UNTERLAGEN zum STOCHASTIKSEMINAR

TEIL 1

Martina Weiß und *Ferdinand Österreicher*
IFFB Fachdidaktik und LehrerInnenbildung, FB Mathematik
der Universität Salzburg

Salzburg, März 2007

Inhaltsverzeichnis

Kapitel 1

BESCHREIBENDE STATISTIK

1.1 Darstellungsformen eindimensionaler Daten

1.1.1 Einführende Beispiele

Beispiel 1: Chuck-a-luck

In den Lehrbüchern der Wahrscheinlichkeitsrechnung aus dem angelsächsischen Raum findet sich häufig das folgende Spielchen, "Chuck-a-luck" genannt, das in geringfügig modifizierter Form auch im Casinospiel "Sic-Bo" integriert ist.

Man setzt einen Einsatz e auf eine Zahl $z \in \{1, ..., 6\}$. Dann werden drei Würfel geworfen. Kommt z unter den drei Augenzahlen nicht vor, ist der Einsatz verloren. Kommt z genau ein-, zwei- oder dreimal vor, gewinnt man den ein-, zwei- bzw. dreifachen Einsatz.

Man spiele 36 Spiele und trage den Gewinn jedes Spieles gegen dessen Nummer auf.

Der Wertebereich des Gewinns X ist, bezogen auf den Einsatz $e = 1$, gleich $W_X = \{-1, 1, 2, 3\}$.

Beispiel 2: Warten bis zum ersten Erfolg

Ein Versuch bestehe darin, eine Münze so oft zu werfen, bis zum ersten Mal das Ereignis "Kopf" eintritt. X bezeichne die Anzahl der nötigen Münzwürfe.

Man führe $2^6 = 64$ Versuche durch und erstelle ein Histogramm für die Häufigkeiten der beobachteten "Wartezeiten" X_i, $i \in \{1, ..., 64\}$, und zwar für die Klassen $K_1 = \{1\}, ..., K_6 = \{6\}$ und $K_7 = \{7, 8, ...\}$.

Der Wertebereich einer "Wartezeit" X ist die abzählbar unendliche Menge $W_X = \mathbb{N} \cup \{\infty\}$.

1.1.2 Tabellarische Darstellung

Wir gehen auch weiterhin davon aus, dass der Wertebereich W_X der zugrundeliegenden Variablen X eine endliche oder abzählbar unendliche Teilmenge der reellen Zahlen ist, dass also gilt $W_X = \{\omega_1, \omega_2, ...\} \subset \mathbb{R}$. Wenn möglich - und das ist vielfach der Fall - nehmen wir an, dass gilt $\omega_1 < \omega_2 < ...$.

Ausgehend von einer Urliste, das ist eine Folge $x_1, x_2, ..., x_n \in W_X$ von Beobachtungswerten der Variablen X erstellen wir eine Häufigkeitstabelle.

Dann sind

$$\begin{aligned} H_j &= |\{i \in \{1, ..., n\} : x_i = \omega_j\}| \\ &= \text{die Anzahl der Versuche mit dem Ausfall } \omega_j \end{aligned}$$

bzw.

$$h_j = \frac{H_j}{n} = \frac{\text{Anzahl der Versuche mit dem Ausfall } \omega_j}{\text{Gesamtanzahl der Versuche}}$$

die *absolute Häufigkeit* bzw. die *relative Häufigkeit* des Ausfalls ω_j , $j \geq 1$.

1.1.3 Graphische Darstellung

a) Die Strichliste

Ist der Wertebereich der Variablen im Vergleich mit dem gewählten Stichprobenumfang klein und ist man von vornherein nicht an der Reihenfolge interessiert, in welcher die Daten erhoben werden, so ist zu erwägen, eine Strichliste anzufertigen [1]. Dabei sind die einzelnen Beobachtungswerte unmittelbar bei deren Erhebung durch je einen waagrechten (oder senkrechten)

[1] Es gibt Anwendungen, bei welchen man sich speziell für die Reihenfolgen von Messwerten interessiert. Typisch dafür sind Folgen von Tagen mit und ohne Niederschlag. Hinsichtlich der Untersuchung sogenannter *Läufe von binären Zufallsfolgen* sei auf die Diplomarbeit von Frau *Radauer* [20] verwiesen.

Strich über (bzw. neben) dem zugehörigen Wert der Variablen in einer Strichliste einzutragen. Zur Wahrung der Übersicht - und um das spätere Abzählen zu erleichtern - formt man in der hergebrachten Weise Blöcke zu je 5 Strichen. Auf diese Weise kann bereits während der Datenerhebung eine graphische Darstellung der Verteilung der Beobachtungswerte entstehen.[2]

b) Das Histogramm[3]

Bei einem Histogramm für die relativen Häufigkeiten der Werte einer metrischen Variablen trägt man über jeder Klasse die relative Häufigkeit pro Einheit der Variablen auf, also

$$\text{Höhe} = \frac{\text{relative Häufigkeit}}{\text{Klassenbreite}}$$

Ein Histogramm beruht auf einer geeigneten Klasseneinteilung. Eine solche ist die Partition eines Intervalls I von $\mathbb{R}$, welches den Wertebereich der Variablen umfasst, in endlich viele Intervalle (Klassen).

Für Variable mit abzählbar unendlichem oder überabzählbarem Wertebereich ist eine Klassenzusammenfassung unumgänglich. (Beispiele: Wartezeit bis zum ersten Erfolg, Anzahl der radioaktiven Zerfälle in einem Zeitintervall vorgegebener Länge. Messergebnisse mit prinzipiell beliebig genauer Messgenauigkeit.)

Bezeichnung: Ist der Wertebereich W_X einer Variablen X eine Teilmenge von $\{\mu+c\cdot z : z \in \mathbb{N}_0\}$ oder $\{\mu+c\cdot z : z \in \mathbb{Z}\}$ mit $c > 0$ und $\mu \in \mathbb{R}$, so sagt man, W_X habe eine *Gitterstruktur mit der Gitterkonstante*[4] c. Seien $h_{\mu+c\cdot z}$ die relativen Häufigkeiten der Ausfälle $\mu + c \cdot z$ eines Versuches. Dann heißt die Funktion

$$f(x) = \sum_{z\in\mathbb{N}_0(\mathbb{Z})} \frac{h_{\mu+c\cdot z}}{c} \cdot 1_{[\mu+c(z-1/2),\mu+c(z+1/2))}(x)\,,\ x \in \mathbb{R}$$

das zugehörige *Histogramm*. (Man stellt das Histogramm zumeist so dar, dass man zu jedem Wert $\mu + c \cdot z$ ein Rechteck mit der Basis $[\mu + c(z -$

[2] Manche Beobachter verzichten bewusst darauf, während des Beobachtungsverlaufs Strichlisten anzufertigen, um jegliche subjektive Erwartungshaltung auszuschließen.

[3] Die Bezeichnung "Histogramm" (histogram) wurde vermutlich vom englischen Statistiker *Karl Pearson* (1857 − 1936) im Jahre 1895 erstmals verwendet.

[4] Da man stets davon ausgehen kann, dass die Messgenauigkeit jedes noch so guten Messinstruments beschränkt ist, ist zumeist auch die Annahme einer Gitterstruktur angebracht.

$1/2), \mu + c(z + 1/2))$ und dem Flächeninhalt $h_{\mu+c\cdot z}$ - und somit der Höhe $h_{\mu+c\cdot z}/c$ - zeichnet.)

Beispiel 3: Radioaktiver Zerfall

Ernest Rutherford (1871 – 1937) und *Hans W. Geiger* (1882 – 1945) verwendeten bei ihrem klassischen Experiment im Jahre 1910 eine Polonium-Quelle und registrierten für 2608 disjunkte Zeitintervalle von je 7.5 Sekunden Dauer die Anzahl der Szintillationen. Im Folgenden sind die beobachteten absoluten Häufigkeiten H_j der Intervalle mit j, $j \in \{0, ..., 14\}$ Szintillationen dargestellt (vgl. beispielsweise [12], S. 36).

j	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	≥ 15
H_j	57	203	383	525	532	408	273	139	45	27	10	4	0	1	1	0

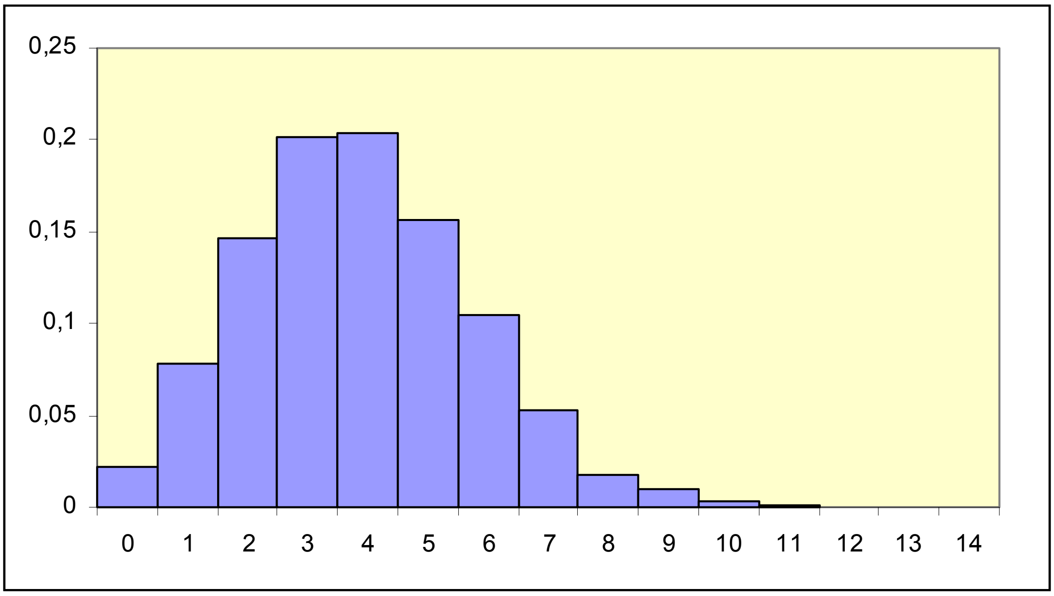


Abbildung: Zugehöriges Histogramm

Allgemeine Hinweise zur Klassenbildung

Generell lässt sich sagen, dass die Wahl der Klassen einige Erfahrung voraussetzt. Verwenden Sie nicht das erstbeste Histogramm. Experimentieren Sie, um eine möglichst charakteristische Form zu erzielen!

- Wahl der Anzahl der Klassen

Diese sollte man nicht zu klein und nicht zu groß wählen. Zu viele Klassen erzeugen ein unerwünscht erratisches Muster der Häufigkeiten, zu wenige ein unangebracht grobes Muster (Siehe Aufgabe 1). Als Orientierungshilfe kann im Fall *unimodaler* (*eingipfliger*) Verteilungen der gerundete Funktionswert $1+\log_2(n)$ des Stichprobenumfangs dienen [5].

- Wahlen von μ und der Klassenbreite c

  Es ist vielfach bequem, $\mu = 0$ zu wählen. Für die Wahl der Klassenbreite c kann ein geeignetes Stängel-Blatt-Diagramm der Daten als Orientierungshilfe dienen. Weitere Orientierungshilfen hinsichtlich der Wahl von Klassenbreite und Klassengrenzen finden Sie in Abschnitt 1.3.3 von [17].

Verwendungszweck von Histogrammen

Histogramme werden verwendet,

(i) um eine Datenmenge zwecks guter visueller Wahrnehmbarkeit von allgemeinen Charakteristika der Verteilung der Beobachtungswerte, wie repräsentativem Wert, Streuungsverhalten und charakteristischer Gestalt zu verdichten,

(ii) um Hinweise zur Wahl eines geeigneten Wahrscheinlichkeitsmodells und nötigenfalls einer geeigneten Variablentransformation zu erhalten - welcher später eine genauere statistische Analyse folgen kann - und

(iii) um ein unerwartetes Verhalten der zugrundeliegenden Variablen festzustellen und/oder ungewöhnliche Beobachtungswerte zu entdecken.

c) **Das Stängel-Blatt-Diagramm** [6]

Ist der Wertebereich einer Variablen vergleichsweise groß und ist man nic ht an der Reihenfolge interessiert, in welcher die Daten erhoben werden, so kann die Anfertigung eines sogenannten *Stängel-Blatt-Diagramms* -

[5] Für Beispiel 2 stimmt $1+\log_2(64) = 7$ mit der vorgeschlagenen Klassenanzahl überein.

[6] Das *Stängel-Blatt-Diagramm* (*stem-and-leaf-display*) ist - wie das *Kastenbild* (*box-plot*) in Abschnitt 1.2.2 eine Darstellungsform der *Explorativen Datenanalyse* (*Exploratory Data Analysis*). Dieser Zweig der modernen beschreibenden Statistik ist eine Schöpfung des US-amerikanischen Statistikers *John W. Tukey* (1915 − 2000), vom dem übrigens auch die Kurzbezeichnung *bit* für *b*inary dig*it* stammt.

wie die einer Strichliste - bereits während der Datenerhebung erfolgen. Im wesentlichen entspricht dies nämlich dem Erstellen einer Strichliste mit zusätzlicher Klassenbildung. Im folgenden sei diese, vom amerikanischen Statistiker *John W. Tukey* (1977), dem Begründer der sogenannten Explorativen Datenanalyse, vorgeschlagene visuelle Darstellung von Daten anhand eines Beispiels erläutert.

Anmerkungen zur Anfertigung eines Stängel-Blatt-Diagramms

für eine Liste von Rohdaten mit Werten beispielsweise aus $\{70, ..., 115\}$

Zunächst werden alle möglichen Zehnerzahlen (Zehnerziffer bzw. Paar, bestehend aus Hunderter- und Zehnerziffer), also $7, ..., 11$ im "Stängel" - d.h. vertikal - eingetragen,

anschließend wird die Liste der Rohdaten der Reihe nach durchgegangen und für jeden Wert der Liste wird hinter der zugehörigen Zehnerzahl im Stängel die Einerziffer als Blatt - d.h. horizontal - eingetragen.

Modifikationsmöglichkeit: Sollte sich die Klassenbildung nach Zehnerzahlen als zu grob erweisen, ist zu erwägen, die Zehnerzahlen im Stängel jeweils zweimal anzuschreiben und hinter der jeweils ersten Zehnerzahl die Einerziffern von 0 bis 4 und hinter der jeweils zweiten Zehnerzahl die Einerziffern von 5 bis 9 als Blatt anzufügen.

d) Die empirische Verteilungsfunktion

Seien $x_1, ..., x_n$ reelle Beobachtungswerte, sei x eine weitere reelle Zahl und sei

$$|\{i \in \{1, ..., n\} : x_i \leq x\}|$$

die absolute Häufigkeit der Beobachtungen, deren Bobachtungswerte kleiner oder gleich x sind. Dann heißt die Abbildung

$$x \longmapsto F_n(x) = \frac{|\{i \in \{1, ..., n\} : x_i \leq x\}|}{n}$$

die *empirische Verteilungsfunktion* der Beobachtungswerte $x_1, ..., x_n$.

Anmerkung: Im folgenden wird die empirische Verteilungsfunktion mit Hilfe des Wertebereichs ihrer Funktionswerte definiert. Zu diesem Zweck ordnet man die Beobachtungswerte ihrer Größe nach: Es bezeichne

$x_{1:n}$	den kleinsten von n Beobachtungswerten,
$x_{2:n}$	den zweit-kleinsten Beobachtungswert,
...	...
$x_{n:n}$	den größten Beobachtungswert.

Dann gelten $x_{1:n} \leq x_{2:n} \leq ... \leq x_{n:n}$, $\{x_{1:n}, ..., x_{n:n}\} = \{x_1, ..., x_n\}$ und

$$F_n(x) = \begin{cases} 0 & \text{für} \quad & & x & < & x_{1:n} \\ \frac{i}{n} & \text{für} \quad & x_{i:n} & \leq & x & < & x_{i+1:n} \,, \quad i \in \{1, ..., n-1\} \\ 1 & \text{für} \quad & x_{n:n} & \leq & x \end{cases}$$

Anmerkung zur "Kurve von Quetelet" [7]

Der nachstehende Bericht des Mathematikers *B.L.* v*an der Waerden* gibt einerseits einen interessanten Einblick in die Ideengeschichte der Statistik und andererseits eine Anregung, eine empirische Verteilungsfunktion plastisch darzustellen.

"Lebhaft erinnere ich mich noch, wie mein Vater mich als Knaben eines Tages an den Rand der Stadt führte, wo am Ufer die Weiden standen, und mich hundert Weidenblätter willkürlich pflücken ließ. Nach Aussonderung der beschädigten Spitzen blieben noch 89 *unversehrte Blätter übrig, die wir dann zu Hause, nach abnehmender Größe geordnet, wie Soldaten in Reih und Glied stellten. Dann zog mein Vater durch die Spitzen eine gebogene Linie und sagte: "Dies ist die Kurve von QUETELET. Aus ihr siehst du, wie die Mittelmäßigen immer die große Mehrheit bilden und nur wenige nach oben hervorragen oder nach unten zurückbleiben." "*

Beispiel 4: Zur "Kurve von Quetelet"

W. *Rohm* (HTL Saalfelden) und *F. Österreicher* haben 1982 gemäß *B.L. van der Waerden*s Beschreibung einer Föhre 100 Nadeln aufs Geratewohl entnommen und deren Längen (in *mm*) gemessen. Die geordneten Messergebnisse sind im nachstehenden Stängel-Blatt-Diagramm dargestellt.

7	044778
8	000134455677777888889999
9	000111112222333444445555566666777778899
10	001122222233333444556667777
11	0115

[7] Der belgische Statistiker *Lambert-Adolphe-Jacques Quetelet* (1796 – 1874) erschloss der Normalverteilung in der Anthropometrie ein gänzlich neues und unvermutetes Anwendungsgebiet und übte damit einen prägenden Einfluss auf *Francis Galton* (1822 – 1911) aus. Auf seinen Einfluss geht auch die Gründung vieler statistischer Behörden in Europa zurück.

