0,458	1,01	0,688	1,41	0,841	1,81	0,930
0,22	0,174	0,62	0,465	1,02	0,692	1,42	0,844	1,82	0,931
0,23	0,182	0,63	0,471	1,03	0,697	1,43	0,847	1,83	0,933
0,24	0,190	0,64	0,478	1,04	0,702	1,44	0,850	1,84	0,934
0,25	0,197	0,65	0,484	1,05	0,706	1,45	0,853	1,85	0,936
0,26	0,205	0,66	0,491	1,06	0,711	1,46	0,856	1,86	0,937
0,27	0,213	0,67	0,497	1,07	0,715	1,47	0,858	1,87	0,939
0,28	0,221	0,68	0,504	1,08	0,720	1,48	0,861	1,88	0,940
0,29	0,228	0,69	0,510	1,09	0,724	1,49	0,864	1,89	0,941
0,30	0,236	0,70	0,516	1,10	0,729	1,50	0,866	1,90	0,943
0,31	0,243	0,71	0,522	1,11	0,733	1,51	0,867	1,91	0,944
0,32	0,251	0,72	0,528	1,12	0,737	1,52	0,871	1,92	0,945
0,33	0,259	0,73	0,535	1,13	0,742	1,53	0,874	1,93	0,946
0,34	0,266	0,74	0,541	1,14	0,746	1,54	0,876	1,94	0,948
0,35	0,274	0,75	0,547	1,15	0,750	1,55	0,879	1,95	0,949
0,36	0,281	0,76	0,553	1,16	0,754	1,56	0,881	1,96	0,950
0,37	0,289	0,77	0,559	1,17	0,758	1,57	0,884	1,97	0,951
0,38	0,296	0,78	0,565	1,18	0,762	1,58	0,886	1,98	0,952
0,39	0,303	0,79	0,570	1,19	0,766	1,59	0,888	1,99	0,953

a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$
2,00	0,955	2,20	0,972	2,40	0,984	2,60	0,991	2,88	0,996
2,01	0,956	2,21	0,978	2,41	0,984	2,61	0,991		
2,02	0,957	2,22	0,974	2,42	0,984	2,62	0,991	2,90	0,996
2,08	0,958	2,28	0,974	2,48	0,985	2,68	0,991	2,92	0,996
2,04	0,959	2,24	0,975	2,44	0,985	2,64	0,992	2,94	0,997
2,05	0,960	2,25	0,976	2,45	0,986	2,65	0,992	2,96	0,997
2,06	0,961	2,26	0,976	2,46	0,986	2,66	0,992	2,98	0,997
2,07	0,962	2,27	0,977	2,47	0,986	2,67	0,992	3,00	0,997
2,08	0,962	2,28	0,977	2,48	0,987	2,68	0,998	3,10	0,998
2,09	0,968	2,29	0,978	2,49	0,987	2,69	0,998	3,20	0,999
2,10	0,964	2,80	0,979	2,50	0,988	2,70	0,998	3,30	0,999
2,11	0,965	2,81	0,979	2,51	0,986	2,72	0,998	3,40	0,999
2,12	0,966	2,82	0,980	2,52	0,988	2,74	0,994	3,50	0,9995
2,18	0,967	2,88	0,980	2,58	0,989	2,76	0,994	3,60	0,9997
2,14	0,968	2,84	0,981	2,54	0,989	2,78	0,995	3,70	0,9998
2,15	0,968	2,85	0,981	2,55	0,989			3,80	0,99986
2,16	0,933	2,86	0,982	2,56	0,990	2,89	0,995	3,90	0,99990
2,17	0,970	2,87	0,982	2,57	0,990	2,82	0,995		
2,18	0,971	2,88	0,988	2,58	0,990	2,84	0,995	4,00	0,99994
2,19	0,971	2,89	0,988	2,59	0,990	2,86	0,996	5,00	0,99999994