

LKAbi08: Stochastik

BW

17. Juni 2007

Für und mit: Philipp, Vanessa, Viktor, Ilja, Alexander, Matthias, Kerstin, Joschka, Thomas K., Julian, Eyüp, Daniel P., Konstantin, Thomas S., Christian, Johannes, Clas, Daniel U., Ivo und Benjamin.

Inhaltsverzeichnis

1	Einführung	3
1.1	Wahrscheinlichkeit und Zufall in der Umwelt	3
1.2	Konkrete Einstiegsbeispiele und Grundbegriffe	3
1.3	Wahrscheinlichkeit	4
1.4	Hilfsmittel zur Berechnung von Wahrscheinlichkeiten	4
1.5	Beispiele	5
1.5.1	Skat und subjektive Wahrscheinlichkeiten	5
1.5.2	Ziehen ohne Zurücklegen	6
1.5.3	Ziehen mit Zurücklegen	6
1.5.4	Würfeln, bis eine 5 kommt	6
2	Zufallsgrößen	7
2.1	Definition und Begriffe	7
2.2	Der Erwartungswert $E(X)$ einer Zufallsgröße X	7
2.3	Linearität des Erwartungswertes	8
2.4	Die Varianz	9
2.5	Bernoullis Gesetz der großen Zahl	10
2.6	Die Näherungsformeln von de Moivre und Laplace und ihr Gebrauch	12
3	Konfidenzintervalle	14
3.1	Einführung	14
3.2	Ein durchgerechnetes Beispiel	16
3.3	Aufgaben	16
3.4	Ein Beispiel für die Bestimmung der notwendigen Stichprobenlänge n	16
4	Tests	17
4.1	Einführungsbeispiel	17
4.2	Schema eines Tests	18
4.3	Test-FAQ	19
5	Stetig verteilte Zufallsgrößen	20
5.1	Einführung	20
5.2	Die Standardnormalverteilung	21
5.3	Erwartungswert und Varianz einer stetig verteilten Zufallsgröße	22
6	Klausuren	22
6.1	Klausur am 7. März 2007	22
6.2	Nachschreibklausur	23
6.3	Klausur am 2. Mai 2007	24
6.4	Nachschreibklausur am 30. Mai 2007	25
7	Komplexe Zahlen	27
7.1	Die Gaußsche Zahlenebene	27
7.2	Summe und Produkt komplexer Zahlen	27
7.3	Die Darstellung $\alpha = a_1 + a_2i$	28
7.4	Die Darstellung $\alpha = re^{i\varphi}$	29
7.5	Konjugiert komplexe Zahl und Division	29
7.6	Übungen	30
7.7	Vom Nutzen der komplexen Zahlen	30
7.8	Funktionen $f : \mathbb{C} \rightarrow \mathbb{C}$	31
7.9	Wirkung von f „im Kleinen“	32
7.10	Ausblick auf Anwendungen komplexer Funktionen	35

1 Einführung

1.1 Wahrscheinlichkeit und Zufall in der Umwelt

Nichts geschieht ohne Grund, alles hat seine Ursache. Klingt gut, nicht? Aber nimm dich in Acht: Wenn du den Gedanken zu Ende denkst, könnte die Sache so ausgehen, dass die Welt funktioniert wie ein Uhrwerk, das in grauer Vorzeit in Gang gesetzt wurde und nun läuft, und dann lebst du dein Leben ab, und alles was du tust oder nicht tust, was dir geschieht oder nicht geschieht, ist vorherbestimmt. Dann gibt es weder Spontaneität noch Verantwortung. Dass dir gerade einfällt, die Hausaufgaben für morgen nicht zu machen und statt dessen ins Kino zu gehen, ist beileibe nicht deine Schuld, sondern ergibt sich aus chemischen Reaktionen in einem Eckchen deines Kopfes, die gerade so ablaufen, weil die Gesamtheit der Einflüsse des Weltlaufs, von genetischen Prägungen über deine frühkindlichen Erlebnisse bis hin zu atmosphärischen Erscheinungen dies unabwendbar erzwingen. Letztlich tragen auch die Schmetterlinge im Hiller Moor durch ihre Flügelschläge dazu bei. Das mit dem Kino mag dir gefallen, aber der, der dir morgen auf der Straße dein Handy wegnimmt, verweist dann auch achselzuckend auf Schicksal und Vorherbestimmung, und das findest du dann vermutlich nicht mehr so gut.

Was ist also mit dem freien Willen? Ich weiß es nicht. Bei Philosophen habe ich keine tragfähigen Antworten gefunden; die Quantenphysik öffnet uns ein Schlupfloch der Unbestimmtheit, aber so attraktiv ist das auch nicht. Ich gebe mich damit zufrieden, dass der Versuch des Verstandes, das Gesamtsystem zu erkennen, von dem er doch ein Teil ist, von vornherein aussichtslos ist. Vielleicht findest du ja eine bessere Antwort.

So stehen die Dinge nun: Keiner weiß letztlich – sage ich jetzt einmal etwas zugespitzt – was Zufall ist und ob es so etwas wie Zufall überhaupt gibt. Dennoch gibt es die Stochastik, die Mathematik des Zufalls, und sie wird sehr erfolgreich angewandt, wie die stattlichen Gebäude der Versicherungen eindrücklich belegen. „We do not know, who will die next year, but we know, how many“, sagt man dort, und wenn wir im Kurs etwas weiter sind, wirst du verstehen, wieso der Satz zutrifft.

Stochastik wird angewandt, wo die Situation zu komplex ist für einfache wenn–dann–Aussagen. Die Luft im Kursraum besteht aus einer riesigen Anzahl von Teilchen, die in wilder, regelloser Bewegung sind. Der Versuch, dieses System mit Mitteln der Mechanik zu behandeln, ist völlig aussichtslos, aber mit Hilfe der Stochastik macht die Statistische Physik Aussagen über das Gesamtsystem. Niemand weiß, wann ein radioaktives Teilchen zerfällt, das geschieht „rein zufällig“, aber bei einer großen Anzahl von Teilchen kann man recht präzise angeben, wie lange es dauert, bis die Hälfte zerfallen ist. Ein Medikament bewirkt bei dem einen dies, bei dem anderen das, bei einem dritten überhaupt nichts. Die Einschätzung, dass es wirksam ist, gewinnt man durch statistische Untersuchungen. Ähnlich sieht es in der Marktforschung aus. Weitere wichtige Anwendungsfelder sind die Qualitätskontrolle in der industriellen Fertigung, Wahlanalysen und, mit Blick auf Espelkamp, die Konstruktion von Glücksspielen.

1.2 Konkrete Einstiegsbeispiele und Grundbegriffe

Lasst uns mit einfachen Beispielen anfangen, mit dem Werfen eines Würfels mit Hilfe eines Würfelbechers und mit dem Ziehen einer Kugel aus einer Urne, die mit drei roten, zwei schwarzen und einer grünen Kugel bestückt ist. Beides sind Beispiele für Zufallsexperimente: Man kann sie im Prinzip beliebig oft wiederholen, und man weiß nicht im Vornherein, welches Ergebnis bei einer Durchführung herauskommt. Man weiß nur, dass das Ergebnis ein Element ω aus einem bestimmten Vorrat Ω ist, dem sogenannten Ergebnisraum. Beim Würfeln ist $\Omega = \{1, 2, 3, 4, 5, 6\}$. Ein Ereignis ist eine Teilmenge $A \subseteq \Omega$, und man sagt, das Ereignis sei eingetreten, wenn ein $\omega \in A$ ausgelost wurde.

1.3 Wahrscheinlichkeit

Der Begriff der Wahrscheinlichkeit ist der zentrale Begriff der Stochastik. Du hast eine gewisse intuitive Vorstellung, was Wahrscheinlichkeit ist, jedenfalls sind dir einschlägige Sprechweisen geläufig. Wenn ich dich frage, wie wahrscheinlich es ist, eine 2 zu würfeln, wirst du ohne zu zögern mit $\frac{1}{6}$ antworten. Aber was bedeutet das nun? Erwartest du eine einfache Antwort? Hören wir einem Gespräch zweier Leute zu. Sie heißen Hinz und Kunz, es sind alte Bekannte von mir. Kunz ist gutwillig und eifrig, aber vielleicht ein wenig unbedarft. Hinz jedoch ist von scharfem Verstand und eher kritisch. Am hilfreichsten sind sie, wenn sie zu zweit auftreten. Kunz wagt mutige Vorschläge und bringt damit die Entwicklung weiter. Mutige Vorschläge enthalten neben brauchbaren Ideen oft Blödsinn, und Hinzens Job ist es, beides von einander zu unterscheiden.

K: Wir lernen jetzt Stochastik. Die Wahrscheinlichkeit, mit einem Würfel eine 2 zu werfen, ist $\frac{1}{6}$. Und wir rechnen dann aus, wie wahrscheinlich es ist, dass unter 6 Würfeln keine 2 ist, oder dass man, wenn man 6-mal gewürfelt hat, alle möglichen Ergebnisse bekommen hat.

H: Klar, wenn du die Gesamtwahrscheinlichkeit auf 1 setzt, wirst du jedem der sechs möglichen Ergebnisse beim Würfeln die Wahrscheinlichkeit $\frac{1}{6}$ zuweisen, schon aus Symmetriegründen. Aber was bedeutet diese Wahrscheinlichkeit $\frac{1}{6}$ einer 2 für einen realen Würfel?

K: Hm, dass jeder 6-te Wurf eine 2 ist?

H: Kunz, du weißt ganz genau, dass das nicht stimmt; du hast doch selbst „Mensch ärgere dich nicht“ gespielt.

K: Klar, praktisch stimmt das nicht, aber ich meine: theoretisch.

H: Schöne Theorie, die etwas vorhersagt, was überhaupt nicht stimmt. Rechne doch die Wahrscheinlichkeit aus, dass von 60 Würfeln eines Würfels genau jeder sechste eine 2 ist. Du wirst sehen, da kommt nicht viel heraus.

Die Wahrscheinlichkeiten beim Werfen eines Würfels, die Kunz ausrechnet, sind Gegenstände einer mathematischen Theorie. Die Theorie sagt nicht, was Wahrscheinlichkeit „ist“, sondern nur, wie man damit umgeht. Auch der L-Würfel unseres Buches ist etwas Ausgedachtes, wie eine Gerade der euklidischen Geometrie. Aber die Geraden der Geometrie sind gewissermaßen Abstraktionen von Gegenständen unserer Umwelt, von Lichtstrahlen meinetwegen. Auch das Konzept der Wahrscheinlichkeit ist angeregt von Umweltphänomenen. Es gibt das **empirische Gesetz der großen Zahl**, eine Erfahrungstatsache, die folgendes besagt: Wirft man einen einwandfreien Würfel sehr oft, dann ist der Anteil der 2-en unter den geworfenen Zahlen etwa $\frac{1}{6}$. Diese Beobachtung ist die Rechtfertigung dafür, dass die Wahrscheinlichkeitsrechnung auf die Realität anwendbar ist. Was wir über unseren L-Würfel ausrechnen, sagt etwas über reale Würfel. Und wir werden sehen, dass sich unser idealer Würfel eben so verhält: der Anteil der 2-en an einer großen Zahl von Würfeln unterscheidet sich mit hoher Wahrscheinlichkeit nur wenig von $\frac{1}{6}$. Das sagt das Bernoullische Gesetz der großen Zahl, ein mathematischer Satz, den wir später sauber formulieren und beweisen wollen.

1.4 Hilfsmittel zur Berechnung von Wahrscheinlichkeiten

Hier sind zunächst einige Zählmethoden zu nennen. Wenn alle Ergebnisse gleich wahrscheinlich sind, braucht man nur die Ergebnisse zu zählen, die zu einem Ereignis $A \subseteq \Omega$ gehören, dann ist

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

Betrachten wir eine Urne mit n unterschiedlichen Kugeln, aus der wir k Kugeln ziehen. Die Anzahl $|\Omega|$ der Elemente des Ergebnisraums Ω ist dann

1. n^k , wenn wir mit Zurücklegen ziehen;

2. $n(n-1)(n-2)\cdots(n-k+1)$, wenn wir ohne Zurücklegen ziehen und auf die Reihenfolge achten, in der die Kugeln gezogen wurden;
3. $\binom{n}{k}$, wenn wir ohne Zurücklegen ziehen und die Reihenfolge, in der die Kugeln gezogen wurden, egal ist. Das ist zum Beispiel beim Lotto so.

Besonders interessant und nützlich sind die **Binomialkoeffizienten** $\binom{n}{k}$. Es ist

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

= Anzahl der k -Teilmengen einer n -Menge.

Oft hilft auch ein Baumdiagramm weiter. Damit bist du vertraut, da gebe ich jetzt kein Beispiel. Manchmal kann man leichter die Gegenwahrscheinlichkeit $P(\bar{A})$ ausrechnen als $P(A)$, dann kommt man mit $P(A) = 1 - P(\bar{A})$ zum Ziel.

Mit diesen wenigen Werkzeugen können wir schon allerhand erreichen.

1.5 Beispiele

1.5.1 Skat und subjektive Wahrscheinlichkeiten

Skat spielt man mit einem Kartenspiel, das aus 32 Karten besteht. Es wird von drei Spielern gespielt; jeder bekommt zehn Karten, die restlichen zwei liegen im Stock. Beim Skat spielen die vier Buben eine besondere Rolle.

Wir diskutieren nun eine Reihe von Fragen.

1. Hinz hat seine zehn Karten noch nicht angeschaut. Er fragt sich, mit welcher Wahrscheinlichkeit genau ein Bube im Stock liegt und mit welcher Wahrscheinlichkeit Mitspieler Kunz genau drei Buben hat. Die Antworten lauten

$$p_1 := \frac{\binom{4}{1} \binom{28}{1}}{\binom{32}{2}}$$

im ersten Fall und

$$p_2 := \frac{\binom{4}{3} \binom{28}{7}}{\binom{32}{10}}$$

im zweiten Fall. Begründung für den zweiten Fall: Es gibt $\binom{32}{10}$ Teilmengen von zehn Karten der Gesamtmenge von 32 Karten, und die sind alle gleich wahrscheinlich. Die Anzahl dieser Teilmengen, die genau drei Buben enthalten, bekommt man so: Man wählt drei der vier Buben und sieben der 28 übrigen Karten. Jede dieser beiden gewählten Mengen kann man zu einer 10-Menge mit drei Buben zusammenstellen, und so erhält man alle möglichen 10-Teilmengen mit genau drei Buben.

2. Nun schaut sich Hinz seine Karten an, und er stellt fest, dass er genau zwei Buben auf der Hand hat. Nun berechnet er die beiden Wahrscheinlichkeiten p_1 und p_2 neu:

$$p'_1 := \frac{\binom{2}{1} \binom{20}{1}}{\binom{22}{2}}$$

und

$$p'_2 := \frac{\binom{2}{3} \binom{20}{7}}{\binom{22}{10}} .$$

Natürlich ist $p_2 = 0$, denn $\binom{2}{3} = 0$.

