

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (<http://www.ub.tuwien.ac.at>).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (<http://www.ub.tuwien.ac.at/englweb/>).



TECHNISCHE
UNIVERSITÄT
WIEN

Vienna University of Technology

DIPLOMARBEIT

Stochastische Simulation im Mathematikunterricht

ausgeführt am

Institut für Statistik und Wahrscheinlichkeitstheorie

der Technischen Universität Wien

unter Anleitung von

Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Heinz Stadler

durch

Daniel Homann

1150 Wien

3. Mai 2013

Danksagung

An dieser Stelle möchte ich mich bei Dr. Heinz Stadler bedanken, mir das Thema der *Stochastischen Simulation* vorgeschlagen, mir Hinweise für deren Anwendung gegeben und mir ferner die Arbeit an diesem spannenden Thema ermöglicht zu haben.

Ein besonderer Dank gebührt meiner Familie für die Unterstützung während des Studiums, meinen Freunden und Freundinnen für die Gespräche, die mich auf neue Ideen für diese Arbeit brachten und meiner Freundin Franziska, die mir im Zuge des gesamten Studiums zur Seite stand.

Zudem möchte ich mich bei meinen Studienkollegen und Studienkolleginnen für die gute Zusammenarbeit im Laufe des Studiums bedanken.

Wien, 3. Mai 2013

Daniel Homann
dhomann@lavabit.com

Inhaltsverzeichnis

Einleitung	1
I. Theoretische Grundlagen	3
1. Wahrscheinlichkeitstheoretische Grundlagen und Zufallszahlen	5
1.1. Motivation	5
1.2. Wahrscheinlichkeitstheoretische Grundlagen	7
1.3. Zufallszahlgeneratoren für die diskrete Gleichverteilung	9
1.4. Graphisches Testen von Zufallszahlen: Erster Versuch	12
1.5. Zufallszahlgeneratoren für die stetige Gleichverteilung	13
1.6. Graphisches Testen von Zufallszahlen: Zweiter Versuch	15
1.7. Durchführung von statistischen Tests	17
1.8. Zufallsgeneratoren bei beliebiger Verteilung	20
1.9. Weitere Grundlagen der Stochastik	23
2. Didaktische Überlegungen	25
2.1. Didaktische Theorie	25
2.1.1. Die genetische Methode	26
2.1.2. Perspektiven der Mathematik	27
2.1.3. Heuristik im Mathematikunterricht	28
2.2. Didaktik der Stochastik	28
2.2.1. Genese der Stochastik	29
2.2.2. Vorverständnis und intellektuelle Entwicklung der Lernenden	30
2.2.3. Frequentistische Wahrscheinlichkeit	31
2.2.4. Stochastische Simulation und Ziele des Stochastikunterrichts	32
2.3. Stochastische Simulationen im Mathematikunterricht	33
2.3.1. Zehn Gründe für den Einsatz von stochastischen Simulationen im Mathematikunterricht	33
2.3.2. Phasen der stochastischen Simulation im Unterricht	35
II. Durchführung von Simulationen	37
3. Veranschaulichung des empirischen Gesetzes der Großen Zahlen	41
4. Beispiele zum Erwartungswert	45
4.1. Einführung des Begriffs <i>Erwartungswert</i> mithilfe von Simulation	45

4.2. Treibjagdproblem	46
4.3. Lotto-Kugel-Paradoxon	49
5. Auflösen von Paradoxien	53
5.1. Drei-Türen-Paradoxon	53
5.2. Geburtstagsparadoxon	55
6. Veranschaulichung von Grenzwertsätzen	59
6.1. Approximation der Binomialverteilung	59
6.2. Zentraler Grenzwertsatz	59
7. Subjektive Wahrscheinlichkeiten und Simulation	65
7.1. Fußball WM-Qualifikation	65
8. Wartezeiten und Warteschlangen	71
8.1. Wartezeiten	71
8.1.1. Warten auf Augenzahl 6	71
8.1.2. Sammlerproblem	72
8.2. Warteschlangen	74
8.2.1. Handy-Reperatur	74
8.2.2. Definition einer Warteschlange	75
8.2.3. Warteschlange an einer Kassa mittels Würfeln	76
8.2.4. Warteschlange an einer Kassa als $M M 1$ -System	80
9. Schließende Statistik	89
9.1. Konfidenzintervalle	89
9.2. Statistische Tests und Resampling	91
9.2.1. Test des Mittelwerts: t-Test vs. Bootstrapping	91
9.2.2. Vergleich von Mittelwerten: t-Test vs. Permutationstest	93
10. Flächenberechnungen mit Monte-Carlo-Methoden	97
10.1. Schätzen von π	97
10.2. Treffermethode	100
10.3. Intervallschätzer und Genauigkeit der Treffermethode	101
10.3.1. Konfidenzintervalle	101
10.3.2. Erreichen einer gewissen Genauigkeit	103
10.4. Güte der Treffermethode	104
10.5. Mittelwertmethode	104
10.6. Güte der Mittelwertmethode	106
10.7. Effizienz - Vergleich der beiden Verfahren	107
Literaturverzeichnis	109
Anhang	111

Einleitung

Sind wir mit Phänomenen, in denen der Zufall eine Rolle spielt, konfrontiert, widerspricht deren Verhalten oft unserer ersten Intuition. Unser Bauchgefühl, das sich im Zuge der von uns gemachten Erfahrungen herausgebildet hat, hilft uns daher oft nicht weiter.

Eine Möglichkeit wäre es sich von diesem ganz abzuwenden und die Probleme rein auf mathematischen Wege mithilfe der Wahrscheinlichkeitstheorie und Statistik zu lösen. In dieser Arbeit wird jedoch ein anderer Weg eingeschlagen: Das zufällige Phänomen wird mehrmals durchgespielt und mithilfe des beobachteten Verhaltens soll das Bauchgefühl geschult werden. Erst danach erfolgt eine Behandlung des Problems auf analytisch-mathematische Weise.

Im ersten Teil der Arbeit werden die theoretischen Grundlagen gelegt: Im ersten Kapitel werden die für die Simulationen benötigten Zufallszahlen auf Schulniveau mathematisch behandelt und wahrscheinlichkeitstheoretische Grundlagen erarbeitet. Anschließend wird im zweiten Kapitel der Frage nachgegangen, welche Vorteile die eben geschilderte experimentelle Vorgangsweise gegenüber einer rein mathematischen Behandlung des Problems haben kann.

Im zweiten Teil werden Simulationen durchgeführt. Nach einer kurzen Wiederholung des *Empirischen Gesetzes der Großen Zahlen* in Kapitel 3 beschäftigen sich die folgenden zwei Kapitel damit, bekannte Beispiele aus Stochastik-Kursen oder dem Mathematikunterricht in einem neuen Gewand, d.h. mithilfe stochastischer Simulation, zu behandeln. Anschließend widmet sich das Kapitel 6 der Veranschaulichung des *Zentralen Grenzwertsatzes*. Nach einem Ausflug in die Welt der Buchmacher und Buchmacherinnen (Kapitel 7) beschäftigen sich die letzten drei Kapitel mit realen Anwendungen von Simulationen, wie dem Verhalten von *Warteschlangen*, Simulationen in der *schließenden Statistik* und Flächenberechnungen mithilfe experimenteller Stochastik. Ferner weisen die Kapitel eine interne Differenzierung auf. Das Niveau der Beispiele der einzelnen Kapitel nimmt mit dem Verlauf zu: Über einführende einfache Beispiele wird schrittweise zu komplexeren Beispielen hingeführt.

Teil I.

Theoretische Grundlagen

Kapitel 1.

Wahrscheinlichkeitstheoretische Grundlagen und Zufallszahlen

1.1. Motivation

Bevor mit der Simulation von Zufallsphänomenen begonnen werden kann, müssen noch einige Vorarbeiten erledigt werden.

Der erste Schritt ist die Erzeugung von Zufallszahlen. Anhand des folgenden Beispiels soll kurz erläutert werden, wie wir Zufallszahlen erzeugen und wozu sie benötigt werden.

Beispiel: Stellen wir uns vor, wir veranstalten eine Tombola auf einem Ball - beispielsweise auf einem Maturaball. Für diesen Zweck haben wir 1000 Lose erstellt und bieten diese zum Verkauf an. Die Lose sind von 1 bis 1000 durchnummeriert. Sind alle Lose verkauft, wird anschließend ein zufälliges gezogen und die Person mit diesem Los gewinnt den Hauptpreis.

Doch wie wählen wir nun ein beliebiges Los aus der Menge aller Lose aus?

Eine Möglichkeit wäre es, die Lose in doppelter Ausführung vorzubereiten, die zweite Serie in eine Urne zu geben und aus dieser ein Los herauszuziehen. Wir wählen also aus den Losen mit den Nummern 1 bis 1000 ein zufälliges aus. Eine solche Ziehung bezeichnen wir als ein *Zufallsexperiment*.

Dabei handelt es sich um einen realen Vorgang, dessen mögliche Ausgänge bereits vor der Versuchsdurchführung feststehen. Welchen Ausgang der Vorgang nimmt, hängt allerdings vom Zufall ab. In unserem Fall wissen wir, dass ein Los mit einer Nummer von 1 bis 1000 gezogen wird. Welches Los und welche Nummer wir ziehen werden, wissen wir allerdings nicht. Wir könnten das Los mit der Nummer *34* oder jenes mit der Nummer *701* oder ein ganz anderes ziehen.

Betrachten wir das Beispiel nun aus einer anderen Perspektive. Statt aus allen von 1 bis 1000 durchnummerierten Losen eines zu ziehen, könnten wir auch sagen, dass wir eine zufällige Zahl von 1 bis 1000 auswählen. Wir erzeugen also mithilfe der Ziehung eines Loses eine zufällige Zahl: *34*. Die gezogene Zahl bezeichnen wir daher als eine *Zufallszahl* und das Zufallsexperiment des Ziehens dient als *Zufallszahlgenerator* - d.h. als Erzeuger von Zufallszahlen.

Oft werden wir allerdings mit dem Fall konfrontiert sein, dass ein solches reales Zufallsexperiment nicht bzw. nur mit großem Aufwand möglich ist.

Veranstaltet beispielsweise eine große Firma eine Lotterie mit 1.000.000 Losen und 10.000 Gewinnen, wird sie vermutlich keine 1.000.000 Lose drucken, jedes einzelne Los zusammenrollen, alle Lose in eine Urne geben und vor allem keine 10.000 händischen Ziehungen wie im obigen Fall durchführen.

Die Verantwortlichen müssen in diesem Fall eine schnellere und einfachere Möglichkeit finden, um ein Siegerlos zu ermitteln. Gäbe es schnelle Verfahren, um eine (scheinbar¹) zufällige Zahl von 1 bis 1.000.000 zu erzeugen, so könnten sie dieses verwenden, um ein bestimmtes Los auszuwählen. Wir würden die reale Ziehung von Losen in diesem Fall anhand von Zufallszahlen *simulieren*. Mithilfe eines Computers kann man sehr schnell (scheinbare) Zufallszahlen erzeugen. Beispielsweise erstellt ein herkömmlicher Computer ohne bemerkbare Rechenzeiten folgende 3 Zufallszahlen von 1 bis 1.000.000: 205779, 185107, 431732

Wollen wir unter allen 1.000.000 Losen 3 Gewinne vergeben, haben wir bereits die Nummern der Gewinnerlose bestimmt.

1. Preis: Los 205779, 2. Preis: Los 185107, 3. Preis: Los 431732

Obwohl dieses Beispiel nur eine Einführung darstellt, kann man die allgemeine Bedeutung von Zufallszahlen für Problemstellungen, in denen der Zufall eine Rolle spielt, vermuten.

Stellen wir uns beispielsweise die Kassen eines Geschäfts vor. Je nachdem wie viele Leute einkaufen, werden eine oder sogar zwei bzw. drei Kassen geöffnet. Unter der Annahme, dass die Kunden zufällig alle paar Minuten eintreffen und eine bestimmte zufällige Zeit im Geschäft verbringen, wird man das Eintreffen mit Zufallszahlen simulieren und davon ableiten, wie viele Kassen man geöffnet haben sollte.

Verfahren, die Zufallszahlen verwenden, werden allgemein als *Monte-Carlo-Algorithmen*² bezeichnet und ein zentrales Thema dieser Diplomarbeit sein. Wir werden sehen, dass diese Verfahren einerseits für Aufgabenstellungen, die etwas Zufälliges in sich tragen, verwendet werden können - z.B. für die Abschätzung der Auslastung der Kassen eines Geschäfts - andererseits gibt es auch Monte-Carlo-Algorithmen für Problemstellungen, die nichts mit dem Zufall zu tun haben. Ein Beispiel hierfür sind Flächenberechnungen mithilfe von Monte-Carlo-Verfahren.

Das Gebiet, welches sich mit diesen Algorithmen beschäftigt, wird *stochastische Simulation*³ genannt. Doch bevor wir uns mit diesem beschäftigen, werden wir im Folgenden,

¹Das Wort Verfahren lässt bereits vermuten, dass keine echt zufällige Zahl erzeugt, sondern eine scheinbar zufällige Zahl berechnet wird. Eine genaue Unterscheidung zwischen echten und scheinbar echten Zufallszahlen erfolgt in nächsten Abschnitt.

²auch: *stochastische Algorithmen* oder *randomisierte Algorithmen*

³auch: *experimentelle Stochastik*

als notwendige Vorbereitung, Zufallszahlen etwas genauer behandeln.

1.2. Wahrscheinlichkeitstheoretische Grundlagen

Im vorherigen Abschnitt haben wir uns mit zufälligen Zahlen z aus der Menge aller natürlichen Zahlen von 1 bis zu einem bestimmten Wert n aus den natürlichen Zahlen, z.B. $n = 1.000$, beschäftigt. Nun wollen wir das Zufallsexperiment, bei dem ein Los aus einer Urne gezogen wird, mathematisch fassen.

Die möglichen Ausgänge des Experiments bezeichnen wir als **Ergebnismenge** Ω . Da in unserem Fall ein beliebiges aus allen n Losen gezogen werden kann, gilt für unsere Ergebnismenge Ω aller möglichen Ziehungen

$$\Omega = \{\text{Los 1 gezogen, Los 2 gezogen, Los 3 gezogen, ..., Los n gezogen}\}$$

Die Elemente ω aus Ω sind nicht die einzelnen Lose, sondern die möglichen Ziehungen einzelner Lose.

Jede Teilmenge A der Ergebnismenge Ω bildet ein mögliches **Ereignis**. Habe ich ein Los gekauft und trägt dieses die Nummer 245, so wird für mich das Ereignis $A = \{\text{Los 245 gezogen}\}$ von Interesse sein. Alle einelementigen Ereignisse der Form

$$A = \{\omega\} = \{\text{Los } i \text{ gezogen}\}$$

nennt man *Elementarereignisse*. Habe ich aber beispielsweise mehrere Lose, z.B. die Lose 134, 875 und 901 gekauft, so ist für mich das Ereignis

$$A = \{\text{Los 134 gezogen, Los 875 gezogen, Los 901 gezogen}\}$$

von Interesse, d.h. dass eines der Lose gezogen wird. Die Menge aller möglichen Ereignisse, also die Menge aller Teilmengen T von Ω , ist daher die Potenzmenge \mathcal{P} von Ω

$$\mathcal{P}(\Omega) = \{T | T \subseteq \Omega\}$$

die wir als **Ereignismenge** bezeichnen werden.

Jetzt wollen wir uns dem Begriff Wahrscheinlichkeit widmen. Betrachten wir wiederum den Fall mit 1000 Losen. Wenn man ein beliebiges Los aus der Menge aller Lose herausnimmt, dann ist die Wahrscheinlichkeit P der Ziehung ω eines bestimmten Loses für jedes Los gleich groß und nach der *Klassischen Wahrscheinlichkeitsdefinition* gleich

$$P(\omega) = \frac{1}{n} = \frac{1}{|\Omega|} = \frac{1}{1000}, \forall \omega \in \Omega$$

Ein beliebiges Los wird mit Wahrscheinlichkeit $\frac{1}{1000} = 0.001 = 0.1\%$ gezogen.

Definieren wir nun die Wahrscheinlichkeit P von einem Ereignis A allgemein als Abbildung von der Ereignismenge $\mathcal{P}(\Omega)$ auf die reellen Zahlen \mathbb{R} :

Definition 1.2.1 (Wahrscheinlichkeitsmaß auf endlicher Ergebnismenge) Sei Ω nicht-leer und endlich. Eine Abbildung

$$P : \mathcal{P}(\Omega) \rightarrow \mathbb{R}$$

heißt Wahrscheinlichkeitsmaß genau dann, wenn gilt:

1. $P(A) \geq 0$ für alle $A \in \mathcal{P}(\Omega)$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B)$ für alle $A, B \in \mathcal{P}(\Omega)$ mit $A \cap B = \emptyset$

Der Funktionswert $P(A)$ heißt die Wahrscheinlichkeit des Ereignisses A .

Das Tripel $(\Omega, \mathcal{P}(\Omega), P)$ - oder auch nur das Paar (Ω, P) - heißt endlicher Wahrscheinlichkeitsraum.

Bemerkung 1.2.2 Das erste Axiom nennt man auch die Nichtnegativität, das zweite die Normierung und das dritte die Additivität des Wahrscheinlichkeitsmaßes. Wenn wir im Folgenden von Wahrscheinlichkeiten sprechen, sind für das Wahrscheinlichkeitsmaß P immer diese drei Axiome gültig.

Die von uns verwendete klassische Wahrscheinlichkeit ist ein solches Wahrscheinlichkeitsmaß:

Satz 1.2.3 (Klassische Wahrscheinlichkeitsdefinition) Sei Ω nichtleer und endlich, $\mathcal{P}(\Omega)$ die Potenzmenge von Ω , $A \in \mathcal{P}(\Omega)$ ein beliebiges Ereignis und alle Elementarereignisse (Fälle) gleich wahrscheinlich, dann ist

$$P : \mathcal{P}(\Omega) \rightarrow \mathbb{R} \text{ mit } P(A) = \frac{\text{Anzahl der für das Ereignis } A \text{ günstigen Fälle}}{\text{Anzahl möglicher Fälle}} = \frac{|A|}{|\Omega|}$$

ein Wahrscheinlichkeitsmaß und $P(A)$ wird mit Laplace-Wahrscheinlichkeit bezeichnet.

Beweis. Durch Überprüfen der Axiome 1 - 3, vgl. Kütting/Sauer [14], S. 98, Anmerkung 2. □

Daher können wir den Wahrscheinlichkeitsraum $(\Omega, \mathcal{P}(\Omega), P)$, mit Ω als Menge aller möglichen Ziehungen, $\mathcal{P}(\Omega)$ als Menge aller Teilmengen von Ω und P als Wahrscheinlichkeitsmaß nach Satz 1.2.3, als Modell für unser Zufallsexperiment verwenden und können Fragestellungen, die vom Zufall abhängen, mathematisch beschreiben und lösen.

Nun haben wir fast alle Grundlagen für die Definition einer Zufallszahl erarbeitet. Das einzige Problem ist, dass wir immer die Ziehungen von Losen als Ereignisse betrachtet haben und nicht die Nummern der gezogenen Lose, die uns eigentlich interessieren.