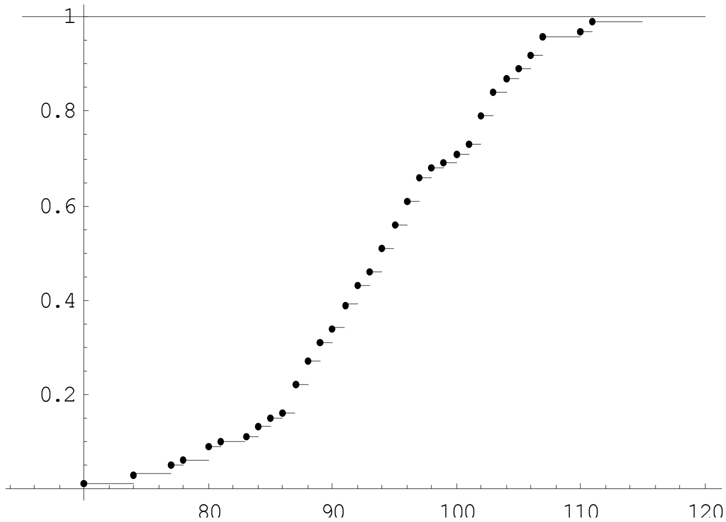


Abbildung: Zugehörige empirische Verteilungsfunktion

1.2 Kenngrößen eindimensionaler Daten: Zentral- und Streumaße

Wir gehen im folgenden - sofern nicht anders angenommen - von Daten $x_1, ..., x_n$ einer eindimensionalen Variablen (d.h. mit Werten aus $\mathbb{R}$), versehen mit der natürlichen Ordnung " $<$ " und dem üblichen Abstand, aus.

1.2.1 Stichprobenmittel und Standardabweichung

a) **Das Stichprobenmittel**

Das *arithmetische Mittel*

$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$$

der Stichprobenwerte heißt *Stichprobenmittel* oder *Mittelwert.*

Anmerkung 1: Der britische Wissenschaftler *Thomas Simpson* ($1710 - 1761$) schlug für die Handhabung von mehrfachen, erfahrungsgemäß verschiedenen Messwerten einer Messgröße die Verwendung des Stichprobenmittels vor. Mit ein Motiv dafür war wohl, die Subjektivität der Beobachter beim Umgang mit solchen Daten hintanzuhalten und durch eine standardisierte Vorgangsweise den Austausch von Messergebnissen zu erleitern. Er schreibt:

"Zusammenfassend scheint es, dass das Bestimmen des arithmetischen Mittels einer Anzahl von Messwerten die Chance kleiner Fehler beträchtlich verringert und nahezu jede Möglichkeit für große ausschließt. Diese Erwägung allein scheint ausreichend, um die Verwendung dieser Methode nicht nur Astronomen zu empfehlen, sondern allen, die Präzisionsmessungen durchführen. Je mehr Beobachtungen oder Experimente gemacht werden, desto weniger werden die Resultate fehleranfällig sein, vorausgesetzt, eine Wiederholung der Messungen ist unter gleichen Bedingungen möglich."

Anmerkung 2: Liegen bereits die relativen Häufigkeiten h_j der einzelnen Ausfälle ω_j , $j \geq 1$, der Variablen X vor, so ist es zweckmäßiger, das Stichprobenmittel gemäß

$$\bar{x}_n = \sum_{j\geq 1} \omega_j \cdot h_j$$

mit deren Hilfe auszudrücken. Falls möglich denken wir uns dabei die ω_j wieder der Größe nach geordnet: $\omega_1 < \omega_2 < \dots$.

Zu den beiden Darstellungen des Stichprobenmittels äußerte sich *Henri Lebesgue*[8], der Schöpfer des nach ihm benannten Lebesgue-Integrals, welches in der höheren Analysis das *Riemann*-Integral ersetzt, wie folgt.

"Man kann auch sagen, dass man sich bei der Verwendung der ersten Berechnungsart wie ein Kaufmann ohne System verhält, der Geldstücke und Banknoten in der zufälligen Reihenfolge zählt, wie er sie in die Hand bekommt. Während wir bei der zweiten Rechenart vorgehen wie ein umsichtiger Kaufmann, der sagt:

Ich habe	$H(E_1)$	Münzen zu einer Krone, macht	$1 \times H(E_1)$,
ich habe	$H(E_2)$	Münzen zu zwei Kronen, macht	$2 \times H(E_2)$,
ich habe	$H(E_5)$	Münzen zu fünf Kronen, macht	$5 \times H(E_5)$, usw.;

ich habe also insgesamt

$$S = 1 \times H(E_1) + 2 \times H(E_2) + 5 \times H(E_5) + \dots ,$$

weil er - wie reich er auch sein mag - nur eine endliche Anzahl von Banknoten zu zählen hat."

Die **Minimaleigenschaft des Stichprobenmittels** bezüglich der Summe der Abweichungsquadrate

Behauptung: Die Funktion

$$f(x) = \frac{1}{n} \sum_{i=1}^{n} (x_i - x)^2$$

ist genau dann minimal, wenn $x = \bar{x}_n$ ist.

Beweisvariante 1: Die Ableitung von $f(x)$ ist

$$f'(x) = -\frac{2}{n} \sum_{i=1}^{n} (x_i - x) = 2\,(x - \bar{x}_n) \begin{cases} < 0 & \text{für } \ x < \bar{x}_n \\ = 0 & \text{für } \ x = \bar{x}_n \\ > 0 & \text{für } \ x > \bar{x}_n \,. \end{cases}$$

[8] *Henri Lebesgue* (1875-1941), französischer Mathematiker

Somit wird das Minimum im Punkt $x = \bar{x}_n$ angenommen.

Beweisvariante 2 bedient sich einer Aussage aus der Mechanik, welche trotz ihrer Einfachheit äußerst nützlich ist. Diese Aussage ist

Der Steinersche Verschiebungssatz[9]**:** Es gilt

$$\sum_{i=1}^{n} (x_i - x)^2 = \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 + n(x - \bar{x}_n)^2 \,.$$

Wegen $n(x - \bar{x}_n)^2 \geq 0$ ergibt sich daraus als **Folgerung** die gewünschte Aussage, nämlich

$$\sum_{i=1}^{n} (x_i - x)^2 \geq \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 \ ,$$

wobei Gleichheit offensichtlich genau dann gilt, wenn $x = \bar{x}_n$ ist.

Anmerkung 3: Interpretiert man die ω_j, $j \geq 1$, als Punkte der x-Achse, in denen die Massen H_j sitzen, und lässt man diese Massen um eine Achse rotieren, die durch den Punkt x der x-Achse geht und normal auf dieser steht, dann ist die Summe

$$\sum_{i=1}^{n} (x_i - x)^2 = \sum_{j \geq 1} (\omega_j - x)^2 \, H_j$$

das zugehörige Trägheitsmoment der Massenverteilung[10]. Die Folgerung des Stein-erschen Verschiebungssatzes besagt also, dass das Trägheitsmoment genau dann kleinstmöglich ist, wenn die Drehachse durch den Massenmittelpunkt $\bar{x}_n$ der Massenverteilung geht.

Beweis des Satzes: Die Anwendung der unmittelbar einsichtigen (und auch geometrisch interpretierbaren) Beziehung

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

für $a = x_i - x$ und $b = x_i - \bar{x}_n$ ergibt

$$(x_i - x)^2 = (x_i - \bar{x}_n)^2 + (\bar{x}_n - x)^2 + 2(\bar{x}_n - x)\,(x_i - \bar{x}_n) \ .$$

[9] parallel–axis theorem, *Jakob Steiner* (1796-1863), Schweizer Geometer

[10] Der Begriff des Trägheitsmoments wurde - wie der des Erwartungswerts in der Wahrscheinlichkeitsrechnung - vom holländischen Wissenschaftler *Christiaan Huygens* (1629 – 1695) geprägt. Er ist überdies der Erfinder der Pendeluhr, der Entdecker der Saturnringe und der Urheber des *Huygensschen Prinzips* in der Optik.

Daraus ergibt sich die Aussage des Verschiebungssatzes durch Summation über $i \in \{1, ..., n\}$ und Berücksichtigen von $\sum_{i=1}^{n}(x_i - \bar{x}_n) = \sum_{i=1}^{n} x_i - n\,\bar{x}_n = 0$.

b) **Stichprobenvarianz und Standardabweichung** [11]

Die Größe

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$$

heißt Stichprobenvarianz. Deren Quadratwurzel

$$s_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}$$

heißt *Standardabweichung der Stichprobe.* Nicht die Stichprobenvarianz, sondern die Standardabweichung ist die eigentliche Maßzahl für die Abweichung der Daten vom Mittelwert. Dies kann man sich etwa mit Hilfe einer Dimensionsbetrachtung überlegen, wie sie in der Physik üblich ist. Sind beispielsweise die x_i die Ergebnisse der Messungen einer Länge, so sind auch die Differenzen $x_i - \bar{x}_n$ Längen. Deren Quadrate $(x_i - \bar{x}_n)^2$ sind demnach die Flächen der zugehörigen Quadrate, sodass auch s_n^2 ein Maß für eine Fläche ist. Erst deren Quadratwurzel, die Standardabweichung s_n, hat wieder die richtige "Dimension", nämlich die einer Länge.

Warum dividiert man durch $n-1$ und nicht durch n?

Vordergründige Antwort: Wegen der Bedingung $\sum_{i=1}^{n}(x_i - \bar{x}_n) = 0$ ist eine der Abweichungen $x_i - \bar{x}_n$ durch die restlichen $n-1$ Abweichungen festgelegt. Da also nur $n-1$ Summanden frei variieren können, dividiert man durch $n-1$. Einer physikalischen Tradition folgend nennt man diese Zahl auch die *Anzahl der Freiheitsgrade.*

Hintergründige Antwort: Dazu bedarf es eines stochastischen Modells. Sei X eine Zufallsvariable mit Erwartungswert μ und unbekannter Varianz σ^2. Wäre nun der Erwartungswert μ bekannt, so würde man σ^2

[11] Die Bezeichnung "*standard deviation*" (Standabweichung) wurde *Karl Pearson* im Jahre 1893 geprägt. Der Begriff selbst wurde jedoch bereits vom deutschen Mathematiker *Carl Friedrich Gauß* (1777 – 1855) im Rahmen der Fehlerrechung verwendet. Bei Gauß hieß diese Größe "*mittlerer zu befürchtender Fehler*".

naturgemäß durch

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

schätzen. Ist jedoch der Erwartungswert unbekannt, so hat man diesen durch seinen Schätzwert $\bar{x}_n$ zu ersetzen. Gemäß der Minimaleigenschaft von $\bar{x}_n$ gilt

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \geq \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x}_n)^2 \ .$$

Damit schätzt man die Varianz "im Durchschnitt etwas zu kurz". Durch die Multiplikation mit dem Faktor $\frac{n}{n-1} > 1$ wird dieses Defizit behoben und man erhält den angegebenen Schätzwert

$$\frac{n}{n-1} \times \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x}_n)^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x}_n)^2 \ .$$

Es lässt sich zeigen, dass der zugehörigen Schätzer S_n^2 *erwartungstreu* (unverfälscht, oder - englisch - *unbiased*) ist [12]. Dazu bedarf es freilich eines stochastischen Modells.

Begründung ohne Hilfsmittel der Wahrscheinlichkeitsrechnung: Die folgende Darstellung der Stichprobenvarianz, welche

$$s_n^2 = \frac{1}{\binom{n}{2}}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\frac{(x_j-x_i)^2}{2}$$

eine Möglichkeit bietet, diese mit Hilfe der wechselseitigen Abweichungsquadrate der Beobachtungswerte und damit ohne die Benützung des Stichprobenmittels zu definieren, begründet die Verwendung des Nenner $n-1$ ohne Hilfsmittel der Wahrscheinlichkeitsrechnung (vgl. auch [15]).

Die obige Darstellung lässt sich folgendermaßen einsehen. Durch Anwendung der bereits beim Nachweis des Steinerschen Verschiebungssatzes benützten Beziehung

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

[12] Der englische Statistiker *William Searly Gosset* $(1876-1937)$, der unter dem Pseudonym "*Student*" für die Brauerei Guiness arbeitete, verwendete aus diesem Grund anstelle des von *Karl Pearson* benützten Nenners n den Nenner $n-1$. Dies veranlasste Karl Pearson zu der Äußerung: "Only naughty brewers take n so small that the difference is not of the order of the probable error!"

auf die Größen $a = x_j - x_i$ und $b = x_j - \bar{x}_n$ lassen sich die wechselseitigen Abstandsquadrate $(x_i - x_j)^2$ mit Hilfe des Stichprobenmittels wie folgt darstellen

$$(x_j - x_i)^2 = (x_j - \bar{x}_n)^2 + (\bar{x}_n - x_i)^2 + 2(x_j - \bar{x}_n)(\bar{x}_n - x_i) .$$

Die Summe der Hälften dieser Größen ist, da die Beiträge für $j = i$ verschwinden und weil bekanntlich $\sum_{j=1}^n (x_j - \bar{x}_n) = 0$ ist, gleich

$$\begin{aligned}
\sum_{i=1}^n \sum_{j=i+1}^n \frac{(x_j - x_i)^2}{2} &= \frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n (x_j - x_i)^2 \\
&= \frac{1}{4} \left[\begin{array}{l} \sum_{i=1}^n \sum_{j=1}^n (x_j - \bar{x}_n)^2 + \sum_{i=1}^n \sum_{j=1}^n (\bar{x}_n - x_i)^2 + \\ + 2 \sum_{j=1}^n (x_j - \bar{x}_n) \sum_{i=1}^n (\bar{x}_n - x_i) \end{array} \right] \\
&= \frac{1}{4} \times 2n \sum_{j=1}^n (x_j - \bar{x}_n)^2 \\
&= \frac{n}{2} \sum_{i=1}^n (x_i - \bar{x}_n)^2 .
\end{aligned}$$

Da es $\binom{n}{2} = \frac{n(n-1)}{2}$ von 0 verschiedene wechselseitige Abstandsquadrate gibt, ist der betrachtete Durchschnittswert tatsächlich gleich der Stichprobenvarianz.

1.2.2 Stichprobenmedian und mittlere absolute Abweichung

a) **Der Stichprobenmedian** [13]

Zur Bestimmung des Stichprobenmedians ist es notwendig, die ursprünglichen Daten $x_1, ..., x_n$ der Größe nach zu ordnen. Seien also $x_{1:n}, ..., x_{n:n}$ die der Größe nach geordneten Daten (somit ist $x_{1:n} \leq x_{2:n} \leq ... \leq x_{n:n}$ und es gilt $\{x_{1:n}, ..., x_{n:n}\} = \{x_1, ..., x_n\}$). Dann ist

$$\tilde{x}_n = \begin{cases} x_{(n+1)/2:n} & \text{für } n \text{ ungerade} \\ (x_{n/2:n} + x_{n/2+1:n})/2 & \text{für } n \text{ gerade} \end{cases}$$

[13] Vom lateinischen Wort *medius 3: der, die, das Mittlere* abgeleitet

der *Median* der Daten. Dabei sind

$x_{(n+1)/2:n}$	der Wert in der Mitte der geordneten Datenliste,
$(x_{n/2:n} + x_{n/2+1:n})/2$	das arithmetische Mittel der beiden Werte in der Mitte.

Anmerkung 1: Sofern n ungerade ist, verwendet man für die Definition des Stichprobenmedians nur die Ordnungsrelation "$<$". Demnach ist die Definition des Stichprobenmedians in diesem Fall auch für ordinale Variable - wie z.B. für Schulnoten - möglich.

Anmerkung 2: Sei $\lceil x \rceil$ die kleinste ganze Zahl $\geq x$, d.h. sei $\lceil x \rceil = \min\{\, i \in \mathbb{Z} : i \geq x \,\}$. Mit Hilfe dieser Bezeichnung kann der Stichprobenmedian ohne Fallunterscheidung definiert werden:

$$\tilde{x}_n = \left(x_{\lceil n/2 \rceil:n} + x_{\lceil (n+1)/2 \rceil:n}\right)/2\,.$$

Anmerkung 3: Wird im Fall $n \geq 3$ ein Beobachtungswert $x_i = x_{n:n}$ durch einen extrem großen Wert ersetzt, so hat dies keinen Einfluss auf den Stichprobenmedian, hingegen einen beträchtlichen auf das Stichprobenmittel. Man sagt: Der Stichprobenmedian ist *robust* (*resistent*) gegen *Ausreißer*. Das Stichprobenmittel hingegen ist *sensitiv* gegen Ausreißer.

Die **Minimaleigenschaft des Stichprobenmedians** hinsichtlich der Summe der Absolutbeträge der Abweichungen

Behauptung: Die Funktion

$$g(x) = \frac{1}{n}\sum_{i=1}^{n} |x_i - x|$$

ist genau dann minimal, wenn $x \in \left[x_{\lceil n/2 \rceil:n}, x_{\lceil (n+1)/2 \rceil:n}\right]$ ist.

Beweis: Dieser erfolgt durch Fallunterscheidung. Eine ausführliche Darstellung findet sich in [17], Abschnitt 1.4.2.

b) **Die mittlere absolute Abweichung** [14] vom Stichprobenmedian ist

$$\tilde{s}_n = \frac{1}{n}\sum_{i=1}^{n} |x_i - \tilde{x}_n|\ .$$

[14] Der französische Mathematiker *Pierre Simon Laplace* (1749−1827) bevorzugte die mittlere absolute Abweichung, wogegen *Carl Friedrich Gauß* die Standardabweichung (unter der Bezeichnung *mittlerer zu befürchtenden Fehler*) vorzog.

Diese ist ein mit der Standardabweichung vergleichbares Maß für die Abweichung der Beobachtungswerte vom zugehörigen Zentralmaß.

c) Das einfachste Streumaß ist die **Spannweite der Stichprobe**, nämlich die Differenz

$$x_{n:n} - x_{1:n}$$

zwischen Stichprobenmaximum und Stichprobenminimum.

In der nachstehenden Definition betrachten wir eine Verallgemeinerung des Medians.

Definition: Sei $\alpha \in (0,1)$. Dann heißt

$$q_{\alpha,n} = \begin{cases} x_{\lceil \alpha \cdot n \rceil : n} & \text{für } \alpha \cdot n \notin \mathbb{N} \\ (1-\alpha) x_{\alpha \cdot n : n} + \alpha x_{\alpha \cdot n + 1 : n} & \text{für } \alpha \cdot n \in \mathbb{N} \end{cases}$$

das α-*Quantil* der Stichprobe.