Ich hoffe schon, dass du jetzt einigermaßen verwirrt bist. Wieso ändert sich denn etwas am Stock, wenn Hinz seine Karten anschaut? Kiebitz Karl, der in den Stock gesehen hat und deshalb weiß, dass Karo-sieben und Pik-acht im Stock sind, hört Hinz sagen, mit der Wahrscheinlichkeit $p_1 = \frac{7}{31}$ liege genau ein Bube im Stock, und er tippt sich natürlich an die Stirn. Was bedeutet das denn nun? Offensichtlich sagen p_1 und p'_1 etwas aus über die Information, die Hinz von der Situation hat, und nichts über die konkreten Karten im Stock; manche reden denn auch von einem subjektiven Wahrscheinlichkeitsbegriff. Man kann die Sache so retten, dass man sagt: Schaut man sich sehr oft den Stock an, wird man in etwa $p_1 = \frac{7}{31}$ der Fälle genau einen Buben im Stock finden. Betrachtet man nun nur noch die Fälle, in denen Hinz zwei Buben auf der Hand hat, ist der Anteil von diesen, bei denen genau ein Bube im Stock liegt, etwa p'_1 . Aber meinetwegen kann man auch subjektive Wahrscheinlichkeiten benutzen, wenn sie hilfreich sind. In der Wahrscheinlichkeitstheorie hat man sowieso bloß ein Ω und ein P dazu, und welcher Zufallsversuch dahintersteckt, ist völlig irrelevant. Der Anwender der Wahrscheinlichkeitstheorie muss sehen, dass er zurechtkommt.

1.5.2 Ziehen ohne Zurücklegen

In einer Urne seien N Kugeln, von denen seien n rot. Wir ziehen m Kugeln ohne zurücklegen und bezeichnen die Anzahl der gezogenen roten Kugeln mit X . Dann ist

$$P(X = k) = \frac{\binom{n}{k} \binom{N-n}{m-k}}{\binom{N}{m}} ,$$

und die Zufallsgröße X heißt **hypergeometrisch verteilt**. Wir hatten für $P(X = k)$ noch einen zweiten Rechenausdruck erhalten, indem wir uns vorstellten, wie die Kugeln der Reihe nach gezogen werden; den schenke ich mir hier.

1.5.3 Ziehen mit Zurücklegen

Der Anteil der roten Kugeln in einer Urne sei p . Aus der Urne ziehen wir n -mal eine Kugel und legen sie immer gleich wieder zurück. Wieder sei X die Anzahl der gezogenen roten Kugeln. Dann ist

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} ,$$

und die Zufallsgröße X heißt **binomial verteilt**.

1.5.4 Würfeln, bis eine 5 kommt

Wie würfeln so lange, bis wir eine 5 haben, und es sei X die Anzahl der benötigten Würfe. Dann ist

$$P(X = k) = \left(\frac{5}{6}\right)^{k-1} \left(\frac{1}{6}\right) .$$

Es wäre übrigens schon eine recht delikate Aufgabe, zu diesem Zufallsversuch ein Ω hinzuschreiben.

2 Zufallsgrößen

2.1 Definition und Begriffe

Dir ist schon mehrfach ein X begegnet, und ich habe auch schon beiläufig erwähnt, dass X Zufallsgröße genannt wird. Jetzt wollen wir die Sache formal etwas sauberer angehen. Wie es sich in der Mathematik gehört, beginne ich mit einer Definition.

1 Definition

Es sei Ω ein Ergebnisraum und $X : \Omega \rightarrow \mathbb{R}$ eine Funktion. Dann heißt X eine Zufallsgröße auf dem Ergebnisraum Ω .

Jedem $\omega \in \Omega$ wird also durch X eine Zahl zugeordnet, wir bezeichnen sie wie gewohnt mit $X(\omega)$. Zahlen kann man der Größe nach anordnen und mit Zahlen kann man rechnen, das werden wir noch ausnutzen. Zum anderen stiftet das X eine Struktur auf Ω , das heißt Beziehungen zwischen den $\omega \in \Omega$. Man steckt natürlicherweise alle ω in einen Topf, die den gleichen X -Wert haben. Genauer: Für jedes $\omega \in \Omega$ ist $k = X(\omega)$ eine Zahl, und

$$X^{-1}(k) := \{\omega \in \Omega \mid X(\omega) = k\}$$

ist die Teilmenge all der $\omega \in \Omega$, die durch X auf dieses k abgebildet werden. Die Wahrscheinlichkeit, dass der Zufallsversuch ein ω mit diesem $k = X(\omega)$ liefert¹, ist folglich

$$P(X = k) := P(X^{-1}(k)) \quad .$$

Nützliche Hilfsmittel bei der Arbeit mit Zufallsgrößen sind:

1. Tabellen, die alle Werte k von X und deren Wahrscheinlichkeit $P(X = k)$ auflisten;
2. Histogramme, das sind bildliche Darstellungen dieser Tabellen;
3. die Verteilungsfunktion F mit $F(x) := P(X < x)$ für alle $x \in \mathbb{R}$.

Weitere folgen in Kürze.

2.2 Der Erwartungswert $E(X)$ einer Zufallsgröße X

Es sei X eine Zufallsgröße auf dem Ergebnisraum Ω . Der Erwartungswert $E(X)$ von X ist, praktisch gesprochen, der Wert, der sich bei sehr häufiger Versuchsdurchführung im Mittel je Durchführung ergibt. Wenn man also den zugehörigen Zufallsversuch sehr oft durchführt, jeweils die X -Werte der ausgelosten $\omega \in \Omega$ berechnet, ergibt der Durchschnittswert dieser $X(\omega)$ etwa den Erwartungswert.

Lasse dich gleich warnen: In der Regel ist der Erwartungswert kein Wert, der als $X(\omega)$ auftritt. Es braucht nicht einmal $X(\omega)$ -Werte in der Nähe des Erwartungswertes zu geben.

Definiert wird der Erwartungswert so:

2 Definition

Es sei X eine Zufallsgröße auf dem Ergebnisraum Ω . Der Erwartungswert $E(X)$ von X ist dann gegeben durch

$$E(X) := \sum_{k \in X(\Omega)} kP(X = k) = \sum_{\omega \in \Omega} X(\omega)P(\omega).$$

¹Hier muss ich eine Anmerkung machen, damit ein Kenner der Materie mich nicht für einen Trottler hält: Man lässt nur solche Funktionen X auf Ω zu, für die jedes $X^{-1}(k)$ auch eine Wahrscheinlichkeit hat. Bisher hatte jedes $A \subseteq \Omega$ auch eine Wahrscheinlichkeit $P(A)$, aber das ist allgemein nicht so. Meinetwegen kannst du das schleunigst wieder vergessen, aber ich will dir ja nichts Falsches beibringen.

Wenn X die Werte x_1, x_2, \dots, x_n annimmt, kann man die erste Summe auch in der Form

$$E(X) = \sum_{i=1}^n x_i P(X = x_i)$$

schreiben.

Anmerkung: Eigentlich kann man X nicht auf Mengen loslassen, sondern nur auf einzelne ω . Aber für $A \subseteq \Omega$ bildet man ganz unbekümmert $X(A)$ und meint damit die Menge der Bilder: $X(A) = \{X(\omega) \mid \omega \in A\}$. Das $X(\Omega)$ oben ist also die Menge aller Funktionswerte von X . In gleicher Weise kann man P nur Ereignisse anwenden, also auf Teilmengen von Ω . Wenn die Teilmenge nur ein einziges Element hat, also $A = \{\omega\}$ ist für ein $\omega \in \Omega$, spricht man von einem **Elementarereignis**, und statt $P(\{\omega\})$, wie es korrekt ist, schreibt man einfach nachlässig $P(\omega)$. Den Fachmann kratzt das nicht, aber du musst aufpassen, dass du nicht durcheinander kommst.

2.3 Linearität des Erwartungswertes

Es sei X eine Zufallsgröße auf dem Ergebnisraum Ω , zum Beispiel die Anzahl der gezogenen roten Kugeln, wenn man aus einer Urne mit zehn roten und fünf weißen Kugeln vier Kugeln auf einmal herausnimmt. Wie üblich ist dann $E(X)$ die Anzahl der roten Kugeln, die bei sehr häufiger Versuchsdurchführung im Mittel je Durchführung gezogen werden. Zahlt man dem Spieler je gezogener roter Kugel drei Euro, ist der Erwartungswert der neuen Zufallsgröße „ Y : an den Spieler gezahlter Betrag“ einfach $3E(X)$; das liegt auf der Hand. Wir notieren den Sachverhalt in formaler Sprache und beweisen ihn:

3 Satz

Es sei X eine Zufallsgröße auf dem Ergebnisraum Ω und $r \in \mathbb{R}$. Durch $\omega \mapsto rX(\omega)$ ist dann eine neue Zufallsgröße auf Ω gegeben, die wir einfach mit rX bezeichnen. Dann gilt

$$E(rX) = rE(X) \quad .$$

Beweis. Man nimmt hier am besten die Definition des Erwartungswertes, bei der über alle $\omega \in \Omega$ summiert wird. Damit ist

$$E(rX) = \sum_{\omega \in \Omega} (rX)(\omega)P(\omega) = \sum_{\omega \in \Omega} r(X(\omega))P(\omega) = r \sum_{\omega \in \Omega} X(\omega)P(\omega) = rE(X) \quad ,$$

und wir sind schon fertig. □

Kehren wir zu unserem Beispiel zurück. Wenn der Anbieter vom Spieler je Spiel einen Einsatz von einem Euro verlangt, ist es sinnvoll, die Zufallsgröße „Gewinn des Spielers“ zu betrachten. Sie ordnet jedem $\omega \in \Omega$ den Wert $X(\omega) - 1$ zu. Ihr Erwartungswert ist natürlich $E(X) - 1$.

4 Satz

Es sei X eine Zufallsgröße auf dem Ergebnisraum Ω und $a \in \mathbb{R}$. Dann ist durch $\omega \mapsto X(\omega) + a$ eine neue Zufallsgröße gegeben, die wir mit $X + a$ bezeichnen. Dann ist

$$E(X + a) = E(X) + a \quad .$$

Beweis. Nun ja, es ist

$$\begin{aligned}
 E(X + a) &= \sum_{\omega \in \Omega} (X + a)(\omega)P(\omega) \\
 &= \sum_{\omega \in \Omega} (X(\omega) + a)P(\omega) \\
 &= \sum_{\omega \in \Omega} (X(\omega)P(\omega) + aP(\omega)) \\
 &= \sum_{\omega \in \Omega} X(\omega)P(\omega) + \sum_{\omega \in \Omega} aP(\omega) \\
 &= E(X) + a \sum_{\omega \in \Omega} P(\omega) \\
 &= E(X) + a \cdot 1 = E(X) + a \quad ,
 \end{aligned}$$

und wir sind fertig. □

Man kann beide Ergebnisse auch so zusammenfassen:

$$E(rX + a) = rE(X) + a \quad \text{für } a, r \in \mathbb{R} \quad (1)$$

2.4 Die Varianz

Der Erwartungswert $E(X)$ einer Zufallsgröße misst, welchen Wert X im Mittel liefert. Die Varianz, die du nun kennenlernst, misst, wie stark die Werte von X streuen. Das grösste Streuungsmaß ist die Spannweite, das ist die Differenz zwischen dem größten und dem kleinsten Wert, den X annimmt, aber sie sagt nichts darüber, wie die Werte von X und ihre Wahrscheinlichkeiten in diesem Bereich verteilt liegen. Die Varianz ist schon etwas subtiler. Hier ist die Definition:

5 Definition

Es sei $X : \Omega \rightarrow \mathbb{R}$ eine Zufallsgröße mit Erwartungswert $\mu := E(X)$. Dann ist die **Varianz** $V(X)$ von X definiert durch

$$V(X) := \sum_{k \in X(\Omega)} (k - \mu)^2 P(X = k) \quad .$$

Die Größe $\sigma := \sqrt{V(X)}$ wird als **Standardabweichung** von X bezeichnet.

Diese Definition macht dich vermutlich zunächst einmal ratlos. Du kannst damit die Varianz einer konkret gegebenen Zufallsgröße X berechnen, und du erhältst schlicht eine Zahl. Aber es ist zunächst nicht klar, was diese Zahl bedeutet. Dazu müssen wir uns die Sache ein wenig genauer anschauen. Also, $V(X)$ ist eine Summe von Summanden der Form

$$(k - \mu)^2 P(X = k) \quad .$$

Es sei nun a eine positive Zahl. Wenn der Wert k von X um weniger als a von μ abweicht, lassen wir den Beitrag von k zur Varianz einfach weg. Überwinde dein Entsetzen ob solchen Handelns, es kommt noch eine zweite Maßnahme, und sie ist nicht weniger rabiat. Ist der Unterschied $|k - \mu|$ zwischen k und μ mindestens a , ersetzen wir ihn durch a . Der Beitrag von k zur Varianz wird dadurch höchstens kleiner:

$$(k - \mu)^2 P(X = k) \geq a^2 P(X = k) \quad \text{für } |k - \mu| \geq a.$$

Fügen wir nun alles zusammen:

$$\begin{aligned}
 V(X) &= \sum_{k \in X(\Omega)} (k - \mu)^2 P(X = k) \\
 &\geq \sum_{k \in X(\Omega), |k - \mu| \geq a} (k - \mu)^2 P(X = k) \\
 &\geq \sum_{k \in X(\Omega), |k - \mu| \geq a} a^2 P(X = k) \\
 &= a^2 \sum_{k \in X(\Omega), |k - \mu| \geq a} P(X = k) \\
 &= P(|X - \mu| \geq a)
 \end{aligned}$$

Insgesamt ergibt sich also

$$V(X) \geq a^2 P(|X - \mu| \geq a) .$$

Diese Ungleichung müssen wir nur noch durch die (positive!) Zahl a^2 dividieren, und wir stehen vor der berühmten **Ungleichung von Tschebyschew**:

$$P(|X - \mu| \geq a) \leq \frac{V(X)}{a^2} \quad \text{für } a > 0 \text{ und } \mu = E(X). \quad (2)$$

Der Term auf der rechten Seite heißt **Tschebyschew-Risiko**, er ist eine obere Schranke für die Wahrscheinlichkeit, dass der ausgeloste Wert von X den Erwartungswert μ um mindestens a verfehlt. Für $a < \sigma$ ist das Tschebyschew-Risiko größer als 1, also unbrauchbar. Aber für $a = 10\sigma$ zum Beispiel ist $V(X)/a^2 = \frac{1}{100}$, und damit kann man schon etwas anfangen. Die Tschebyschew-Ungleichung zeigt übrigens, dass es sinnvoll ist, die Abweichung von μ in der Einheit σ zu messen.

Anmerkung. Die Tschebyschew-Ungleichung gilt für jede Zufallsgröße, die überhaupt eine Varianz hat. Man kann sie nicht verbessern, wenn sie allgemein gültig bleiben soll, aber in der Regel ist das Tschebyschew-Risiko deutlich größer als die Wahrscheinlichkeit $P(|X - \mu| \geq a)$ selbst.