Wir bilden daher die Ziehungen ω der einzelnen Lose mithilfe der Funktion X auf die Nummer des gezogenen Loses $X(\omega) \in \mathbb{N}$ ab:

ω	$X(\omega)$
Los 1 gezogen	1
Los 2 gezogen	2
Los 3 gezogen	3
...	...
Los 1000 gezogen	1000

Definieren wir diese Abbildung X nun allgemein:

Definition 1.2.4 (Diskrete Zufallsvariable) Sei $(\Omega, \mathcal{P}(\Omega), P)$ ein Wahrscheinlichkeitsraum mit endlicher Ergebnismenge Ω . Dann heißt jede Abbildung

$$X : \Omega \rightarrow \mathbb{R} \text{ mit } \omega \mapsto X(\omega)$$

eine diskrete Zufallsvariable oder diskrete Zufallsgröße auf Ω .

In Zukunft werden wir nur mehr diese Zufallsvariable X betrachten. Anstatt von der Ziehung des i -ten Loses zu sprechen, sagen wir, dass die Zufallsvariable X den Wert i annimmt. Wollen wir nun beispielsweise in unserem Modell der *Klassischen Wahrscheinlichkeitsdefinition* ermitteln, wie groß die Wahrscheinlichkeit ist, dass das Los 345 gezogen wird, d.h. dass X den Wert 345 annimmt, so schreiben wir

$$P(X = 345) = P(\{\text{Los 345 gezogen}\}) = \frac{1}{|\Omega|} = \frac{1}{1000}$$

Allgemein definieren wir die Wahrscheinlichkeiten von Zufallsvariablen folgendermaßen:

Definition 1.2.5 (Wahrscheinlichkeitsverteilung) Sei $(\Omega, \mathcal{P}(\Omega), P)$ ein Wahrscheinlichkeitsraum mit endlicher Ergebnismenge, X eine Zufallsvariable auf Ω . Dann bezeichnen wir mit

$$p(x) = P(X = x) = P(\{\omega \in \Omega | X(\omega) = x\})$$

die Wahrscheinlichkeit, dass X den Wert x annimmt und die Gesamtheit aller $p(x)$ aus der Bildmenge $X(\Omega)$ von X als Wahrscheinlichkeitsverteilung der Zufallsgröße X .

Mithilfe der theoretischen Vorarbeit in diesem Abschnitt können wir im Folgenden Zufallszahlen definieren.

1.3. Zufallszahlgeneratoren für die diskrete Gleichverteilung

Bei unserem Zufallsexperiment ziehen wir eines von 1000 durchnummerierten Loses. Am Ende des letzten Abschnitts haben wir den Begriff der *diskreten Zufallsvariable* eingeführt: Mithilfe einer Funktion X können wir die Ziehung ω eines Loses auf die Nummer des Loses $X(\omega) = i$ abbilden und in Zukunft nur mehr diese Zufallsvariable X betrachten. In unserem Fall nimmt X Werte im Bereich von 1 bis 1000 an. Bei der ersten Versuchsdurchführung könnte es die Zahl 65 bei der zweiten die Zahl 873 und bei der dritten wiederum die Zahl 65⁴ sein. Doch welche Eigenschaften hat unsere Zufallsgröße X , dass wir sie als Zufallszahl bezeichnen können?

⁴Unter der Annahme, dass man die gezogenen Lose wieder in die Urne zurücklegt

Definition 1.3.1 (Diskrete Gleichverteilung) Eine diskrete Zufallsgröße X heißt diskret gleichverteilt auf M , ($X \sim D_n$), wenn

$$M = \{1, \dots, n\} \text{ und } p(x) = \frac{1}{n} \forall x \in M$$

Im vorherigen Abschnitt haben wir gesehen, dass jedes Los mit gleicher Wahrscheinlichkeit $P(\{\text{Los } i \text{ gezogen}\}) = \frac{1}{1000}$ gezogen wird. Daher gilt das auch für die einzelnen Ausprägungen unserer Zufallsvariable X , denn $p(i) = P(X = i) = P(\{\text{Los } i \text{ gezogen}\})$, und X ist diskret gleichverteilt.

Bildlich kann man sich diese Eigenschaft so vorstellen, dass bei der Ziehung der Lose kein Los bevorzugt wird. Es werden die Lose in der Urne vermischt und anschließend, ohne eine Auswahl zu treffen, ein beliebiges Los herausgezogen. Wir wollen die Eigenschaft der diskreten Gleichverteilung als Definition einer Zufallszahl aus der Menge $\{1, \dots, n\}$ heranziehen:

Definition 1.3.2 (Echte Zufallszahl aus $M = \{1, \dots, n\}$) Eine Zufallsvariable X heißt echte Zufallszahl aus $M = \{1, \dots, n\}$, wenn sie diskret gleichverteilt auf M ist.

Echte Zufallszahlen erzeugt man mithilfe von Zufallsexperimenten. Beispiele hierfür sind Glücksspiele wie Würfeln oder Roulette bzw. physikalische Experimente. Da in der Praxis oft eine große Anzahl an Zufallszahlen benötigt wird und die Durchführung von Zufallsexperimenten in diesem Fall zu aufwendig oder teuer wäre, erzeugt man sogenannte **Pseudozufallszahlen**.

Diese werden mithilfe von Algorithmen erzeugt. Das bringt den Vorteil mit sich, dass sehr viele Pseudozufallszahlen in sehr kurzer Zeit erzeugbar sind, da man die Generierung durch den Computer automatisieren kann. Andererseits sind sie in der Regel nicht diskret gleichverteilt und streng genommen nicht zufällig.

In dieser Arbeit wird nur kurz auf die einzelnen Generatoren eingegangen und nur einige exemplarisch erwähnt. Der interessierte Lesende sei auf Hagmann [9], Kapitel 3, und Kolonko [11], Kapitel 2 bis 5, verwiesen.

Zur Erzeugung einer Folge von Pseudozufallszahlen verwendet man für gewöhnlich rekursive Algorithmen. Ausgehend von einem Startwert erzeugt man Schritt für Schritt mithilfe der aktuellen Zahl eine neue.

Beispielgeneratoren

$$\begin{array}{ll} \text{Gemischter Kongruenzgenerator} & x_i \equiv a \cdot x_{i-1} + b \pmod{m} \quad 0 \leq x_i \leq m \quad m \in \mathbb{N} \\ \text{Multiplikativer Kongruenzgenerator} & x_i \equiv a \cdot x_{i-1} \pmod{m} \quad 0 \leq x_i \leq m \quad m \in \mathbb{N} \\ \text{Fibonacci Generator} & x_i = x_{i-1} + x_{i-2} \pmod{m} \quad 0 \leq x_i \leq m \quad m \in \mathbb{N} \end{array}$$

Da in jedem Fall $\text{mod } m$ gerechnet wird und m endlich ist, sind mit diesen Algorithmen nur endlich viele Zufallszahlen erzeugbar. In anderen Worten bedeutet das, dass in der

Folge von Zufallszahlen irgendwann Wiederholungen auftreten. Wählen wir beispielsweise den gemischten Kongruenzgenerator $x_i \equiv 2 \cdot x_{i-1} + 5 \pmod{10}$ mit dem Startwert $x_0 := 3$ und sieht man sich die ersten 10 Werte an, kann man sofort das periodische Verhalten erkennen.

i	0	1	2	3	4	5	6	7	8	9	10
x_i	3	1	7	9	3	1	7	9	3	1	7

Erhalte ich somit für irgendeine Stelle x_i , hier x_4 , der Folge einen bereits in der Folge enthaltenen Wert a , hier $a = 3$, dann werden im Folgenden keine neuen Zufallszahlen erzeugt.

Mithilfe von zahlentheoretischen Überlegungen wird man daher versuchen Zufallsgeneratoren zu konstruieren, die folgende Bedingungen erfüllen.

Forderungen an Pseudozufallszahlen (sinngemäß nach Hagmann [9], S. 11)

1. Die Periode des Generators sollte möglichst lange sein.
2. Die relativen Häufigkeiten der möglichen Werte der Zufallszahlen sollten annähernd diskret gleichverteilt sein.⁵
3. Zwei Zufallszahlen sollten voneinander unabhängig sein.

Für den dritten Punkt müssen wir noch definieren, was Unabhängigkeit bedeutet. Diese Eigenschaft bedeutet, dass, wenn wir uns eine Folge von erzeugten Zufallszahlen $(X_n)_{n \geq 0}$ vorstellen, die Werte von zwei Zufallszahlen X_i und X_j mit $i \neq j$ nicht voneinander abhängen. Hat X_{25} beispielsweise den Wert 3, so hat dieser Wert keinen Einfluss auf den Wert von X_{12} . Da wir Zufallszahlen als Zufallsvariablen aufgefasst haben, können wir diese Eigenschaft auch mathematisch definieren:

Definition 1.3.3 (Unabhängigkeit von Zufallsvariablen) Seien X und Y zwei Zufallsvariablen über demselben endlichen Wahrscheinlichkeitsraum (Ω, P) . X und Y heißen unabhängig, wenn für alle $a \in X$ und $b \in Y(\Omega)$ gilt

$$P(X = a | Y = b) = P(X = a)$$

Obwohl wir nun alle Forderungen verstehen, stellen wir uns die Frage, woher wir wissen wollen, dass der verwendete Zufallsgenerator diese Forderungen erfüllt? Man wird eine (ausreichend lange) Folge von Pseudozufallszahlen erzeugen und diese mithilfe von statistischen Tests auf die gewünschten Eigenschaften überprüfen. Bestehen diese Zahlen eine festgelegte Reihe von Tests, wird man sie als ausreichend zufällig erachten und den Zufallsgenerator akzeptieren.

Zu beachten ist, dass diese Tests keine Bestätigung der Qualität von Zufallsgeneratoren liefern - sie sortieren nur schlechte Generatoren aus.

⁵Will man z.B. Zufallszahlen aus $\{1, \dots, 10\}$ erzeugen, dann sollten ungefähr $\frac{1}{10}$ aller Zahlen den Wert 1, ungefähr $\frac{1}{10}$ aller Zahlen den Wert 2, usw. besitzen.

1.4. Graphisches Testen von Zufallszahlen: Erster Versuch

Wir wollen uns nun mit der Qualität von Zufallsgeneratoren beschäftigen und nähern uns diesem Problem zuerst mit graphischen Mitteln.

Visualisierung der Häufigkeiten

Betrachten wir nochmals den Kongruenzgenerator $x_i \equiv 2 \cdot x_{i-1} + 5 \pmod{10}$ mit dem Startwert $x_0 := 3$ aus dem vorigen Abschnitt. Erzeugt man eine Folge von 10 Zufallszahlen erhält man folgendes Ergebnis:

i	0	1	2	3	4	5	6	7	8	9	10
x_i	3	1	7	9	3	1	7	9	3	1	7

Bereits anhand dieser ersten zehn Werte erkennt man, dass der Zufallsgenerator sicher keine gleichverteilten Zufallszahlen erzeugt. Das Säulendiagramm des Zufallsgenerators für die ersten 100 Werte in Abbildung 1.1 führt zum selben Schluss.

Wie sieht es allerdings für den Kongruenzgenerator $x_i \equiv x_{i-1} + 1 \pmod{10}$ mit dem Startwert $x_0 := 3$ aus? Die ersten zehn Werte sehen sehr vielversprechend aus und da die Werte eine Periode von der Länge 10 aufweisen, erkennt man in Abbildung 1.1, dass die Werte gleichmäßig verteilt sind. Ist dieser Kongruenzgenerator somit gut?

Diese Frage müssen wir mit Nein beantworten - denn die einzelnen Werte sind, wie auch beim ersten Kongruenzgenerator, voneinander abhängig. Habe ich die Zufallszahl 1 erhalten, weiß ich, dass die nächste Zufallszahl den Wert 2 haben wird. Auf 2 folgt 3, auf 3 der Wert 4, usw. Die alleinige Überprüfung der (relativen) Häufigkeiten wird daher für

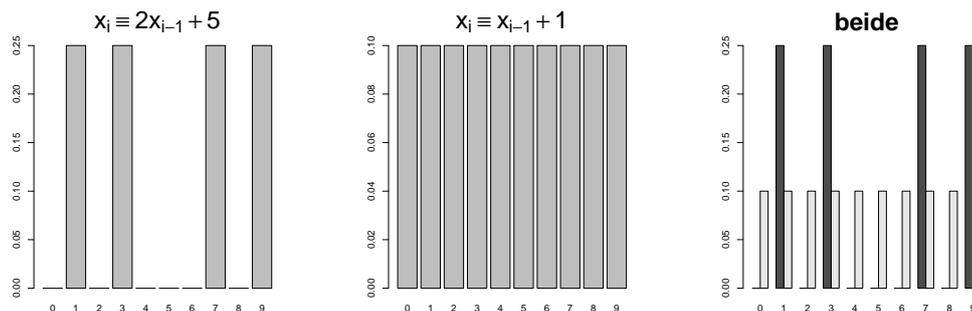


Abbildung 1.1.: Säulendiagramme der Kongruenzgeneratoren

unseren Zweck nicht ausreichend sein, weil die Unabhängigkeit außer Acht gelassen wird. Es gibt eine bessere Variante - doch dazu müssen wir Zufallszahlen auf dem Intervall $[0,1]$ betrachten.

1.5. Zufallszahlgeneratoren für die stetige Gleichverteilung

Wir wollen eine Zahl aus dem Intervall $[0,1]$ als zufällig betrachten, wenn sie wiederum gleichverteilt auf dem Intervall $[0,1]$ ist. Der Unterschied zur Menge $\{1, \dots, n\}$ ist, dass sich nicht endlich viele, sondern sogar überabzählbar unendlich viele Zahlen im Intervall $[0,1]$ befinden. Wir haben also keine endliche Ereignismenge Ω mehr und kommen daher mit unseren Definitionen aus Abschnitt 1.2 nicht weiter.

Dichte- und Verteilungsfunktion

Da man bei überabzählbar-unendlichen Mengen nicht jeder Teilmenge eine Wahrscheinlichkeit zuordnen kann, bedürfen die Definitionen ausführlicher Überlegungen. Eine ausführliche Behandlung, wie in Abschnitt 1.2, ist für das Verständnis dieser Arbeit nicht zwingend erforderlich und würde den Rahmen dieser Arbeit sprengen. Deshalb wird an dieser Stelle mit dem Verweis auf die Literatur der elementaren Wahrscheinlichkeitsrechnung, wie Kütting/Sauer [14], S. 286ff, auf eine exakte Definition verzichtet.

Betrachten wir eine Zufallsvariable X , die alle Werte aus einem Bereich M der reellen Zahlen \mathbb{R} , $(-\infty, c]$, $[c, \infty)$ oder $[a, b]$ mit $a, b, c \in \mathbb{R}$, annehmen kann. Nun können wir nicht jedem Wert der Zufallsvariable eine Wahrscheinlichkeit zuordnen, weil dann die Summe aller Wahrscheinlichkeiten niemals 1 wäre. Deswegen definieren wir ein Wahrscheinlichkeitsmaß, das Intervallen aus M Wahrscheinlichkeiten zuordnet:

Definition 1.5.1 (Dichtefunktion) Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}_0^+$ heißt Dichtefunktion, wenn gilt

1. f ist integrierbar
2. $\int_M f(x) dx = 1$

Mithilfe der Dichtefunktion können wir die Wahrscheinlichkeit berechnen, dass die Zufallsvariable X in irgendeinem Intervall $[a, b] \subseteq M$ liegt.

$$P(a \leq X \leq b) = \int_a^b f(x) dx \text{ für } [a, b] \subseteq M$$

Bildlich kann man sich die Wahrscheinlichkeit, dass eine Zufallsvariable X im Bereich zwischen a und b liegt, als die Fläche zwischen a und b unter der Dichtefunktion $f(x)$ vorstellen. Um diese Wahrscheinlichkeit berechnen zu können, muss $f(x)$ integrierbar sein (Punkt 1). Da X irgendeinen Wert aus M annehmen muss, ist die Wahrscheinlichkeit von $X \in M$ gleich 1 und daher muss auch die Fläche unter $f(x)$ über dem gesamten Bereich M gleich 1 sein (Punkt 2). In Abbildung 1.2 ist dieser Sachverhalt graphisch dargestellt.

Da man zur Ermittlung einzelner Wahrscheinlichkeiten immer ein bestimmtes Integral berechnen muss, tut man sich leichter, wenn man das unbestimmte Integral, die Stammfunktion, bereits kennt. Wir definieren daher das Integral der Dichtefunktion, die sogenannte Verteilungsfunktion, wie folgt:

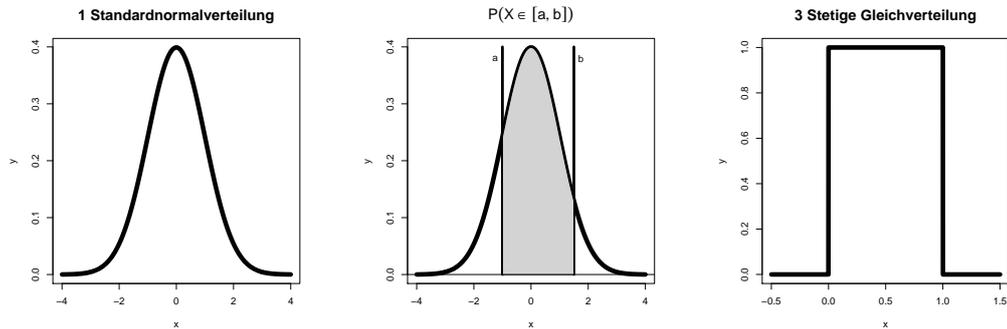


Abbildung 1.2.: Graphische Darstellung von Dichtefunktionen

Definition 1.5.2 (Verteilungsfunktion) Das Integral der Dichtefunktion f heißt Verteilungsfunktion F

$$F(x) := \int_{-\infty}^x f(t) dt$$

Kennt man die Verteilungsfunktion, kann man Wahrscheinlichkeiten ohne Integralbildung berechnen:

$$P(a \leq X \leq b) = \int_a^b f(x) dx = F(b) - F(a)$$

oder

$$P(X \leq x) = F(x)$$

In Abbildung 1.3 sind die Verteilungsfunktionen zu den Dichtefunktionen aus Abbildung 1.2 zu sehen. Wie die Dichtefunktionen (siehe Definition 1.5.1) besitzen auch Verteilungs-

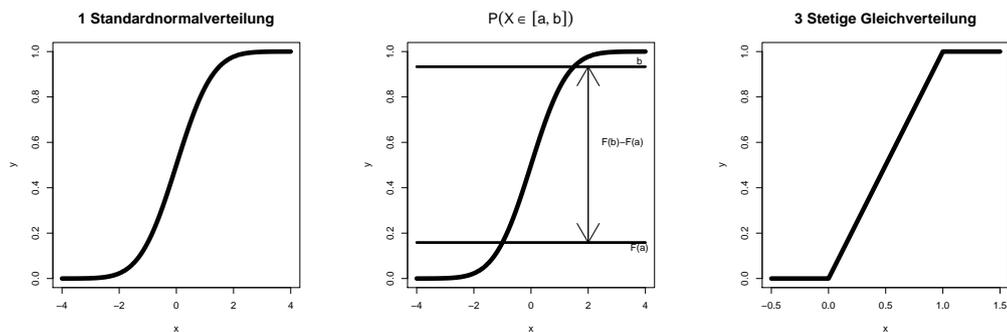


Abbildung 1.3.: Graphische Darstellung von Verteilungsfunktionen

funktionen bestimmte Eigenschaften, die man in Abbildung 1.3 erkennen kann und die in folgendem Satz zusammengefasst sind:

Satz 1.5.3 (Eigenschaften von Verteilungsfunktionen) Eine Verteilungsfunktion $F : \mathbb{R} \rightarrow \mathbb{R}_0^+$ besitzt folgende drei Eigenschaften:

1. F ist monoton wachsend
2. F ist rechtsseitig stetig
3. $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$

Beweis. siehe Kütting/Sauer [14], S. 300 □

Zufallszahlen aus [0,1]

In Definition 1.3.1 haben wir die diskrete Gleichverteilung so definiert, dass jeder Wert der Zufallsvariable X dieselbe Wahrscheinlichkeit besitzt. Das legen wir jetzt auf unsere stetige Zufallsvariable um, indem wir sagen, dass jedes gleich lange Intervall, in dem sich X befinden kann, dieselbe Wahrscheinlichkeit besitzt:

Definition 1.5.4 (Stetige Gleichverteilung) Eine Zufallsvariable X heißt stetig gleichverteilt auf $[a, b]$, $X \sim S_{a,b}$, wenn

$$M = [a, b] \text{ und } f(x) = \frac{1}{b-a} \text{ für } x \in [a, b]$$

Und daher definieren wir analog zu Definition 1.3.2 eine echte Zufallszahl aus $[0,1]$:

Definition 1.5.5 (Echte Zufallszahl aus $[0,1]$) Eine Zufallsvariable X heißt echte Zufallszahl aus $[0,1]$, wenn sie stetig gleichverteilt auf $[0,1]$ ist.