Die Spezialfälle für $\alpha \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$, nämlich das *1. Qartil*, der Median und das *3. Quartil*, lassen sich folgendermaßen ohne Fallunterscheidung definieren

$$\begin{array}{lll} \text{das 1. Quartil} & q_{1/4,n} & = \left(3x_{\lceil n/4 \rceil : n} + x_{\lceil (n+1)/4 \rceil : n}\right)/4 \\ \text{der Median} & \tilde{x}_n = q_{1/2,n} & = \left(x_{\lceil n/2 \rceil : n} + x_{\lceil (n+1)/2 \rceil : n}\right)/2 \\ \text{das 3. Quartil} & q_{3/4,n} & = \left(x_{\lceil 3n/4 \rceil : n} + 3x_{\lceil (3n+1)/4 \rceil : n}\right)/4 \,. \end{array}$$

Zusammen mit dem Stichprobenminimum $x_{1:n}$ und dem Stichprobenmaximum $x_{n:n}$ benötigt man diese zum Anfertigen eines Kasten-Bilds [15].

Ein Streumaße, welches - im Unterschied zur Spannweite - robust gegen Ausreißer ist, ist der **Interquartilabstand**

$$q_{3/4,n} - q_{1/4,n} \,.$$

[15] Das *Kastenbild* (*box plot*) ist eine knappe visuelle Darstellung der Verteilung der Stichprobenwerte. Es stammt, wie das *Stängel-Blatt-Diagramm* aus Abschnitt 1.1.3, von *John W. Tukey.*

1.3 Weitere Mittelwerte

1.3.1 Das geometrische Mittel

Beispiel 1: Man bestimme die Wachstumsrate für das im Schaubild in *Küttings* Artikel [14] dargestellte Wirtschaftswachstum. Offensichtlich wächst die Wirtschaft im angegebenen Zeitraum von 6 Jahren mit dem Faktor

$$(1+0.044)\cdot(1+0.004)\cdot(1-0.02)\cdot(1+0.057)\cdot(1+0.026)\cdot(1+0.034) = 1.1519\,.$$

Um dieses Resultat bei einer für jedes Jahr gleichen Wachstumsrate r zu erreichen, muss gelten

$$(1+r)\cdot(1+r)\cdot(1+r)\cdot(1+r)\cdot(1+r)\cdot(1+r) = (1+r)^6 = 1.1519$$

und somit

$$1+r = \sqrt[6]{1.1519} = 1+0.0238\,.$$

Anmerkung 1: Es gilt

$$1+\frac{0.044+0.004-0.02+0.057+0.026+0.034}{6} = 1+0.143 > 1+0.0238\,.$$

Anders ausgedrückt: Bei einem durchschnittlichen Prozentsatz von 14.3 % würde die Wirtschaft gemäß $1.143^6 = 2.2299$ mit dem Faktor 2.2299 wachsen!

Definition: Seien $x_1,...,x_n \geq 0$. Dann nennt man

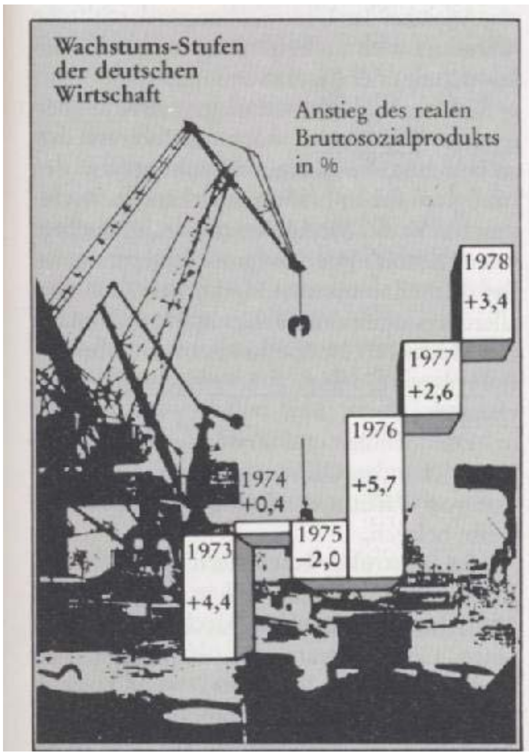
$$\hat{x}_n = \sqrt[n]{x_1\cdot...\cdot x_n}$$

das *geometrische Mittel* der Zahlen $x_1,...,x_n$.

Bezeichnung: Für $x_i = 1+r_i$, $i \in \{1,...,n\}$ nennt man die Größe

$$\hat{x}_n - 1 = \sqrt[n]{(1+r_1)\cdot...\cdot(1+r_n)} - 1$$

die *Wachstumsrate* (*growth rate*).



Anmerkung 2: Das geometrische Mittel eignet sich für Größen, die multiplikativ verknüpft werden. Typisch dafür ist der durchschnittliche *Aufzinsungsfaktor* in der Zinseszinsrechnung. Demgemäß wird es häufig als Durchschnittswert bei zeitlich aufeinanderfolgenden Veränderungsraten verwendet, also etwa bei jährlich erhobenen Kosten- und Preisindizes, den Quotienten der Börsenkurse von Ende zu Beginn eines Tages und den Wachstumsraten von Individuen und Populationen.

Behauptung 1: Es gilt

$$\sqrt[n]{\prod_{i=1}^{n} x_i} \leq \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}_n$$

mit Gleichheit genau dann, wenn gilt $x_1 = ... = x_n$.

Beweis: Ist eines der x_i gleich 0 , so ist die Gültigkeit der Ungleichung unmittelbar einsichtig. Im weiteren seien also die $x_1, ..., x_n > 0$ vorausgesetzt. Wir gehen vom folgenden **fundamentalen Sachverhalt** aus:

$\ln u \leq u - 1 \;\; \forall\, u \in (0, \infty)$ mit Gleichheit genau dann, wenn $u = 1$ ist.

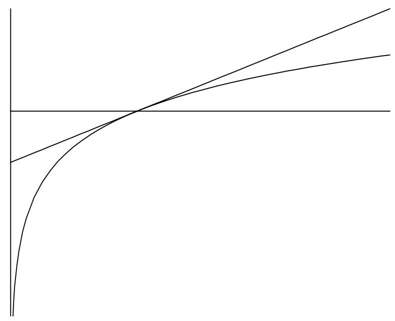


Abbildung der Funktion $u \mapsto \ln u$ und deren Tangente $u \mapsto u - 1$ im Punkt $(1, 0)$

Wendet man diesen Sachverhalt auf die Quotienten $\frac{x_i}{\bar{x}_n}$, $i \in \{1, ..., n\}$ an, summiert über alle i und dividiert schließlich durch n, so erhält man

$$\frac{1}{n} \sum_{i=1}^{n} \ln(\frac{x_i}{\bar{x}_n}) \leq \frac{1}{n} \sum_{i=1}^{n} (\frac{x_i}{\bar{x}_n} - 1) = \frac{1}{\bar{x}_n} \cdot \frac{1}{n} \sum_{i=1}^{n} x_i - 1 = \frac{\bar{x}_n}{\bar{x}_n} - 1 = 0\,,$$

wobei Gleichheit genau dann gilt, wenn alle Quotienten $\frac{x_i}{\bar{x}_n}$ gleich 1 sind, oder gleichbedeutend, wenn $x_1 = ... = x_n = \bar{x}_n$ ist.

Wegen der Additivität des Logarithmus ist die obige Ungleichung gleichbedeutend mit

$$\frac{1}{n} \sum_{i=1}^{n} \ln x_i \leq \ln \bar{x}_n\,.$$

Wendet man auf diese Ungleichung die Exponentialfunktion an und berücksichtigt die Monotonie derselben, so erhält man die Behauptung. $\square$

Der oben verwendete **fundamentale Sachverhalt** ist äquivalent mit folgender

Aussage: Die durch

$$g(u) = u - 1 - \ln u$$

auf dem Intervall $(0, \infty)$ definierte Funktion besitzt an der Stelle $u = 1$ ihr eindeutiges Minimum $g(1) = 0$. Diese Aussage ist aus der Betrachtung der Ableitung $g'(u) = 1 - \frac{1}{u}$ unmittelbar einsichtig. Denn letztere ist

$$1 - \frac{1}{u} = \begin{cases} < 0 & \text{für} \quad u < 1 \\ = 0 & \text{für} \quad u = 1 \\ > 0 & \text{für} \quad u > 1 \, . \end{cases}$$

1.3.2 Das harmonische Mittel

Beispiel 2: Die ersten 100 km legt ein Zug mit einer konstanten Geschwindigkeit von 70 km/h zurück, die zweiten 100 km mit einer konstanten Geschwindigkeit von 110 km/h. Mit welcher Durchschnittsgeschwindigkeit fährt der Zug?

Bezeichnen $s_1, s_2; v_1, v_2; t_1, t_2$ Weg, Geschwindigkeit und Zeit der jeweiligen Teilstrecken, und $\tilde{v}$ die gesuchte Durchschnittsgeschwindigkeit, dann gilt wegen der Beziehungen $s = \tilde{v} \cdot t$ (Weg = Geschwindigkeit×Zeit), $s = s_1 + s_2$ und $t = t_1 + t_2$

$$\begin{aligned} \tilde{v} &= \frac{s}{t} = \frac{s_1 + s_2}{t_1 + t_2} = \frac{s_1 + s_2}{\frac{s_1}{v_1} + \frac{s_2}{v_2}} \\ &= \frac{1}{\frac{s_1}{s_1+s_2} \cdot \frac{1}{v_1} + \frac{s_2}{s_1+s_2} \cdot \frac{1}{v_2}} \, . \end{aligned}$$

Die numerische Lösung ist daher $\tilde{v} = \frac{1}{\frac{1}{2} \cdot \frac{1}{70} + \frac{1}{2} \cdot \frac{1}{110}} = 85,56$ km/h.

Anmerkung 1: Die durchschnittliche Geschwindigkeit wäre 90 km/h, das geometrische Mittel der beiden Geschwindigkeiten wäre 87, 75 km/h!

Definition: Seien $x_1, ..., x_n > 0$. Dann nennt man

$$\frac{1}{\frac{1}{n} \sum\limits_{i=1}^{n} \frac{1}{x_i}}$$

das *harmonische Mittel* der Zahlen $x_1, ..., x_n$.

Anmerkung 2: Das harmonische Mittel ist der Kehrwert des arithmetischen Mittels der Kehrwerte der x_i . Es eignet sich als Mittelwert von Größen, die indirekt proportional zu anderen Größen sind. Anwendungsmöglichkeiten sind z.B.: Absolutbetrag der Krümmung (indirekt proportional zum Krümmungsradius, $|\kappa(\rho)| = 1/\rho$), Frequenz (indirekt proportional zur Wellenlänge, $\nu(\lambda) = 1/\lambda$), Geschwindigkeit (indirekt proportional zur Zeit, $v(t) = \frac{s}{t}$), Dichte (indirekt proportional zum Volumen, $\rho(V) = \frac{m}{V}$), Stromstärke (indirekt proportional zum Widerstand, $I(R) = \frac{U}{R}$), Druck eines idealen Gases (indirekt proportional zum Volumen, $p(V) = \frac{RT}{V}$).

Behauptung 2: Es gilt

$$\frac{1}{\frac{1}{n}\sum\limits_{i=1}^{n}\frac{1}{x_i}} \leq \sqrt[n]{\prod_{i=1}^{n} x_i}$$

mit Gleichheit genau dann, wenn gilt $x_1 = ... = x_n$.

Beweis: Diese Ungleichung ist gleichbedeutend mit

$$\sqrt[n]{\prod_{i=1}^{n}\frac{1}{x_i}} \leq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}\,,$$

was identisch mit der Ungleichung zwischen dem geometrischen und dem arithmetischen Mittel der Kehrwerte $y_i = \frac{1}{x_i}$, $i \in \{1, ..., n\}$ ist. $\square$

Aus den Behauptungen 2 und 3 ergibt sich zusammenfassend

Behauptung 3: Für $x_i \in (0, \infty)$, $i \in \{1, ..., n\}$ gilt

$$\min(x_1, ..., x_n) \leq \frac{1}{\frac{1}{n}\sum\limits_{i=1}^{n}\frac{1}{x_i}} \leq \hat{x}_n \leq \bar{x}_n \leq \max(x_1, ..., x_n)$$

mit Gleichheit jeweils genau dann, wenn gilt $x_1 = ... = x_n$.

1.3.3 Der Spezialfall n = 2

Für $0 < x < y < \infty$ bezeichne $a = a(x, y)$, $g = g(x, y)$ und $h = h(x, y)$ das arithmetische, geometrische bzw. harmonische Mittel von x und y. Der nachstehenden Abbildung entnimmt man - vermittels der Ähnlichkeit jeweils zweier Dreiecke -

$$\frac{x}{g} \; (= \tan\alpha) \; = \; \frac{g}{y} \quad \text{und daher} \quad g^2 \; = \; x \cdot y \quad ... \quad (1_\alpha)$$

und

$$\frac{h}{g} \; (= \cos\beta) \; = \; \frac{g}{a} \quad \text{und daher} \quad g^2 \; = \; h \cdot a \quad ... \quad (1_\beta)$$

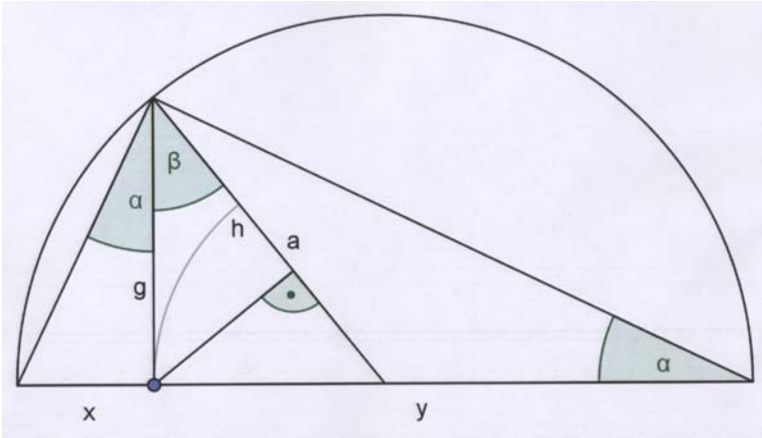


Abbildung: **a**rithmetisches, **g**eometrisches und **h**armonisches Mittel von x und y

Also gilt

$$h(x,y)\cdot a(x,y) = xy \tag{2}$$

und somit im Hinblick auf (1_α) $g(x,y)=\sqrt{xy}$ und wegen $a(x,y)=\frac{x+y}{2}$ im Hinblick auf (2)

$$h(x,y) = \frac{1}{\frac{1}{2}(\frac{1}{x}+\frac{1}{y})}\,.$$

Ferner entnimmt man der Abbildung die Gültigkeit der Ungleichungskette

$$x < h(x,y) < g(x,y) < a(x,y) < y\,. \tag{3}$$

Für den Fall $x=y$ gilt offensichtlich überall das Gleichheitszeichen.

Auf den beiden Beziehungen (2) und (3) beruht der sogenannte *babylonische Wurzelalgorithmus* zur numerischen Approximation der Quadratwurzel (siehe Aufgabe 13). Hinsichtlich detaillierter Ausführungen dazu sei auf die Diplomarbeit [22] von Frau *Reichenberger* verwiesen.

Anmerkung 3: Analog zu (1_β) erhält man auch

$$(a-x)^2 = (a-h)a\,.$$

Daraus folgt zusammen mit (1_β) der Pythagoräische Lehrsatz

$$g^2 + (a-x)^2 = a^2\,.$$

1.3.4 Ausblick: Zum Ursprung der Mittelwerte

Der Geschichtsschreiber *Jamblichos von Chalkis* (ca. 250-330 n.Chr.) berichtet, dass *Pythagoras* (Samos um 570 - Metapont um 496 v.Chr.) von einem Aufenthalt in Mesopotamien (dem Gebiet des heutigen Irak) die Kenntnisse der drei "*musikalischen Proportionen*" mitgebracht habe, welche schon ca. 2000 v.Chr. von den Babyloniern für ihren Quadratwurzelalgorithmus benutzt worden seien. *Pythagoras* gründete in Kroton, einer griechischen Kolonie in Süditalien, eine religiös-kultisch orientierte Lebensgemeinschaft, deren Mitglieder sich um die Erfüllung der asketischen, auf Reinhaltung der Seele (*katharsis*) abzielende Verhaltensregeln des Meisters bemühten.

Im Fragmente 2 der Harmonik des Pythagoräers *Archytas von Tarent* (428 - 365 v.Chr.), einem Freund von *Platon*, werden die drei behandelten Mittel erstmals explizit beschrieben. Und zwar vermittels der für die Griechen typischen Proportionen, d.h. durch Verhältnisse natürlicher Zahlen.

"Es gibt aber drei Proportionen in der Musik: einmal die arithmeti-sche, zweitens die geometrische, drittens die entgegengesetzte, sogenannte harmonische.

Die arithmetische, wenn drei Zahlbegriffe analog folgende Differenz aufweisen: um wieviel der erste den zweiten übertrifft, um soviel übertrifft der zweite den dritten. ... ."

"Die geometrische: wenn sich der erste Begriff zum zweiten, wie der zweite zum dritten verhält. Die größeren von ihnen haben das gleiche Verhältnis wie die geringeren."

"Die entgegengesetzte, sogenannte harmonische Proportion, wenn sich die Begriffe so verhalten: um den wievielten Teil der eigenen Größe der erste Begriff den zweiten übertrifft, um diesen Teil des dritten übertrifft der Mittelbegriff den zweiten. ... ."

Zusammenfassend werden das arithmetische, geometrische und harmonische Mittel durch die Proportionen

$$\begin{array}{lcl} a - b_1 = b_1 - c & \Leftrightarrow & b_1 = \frac{a+c}{2} \\ a : b_2 = b_2 : c & \Leftrightarrow & b_2 = \sqrt{a \cdot c} \\ (a - b_3) : (b_3 - c) = a : c & \Leftrightarrow & b_3 = \frac{1}{\frac{1}{2}\left(\frac{1}{a}+\frac{1}{c}\right)} \end{array}$$

definiert. Das arithmetische und das harmonische Mittel entsprechen einander im folgenden Sinn

$$a : b_1 = b_3 : c\,.$$

Hinsichtlich detaillierterer Ausführungen zur Musik sei auf Abschnitt 1.5.5 in [17] und die fächerübergreifende Diplomarbeit [21] von Frau *Fritz* verwiesen.

Kapitel 2

KOMBINATORISCHE GRUNDLAGEN

Jakob Bernoulli's im Jahre 1713 posthum erschienene Werk *Ars Conjectandi* [2] war grundlegend für die Entwickung der Stochastik [1]. In dessen zweiten Teil entwickelt er unter dem Titel *Permutations- und Combinationslehre* die Grundlagen der Wahrscheinlichkeitsrechnung.