2.5 Bernoullis Gesetz der großen Zahl

Es sei $X : \Omega \rightarrow \mathbb{R}$ eine Zufallsgröße. Wenn wir den Zufallsversuch, der X zugrundeliegt, n -mal so wiederholen, dass die Durchführungen sich gegenseitig nicht beeinflussen, erhalten wir einen Satz von n Werten von X . Durch den X -Wert, der sich bei der k -ten Durchführung des Versuches ergibt, ist wieder eine Zufallsgröße gegeben, sozusagen eine Kopie von X , die wir mit X_k bezeichnen wollen. Wir bilden die Zufallsgrößen

$$S := \sum_{k=1}^n X_k \quad \text{und} \quad \bar{X} := \frac{1}{n} \sum_{k=1}^n X_k = \frac{1}{n} S .$$

Damit ich zuversichtlich sein kann, dass du die Konstruktion verstanden hast, schiebe ich einige Zwischenfragen ein:

1. Nimm an, der Zufallsversuch bestehe darin, dass ein Würfel geworfen und die gewürfelte Augenzahl ermittelt wird. Wir würfeln zehnmal. Was ist dann X_7 ? Wie sieht der Ergebnisraum aus, auf dem S und die X_k definiert sind? Gib ein Ergebnis ω und $X_7(\omega)$ an.
2. Was bedeutet das \bar{X} ?
3. Wir ziehen aus unserer Standardurne dreimal ohne Zurücklegen und zählen die Anzahl der gezogenen roten Kugeln. Hier liegt **keine** unabhängige Wiederholung eines Zufallsversuches vor, denn was in der zweiten Ziehung geschieht, wird vom Ergebnis der ersten Ziehung beeinflusst.

4. Nimm einen Bernoulliversuch und wiederhole ihn n -mal unabhängig. Das ergibt natürlich eine Bernoulli-Kette. Was bedeuten S , \bar{X} und die X_k ?

5. Es sei $\mu = E(X)$. Was ist dann $E(X_k)$, $E(S)$, $E(\bar{X})$?

Unter den beschriebenen Bedingungen² gilt

$$V(S) = V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k) = \sum_{k=1}^n V(X) = nV(X) \quad .$$

Ferner ist

$$V(\bar{X}) = V\left(\frac{1}{n}V(S)\right) = \frac{1}{n^2}V(S) = \frac{1}{n^2}nV(X) = \frac{1}{n}V(X) \quad .$$

Wir schreiben nun die Tschebyschew-Ungleichung für \bar{X} hin:

$$P(|\bar{X} - \mu| \geq a) \leq \frac{V(\bar{X})}{a^2} = \frac{V(X)}{na^2}$$

Das Tschebyschew-Risiko auf der rechten Seite strebt für $n \rightarrow \infty$ offensichtlich gegen Null. Das heißt, dass die Wahrscheinlichkeit, dass der Durchschnittswert \bar{X} der X -Werte den Erwartungswert μ von X mindestens um ein (noch so kleines positives) a verfehlt, strebt für n gegen unendlich gegen Null. Anders gesagt, der Durchschnittswert je Durchführung der X -Werte liegt für große n nahe am Erwartungswert von X . Eben dies ist die Aussage des Bernoullischen Gesetzes der großen Zahl. Ich formuliere es noch einmal explizit.

6 Satz (Bernoullis Gesetz der großen Zahl)

Es sei $X_1, X_2, \dots, X_n, \dots$ unabhängige Kopien der Zufallsvariablen X mit dem Erwartungswert μ und der Varianz $V(X)$. Dann gilt

$$P\left(\left|\frac{1}{n}\sum_{k=1}^n X_k - \mu\right| \geq a\right) \xrightarrow{n \rightarrow \infty} 0$$

für jedes $a > 0$.

Wir betrachten einen beliebigen Zufallsversuch mit Ergebnisraum Ω . Für ein festes $\omega_0 \in \Omega$ setzen wir

$$X(\omega) = \begin{cases} 1 & \text{für } \omega = \omega_0 \\ 0 & \text{sonst} \end{cases} \quad ,$$

dann ist $\mu = E(X) = P(\{\omega\}) =: p$. Wie oben führen wir den Zufallsversuch sehr oft durch. Die Zufallsgröße S zählt dann, wie oft ω_0 bei n Durchführungen des Versuches eingetreten ist, und die Zufallsgröße \bar{X} gibt an, wie oft das im Mittel je Durchführung geschehen ist. Der Wert von \bar{X} wird dann für große n nahe bei p liegen.

Wenn dich das an das empirische Gesetz der großen Zahl erinnert, ist das völlig in Ordnung. Wenn du allerdings denkst, hier sei nichts Neues, liegst du falsch. Das empirische Gesetz der großen Zahl ist ein empirisches Gesetz und somit eine Erfahrungstatsache. Hier haben wir festgestellt, dass sich die ausgelosten Werte unserer Zufallsgrößen, also innermathematische Konstruktionen, genauso verhalten wie die Werte in Beobachtungsreihen, und das ist überaus bemerkenswert. Weil das funktioniert, kann man überhaupt erst erwarten, dass mathematische Modelle von Zufallsercheinungen brauchbare Einsichten liefern.

Praktisch ist der Apparat, den wir aufgebaut haben, allerdings nur sehr eingeschränkt brauchbar. Das liegt daran, dass das Tschebyschew-Risiko in der Regel deutlich größer ist als die Wahrscheinlichkeit, die es abschätzen soll. Die n , die man auf diese Weise berechnet, sind in aller Regel viel zu groß.

²Hier benutzen wir Regeln für die Varianz, nämlich $V(aX) = a^2V(X)$ und $V(X + Y) = V(X) + V(Y)$. Die erste Regel gilt stets, und der Beweis ist nicht schwierig. Die zweite Regel hat als Voraussetzung, dass X und Y unabhängig von einander sind. Es wäre ein ziemlicher Aufwand, wenn wir das jetzt genauer untersuchen wollten.

2.6 Die Näherungsformeln von de Moivre und Laplace und ihr Gebrauch

In Anwendungen interessiert man sich häufig für Wahrscheinlichkeiten bei binomialverteilten Zufallsgrößen. Zwar können wir Terme für die exakte Wahrscheinlichkeit hinschreiben, aber es ist nur für kleine n möglich, die Wahrscheinlichkeiten wirklich auszurechnen. Für große n haben wir uns mit der Ungleichung von Tschebyschew beholfen, aber anstelle der Wahrscheinlichkeit selbst bekommen wir dann das Tschebyschew-Risiko, und das überschätzt die gesuchte Wahrscheinlichkeit in aller Regel deutlich.

Du sollst nun eine Methode lernen, für $B(n, p)$ -verteiltes X bei großen n Wahrscheinlichkeiten zu berechnen. Dazu standardisiert man X , das heißt, man geht von X über zur Zufallsgröße

$$Z := \frac{X - \mu}{\sigma} .$$

Es ist dann $E(Z) = 0$ und $V(Z) = 1$, und in der Tat sehen die Histogramme der Z ziemlich gleich aus, egal, bei welchem X man gestartet ist. Sie werden sehr gut durch die Gaußsche Glockenkurve angenähert, die du schon von früher kennst, und man berechnet die gesuchte Wahrscheinlichkeit, sagen wir, $P(X \leq k)$, mit Hilfe eines Integrals über die Gauß-Funktion.

Grundlage ist der folgende Satz.

7 Satz (Grenzwertsatz von de Moivre und Laplace)

Es sei X $B(n, p)$ -verteilt, $Z = \frac{X - \mu}{\sigma}$ und $b \in \mathbb{R}$. Dann gilt

$$P(Z \leq b) \xrightarrow{n \rightarrow \infty} \int_{-\infty}^b \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx := \int_{-\infty}^b \varphi(x) dx =: \Phi(b) .$$

Wir machen keinerlei Versuche, uns dem Beweis den Satzes zu nähern. Praktisch benutzt man den Satz so, dass man die gesuchte Wahrscheinlichkeit, die man ja als Flächeninhalt im Histogramm von Z sehen kann, einfach mit Hilfe der Gauß-Funktion

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

berechnet. Das ist zulässig, wenn sich das Histogramm von Z und der Graph der Gauß-Funktion φ genügend ähnlich sind, und das ist der Fall, wenn die Varianz npq von X groß genug ist: Eine in der Praxis bewährte Faustformel sagt, dass man mit der Näherung bedenkenlos arbeiten kann, wenn

$$npq \gtrsim 9 \tag{3}$$

ist. Das Zeichen \gtrsim sagt dabei, dass die Zahl 9 keineswegs eine scharfe Grenze ist; wenn npq etwas kleiner als 9 ist, ist das nicht dramatisch.

Anmerkung und Beispiele

1. Berechnung von $P(X \leq k)$

Es sei X $B(500, \frac{1}{5})$ -verteilt. Wir suchen $P(X \leq 110)$: Es ist

$$P(X \leq 110) = P\left(Z \leq \frac{110 - 100}{\sqrt{80}}\right) \approx \Phi\left(\frac{10}{\sqrt{80}}\right) \approx \Phi(1.12) \approx 0.8686 .$$

2. Die Stetigkeitskorrektur

Schauen wir uns die Histogramme genauer an, erkennen wir, dass wir bei dieser Rechnung das letzte halbe Kästchen (auf $k = 110$) einfach abgeschnitten haben. Wenn es nur wenige Kästchen sind und das letzte Kästchen vielleicht auch noch in der Nähe des Erwartungswertes liegt, kann das zu deutlichen Fehlern führen. Will man dies vermeiden, muss man mit 110.5 rechnen:

$$P(X \leq 110) = P\left(Z \leq \frac{110.5 - 100}{\sqrt{80}}\right) \approx \Phi(1.17) \approx 0.879 .$$

3. Vergleich mit exakten Werten

Zur Kontrolle: Maple gibt als dezimalen Näherungswert für $P(X \leq 110)$ an:

$$\sum_{k=0}^{110} \binom{500}{k} \left(\frac{1}{5}\right)^k \left(\frac{4}{5}\right)^{500-k} \approx 0.8790248638 .$$

4. Berechnung von $P(a \leq X \leq b)$

Berechnen wir noch $P(90 \leq X \leq 110)$:

$$\begin{aligned} P(90 \leq X \leq 110) &\approx \Phi\left(\frac{110.5 - 100}{\sqrt{80}}\right) - \Phi\left(\frac{89.5 - 100}{\sqrt{80}}\right) \approx \\ &\approx \Phi(1.17) - \Phi(-1.17) \approx 0.879 - 0.121 = 0.758 \end{aligned}$$

Also ist die Wahrscheinlichkeit, dass sich der beobachtete Wert von X vom Erwartungswert um höchstens zehn unterscheidet, etwa 0.758.

5. Berechnung von $P(X = k)$

Es sei nun X $B(n, p)$ -verteilt und $k \in \{0, 1, \dots, n\}$. Dann ist

$$P(X = k) = P(X \leq k) - P(X \leq k - 1) \approx \int_{\frac{k-\mu-1/2}{\sigma}}^{\frac{k-\mu+1/2}{\sigma}} \varphi(x) dx \approx \frac{1}{\sigma} \varphi\left(\frac{k-\mu}{\sigma}\right) .$$

Konkret für unser $B(500, \frac{1}{5})$ -verteiltes X von oben erhalten wir zum Beispiel

$$P(X = 100) \approx \frac{1}{\sigma} \varphi\left(\frac{100 - \mu}{\sigma}\right) = \frac{1}{\sqrt{80}} \varphi(0) \approx 0.0446031 .$$

Als dezimalen Näherungswert für $P(X = 100)$ gibt Maple 0.04456409 an, der mit φ bestimmte Wert passt also ganz gut.

6. Bestimme k minimal so, dass $P(X \leq k) \geq 1 - \alpha$ ist

Beispiel: Bestimme für unser $B(500, \frac{1}{5})$ -verteiltes X eine möglichst kleine Zahl k so, dass $P(X \leq k) \geq 0.95$ ist.

Wir bilden

$$\Phi^{-1}(0.95) \approx 1.645 .$$

Das gesuchte k finden wir nun, indem wir die Standardisierung rückwärts ausführen:

$$1.645 \longrightarrow \mu + 1.645\sigma = 100 + 1.645\sqrt{80} \approx 114.712$$

Das heißt: Für k können wir den Wert 114 nehmen, dann ist $P(X \leq k)$ allerdings eine Idee kleiner als 0.95. Soll $P(X \leq k) \geq 0.95$ streng gelten, müssen wir $k = 115$ nehmen.

Was haben wir nun erreicht? Wir können vorhersagen, dass man höchstens 115 Erfolge beobachten wird, wenn man den Versuch 500-mal durchführt – bei einer Irrtumswahrscheinlichkeit α von höchstens 0.05.

7. Bestimmung von n

Wenn man oft würfelt, sollte der Anteil der Fünfen in der Nähe von $\frac{1}{6}$ liegen. Das sagen wir so, aber was heißt hier „oft“? Auf diese Frage können wir nun eine brauchbare Antwort geben, wenn wir sie in der folgenden Weise präzisieren: Wie oft muss man würfeln, damit der Anteil der Fünfen höchstens mit der Wahrscheinlichkeit $\alpha = 0.05$ um $a = 0.1$ oder mehr von $\frac{1}{6}$ abweicht?

Wir bezeichnen die Anzahl der Fünfen bei n Würfeln eines ordentlichen Würfels mit X . Dann ist X $B(n, p)$ -verteilt mit $p = \frac{1}{6}$, und verlangt ist, dass

$$P(|X - np| \geq na) \leq \alpha$$

sein soll. Wir formen dies um zu

$$P(X < np + na) \geq 1 - \frac{1}{2}\alpha .$$

Dies heißt aber, dass

$$\frac{na}{\sqrt{np(1-p)}} \geq \Phi^{-1}\left(1 - \frac{1}{2}\alpha\right)$$

sein soll, folglich

$$na \geq \sqrt{np(1-p)}\Phi^{-1}\left(1 - \frac{1}{2}\alpha\right) .$$

Wir dividieren durch $a\sqrt{n}$ und quadrieren, dann haben wir das Ergebnis:

$$n \geq p(1-p) \left(\frac{1}{a}\Phi^{-1}\left(1 - \frac{1}{2}\alpha\right)\right)^2$$

Für unsere konkreten Daten ergibt sich $n \geq 37.6$. Nehmen wir $n = 38$. Die Varianz $np(1-p)$ ist nur etwa 5, also nach der Faustformel zu klein. Schauen wir uns an, ob das Ergebnis brauchbar ist: Es sei X $B(38, \frac{1}{6})$ -verteilt. Dann ist $E(X) = \frac{38}{6} = 6\frac{1}{3}$ und $38 \cdot 0.1 = 3.8$. Eine exakte Rechnung ergibt $P(3 \leq X \leq 10) \approx 0.922$, das ist in Ordnung.

3 Konfidenzintervalle

3.1 Einführung

„Jeder Bolzen aus einer Serienfertigung ist mit der Wahrscheinlichkeit p unbrauchbar, ...“ – so fängt manche Aufgabe im Rechenbuch an, und du sollst dann irgendetwas dazu ausrechnen. Das kannst du ja auch, die Anzahl der unbrauchbaren Bolzen unter n zufällig gewählten Bolzen ist eine $B(n, p)$ -verteilte Zufallsgröße, und über solche wissen wir eine ganze Menge. Das Rechnen ist nicht das Problem, aber woher weiß der Verfasser des Schulbuches denn das p ? Die Wahrscheinlichkeit einer fünf beim Würfeln können wir uns ausdenken, aber die Wahrscheinlichkeit, dass ein Butterbrot mit der beschmierten Seite auf den Boden fällt, dass ihr in der nächsten Stunde alle da seid, dass ein zufällig gewählter Bolzen defekt ist, kann man sich nicht ausdenken, die muss man versuchen zu messen.