Abgesehen davon, dass wir im nächsten Kapitel einen weiteren graphischen Test mithilfe der Zufallszahlen aus $[0,1]$ durchführen werden, können diese dazu verwendet werden, um jede beliebige Verteilung auf \mathbb{R} zu simulieren. Waren aufgrund dieser Vorteile unsere Überlegungen zu Zufallszahlen aus $\{1, \dots, n\}$ überflüssig?

Überlegen wir uns zuerst, wie wir unendlich viele Pseudofallszahlen aus dem Bereich $[0,1]$ generieren. Da der Computer nur endlich viele Zahlen kennt bzw. wir uns jede erzeugte Zahl aufschreiben müssen, können wir nur eine Diskretisierung des Intervalls $[0,1]$ erzeugen. Haben wir Zufallszahlen $z_i \in \{1, \dots, n\}$ erzeugt, können wir diese auf das Intervall $[0,1]$ transformieren, indem wir sie durch $n+1$ dividieren: $x_i := \frac{z_i}{n+1} \in [0,1]$. Die x_i liegen somit in einer Diskretisierung von $[0,1]$ und für große n wird man auch gute Ergebnisse erhalten. Diese Vorgehensweise kann man damit rechtfertigen, dass die x_i , wenn n gegen unendlich läuft, gegen eine stetige Gleichverteilung auf $[0,1]$ streben. (Siehe Kolonko [11], Satz 2.1, S. 11)

1.6. Graphisches Testen von Zufallszahlen: Zweiter Versuch

Visualisierung als Punktwolke im Einheitsquadrat

Erzeugt man nun eine Folge von Zufallszahlen aus $[0,1]$ und betrachtet zwei gleich lange Teilintervalle, z.B. $[0.2, 0.4]$ und $[0.7, 0.9]$, so sollten ungefähr gleich viele Zufallszahlen in den beiden Intervallen liegen - siehe Abbildung 1.4.

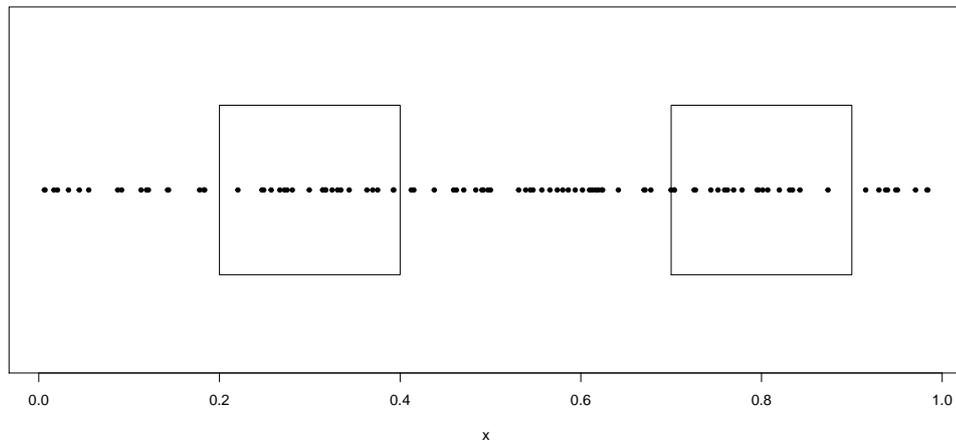


Abbildung 1.4.: Intervalle $[0.2, 0.4]$ und $[0.7, 0.9]$

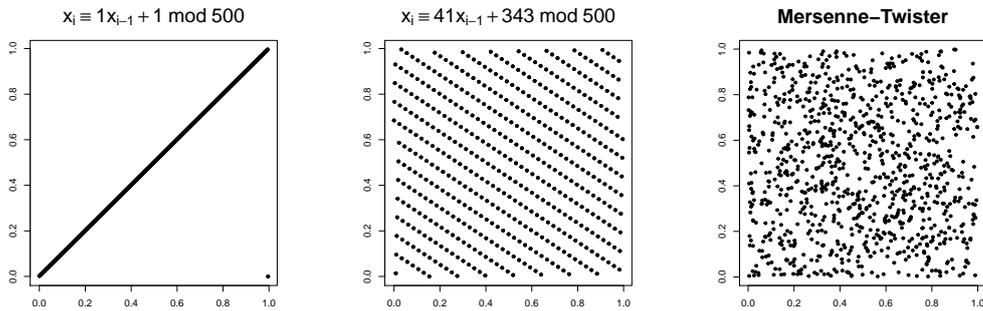
Diese Überlegung liefert uns ein zweites graphisches Testverfahren, mit dem wir zusätzlich noch die Abhängigkeit von zwei Zufallszahlen überprüfen können. Wir stellen uns das Einheitsquadrat $[0, 1] \times [0, 1]$ vor. Dann kann man zwei aufeinanderfolgende Zufallszahlen x_i und x_{i+1} aus einer Folge von Zufallszahlen $(x_n)_{n \geq 0}$ als Punkte $(x_i, x_{i+1}) = (x, y)$ in diesem Einheitsquadrat auffassen. Wie müssen diese Punkte nun im Quadrat verteilt sein, damit sie unabhängigen, stetig gleichverteilten Zufallsvariablen möglichst nahe kommen?

Erstens müssen die x - und y -Werte gleichmäßig auf dem $[0,1]$ -Intervall der jeweiligen Achse verteilt sein und zweitens müssen die Punkte, sofern die Werte x_i und x_{i+1} unabhängig sind, gleichmäßig im Quadrat verteilt sein. In Abbildung 1.5, in welcher Folgen von 1000 Zufallszahlen unterschiedlicher Zufallsgeneratoren⁶ im Einheitsquadrat dargestellt sind, kann man bereits schlechte Generatoren erkennen. Bei den beiden Kongruenzgeneratoren⁷ sind die x - und y -Werte zwar gleichmäßig auf den Intervallen $[0,1]$ der jeweiligen Koordinatenachsen verteilt, doch aufeinanderfolgende Werte voneinander abhängig. Das dritte Bild des Zufallsgenerators *Mersenne-Twister*⁸ lässt bereits ein besseres Verhalten bezüglich Unabhängigkeit vermuten.

⁶Die von den beiden Kongruenzgeneratoren erzeugten Folgen von Zufallszahlen wurden durch Division von $m + 1$, wobei m dem Modul entspricht, auf das Intervall $[0,1]$ transformiert.

⁷Es gibt auch gute Kongruenzgeneratoren.

⁸Dieser Generator wird standardmäßig von der Programmiersprache R und daher auch in dieser Arbeit verwendet, um Pseudo-Zufallszahlen aus $[0,1]$ zu erzeugen.

Abbildung 1.5.: Punktwolken verschiedener Zufallsgeneratoren auf $[0,1]$

1.7. Durchführung von statistischen Tests

Um feststellen zu können, ob Pseudozufallsgeneratoren Zahlen erzeugen, die sich annähernd wie echte Zufallszahlen verhalten, werden Tests durchgeführt. Dabei wird eine möglichst lange vom Generator erzeugte Zahlenfolge untersucht. Besteht die betrachtete Zahlenfolge eine Reihe von Tests, wird der Generator akzeptiert. Wir müssen also eine Entscheidung treffen: Entweder wir akzeptieren den Zufallszahlgenerator oder wir verwerfen ihn. Diese Entscheidung kann mithilfe eines statistischen Tests getroffen werden. Weil statistische Tests auch in Kapitel 9 thematisiert werden, machen wir noch einen kurzen Exkurs zu statistischer Testtheorie.

Testtheorie

Ausgangslage

Ausgangspunkt eines statistischen Tests ist eine zu untersuchende *Grundgesamtheit*. In unserem Fall sind das alle Zahlen, die von einem bestimmten Zufallsgenerator GEN erzeugt werden. Da es unmöglich ist, alle Zahlen zu ermitteln, wird man sich mit einer *Stichprobe*, d.h. einer endlichen Folge von erzeugten Zahlen $(x_n)_{N \geq n \geq 0}$, zufrieden geben und anhand dieser versuchen, eine Entscheidung zu treffen: Entweder wir akzeptieren den Zufallsgenerator oder wir verwerfen ihn.

Wir erzeugen für diesen Zweck 100 Zahlen mit dem Kongruenzgenerator $x_i \equiv 41 \cdot x_{i-1} + 343 \pmod{500}$ mit dem Startwert 251 - siehe Abbildung 1.6.

Hypothesen

Wir stellen nun eine Behauptung, die *Nullhypothese* H_0 , auf, die es gegen die *Alternative* H_1 zu überprüfen gilt. X sei eine Zufallsvariable, die alle vom Zufallsgenerator erzeugten Werte annimmt:

$$H_0 = X \text{ ist stetig gleichverteilt auf } [0,1]$$

$$H_1 = X \text{ folgt einer anderen Verteilung}$$

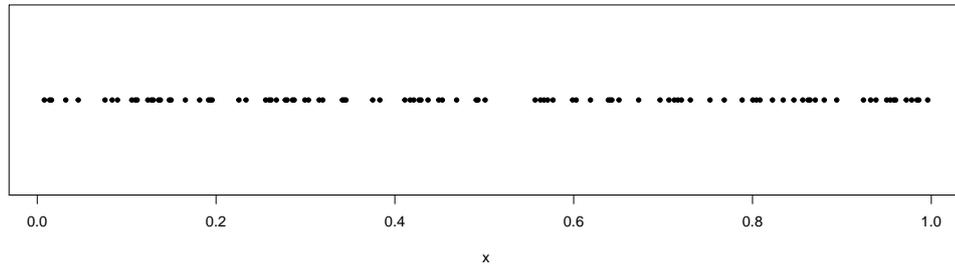


Abbildung 1.6.: 100 Zufallszahlen - mithilfe eines Kongruenzgenerators erstellt

Diese Hypothese gilt es nun zu überprüfen. Meist verwendet man die Daten der Statistik nicht direkt dafür, sondern berechnet aus diesen eine **Teststatistik**⁹.

Der χ^2 -Test

Wir wollen im Zuge dieser Einführung in die Thematik einen speziellen Test diskutieren: Den χ^2 -Test (*Chi-Quadrat-Test*) mit der Prüfgröße

$$\chi^2 = \sum_{j=1}^k \frac{(N_i - E_i)^2}{E_i}$$

und verwenden daher den Wert von χ^2 als Teststatistik.

Wir teilen dazu das Intervall $[0,1]$ in k Klassen und ermitteln die Häufigkeiten N_i von Zahlen in der i -ten Klasse. In unserem Fall teilen wir das Intervall in 4 gleich große Klassen. Die Klasseneinteilung und die Häufigkeiten N_i sind in Abbildung 1.7 dargestellt.

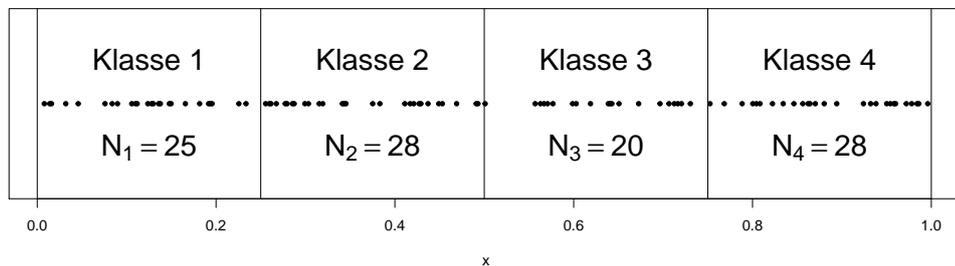


Abbildung 1.7.: 100 Zufallszahlen - Einteilung in 4 Klassen

⁹auch: Prüfgröße oder Testgröße

Nun vergleichen wir die ermittelten Häufigkeiten N_i mit den erwarteten Häufigkeiten E_i . Die E_i bestimmen wir unter Verwendung unserer Nullhypothese: Trifft diese zu, müssten die Zahlen stetig auf dem Intervall $[0,1]$ gleichverteilt sein und somit jede Klasse gleich viele Zahlen - also $\frac{1}{4}$ aller Zahlen - beinhalten: $E_i = \frac{100}{4} = 25$. Wir erkennen, dass die N_i schon recht nahe bei den E_i liegen. Ist das für unseren Zweck bereits ausreichend?

Sicherheit eines Tests und Annahmebereich

Bevor wir diese Frage beantworten können, müssen wir noch einige Überlegungen anstellen: Wie zu Beginn bereits erwähnt, wollen wir anhand einer Stichprobe (Folge von Zufallszahlen) eine Aussage über die Gesamtheit (alle vom Zufallsgenerator erzeugten Zahlen) treffen. Dieser Schluss vom Speziellen auf das Allgemeine ist mit Fehlern behaftet. Wir können unsere Entscheidung nicht mit absoluter Sicherheit treffen und es wäre daher gut, den von uns gemachten Fehler anzugeben. Mit $1 - \alpha$ bezeichnen wir daher die **Sicherheit des Tests** und mit α das **Signifikanzniveau**. Für α werden meist Standardwerte wie 1% oder 5% verwendet - je kleiner α ist, umso sicherer ist unser Test. Was bedeutet das nun in unserem Fall?

Wir wählen α als $\alpha = 5\% = 0.05$ und fragen uns nun, in welchem Bereich $1 - \alpha = 0.95 = 95\%$ aller χ^2 liegen. Ohne hier die Hintergründe zu erläutern, sei bekannt, dass der von uns ermittelte Wert χ^2 näherungsweise χ_{k-1}^2 -verteilt ist, wobei $k - 1 = 3$ der Anzahl an Freiheitsgraden entspricht. Sehen wir uns hierfür die Dichtefunktion dieser Verteilung in Abbildung 1.8 an. 95% aller Werte von χ^2 liegen im Intervall $[0, 7.81]$ und die restlichen 5% außerhalb. Wir bezeichnen das Intervall $[0, 7.81]$ daher als **Annahmebereich**

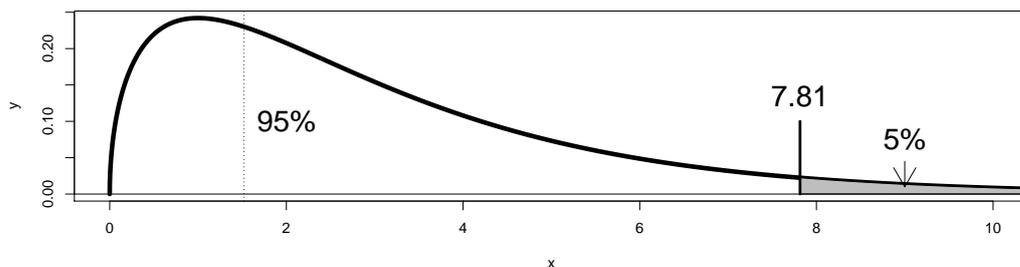


Abbildung 1.8.: Dichtefunktion der χ_3^2 -Verteilung

und $(7.81, \infty)$ als **kritischen Bereich**. Die Berechnung von unserem χ^2 -Wert ergibt 1.52 (strichlierte Linie in Abbildung 1.8). Da dieser Wert im Annahmebereich liegt, werden wir die Nullhypothese annehmen. Unsere Aussage hat eine Sicherheit von 95%. Das ist das selbe Ergebnis, das wir in Abbildung 1.5, Mitte, mithilfe graphischer Mittel erhalten haben. Dort haben wir allerdings bereits erkannt, dass die Zahlen nicht unabhängig voneinander sind. Wir werden daher noch weitere statistische Test benötigen.

Statistische Tests zur Überprüfung von Zufallszahlen

Der χ^2 -Test aus dem vorherigen Unterabschnitt ist ein sogenannter *Anpassungstest*, der überprüft, ob eine Stichprobe einer bestimmten Verteilung, in unserem Fall der stetigen Gleichverteilung auf $[0,1]$, entspricht.

Da wir neben der Verteilung auch die Unabhängigkeit der Zufallszahlen überprüfen wollen, werden wir auch für diesen Zweck statistische Tests benötigen. Auf eine Erläuterung von Beispielen solcher *Unabhängigkeitstests* wird an dieser Stelle mit einem Verweis auf Hagmann [9] und Kolonko [11] verzichtet. Diese Tests wenden oft den eben erläuterten χ^2 -Test auf bestimmte Eigenschaften der erzeugten Zufallszahlen an, um die Unabhängigkeit zu überprüfen.

1.8. Zufallsgeneratoren bei beliebiger Verteilung

Wir gehen nun im Anschluss an die vorhergehenden Abschnitte davon aus, dass wir Pseudozufallszahlen aus dem Intervall $[0,1]$ erzeugen können, die eine Reihe von Tests bestehen und daher einen hohen Grad an Unabhängigkeit und stetiger Gleichverteilung aufweisen.

Für Simulationen benötigt man allerdings oft Zufallszahlen, die einer anderen Verteilung entstammen. Es wurden daher mehrere Verfahren zur Transformation von auf $[0,1]$ stetig gleichverteilten Zufallsvariablen auf Zufallsvariablen beliebiger Verteilungen entwickelt.