Der Aufbau des die Wahrscheinlichkeitsrechnung beinhaltenden Teils unseres Seminars folgt Jakob Bernoulli's Zugang und vermeidet ganz bewusst den im geltenden Lehrplan der AHS vorgesehen Weg über bedingte Wahrscheinlichkeiten und Unabhängigkeit. Der letztgenannte Begriff ist zwar zentral aber schwierig und bedarf daher einer sorgfältigen Behandlung.

Die Anzahl aller oder bestimmter Ergebnisse eines Zufallsexperiments zu bestimmen, spielt in der Wahrscheinlichkeitsrechnung eine entscheidende Rolle. Die Kombinatorik liefert dazu die nötigen Hilfsmittel.

---

[1] Zum Begriff "Stochastik"

"Wenn jemand von den Fertigkeiten und Künsten die Rechenkunst, die Messkunst und die Kunst des Wägens wegnimmt, so bleibt, um es offen zu sagen, nur etwas übrig, was fast minderwertig ist [...]. Es bleibt nichts anderes übrig, als ein Errarten, ein Schließen durch Vergleichen und ein Schärfen der Sinneswahrnehmung durch Erfahrung und durch eine gewisse Übung, wobei man die - von vielen als Künste titulierten - Fähigkeiten des geschickten Vermutens (στωχαστική sc. τέχνη) benützt, die durch stete Handhabung und mühevolle Arbeit herangebildet werden."

Platon, Philebos

Ars Conjectandi (die Mutmaßungskunst) ist die lateiniscche Entsprechung des Wortes Stochastik.

Bezeichnung:
Die Anzahl der Elemente einer endlichen Menge M bezeichnen wir mit $|M|$.

2.1 Permutationen von n Elementen

Beispiel:
Auf wieviel verschiedene Arten können sich n Personen nebeneinander aufstellen?

Antwort:
Wir müssen n Personen auf n Plätze so verteilen, dass keine Person gleichzeitig zwei Plätze besetzt. Dem ersten Platz kann jede der n Personen zugewiesen werden. Ist dieser besetzt, stehen für den zweiten Platz nur noch $n-1$ Personen zur Verfügung. Da jede der n Möglichkeiten für den ersten Platz mit jeder der $n-1$ Möglichkeiten für den zweiten Platz kombiniert werden kann, gibt es für die Besetzung des ersten und zweiten Platzes $n \cdot (n-1)$ Möglichkeiten. Nach der Besetzung des n-ten Platzes beträgt daher die Anzahl der Möglichkeiten

$$n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 3 \cdot 2 \cdot 1 .$$

Für dieses Produkt ist in der Mathematik das Symbol $n!$ - gelesen "n-Fakultät"oder "n-Faktorielle"- gebräuchlich. Es gibt die Anzahl der Möglichkeiten an, n Elemente verschieden anzuordnen. Der mathematische Fachausdruck dafür ist "Permutationen von n Elementen". Als Präzisierung des Begriffs Permutation geben wir folgende

Formal gilt folgende rekursive

Definition: Sei $n \in \mathbb{N}_0$. Dann sind

$$0! := 1 \quad \text{und} \quad n! := n \cdot (n-1)! \, , \ n \in \mathbb{N} .$$

Die Frage nach der Anzahl der Möglichkeiten, n Elemente verschieden anzuordnen, entspricht der Frage, wie oft man n verschiedene Elemente vertauschen kann. Daher erklärt sich der Ausdruck Permutation, der vom lateinischen *permutare* (vertauschen) kommt.

Bermerkung:
Eine **Permutation** ist eine eineindeutige Abbildung einer Menge auf sich.

2.2 Anzahl der k-Tupel aus einer Menge mit n Elementen

Beispiel:
Berechne die Anzahl der vierstelligen Zahlen, die sich aus den Ziffern 1 bis 9 bilden lassen, wenn jede Ziffer auch mehrmals verwendet werden darf !

Lösung:
Für jede Stelle stehen 9 Ziffern zur Besetzung zur Vefügung. Da alle Besetzungen miteinander kombiniert werden können, ist die gefragte Anzahl gleich $9 \cdot 9 \cdot 9 \cdot 9 = 9^4$.
Allgemein handelt es sich hier um die Frage nach der Anzahl der Elemente des k-fachen kartesischen Produkts

$$M^k := \underbrace{M \times ... \times M}_{k-\mathrm{mal}} = \{(x_1, ..., x_k) : x_j \in M\,,\ 1 \le j \le k\}$$

einer Menge M. Entsprechend der Begründung im obigen Beispiel gilt für eine Menge M mit n Elementen:

$$\left|M^k\right| = n^k\,.$$

2.3 Anzahl der k-Tupel mit verschiedenen Elementen aus einer Menge mit n Elementen

Beispiel:
Berechne die Anzahl aller vierstelligen Zahlen, die sich mit den Ziffern 1 bis 9 bilden lassen, wenn in jeder dieser vierstelligen Zahlen keine Ziffer mehrfach auftritt !

Lösung:
Für die Tausenderstelle hat man 9 Ziffern zur Besetzung zur Verfügung; da keine der Ziffern mehrfach auftreten darf, bleiben für die Besetzung der Hunderterstelle nur noch 8 Ziffern übrig, wenn die Tausenderstelle bereits besetzt ist. Nach der Besetzung der Hunderterstelle bleiben für die Zehner- bzw. Einerstelle 7 bzw. 6 Ziffern. Da man alle Möglichkeiten miteinander kombinieren kann, ergibt sich die Gesamtanzahl aller derartigen vierstelligen Zahlen durch das Produkt $9 \cdot 8 \cdot 7 \cdot 6$.

Allgemein handelt es sich um die Frage nach der Anzahl der k-Tupel mit verschiedenen Elementen aus einer Menge mit n Elementen.
Es ergibt sich diese Anzahl durch das Produkt

$$n \cdot (n-1) \cdot \ldots \cdot (n-k+2) \cdot (n-k+1) \quad (k, n \in \mathbb{N},\ 1 \leq k \leq n)\,.$$

Unter Verwendung des in Abschnitt 2.1 eingeführten Symbols

$$n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$$

erhält man für $k, n \in \mathbb{N}$, $1 \leq k \leq n$,

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

2.4 Binomialkoeffizienten

Bezeichnung:
Zur Angabe der Anzahl der k-elementigen Teilmengen einer n-elementigen Menge ist in der Mathematik das Symbol $\binom{n}{k}$ - gelesen "n über k" - gebräuchlich. Die Anzahlen $\binom{n}{k}$ werden **Binomialkoeffizienten** genannt.
Um zu einer Formel für $\binom{n}{k}$ zu gelangen, betrachten wir folgendes

Einführungsbeispiel:
Gesucht ist die Anzahl der 3-elementigen Teilmengen einer Menge M mit 5 Elementen: $M = \{\square, \emptyset, \triangle, \Diamond, \nabla\}$.
Als Illustration dient folgende Tabelle:

$\square\emptyset\triangle$	$\square\triangle\emptyset$	$\emptyset\square\triangle$	$\emptyset\triangle\square$	$\triangle\square\emptyset$	$\triangle\emptyset\square$
$\square\emptyset\Diamond$	$\square\Diamond\emptyset$	$\emptyset\square\Diamond$	$\emptyset\Diamond\square$	$\Diamond\square\emptyset$	$\Diamond\emptyset\square$
$\square\emptyset\nabla$	$\square\nabla\emptyset$	$\emptyset\square\nabla$	$\emptyset\nabla\square$	$\nabla\square\emptyset$	$\nabla\emptyset\square$
$\square\triangle\Diamond$	$\square\Diamond\triangle$	$\triangle\square\Diamond$	$\triangle\Diamond\square$	$\Diamond\square\triangle$	$\Diamond\triangle\square$
$\square\triangle\nabla$	$\square\nabla\triangle$	$\triangle\square\nabla$	$\triangle\nabla\square$	$\nabla\square\triangle$	$\nabla\triangle\square$
$\square\Diamond\nabla$	$\square\nabla\Diamond$	$\nabla\square\Diamond$	$\nabla\Diamond\square$	$\Diamond\square\nabla$	$\Diamond\nabla\square$
$\emptyset\triangle\Diamond$	$\emptyset\Diamond\triangle$	$\triangle\emptyset\Diamond$	$\triangle\Diamond\emptyset$	$\Diamond\emptyset\triangle$	$\Diamond\triangle\emptyset$
$\emptyset\triangle\nabla$	$\emptyset\nabla\triangle$	$\triangle\emptyset\nabla$	$\triangle\nabla\emptyset$	$\nabla\emptyset\triangle$	$\nabla\triangle\emptyset$
$\emptyset\Diamond\nabla$	$\emptyset\triangle\Diamond$	$\Diamond\emptyset\triangle$	$\Diamond\triangle\emptyset$	$\nabla\emptyset\Diamond$	$\nabla\Diamond\emptyset$
$\triangle\Diamond\nabla$	$\triangle\nabla\Diamond$	$\Diamond\triangle\nabla$	$\Diamond\nabla\triangle$	$\nabla\triangle\Diamond$	$\nabla\Diamond\triangle$

In der ersten Spalte stehen die verschiedenen Teilmengen mit 3 Elementen aus der Menge M. In einer Zeile stehen jeweils die Möglichkeiten, eine Menge verschieden anzuordnen, das sind $3!$.

Man kann die Bildung der Elemente der Tabelle interpretieren als 3-maliges Ziehen ohne Zurücklegen aus der Menge M.

Für die Wahl des ersten Elements hat man 5 Möglichkeiten, für die Wahl des zweiten 4 und für die Wahl des dritten 3 Möglichkeiten. Insgesamt ergibt sich also für die Anzahl der Elemente der Tabelle $5 \cdot 4 \cdot 3$ (Anzahl der Tripel mit verschiedenen Elementen aus einer Menge mit 5 Elementen).
Man kann nun die Gesamtanzahl der Elemente der Tabelle auch dadurch erhalten, dass man das Produkt aus der Anzahl der Spalten und der Anzahl der Zeilen bildet. Dies führt auf die Beziehung

$$3!\binom{5}{3} = 5 \cdot 4 \cdot 3\,.$$

Wir haben damit gefunden, dass

$$\binom{5}{3} = \frac{5 \cdot 4 \cdot 3}{3!} = \frac{5!}{3! \cdot 2!}\,.$$

Verallgemeinert man diesen Sachverhalt, erhält man folgende

Behauptung:
Für $n \in \mathbb{N}_0$ und $k \in \{0, 1, ..., n\}$ gelten:

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

Formal gelten folgende

Gleichwertige Definitionen: Sei $n \in N_0$. Dann ist

$$\binom{n}{k} := \begin{cases} \frac{n!}{k!(n-k)!} & \text{für } \ k \in \{0, ..., n\} \\ 0 & \text{für } \ k \geq n+1 \end{cases} \quad ... \ (0_1)$$

oder, gleichbedeutend,

$$\binom{n}{k} := \frac{n \cdot ... \cdot (n-k+1)}{k!} \quad ... \ (0_2)\,.$$

Bemerkung:

· Für die Definition (0_2) erübrigt sich die Fallunterscheidung, da die rechte Seite für $k \geq n+1$ "automatisch" den Wert 0 hat (weil der $n+1$-te Faktor $= 0$ ist).

· Sie hat gegenüber (0_1) zudem den Vorteil , dass sie nur $2k-1$ Rechenoperationen erfordert, während die Definition (0_1) $2n-1$ Rechenoperationen benötigt $(n, k \in \mathbb{N})$.

Wir notieren einige **grundlegende Eigenschaften** der Binomialkoeffizienten.

$$\binom{n}{0} = \binom{n}{n} = 1 \tag{1}$$

Das ist anschaulich klar, da es unter den Teilmengen einer Menge genau einmal die leere Menge gibt und genau einmal die Menge selbst.

$$\binom{n}{k} = \binom{n}{n-k} \tag{2}$$

Die Formel besagt, dass die Anzahl der k-elementigen Teilmengen einer Menge mit n Elementen gleich der Anzahl der $(n-k)$-elementigen Teilmengen dieser Menge ist. Wir können dies folgend begründen: Jede Bildung einer k-elementigen Teilmenge aus einer Menge mit n Elementen hat eine $(n-k)$-elementige Teilmenge von M als Komplement und umgekehrt. Man kann die Formel natürlich auch rechnerisch bestätigen:

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

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \quad (1 \leq k \leq n-1) \tag{3}$$

Inhaltlich kann man diese Formel folgendermaßen begründen. Ist M eine n-elementige Menge mit den Elementen $a_1, a_2, ..., a_n$, so gibt $\binom{n-1}{k-1}$ die Anzahl der k-elementigen Teilmengen, die a_n enthalten, und $\binom{n-1}{k}$ die Anzahl der k-elementigen Teilmengen, die a_n nicht enthalten, an. Die Anzahl der k-elementigen Teilmengen von M ergibt sich daher durch die Summe: $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

Rechnerische Bestätigung der Formel:

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

Eine Illustration für diese Rekursionsformel bietet das sog.

Pascalsche Dreieck

$$
\begin{array}{llccccccccccccccl}
\text{Zeilen:} & & & & & & & & & \text{0.} & & & & & & & \text{Spalten:} \\
 & & & & & & & & \nearrow & & \text{1.} & & & & & & \\
\text{0.} & \leftarrow & & & & & & 1 & & \nearrow & & \text{2.} & & & & & \\
\text{1.} & \leftarrow & & & & & 1 & & 1 & & \nearrow & & \text{3.} & & & & \\
\text{2.} & \leftarrow & & & & 1 & & 2 & & 1 & & \nearrow & & \text{4.} & & & \\
\text{3.} & \leftarrow & & & 1 & & 3 & & 3 & & 1 & & \nearrow & & \text{5.} & & \\
\text{4.} & \leftarrow & & 1 & & 4 & & 6 & & 4 & & 1 & & \nearrow & & & \\
\text{5.} & \leftarrow & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & & &
\end{array}
$$

Das Dreieck entsteht nach folgender Gesetzmäßigkeit: Numeriert man die Zeilen und Spalten wie angegeben, so ergibt sich die Zahl, die am Kreuzungspunkt der n-ten Zeile und k-ten Spalte steht, durch die Summe der beiden darüberstehenden Zahlen, sofern $1 \leq k \leq n-1$ ist. Jede Zeile beginnt und endet mit einer Eins.

Aufgrund der Rekursionsformel 3 steht am Kreuzungspunkt der n-ten Zeile und k-ten Spalte die Anzahl $\binom{n}{k}$, die sich aus den beiden darüberliegenden zusammensetzt, d.h. aus der Anzahl am Kreuzungspunkt der $(n-1)$-ten Zeile und $(k-1)$-ten Spalte und der Anzahl am Kreuzungspunkt der $(n-1)$-ten Zeile und k-ten Spalte. Die Einsen an den seitlichen Rändern des Dreiecks erklären sich aus den Bedingungen $\binom{n}{0} = \binom{n}{n} = 1$.

Im folgenden wollen wir uns überlegen, auf wieviel verschiedenen Pfaden man von der Spitze des Pascalschen Dreiecks zum Kreuzungspunkt der 5. Zeile und 3. Spalte kommt:

								0.						
Zeilen							↗		1.					Spalten
0.						**1**		↗		2.				
1.					1		1		↗		3.			
2.				1		2		1		↗		4.		
3.			1		3		3		1		↗		5.	
4.		1		4		6		4		1		↗		
5.	1		5		10		**10**		5		1			

Teilt man den Pfad in Schritte, die entweder nach links oder rechts führen, so sind z.B. folgende Pfade möglich:

links	links	rechts	rechts	rechts
links	rechts	rechts	rechts	links
links	rechts	links	rechts	rechts

Man sieht, dass sich ein Pfad immer in 5 Schritten bewältigen lässt, wenn man nur "Abwärtsbewegungen" zulässt, und dass die Anzahl der Schritte nach rechts in jedem Pfad gleich ist. An welchen Stellen die Schritte nach rechts gemacht werden, ändert sich von Pfad zu Pfad.
Wir interessieren uns daher für die Anzahl der Pfade der Schrittlänge 5 mit genau 3 Schritten nach rechts.
Wählen wir nun aus der Menge

$$M = \{1.\text{ Schritt}, 2.\text{ Schritt}, 3.\text{ Schritt}, 4.\text{ Schritt}, 5.\text{ Schritt}\}$$

eine Teilmenge mit 3 Elementen für die 3 Schritte nach rechts aus, z.B.

$$\{2.\text{ Schritt}, 3\text{ Schritt}, 5.\text{ Schritt}\}\ ,$$

so teilen wir den Schritten nach rechts eindeutig die Plätze zu. Wie wir bereits wissen, gibt es $\binom{5}{3}$ Teilmengen mit 3 Elementen aus einer 5-elementigen Menge. $\binom{5}{3}$ ist daher die Anzahl der Pfade, die zum Punkt in der 5. Zeile und 3. Spalte führen.

Indem wir nun links mit "0" und rechts mit "1" bezeichnen, erhalten wir die folgende für uns wichtige Interpretation der Binomialkoeffizienten.

Weitere Interpretation:
$\binom{n}{k}$ **ist die Anzahl der** $(0,1)$**-Folgen der Länge** n **mit genau** k **Einsen**.

Bemerkung:
Den Namen haben die Binomialkoeffizienten vom **Binomischen Lehrsatz**. Dieser lautet:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k} \quad (a,b \in \mathbb{R}\,;\, n = 0,1,2,...)$$

Das heißt: Potenziert man ein *Binom*, so geben die Zahlen $\binom{n}{k}$ die Koeffizienten der fallenden Potenzen von a bzw. steigenden Potenzen von b an.

Zusammenfassend ergibt sich folgende Übersicht:

Stichprobe	Beachtung der Reihenfolge	Nichtbeachtung der Reihenfolge
ohne Zurücklegen	$n \cdot ... \cdot (n-k+1) = \frac{n!}{(n-k)!} = \binom{n}{k} \cdot k!$	$\binom{n}{k}$
mit Zurücklegen	n^k	$\binom{n+k-1}{n}$ ... (4)

Bemerkung: Der Fall (4), dessen Motivation methodisch am anspruchsvollsten ist, wird hier nicht behandelt. Außerdem verzichten wir bewusst auf die Bezeichnungen "Variationen ohne/mit Wiederholung" und "Kombinationen ohne/mit Wiederholung".