Eine solche Messung führt man durch, indem man eine Stichprobe zieht. Man entnimmt der Produktion zehn oder hundert oder tausend Bolzen und prüft sie, und man findet unter den geprüften Bolzen meinetwegen zwei oder zwanzig oder zweihundert defekte Bolzen. Dann wird man jeweils $\frac{1}{5}$ als Schätzwert für die gesuchte Wahrscheinlichkeit p nehmen, was soll man auch anderes tun. Aber natürlich wird man dem durch die Prüfung von den tausend Bolzen gewonnenen Wert mehr Vertrauen entgegenbringen als dem durch die Prüfung von zehn gewonnenen.

Den Schätzwert darf man nicht als scharfen Wert ansehen, man meint eigentlich eher, dass das p in der Nähe des Schätzwertes liegen sollte. Aber Nähe ist ein dehnbare Begriff, der muss noch präzisiert werden. Um zu zeigen, wie das geht, brauche ich nur auszuführen, was Matthias vorgeschlagen hat:

Man setzt eine Wahrscheinlichkeit α fest, die sogenannte **Irrtumswahrscheinlichkeit**. Dann bestimmt man zu jedem p einen möglichst kleinen Radius $a = a(p)$ so, dass

$$P(|X_p - np| \geq a) \leq \alpha$$

ist; X_p steht dabei für eine $B(n, p)$ -verteilte Zufallsgröße. Alle p , für die der beobachtete Wert k in dieser Umgebung liegt, bilden das **Konfidenzintervall** für p zur Irrtumswahrscheinlichkeit α . Dieses Intervall gibt man als Ergebnis der Schätzung an.

Eine kritische Würdigung des Ergebnisses wird später noch erfolgen, jetzt schauen wir uns nur den Mechanismus an. Das $a(p)$ ist nicht schwierig auszurechnen, es ist

$$a(p) = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \sqrt{np(1-p)} .$$

Demnach liegt ein ausgeloster Wert einer $B(n, p)$ -verteilten Zufallsgröße X_p mindestens mit der Wahrscheinlichkeit $1 - \alpha$ in dem Intervall $[np - a(p), np + a(p)]$. Dieses Intervall ist für jedes p in den folgenden Diagrammen dargestellt.

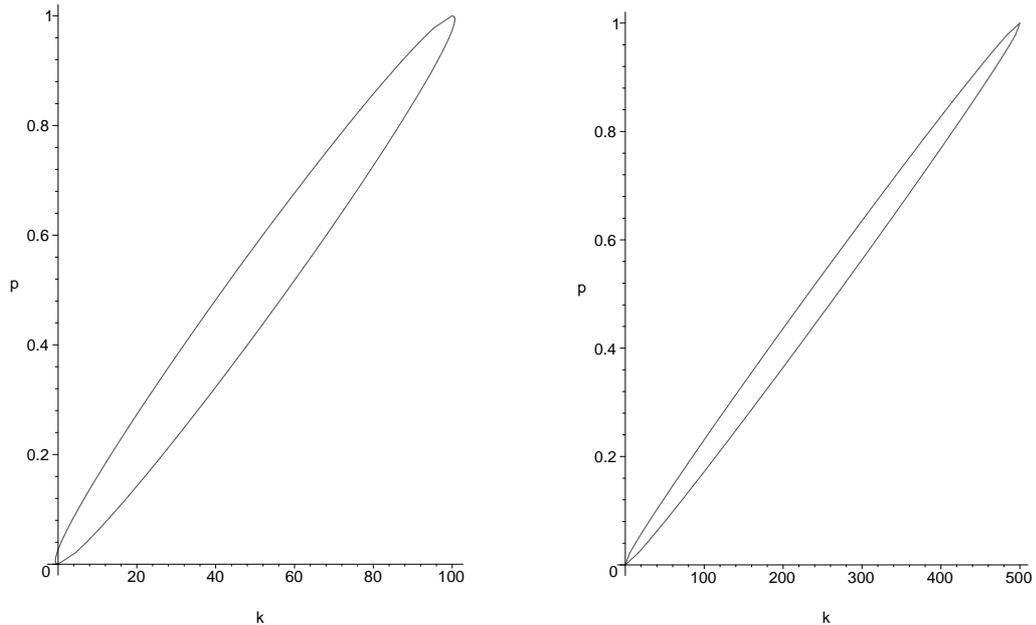


Abbildung 1: Zur Bestimmung von Konfidenzintervallen zur Irrtumswahrscheinlichkeit $\alpha = \frac{1}{10}$, links für $n = 100$, rechts für $n = 500$

An der Hochachse ist p angetragen, an der Rechtsachse der beobachtete Ausgang k von X_p . Nur mit der Wahrscheinlichkeit α wird ein k ausgelost, für das der Punkt (k, p) außerhalb des ellipsenförmigen Gebietes liegt, mit der Wahrscheinlichkeit $1 - \alpha$ jedoch ein k , für das (k, p) im Innern des Gebietes liegt. Zum Konfidenzintervall gehören alle p , für die (k, p) im Innern des Gebietes liegt.

Man bestimmt das Konfidenzintervall zur Irrtumswahrscheinlichkeit α und dem beobachteten Wert k praktisch so: Senkrecht über dem beobachteten Wert k liegen zwei Punkte der Kurve. Der y -Wert p_{max} des oberen Punktes ist der größte p -Wert, für den k noch in der $1 - \alpha$ -Umgebung des Erwartungswertes liegt:

$$k = np_{max} - a(p_{max})$$

Der y -Wert p_{min} des unteren Punktes ist der kleinste p -Wert, für den k noch in der $1 - \alpha$ -Umgebung des Erwartungswertes liegt:

$$k = np_{min} - a(p_{min})$$

Das gesuchte Konfidenzintervall ist das Intervall $[p_{min}; p_{max}]$.

Wie zu erwarten war, ist das Gebiet für $n = 500$, das im rechten Bild dargestellt ist, relativ schmaler als das Gebiet im linken Bild, das zu $n = 100$ gehört. Je größer der Stichprobenumfang n ist, desto kleiner ist das Konfidenzintervall bei gleicher beobachteter relativer Häufigkeit k/n .

3.2 Ein durchgerechnetes Beispiel

Im nordrhein-westfälischen Kultusministerium wird überlegt, ob an den Schulen des Landes Schuluniformen eingeführt werden sollen. Vorsichtig, wie man ist, fragt man erst einmal nach der Meinung im Lande zu dieser Frage. Alle Einwohner zu fragen, ist nicht praktikabel, deshalb zieht man eine Stichprobe. Nun, das Ziehen einer Stichprobe ist eine hohe Kunst; man muss Sorge tragen, dass alle Bevölkerungsschichten angemessen berücksichtigt werden. Es wäre zum Beispiel ein grober Fehler, befragte man nur alle Soldaten in einigen Kasernen, obwohl das recht einfach zu machen wäre. Gehen wir also davon aus, dass $n = 800$ Leute ordentlich befragt wurden und dass $k = 312$ Befragte für Schuluniformen waren. Als Irrtumswahrscheinlichkeit nehmen wir $\alpha = 0.05$. Nun müssen wir p_{min} und p_{max} so bestimmen, dass

$$np_{min} + a(p_{min}) = k = np_{max} - a(p_{max})$$

ist. Das läuft darauf hinaus, dass man die quadratische Gleichung

$$(k - np)^2 = (\Phi^{-1}(1 - \frac{\alpha}{2}))^2 np(1 - p)$$

lösen muss. Das ist nicht tiefsinnig, aber leider sehr fehlerträchtig, deshalb nimmt man gern näherungsweise für $a(p_{min})$ und $a(p_{max})$ einfach $a(k/n)$; das ist in der Regel gut genug. Das Konfidenzintervall mit diesen Grenzen heißt **Näherungskonfidenzintervall**. Hier erhalten wir folgende Ergebnisse:³

Näherungskonfidenzintervall: [0.3562; 0.4238], echtes Konfidenzintervall: [0.3568; 0.4243]

3.3 Aufgaben

1. Studiere den Text zum Konfidenzintervall genau.
2. Unter 500 Befragten gaben 30 an, die Partei ABC wählen zu wollen. Dagegen wollten 260 die Partei DEF wählen. Bestimme Konfidenzintervalle zur Irrtumswahrscheinlichkeit 0.1 für die Anteile der Wähler der beiden Parteien in der Gesamtbevölkerung.
3. Was ist die kleinste Wahrscheinlichkeit, für die der beobachtete Wert 30 unter 500 noch in der 99%–Umgebung des Erwartungswertes liegt?
4. Wie groß ist die Wahrscheinlichkeit, 30 oder noch mehr ABC–Wähler unter 500 Befragten anzutreffen, wenn der Anteil der ABC–Wähler in der Gesamtbevölkerung genau 5% ist?
5. Die Parteivorstände der beiden Parteien wollen ihren Wähleranteil auf einen halben Prozentpunkt genau wissen. Wieviele Leute muss man fragen?
6. „Wir sollten herausfinden, welcher Anteil der Bevölkerung für Schuluniformen ist. Unsere Umfrage ergab als Konfidenzintervall zur Irrtumswahrscheinlichkeit 1% für den Anteil der Befürworter das Intervall [0.47; 0.49].“ So doziert der junge Consultant bei der Präsentation. Was bedeutet seine Aussage? Formuliere deine Antwort so, dass sie im Kurs akzeptabel ist – für deine Mitschüler und für mich.

3.4 Ein Beispiel für die Bestimmung der notwendigen Stichprobenlänge

n

Die ABC ist eine Partei, deren Stimmenanteil so um 6% herumkrebst. Die nächste Wahl steht vor der Tür, und der Vorstand will wissen, wie die Partei in der Gunst der Bevölkerung steht. Ein Meinungsforschungsinstitut wird beauftragt, den aktuellen Anteil der ABC–Wähler zu messen.

³In „stochastik.mws“ findest du Hilfsmittel, mit denen du die nötigen Rechnungen sehr bequem erledigen kannst.

Wie du weißt, kann das Institut nur ein Konfidenzintervall bestimmen; mehr ist nicht drin, wenn man nicht alle Wähler befragen will, und das wäre unbezahlbar. Der Vorstand möchte ein möglichst kleines Konfidenzintervall mit einer möglichst kleinen Irrtumswahrscheinlichkeit, das ist verständlich, aber beide Wünsche treiben die notwendige Stichprobenlänge und damit die Kosten der Erhebung in die Höhe. Bleiben wir auf dem Boden und geben wir als Irrtumswahrscheinlichkeit $\alpha = 5\%$ und als Radius des Konfidenzintervalls 0.01 vor. Welche Stichprobenlänge ist dann erforderlich?

Die Grenzen des Konfidenzintervalls bezeichnen wir mit p_{min} und p_{max} . Der Durchmesser des Intervalls ist dann⁴

$$\begin{aligned} p_{max} - p_{min} &= \frac{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sigma(p_{min}) + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sigma(p_{max})}{n} \\ &= \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \frac{\sqrt{np_{min}(1-p_{min})} + \sqrt{np_{max}(1-p_{max})}}{n} \\ &= \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \frac{\sqrt{p_{min}(1-p_{min})} + \sqrt{p_{max}(1-p_{max})}}{\sqrt{n}} . \end{aligned}$$

Dieser Wert soll höchstens 0.02 sein. Das führt auf die Ungleichung

$$n \geq \left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \frac{(\sqrt{p_{min}(1-p_{min})} + \sqrt{p_{max}(1-p_{max})})}{0.02} \right)^2 .$$

Dumm ist nur, dass n vor der Untersuchung festlegen muss, man p_{min} und p_{max} aber erst durch die Untersuchung erhält. Was nun? Es gibt zwei Möglichkeiten:

1. Man geht auf Nummer sicher. Es gilt

$$\sqrt{p(1-p)} \leq \frac{1}{2} . \tag{4}$$

Damit bekommt man

$$n \geq \left(\frac{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}{0.02} \right)^2 \approx 6763.86 ,$$

das ist sehr viel.

2. Man benutzt das Vorwissen $p \approx 0.06$ und setzt näherungsweise

$$\sqrt{p_{min}(1-p_{min})} + \sqrt{p_{max}(1-p_{max})} \approx 2\sqrt{0.06(1-0.06)} .$$

Dies führt zu

$$n \geq \left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \frac{2\sqrt{0.06(1-0.06)}}{0.02} \right)^2 \approx 1525.93 .$$

In der Praxis wird man wohl mit dem zweiten Wert arbeiten.

4 Tests

4.1 Einführungsbeispiel

Der Schulkonferenz liegt folgender Antrag zur Beschlussfassung vor: Wer bei zwanzig Kontrollen öfter als sechsmal negativ auffällt, muss die Schule verlassen. Schätze jeder seine persönliche

⁴Wenn dir nicht klar ist, dass man durch n dividieren muss, fragst du am besten Christian. Er erklärt dir, dass man es mit zwei Skalen zu tun hat: der der Merkmalsachse und der der relativen Häufigkeiten.

Wahrscheinlichkeit p , bei einer Kontrolle zu einem zufällig gewählten Zeitpunkt Grund zur Beanstandung zu geben – wegen nicht gemachter Hausaufgaben oder wegen mangelnder Vorbereitung oder fehlender Mitarbeit oder ähnlichen Delikten. Anhand von Abbildung 2 auf Seite 18 kann man sich dann über die Aussichten informieren, in einem halben Jahr noch an der Schule zu sein, wenn der Antrag beschlossen wird. Sie zeigt den Graphen der Funktion

$$f : p \mapsto P(X_p > 6) \quad \text{für } B(20, p)\text{-verteiltes } X_p ;$$

$f(p)$ ist also die Wahrscheinlichkeit, dass ein Schüler mit der persönlichen Wahrscheinlichkeit p nächstens die Schule verlassen muss.

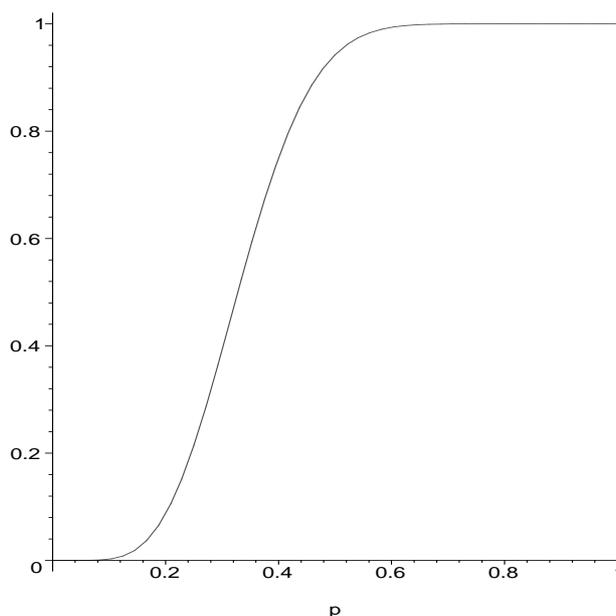


Abbildung 2: Graph der Funktion $f : p \mapsto P(X_p > 6)$ für $B(20, p)$ -verteiltes X_p

Was du hier vor dir hast, ist ein einfacher **Test**. Getestet wird die **Hypothese** H_0 , dass der gerade begutachtete Schüler brav seine Pflicht tut. Die **Entscheidungsregel** des Tests besagt, dass diese Hypothese verworfen wird, wenn der Schüler bei zwanzig Kontrollen mehr als sechsmal auffällt. Wenn du einem Schüler die persönliche Tunichtgutquote $p = 0.2$ zubilligen willst, siehst du vielleicht mit Schrecken, dass der Schüler doch mit der Wahrscheinlichkeit 9% der Schule verwiesen wird. Ein Bösewicht mit der Quote $p = 0.6$ bleibt dagegen nur etwa mit der Wahrscheinlichkeit 0.6% an der Schule.