Die statistische Programmiersprache R, die in dieser Arbeit zur Anwendung kommt, bringt bereits Zufallsgeneratoren für die in dieser Arbeit notwendigen Verteilungen mit. Wir werden daher im Folgenden diese verwenden und nur eines der möglichen Transformationsverfahren theoretisch betrachten. Für eine Erläuterung anderer Verfahren sei auf Kolonko [11] und Müller et. al. [19] verwiesen.

Inversionsmethode

Mit dieser Methode lässt sich theoretisch jede beliebige Verteilung mit Verteilungsfunktion F simulieren. Für die Simulation ist allerdings die zur Verteilungsfunktion F inverse Funktion F^{-1} vonnöten, die mitunter nicht berechenbar ist. Überlegen wir uns zuerst, wann zu einer Funktion f die Umkehrfunktion f^{-1} existiert.

Betrachten wir eine stetige und streng monotone Funktion, wie beispielsweise die Exponentialfunktion $f(x) = e^x$. In Abbildung 1.9 wird dieser Sachverhalt dargestellt: Erhöht man den Wert von x um Δ , so ergibt sich ein anderer Funktionswert $f(x + \Delta)$ und es gilt $f(x + \Delta) > f(x)$. Die Exponentialfunktion ist streng monoton steigend. Der rechte Funktionsgraph stellt den natürlichen Logarithmus $\ln(y)$ dar. Dieser ist die Umkehrfunktion von e^x , da er zu jedem Wert $y = e^x$ den Wert x zurückliefert. Das ist möglich, weil die Exponentialfunktion streng monoton steigend (und stetig) ist. Die Abszisse und Ordinate der Graphen in Abbildung 1.9 wurden mit Absicht jeweils unterschiedlich bezeichnet, da man sich die Inversenbildung bildlich als Vertauschung der x - und y -Achse vorstellen kann.

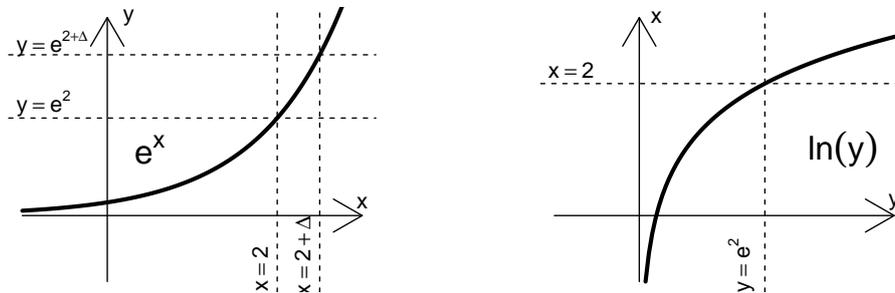


Abbildung 1.9.: Exponentialfunktion e^x und die dazu inverse Logarithmusfunktion $\ln(y)$

Betrachten wir nun als Beispiel die Verteilungsfunktion der Standardnormalverteilung

$$F(x) = \int_{-\infty}^x f(t)dt \quad \text{mit der Dichte } f(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}$$

Für alle $t \in \mathbb{R}$ gilt $f(t) > 0$, da $e^{-\frac{1}{2}t^2}$ nur positive Werte annimmt. Daher ist die Verteilungsfunktion $F(x)$ streng monoton steigend, denn: Als $F(x)$ stelle man sich die Fläche unter $f(t)$ bis $t = x$ vor. Erhöht man x ein wenig um Δ , so erhöht sich auch die Fläche unter der Kurve $f(t)$, d.h. das Integral $F(x)$. Die Verteilungsfunktion F ist deswegen umkehrbar - siehe Abbildung 1.10. Was macht nun die Inverse F^{-1} genau? Angenommen X ist eine standardnormalverteilte Zufallsvariable. Wenn wir auf irgendeinen Wert x , den die Zufallsvariable annimmt, die Verteilungsfunktion F anwenden, erhalten wir die Wahrscheinlichkeit, dass X kleiner oder gleich unserem Wert x ist: $P(X \leq x) = F(x)$. Diese Wahrscheinlichkeit liegt im Intervall $[0,1]$. Die Inverse F^{-1} geht in die Gegenrichtung vor: Wahrscheinlichkeiten, Werte aus $[0,1]$, werden auf die Ausprägungen x_i von X abgebildet.

Aus dieser Überlegung heraus könnte man die Vermutung anstellen, dass, wenn man auf alle Werte von $[0,1]$, also auf eine im Intervall $[0,1]$ stetig gleichverteilte Zufallsgröße Z , die Funktion F^{-1} anwendet, als Ergebnis lauter Werte x_i erhält, die, als Ausprägungen einer Zufallsvariable X aufgefasst, die Verteilung F besitzen.

Satz 1.8.1 Sei $Z \sim S_{0,1}$ und F eine Verteilungsfunktion, dann besitzt

$$X := F^{-1}(Z)$$

die Verteilungsfunktion F . Es gilt daher $P(X \leq x) = F(x)$ für $x \in \mathbb{R}$.

Beweis. siehe Kolonko [11], INVERSIONSPRINZIP, Satz 8.3, S. 87

□

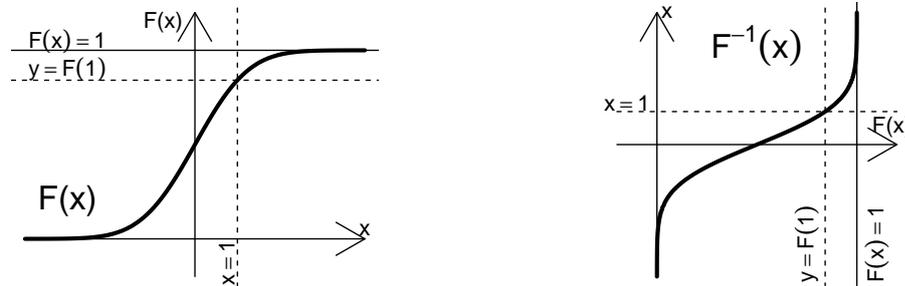


Abbildung 1.10.: Verteilungsfunktion der Standardnormalverteilung und ihre Inverse

Mithilfe dieses Satzes können wir, sofern wir die Inverse der Verteilungsfunktion F kennen oder berechnen können, beliebige Verteilungen simulieren. Doch was machen wir, wenn wir die Inverse nicht ermitteln können?

Verteilungsfunktionen, wie wir sie in Abschnitt 1.5 eingeführt haben, besitzen nach Satz 1.5.3 im Allgemeinen nur die Eigenschaft des monotonen Wachstums - nicht jedoch die des strengen monotonen Wachstums. Daher sind Verteilungsfunktionen allgemein nicht umkehrbar. Es gibt trotzdem eine Möglichkeit die Inversion durchzuführen, indem man eine verallgemeinerte Inverse definiert. Da es nicht für jeden Funktionswert $y = F(x)$ ein eindeutiges x geben muss, sodass die Gleichung erfüllt ist, definiert man $F^{-1}(y)$ als das kleinste x , das obige Gleichung erfüllt.

Definition 1.8.2 (Verallgemeinerte Inverse) Ist F eine Verteilungsfunktion, dann ist

$$G(y) = \inf\{x \in \mathbb{R} | y \leq F(x)\}$$

die verallgemeinerte Inverse von F .

Ist F invertierbar, so ist die Funktion G gleich der Inversen F^{-1} .

Zum Abschluss wollen wir versuchen, mithilfe der Inversionsmethode eine Verteilung zu simulieren. Wir werden dabei die Exponentialverteilung simulieren, da ihre Verteilungsfunktion umkehrbar und die entstehende Transformationsformel nicht so einfach intuitiv fassbar, wie beispielsweise die Transformation auf die Standardnormalverteilung, ist.

Beispiel 1.8.3 Eine exponentialverteilte Zufallsvariable hat die Verteilungsfunktion

$$F(t) = 1 - e^{-\lambda t} \text{ für } \lambda, t > 0$$

Diese Funktion ist streng monoton wachsend und stetig und daher umkehrbar. Die Umkehrfunktion F^{-1} ist gleich

$$F^{-1}(y) = -\frac{1}{\lambda} \cdot \ln(1 - y)$$

Wendet man die Umkehrfunktion auf eine auf $[0,1]$ stetig gleichverteilte Zufallsvariable X an, so erhält man eine exponentialverteilte Zufallsvariable Z .

$$Z = -\frac{1}{\lambda} \cdot \ln(1 - X)$$

1.9. Weitere Grundlagen der Stochastik

Erwartungswert und Varianz

	Erwartungswert	Varianz
X diskret	$E(X) = \sum_i^n x_i \cdot P(X = x_i)$	$V(X) = \sum_i^n (x_i - E(X))^2 \cdot P(X = x_i)$
allgemein $X \in M \subseteq \mathbb{R}$	$E(X) = \int_M x \cdot f(x) dx$	$V(X) = \int_M (x - E(X))^2 \cdot f(x) dx$

Diskrete Verteilungen

Name	Bez.	Parameter	W-Funktion	EW	Varianz
Gleichverteilung	D_n	$n \in \mathbb{N}$	$p(x) = \frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Alternativverteilung	A_θ	$\theta \in (0, 1)$	$p(1) = \theta$ $p(0) = 1 - \theta$	θ	$\theta(1 - \theta)$
Binomialverteilung	$B_{n,\theta}$	$\theta \in (0, 1)$ $n \in \mathbb{N}$	$p(x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}$	$n\theta$	$n\theta(1 - \theta)$
Geometrische Verteilung	G_θ	$\theta \in (0, 1)$	$p(x) = (1 - \theta)^{x-1} \theta$	$\frac{1}{\theta}$	$\frac{1-\theta}{\theta^2}$
Poissonverteilung	P_λ	$\lambda \in \mathbb{R}_+$	$p(x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda}$	λ	λ

Stetige Verteilungen

Name	Bez.	Parameter	Dichte	EW	Varianz
Gleich- verteilung	$S_{a,b}$	$a, b \in \mathbb{R}$	$f(x) = \frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential- verteilung	$Ex(\lambda)$	$\lambda > 0$	$f(x) = \lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normal- verteilung	$N(\mu, \sigma^2)$	$\mu \in \mathbb{R}$ $\sigma > 0$	$f(x) = \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2}$	μ	σ^2

Gütekriterien eines Punktschätzers

Schätzt man einen Parameter p mithilfe einer Stichprobe der Größe n als \tilde{p}_n , dann sind folgende drei Eigenschaften wünschenswert.

Erwartungstreue	Ein Schätzer \tilde{p}_n heißt <i>erwartungstreu</i> , wenn $E(\tilde{p}_n) = p$.
Effizienz	Ein Schätzer \tilde{p}_n heißt <i>effizient</i> , wenn er eine geringere Varianz als ein alternativer Schätzer aufweist.
Konsistenz	Ein Schätzer \tilde{p}_n heißt <i>konsistent</i> , wenn für alle $\varepsilon > 0$ gilt, dass $\lim_{n \rightarrow \infty} P(\tilde{p}_n - p \leq \varepsilon) = 1$.

Der folgende Satz ermöglicht eine einfache Überprüfung der Konsistenz eines Punktschätzers:

Kriterium für Konsistenz Sei \tilde{p}_n ein erwartungstreuer Schätzer und gilt $\lim_{n \rightarrow \infty} V(\tilde{p}_n) = 0$, dann ist \tilde{p}_n konsistent.

Kapitel 2.

Didaktische Überlegungen

Nachdem wir im vorangegangenen Kapitel erste mathematische Grundlagen für stochastische Simulationen gelegt haben, geht es in diesem Kapitel darum, die didaktischen Grundlagen für die Verwendung dieser im Schulunterricht oder in Mathematikkursen allgemein zu behandeln.

Vielleicht haben Sie sich im vorhergehenden Kapitel gewundert, warum die Erzeugung von Zufallszahlen, die annähernd diskret gleichverteilt sind, behandelt wurde, wenn eigentlich nur Zufallszahlen aus $[0,1]$, die approximativ stetig gleichverteilt sind, benötigt werden, da man aus diesen mittels Transformation beliebige Verteilungen erzeugen kann. Eventuell fragen Sie sich auch, warum so viel Zeit, in diesem Fall ausführliche und lange Texte, dafür aufgewandt wurde, um Dinge zu erklären, die schneller und einfacher mit der Formulierung der dahinterliegenden Definitionen und Sätze vollzogen werden könnten.

Die Gründe sind didaktischer Natur und werden in diesem Kapitel erläutert.

2.1. Didaktische Theorie

Will man einen bestimmten Lehrstoff vermitteln, ist es notwendig diesen strukturell zu gliedern. Man steht vor der Entscheidung, in welcher Reihenfolge man welche Teilgebiete vermittelt und in welchen Zwischenschritten man sich dem erwarteten Lernziel nähert, da das gesamte Themengebiet nicht auf einmal vom Lernenden verstanden und gefasst werden kann. Es ist daher notwendig, sich für jedes Thema eine solche „*Sequenz*“ (Wittmann [30], S. 130) zu überlegen.

In Mathematikkursen an Hochschulen wählt man meistens den Ansatz des synthetischen Aufbaus: Ausgehend von ein paar Axiomen werden (Hilfs-)Sätze bewiesen, um darauf aufbauend wiederum (Hilfs-)Sätze zu beweisen und zwar solange bis man die gewünschte Aussage, den gewünschten Satz, erhalten hat.

Dabei orientiert man sich allerdings am aktuellen Wissensstand und die Struktur kann sich stark von der Genese, der Entstehung, des Themengebiets unterscheiden. Dass dieser Ansatz für die Entwicklung von Mathematikkursen im Schulkontext oder einem vergleichbaren Umfeld mitunter nicht sinnvoll ist, wurde bereits Anfang des 19. Jahrhunderts kritisiert.

Im Folgenden werden neben einer alternativen Vorgehensweisen noch weitere didaktische Überlegungen formuliert, um später Gründe für die Verwendung von Simulationen im Stochastikunterricht abzuleiten.

2.1.1. Die genetische Methode

Die genetische Methode schlägt bei der Wahl des Ablaufs eines Mathematikurses die Orientierung an dem Entstehungsprozess der jeweiligen Teildisziplin der Mathematik vor.

Veränderungen gingen immer vom damaligen Stand der Forschung aus und wurden Schritt für Schritt vorangetrieben. Geht man allerdings bei der Vermittlung von neuen Inhalten vom aktuellen Forschungsstand aus, nimmt man dem Lernenden die Chance, den Prozess der Entwicklung der Wissenschaft mitzuerleben. Er wird vor vollendete Tatsachen gestellt, die für ein tiefgreifendes Verständnis oft hinderlich sind, weil der Sinn des Lehrstoffs für den Lernenden nicht klar ist.

Betrachten wir das Beispiel des Grenzwertbegriffs: Erst wenn der Lernende auf Probleme stößt, die ohne eine exakte Formulierung des Grenzwertbegriffs nicht lösbar sind, wird für ihn dieser Begriff eine Bedeutung haben. Davor wird er sich mit einem intuitiven Grenzwertbegriff durchzuschlagen versuchen, wie es auch in der Entwicklung der Mathematik erkennbar ist.

Daher sollte man im Unterricht versuchen über intuitive und heuristische Überlegungen ins Themengebiet einzuführen und mit dem Voranschreiten des Unterrichts schrittweise die Begriffe zu exaktifizieren bis die angestrebte Formulierung erreicht ist. Mathematik ist als Prozess (und nicht als Fertigprodukt) zu verstehen und auch als solcher zu vermitteln. Ähnlich wie die Mathematiker und Mathematikerinnen, die an der Genese der mathematischen Teildisziplinen beteiligt waren, aufbauend auf ihr begrenztes Hintergrundwissen die Mathematik weiterentwickelt haben, soll im Unterricht auf dem Vorverständnis des Lernenden aufgebaut werden. In diesem Fall werden die Lernenden nicht zu Beginn mit einer Fülle an neuen Begriffen überfordert, sondern gewöhnen sich langsam an die Sprechweisen und Begrifflichkeiten und verstehen die anschließenden exakten Formulierungen besser.

Aus dieser Schilderung geht der große Kritikpunkt dieser Methode bereits hervor: Man benötigt viel Zeit. Im Verhältnis zu einem Unterricht, in dem direkt die Ergebnisse präsentiert werden, oft sogar sehr viel mehr Zeit. Doch diese Zeit, so interpretiert Claus in seinem Buch (Claus [6], S. 201f) Wagenschein¹, darf nicht als Verschwendung angesehen werden, denn nur in diesem Fall hat der Lernende die Möglichkeit, selbst Einsichten in die Thematik zu erlangen. Die Erfahrungen die durch Selbsttätigkeit und Eigenständigkeit des Lernenden in genetischem Unterricht gemacht werden, verursachen neben tiefgreifender Erkenntnis des Themas selbst auch eine Steigerung des „*Selbstbewusstseins, das sich auf geistige Leistungen außerhalb der Mathematik bezieht*“ ([6], S. 202).

¹ *Martin Wagenschein* (1896 - 1988): Deutscher Pädagoge, der sich mit der Didaktik der Mathematik und der genetischen Methode im Speziellen beschäftigte.

Felix Klein² schlug bereits 1911 die Abkehr vom logisch-systematischen Unterricht von Mathematik, wie er heute noch an Hochschulen durchgeführt wird, vor und forderte einen genetischen Unterricht, der eine „*historische und [eine] psychologische Komponente*“ ([6], S. 185) besitzt. Die psychologische Komponente schließt unter anderem das Anknüpfen am Vorverständnis und an der intellektuellen Entwicklung des Lernenden mit ein. Als theoretische Grundlage dieser Komponente können beispielsweise die genetische Erkenntnistheorie Piagets³ oder die Weiterentwicklung dieser durch Bruner⁴ angesehen werden. Die historische Komponente, die ich als den Entwicklungsverlauf einer Wissenschaft bezeichnen würde, ist für jeden Themenbereich einzeln zu untersuchen. Dabei müssen im genetischen Unterricht nicht alle Irrwege, die es im Laufe der Entwicklung gegeben hat, auch wirklich vom Lernenden gegangen werden. Die Lehrperson kann, entsprechend der „*indirekte[n] genetische[n] Methode*“ (Claus [6], S. 185), die entscheidenden Ereignisse und Entwicklungen in der Geschichte analysieren und aufbereiten. Heute würde man in diesem Zusammenhang statt von einem „*historischen Prozess*“ (Leuders [17], S. 16) von den „*historischen Kernideen*“ ([17], S. 16) sprechen.

2.1.2. Perspektiven der Mathematik

Die sehr abstrakte Wissenschaft der Mathematik bietet einen großen Spielraum an Betrachtungsweisen. Wir werden uns im Folgenden in Anlehnung an Leuders ([17] S. 19 - 29) drei dieser Blickwinkel ansehen, um später deren Rolle für den Stochastik-Unterricht zu erläutern.