Kapitel 3

WAHRSCHEINLICHKEITS-THEORETISCHE GRUNDLAGEN

3.1 Grundraum und Ereignisse

Vor der Erklärung dieser Begriffe wollen wir die Frage stellen: Was ist ein Zufallsexperiment? Dies kann man am besten anhand von Beispielen beantworten. Die einfachsten kommen aus dem Glücksspielbereich: Roulette, Münzwurf, Würfelwurf. Wir beschränken uns hier auf Zufallsexperimente mit höchstens abzählbar vielen möglichen Ausfällen

Definition:
Die Menge $\Omega = \{\omega_1, \omega_2, ...\}$ der möglichen (berücksichtigten) Ausfälle oder Ergebnisse eines Zufallsexperiments heißt **Grundmenge** oder **Grundraum**. Die Teilmengen von Ω heißen **Ereignisse**. Sie werden mit Großbuchstaben bezeichnet: $E, A, B, ...$.
Beim Roulette ist der Grundraum die Menge $\Omega = \{0, 1, 2, ..., 36\}$. Beim einmaligen Münzwurf ist der Grundraum die Menge $\Omega = \{1, 0\}$, wenn man die "Zahlseite" mit 1 und die "Wappenseite" mit 0 bezeichnet. Beim Werfen eines Würfels ist der Grundraum die Menge $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Beispiele für Ereignisse sind beim Roulette die Menge der ungeraden Zahlen $A = \{1, ..., 35\}$ oder die Zahlen der ersten Kolonne $B = \{1, 4, 7, 10, 13, ..., 34\}$.
Ein Beispiel für ein Ereignis beim Werfen eines Würfels ist das Werfen einer

geraden Augenzahl: $A = \{2, 4, 6\}$.

3.2 Wahrscheinlichkeit

Nach der Erklärung von Ausfall, Grundraum und Ereignis wollen wir uns mit der Frage beschäftigen, was wir unter der "Wahrscheinlichkeit eines Ereignisses" verstehen.
Wir führen ein Zufallsexperiment wiederholt unter stets gleichen Bedingungen durch und notieren uns die relativen Häufigkeiten des Ausfalls ω

$$h_n(\omega) = \frac{\text{Anzahl der Versuche mit Ausfall } \omega}{n} \quad (n = 1, 2, 3, ...) \,.$$

Dabei machen wir folgende Beobachtung: Bei wachsendem n verhalten sich die Zahlen $h_n(\omega)$ so, als würden sie sich auf eine feste Zahl $p(\omega)$ einpendeln, d.h. es is $h_n(\omega) \cong p(\omega)$ für große n. Diese Erfahrungstatsache bezeichnet man als das **empirische Gesetz der großen Zahlen**. Wir ordnen nun in diesem Sinne jedem $\omega \in \Omega$ eine Zahl $p(\omega)$ zu und bezeichnen sie als "Wahrscheinlichkeit des Eintretens von ω". Diese Zahlen erhält man in der Praxis durch Schätzungen, Annahmen, sowie aus der Kombination von Schätzungen und Annahmen.

Wir stellen an die Zahlen $p(\omega)$, $\omega \in \Omega$, folgende Forderungen, die sich aus den Eigenschaften der relativen Häufigkeiten ergeben.

$$p(\omega) \geq 0 \,, \ \forall \ \omega \in \Omega \tag{i}$$

$$\sum_{\omega \in \Omega} p(\omega) = 1 \,. \tag{ii}$$

Ist dies durchgeführt, dann werden beliebigen Ereignissen Wahrscheinlichkeiten aufgrund folgender Überlegung zugeordnet: $A \subseteq \Omega$ stelle ein beliebiges Ereignis dar. Die relative Häufigkeit von A wird bestimmt durch

$$h_n(A) = \frac{\text{Anzahl der Versuche, in denen } A \text{ eintritt}}{n} \quad (n = 1, 2, 3, ...) \,.$$

Da $h_n(A) = \sum_{\omega \in \Omega} h_n(\omega)$ ist, wird für große n gelten:

$$h_n(A) \cong \sum_{\omega \in A} p(\omega) \,.$$

Folgende Festlegung ist daher sinnvoll:

$$P(A) := \sum_{\omega \in A} p(\omega) \, .$$

Man sagt, das Ereignis A tritt mit Wahrscheinlichkeit $P(A)$ ein. Auf diesem Hintergrund treffen wir nun folgende

Definition:
Ω sei die zu einem Zufallsexperiment gehörige Grundmenge.

Man nennt $(p(\omega) : \omega \in \Omega)$ eine **Wahrscheinlichkeitsverteilung** auf Ω, wenn

$$p(\omega) \geq 0 \, , \; \forall \; \omega \in \Omega \quad \text{und} \quad \sum_{\omega \in \Omega} p(\omega) = 1 \, .$$

Ist eine Wahrscheinlichkeitsverteilung $(p(\omega) : \omega \in \Omega)$ gegeben, dann definiert man die **Wahrscheinlichkeit eines beliebigen Ereignisses** A durch

$$P(A) := \sum_{\omega \in A} p(\omega) \, .$$

3.3 Laplace - Experimente

Um Wahrscheinlichkeiten festzulegen, muß man nicht immer den Weg über die relativen Häufigkeiten wählen. Es gibt Experimente, bei denen man von vornherein annehmen wird, daß die möglichen Ausfälle bei vielen Durchführungen ungefähr gleich oft auftreten werden. Ein Beispiel dafür ist ein sorgfältig gearbeiteter Würfel. Jeder Augenzahl wird man in diesem Fall aus Symmetriegründen die Wahrscheinlichkeit $p = \frac{1}{6}$ zuordnen.

Definition:
Gegeben sei ein Zufallsexperiment mit endlicher Grundmenge $\Omega = \{\omega_1, \omega_2, ..., \omega_N\}$. Falls die Annahme getroffen wird, daß $p(\omega_1) = p(\omega_2) = ... = p(\omega_N) = \frac{1}{N}$, sprechen wir von einem **Laplace - Experiment**. Die Wahrscheinlichkeit eines Ereignisses A ist somit gegeben durch

$$P(A) = \frac{|A|}{|\Omega|} \, ,$$

bzw. in der klassischen Terminologie

$$P(A) = \frac{\text{Anzahl der für } A \text{ günstigen Fälle}}{\text{Anzahl der möglichen Fälle}} \, .$$

Die Wahrscheinlichkeitsverteilung $(p(\omega_i), 1 \leq i \leq N)$ mit $p(\omega_i) = \frac{1}{N}$ $(1 \leq i \leq N)$ heißt **Gleichverteilung** auf Ω.

Bemerkung:
Eine Münze, von der man annimmt, daß beide Seiten mit Wahrscheinlichkeit $\frac{1}{2}$ fallen, nennt man eine faire Münze.
Ein Würfel, von dem man annimmt, daß die Wahrscheinlichkeiten für die Augenzahlen gleich sind, heißt symmetrischer Würfel, idealer Würfel, manchmal auch Laplace - Würfel.

Beispiel:
Ein symmetrischer Würfel wird einmal geworfen. Wie groß ist die Wahrscheinlichkeit für das Eintreten einer geraden Augenzahl?

Antwort:
Der Grundraum ist die Menge $\Omega = \{1, 2, 3, 4, 5, 6\}$. Wir betrachten das Ereignis "Eintreten einer geraden Augenzahl". Dieses ist repräsentiert durch die Menge $A = \{2, 4, 6\}$. Unter der Annahme, daß es sich um einen symmetrischen Würfel handelt, sind die Wahrscheinlichkeiten aller Ausfälle gleich, d.h.

$$p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6}.$$

Es gilt

$$P(A) = \sum_{\omega \in A} p(\omega) = p(2) + p(4) + p(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6},$$

also

$$P(A) = \frac{3}{6} = \frac{\text{Anzahl der für } A \text{ günstigen Fälle}}{\text{Anzahl der möglichen Fälle}}.$$

3.4 Zufallsvariable, Erwartungswert und Varianz

Beispiel:
Beim Werfen zweier Würfel betrachte man die auftretende Augensumme. Die Ausfälle und die dazugehörigen Augensummenwerte sind in der folgenden

Tabelle angegeben.

Mögliche Wurfkombinationen						mögliche Augensumme
(1/1)						2
(1/2)	(2/1)					3
(1/3)	(2/2)	(3/1)				4
(1/4)	(2/3)	(3/2)	(4/1)			5
(1/5)	(2/4)	(3/3)	(4/2)	(5/1)		6
(1/6)	(2/5)	(3/4)	(4/3)	(5/2)	(6/1)	7
(2/6)	(3/5)	(4/4)	(5/3)	(6/2)		8
(3/6)	(4/5)	(5/4)	(6/3)			9
(4/6)	(5/5)	(6/4)				10
(5/6)	(6/5)					11
(6/6)						12

In diesem Beispiel haben wir ein bestimmtes Merkmal, das wir Variable nennen, untersucht. Diese Variable kann bestimmte Werte annehmen.

Variable X	Wertebereich W_X
Augensumme	$2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$

Jedem Ausfall des Zufallsversuchs wird ein bestimmter Variablenwert zugeordnet. In diesem Beispiel wird etwa dem Ausfall (5/1) der Augensummenwert (Variablenwert) 6, dem Ausfall (3/2) der Augensummenwert 5 zugeordnet. Welchen Wert die betrachtete Variable in der Versuchsdurchführung jeweils annehmen wird, hängt vom Zufall ab; deshalb bezeichnet man Variable solcher Art als Zufallsvariable.

Definition:
Eine **Zufallsvariable** ist eine Funktion, die jedem Ausfall eines Zufallsexperiments eine reelle Zahl zuordnet.

Als Bezeichnung für Zufallsvariable sind Großbuchstaben, meistens X, Y, S, $T, ...$ üblich; die Werte, die die Zufallsvariablen annehmen, werden oft mit den entsprechenden Kleinbuchstaben x_j, h_j, s_j, t_j, ..., $j \in I$ (I Indexmenge) bezeichnet.

3.4.1 Die Verteilung einer Zufallsvariablen

Es sei X eine Zufallsvariable mit Wertebereich $W_X = \{x_1, x_2, x_3, ...\}$. Uns interessiert nun die Frage, mit welcher Wahrscheinlichkeit die Zufallsvariable

X den Wert x_j annimmt. Für das Ereignis $\{\omega \in \Omega : X(\omega) = x_j\}$ verwendet man dabei die Schreibweise $\{X = x_j\}$, die zugehörige Wahrscheinlichkeit bezeichnet man kurz mit $P(X = x_j)$.

Wir betrachten nocheinmal die möglichen Augensummen beim Werfen zweier Würfel, die wir nun symmetrisch voraussetzen. Der Grundraum $\Omega = \{\omega = (\omega_1, \omega_2) : \omega_i \in \{1,2,3,4,5,6\},\ i = 1,2\} = \{1,2,3,4,5,6\}^2$, mit $|\Omega| = 6^2$, wird durch die Frage nach den Augensummen in Ereignisse $E_2, ..., E_{12}$ zerlegt. Das Eintreten des Ereignisses E_k bedeute das Auftreten der Augensumme k, $(k = 2, ..., 12)$. Die Wahrscheinlichkeit $P(E_k)$ kann man unter der Laplace-Annahme auf Ω,die aufgrund der Symmetrie der Würfel adäquat ist, folgendermaßen berechnen:

$$P(E_k) = \frac{\text{Anzahl der für } E_k \text{ günstigen Ausfälle}}{\text{Anzahl der möglichen Ausfälle}} .$$

Die folgende Tabelle listet diese Wahrscheinlichkeiten auf:

k	2	3	4	5	6	7	8	9	10	11	12
$P(E_k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Betrachtet man nun die Augensumme als Zufallsvariable X, so hat man in der für Zufallsvariable üblichen Schreibweise:

$$P(E_k) = P(X = k), \quad k \in \{2, ..., 12\} .$$

Durch diese Überlegungen kommen wir zu folgender

Definition:
$X : \Omega \mapsto R$ sei eine Zufallsvariable mit den verschiedenen Werten $x_1, x_2, ...$. Dann heißt die Funktion

$$x_j \longmapsto P(X = x_j) \quad (j = 1, 2, ...)$$

die **Verteilung der Zufallsvariablen** X .

3.4.2 Erwartungswert

In diesem Abschnitt wollen wir uns mit dem in der Wahrscheinlichkeitsrechnung sehr wichtigen Begriff des Erwartungswerts beschäftigen. Folgendes Beispiel dient der Motivation.

Beispiel (Chuck - a - luck):
Chuck-a-luck ist einWürfelspiel, bei dem man den Einsatz e auf eine Zahl $z \in \{1,2,3,4,5,6\}$ setzt. Dann werden 3 Würfel geworfen. Kommt z auf keinem der Würfel vor, ist der Einsatz verloren. Kommt z genau k-mal vor ($k \in \{1,2,3\}$), gewinnt man den Betrag $k\,e$. Jemand spielt sehr oft. Wie groß wird sein mittlerer Gewinn pro Spiel sein?

Antwort:
Werden die drei Würfel n-mal geworfen, dann werden die Werte $-e, e, 2e, 3e$ mit bestimmten relativen Häufigkeiten auftreten. Wir bezeichnen diese mit $h_n(-e),\ h_n(e),\ h_n(2e),\ h_n(3e)$.
Der Mittelwert g_n aus den Ergebnissen aller Spiele ist dann

$$g_n = -e\,h_n(-e) + e\,h_n(e) + 2e\,h_n(2a) + 3e\,h_n(3e)\,.$$

Ist n groß, dann stimmen die relativen Häufigkeiten aufgrund des empirischen Gesetzes der großen Zahlen näherungsweise mit den entsprechenden Wahrscheinlichkeiten überein. Es wird daher für große n gelten:

$$g_n \cong -e\,P(X=-e) + e\,P(X=e) + 2e\,P(X=2e) + 3e\,P(X=3e) := E(X).$$

Diesen Wert werden wir als "Erwartungswert von X" bezeichnen.

Wir berechnen die Wahrscheinlichkeiten unter der Voraussetzung, daß die Würfel symmetrisch sind. Diese führt zur Laplace-Annahme auf $\Omega = \{1,...,6\}^3$.
Die Anzahl der möglichen Fälle ist somit 6^3.

1) $P(X=-e) = P(\text{Kein Würfel zeigt die Zahl } z)$:
Die Anzahl der günstigen Fälle für das Ereignis "Kein Würfel zeigt die Zahl z" entspricht der Anzahl der Tripel in der Menge M^3 mit $M = \{1,2,3,4,5,6\}\backslash\{z\}$.
Das sind $5^3 = 125$. Es ist daher $P(X=-e) = \frac{5^3}{6^3} = \frac{125}{216}$.

2) $P(X=e) = P(\text{Genau ein Würfel zeigt die Zahl } z)$:
Die Anzahl der für das betrachtete Ereignis günstigen Fälle ergibt sich folgend. Für zwei Würfel hat man je 5 Möglichkeiten; mit dem verbleibenden Würfel muß die Zahl z geworfen werden. Da jeder der drei Würfel der "verbleibende" sein kann, ergibt sich für die gesuchte Anzahl : $3\cdot5\cdot5\cdot1 = 75$.
Daraus folgt: $P(X=e) = \frac{3\cdot5\cdot5}{6^3} = \frac{75}{216}$.

3) $P(X=2e) = P(\text{Genau 2 Würfel zeigen die Zahl } z)$:
In den günstigen Fällen hat man für zwei Würfel je eine Möglichkeit, für den verbleibenden 5 Möglichkeiten. Da wiederum jeder der drei Würfel der

"verbleibende" sein kann, ergibt sich für die Anzahl der günstigen Fälle: $3 \cdot 5 \cdot 1 \cdot 1 = 15$. Daraus folgt: $P(X = 2e) = \frac{15}{216}$.

4) $P(X = 3e) = P(\text{Alle drei Würfel zeigen die Zahl } z)$:
Es gibt nur einen günstigen Fall. Daher ist $P(X = 3e) = \frac{1}{216}$.

Für den Mittelwert g_n folgt daher:

$$g_n \cong -e\,\frac{125}{216} + e\,\frac{75}{216} + 2e\,\frac{15}{216} + 3e\,\frac{1}{216} = -\frac{17}{216}\,e\,.$$

Aus dem errechneten Mittelwert kann man schließen, daß ein Spieler pro Spiel einen geringfügigen Prozentsatz des Einsatzes verliert. Der Verlust pro Spiel ist so klein, dass die Benachteiligung erst nach sehr vielen Spielen auffallen wird. Spiele mit geringfügig negativen Gewinnerwartungen anzubieten, gehört zur Strategie der Casinobetreiber. Ein Spiel ist fair, wenn der Erwartungswert gleich 0 ist.

Diese Überlegungen führen uns zu folgender Definition, wobei wir uns auf den Fall endlicher Wertebereiche beschränken.

Definition:
Es sei $X : \Omega \mapsto \mathbb{R}$ eine Zufallsvariable mit dem Wertebereich $W_X = \{x_1, ..., x_m\}$ und der Verteilung $x_j \mapsto P(X = x_j),\ j \in \{1, ..., m\}$.
Dann heißt die Größe

$$E(X) = \sum_{j=1}^{m} x_j\, P(X = x_j)$$

der **Erwartungswert** oder **Mittelwert** von X.

Der Erwartungswert einer Zufallsvariablen X wird oft mit μ (genauer μ_X) bezeichnet.
Anmerkung: Zur Motivation mit Hilfe der Darstellung des Stichprobenmittels siehe Anmerkung 2 des Abschnittes 1.2.1

Erwartungswert von Funktionen von Zufallsvariablen

Es seien X eine Zufallsvariable und f eine reellwertige Funktion auf W_X. Um den Erwartungswert der Zufallsvariablen $Y := f(X)$ zu berechnen, ist es nicht notwendig, die Wahrscheinlichkeitsverteilung von Y zu kennen.

Es gilt nämlich folgende

Behauptung: Besitzt die Zufallsvariable X den Wertebereich $\{x_1, ..., x_m\}$ und die Wahrscheinlichkeitsverteilung $x_j \mapsto P(X = x_j)$, $1 \leq j \leq m$, so gilt für den Erwartungswert der Zufallsvariablen $f(X)$

$$E(f(X)) = \sum_{j=1}^{m} f(x_j)P(X = x_j)\,.$$

Wir verzichten im Rahmen dieser Fortbildung auf den Beweis.