4.2 Schema eines Tests

Zwar gefällt mir, dass ihr Testprobleme souverän und kreativ angeht, aber es gibt eine Reihe von Fachbegriffen, die ihr unbedingt beherrschen müsst. Ich erkläre sie an dem folgenden **Beispiel**: Sagen wir, der Wähleranteil einer Partei liege so bei 40%. Nun hat die Partei eine neue Spitzenkandidatin, und sie behauptet, durch sie sei der Wähleranteil der Partei wesentlich gestiegen. Ihre Behauptung soll durch einen Test überprüft werden.

Die Sache läuft so ab:

1. Man formuliert zwei Hypothesen über den Wähleranteil p . Die eine ist die sogenannte **Nullhypothese**, sie wird mit H_0 bezeichnet:

$$H_0 : p \leq 0.4$$

Die zweite ist die Gegenhypothese

$$H_1 : p > 0.4 .$$

2. Dann legt man eine **Irrtumswahrscheinlichkeit** α fest, häufig ist $\alpha = 0.05$.
3. Drittens braucht man eine **Entscheidungsregel**, die sagt, wie der Test laufen soll. Sie könnte hier so aussehen: Wenn in einer Zufallsstichprobe der Länge n mehr als g Wähler der Partei angetroffen werden, verwirft man H_0 zugunsten von H_1 und glaubt der Kandidatin. Andernfalls wird die Behauptung der Kandidatin nicht durch den Test „nachgewiesen“.

Es wird von dir erwartet, dass du zu gegebenem H_0 und α eine passende Entscheidungsregel konstruierst. Dazu müssen n und g festgelegt werden. Wir gehen jetzt davon aus, dass auch n schon vorgegeben ist, meinetwegen $n = 300$; dann ist nur noch g zu bestimmen. Das geschieht auf folgende Weise: Mit X_p bezeichnen wir die Anzahl der Erfolge in einer Bernoullikette der Länge n mit Erfolgswahrscheinlichkeit p . Wenn H_0 gilt, also $p \leq 0.4$ ist, ist

$$P(X_p > g)$$

die Wahrscheinlichkeit, dass H_0 zu Unrecht verworfen wird. Diese Wahrscheinlichkeit heißt **Risiko erster Art**. Du musst ein möglichst kleines g so bestimmen, dass dieses Risiko höchstens α ist:

$$P(X_p > g) \leq \alpha \quad \text{für alle } p \leq 0.4$$

Nun ist $P(X_p > g) \leq P(X_{0.4} > g)$, wie du dir anhand der Histogramme plausibel machen kannst (man kann es aber auch streng beweisen). Deshalb muss ein möglichst kleines g so bestimmt werden, dass

$$P(X_{0.4} > g) \leq \alpha$$

ist. Mit Hilfe der Gaußfunktion findet man

$$\begin{aligned} g &\geq \sigma \Phi^{-1}(1 - \alpha) + \mu - \frac{1}{2} \\ &\approx 1.645 \sqrt{300 \cdot 0.4 \cdot 0.6} + 120 - 0.5 \\ &\approx 133.457 \quad , \end{aligned}$$

das rechne ich hier nicht Schritt für Schritt vor. Das Ergebnis besagt: Wenn mehr als 134 der 300 zufällig ausgewählten befragten Wahlberechtigten Anhänger der Partei sind, wird H_0 verworfen und die Behauptung der Kandidatin gilt als statistisch bewiesen (Zusatz für Fachleute: auf dem Signifikanzniveau $\alpha = 0.05$). Übrigens heißt die Menge $K := \{g+1, g+2, \dots, n\}$ der **Ablehnungsbereich** des Tests. Die restlichen Ergebnisse $1, 2, \dots, g$ sind die Elemente des Annahmebereiches.

4.3 Test-FAQ

Hier liste ich in loser Folge Fragen auf, die uns beschäftigt haben. Bei den Antworten greife ich häufig auf das Beispiel des vorigen Abschnitts zurück.

1. Gibt es auch ein Risiko zweiter Art?

Na klar, sonst hätte das Risiko erster Art keinen besonderen Namen. Es mag ja sein, dass H_1 gilt, also in unserem Beispiel $p > 0.4$ ist, aber H_0 dennoch nicht verworfen wird. Die Wahrscheinlichkeit, die man betrachten muss, ist $P(X_p \leq g)$ mit $p > 0.4$. Man bestimmt nun ein möglichst kleines β so, dass

$$P(X_p \leq g) \leq \beta \quad \text{ist für alle } p > 0.4 .$$

Dieses β ist das Risiko zweiter Art des Tests.

2. Wieso braucht man eigentlich zwei Hypothesen?

Man hat eine Menge zulässiger Parameterwerte, im Beispiel ist p ein Element des Intervalls $[0; 1]$. Diese Menge wird in zwei Teile geteilt, nämlich in die der $p > 0.4$ und die der $p \leq 0.4$. So kommen die beiden Teile zustande; du kannst die Hypothesen H_0 und H_1 sogar als Mengen von p 's ansehen.

Das war ein formales Argument. Ich füge ein zweites Argument hinzu. Es gilt, die Behauptung $p > 0.4$ durch Erhebung einer Stichprobe zu testen. Der Test macht das so, dass er prüft, ob sich das Ergebnis der Stichprobe mit der Annahme verträgt, das Gegenteil der Behauptung sei richtig. Nun ja, die Behauptung ist H_1 , ihr Gegenteil H_0 .

3. Herrscht Symmetrie zwischen H_0 und H_1 ?

Nein. Der Test behandelt H_0 und H_1 verschieden. Das Risiko, H_0 zu Unrecht zu verwerfen, wird durch α gedeckelt, und diese Setzung legt, bei gegebenem n , den Ablehnungsbereich fest. Das Risiko zweiter Art hat dabei keinen Einfluss. Man kann es ausrechnen und anschauen; wenn man es zu groß findet, ist der Test nicht geeignet.

Vertauscht man H_1 und H_0 , erhält man als Ablehnungsbereich des neuen Tests keineswegs des Annahmebereich des alten Tests, sondern eine viel kleinere Menge.

4. Ist H_0 bewiesen, wenn der Test nicht zur Verwerfung von H_0 führt?

Keineswegs. Man stellt nur fest, dass man H_0 nicht verwirft. Wenn der beobachtete Wert von X nur wenig größer ist als $0.4n$, wird man H_0 nicht verwerfen können. Aber wer wird eine beobachtete relative Häufigkeit $X/n > 0.4$ als Bestätigung der Hypothese $p \leq 0.4$ ansehen wollen? Wenn freilich X/n deutlich kleiner als 0.4 ist, wird man vielleicht neu überlegen, aber das gehört nicht mehr zu dem alten Test.

5. Kann ich nicht zum Stichprobenergebnis das Konfidenzintervall ausrechnen?

Natürlich kannst du das. Und es sagt gegebenenfalls nicht nur, dass man tatsächlich von $p > 0.4$ ausgehen sollte, sondern es liefert auch gleich Informationen, wo man das neue p vermuten sollte. Aber der Aufwand ist auch größer. Wenn zum Beispiel regelmäßig eingehende Ware geprüft werden soll, ist ein Test mit fester Entscheidungsregel praktischer.

5 Stetig verteilte Zufallsgrößen

5.1 Einführung

Die Zufallsgrößen, die wir bisher behandelt haben, nennt der Fachmann diskret. Sie nahmen in der Regel nur endlich viele Werte an, und wenn es unendlich viele Werte gab, lagen sie isoliert von einander. Niemals füllten die Werte, die X annehmen konnte, ein ganzes Intervall. Aber dieser Fall ist in Anwendungen durchaus wichtig, wie einige Beispiele zeigen mögen.

1. Es sei X eine zufällig gewählte Zahl aus dem Intervall $[0; 2]$.
2. Es sei X die Lebensdauer eines zufällig gewählten Teilchens einer radioaktiven Substanz.
3. Aus einer Serienfertigung von Nägeln wird ein Nagel zufällig gewählt und seine Länge gemessen. Es sei X die Länge des Nagels in Millimetern. Wenn die Nägel die Länge 90 mm haben sollen, sollte der Wert von X in der Nähe von 90 liegen.
4. Im Physikunterricht wird die Schallgeschwindigkeit gemessen. Es sei X die gemessene Geschwindigkeit in Metern pro Sekunde.

Natürlich kommt man in der Praxis immer mit endlich vielen Werten aus, aber das können sehr viele Werte sein, und dann ist die Vorstellung einer zufällig gewählten Zahl wie in Beispiel 1 durchaus auch von praktischem Interesse.

Beginnen wir mit Beispiel 1, da fehlt jegliches Beiwerk. Alle Zahlen sollen gleichberechtigt sein, das heißt gleich möglich sein: Es ist

$$P(X = 1) = P(X = \frac{3}{4}) = P(X = \frac{5}{7}) = P(X = \sqrt{2}) = P(X = \frac{1}{e}) =: p .$$

Es folgt sofort, dass $p = 0$ sein muss, denn es gibt ja unendlich viele Zahlen. Das sieht nicht gut aus, aber man darf die Flinte nicht vorschnell ins Korn werfen. Es ist nämlich für $0 \leq a \leq b \leq 2$

$$P(a \leq X \leq b) = \frac{b-a}{2} .$$

Damit haben wir diesen Fall eigentlich ganz gut im Griff.

Gehen wir zum dritten Beispiel. Wenn wir X angemessen beschreiben wollen, zeichnen wir eine Kurve $y = f(x)$, die wie eine transformierte Gaußkurve $y = \varphi(x)$ aussieht. Der Gipfel sollte bei 90 sein, und wenn die Produktion etwas taugt, sollte sie sehr schmal und hoch sein. Natürlich ist dann der Wert 90 vor 89.9 bevorzugt, aber dennoch ist $P(X = 90) = P(X = 89.9) = 0$. Aber die Wahrscheinlichkeit, dass die Länge des Nagels in einem sehr kleinen Intervall der Länge Δx um 89.9 liegt, ist kleiner als die Wahrscheinlichkeit, dass dieser Wert in einem Intervall der gleichen Länge um 90 liegt. Das legt nahe, die Wahrscheinlichkeit, dass X einen Wert in dem kleinen Intervall der Länge Δx um eine Zahl x annimmt, durch $f(x)\Delta x$ zu messen. Zu Ende gedacht, kommt man zu

$$P(a \leq X \leq b) := \int_a^b f(x) dx . \quad (5)$$

Damit das funktioniert, muss

$$f(x) \geq 0 \text{ sein für alle } x \text{ und } \int_{-\infty}^{\infty} f(x) dx = 1 \text{ gelten.} \quad (6)$$

Die Zufallsgröße X heißt dann **stetig verteilt mit der Dichtefunktion f** .

Die Zufallsgröße des ersten Beispiels hat die Dichtefunktion

$$f(x) = \begin{cases} \frac{1}{2} & \text{für } 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

und die des zweiten Beispiels so etwas wie

$$f(x) = \begin{cases} e^{-x} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases} .$$

5.2 Die Standardnormalverteilung

Da kommt es knüppeldick, gleich „standard“ und „normal“ soll sie sein. Nun, du triffst eigentlich einen alten Bekannten: Die Zufallsgröße X mit der Dichtefunktion

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

heißt (standard) normalverteilt. Du hast ihren Graphen schon als Näherung für die Histogramme der Standardisierungen binomial verteilter Zufallsgrößen benutzt.

5.3 Erwartungswert und Varianz einer stetig verteilten Zufallsgröße

An dieser Stelle gebe ich nur die Formeln an: Für eine stetig verteilte Zufallsgröße X mit Dichtefunktion f ist

$$\mu = E(X) := \int_{-\infty}^{\infty} xf(x) dx \quad (7)$$

und

$$\sigma^2 = V(X) := \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \quad (8)$$

Du erkennst unschwer die Verwandtschaft zu den vertrauten Formeln.

6 Klausuren

6.1 Klausur am 7. März 2007

- Wir werfen einen Würfel und bilden die Zufallsgröße

$$X : \frac{60}{\text{gewürfelte Augenzahl}} \quad .$$

- Berechne $E(X)$ und schreibe hin, was dieser Wert bedeutet.
- Kunz behauptet ohne erst zu rechnen, $E(X)$ sei $\frac{60}{3,5}$. Wie kommt er darauf? Ist sein Schluss richtig?
- Es sei

$$Y : \frac{1}{\text{gewürfelte Augenzahl}} \quad .$$

Kann man ohne neue Rechnung $E(Y)$ angeben, wenn man $E(X)$ schon berechnet hat?

- Hinz, Kunz und Tobias sitzen beim Skat. Gerade hat Tobias die 32 Karten verteilt; vor jedem liegen zehn Karten, zwei Karten liegen verdeckt in der Mitte.
 - Mit welcher Wahrscheinlichkeit ist die erste Karte, die Tobias aufnimmt, eine Kreuzkarte?
 - Mit welcher Wahrscheinlichkeit liegt mindestens eine Kreuzkarte in der Mitte?
 - Jetzt schaut Tobias seine Karten an, und er nimmt auch die zwei aus der Mitte auf. Unter seinen zwölf Karten ist kein einziger Bube. Es wäre sehr fatal für Tobias, wenn einer der beiden Mitspieler alle vier Buben auf der Hand hätte. Mit welcher Wahrscheinlichkeit ist das der Fall?
 - Wenn du die vorige Teilaufgabe gerechnet hast, hast du als Ergebnis eine Zahl erhalten, meinetwegen $1/3$. Sag mal: Was bedeutet die Zahl überhaupt, und was kann Tobias mit der Zahl anfangen? Entweder hat ein anderer Spieler alle Buben auf der Hand oder nicht, und es steht sogar schon fest, ob es so ist oder nicht?!
 - Hinz ist nun des Skatspielens müde, und Tobias muss ohnehin nach Hause. Da schlägt Hinz dem Kunz ein neues Spiel vor: Kunz zieht aus dem gemischten Kartenstapel einzeln Karten, so viele er will. Solange er keine Kreuzkarte zieht, will Hinz an ihn zahlen, und zwar einen Euro für die erste gezogene Karte, zusätzlich zwei für die zweite, zusätzlich drei für die dritte, und so weiter. Kunz darf so oft ziehen, wie er will, mindestens aber dreimal. Zieht er irgendwann eine Kreuzkarte, verfällt sein Gewinn.
 - Wieviel zahlt Hinz dem Kunz im Mittel, wenn Kunz dreimal zieht?
 - Sagen wir, Kunz hat n -mal erfolgreich gezogen, bisher war kein Kreuz dabei. Berechne, wieviel Geld Hinz dem Kunz jetzt zahlen muss und ob es für Kunz ratsam ist, ein weiteres Mal zu ziehen.