Platonismus

Der Platonismus geht davon aus, dass Mathematik etwas nicht Fassbares, etwas Abstraktes ist, das „*außerhalb von Zeit und Raum*“ ([17], S. 20) existiert. Es handelt sich dabei um Strukturen und Muster, die es zu entdecken gilt. Man kann aber auch eine Brücke zur Realität schlagen, indem man, wie Leuders Stewart zitiert, Mathematik als „*die Wissenschaft der Muster*“ ([17], S. 20) und die Natur als Anwendung „*so ungefähr jedes existierende[n] Muster[s]*“ ([17], S. 20) ansieht. Wie wir bereits im Zuge der genetischen Methode gesehen haben, entwickelt sich die Beschreibung dieser Muster schrittweise und ist als Prozess zu sehen. Dieser Prozess ist aber nicht eindeutig: Es ist möglich verschiedene Wege zu gehen. Daher greift es nach Leuders zu kurz, die „*mathematische Wahrheit als auskristallisiert in den Definitionen und Sätzen der etablierten Fachgebiete*“ ([17], S. 20) und daher als abgeschlossen anzusehen. Entgegen dieser Sichtweise, sollte Mathematik im Unterricht nicht einfach nachvollzogen, sondern auch der Erfindungsprozess, das schrittweise Entwickeln von Definitionen, thematisiert werden.

² Felix Christian Klein (1849 - 1925): Deutscher Mathematiker, der sich im Besonderen mit der genetischen Methode im Mathematikunterricht beschäftigte und die Bedeutung dieser hervorhob.

³ Jean Piaget (1896 - 1980): Schweizer Psychologe, dessen genetische Epistemologie als eine der bedeutendsten Theorien des Lernens angesehen wird.

⁴ Jérôme Seymour Bruner (1915 -): US-amerikanischer Psychologe

Physikalismus

Der Physikalismus betrachtet die Strukturen und Muster, im Gegensatz zum oder als Weiterentwicklung des Platonismus, als aus der Natur abstrahierte Objekte. Aus diesem Blickwinkel kann man reale Beobachtungen und Erfahrungen nach Leuders dazu verwenden, um einen formalen Nachweis zu ersetzen, weil man in der Schule „*nicht immer zu den mathematischen Fundamenten vorstößt*“ ([17], S. 22). Wobei man, meiner Meinung nach, thematisieren sollte, dass es sich um keine absolut gewonnenen Wahrheiten handelt und eine formale Behandlung notwendig wäre.

Konstruktivismus

Der Konstruktivismus betrachtet Mathematik, zumindest in dem Sinne in dem sie beschrieben wird, als Erfindung und nicht als Entdeckung. Sie könnte demnach auch ganz anders aussehen. Dieser Sachverhalt ist auch in der Genese der Mathematik erkennbar. Leibniz definierte ein Differential anders als wir es heute (im klassischen Sinne) tun. Erst später konnten die Probleme seiner Theorie, die infinitesimalen Elemente, durch die Non-Standard-Analyse gelöst werden, sodass man Leibniz'sche Differentiale ohne Zweifel verwenden kann. Es ist also möglich einen Sachverhalt auf verschiedene Arten zu charakterisieren.

Konstruktivismus im Bezug auf Lernen von Mathematik bedeutet, dass jeder Lernende sein Wissen, seine Vorstellung, von Mathematik selbst konstruiert. Dieses Konzept ist in Einklang mit aktuellen Bestrebungen der Didaktik, Wissen „*selbst [zu] erfahren, aus[zu]probieren, [zu] experimentieren und [es] immer in eigene Konstruktionen ideeller oder materieller Art zu überführen*“ (Reich [21]).

2.1.3. Heuristik im Mathematikunterricht

Der Mathematikunterricht auf Universitäten ist meist nach der Sequenz „Definiton-Satz-Beweis“ aufgebaut. Doch entgegen der Vermutung, die durch diese Vorgangsweise entstehen könnte, steht ein mathematischer Beweis „*nicht am Anfang der mathematischen Arbeit, sondern am Ende*“ (Leuders [16], S. 27). Auch Mathematiker und Mathematikerinnen bedienen sich zu Beginn heuristischer Werkzeuge, „*bildlicher Darstellungen*“ ([16], S. 28) und suchen „*Beispiele, um die Plausibilität von Vermutungen zu prüfen*“ ([16], S. 28). Sie experimentieren daher, bevor sie streng formal arbeiten. Deswegen kann man ohne Bedenken auch im Schulunterricht heuristische Methoden zur ersten Entwicklung einer Idee des Themas anwenden.

2.2. Didaktik der Stochastik

In diesem Abschnitt wollen wir versuchen, die eben erarbeiteten didaktischen Theorien auf die Vermittlung der Stochastik allgemein bzw. auf die Verwendung stochastischer Simulationen im Unterricht im Speziellen umzulegen.

Zuerst wollen wir dazu die Entwicklung der Stochastik als Wissenschaft betrachten.

2.2.1. Genese der Stochastik

Es ist anzunehmen, dass, seitdem es Spiele und Orakel gibt, auch über Wahrscheinlichkeiten nachgedacht wird. Erste Zufallsgeräte waren nach Tietze et. al. [27] Astragali⁵ und Würfel. Diese wurden bereits in prähistorischen Zeiten zum Spielen und zur Zukunfts- und Entscheidungsfindung, als sogenanntes Orakel, verwendet.

Die Geburtsstunde der Wahrscheinlichkeitsrechnung wird dennoch meist im Briefwechsel zwischen *Pascal*⁶ und *Fermat*⁷ im 17. Jahrhundert begründet. Sie entstand in der „*Außeinandersetzung mit praktischen Problemen*“ ([27], S. 78) und wurde durch die Verbreitung von Glücks- und Kartenspielen vorangetrieben.

Ein Beispiel für die Bedeutung von Spielen in der Entwicklung der Stochastik ist das *Drei-Würfel-Problem* von *Chevalier de Méré*⁸, der sich wunderte, warum beim Wurf dreier Würfel die Augensummen 11 und 12 nicht gleich häufig waren, obwohl es, seiner Behauptung nach, jeweils sechs verschiedene Möglichkeiten gab, aus drei Würfeln diese Augensummen zu erhalten. *Pascal* löste das Problem, indem er auch die Permutationen der drei Würfel berücksichtigte.

Ein weiteres Beispiel ist das *Teilungsproblem*. Man stellte sich die Frage, wie man in einem Spiel, das aufgrund einer höheren Macht zu beenden ist, am Ende die Einsätze fair aufteilt. *Pascal* und *Fermat* schlugen als Erste Vorgehensweisen vor, die die möglichen Ausgänge einer weiteren Partie des Spiels, also die Zukunft, in die Aufteilung miteinbezogen.

In den darauffolgenden 300 Jahren wurde eine Menge an neuen Erkenntnissen geliefert: *Huygens*⁹ publizierte seinen „*Traktat über Glücksspiele*“. *Bernoulli*¹⁰ verfasste kombinatorische Abhandlungen und formulierte das nach ihm benannte *Gesetz der Großen Zahlen*. *De Moivre*¹¹, der als „*Experte für Glücksspiele*“ ([27], S. 79) eine beratende Rolle einnahm, entdeckte die *Normalverteilung* und formulierte den nach ihm benannten Grenzwertsatz. Es folgten *Laplace*¹² mit seinem Wahrscheinlichkeitsbegriff bei Symmetrievoraussetzungen, der *Laplace-Wahrscheinlichkeit*, und *von Mises*¹³, der Wahrscheinlichkeit, unter Beachtung von Einschränkungen, als Grenzwert von relativen Häufigkeiten betrachtete.

Diese Entwicklungen geschahen, ohne dass die Wahrscheinlichkeitsrechnung auf einem formalen Fundament stand. Die Erkenntnisse entstanden alle ohne eine mathematische Fundierung der Überlegungen. Aus diesem Grund nahm *Hilbert*¹⁴ im Jahr 1900 diese Lücke,

⁵Nach [27] handelt es sich dabei um Knochen von Schafen und Ziegen

⁶*Blaise Pascal* (1623 - 1662): französischer Mathematiker

⁷*Pierre de Fermat* (1607 - 1665): französischer Mathematiker

⁸*Chevalier de Méré* (1607 - 1684): französischer Glücksspieler

⁹*Christiaan Huygens* (1629 - 1695): niederländischer Mathematiker

¹⁰*Jakob Bernoulli* (1654 - 1705): schweizer Mathematiker

¹¹*Abraham de Moivre* (1667 - 1754): französischer Mathematiker

¹²*Pierre Simon Laplace* (1749 - 1827): französischer Mathematiker

¹³*Richard Edler von Mises* (1883 - 1953): österreichischer Mathematiker

¹⁴*David Hilbert* (1862 - 1943): deutscher Mathematiker

unter dem Namen „*Mathematische Behandlung der Axiome der Physik*“, in seine Liste der 23 wichtigsten ungelösten Probleme der Mathematik auf. (vgl. Kütting [13], S. 54) Der Name deutet bereits darauf hin, dass die Wahrscheinlichkeitsrechnung damals noch nicht der Mathematik, sondern der Physik, zugeordnet wurde.

Man war im Folgenden darum bemüht, „*mathematische[r] Begriffe von inhaltlichen Begriffen aus der Realität [loszulösen]*“ (Tietze et.al. [27], S. 78). Von Mises scheiterte mit seiner Definition von Wahrscheinlichkeit als (analytischem) Grenzwert von relativen Häufigkeiten. Erst 1933 schuf Kolmogoroff¹⁵ eine axiomatische Grundlage für die Wahrscheinlichkeitsrechnung. (vgl. Definitionen im ersten Kapitel dieser Arbeit)

2.2.2. Vorverständnis und intellektuelle Entwicklung der Lernenden

Erste empirische Untersuchungen an Kindern bezüglich der Entwicklung eines Wahrscheinlichkeitsbegriffs wurden nach Kütting [13] von Piaget und Inhelder¹⁶ durchgeführt und die Ergebnisse 1951 in ihrem Buch „*La genèse de l'idée de hasard chez l'enfant*“ publiziert. Auf dieser Veröffentlichung aufbauend wurden später noch weitere Untersuchungen durchgeführt. Manche Ergebnisse Piagets und Inhelders wurden revidiert, andere bestätigt und wiederum andere sowohl revidiert als auch bestätigt. Wir wollen im Folgenden trotzdem versuchen, allgemeine Aussagen aus der Fülle an Untersuchungen herauszulösen.

Als Zusammenfassung der im Laufe der Zeit entstandenen Studien kann man festhalten, dass Kinder im schulfähigen Alter¹⁷ zwischen zufälligen und deterministischen Vorgängen unterscheiden können. Es ist allgemein zu erkennen, dass Kinder bereits früh mit dem Phänomen Zufall vertraut sind und daher anzunehmen, dass „*[d]iese Bekanntschaft [...] nur aus Begegnungen im Alltagsleben herrühren [kann]*“ ([13], S. 102). Stochastikunterricht sollte daher in jeder Schulstufe möglich sein.

Andererseits weisen die Untersuchungen einen signifikanten Unterschied zwischen den Klassenstufen der Kinder aus. Denn im Gegensatz zur eben geschilderten *ersten Intuition*¹⁸ ist die Bildung eines *quantitativen Wahrscheinlichkeitsbegriffs* erst in Piagets *Phase der formalen Operationen*¹⁹ möglich, da hierfür das Verständnis von Proportionen und kombinatorischen Operationen notwendig ist und dieses erst im formal-operativen Stadium entwickelt wird.

Ein weiterer, für diese Arbeit wichtiger, Aspekt, den die Untersuchungen zutage gebracht haben, ist, dass sich durch geeignete Versuche, das heißt durch Experimente mit Kindern, die Wahrscheinlichkeitseinschätzungen dieser Kinder verbessern können.

¹⁵Alexander Nikolajewitsch Kolmogoroff (1903 - 1987): russischer Mathematiker

¹⁶Bärbel Inhelder (1913 - 1997): Schweizer Psychologin

¹⁷In manchen Publikationen ist 4 Jahre und in anderen 7 Jahre als Alter für Kinder, ab dem Zufälligkeit für sie zugänglich ist, zu finden.

¹⁸Nach [13] forderte aufgrund der unterschiedlichen Entwicklung der beiden Aspekte eine Unterscheidung zwischen einer *ersten Intuition* und dem eigentlichen *Begriff*.

¹⁹Piaget teilt die Entwicklung der Intelligenz in vier aufeinander aufbauende Phasen. Beim Stadium der formalen Operationen handelt es sich um die letzte, und damit elaborierteste, aller Phasen.

2.2.3. Frequentistische Wahrscheinlichkeit

Können wir aufgrund von Symmetrieanahmen davon ausgehen, dass jedes *Elementarereignis* die gleiche Wahrscheinlichkeit des Auftretens besitzt, dann wird meist auf die *Klassische Wahrscheinlichkeitsdefinition* (Definition 1.2.3, Abschnitt 1.2, Kapitel 1) zurückgegriffen, um Ereignissen Wahrscheinlichkeiten zuzuordnen. Ein Beispiel hierfür ist ein Würfel, dessen Seitenflächen alle die gleiche Wahrscheinlichkeit besitzen, gewürfelt zu werden oder die Urne mit 1000 Losen zu Beginn dieser Arbeit (Abschnitt 1.1, Kapitel 1). Was macht man allerdings, wenn diese sogenannte *Laplace-Annahme* der Gleichwahrscheinlichkeit, wie beispielsweise bei einem unförmigen Würfel, nicht getroffen werden kann?

Einen Ausweg liefern in diesem Fall die *Gesetze der Großen Zahlen*. Bereits im 18. Jahrhundert wurde das Phänomen beobachtet, dass sich die *relativen Häufigkeiten* von Ereignissen *stabilisieren*. Beispielsweise kommen bei einem Würfel die verschiedenen Augenzahlen, auf lange Sicht gesehen, ungefähr gleich oft vor - ihre relativen Häufigkeiten nähern sich dem Wert $\frac{1}{6}$.

Dieses Phänomen wird als *empirisches Gesetz der Großen Zahlen* bezeichnet und kann dazu verwendet werden, die objektiven Wahrscheinlichkeiten, die den Ereignissen innewohnen, abzuschätzen. Dazu führt man das Zufallsexperiment mehrmals durch, in anderen Worten: man *simuliert* es, und legt dann die Wahrscheinlichkeiten mithilfe der relativen Häufigkeiten fest. Man könnte irgendein unsymmetrisches Objekt, wie einen Reißnagel, hundert Mal in die Luft werfen und die Wahrscheinlichkeit, dass er auf die Spitze fällt, mithilfe der relativen Häufigkeit dieses Ereignisses schätzen. Man erkennt die, aus mathematischer Sicht unübliche, Vorgangsweise: Die „*Wahrscheinlichkeit wird wie eine physikalische Größe gemessen*“. (Tietze et. al. [27], S. 84) und die „*Ergebnisse der Messungen werden genauer je größer der Umfang der Messserie ist*“ ([27], S. 84). Gewinnt man Wahrscheinlichkeiten auf diese Art, so spricht man von *frequentistischen Wahrscheinlichkeiten*.

Man kann diesen Ansatz natürlich auch in Fällen, in denen die *Laplace-Annahme* zutrifft, verwenden, um schneller zu einem Ergebnis zu kommen. Es wird ein Zufallsexperiment mehrfach simuliert und anschließend die Wahrscheinlichkeit des interessierenden Ereignisses geschätzt. Aufgrund dieser experimentellen Herangehensweise wird die *stochastische Simulation* auch als *experimentelle Stochastik* bezeichnet. Wir werden solche Simulationen im nächsten Kapitel durchführen. Dabei werden wir zuerst Schätzwerte für die Wahrscheinlichkeiten berechnen, um diese anschließend mit den analytisch berechneten Werten zu vergleichen.

Es muss uns allerdings bewusst sein, dass es sich beim *empirischen Gesetz der Großen Zahlen* um kein Gesetz der Stochastik, sondern um „*eine von der Theorie unabhängige Beschreibung der Beobachtungen [...], die bei Wiederholungen von Zufallsexperimenten gemacht werden können*“ ([27], S. 84), handelt. Wir müssen uns auch vor Augen halten, dass unsere Schätzung subjektiv ist und von unserer Versuchsreihe abhängt: „*Ein anderer Experimentator würde vermutlich eine etwas andere Versuchsserie bekommen und damit*

einen etwas anderen Schätzwert für p wählen.“ (Büchter [5], S. 177)

Doch was bedeutet es, dass sich die Werte stabilisieren? Eine mathematische Grundlage dieser Vorgangsweise liefert das (*Starke*) *Gesetz der Großen Zahlen*, das eine *fast sichere* Konvergenz der relativen Häufigkeiten gegen die Wahrscheinlichkeiten postuliert. Es handelt sich um eine *stochastische* Konvergenz, also keine Konvergenz im analytischen Sinn: Es ist nicht sicher, dass die relativen Häufigkeiten gegen die Wahrscheinlichkeit konvergieren - es ist nur sehr wahrscheinlich.

2.2.4. Stochastische Simulation und Ziele des Stochastikunterrichts

Vor über 20 Jahren wies Trauerstein ([28], 1990) bereits auf die gute Eignung von stochastischer Simulation im Unterricht der Sekundarstufe I und II hin. In seinen Augen ermöglicht diese mit „*relativ elementaren Methoden*“ ([28], S. 2) die Bearbeitung von Problemen, die sonst (im Schulkontext) nicht lösbar wären. Man benötigt nur den Begriff der (relativen) Häufigkeit und des (arithmetischen) Mittelwerts. Daher könnte man sie bereits früh einführen.

In einer aktuellen Publikation von Biehler et. al. ([4], 2010), die sich mit den Leitideen im Stochastikunterricht beschäftigt und auf einer „*Expertise von Fachleuten in der Stochastik verschiedener Bereiche*“²⁰ ([4], S. 1) aufbaut, wird *stochastische Simulation* als eigene Kompetenz verstanden. Diese wird als Fähigkeit der Lernenden verstanden, Simulationen zu nutzen, „*um mit stochastischen Situationen zu experimentieren und Näherungslösungen in komplexen Situationen zu gewinnen*“ ([4], S. 3).

Von der eigenen Kompetenz abgesehen, spielt stochastische Simulation in zwei der anderen sechs Grundkompetenzen, bei der *Beurteilenden Statistik* und dem *Gesetz der Großen Zahlen*, eine bedeutende Rolle. Ich nehme an, dass diese Publikation die zentralen Ideen der Stochastik der verschiedensten Bereichen abbildet und behauptet daher, dass die stochastische Simulation aufgrund ihrer häufigen Nennung als Grundidee im Schulunterricht Einzug finden sollte bzw. müsste.

Der Blick in den österreichischen Lehrplan zeigt uns allerdings ein anderes Bild: Obwohl bereits in der 6. Schulstufe relative Häufigkeiten gelehrt werden, ist an keiner Stelle des Lehrplans der Unter- oder Oberstufe Allgemeinbildender Höherer Schulen stochastische Simulation erwähnt. Ebenso ist in den Schulbüchern nie explizit von stochastischer Simulation die Rede. Ich vermute allerdings, dass im Schulumfeld oft stochastische Simulation betrieben wird, ohne dass man sich darüber bewusst ist oder ohne explizit von dieser zu reden und diese implizit im Lehrplan verankert ist.

²⁰schulpraktische, universitär-fachstochastische und wirtschaftswissenschaftliche Perspektiven wurden miteinbezogen

2.3. Stochastische Simulationen im Mathematikunterricht

Ausgehend von den bisherigen Überlegungen versuchen wir nun Gründe zu finden, warum man stochastische Simulationen im Unterricht einbringen soll.