3.4.3 Varianz

Die Kenntnis des Erwartungswerts der Zufallsvariablen X stellt eine wichtige Information über die zugehörige Verteilung dar. In statistischen Berechnungen ist es oft entscheidend zu wissen, wie die Werte der Zufallsvariablen um den Erwartungswert streuen. Man versucht daher, eine Maßzahl für die "mittlere Abweichung" der Werte der Zufallsvariablen vom Erwartungswert zu definieren. Überlegungen in den Abschnitten 1.2.1 und 1.2.2 legen folgende Definition nahe.

Definition:
Als **Varianz** ($V(X)$) der Zufallsgröße X definiert man den Erwartungswert der quadratischen Abweichungen von μ :

$$V(X) := E[(X - \mu)^2]\,.$$

Man kann die Varianz von X über die Verteilung von X berechnen, ohne die Verteilung von $(X - \mu)^2$ bestimmen zu müssen:

$$V(X) = \sum_{j=1}^{m} (x_j - \mu)^2 \cdot P(X = x_j)\,.$$

Als **Standardabweichung** der Zufallsgröße X bezeichnet man die Zahl

$$\sigma_X := \sqrt{V(X)} = \sqrt{E[(X - \mu)^2]}\,.$$

Nützliche Berechnungsmöglichkeiten für die Varianz sind

$$\begin{aligned} V(X) &= E(X^2) - \mu^2 \\ &= E[X(X-1)] - \mu(\mu - 1)\,. \end{aligned}$$

Bemerkung:
Wie bereits in Abschnitt 1.2.2 angemerkt, bevorzugte *Laplace* die mittlere absolute Abweichung, wogegen *Gauß* die Standardabweichung vorzog. Laplace's Abweichungsmaß - welches in moderner Gestalt die Form $E[|X-\mu|]$ hat - hätte übrigens den Vorteil, dass es die Dimension einer Länge besitzt, während die Varianz die Dimension einer Fläche hat und somit erst die Quadratwurzel der Varianz wieder die Dimension einer Länge besitzt. Da jeder möglichen Definition eines Abweichungsmaßes eine gewisse Willkür innewohnt, dauerte es eine beträchtliche Zeit, bis sich die Standardabweichung durchsetzte. Der Grund für Letzteres ist, dass "die Varianz der Summe stochastisch unabhängiger Zufallsvariablen die Summe der Varianzen ist". Dieser Sachverhalt wiederum beruht - wie der Satz von Pythagoras - auf dem Begriff der Orthogonalität. (In knapper Form: "Unabhängigkeit impliziert Orthogonalität.")

Kapitel 4

Die Binomialverteilung

Die Binomialverteilung ist eine der wichtigsten Wahrscheinlichkeitsverteilungen.
Typisch für ihr Auftreten ist folgende Situation:

Ein Zufallsexperiment, bei dem genau 2 einander ausschließende Ausfälle A und A^c eintreten können, werde n-mal durchgeführt. Der Ausfall A tritt jeweils mit der Wahrscheinlichkeit p ein, und die Versuchsausgänge beeinflussen sich gegenseitig nicht. A wird als Erfolg oder Treffer bezeichnet. Der Gegenausfall A^c tritt jeweils mit der Wahrscheinlichkeit $q = 1 - p$ auf und wird als Misserfolg oder Niete bezeichnet.
Wir betrachten nun die Zufallsvariable X, die angibt, wie oft unter den n Versuchen der Ausfall A eingetreten ist. Die Verteilung dieser Zufallsvariablen wird als Binomialverteilung bezeichnet

Beispiele für das Auftreten der Binomialverteilung sind

1. Bei einem Automaten gewinnt man bei einem Spiel mit der Wahrscheinlichkeit 0.7 . Wie groß ist die Wahrscheinlichkeit, bei 20 Spielen a) genau 12 mal, b) mindestens 12 mal, c) höchstens 18 mal, d) mindestens 12 mal und höchstens 18 mal zu gewinnen? (Nr. 5.11, S. 103 aus [50])

2. Eine Maschine stellt Stanzteile mit einem Ausschußanteil von 5% her. Wie groß ist die Wahrscheinlichkeit, daß 4 zufällig ausgewählte Teile ausnahmslos in Ordnung sind? (Nr. 2, S. 261 aus [51])

3. Zwei Schachspieler, von denen der eine den anderen erfahrungsgemäß mit der Wahrscheinlichkeit 0.6 schlägt, beschließen, 5 Spiele zu spielen.

Mit welcher Wahrscheinlichkeit gewinnt der bessere Spieler mehr als die Hälfte der Spiele? (Nr. 9.104, S. 270 aus [48])

Diese Beispiele lassen sich auf folgendes Urnenmodell zurückführen:

Eine Urne enthalte N Kugeln, davon r rote und w weiße, welche von ihrer Farbe abgesehen ununterscheidbar sind. Es werden n Kugeln mit Zurücklegen zufällig gezogen. Wie groß ist die Wahrscheinlichkeit, daß die Stichprobe genau k rote Kugeln enthält?

4.1 Einführung der Binomialverteilung mittels des Urnenmodells

Wir beziehen uns bei der Herleitung der Binomialverteilung direkt auf das Urnenmodell in Beispiel 4 und verwenden das Prinzip des Laplace-Experiments. Die explizite Verwendung des Begriffes der stochastischen Unabhängigkeit wird dabei vermieden (siehe dazu Abschnitt 2.6 in [18]).

Wir betrachten das Urnenbeispiel mit $r = 3,\ w = 7$ (also $N = 10$), $n = 5$ und $k = 2$.
Der Grundraum (die Menge aller möglichen Ausfälle) des Zufallsexperiments ist die Menge aller 5-Tupel aus einer Menge mit 10 Elementen. Das heißt:

$$\Omega = \{\, (\omega_1, \omega_2, \omega_3, \omega_4, \omega_5) : \omega_i \in U,\ 1 \leq i \leq 5\} \ ,$$

wobei U die Menge der Kugeln in der Urne und ω_i das Ergebnis des i-ten Zuges sind. Die Kugeln seien von 1 bis 10 folgendermaßen numeriert:

$$\underbrace{1 \quad 2 \quad 3}_{\text{rot}} \quad \underbrace{4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10}_{\text{weiß}}$$

Die zunächst noch vage Angabe, daß "zufällig" gezogen wird, präzisieren wir durch die **Laplace-Annahme:**
<u>Alle Elemente von Ω haben dieselbe Wahrscheinlichkeit einzutreten.</u>

E_2 sei das Ereignis, bei 5-maligem Ziehen mit Zurücklegen genau 2 rote Kugeln zu erhalten. Aufgrund der Laplace-Annahme können wir die Wahrscheinlichkeit $P(E_2)$ berechnen mittels der Formel

$$P(E_2) = \frac{\text{Anzahl der für } E_2 \text{ günstigen Fälle}}{\text{Anzahl der möglichen Fälle}} \ .$$

Die Anzahl der möglichen Fälle erhalten wir folgendermaßen. Da sich in der Urne beim 5-maligen Ziehen mit Zurücklegen vor jedem Zug alle Kugeln in der Urne befinden, hat man für den 1. bis 5. Zug je 10 Möglichkeiten, die man miteinander kombinieren kann. Es ergibt sich also für die gesuchte Anzahl 10^5 .
Zur Ermittlung der Anzahl der für E_2 günstigen Fälle bestimmen wir zuerst die Anzahl der Möglichkeiten, zwei Plätze für die roten Kugeln auszuwählen. Treten in einer Reihe von 5 Kugeln genau 2 rote ($\otimes$) und 3 weiße ($\bigcirc$) auf, so können diese folgendermaßen plaziert sein:

$$\begin{array}{ccc}
\otimes\otimes\bigcirc\bigcirc\bigcirc & \qquad & \bigcirc\otimes\bigcirc\otimes\bigcirc \\
\otimes\bigcirc\otimes\bigcirc\bigcirc & & \bigcirc\otimes\bigcirc\bigcirc\otimes \\
\otimes\bigcirc\bigcirc\otimes\bigcirc & & \bigcirc\bigcirc\otimes\otimes\bigcirc \\
\otimes\bigcirc\bigcirc\bigcirc\otimes & & \bigcirc\bigcirc\otimes\bigcirc\otimes \\
\bigcirc\otimes\otimes\bigcirc\bigcirc & & \bigcirc\bigcirc\bigcirc\otimes\otimes
\end{array}$$

Schreiben wir "1" für rot und "0" für weiß, so erhalten wir die Menge der (0,1)-Folgen der Länge 5 mit genau 2 Einsen. Die gesuchte Anzahl ist daher gleich $\binom{5}{2}$.
Im zweiten Schritt bestimmen wir nun die Anzahl der Möglichkeiten, die Plätze entsprechend den zugewiesenen Farben zu besetzen. Dabei benützen wir die Numerierung der Kugeln. Für die "roten Plätze" stehen uns jeweils die Kugeln mit den Nummern 1 bis 3, für die "weißen Plätze" die Kugeln mit den Nummern 4 bis 10 zur Verfügung. Die Anzahl der Möglichkeiten der Besetzung einer Reihe aus 5 Kugeln mit genau 2 roten beträgt deshalb

$$3 \cdot 3 \cdot 7 \cdot 7 \cdot 7 = 3^2 \cdot 7^3 .$$

Da es $\binom{5}{2}$ solche Reihen gibt, erhalten wir als Gesamtzahl der für E_2 günstigen Fälle

$$\binom{5}{2} \cdot 3^2 \cdot 7^3 .$$

Somit ergibt sich für $P(E_2)$:

$$P(E_2) = \frac{\binom{5}{2} \cdot 3^2 \cdot 7^3}{10^5} = \binom{5}{2} \cdot \left(\frac{3}{10}\right)^2 \cdot \left(\frac{7}{10}\right)^3 .$$

In diesem Beispiel ist die Wahrscheinlichkeit, bei einem Zug rot zu ziehen, $\frac{3}{10}$, bei einem Zug weiß zu ziehen, $\frac{7}{10}$. Gemäß unserer Bezeichnungsweise gilt daher: $p = \frac{3}{10}$, $q = \frac{7}{10}$.
Wir haben somit gefunden

$$P(E_2) = \binom{5}{2} p^2 q^3 .$$

Für den allgemeinen Fall ergibt die analoge Überlegung

$$P(E_k) = \binom{n}{k} p^k q^{n-k} \text{ mit } p = \frac{r}{N} , \; q = \frac{w}{N} ; \; k \in \{0, ..., n\} , \; n \in \mathbb{N} .$$

Definition:
Die durch

$$p_k = \binom{n}{k} p^k q^{n-k} , \quad 0 \leq k \leq n$$

gegebene Wahrscheinlichkeitsverteilung heißt **Binomialverteilung** mit den Parametern $n \in \mathbb{N}$ und $p \in [0, 1]$.
Die Kurzbezeichnung dieser Verteilung ist $B_{n,p}$-Verteilung.

Bemerkung:

Es ist klar, daß $(p_k : 0 \leq k \leq n)$ eine Wahrscheinlichkeitsverteilung definiert, da

$$\cup_{k=0}^{n} E_k = \Omega \quad \text{und} \quad E_i \cap E_j = \emptyset \quad \text{für} \quad i \neq j .$$

Der binomischen Lehrsatz liefert dies direkt:

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

4.2 Erwartungswert

In diesem Abschnitt wollen wir eine einfache Berechnungsweise für den Erwartungswert einer binomialverteilten Zufallsvariable kennenlernen. Dazu rechnen wir folgendes

Beispiel :
Es sei X eine binomialverteilte Zufallsvariable mit den Parametern p und a) $n = 2$. Berechne $E(X)$ mit Hilfe der Definition des Erwartungswerts. Ergibt sich eine allgemeine Vermutung? (Nr. 9.66 a), S. 266 aus [48])

Da $n = 2$ ist, kann die Zufallsvariable die Werte $0, 1, 2$ annehmen. Die Formel für den Erwartungswert lautet daher

$$E(X) = \sum_{k=0}^{2} k \cdot P(X = k)\,.$$

Für die Wahrscheinlichkeiten $P(X = k)$ ergeben sich:

$$\begin{array}{lclcl} P(X=0) & = & \binom{2}{0} p^0 q^2 & = & (1-p)^2 \\ P(X=1) & = & \binom{2}{1} p^1 q^1 & = & 2p(1-p) \\ P(X=2) & = & \binom{2}{2} p q^0 & = & p^2 \end{array}$$

Daher ist

$$E(X) = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) = 2\,p(1-p) + 2 \cdot p^2 = 2p\,.$$

Für $n = 3$ ergibt sich analog $E(X) = 3p$. Aufgrund dieser Beispiele kann man den folgenden Satz vermuten.

Behauptung 1
Ist S_n eine binomialverteilte Zufallsvariable mit den Parametern n und p, dann gilt:

$$E(S_n) = np\,.$$

Beweis:
Es sei S_n eine binomialverteilte Zufallsvariable mit den Parametern n und p. Dann sind

$$W_{S_n} = \{0, 1, 2, ..., n\}$$

und

$$P(S_n = k) = \binom{n}{k} p^k q^{n-k}\,, \quad k \in W_{S_n}\,,\ q = 1 - p\,.$$

Also ist

$$E(S_n) = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^{n} k \binom{n}{k} p^k q^{n-k}\,,$$

da der Wert $k = 0$ keinen Beitrag zur Summe leistet.
Es ist nun für $1 \le k \le n$ wegen $k! = k \cdot (k-1)!$

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

Also ist

$$\begin{aligned} E(S_n) &= \sum_{k=1}^{n} k \cdot \frac{n}{k} \binom{n-1}{k-1} p^k q^{n-k} \\ &= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-1-(k-1)} \\ &= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{(n-1)-j} \end{aligned}$$

Da

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

ist, folgt

$$E(S_n) = np \quad \text{q.e.d.}$$

4.3 Varianz

Wir wollen jetzt eine Formel für die Varianz einer binomialverteilten Zufallsvariablen kennenlernen. Dazu rechnen wir Folgendes

Beispiel:
Es sei X eine binomialverteilte Zufallsvariable mit den Parametern p und a) $n = 2$. Berechne $V(X)$ mit Hilfe der Definition der Varianz. Ergibt sich eine allgemeine Vermutung? (Nr. 9.66 a), S. 266 aus [48])

Für die Zufallsvariable X gelten

$$W_X = \{0, 1, 2\},$$

$$P(X = k) = \binom{2}{k} p^k q^{2-k}, \ 0 \le k \le 2, \ q = 1 - p,$$

$$E(X) = 2p.$$

Daher folgt für die Varianz

$$\begin{aligned} V(X) &= E(X^2) - [E(X)]^2 = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2^2 \cdot P(X = 2) - (2p)^2 \\ &= 2pq + 2^2 p^2 - (2p)^2 = 2pq. \end{aligned}$$

Für $n = 3$ ergibt sich analog $V(X) = 3pq$. Aufgrund dieser Ergebnisse wird man die allgemeine Formel bereits erraten.

Behauptung 2
Ist S_n eine binomialverteilte Zufallsvariable mit den Parametern n und p, dann gilt

$$V(S_n) = npq\,.$$

Beweis:
Sei S_n eine binomialverteilte Zufallsvariable mit den Parametern n und p. Dann sind

$$W_{S_n} = \{0, 1, 2, ..., n\}\,,$$
$$P(S_n = k) = \binom{n}{k} p^k q^{n-k}\,,\ 0 \le k \le n\,,\ q = 1 - p\,,$$
$$E(S_n) = np\,.$$

In diesem Fall ist nachstehende Formel aus Abschnitt 3.4.3 zweckmäßig:

$$V(X) = E[X(X-1)] - \mu(\mu - 1)\,.$$

Berücksichtigt man - wie im Beweis von Behauptung 1 - dass $k = 0$ keinen Beitrag liefert und $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$ gilt, so erhält man

$$\begin{aligned} E[S_n(S_n-1)] &= \sum_{k=0}^{n} (k-1)k \cdot P(S_n = k) \\ &= np \sum_{k=1}^{n} (k-1) \cdot k \binom{n}{k} p^{k-1} q^{n-k} \\ &= np \sum_{k-1=0}^{n-1} (k-1) \cdot \binom{n-1}{k-1} p^{k-1} q^{(n-1)-k} \\ &= np \sum_{j=1}^{n-1} j \cdot P(S_{n-1} = j)\,, \end{aligned}$$

wobei S_{n-1} eine binomialverteilte Zufallsvariable mit den Parametern $n-1$ und p ist. Die letzte Summe ist die Formel für den Erwartungswert der Zufallsvariablen S_{n-1}, das heißt, es ist

$$E(S_{n-1}) = (n-1)p\,.$$

Es gilt daher

$$E[S_n(S_n-1)] = np \cdot (n-1)p$$

und somit

$$V(S_n) = np \cdot (n-1)p - np \cdot (np-1) = npq \quad \text{q.e.d.}$$

4.4 Eigenschaften der Binomialverteilung

Jede Binomialverteilung ist durch die beiden Größen n (Anzahl der Versuche) und p (Trefferwahrscheinlichkeit) bestimmt. In diesem Abschnitt wollen wir einige Bemerkungen zur Abhängigkeit der Verteilung von diesen Größen machen.

4.4.1 Monotonieverhalten und Maximalstellen

Alle Binomialverteilungen haben folgende Grundform: Die Wahrscheinlichkeiten nehmen mit wachsendem k zuerst zu und dann ab, wobei das Maximum an zwei (benachbarten) Stellen angenommen werden kann. Im folgenden wollen wir diesen Sachverhalt nachweisen.
Die Stelle des Maximums einer Binomialverteilung kann man nicht mit Hilfe der Differentialrechnung bestimmen, da die Funktion $k \mapsto p_k$, $k \in \{0,1,2,...,n\}$ diskret ist. Anstelle der Ableitung verwenden wir den Vergleich benachbarter Funktionswerte mittels Quotientenbildung. Je nachdem, ob der Wert von $\frac{p_k}{p_{k-1}}$ größer, gleich oder kleiner als 1 ist, wächst die Binomialverteilung beim Übergang von $k-1$ zu k, bleibt gleich oder fällt.

Es seien also $0 < p < 1$, $q = 1-p$, $k \in \{1,...,n\}$. Dann ist

$$\frac{p_k}{p_{k-1}} = \frac{\binom{n}{k}p^k q^{n-k}}{\binom{n}{k-1}p^{k-1}q^{n-(k-1)}} = \frac{n-k+1}{k} \cdot \frac{p}{q} = 1 + \frac{p(n+1)-k}{kq}$$

Da das Produkt $k \cdot q$ stets positiv ist, entscheidet der Zähler $p(n+1)-k$, ob die Binomialverteilung beim Übergang von $k-1$ zu k wächst, konstant bleibt oder fällt.