3. Die 30 lieben Heranwachsenden der Klasse 10a stöhnen, dass ihre Physikbücher so schwer sind. Der gutherzige Physiklehrer hat ein Einsehen, und er ordnet an, dass immer jeder zweite sein Buch mitbringen soll, damit jeder in ein Buch schauen kann. Er meint, dann müssten ja im Mittel immer 15 Bücher vor Ort sein, und selten weniger als acht. Es klappt natürlich überhaupt nicht, es sind nie mehr als zehn Bücher da.
 - (a) Der Lehrer hatte an eine Bernoullikette gedacht. Beschreibe den Bernoulliversuch präzise, definiere die natürliche Zufallsgröße X und schreibe Terme für $P(X = 15)$ und für $P(X \leq 7)$ hin.
 - (b) Berechne den Wert von $P(X = 15)$ und schätze $P(X \leq 7)$ ab, indem du die Tschebyschew-Ungleichung geschickt anwendest.
 - (c) Was könnte der Lehrer anordnen, damit das Einpacken der Bücher zu Hause durch die Schüler tatsächlich als Bernoullikette abläuft?
 - (d) Schreibe die Formel für $P(X = k)$ für eine $B(n, p)$ -verteilte Zufallsgröße hin und erkläre, wie sie zustande kommt.
4. Aus unserer Urne mit fünf roten, drei weißen und zwei grünen Kugeln werden drei Kugeln ohne Zurücklegen gezogen.
 - (a) Es sei X die Anzahl der gezogenen roten Kugeln. Bestimme die Wahrscheinlichkeitsfunktion von X und zeichne ein Histogramm.
 - (b) Bilden wir ein bequemes Ω zu dem Versuch, und zwar nehmen wir die Menge aller zulässigen Wörter abc mit $a, b, c \in \{r, w, g\}$. Nun ist unser X eine Abbildung $X : \Omega \rightarrow \mathbb{R}$. In gleicher Weise könnten wir als Y : „Anzahl der gezogen gelben Kugeln“ eine zweite Zufallsgröße auf demselben Ω bilden, und anschließend deren Summe $X + Y : \Omega \rightarrow \mathbb{R}$. Das ist dann wieder eine Zufallsgröße auf Ω . Sage, im Beispiel und allgemein, was $(X + Y)(\omega)$ für $\omega \in \Omega$ bedeutet, und beweise allgemein, dass $E(X + Y) = E(X) + E(Y)$ ist.

6.2 Nachschreibklausur

1. Hinz, Kunz und Tobias sitzen wieder beim Skat. Das Spiel hat 32 Karten, jeder bekommt zehn und zwei kommen in die Mitte. Es folgt eine Reihe von Fragen zum Skatspiel. Als Antwort genügt jeweils ein durchsichtiger Rechenausdruck, du brauchst die Werte nicht auszurechnen.
 - (a) Wie wahrscheinlich ist es, dass Kunz als erste Karte das Herz As aufnimmt?
 - (b) Kunz überlegt, bevor er beginnt, seine Karten aufzunehmen, ob er wohl damit rechnen kann, (genau) zwei Buben zu bekommen. Wie wahrscheinlich ist das?
 - (c) Nun hat Kunz schon sechs Karten aufgenommen, und es ist noch kein Bube dabei. Wie wahrscheinlich ist es denn jetzt, dass er noch seine zwei Buben bekommt?
 - (d) Es ist wieder neu gegeben worden. Wie wahrscheinlich ist es, dass Tobias drei Buben, sechs Kreuz und das Herz As bekommt?
 - (e) Wie wahrscheinlich ist es, dass Tobias weder Buben noch Kreuz auf der Hand hat?
 - (f) Wie wahrscheinlich ist es, dass Tobias höchstens zwei der sieben Herzkarten bekommt?
 - (g) Tobias hat auch die zwei Karten in der Mitte bekommen. Er schaut auf sein Blatt: er hat zwei Buben. Wie wahrscheinlich ist es, dass Hinz oder Kunz die beiden anderen Buben haben?

- (h) Während einer Spielpause überlegen die drei, wieviele verschiedene Verteilungen der Karten es wohl gibt, weil sie wissen wollen, ob man damit rechnen muss, dass sich genau die gleiche Verteilung wiederholt. Folgende Terme sind im Gespräch:

$$32! \quad \frac{32!}{10! \cdot 10! \cdot 10! \cdot 2!} \quad \binom{32}{10} \cdot \binom{22}{10} \cdot \binom{12}{10} \cdot \binom{2}{2} \quad 32^{32}$$

Beurteile diese Ansätze.

- (i) Es sei nun X die Anzahl der Buben unter den beiden Karten in der Mitte. Berechne $E(X)$ und zeichne das Histogramm und den Graphen der Funktion F mit $F(x) = P(X < x)$.
2. Arne und Björn spielen ein Würfelspiel. Sie würfeln abwechselnd, Arne beginnt. Arne gewinnt, wenn er eine fünf oder eine sechs wirft. Björn gewinnt, wenn er höchstens eine vier wirft. Mit welcher Wahrscheinlichkeit gewinnt Arne das Spiel?
3. Wie wahrscheinlich ist es, dass man höchstens 150 oder mindestens 250 Fünfen bekommt, wenn man 1200-mal würfelt? Schreibe einen Term für die Wahrscheinlichkeit hin und gib eine möglichst gute obere Schranke für die Wahrscheinlichkeit an.
4. „Die Wahrscheinlichkeit, dass es außerirdisches Leben gibt, liegt bei 14%“, tönt Kunz, der Simpel. Was sagst du als Kenner des Begriffs der Wahrscheinlichkeit dazu? Ist das völliger Blödsinn, oder ist etwas daran zu retten? Aber sei nicht zu grob zu Kunz.
5. Es sei X eine Zufallsgröße mit der Varianz $V(X)$, und es seien a und b reelle Zahlen. Beweise, dass $V(aX) = a^2V(X)$ und dass $V(X + b) = V(X)$ ist.

6.3 Klausur am 2. Mai 2007

1. Binomial verteilte Zufallsgrößen

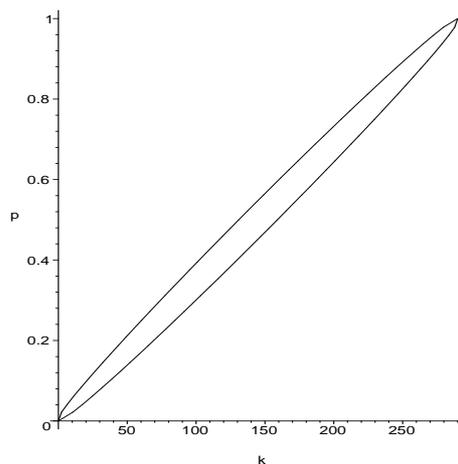
- (a) Gib ein konkretes Beispiel für eine binomial verteilte Zufallsgröße.
- (b) Es sei X $B(800, \frac{1}{4})$ -verteilt. Berechne:
- $P(X = 0)$
 - $P(X \leq 205)$
 - $P(X > 195)$
 - ein möglichst kleines k so, dass $P(X \leq k) \geq \frac{9}{10}$ ist;
 - ein möglichst kleines a so, dass $P(|X - \mu| \geq a) \leq \frac{1}{20}$ ist.
- (c) Es sei $p = \frac{1}{3}$. Wie groß sollte n mindestens sein, damit man bedenkenlos mit der Näherungsformel von de Moivre und Laplace rechnen kann?
- (d) Es sei X $B(2n, \frac{1}{2})$ -verteilt. Gib mit Hilfe der Gauß-Funktion φ einen Term für $P(X = n)$ an. Wohin strebt der Term für $n \rightarrow \infty$? Was sagt das über das Histogramm von X ?
- (e) Ein Hof ist mit einhundert quadratischen Platten gleicher Größe gepflastert. Bei einem leichten Regen fallen eintausend Tropfen auf den Hof. Fasse den Vorgang in geeigneter Weise als Zufallsversuch auf und gib eine Formel für $P(X = k)$ für die Zufallsgröße X : Anzahl der Tropfen, die auf die erste Platte fallen.
- (f) Es sei X_p $B(100, p)$ -verteilt. Skizziere, wie du dir den Graphen der Funktion

$$f : p \mapsto P(X \leq 20)$$

vorstellst.

2. Konfidenzintervalle 1

- (a) Bei einer Umfrage der SV, ob die Aktion zum Valentinstag auch im nächsten Jahr durchgeführt werden soll, stimmten 73 von 120 Befragten mit „Nein“. Berechne das Näherungskonfidenzintervall für den Anteil der Valentinsmuffel unter den Schülern zur Irrtumswahrscheinlichkeit $\alpha = \frac{1}{10}$.
- (b) Hier siehst du unsere Konfidenzellipse zu $n = 290$ zu einer bestimmten Irrtumswahrscheinlichkeit α . Lies das Konfidenzintervall zum beobachteten Wert $k = 50$ ab und notiere, was du an der Kurve beispielsweise zum p -Wert $\frac{1}{2}$ ablesen kannst. Wie groß mag α sein?



3. Konfidenzintervalle 2

Hinz simuliert auf dem Rechner eine $B(100; \frac{3}{10})$ -verteilte Zufallsgröße X . Zum beobachteten Wert k berechnet er sich jeweils das echte Konfidenzintervall zur Irrtumswahrscheinlichkeit $\alpha = 0.05$ und schaut nach, ob $\frac{3}{10}$ im Intervall liegt.

- (a) Mit welcher Wahrscheinlichkeit überdeckt das Intervall zum nächsten ausgelosten Wert k von X das wahre p ?
- (b) Hinz bestimmt 160 Intervalle. Wieviele davon sollten das wahre $p = \frac{3}{10}$ enthalten?
- (c) Hinz bestimmt fünf Intervalle. Mit welcher Wahrscheinlichkeit enthalten sie alle das wahre p ?
- (d) Wie viele Intervalle darf Hinz höchstens bilden, wenn sie mit 50%-iger Wahrscheinlichkeit alle das wahre p enthalten sollen?
- (e) Kunz schaut herein. Auf dem Bildschirm sieht er die Ausgabe $[0.22661, 0.31828]$. Aha, ein Konfidenzintervall. Er erkundigt sich nach der Irrtumswahrscheinlichkeit und trompetet: „In diesem Intervall liegt das wahre p mit der Wahrscheinlichkeit 0.95.“ Hinz seufzt. Schon wieder dieser Blödsinn. Er bittet dich, Kunz behutsam zu belehren. Was sagst du zu Kunz?
- (f) Hinzens Rechner gibt Hinz auf die Eingabe „rand(0..20000)()“; eine Zufallszahl zwischen 0 und 20000. Wie kann Hinz damit sein X simulieren?

6.4 Nachschreibklausur am 30. Mai 2007

1. Test (praktisch)

Durch die beschlossene Verkürzung der Schulzeit auf zwölf Jahre fallen mehr Wochenstunden an, und die müssen irgendwie im Plan untergebracht werden. Deshalb wurde vorgeschlagen, wieder Samstagsunterricht einzuführen. Nun soll festgestellt werden, ob es in der Schulgemeinde eine Mehrheit für Samstagsunterricht gibt. Dazu wird eine Umfrage unter

120 Mitgliedern der Schulgemeinde durchgeführt, und du sollst die statistische Auswertung übernehmen. Keine Sorge, Ivo, dich betrifft das nicht mehr; du kannst ruhig korrekt rechnen.

- (a) Die SV will die Hypothese

$$H_0 : p \leq \frac{1}{2}$$

über den Anteil der Befürworter von Samstagsunterricht in der Schulgemeinde testen. Die Irrtumswahrscheinlichkeit α soll 10% betragen. Konstruiere die Entscheidungsregel.

- (b) Die Elternvertretung will mit dem gleichen α die Hypothese

$$H_0 : p > \frac{1}{2}$$

testen. Stelle auch dafür die Entscheidungsregel auf.

- (c) Vergleiche die beiden Entscheidungsregeln. Was lassen die Wahlen der Hypothesen über die Motive der beiden Gruppen vermuten? Begründe deine Antwort.
- (d) Berechne das Risiko zweiter Art beim Test der Schüler und schreibe hin, was es genau bedeutet.
- (e) Berechne den p -Wert $\pi(69)$ des Umfrageergebnisses „69 Befürworter von Samstagsunterricht unter 120 Befragten“ aus Sicht der Schüler.
- (f) Berechne das Konfidenzintervall zum Ergebnis 69 zur Irrtumswahrscheinlichkeit $\alpha = 5\%$ und schreibe hin, was es bedeutet.
- (g) Die Schulkonferenz, die die Entscheidung über Samstagsunterricht trifft, hatte früher 24 stimmberechtigte Mitglieder. Nehmen wir an, vier von ihnen seien definitiv für Samstagsunterricht. Mit welcher Wahrscheinlichkeit stimmt die Mehrheit für Samstagsunterricht, wenn jeder der übrigen Stimmberechtigten mit der Wahrscheinlichkeit $\frac{1}{2}$ „dafür“ stimmt?

2. Test (theoretisch)

Um eine Hypothese zu testen, führt das Institut für Empirische Sozialforschung zahlreiche Tests durch. Wir gehen einmal davon aus, dass die Hypothese stimmt, sie aber trotzdem mit der Wahrscheinlichkeit $\frac{1}{20}$ verworfen wird.

- (a) Wie groß ist die Wahrscheinlichkeit, dass die Hypothese bei keinem einzigen von zwanzig durchgeführten Tests verworfen wird?
- (b) Wieviele Tests muss man höchstens durchführen, bis man mit mindestens 99-prozentiger Wahrscheinlichkeit einmal ein Testergebnis hat, das zur Verwerfung der Hypothese führt?
- (c) Tendenziell liest du nur von Untersuchungsergebnissen in der Zeitung, bei denen ein Test zur Verwerfung der Hypothese führte. Kommentiere.
- (d) Es sei X_p eine $B(100, p)$ -verteilte Zufallsgröße. Kunz sagt, dass man zwar $P(X_p > 15)$ für einen konkreten Wert von p berechnen kann, dass man aber über

$$P(X_p > 15) \quad \text{für } 0 \leq p \leq 0.1$$

nichts sagen könne. Was sagst du dazu?

3. Gütefunktion

In Abbildung 3 auf Seite 27 siehst du den Graphen der Gütefunktion $f : p \mapsto P(X_p < g)$ eines Tests.

- (a) Was kannst du an der Gütefunktion ablesen?
- (b) Wie ändert sich die Kurve, wenn man die Stichprobenlänge n und g beide verdoppelt? Es ist nur eine qualitative Antwort verlangt, du brauchst nichts zu rechnen.
- (c) Die Kurve hier habe ich mit dem Wert $n = 120$ berechnet. Wie groß war g ? Lege auch offen, wie du die Antwort gefunden hast.