2.3.1. Zehn Gründe für den Einsatz von stochastischen Simulationen im Mathematikunterricht

Ausrichtung an der Genese der Stochastik Die Stochastik hat sich anhand praktischer Probleme und dem experimentellen Umgang mit diesen entwickelt. Aus damaliger Sicht verblüffende Unregelmäßigkeiten bei oftmaligem Würfeln mit mehreren Würfeln oder bei (Glücks-)Spielen im Allgemeinen haben den Anstoß zu einer abstrakt-formalen Auseinandersetzung mit der Thematik gegeben.

Es wurde von Anfang an stochastische Simulation betrieben: Zuerst hat man durch lange Versuchsreihen und mithilfe der beobachteten relativen Häufigkeiten bestimmte Regelmäßigkeiten vermutet und sich im Folgenden abstrakt mit dem Problem auseinandergesetzt.

Will man Unterricht an der Genese der Stochastik ausrichten, kommt man nicht daran vorbei, die Lernenden zuerst experimentieren zu lassen. Erst danach ist es notwendig und sinnvoll eine Theorie zu entwickeln, die die Beobachtungen aus den Versuchen beschreibt.

Heuristische Arbeitsweise in der Mathematik Wie bereits erwähnt (Unterabschnitt 2.1.3), ist ein Beweis nicht der Ausgangspunkt mathematischer Überlegungen sondern deren Ziel. Am Weg zu diesem Ziel bedienen sich Mathematikerinnen und Mathematiker experimenteller Mittel. Simulationen eignen sich daher besonders, um die Arbeitsweise eines Wissenschaftlers im Schulunterricht zu thematisieren und zu integrieren, sowie diese dadurch zu verstehen.

Anknüpfen an Vorverständnis und Alltag der Lernenden Wie wir in Unterabschnitt 2.2.2 über das Vorverständnis der Lernenden gesehen haben, bringen bereits Kinder im Alter von (spätestens) 7 Jahren ein Verständnis für zufällige Phänomene mit. Da diesen Kindern eine formal-abstrakte Denkweise noch Schwierigkeiten bereitet, ist anzunehmen, dass dieses Zufallsbewusstsein in gemachten Beobachtungen im Alltag zu begründen ist.

Stochastische Simulationen können in diesem Fall sehr gut am Vorverständnis der Lernenden anknüpfen. Sie ermöglicht eine Beschäftigung mit stochastischen Sachverhalten in der Art, die dem Lernenden vertraut ist - durch Beobachtung. Ich bezeichne Simulationen daher als die intuitivste und natürlichste Herangehensweise, um sich mit der Thematik auseinanderzusetzen.

Aufdecken von Fehlern Da wir unser erstes Bewusstsein von Zufälligkeit aufgrund von (wenigen) Beobachtungen entwickeln, ziehen wir oft zu schnelle Schlüsse und es entstehen fehlerhafte Vorstellungen.

Um diesen Fehlbildern, die durch Beobachtungen entstanden sind, entgegenzuwirken, ist es sicherlich vorteilhaft diese mit neuen entgegengesetzten Beobachtungen kollidieren zu lassen. Für die Erzeugung solcher den eigenen Fehlvorstellungen widersprechenden Beobachtungen können Simulationen durchgeführt werden.

In diesem Fall liegt der Vorteil von Simulationen gegenüber einer analytischen Behandlung wiederum in der Anknüpfung am Vorverständnis und an der gewohnten und natürlichen Herangehensweise an stochastische Probleme. Man beseitigt Fehlvorstellungen nicht durch eine neue Herangehensweise an das Problem, sondern mit denselben Mitteln, die für die falschen Bilder verantwortlich sind: durch Beobachtungen.

Interesse wecken und motivieren Simulationen von Fragestellungen aus dem Umfeld der Lernenden können dazu beitragen, das Interesse dieser zu wecken bzw. diese zu motivieren. Aufgrund der gemachten Beobachtungen können Fragen entstehen, die an den Hintergründen der Beobachtungen ausgerichtet sind: Warum haben wir diese Beobachtungen gemacht und wie sind sie entstanden?

Eine besondere Stellung können in diesem Fall Paradoxien, d.h. Sachverhalte, die im Widerspruch zu unseren Erwartungen stehen, einnehmen. Die Beobachtungen sind oft so verblüffend, dass sie ein sehr großes Motivationspotential in sich tragen können und man dazu verleitet wird, dem Grund nachzugehen. Stochastische Simulationen spielen dabei oft eine Vorreiterrolle: Mit ihrer Hilfe kann man das paradoxe Verhalten aufzeigen, um anschließend die kausalen Zusammenhänge zu untersuchen und das Problem gegebenenfalls analytisch zu lösen.

Sinn der Wahrscheinlichkeitsrechnung verstehen Durch einführende Simulationen und anschließende analytische Lösung des Problems, sollte man ziemlich bald die Schnelligkeit und Eleganz der analytischen Herangehensweise erkennen. Diesen aufgezeigten Effizienzgewinn kann man dazu verwenden, die klassische Wahrscheinlichkeitsrechnung rechtzuerfüllen.

Tieferer Einblick und Verständnis der Wahrscheinlichkeitsrechnung Simulationen schlagen eine Brücke zwischen Theorie und Praxis, zwischen Stochastik (Wahrscheinlichkeitstheorie und Statistik) und den gemachten Beobachtungen. Sie kann daher auch dazu beitragen, ein vertieftes Verständnis in die Thematik zu erlangen.

Man ist beispielsweise in der Lage den Erwartungswert neben der formalen Definition auch als solchen zu verstehen: Wenn ich hundert Mal A durchführe, dann habe ich im Durchschnitt $E(A)$ zu erwarten.

Wie wir im nächsten Teil der Arbeit sehen werden, können wir auch andere Inhalte der Theorie wie die Grenzwertsätze, Hypothesentests und Konfidenzintervalle frequentistisch deuten und erhalten somit einen Zusammenhang zu in der Realität gemachten Beobachtungen.

Modellbildung im Mathematikunterricht „Die Simulation ist ein wichtiges Verfahren zur Modellbildung in Theorie und Praxis“ (Tietze et.al. [27], S. 130) und schlägt eine Brücke zwischen Empirie und stochastischer Theorie.

Neue Türen öffnen Mithilfe stochastischer Simulation sind Probleme lösbar, deren analytische Behandlung nicht möglich ist. Im Schulkontext kann man sie auch dazu verwenden, früh in den Themenbereich der Stochastik einzuführen, da nur relative Häufigkeiten

und arithmetische Mittel zur Berechnung der geschätzten Wahrscheinlichkeiten benötigt werden. Simulationen ermöglichen daher eine Auseinandersetzung mit stochastischen Fragestellungen basierend auf den mathematischen Grundlagen der Sekundarstufe I. Ferner ist es möglich Modelle zu untersuchen, die im Schulunterricht nicht oder nur rudimentär behandelt werden, wie Wartezeitprobleme oder geometrische Wahrscheinlichkeiten.

Perspektiven der Mathematik Simulationen lassen eine Betrachtung der Stochastik aus den drei zuvor angeführten Perspektiven zu.

Durch Beobachtung und Experimentieren versucht man Muster und Strukturen zu erkennen. Diese Muster sind aus der Natur abstrahierte Objekte und die Erkenntnisse aus den Beobachtungen können (vorerst) eine formale Auseinandersetzung mit der Thematik ersetzen (Physikalismus).

Ausgehend von diesen empirischen Erkenntnissen und erneuten Simulationen versucht man Schritt für Schritt die Beobachtungen zu beschreiben (Platonismus).

Aufgrund der Verbindung zwischen Theorie und Empirie kann jeder Lernende seine eigene Vorstellung von der Thematik entwickeln (Konstruktivismus).

2.3.2. Phasen der stochastischen Simulation im Unterricht

Tietze et.al. ([27], S. 131) teilen den Einsatz stochastischer Simulationen im Unterricht in verschiedene Phasen. Da eine detaillierte Gliederung der Simulationsbeispiele im folgenden Kapitel in diese Phasen den Umfang dieser Arbeit sprengen würde und weil die Beispiele nur Ideen darstellen sollen, die Lehrenden noch genügend Gestaltungsspielraum einräumen, wird darauf verzichtet.

Für Lehrende, die auf die strukturgebenden Phasen von Tietze et.al. zurückgreifen wollen, sind sie im Folgenden aufgelistet ([27], S. 131):

1. Vorüberlegungen und Vermutungen über die stochastische Struktur des Problems
2. Aufstellen eines vorläufigen Modells
3. Auswahl eines geeigneten Zufallsgenerators
4. Durchführung der Simulation
5. Auswertung der Ergebnisse (Schätzwerte)
6. Interpretation der Ergebnisse

Teil II.

Durchführung von Simulationen

Vorbereitungen für Simulationen im Schulunterricht

Wenn Sie denken, dass stochastische Simulationen im Schulunterricht mit viel Aufwand und dem Einsatz von Computern verbunden sind, dann hoffe ich Sie im Folgenden vom Gegenteil zu überzeugen, denn stochastische Simulationen sind mit elementaren Mitteln durchführbar.

Für viele Aufgaben benötigen wir keine Zufallsgeneratoren, wie aus Kapitel 1, die in kürzester Zeit Unmengen an (Pseudo)-Zufallszahlen erzeugen. In der Regel sind Würfel ausreichend. Es sind Würfel mit 4, 6, 8, 12 und 20 durchnummerierten Seitenflächen²¹ in Spielzeugläden erhältlich (vgl. Abbildung 2.1)²² Zusammen mit einer Münze, mit den

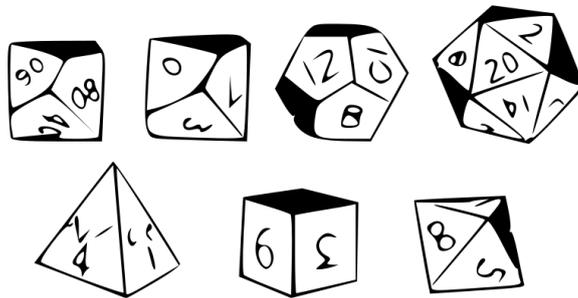


Abbildung 2.1.: Würfel mit den Seitenflächen: 10, 10, 12, 20 (v.l) sowie 4, 6, 8 (v.l)

zwei Möglichkeiten *Kopf* und *Zahl*, kann man somit auf einfachem Weg 2, 4, 6, 8, 12 und 20 verschiedene zufällige Werte erzeugen.

Benötigt man eine andere Anzahl an verschiedenen zufälligen Werten, kann man mit einer Urne und darin enthaltenen nummerierten Zetteln Abhilfe schaffen. Ist der Aufwand für die Anfertigung der Zettel aufgrund der Anzahl zu groß, gibt es die Möglichkeit, im Unterricht Tabellen mit Zufallszahlen zu verwenden.

Erst wenn sehr viele Zufallszahlen benötigt oder diese für ausführliche Berechnungen verwendet werden, sollte man den Computer einsetzen, um diese zu erzeugen bzw. automatisiert bestimmte Kenngrößen zu berechnen.

Im Folgenden werden Beispiele zur stochastischen Simulation, thematisch in Kapiteln gegliedert, vorgestellt. Die Grundlagen zu den Zufallszahlen aus Kapitel 1 sind für den Lehrenden sicherlich von Interesse, für den Lernenden hingegen nicht erforderlich. Unter den Beispielen sind nicht nur jene, die Stochastik-Grundlagen anwenden, sondern auch ein paar, die keine Kenntnisse der Stochastik voraussetzen und Brücken zu anderen Teilgebiete

²¹ Geometrisch betrachtet handelt es sich bei diesen Würfeln um vollkommen regelmäßige Polyeder, die sogenannten platonischen Körper: Tetraeder (Vierflächer), Hexaeder = Kubus (Sechsfächer), Oktaeder (Achtflächer), Dodekaeder (Zwölfächer) und Icosaeder (Zwanzigflächer)

²² Sämtliche Abbildungen nicht-mathematischer Sachverhalte wurden mithilfe des Programms *tuxpaint* (tuxpaint.org), den darin enthaltenen Stempeln und freien Cliparts von openclipart.org erstellt.

ten der Mathematik schlagen.

Kapitel 3.

Veranschaulichung des empirischen Gesetzes der Großen Zahlen

Zu Beginn wollen wir das in Unterabschnitt 2.2.3 erläuterte *empirische Gesetz der Großen Zahlen* graphisch darstellen. In diesem Unterabschnitt haben wir die Möglichkeit, Wahrscheinlichkeiten ausgehend von langen Versuchsreihen zu schätzen, kurz erörtert und die so erhaltenen Wahrscheinlichkeiten *frequentistisch* genannt.

Wir wollen nun das Werfen einer Münze mit den möglichen Ereignissen *Kopf* (K) und *Zahl* (Z) aus diesem Blickwinkel betrachten.

Man könnte eine Klasse von 28 Lernenden in 4 Gruppen aufteilen, anschließend jede dieser Gruppen 50 Mal eine Münze werfen und nach jedem Wurf die relative Häufigkeit der Köpfe

$$r_n(K) = \frac{H_n(K)}{n}$$

bestimmen lassen, wobei $H_n(K)$ das absolute Auftreten von Kopf bei n erfolgten Versuchen ist. Trägt man diese relativen Häufigkeiten graphisch auf, könnte sich ein ähnliches Bild wie in Abbildung 3.1 auf Seite 43 ergeben. Die relativen Häufigkeiten für das Ereignis *Kopf* liegen am Ende nahe bei der Wahrscheinlichkeit $p = \frac{1}{2}$, die man dem Ereignis mithilfe der *klassischen Wahrscheinlichkeitsdefinition* (Satz 1.2.3) zuschreiben würde. Wenn wir weitere Versuche durchführen, verhalten sich die relativen Häufigkeiten meist ähnlich wie in Abbildung 3.2, Seite 44. Es ist eine Stabilisierung der relativen Häufigkeiten $r_n(K)$ zu erkennen. Diese Beobachtung nennt man *empirisches Gesetz der Großen Zahlen*.

Empirisches Gesetz der Großen Zahlen Mit wachsender Versuchszahl n stabilisieren sich die relativen Häufigkeiten $r_n(K)$.

Wir erhalten also in der Regel mit mehr Versuchsdurchführungen immer genauere Schätzwerte für p . Bei diesem Gesetz handelt es sich allerdings um keine mathematische Aussage, da es alleine durch Beobachtungen entstanden ist.

Die mathematische Exaktifizierung bildet das *Starke Gesetz der Großen Zahlen*, das besagt, dass die relativen Häufigkeiten $r_n(K)$ für ins Unendliche wachsende Versuchsdurchführungen n mit Wahrscheinlichkeit 1 gegen den Erwartungswert $\mu = E(r_n(K)) = p$ konvergieren. Betrachten wir eine Folge von Zufallsvariablen $(X_i)_{i \in \mathbb{N}}$, die das Auftreten

von Kopf im i -ten Versuch beschreiben und die Zufallsvariable \bar{X}_n , die dem arithmetischen Mittelwert von $(X_i)_{i \leq n}$ entspricht und somit die relativen Häufigkeiten $r_n(K)$ darstellt, können wir folgenden Satz formulieren:

Satz 3.0.1 (Starke Gesetz der Großen Zahlen) Sei $(X_i)_{i \in \mathbb{N}}$ eine Folge von Zufallsvariablen mit $E(X_i) = \mu$ und \bar{X}_n das arithmetische Mittel der X_i . Dann gilt

$$P\left(\lim_{n \rightarrow \infty} \bar{X}_n = \mu\right) = 1$$

Man sagt: \bar{X}_n konvergiert fast sicher gegen μ .

Wenn wir also im Folgenden Wahrscheinlichkeiten mithilfe von relativen Häufigkeiten stochastischer Simulationen schätzen, berufen wir uns auf die *Gesetze der Großen Zahlen*.

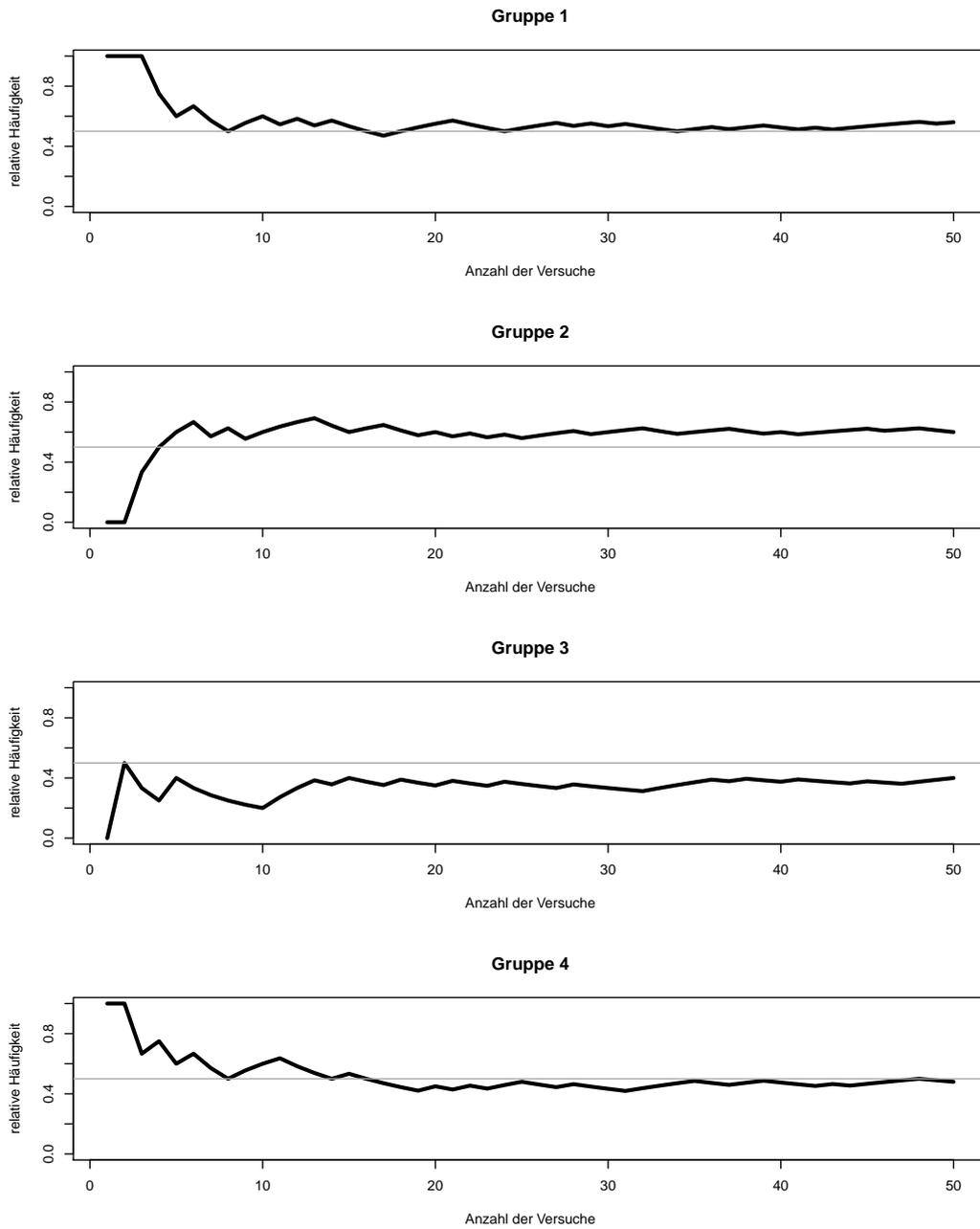


Abbildung 3.1.: Gruppenarbeit - relative Häufigkeit von Kopf beim Werfen einer Münze