Es gilt

$$p_k > p_{k-1} \iff \frac{p_k}{p_{k-1}} > 1 \iff p(n+1)-k > 0 \iff k < (n+1)p\,.$$

Sei

$$k_0 := [(n+1)p]$$

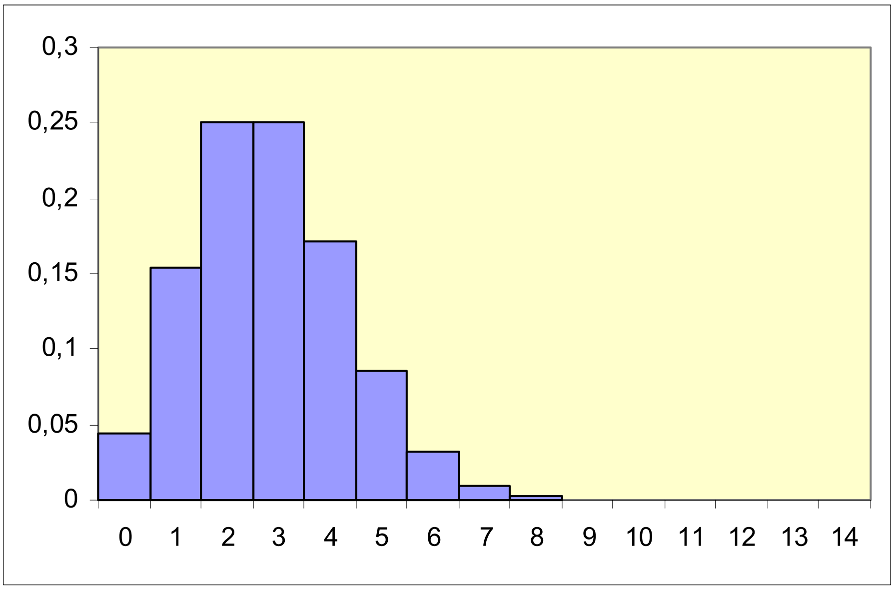
wobei $[x]$ die größte ganze Zahl kleiner oder gleich x bezeichnet (Gauß Klammer).

Ist $(n+1)p$ nicht ganzzahlig, so ist p_{k_0} das eindeutig bestimmte Maximum der Binomialverteilung, und es gilt: Die Binomialverteilung ist streng monoton steigend auf dem Intervall $[0, k_0] \cap \mathbb{N}_0$ und streng monoton fallend auf $[k_0, n] \cap \mathbb{N}_0$.
Ist $(n+1)p$ ganzzahlig, also $(n+1)p = k_0$, so folgt

$$\frac{p_{k_0}}{p_{k_0-1}} = 1 \quad \text{oder} \quad p_{k_0-1} = p_{k_0} \, .$$

In diesem Falle sind $p_{k_{0-1}}$ und p_{k_0} beide maximal, und die Binomialverteilung ist streng monoton steigend auf $[0, k_{0-1}] \cap \mathbb{N}_0$ und streng monoton fallend auf $[k_0, n] \cap \mathbb{N}_0$.

Die Abbildung zeigt das Histogramm der $B_{14,0.2}$-Verteilung.



Daraus ergeben sich für die Maximalstelle(n)der Binomialverteilung folgende Eigenschaften:

Die Maximalstelle(n) liegt (liegen) im Intervall $[np - q, np + p]$ .. Die relati-ve Abweichung vom Erwartungswert strebt für $n \to \infty$ nach 0 .

Die Maximalstelle(n) der Verteilung rückt (rücken) mit wachsendem n bei festem p nach rechts.

Die Maximalstelle(n) der Verteilung rückt (rücken) mit wachsendem p bei festem n nach rechts.

4.4.2 Weitere Eigenschaften

Für eine binomialverteilte Zufallsvariable X, die durch n und p festgelegt ist, gelten:

$$\begin{aligned} E(X) &= np \\ V(X) &= npq \\ \sigma_X &= \sqrt{V(X)} \;=\; \sqrt{npq} \end{aligned}$$

Daraus ergeben sich folgende Eigenschaften:

Der Erwartungswert wächst mit n monoton bei festem p .

Der Erwartungswert wächst mit p monoton bei festem n .

Die Varianz von X und somit die Standardabweichung wird mit wachsendem n bei festem p größer. Dies äußert sich darin, daß die Histogramme der Verteilungen mit wachsendem n bei festem p breiter werden.

Formt man die Formel für die Varianz um, so erhält man bei festem n den Funktionsterm einer nach unten offenen Parabel: $p \longmapsto -np^2 + np$, deren Hochpunkt bei $(\frac{1}{2} \mid \frac{1}{4} \cdot n)$ liegt. Das bedeutet, dass die Varianz und somit die Standardabweichung bei festem n für $p = \frac{1}{2}$ maximal wird. Für die Histogramme heißt dies: Wählt man n fest und vergrößert man p, so werden diese bis zum Wert $p = \frac{1}{2}$ breiter, ab $p = \frac{1}{2}$ wieder schmäler.

Kapitel 5

ÜBUNGSAUFGABEN

5.1 Aufgaben zur beschreibenden Statistik

1. **Augensumme beim Wurf mit drei Würfeln**

   Man werfe $6^3/2 = 108$ mal drei symmetrische Würfel, stelle die jeweilige Augensumme X fest und fertige eine Strichliste an. Danach erstelle man je ein Histogramm für die Häufigkeit der Ausfälle von X
   a) für die einzelnen Werten $j \in \{3, ..., 18\}$ des Wertebereichs W_X
   b) für die Klassen $K_1 = \{3, 4\}$, $K_2 = \{5, 6\}$, $K_3 = \{7, 8\}$, $K_4 = \{9, 10\}$, $K_5 = \{11, 12\}$, $K_6 = \{13, 14\}$, $K_7 = \{15, 16\}$, $K_8 = \{17, 18\}$
   c) für die Klassen $K_1 = \{3, 4, 5, 6\}$, $K_2 = \{7, 8, 9, 10\}$, $K_3 = \{11, 12, 13, 14\}$, $K_4 = \{15, 16, 17, 18\}$
   Der Wertebereich der Augensumme X ist die Menge $W_X = \{3, ..., 18\}$.

2. Beispiel 2 aus Abschnitt 1.1.1

3. **Schwangerschaftsdauern** (siehe [8], S. 407)

   Die nachstehende Tabelle zeigt die in einem der *County General Hospitals* in den U.S.A. im Jahre 1978 erhobenen Dauern von 70 Schwangerschaften.
   a) Erstellen Sie ein *Stängel-Blatt-Diagramm* für die ihrer Größe nach geordneten Daten, b) zeichnen Sie die zugehörige empirische Verteilungs-

funktion und c) fertigen Sie ein Kastenbild an.

251	264	234	283	226	244	269	241	276	274
263	243	254	276	241	232	260	248	284	253
265	235	259	279	256	256	254	256	250	269
240	261	263	262	259	230	268	284	259	261
268	268	264	271	263	259	294	259	263	278
267	293	247	244	250	266	286	263	274	253
281	286	266	249	255	233	245	266	265	264

4. Am National Bureau of Standards in Washington, U.S.A., wurden in den Jahren 1962 – 1963 Messungen des dortigen Standardgewichts NB 10 (Nominalwert von 10 Pond) mit einem der besten Meßgeräte und unter Gewährleistung von möglichst gleichbleibenden Bedingungen durchgeführt. Da alle Messwerte etwa 400 Mikropond unter dem Nominalwert lagen, war es vorteilhaft, die Differenzen der Messwerte vom Nominalwert in Mikropond anzugeben und als Stichprobenwerte zu verwenden. (So ist z.B. für einen Messwert von 9.999591 Pond diese Differenz $10 - 9.999591 = 0.000409$ Pond oder 409 Mikropond.) Bei 100 aufeinanderfolgenden Messungen ergaben sich folgende Stichprobenwerte (vgl. [3], § 6 Measurement Error).

409	400	406	299	402	406	401	403	401	403
398	403	407	402	401	399	400	401	405	402
408	399	399	402	399	397	407	401	399	401
403	400	410	401	407	423	406	406	402	405
405	409	399	402	407	406	413	409	404	402
404	406	407	405	411	410	410	410	401	402
404	405	392	407	406	404	403	408	404	407
412	406	409	400	408	404	401	404	408	406
408	406	401	412	393	437	418	515	404	401
401	407	412	375	409	406	398	406	403	404

a) Erstellen Sie ein *Stängel-Blatt-Diagramm* für die ihrer Größe nach geordneten Daten, b) zeichnen Sie die zugehörige empirische Verteilungsfunktion und c) fertigen Sie ein Kastenbild an.

5. Zu Beispiel 4 in Abschnitt 1.1.3: a) Ermitteln Sie das Stichprobenmittel $\bar{x}_{100}$, die Stichprobenvarianz s_{100}^2 und die geordneten zentrierten und stan-

dardisierten Daten

$$\tilde{x}_{i:100} = \frac{x_{i:100} - \bar{x}_{100}}{s_{100}}\ ,\ i \in \{1, ..., 100\}\,.$$

Sei nun

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})dt\ ,\ x \in \mathbb{R}\,,$$

die Verteilunsfunktion der $N(0,1)$-Verteilung und Φ^{-1} die zugehörige Inverse. Dann gilt offensichtlich $\Phi^{-1}(\Phi(x)) = id(x)$. Darauf beruht das sogenannte *Wahrscheinlichkeitsnetz* (*normal quantile plot*) [1], in dem man die $\Phi^{-1}(i/n)$ gegen die $x_{i:n}$ aufträgt. Die Punktepaare

$$(x_{i:n}, \Phi^{-1}(i/n))\ ,\ i \in \{1, ..., n\}\,,$$

liegen erwartungsgemäß annähernd auf einer Geraden, wenn die Daten aus einer Normalverteilten Grundgesamtheit stammen.

b) Zeichnen Sie die "Punktwolke"

$$(\tilde{x}_{i:100}, \Phi^{-1}(i/n))\ ,\ i \in \{1, ..., 100\}\,,$$

und vergleichen Sie diese mit der Geraden $y = x$.

6. Eine Bakterienkultur wächst während der ca. 16 Tagstunden um 20 % pro Stunde und während der Nacht um nur 12 % pro Stunde. Berechne das durchschnittliche Wachstum pro Stunde! (Beispiel E 1418 aus [47])

7. Bestimmen Sie a) die Fläche des flächengrößten Rechtecks mit vorgegebenem Umfang u und b) das Volumen des volumsgrößten Quaders, dessen Summe der Kantenlängen k ist.

8. Seien $ATX_i \in (0, \infty)$, $i \in \{0, 1, ..., 20\}$ die ATX-Werte [2] von 21 aufeinanderfolgenden Börsetagen, $r_i = ATX_i/ATX_{i-1}$, $i \in \{1, ..., 20\}$. Dann heißt

$$s_{20} = \sqrt{\frac{1}{19}\sum_{i=1}^{20}(\ln(r_i) - \mu_{\ln})^2}$$

[1] Dieses ist im Handel erhältlich; beispielsweise bei Fa. Schleicher & Schüll GmbH, Grimsehlstraße 23, D-37574 Einbeck.

Weiters ist es unter http://www.statistik.tuwien.ac.at/lv-guide/ auf der Homepage des Institus für Statistik & Wahrscheinlichkeitstheorie der TU Wien verfügbar (26.2.2007).

[2] *Austrian Trended Index* (*ATX*)

die *Volatilität*[3] pro Börsentag. Dabei ist

$$\mu_{\ln} = \frac{1}{20} \sum_{i=1}^{20} \ln(r_i) = \ln\left(\sqrt[20]{ATX_{20}/ATX_0}\right)$$

die *durchnittliche logarithmische Tagesrendite.*

a) Entnehmen Sie der Homepage der Österreichischen Nationalbank (www.oenb.at unter dem link: Statistik und Meldeservice/Statistische Daten/Wertpapiere/Aktien/Tägliche Internationale Aktienindizes) die ATX-Werte für die aufeinanderfolgenden Tagen eines Monats und ermitteln Sie die zugehörige Werte der durchschnittlichen logarithmischen Tagesrendite und der Volatilität.

b) Stellen Sie die Wertepaare $(i, ATX_i/ATX_0),\ i \in \{0, ..., 20\}$ graphisch dar und vergleichen Sie die Werte ATX_i/ATX_0 mit den zugehörigen Funktionswerten der Funktionen

$$f_k(t) = e^{\mu_{\ln} t + k s_{30} \sqrt{t}}\,,\ t \in [0, 20]\,,\quad k \in \{-1, 0, 1\}\,.$$

9. Bei einem Autorennen sind 5 Runden zu fahren. Die mittleren Geschwindigkeiten in den einzelnen Runden betragen für einen Fahrer: $183, 210, 201, 180, 182$ (km/h). Wie groß ist die mittlere Geschwindigkeit für alle 5 Runden? (Beispiel E 1419 aus [47]).

10. Von zwei Körpern mit einem Gewicht von jeweils $12\,kg$ verdrängt der eine $2\,l$ und der andere $1\,l$ Wasser. Die Dichten der beiden Körper sind daher $6\,kg/dm^3$ bzw. $12\,kg/dm^3$. Berechnen Sie die durchschnittliche Dichte der beiden Körper (d.i. die Dichte jenes Körpers, welcher durch Zusammenfügen der beiden einzelnen Körper entsteht.)

11. Ermitteln Sie die mittlere Krümmung in den Scheitelpunkten einer Ellipse.

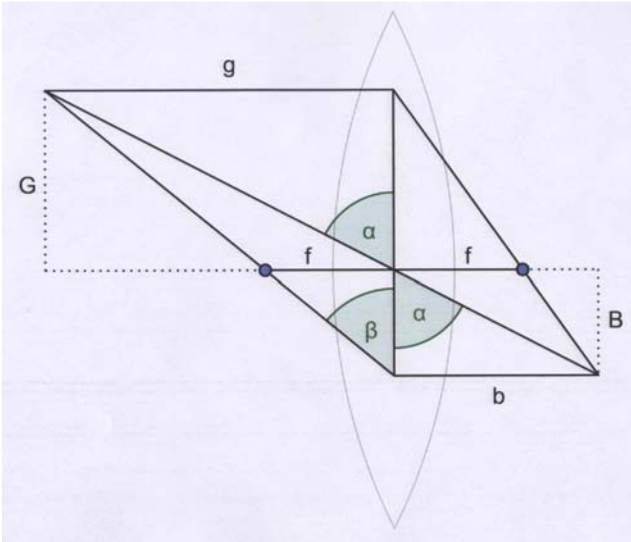
12. Für eine *bikonvexe Linse* bezeichne

    G ... die Gegenstandsgröße, g ... die Gegenstandsweite,

    B ... die Bildgröße, b ... die Bildweite und

[3] abgeleitet vom Italienischen *volare* (fliegen), also "Flatterhaftigkeit", "Beweglichkeit".

f ... die Brennweite.



a) Entnehmen Sie der obigen Abbildung - vermittels Ähnlichkeit zweier Dreiecke - zunächst die Beziehung (1_α) und leiten Sie mit deren Hilfe aus einer weiteren analog entnommenen Beziehung (1_β) die Abbildungsgleichung (2) her:

Abbildungsmaßstab	$\frac{G}{g} = \frac{B}{b}$	(1_α)
Abbildungsgleichung	$\frac{1}{f} = \frac{1}{g} + \frac{1}{b}$	(2)

b) Überzeugen Sie sich davon, dass $2f$ das harmonische Mittel von g und b ist.

13. Zwei gerade Gänge der Breite $b = 1$ verlaufen normal zueinander und münden in einer Ecke zusammen. Jemand möchte einen Stab in waagrechter Lage von einem Gang in den anderen befördern ohne ihn an der Ecke zu biegen oder zu knicken. Wie lang kann der Stab maximal sein ?

14. **Die Babylonische Quadratwurzelapproximation:** Seien $0 < x < y < \infty$, und $a(x,y)$, $g(x,y)$ und $h(x,y)$ das arithmetische, geometrische bzw. harmonische Mittel von x und y. Seien weiters $x_0 := 1$ und $y_0 := c \in (1, \infty)$ und es gelten folgende Rekursionsformeln

$$x_{n+1} := h(x_n, y_n) \qquad \text{und} \qquad y_{n+1} := a(x_n, y_n)\,,\ n \in \mathbb{N}_0\,.$$

Zeigen Sie, dass für alle $n \in \mathbb{N}$ gelten

$$\text{a)}\ \ 1 < x_n < x_{n+1} < \sqrt{c} < y_{n+1} < y_n < c\,, \qquad \text{b)}\ \ y_n - x_n < \frac{c-1}{2^n}$$

und

$$c)\ \ y_{n+1} = \frac{1}{2}\left(y_n + \frac{c}{y_n}\right) \qquad \text{bzw.} \qquad x_{n+1} = \frac{2c}{x_n + \frac{c}{x_n}}\,.$$

5.2 Aufgaben zur Kombinatorik

1. In einem Eissalon gibt es acht verschiedene Eissorten. Um 2.50 Euro bekommt man eine Tüte mit drei Kugeln. Wieviel Geld braucht man, um alle verschiedenen Möglichkeiten durchzuprobieren (gute Gesundheit vorausgesetzt)?
   Beachte beim Lösen der Aufgabe, dass vorher geklärt werden muss, was "Möglichkeiten" bedeutet.
   a) alle drei Kugeln sind verschieden.
   b) zwei Kugeln sollen (dürfen) von derselben Sorte sein
   c) alle drei Kugeln müssen (dürfen) von derselben Sorte sein.
   d) Kommt es auch auf die Reihenfolge an, in der die Kugeln gesetzt werden? (Bsp. 3.1, S. 66 aus [11])

2. Ein Buchhändler hat 10 Bücher, je fünf von zwei Titeln, auf ein Bücherbrett zu stellen. Er stellt sie in beliebiger Weise darauf.
   Wie groß ist die Wahrscheinlichkeit
   a) dass auf fünf Bücher des ersten Titels fünf Bücher des zweiten Titels folgen?
   b) dass sich die beiden Titel auf dem Brett abwechseln.
   Welches Ergebnis wird (vor dem Rechnen) wahrscheinlicher sein? Schätze vorher! (Bsp. 3.10, S. 75 aus [11])

3. a) Wieviele Sitzordnungen gibt es auf einer viersitzigen Bank für vier Personen?
   b) Wieviel verschiedene Sitzplatzanordnungen gibt es in einer Klasse mit 39 Schülern und 30 Plätzen?
   Wie oft müsste man pro Tag (250 Schultage pro Schuljahr) die Sitzordnung ändern, um alle Sitzordnungen innerhalb eines Schuljahres durchgespielt zu haben? (Bsp. 3.3, S. 69 aus [11])

4. Als Beispiel zum Thema Binomialkoeffizienten überlegen wir uns, auf wieviele Arten man im folgenden "Rechteck" das Wort "Mathematik"

lesen kann (aus [18], S. 9):

$$\begin{array}{ccccc} M & A & T & H & E \\ A & T & H & E & M \\ T & H & E & M & A \\ H & E & M & A & T \\ E & M & A & T & I \\ M & A & T & I & K \end{array}$$

Anleitung zur Lösung: Eine Möglichkeit, "Mathematik" zu lesen wäre:

$$\begin{array}{ccccc} M & A & T & \diamondsuit & \diamondsuit \\ \diamondsuit & \diamondsuit & H & \diamondsuit & \diamondsuit \\ \diamondsuit & \diamondsuit & E & M & A \\ \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & T \\ \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & I \\ \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & K \end{array}$$

5. In einer Jugendherberge ist in den Zimmern $1, 2, 4, 7, 8$ und 9 je ein Bett frei. Berechne, auf wie viele Arten vier Wandere auf diese Zimmer aufgeteilt werden können! (Nr.: 686 aus [35])

6. a) Wieviele Möglichkeiten gibt es, bei einer Qualitätskontrolle zufällig aus einem Karton mit 100 Glühbirnen vier Glühbirnen herauszugreifen?

   b) Wieviele Möglichkeiten gibt es für sieben Personen, an den sieben Wochentagen Geburtstag zu haben?
   c) Wieviele Möglichkeiten gibt es für sieben Personen, an sieben verschiedenen Wochentagen Geburtstag zu haben?