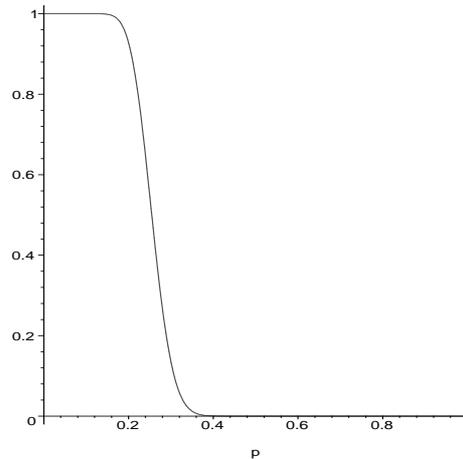


Abbildung 3: Gütefunktion eines Tests

7 Komplexe Zahlen

Weil im Kurs Interesse bestand, versuche ich, euch in aller Kürze einen Eindruck davon zu geben, was es mit komplexen Zahlen auf sich hat.

7.1 Die Gaußsche Zahlenebene

Du weißt, dass es eine enge Beziehung zwischen reellen Zahlen und den Punkten auf der Zahlengeraden gibt: zu jedem Punkt der Zahlengeraden gibt es genau eine reelle Zahl und umgekehrt. Das ist ein Glück, denn die Zahlengerade ist ein wesentlich anschaulicheres Objekt als die Menge der reellen Zahlen, die man zum Beispiel als Menge von Äquivalenzklassen von Intervallschachtelungen abgeschlossener Intervalle mit rationalen Endpunkten konstruieren könnte – das ist in etwa so mühsam, wie es sich anhört. Zum Glück gibt es auch für die Menge \mathbb{C} der komplexen Zahlen ein geometrisches Objekt, das sie darstellt: die Menge der Punkte der Ebene. **Zu jedem Punkt der Ebene gehört genau eine komplexe Zahl und umgekehrt.**

Punkte der Ebene erfassen wir mit Hilfe von Koordinaten. In der Regel verwenden wir ein kartesisches Koordinatensystem mit x - und y -Achse, die einen rechten Winkel bilden. Dann gehört zu jedem Punkt P der Ebene ein Zahlenpaar (x, y) :

$$P \longleftrightarrow (x, y)$$

Das kennst du gut. Man kann P auch mit Hilfe von **Polarkoordinaten** beschreiben: Man gibt die Entfernung r an, die P vom Ursprung hat, und den Winkel φ , den der Strahl, der im Nullpunkt beginnt und der durch P geht, mit der positiven x -Achse bildet. Am besten stellst du dir vor, die positive x -Achse werde gegen den Uhrzeigersinn gedreht, bis sie durch P geht. Der Winkel, den sie dabei überstreicht, ist φ .

Wenn man zwei Darstellungen von P hat, muss man wissen, wie sie zusammenhängen. Hier gilt offensichtlich Folgendes:

$$x = r \cos(\varphi) \tag{9}$$

$$y = r \sin(\varphi) \tag{10}$$

$$r = \sqrt{x^2 + y^2} \tag{11}$$

7.2 Summe und Produkt komplexer Zahlen

Komplexe Zahlen bezeichnen wir mit kleinen griechischen Buchstaben, dabei vermeiden wir allerdings φ . Da wir \mathbb{C} nur in der Gestalt der Ebene vor uns haben, schreiben wir α und β un-

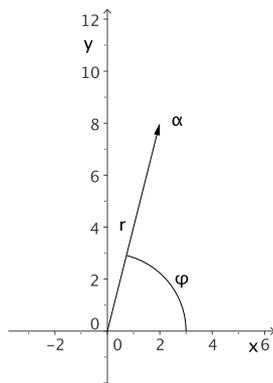


Abbildung 4: $\alpha \longleftrightarrow (r, \varphi)$

bekümmert an zwei Punkte der Ebene. Was wir jetzt wissen müssen, ist, wie man mit komplexen Zahlen rechnet. Für die Summe $\alpha + \beta$ von α und β ist das schnell gesagt: Wenn $\alpha \longleftrightarrow (a_1, a_2)$ und $\beta \longleftrightarrow (b_1, b_2)$ ist, dann ist

$$\alpha + \beta \longleftrightarrow (a_1 + b_1, a_2 + b_2) .$$

Man addiert also einfach die Koordinatenpaare komponentenweise. Geometrisch gesprochen läuft das auf die Vektoraddition hinaus, die dir in der Mittelstufe beim Kräfteparallelogramm begegnet sein mag.

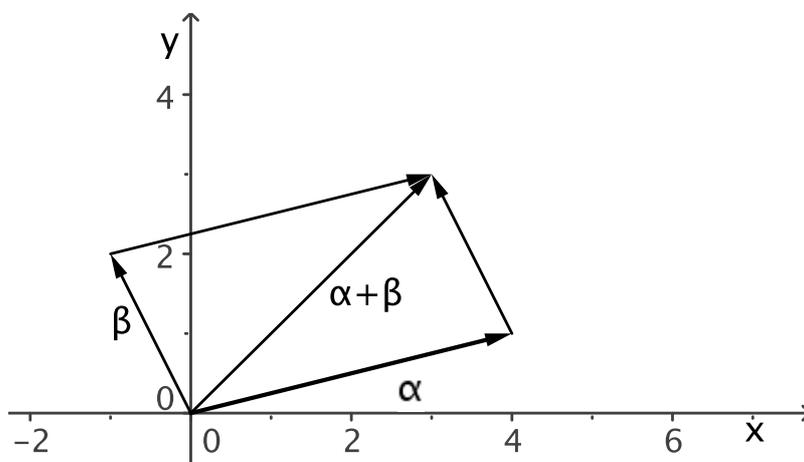


Abbildung 5: Summe komplexer Zahlen

Mit dem Produkt ist das schon etwas schwieriger. Hier benutzt man am besten Polarkoordinaten. Wenn $\alpha_1 \longleftrightarrow (r_1, \varphi_1)$ und $\alpha_2 \longleftrightarrow (r_2, \varphi_2)$ zwei komplexe Zahlen sind, ist das Produkt gegeben durch

$$\alpha_1 \alpha_2 \longleftrightarrow (r_1 r_2, \varphi_1 + \varphi_2) .$$

Man multipliziert also die r -Werte und addiert die φ -Werte.

7.3 Die Darstellung $\alpha = a_1 + a_2 i$

Es sei $\alpha \in \mathbb{C}$ und $\alpha \longleftrightarrow (a_1, a_2)$. Dann schreibt man

$$\alpha = a_1 + a_2 i$$

Dabei ist i die sogenannte **imaginäre Einheit**. Mit diesen Termen rechnest du nach den üblichen Regeln, die du früher gelernt hast, und du siehst dabei das i als eine Variable an, wie x etwa. Das heißt, du kannst etwa so rechnen:

$$\alpha + \beta = (a_1 + a_2i) + (b_1 + b_2i) = a_1 + a_2i + b_1 + b_2i = a_1 + b_1 + (a_2 + b_2)i .$$

Mit Hilfe unserer Festlegung der Multiplikation überlegst du dir, dass

$$i^2 = -1 \tag{12}$$

ist, und damit ist

$$\alpha\beta = (a_1 + a_2i)(b_1 + b_2i) = a_1b_1 + a_1b_2i + a_2b_1i + a_2b_2i^2 = a_1b_1 - a_2b_2 + (a_1b_2 + a_2b_1)i .$$

7.4 Die Darstellung $\alpha = re^{i\varphi}$

Aus der Analysis kennst du die Taylorreihen der Exponentialfunktion, der Sinusfunktion und der Kosinusfunktion; zum Beispiel ist

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Der große Euler hat im Term der Taylorreihe der Exponentialfunktion für x einfach $i\varphi$ eingesetzt, das heißt, er hat formal $e^{i\varphi}$ gebildet.⁵ Wir haben Eulers Rechnung nachvollzogen und sind zu dem Ergebnis gekommen, dass

$$e^{i\varphi} = \cos(\varphi) + i \sin(\varphi) \tag{13}$$

ist. Damit können wir die komplexe Zahl $\alpha \longleftrightarrow (r\varphi)$ so schreiben:

$$\alpha = r \cos(\varphi) + r \sin(\varphi)i = r(\cos(\varphi) + \sin(\varphi)i) = re^{i\varphi} \tag{14}$$

Die etwas gewöhnungsbedürftige Festlegung der Multiplikation komplexer Zahlen steht dann in völligem Einklang mit den Potenzgesetzen:

$$r_1 e^{i\varphi_1} r_2 e^{i\varphi_2} = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}$$

7.5 Konjugiert komplexe Zahl und Division

Für $\alpha = a + bi \in \mathbb{C}$ ist die konjugiert komplexe Zahl $\bar{\alpha}$ definiert als

$$\bar{\alpha} = \overline{a + bi} := a - bi . \tag{15}$$

Geometrisch gesprochen erhält man $\bar{\alpha}$ zu α durch Spiegelung an der reellen Achse.

Die Abbildung $\zeta \mapsto \bar{\zeta}$ verträgt sich ausgezeichnet mit der Addition und der Multiplikation in \mathbb{C} . Es gilt nämlich

$$\overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta} \quad \text{und} \quad \overline{\alpha \cdot \beta} = \bar{\alpha} \cdot \bar{\beta} \quad \text{für alle } \alpha, \beta \in \mathbb{C} . \tag{16}$$

Ferner ist

$$\alpha \bar{\alpha} = (a + bi) \cdot \overline{a + bi} = a^2 + b^2 \in \mathbb{R} \quad \text{für alle } \alpha \in \mathbb{C}, \tag{17}$$

und damit erhält man bequem den Kehrwert einer komplexen Zahl $\alpha \neq 0$, die in der Form $\alpha = a + bi$ gegeben ist:

$$\frac{1}{\alpha} = \frac{\bar{\alpha}}{\alpha \bar{\alpha}} = \frac{a}{a^2 + b^2} + \frac{b}{a^2 + b^2}i$$

⁵Sei gewarnt: Wenn du Schritte solcher Kühnheit unternimmst, geht das mit ziemlicher Sicherheit schief. Als wir die Taylorreihe entwickelt haben, war von komplexen Zahlen keine Rede. Dass die Reihe auch für komplexe Argumente Vernünftiges ergibt, liegt keineswegs auf der Hand.

7.6 Übungen

1. Schreibe um in die jeweils andere Darstellung und zeichne.

$$3 + 4i \quad 2 - 2i \quad 2 + 2i \quad 2e^{i\frac{\pi}{3}}$$

2. Berechne.

$$(5 - 4i) + (5 + 4i) \quad (2 - i)^2 \quad i^{126} \quad (3 + 2i)(3 - 2i) \quad (1 + i)^3 \quad (1 - i)^{24} \quad e^{i\pi}$$

3. Bestimme alle komplexen Zahlen α mit $\alpha^6 = 1$.
4. Stelle $\alpha_n = (1 + i)^n$ für $n = 0, 1, 2, 3, \dots$ graphisch dar.
5. Löse die Gleichungen $x^2 + 1 = 0$ und $x^2 + 6x + 13 = 0$.
6. Berechne e^α für $\alpha = 2 - 3i$.
7. Berechne $\frac{1}{3-4i}$ und $\frac{1+i}{1-2i}$.
8. Begründe, dass $\overline{e^{i\varphi}} = e^{-i\varphi}$ ist.

7.7 Vom Nutzen der komplexen Zahlen

Dass man mit komplexen Zahlen faszinierende Dinge anstellen kann, ist für Mathematiker Grund genug, sich damit zu beschäftigen. Aber komplexe Zahlen sind auch für Anwender interessant.⁶ Es wäre halber Kram, versuchte ich nicht, euch etwas davon erfahrbar zu machen. Eine erste Portion gibt es sofort, eine zweite etwas später.

Ein Merkmal von \mathbb{C} ist die **Vollständigkeit**. Damit ist folgendes gemeint: Du kennst Gleichungen n -ten Grades, also so etwas:

$$\sum_{k=0}^n a_k x^k = 0 \quad \text{mit } a_n \neq 0$$

Du bist solchen Gleichungen schon begegnet; die Koeffizienten a_k waren reelle Zahlen, und die Gleichungen hatten zwischen 0 und n (reelle) Lösungen. In \mathbb{C} ist die Sache klar und einfach:⁷

8 Satz (Fundamentalsatz der Algebra)

Jede Gleichung n -ten Grades mit Koeffizienten in \mathbb{C} hat genau n (nicht notwendig verschiedene) komplexe Lösungen $\alpha_1, \alpha_2, \dots, \alpha_n$. Insbesondere lässt sich die linke Seite der Gleichung stets als Produkt von Linearfaktoren schreiben:

$$\sum_{k=0}^n a_k x^k = a_n (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$$

Nun sind die reellen Zahlen in den komplexen enthalten; daraus ergibt sich, dass jede unserer Gleichungen n -ten Grades mit reellen Koeffizienten in \mathbb{C} immerhin n Lösungen hat, von denen keine bis alle reell sein können.

Wo wir nun schon hier sind, will ich dir folgenden kleinen Satz nicht vorenthalten:

⁶Jänichs Buch „Analysis für Physiker und Ingenieure“ besteht aus drei großen Kapiteln, und das erste davon handelt von Funktionentheorie, also von der Analysis von Funktionen $\xi = f(\zeta)$, bei denen Argument und Funktionswert komplex sind.

⁷Dieser Satz wurde von Gauß bewiesen. Ein Beweis ist für uns nicht grundsätzlich unerreichbar, Johannes hat sich in seiner Facharbeit damit beschäftigt, aber es wäre ein ziemlich weiter Weg.

9 Satz

Es sei

$$p(x) = \sum_{k=0}^n a_k x^k = 0$$

eine Gleichung n -ten Grades mit reellen Koeffizienten, und es sei $p(\alpha) = 0$ für ein $\alpha \in \mathbb{C}$. Dann ist auch $p(\bar{\alpha}) = 0$.

Beweis Beachte, dass für jede Zahl $\beta \in \mathbb{C}$ dann und nur dann $\bar{\bar{\beta}} = \beta$ ist, wenn β reell ist. Damit ist

$$0 = \bar{0} = \overline{\sum_{k=0}^n a_k \alpha^k} = \sum_{k=0}^n \overline{a_k} \cdot \bar{\alpha}^k = \sum_{k=0}^n a_k \bar{\alpha}^k, \quad ,$$

fertig. □

Dass man sich darauf verlassen kann, dass Gleichungen n -ten Grades auch Lösungen haben, ist unglaublich wichtig. Das folgende kleine Beispiel mag dir das zeigen:

Physiker haben oft mit Differentialgleichungen zu tun, dann suchen sie zum Beispiel⁸ eine Lösung $y = y(x)$ von

$$y'' + ay' + y = 0 \quad \text{mit festen reellen Zahlen } a \text{ und } b.$$

Eine Funktion

$$y = e^{\lambda x}$$

ist genau dann Lösung der Differentialgleichung, wenn

$$\lambda^2 + a\lambda + b = 0$$

ist.⁹ Reell gesehen findet man so nur Lösungen der Differentialgleichung, wenn $a^2 - 4b \geq 0$ ist. Lässt man $\lambda \in \mathbb{C}$ zu, findet man immer welche, auch zum Beispiel für $a = 2$ und $b = 5$. Wir besprechen das im Unterricht genauer.

Das Beispiel ist einigermaßen typisch. Geht man von \mathbb{R} zu \mathbb{C} über, werden die Verhältnisse in der Regel übersichtlicher.