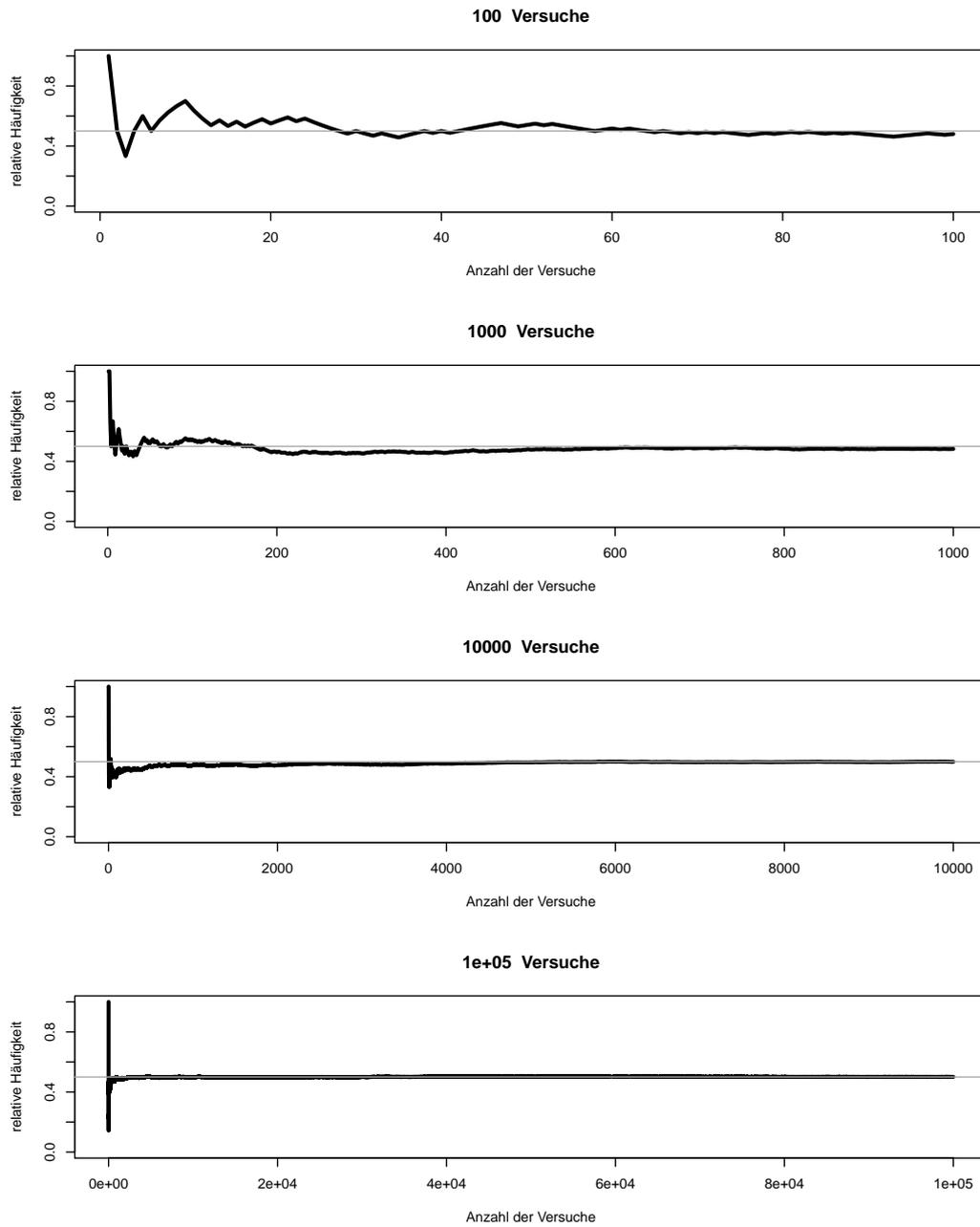


Abbildung 3.2.: Relative Häufigkeit von Kopf bei 100, 1.000, 10.000 und 100.000 Münzwürfen

Kapitel 4.

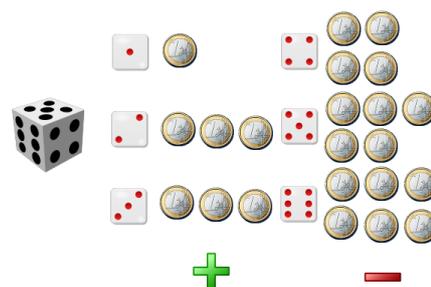
Beispiele zum Erwartungswert

Im Folgenden werden drei Beispiele zum Schätzen von Erwartungswerten mithilfe stochastischer Simulation vorgestellt.

Während das erste Beispiel zur Einführung des Begriffs *Erwartungswert* gedacht ist und geschulte Stochastiker die Fragestellung fast aus dem Stehgreif beantworten können, ist beim zweiten Beispiel, ohne zuvor Überlegungen anzustellen, kaum mit einer zufriedenstellenden Lösung zu rechnen. Das Ergebnis des letzten Beispiels widerspricht sogar unserer Intuition. Mit dem Schwierigkeitsgrad des Beispiels steigt auch der analytische Lösungsaufwand. Der Aufwand der Simulationen ist hingegen bei allen Beispielen in etwa gleich groß.

4.1. Einführung des Begriffs Erwartungswert mithilfe von Simulation

Aufgabenstellung Jemand bietet Ihnen ein Spiel an: Wenn Sie 1, 2 oder 3 würfeln, zahlt Ihnen diese Person die Augenzahl in Euro. Wenn Sie allerdings 4, 5 oder 6 würfeln, müssen Sie ihr die Augenzahl bezahlen. Wie hoch ist Ihr zu erwartender Gewinn? Ist das Spiel fair?



Lösung mithilfe von Simulation

Dass das Spiel nicht fair ist, liegt auf der Hand: Wir gewinnen gleich wahrscheinlich, wie wir verlieren, müssen bei einer Niederlage allerdings immer mehr bezahlen, als wir bei

einem Sieg erhalten. Doch mit welchem Verlust wir jede Runde zu rechnen haben, ist für Neulinge auf dem Gebiet der Stochastik nicht sofort ersichtlich.

Um ein Gefühl für Erwartungswerte zu bekommen, könnte man das Spiel spielen. Eine Klasse von 28 Lernenden wird in 14 Zweier-Gruppen aufgeteilt und jede dieser Gruppe spielt das Spiel 10 Mal. Ausgehend von den Gewinnen X in jedem Spiel, kann jede Gruppe den Gesamtgewinn und den durchschnittlichen Gewinn pro Spiel $\bar{X}^{(10)}$ berechnen.

Sammelt man, wie in folgender Tabelle, alle Gruppenergebnisse und berechnet man den

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
GG	-29	-7	-9	-12	-25	-7	1	-14	-8	-15	-13	-8	-4	-34
G/S	-2.9	-0.7	-0.9	-1.2	-2.5	-0.7	0.1	-1.4	-0.8	-1.5	-1.3	-0.8	-0.4	-3.4

Mittelwert der Gewinne pro Spiel (G/S), erhält man einen noch besseren Schätzwert $\tilde{E}(X) = \bar{X}^{(150)}$ für den zu erwartenden Gewinn $E(X)$.

Wir erhalten $\bar{X}^{(140)} \approx -1.31$ und wissen daher, dass wir jedes Spiel in etwa -1.31 Euro verlieren. Das Spiel ist somit nicht fair und wir werden nicht darauf einsteigen, es sei denn, wir erhalten vor jedem Würfelwurf 1.31 Euro von unserem Gegner bzw. unserer Gegnerin.

Analytische Lösung

Man erkennt sofort die Eleganz und Schnelligkeit der analytischen Lösung:

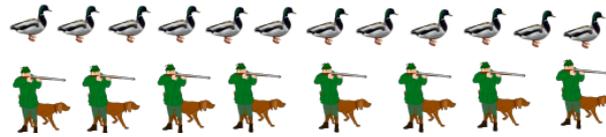
$$E(X) = (1 + 2 + 3 - 4 - 5 - 6) \cdot \frac{1}{6} = \frac{1}{6} \cdot (-9) = -\frac{9}{6} = -\frac{3}{2} = -1.5$$

Die Simulation der Fragestellung dient in erster Linie dem Erlangen einer außermathematischen Interpretation des Erwartungswertes.

4.2. Treibjagdproblem

Aufgabenstellung 8 Jäger schießen gleichzeitig auf 12 Enten. Die Jäger treffen stets, aber sie vereinbaren vorher nicht, wer auf welche Ente schießt. Wie viele Enten werden wahrscheinlich überleben?¹

¹Dieses Beispiel ist oft in der Variante mit 8 Jägern und 10 Enten zu finden. Zur besseren Durchführung in der Schule wurde es zu dieser Variante abgewandelt: Man wirft einen Würfel und eine Münze. Bei Kopf entspricht die Augenzahl der Zufallszahl ($Kopf + 3 = 3$), bei Zahl wird die Augenzahl zusätzlich um den Wert 6 erhöht ($Zahl + 4 = 6 + 4 = 10$) und somit alle Zahlen von 1 bis 12 erzeugt.



Lösung mithilfe von Simulation

Stellt man sich die Enten in einer Reihe aufgereiht und von 1 bis 12 durchnummeriert vor und nimmt man an, dass die Schüsse unabhängig voneinander sind und keine Ente bevorzugt getroffen wird, kann man man das Experiment recht einfach simulieren: Jeden Schuss bildet man nach, indem man eine Zufallszahl von 1 bis 12 erzeugt. Die Zahl entspricht der Ente, die vom jeweiligen Jäger getroffen wurde. Um acht Schüsse zu simulieren erzeugt man demnach acht solcher Zufallszahlen. Anschließend kann man recht einfach die Anzahl der überlebenden Enten ermitteln.

Führt man diesen Vorgang zehn Mal, z.B. in zehn Gruppen, durch, kann man die Ergebnisse wie in Tabelle 4.1 sammeln. Entspricht X den Anzahlen der überlebenden Enten in einem Versuch, dann kann man deren Mittelwert \bar{X} als Schätzung für den Erwartungswert $E(X)$ heranziehen. Wir erhalten $\bar{X} = 6.1$. Es werden daher voraussichtlich 6 Enten

	J1	J2	J3	J4	J5	J6	J7	J8	E tot	E lebend
Versuch 1	11	4	5	7	11	3	11	12	6	6
Versuch 2	8	8	1	3	3	9	5	10	6	6
Versuch 3	6	9	12	5	10	12	3	8	7	5
Versuch 4	2	4	5	1	5	11	5	6	6	6
Versuch 5	8	6	3	10	9	10	2	9	6	6
Versuch 6	5	10	8	10	7	7	10	1	5	7
Versuch 7	6	9	9	6	11	6	3	1	5	7
Versuch 8	2	4	7	8	5	11	4	6	7	5
Versuch 9	4	8	4	6	10	2	11	5	7	5
Versuch 10	11	5	5	6	11	11	5	10	4	8

Tabelle 4.1.: Tabelle mit den Ergebnissen der Simulation und der daraus ermittelten Anzahl an toten und überlebenden Enten

überleben.

Analytische Lösung

Zur analytischen Lösung versetzen wir uns in die Lage einer Ente: Ich überlebe einen Schuss, wenn irgendeine andere Ente getroffen wird und sterbe, wenn ich getroffen werde. Es handelt sich bei einem Schuss daher um einen Alternativversuch mit den Alternativen *überleben* bzw. *sterben* und den Wahrscheinlichkeiten

$$p(\text{Ente überlebt}) = \frac{11}{12} \quad p(\text{Ente stirbt}) = \frac{1}{12}$$

Betrachten wir nun diesen Alternativversuch acht Mal, d.h. für jeden Jäger einmal, durchgeführt, handelt es sich wegen der Unabhängigkeit der einzelnen Schüsse um einen Binomialversuch mit dem Baumdiagramm aus Abbildung 4.1. Um insgesamt zu überleben,

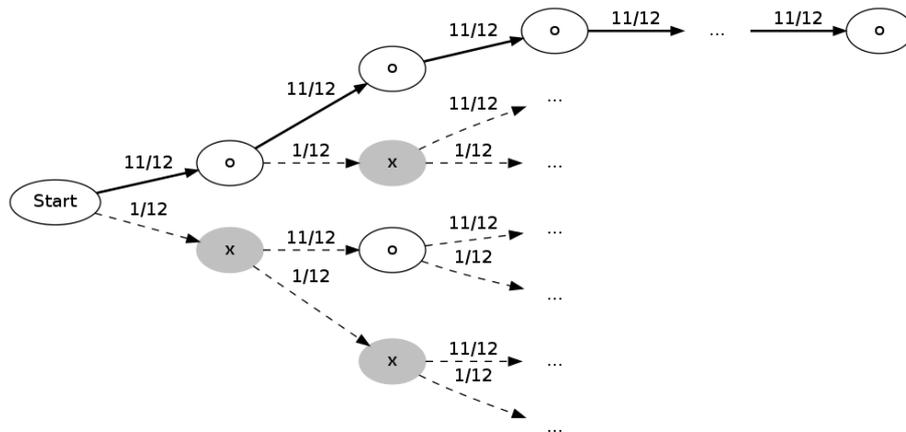


Abbildung 4.1.: Baumdiagramm des Zufallsexperiments aus Sicht einer Ente. o - Schuss verfehlt Ente, x - Ente wird getroffen

muss man jeden einzelnen Schuss überleben. Das entspricht dem obersten Pfad in der Abbildung. Eine Ente überlebt daher mit einer Wahrscheinlichkeit von

$$\left(\frac{11}{12}\right)^8 \approx \frac{1}{2}$$

Bezeichnen wir mit X die Anzahl der überlebenden Enten und definieren wir diese als Summe von Indikatorvariablen

$$X = \sum_{i=1}^{12} X_i \quad X_i = \begin{cases} 1 & \text{Ente überlebt} \\ 0 & \text{Ente stirbt} \end{cases}$$

mit folgenden Wahrscheinlichkeiten

$$P(X_i = 1) = \left(\frac{11}{12}\right)^8 \approx \frac{1}{2} \quad P(X_i = 0) = 1 - \left(\frac{11}{12}\right)^8 \approx \frac{1}{2}$$

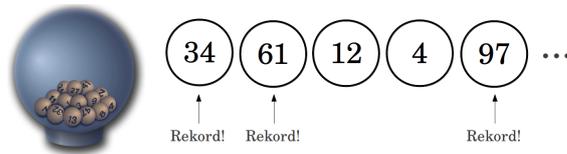
dann können wir den Erwartungswert von X folgendermaßen berechnen

$$\begin{aligned} E(X) &= E\left(\sum_{i=1}^{12} X_i\right) = \sum_{i=1}^{12} E(X_i) = \\ &= \sum_{i=1}^{12} \left(1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0)\right) \approx \sum_{i=1}^{12} \left(\frac{1}{2}\right) = 6 \end{aligned}$$

Es ist zu erwarten, dass ungefähr die Hälfte der Enten überlebt. Das ist das selbe Ergebnis, das wir mittels Simulation erhalten haben.

4.3. Lotto-Kugel-Paradoxon

Aufgabenstellung Wir haben eine Lotterie-Urne mit Kugeln, die von 1 bis 1000 durchnummeriert sind, zur Verfügung. Der Reihe nach ziehen wir die Kugeln heraus, bis sich keine mehr in der Urne befindet. Jedes Mal, wenn wir eine Nummer ziehen, die höher als alle bereits gezogenen Nummern ist, rufen wir „Rekord“. Mit wie vielen „Rekord“-Rufen ist in etwa zu rechnen? Mit wie vielen „Rekord“-Rufen ist bei von 1 bis n durchnummerierten Kugeln allgemein zu rechnen?



Anmerkung Überlegen Sie sich kurz mit wie vielen „Rekord“-Rufen bei $n = 1000$ zu rechnen ist und notieren Sie sich, für einen späteren Vergleich mit der analytischen Lösung, ihren Tipp.

Lösung mithilfe von Simulation

Die Simulation des Beispiels für beliebige n ist relativ einfach durchzuführen, dauert für große n aber mitunter relativ lange.

Wir werden nun das Beispiel für unterschiedliche n ($n = 10, 100$ mittels Urne² und $n = 1000$ am Computer) simulieren und versuchen Gesetzmäßigkeiten in den erhaltenen Ergebnissen zu finden. Das Ziel sollte sein, anhand der vier Simulationen eine allgemeine Formel für die „Rekord“-Rufe für beliebige n zu vermuten.

Die geschätzten Anzahlen X von „Rekord“-Rufen für verschieden viele Kugeln n in der Urne aufgrund von Versuchsreihen der Länge 100 sind folgender Tabelle zu entnehmen.

	n=10	n=100	n=1000
X	3	5	7

Nun könnte man diese Werte in einem Koordinatensystem auftragen (Abbildung 4.2). Um die Frage nach einer allgemeinen Formel zu beantworten, sucht man eine Funktion, die die Punktvolke am besten beschreibt. Der Logarithmus-Naturalis der Anzahl an Kugeln n , d.h. $\ln(n)$, liefert eine gute Näherung. Wir vermuten daher, dass bei Ziehen aus n Lottokugeln mit $\ln(n)$ „Rekord“-Rufen zu rechnen ist.

Anmerkung Vergleichen Sie den Wert $\ln(1000) \approx 7$ mit Ihrem anfangs geschätzten Wert. Die meisten irren sich in ihrer Schätzung gewaltig.

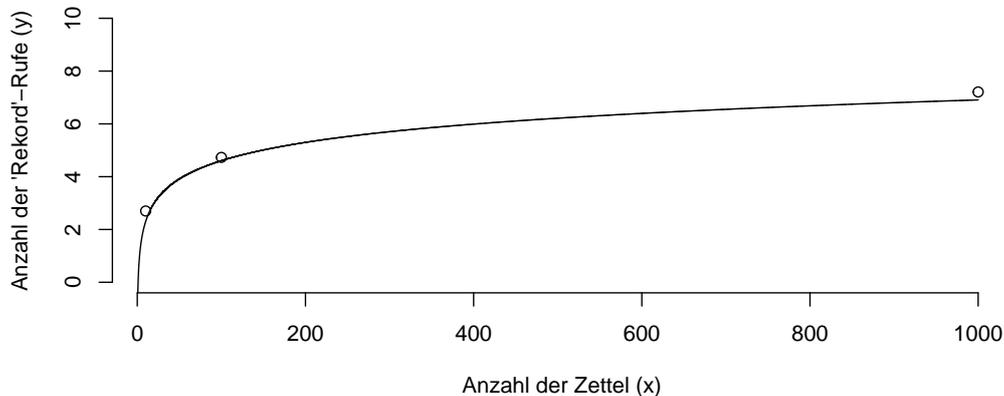


Abbildung 4.2.: Anzahl der „Rekord“-Rufe (y -Werte) in Abhängigkeit von der Anzahl der Zettel (x -Werte) als Punkte (x, y) aufgetragen. Näherung durch den Funktionsgraphen $\ln(x)$.