   [aus: http://www.matheprisma.uni-wuppertal.de/Module/Kombin/ (5.2.2007)]

5.3 Aufgaben zur Binomialverteilung

1. In einer Urne befinden sich 2 schwarze und eine weiße Kugel. Aus dieser Urne wird 6 mal zufällig mit Zurücklegen gezogen. Dabei kann schwarz 1 mal, 2 mal, ..., 6 mal auftreten. Was ist am wahrscheinlichsten? Vermute und rechne nach! (Nr. 2222, S. 120 aus [49])

2. Eine Firma verkauft Zwiebel einer bestimmten Blumensorte in Zehnerpackungen. Auf jeder Packung steht: "Von 10 Zwiebeln keimen mindestens 9." Man weiß jedoch aus Versuchen mit einer großen Anzahl von Zwiebeln dieser Blumensorte, daß ca. 95% der Zwiebeln keimen. Falls jemand 500 Zehnerpackungen kauft, wieviele Packungen werden voraussichtlich das Garantieversprechen nicht erfüllen? (Nr. 5.14, S. 103 aus [50])

3. Bei einer Glühlampenproduktion ist bekannt, daß auf lange Sicht ca. 1% defekte Lampen produziert werden. Zur Kontrolle werden stündlich 10 Glühlampen entnommen. Ergibt sich dabei eine defekte Glühlampe, wird der Produktionsprozeß gestoppt und überprüft. Wie groß ist die Wahrscheinlichkeit, daß bei einer Kontrolle der Produktionsvorgang gestoppt wird? (Nr. 5.15, S. 104 aus [50])

4. Auf einer Hühnerfarm werden Eier in Schachteln zu 12 Stück verpackt. Jedes Ei ist mit der Wahrscheinlichkeit $p = \frac{1}{12}$ angebrochen.

   a) Mit welcher Wahrscheinlichkeit enthält eine Schachtel nur gute Eier?
   b) Mit welcher Wahrscheinlichkeit enthält eine Schachtel zwei oder mehr angebrochene Eier?

   c) 10 Schachteln werden an 10 Einzelhändler verkauft. Mit welcher Wahrscheinlichkeit erhalten 2 Einzelhändler je eine Schachtel mit nur guten Eiern? (Nr. 14, S. 136 aus [53])

5. Eine Firma, die einen Massenartikel in Paketen zu je 15 Stück an den Einzelhandel vertreibt, vereinbart, daß Pakete mit mehr als 2 schadhaften Stücken nicht berechnet werden. Wieviel Prozent der ausgelieferten Pakete muß die Firma als unberechnet kalkulieren, wenn ihr bekannt ist, daß durchschnittlich nur 2% der Artikel schadhaft sind. (Nr. 10, S. 136 aus [53])

6. Ein Beispiel zur Bestimmung des Stichprobenumfanges: Die folgende Fragestellung findet man in vielen Lehrbüchern der Wahrscheinlichkeitsrechnung.

   Einem Teig für einen Guglhupf werden n Rosinen beigegeben. Der Guglhupf wird in 16 gleiche Teile geteilt. Wieviele Rosinen muß man als Zutat nehmen, damit die Wahrscheinlichkeit, daß in einem bestimmten

Stück des Guglhupfs mindestens eine Rosine ist, größer als 99% wird? Bestimme die gesuchte Anzahl mit Hilfe der Binomialverteilung.

Anleitung:

Um das Beispiel einer mathematischen Behandlung zugänglich zu machen, muß es in ein Urnenmodell übersetzt werden, etwa folgendermaßen:

In einer Urne befinden sich 16 Kugeln: 15 weiße und eine schwarze. Wie oft muß man zufällig und mit Zurücklegen ziehen, daß mit 99% Sicherheit mindestens einmal schwarz gezogen wird?

Literaturverzeichnis

[1] **Bücher**

[2] *Bernoulli, J.*: Wahrscheinlichkeitsrechnung (Ars Conjectandi). Erstpublikation 1713; Übersetzung von *R. Hausner*: Ostwalds Klassiker der Exakten Wissenschaften, Band 107; Verlag Harri Deutsch, Frankfurt am Main 1999

[3] *Freedman, D., Pisani, R.* and *R. Purves*: Statistics. Norton & Co., New York 1978

[4] *Hartung, J., Elpelt, B.* und *K-H. Klösener*: Statistik: Lehr- und Handbuch der angewandten Statistik. Oldenbourg Verlag, München - Wien 1991

[5] *Krämer, W.*: Statistik verstehen: Eine Gebrauchsanweisung. Campus Verlag, Frankfurt - New York, 1999

[6] *Krämer, W.*: So lügt man mit Statistik. Campus Verlag, Frankfurt - New York 1997

[7] *Kröpfl, B., Peschek, W., Schneider, E.* und *A. Schönlieb*: Angewandte Mathematik - Eine Einführung für Wirtschaftswissenschaftler und Informatiker. Carl Hanser Verlag, München - Wien 1994

[8] *Larsen, R.J.* and *M.L. Marx*: An Intoduction to Mathematical Statistics and its Applications. Prentice-Hall, Englewood Cliffs, New Yersey 1986

[9] *Moore, D.S.*: Statistics: Concepts and Controversies. W.H. Freeman & Co., San Francisco 1979

[10] *Österreicher, F.*: Die Normalverteilung in Wort und Bild - Einführung in die stochastische Modellbildung am Beispiel der Meßfehler. In: Ausflüge in die Mathematik - 21 Jahre Institut für Mathematik, Abakus Verlag, Salzburg 1988

[11] *Reichel, H.Ch., G. Hanisch* und *R. Müller*: Wahrscheinlichkeitsrechung und Statistik. In: Mathematik für Schule und Praxis. Band 1, Verlag Hölder Pichler Tempsky, Wien 1989

[12] *Topsoe, F.*: Spontane Phänomene - Stochastische Modelle und ihre Anwendungen. Vieweg, Braunschweig - Wiesbaden 1983

Zeitschriften

[13] *Engel, A.*: Statistik in der Schule: Ideen und Beispiele aus neuerer Zeit. Der Mathematikunterricht, Jahrgang 28, Heft 1 (1982), S. 57 – 85

[14] *Kütting, H.*: Stochastisches Denken in der Schule - Grundlegende Ideen und Methoden. Der Mathematikunterricht, Heft 4 (1985), S. 87 – 106

[15] *Diepgen, R.*: Warum nur $n-1$ und nicht n? Erwartungstreue - leicht gemacht. Stochastik in der Schule **19** (1999), Nr. 1, S. 10 – 13

[16] *Hirscher, H.*: Mittelwertfolgen - oder: Mitten inmitten von Mitten. Der Mathematikunterricht, Jahrgang 50, Heft 5 (2004), S. 42 – 55

Skriptum

[17] *Österreicher, F.*: Skriptum zur Lehrveranstaltung Statistik für Lehramt. Salzburg 2001

Diplomarbeiten

[18] *Weiß, M.*: Binomialverteilung und Normalapproximation: Grundlegendes und Hintergrundinformation für den Stochastikunterricht. Salzburg 1995

[19] *Jandl, M.*: Computereinsatz im Stochastikunterricht. Salzburg 1997

[20] *Radauer, Ch.*: Läufe in binären Zufallsfolgen - Beurteilung und Kompression von Zufallsdaten. Salzburg 1998

[21] *Fritz, F.*: Proportionen in Mathematik und Musik - Kunst und Wissenschaft ergänzen einander. Salzburg 1998

[22] *Reichenberger, S.*: Mittelwerte der Pythagoreer und babylonischer Wurzelalgorithmus. Salzburg 2007

Gängige Schulbücher für die AHS

Unterstufe

[23] *Lindbichler·Bartl·Hartmann·Janeschitz·Varelija*: Querschnitt Mathematik 2, Westermann, Wien 2001

[24] *Lindbichler·Bartl·Hartmann·Janeschitz·Varelija*: Querschnitt Mathematik 3, Westermann, Wien 2001

[25] *Lindbichler·Bartl·Hartmann·Janeschitz·Varelija*: Querschnitt Mathematik 4, Westermann, Wien 2001

[26] *Reichel·Litschauer·Groß*: Das ist Mathematik 2, öbvhtp Verlagsgesellschaft, Wien 2000

[27] *Reichel·Litschauer·Groß*: Das ist Mathematik 3, öbvhtp Verlagsgesellschaft, Wien 2006

[28] *Reichel·Litschauer·Groß*: Das ist Mathematik 4, öbvhtp Verlagsgesellschaft, Wien 2006

[29] *Lewisch*: Mathematik: Verstehen - Üben - Anwenden, Band 2, Oldenbourg, Wien 2004

[30] *Lewisch*: Mathematik: Verstehen - Üben - Anwenden, Band 3, Oldenbourg, Wien 2005

[31] *Lewisch*: Mathematik: Verstehen - Üben - Anwenden, Band 4, Oldenbourg, Wien 2006

Oberstufe

[32] *Malle·Ramharter·Ulovec·Kandl*: Mathematik verstehen 6, öbvhtp Verlagsgesellschaft, Wien 2005

[33] *Malle·Ramharter·Ulovec·Kandl*: Mathematik verstehen 7, öbvhtp Verlagsgesellschaft, Wien 2006

[34] *Malle·Ramharter·Ulovec·Kandl*: Mathematik verstehen 8, öbvhtp Verlagsgesellschaft, Wien 2007

[35] *Götz·Reichel·R.Müller·Hanisch·Hederer·Wenzel·M.Müller*: Mathematik Lehrbuch 6, öbvhtp Verlagsgesellschaft, Wien 2005

[36] *Götz·Reichel·R.Müller·Hanisch·Hederer·Wenzel·M.Müller*: Mathematik Lehrbuch 7, öbvhtp Verlagsgesellschaft, Wien 2006

[37] *Götz·Reichel·R.Müller·Hanisch·Hederer·Wenzel·M.Müller*: Mathematik Lehrbuch 8, öbvhtp Verlagsgesellschaft, Wien 2007

[38] *Taschner*: Mathematik 2 - Übungs- und Lehrbuch für die 6. Klasse AHS, Oldenbourg, Wien 1999

[39] *Taschner*: Mathematik 3 - Übungs- und Lehrbuch für die 7. Klasse AHS, Oldenbourg, Wien 2000

[40] *Taschner*: Mathematik 4 - Übungs- und Lehrbuch für die 8. Klasse AHS, Oldenbourg, Wien 2001

[41] *Geretschläger·Griesel·Postel*: Elemente der Mathematik 6, E.Dorner, Wien 2005

[42] *Geretschläger·Griesel·Postel*: Elemente der Mathematik 7, E.Dorner, Wien 2006

[43] *Geretschläger·Griesel·Postel*: Elemente der Mathematik 8, E.Dorner, Wien 2007

[44] *Steiner·Novak*: MatheMaster 6 - Mathematik für die 6. Klasse AHS, Reniets Verlag, Wien 2005

[45] *Steiner·Novak*: MatheMaster 7 - Mathematik für die 7. Klasse AHS, Reniets Verlag, Wien 2006

[46] *Steiner·Novak*: MatheMaster 8 - Mathematik für die 8. Klasse AHS, Reniets Verlag, Wien 2007

Weitere Schulbücher für die AHS und BHS

[47] *Kronfellner, M. Peschek·Blasonig·Fischer·Kronfellner, J.*: Angewandte Mathematik 4. Verlag Hölder Pichler Tempsky, Wien 1997

[48] *Bürger·Fischer·Malle·Kronfellner·Mühlgassner·Schörghofer*: Mathematik Oberstufe 3. Verlag Hölder Pichler Tempsky, Wien 1991

[49] *Bürger·Fischer·Malle·Cejnek·Mühlgassner*: Mathematik Oberstufe 4. Verlag Hölder Pichler Tempsky, Wien 1987

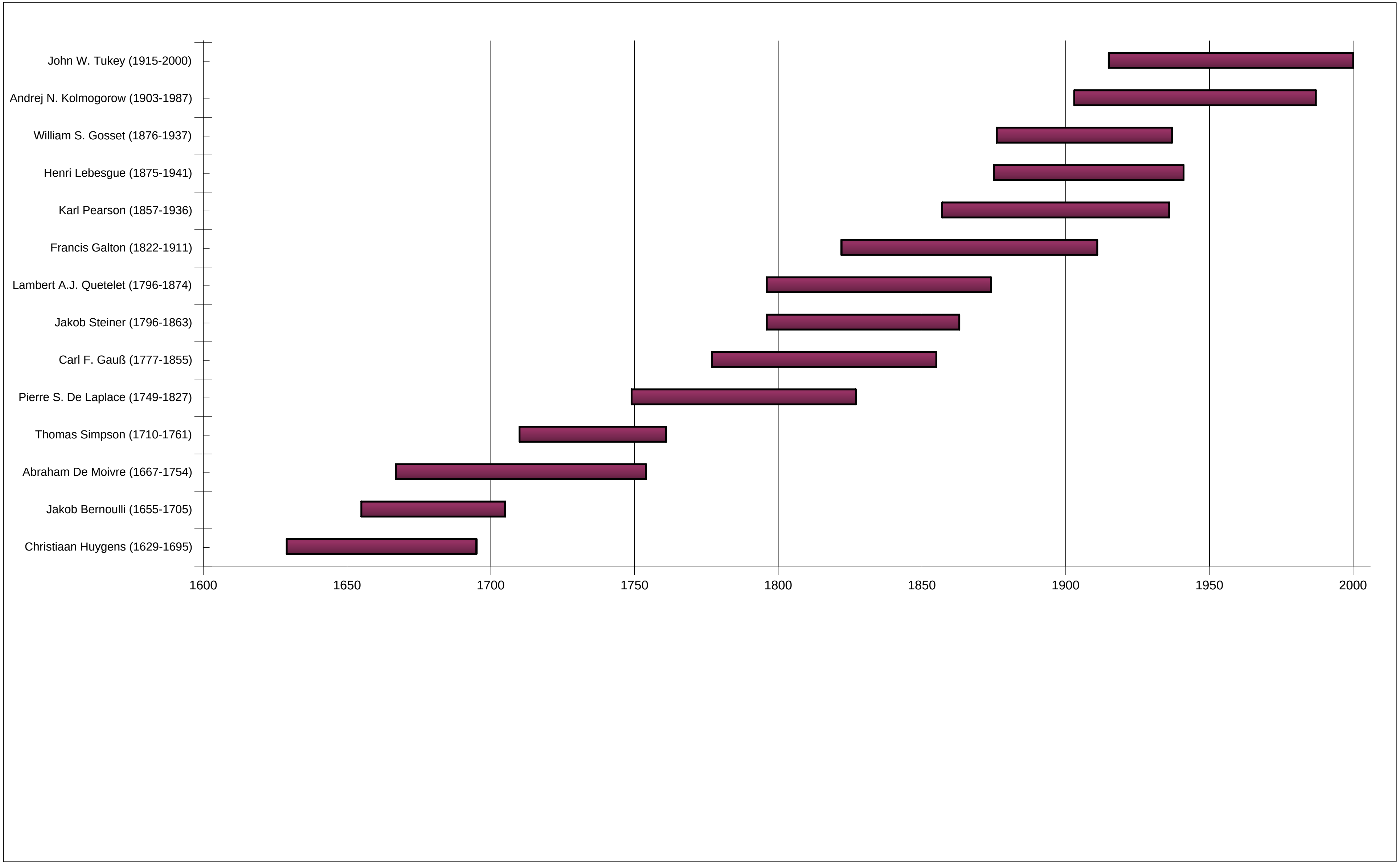
[50] *Bürger·Fischer·Malle·Kronfellner·Mühlbacher·Schlöglhofer*: Mathematik Oberstufe 4. Verlag Hölder Pichler Tempsky, Wien 1992

Gängige Schulbücher aus Deutschland

[51] *Barth·Haller*: Stochastik Leistungskurs. Ehrenwirth Verlag, München 1983

[52] *Heigl·Feuerpfeil*: Stochastik Leistungskurs. Bayrischer Schulbuch Verlag, München 1987

[53] *Lambacher·Schweizer*: Stochastik Leistungskurs. Ernst Klett Schulbuchverlag, Stuttgart 1988



John W. Tukey (1915-2000)
Andrej N. Kolmogorow (1903-1987)
William S. Gosset (1876-1937)
Henri Lebesgue (1875-1941)
Karl Pearson (1857-1936)
Francis Galton (1822-1911)
Lambert A.J. Quetelet (1796-1874)
Jakob Steiner (1796-1863)
Carl F. Gauß (1777-1855)
Pierre S. De Laplace (1749-1827)
Thomas Simpson (1710-1761)
Abraham De Moivre (1667-1754)
Jakob Bernoulli (1655-1705)
Christiaan Huygens (1629-1695)
1600
1650
1700
1750
1800
1850
1900
1950
2000