7.8 Funktionen $f : \mathbb{C} \rightarrow \mathbb{C}$

Wir sind gewohnt, uns Funktionen mit Hilfe von Graphen zu veranschaulichen. Mit Hilfe eines Koordinatensystems geht das nicht mehr, weil sowohl der Definitionsbereich als auch der Wertebereich Flächenstücke sind. Deshalb geht man so vor: Man nimmt eine Ebene für den Definitionsbereich und eine Ebene für den Wertebereich. Dann markiert man ein Stück des Definitionsbereiches, zum Beispiel ein Rechteck, und zeichnet ein Gitternetz ein. In einer zweiten Ebene stellt man dann dar, was die Funktion aus diesem Netz macht. Wir schauen uns das an einigen Beispielen an:

1. $f(\zeta) = \zeta + \alpha$ für ein festes $\alpha \in \mathbb{C}$ (siehe Abbildung 6)
Geometrisch ist das eine Verschiebung der Ebene um α , wenn man sich α als Pfeil vorstellt.
2. $f(\zeta) = \alpha \cdot \zeta$ für ein festes $\alpha \in \mathbb{C}$, $\alpha \neq 0$
Wenn du α in der Form $re^{i\varphi}$ schreibst, siehst du, dass f eine Drehung der Ebene um den Winkel φ um den Nullpunkt und anschließend eine zentrische Streckung mit dem Faktor r bewirkt. Die Wirkung auf das Einheitsquadrat ist in Abbildung 7 dargestellt.

⁸Differentialgleichung eines gedämpften harmonischen Oszillators.

⁹Das kannst du selbst nachrechnen: setze einfach $e^{\lambda x}$ für y in die Differentialgleichung ein.

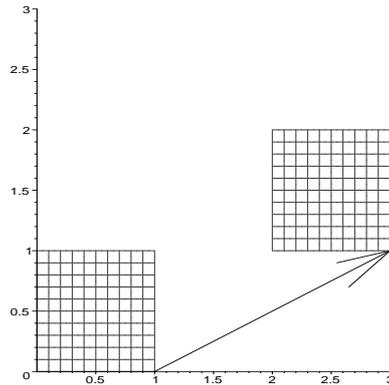


Abbildung 6: Wirkung von $f(\zeta) = \zeta + \alpha$ mit $\alpha = (2 + i)$ auf das Einheitsquadrat

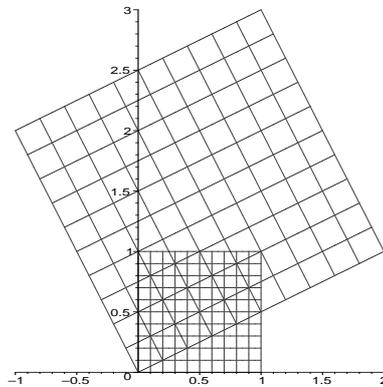


Abbildung 7: Wirkung von $f(\zeta) = \alpha \cdot \zeta$ mit $\alpha = (2 + i)$ auf das Einheitsquadrat

3. $f(\zeta) = \frac{1}{\zeta}$

Da die Funktion an der Stelle 0 nicht definiert ist, ist in Abbildung 8 die Wirkung der Funktion auf das Quadrat mit der linken unteren Ecke $\frac{1}{2}$ und der rechten oberen Ecke $\frac{3}{2} + i$ dargestellt. Du verstehst, was die Funktion bewirkt, wenn du dir überlegst, wie sie auf $\zeta = re^{i\varphi}$ wirkt:

$$f : re^{i\varphi} \mapsto \frac{1}{r}e^{-i\varphi}$$

Die Strahlen vom Nullpunkt durch ζ und durch $f(\zeta)$ liegen symmetrisch zur x -Achse, und die Entfernungen zum Nullpunkt sind Kehrwerte zu einander. Geometrisch handelt es sich um eine Inversion am Einheitskreis. Interessanterweise geht dabei eine Kreis, der nicht durch den Nullpunkt geht, auf einen Kreis.

7.9 Wirkung von f „im Kleinen“

Wir haben uns angeschaut, wie die Funktion f mit $f(\zeta) = \zeta^3 - 2\zeta^2 + 7\zeta - 5$ auf die komplexe Ebene wirkt. Natürlich kann man zu jedem $\zeta \in \mathbb{C}$ den Funktionswert $f(\zeta)$ ausrechnen, zum Beispiel ist $f(2i) = 3 + 6i$. Man kann auch eine Kurve Punkt für Punkt abbilden und sich anschauen, welches Bild die Gesamtheit der Bildpunkte bietet. Zwei Beispiele siehst du in Abbildung 9, links ist ein Bild des Einheitskreises und rechts ein Bild des Kreises um den Nullpunkt mit dem Radius 8.

Du siehst, die Funktion stellt mit den Objekten schon einiges an, aber immerhin gehen die beiden geschlossenen Kurven wieder auf geschlossene Kurven. Wie ich euch erzählt habe, verhält sich die Sache so: Denke dir die Originalebene mit einer dünnen elastischen Folie überzogen. Die Funktion zieht die Folie ab, dehnt sie mancherorts, staucht sie andernorts, und legt sie dann auf

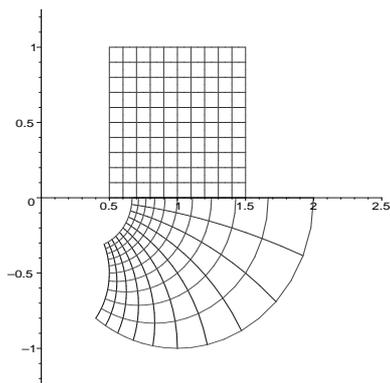


Abbildung 8: Wirkung von $f(\zeta) = \frac{1}{\zeta}$ auf ein Quadrat

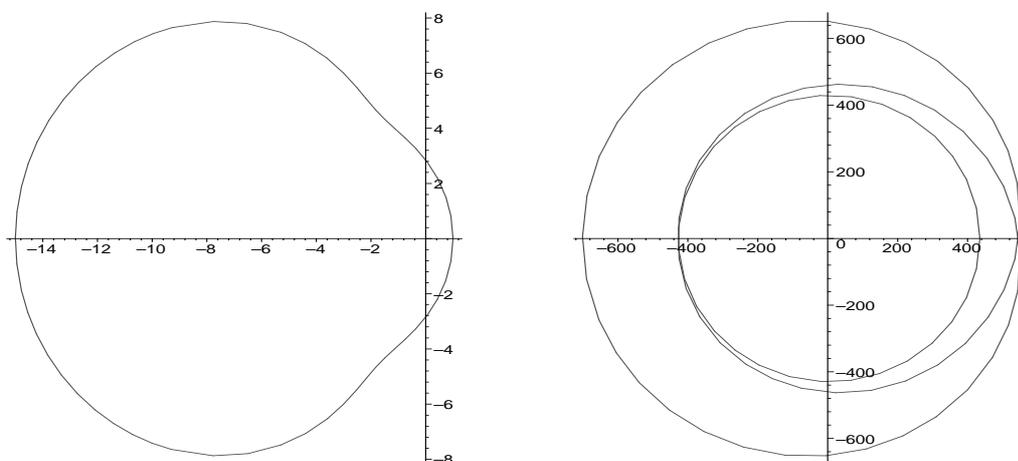


Abbildung 9: Bilder der Kreise um den Nullpunkt mit den Radien 1 und 8 unter f

die Bildebene, teilweise auch doppelt oder dreifach, und dabei kommt jeder Folienpunkt auf seinen Bildpunkt zu liegen. Die Funktion zerreit die Folie dabei nicht; Originalpunkte, die gengend nahe bei einander liegen, gehen auch auf Bildpunkte, die nahe bei einander liegen. Diese Eigenschaft der Funktion heit **Stetigkeit**. Die Tangensfunktion zum Beispiel ist nicht stetig, sie zerreit die reelle Achse in Stcke. Wenn du weit, dass $x \approx \frac{1}{2}\pi$ ist, weit du berhaupt nicht, wie gro $\tan(x)$ ist.

Unsere Funktion f ist nicht nur stetig, sondern auch noch (komplex) differenzierbar, und dadurch verhlt sie sich „im Kleinen“ noch erheblich gutartiger. Das will ich dir anhand einer Folge von Bildern zeigen. Wir betrachten die Funktion an der Stelle $\zeta = 2i$ und lassen uns vom Rechner zeichnen, was sie aus einem Quadrat der Kantenlnge d macht, dessen Seiten achsenparallel sind und dessen untere linke Ecke $2i$ ist. Dabei beginnen wir mit $d = 5$ und verkleinern d immer weiter. In Abbildung 10 auf Seite 34 siehst du die Bildfiguren fr die d -Werte 5, 3, 1, $\frac{1}{2}$, $\frac{1}{5}$, $\frac{1}{10}$ und $\frac{1}{20}$, und zwar sind immer die Figuren zu einem d -Wert und zum nchstkleineren d -Wert in einem Bild gezeigt.

Du siehst, dass die Bildfiguren fr kleine d praktisch Quadrate sind, die gegen das Originalquadrat gestreckt und gedreht sind. In der Tat wird diese Drehstreckung durch Multiplikation mit der komplexen Zahl $f'(2i) = -5 - 8i$ bewirkt: $\zeta \mapsto (-5 - 8i) \cdot \zeta$. Es ist hier genau so, wie im Reellen, da hatten wir

$$f(x+h) \approx f(x) + f'(x)h \quad ,$$

eine kleine Umgebung von x wurde durch f „im Kleinen“ einfach gestreckt und bei $f(x)$ festge-

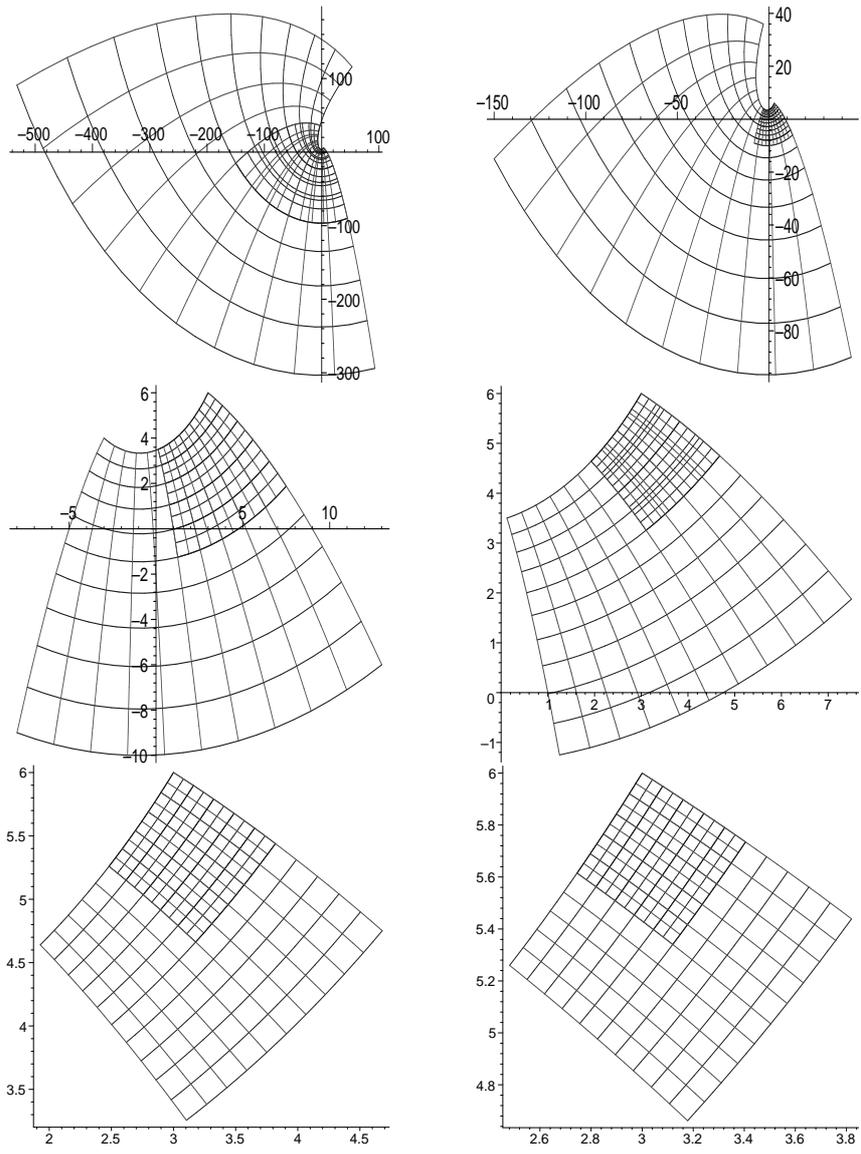


Abbildung 10: Bilder der Quadrate mit den Ecken $2i$, $d+2i$, $d+(2+d)i$ und $(2+d)i$ für verschiedene d -Werte

macht. Eine Umgebung von $\zeta \in \mathbb{C}$ kannst du dir als kleine Kreisscheibe um ζ vorstellen. Diese Kreisscheibe wird gestreckt und gedreht und bei $f(\zeta)$ festgemacht. Insbesondere ist die Abbildung f „im Kleinen“ winkeltreu, die Schnittwinkel von Kurven durch ζ bleiben bei der Abbildung erhalten.

7.10 Ausblick auf Anwendungen komplexer Funktionen

Hier zeige ich dir aber nur noch die Abbildung 11 auf Seite 35, die ich mit Hilfe der Funktion

$$f : \zeta \mapsto \zeta - \sqrt{\zeta^2 - 1}$$

erstellt habe. So etwas verwenden Leute, die sich mit strömenden Flüssigkeiten beschäftigen. Hier umströmt (trockenes!) Wasser ein kreisförmiges Hindernis. Das Wasser halte ich für trocken, weil sich hinter dem Hindernis keine Wirbel bilden; vermutlich wurden vereinfachende Annahmen gemacht. Ich bin kein Experte auf diesem Gebiet, deshalb sage ich das mit einiger Vorsicht. Immerhin hast du einen kleinen Beleg für die Behauptung, dass komplexe Zahlen vielerorts angewandt werden.

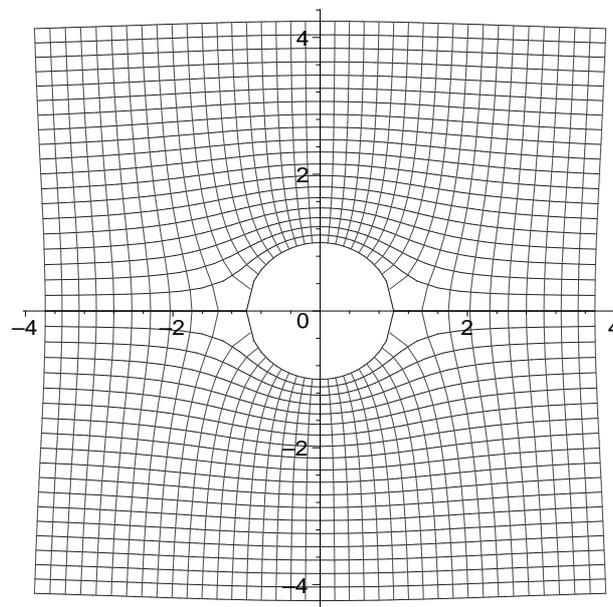


Abbildung 11: Anwendungsbeispiel aus der Strömungsdynamik