Analytische Lösung

Betrachten wir die Anzahl der „Rekord“-Rufe X als Summe von Indikatorvariablen

$$X = \sum_{i=1}^n X_i \quad X_i = \begin{cases} 1, & \text{Rekord bei } i\text{-ter Zahl} \\ 0, & \text{sonst} \end{cases}$$

mit den Wahrscheinlichkeiten

$$P(X_i) = \frac{1}{i}$$

²In dieser Arbeit werden alle Simulationen mithilfe des Computers durchgeführt. Hat man allerdings keinen Computer zur Verfügung, kann man die Simulation für kleine n , sagen wir $n \leq 100$, durch Ziehen aus einer Urne durchführen und die Ergebnisse für große n direkt angeben.

und den Erwartungswerten

$$E(X_i) = 1 \cdot P(X_i = 1) + \overbrace{0 \cdot P(X_i = 0)}^{=0} = \frac{1}{i}$$

Daraus kann man den Erwartungswert von X wie folgt berechnen

$$E(X) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n \frac{1}{i}$$

Da das händische Berechnen der Summen sehr aufwendig ist, approximieren wir die Summe durch das Integral der Funktion $f(x) = \frac{1}{x}$ (Abbildung 4.3)

$$\sum_{i=1}^n \frac{1}{i} \approx \int_1^n \frac{1}{x} dx = \ln(n)$$

Mithilfe des Logarithmus kann man sehr einfach die erwartete Anzahl an Rekordrufen

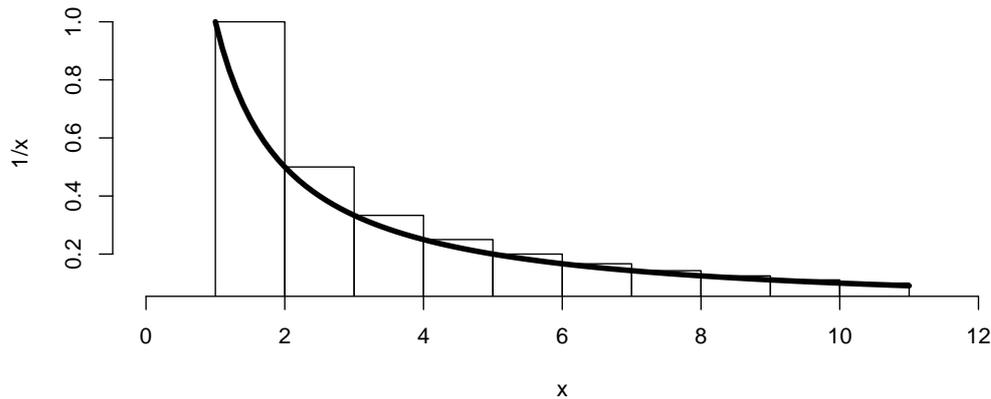


Abbildung 4.3.: Approximation von $\sum_{i=1}^n \frac{1}{i}$ durch $\int_1^n \frac{1}{x} dx = \ln(n)$.

berechnen. Weil der geschätzte Wert immer kleiner als der exakte ist (vgl. Abbildung 4.3), verstehen sich die folgenden Werte als auf die nächste ganze Zahl aufgerundet.

	n=10	n=100	n=1000
X	3	5	7

Kapitel 5.

Auflösen von Paradoxien

„Einer der großen Vorteile der Wahrscheinlichkeitsrechnung ist der, daß man lernt, dem ersten Anschein zu misstrauen“

Pierre-Simon Laplace
(Kütting [14], S.71)

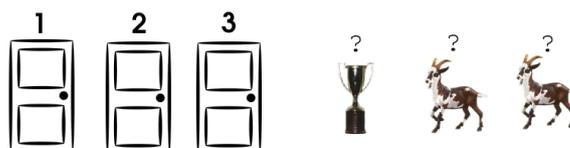
Paradoxien wird, wie bereits im Abschnitt über stochastische Simulationen im Mathematikunterricht (Abschnitt 2.3) erwähnt, ein besonderes Motivationspotenzial zugeschrieben, da man, aufgrund der verblüffenden Ergebnisse, dem Grund dieser auf die Spur kommen will.

Ich bezweifle, dass eine rein analytische Auflösung eines Paradoxons für alle Lernenden zugänglich ist und denke, dass Simulationen der Nachvollziehbarkeit der Lösung dienen können. Im Folgenden wird exemplarisch an zwei Paradoxien gezeigt, wie man diese mithilfe von Simulationen auflösen kann.

Als erstes Beispiel wird ein bekanntes Paradoxon aus dem Umfeld der bedingten Wahrscheinlichkeiten und der Unabhängigkeit von Ereignissen vorgestellt und anschließend geht das zweite Beispiel der Frage nach, wie wahrscheinlich es ist, dass Menschen, um Sie herum (in Schule, Beruf oder Freizeit) am selben Tag Geburtstag haben?

5.1. Drei-Türen-Paradoxon

Aufgabenstellung Ein Moderator oder eine Moderatorin bietet Ihnen drei geschlossene Türen zur Auswahl an: Hinter einer der drei Türen befindet sich ein Preis, hinter den anderen beiden allerdings nicht. Nachdem Sie eine Tür gewählt haben, öffnet der Moderator bzw. die Moderatorin eine andere Tür hinter der sich kein Preis befindet und Sie haben nun zwei Möglichkeiten: Entweder Sie bleiben bei Ihrer Auswahl (Strategie A) oder Sie wechseln auf die andere ungeöffnete Tür (Strategie B). Mit welcher Variante ist eher ein Gewinn zu erwarten?



Tür mit Geschenk			Gewählte Tür			Gezeigte Tür		
Tür 1	Tür 2	Tür 3	Tür 1	Tür 2	Tür 3	Tür 1	Tür 2	Tür 3
○	○	○	○	○	○	○	○	○
Tür bei Wechsel			Wechsel		Kein Wechsel			
Tür 1	Tür 2	Tür 3	gewonnen	verloren	gewonnen	verloren		
○	○	○	○	○	○	○		

Tabelle 5.1.: Simulationstafel für das Drei-Türen-Problem

Anmerkung Wir wollen also wissen, ob das Öffnen einer Tür durch den Moderator oder die Moderatorin meine Gewinnchancen beeinflusst. Oder anders formuliert: Habe ich dadurch einen Erkenntnisvorsprung?

Lösung mithilfe von Simulation

Wir werden das Spiel unter der Annahme, dass jede Tür mit gleicher Wahrscheinlichkeit den Preis versteckt, mehrmals durchführen und uns ansehen, welche Strategie auf Dauer die bessere ist. Dazu verwenden wir die Tabelle 5.1, welche zum Ausdrucken im Anhang auf Seite 111 zu finden ist.

Einerseits stellt man bei längeren Versuchsreihen fest, dass die relative Häufigkeit der Gewinne bei einem Wechsel (Strategie B) größer ist als bei keinem (Strategie A). Sie liegen im ersten Fall in etwa bei $2/3$ und im zweiten bei ungefähr $1/3$.

Andererseits erkennt man auch ein Muster bei Strategie B: Wenn man anfangs die falsche Tür wählt, gewinnt man bei einem Wechsel immer. Der Moderator oder die Moderatorin führt einen indirekt zur richtigen Tür, indem er oder sie die zweite falsche Tür ausschließt. Nur wer anfangs die richtige Tür wählt, verliert.

Das Öffnen einer Tür erzeugt daher einen Erkenntnisvorsprung und der Moderator bzw. die Moderatorin verrät mir mehr als auf den ersten Blick ersichtlich.

Analytische Lösung

Erkennt man im Zuge der Simulation das zuvor genannte Muster der Strategie B, gelangt man auch schneller zu einer analytischen Lösung.

Strategie A Ich gewinne, wenn ich anfangs die Tür mit dem Preis wähle.

$$P(\text{Gewinn}) = \frac{1}{3}$$

Strategie B Ich gewinne, wenn ich anfangs eine der beiden Türen ohne Preis wähle.

$$P(\text{Gewinn}) = \frac{2}{3}$$

Fehler der üblichen Argumentation

Die zuvor erhaltenen Ergebnisse widersprechen wiederum unserer Intuition. Üblicherweise würde man argumentieren, dass bei einem Wechsel das ursprüngliche Verhältnis der Wahrscheinlichkeiten (jede Tür versteckt mit Wahrscheinlichkeit $1/3$ den Preis) erhalten bleibt. Wir müssen aber auch die Information, die wir durch das Öffnen einer Tür erhalten, miteinbeziehen. Es erfolgt ein Übergang zu bedingten Wahrscheinlichkeiten und das ursprüngliche Verhältnis muss nicht gleich bleiben.

Betrachten wir den Fall, dass wir die Tür 1 gewählt haben und die Tür 2 geöffnet wurde. Die Wahrscheinlichkeiten, dass sich der Gewinn hinter Tür 1 (G_1), 2 (G_2) oder 3 (G_3) versteckt,

$$P(G_1) = \frac{1}{3} \quad P(G_2) = \frac{1}{3} \quad P(G_3) = \frac{1}{3}$$

gehen nun über zu folgenden bedingten Wahrscheinlichkeiten

$$P(G_1|O_2) = \frac{1}{3} \quad P(G_2|O_2) = 0 \quad P(G_3|O_2) = \frac{2}{3}$$

mit der Bedingung O_2 , dass die Tür 2 geöffnet wurde.

Wenn die zweite Tür geöffnet wurde, befindet sich der Schatz mit Sicherheit nicht dahinter ($P(G_2|O_2) = 0$). Die beiden anderen Wahrscheinlichkeiten kann man mithilfe des *Bayes'schen Theorems* berechnen. Man erkennt, dass sich das ursprüngliche Verhältnis der Wahrscheinlichkeiten verändert hat.

5.2. Geburtstagsparadoxon

Aufgabenstellung Wie groß ist die Wahrscheinlichkeit, dass in einer Klasse von 25 Schülern und Schülerinnen zwei (oder mehr) am selben Tag Geburtstag haben?¹

Anmerkung Nicht zu verwechseln mit: Wie groß ist die Wahrscheinlichkeit, dass jemand anderes am selben Tag wie ich Geburtstag hat!

¹Das Beispiel kann je nachdem, wie viele Schüler und Schülerinnen eine Klasse umfasst, abgewandelt werden.



Lösung mithilfe von Simulation

Wir wollen nun unter den Annahmen, dass alle Geburtstage unabhängig voneinander sind und jeder Geburtstag gleichwahrscheinlich ist, mittels Simulation eine Antwort finden.

Zwei beliebige Personen

Zuerst könnte man in der Klasse überprüfen, ob zwei Leute am gleichen Tag Geburtstag haben. Wie wir später sehen werden, ist dieser Fall bei Klassengrößen ab 23 SchülerInnen gar nicht so unwahrscheinlich, wie man es vorab erwarten würde. Die motivierende Frage könnte daher lauten: Ist es Zufall, dass zwei Leute am gleichen Tag Geburtstag haben oder war es zu erwarten?

Da man kaum die Möglichkeit hat, die Überprüfung in weiteren Klassen durchzuführen, schafft eine Simulation mithilfe von Zufallszahlen Abhilfe.

Wir wollen im Folgenden den 29. Februar ignorieren und nummerieren die Tage vom 1. Januar bis zum 31. Dezember von 1 bis 365 durch. Mithilfe eines Computers oder anhand der Zufallstafel² auf Seite 112 erzeugt man anschließend eine Folge von 25, bzw. der Klassengröße entsprechend vielen, Zufallszahlen und überprüft, ob es in dieser Stichprobe Doppelte gibt.

Führt man den Vorgang beispielsweise 1000 Mal³ durch, ergeben sich folgende Häufigkeiten.

	Doppelte	keine Doppelten
absolute Häufigkeiten	576	424
relative Häufigkeiten	0.576	0.424

Die Wahrscheinlichkeit für mindestens zwei übereinstimmende Geburtstage beträgt daher in etwa 0.576.

Der Grund für die Diskrepanz zwischen Intuition und tatsächlichem Ergebnis liegt, wie es Stadler in [26] anmerkt, vermutlich in unserem Begriff von Zufälligkeit. Wir assoziieren mit ihm meist eine relativ regelmäßige Verteilung.

²In der Zufallstafel wählt man eine Folge von 25 aufeinanderfolgenden Zahlen

³Bei einer Klassengröße von 25 Lernenden genügt es, wenn jeder die Simulation vier Mal durchführt.

Außerdem ist zu vermuten, dass man zu sehr von sich selbst ausgeht.

Ich und eine andere Person

Gesucht ist im vorherigen Beispiel nämlich die Wahrscheinlichkeit, dass *beliebige* zwei Personen am selben Tag Geburtstag haben und nicht, wie groß die Wahrscheinlichkeit ist, dass jemand anderes auch denselben Geburtstag wie *ich* hat. Ich gebe meinen Geburtstag als den 4. Januar (=4) vor und überprüfe tausend Mal bei einer Gruppengröße von 25, ob eine der anderen 24 Personen, deren Geburtstage ich wiederum zufällig erzeuge, am selben Tag wie ich Geburtstag hat. In der folgenden Tabelle sind die Anzahlen der beiden Fälle aufgelistet.

	Doppelte	keine Doppelten
absolute Häufigkeiten	66	934
relative Häufigkeiten	0.066	0.934

Die Wahrscheinlichkeit beträgt in etwa nur 0.066, d.h. 6.6%, dass jemand am selben Tag Geburtstag hat wie ich.

Analytische Lösung

Zwei beliebige Personen

Zuerst ermitteln wir mithilfe der klassischen Wahrscheinlichkeitsdefinition die Wahrscheinlichkeit, dass bei n Personen alle an verschiedenen Tagen Geburtstag haben.

$$P(A_n^c) = \frac{365 \cdot 364 \cdot \dots \cdot (365 - n + 1)}{365^n}$$

Da dies dem Komplimentärereignis des gesuchten Ereignisses entspricht, können wir die gesuchte Wahrscheinlichkeit folgendermaßen ermitteln.

$$P(A_n) = 1 - P(A_n^c)$$

Für $n = 25$ gilt

$$P(A_{25}) \approx 0.569$$

Ich und eine andere Person

Mein Geburtstag entspricht irgendeiner Zahl von 1 bis 365. Bei einer Gruppe von 25 Leuten, muss ich daher 24 Mal überprüfen, ob jemand am selben Tag Geburtstag hat. Unter den anfangs gemachten Annahmen geschieht das bei jedem Vergleich (Alternativversuch) mit einer Wahrscheinlichkeit von $1/365$ und ich kann die einzelnen Vergleiche als Kette unabhängiger Alternativversuche ansehen. Die Anzahl $X^{(n)}$ der Übereinstimmungen bei einer Gruppengröße von $n + 1$ ist daher binomialverteilt mit Parameter $p = 1/365$.

Zuerst berechnet man wiederum die Gegenwahrscheinlichkeit, d.h. dass es keine Übereinstimmungen gibt

$$P(X^{(n)} = 0) = \binom{n}{0} \cdot p^0 \cdot (1-p)^n = (1-p)^n$$

und erhält folgendermaßen die gesuchte Wahrscheinlichkeit

$$P(X^{(n)} > 0) = 1 - P(X^{(n)} = 0) = 1 - (1-p)^n$$

Für $n = 24$ gilt

$$P(X^{(24)} > 0) \approx 0.064$$

Kapitel 6.

Veranschaulichung von Grenzwertsätzen

Da die Grenzwertsätze in der Schule meist nicht bewiesen werden, wollen wir nun mithilfe stochastischer Simulationen ein intuitives Verständnis dieser Sätze erarbeiten.

6.1. Approximation der Binomialverteilung

Da die Berechnung der Koeffizienten der Binomialverteilung $B_{n,p}$ händisch sehr mühsam und sogar mit dem Computer für große n sehr aufwendig ist, versucht man die Binomialverteilung durch eine stetige Verteilung zu approximieren und somit Wahrscheinlichkeiten einfacher und schneller (näherungsweise) ermitteln zu können. Für diesen Zweck bietet sich die Normalverteilung an, da die Binomialverteilung für ins Unendliche wachsende n im Grenzfall dieser Verteilung entspricht.

Satz 6.1.1 (Grenzverteilungssatz von de Moivre - exakt) Für eine Zufallsgröße $S_n \sim B_{n,p}$ und $-\infty \leq a < b \leq \infty$ gilt

$$\lim_{n \rightarrow \infty} P \left(a \leq \frac{S_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \leq b \right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Dieser Satz ermöglicht die Approximation der *Binomialverteilung* durch die *Normalverteilung*, die im folgenden für den Schulunterricht relevanten locker formulierten Satz dargestellt wird.

Satz 6.1.2 (Approximationssatz von de Moivre, [18], S. 95) Wenn n genügend groß ist, dann kann eine Binomialverteilung mit den Parametern n und p näherungsweise durch eine Normalverteilung mit den Parametern $\mu = n \cdot p$ und $\sigma = \sqrt{n \cdot p \cdot (1-p)}$ ersetzt werden.

Wenn man die Binomialverteilung $B_{n,p}$ betrachtet, kann man dieser die Normalverteilung $N(n \cdot p, n \cdot p \cdot (1-p))$ überstülpen. Umso mehr n erhöht wird, desto genauer wird die Näherung (vgl. Abbildung 6.1).

6.2. Zentraler Grenzwertsatz

Eine Verallgemeinerung dieser Überlegungen stellt der zentrale Grenzwertsatz dar, den wir nun mithilfe stochastischer Simulation veranschaulichen wollen.

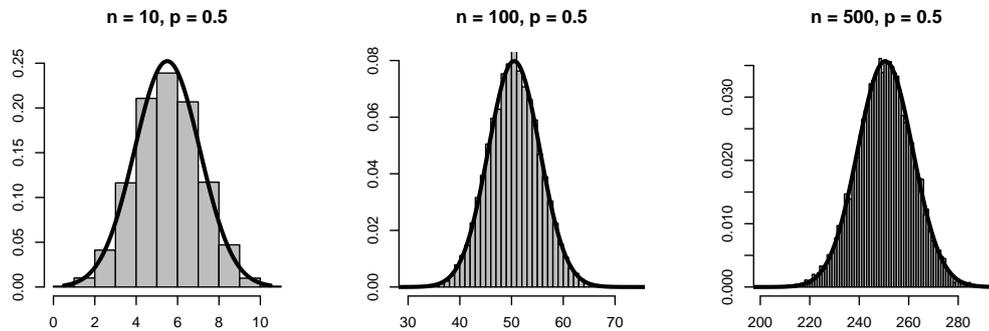


Abbildung 6.1.: Approximation der Binomialverteilung durch die Normalverteilung

Die zuvor in Satz 6.1.1 verwendete binomialverteilte Zufallsgröße S_n ist als Summe von n unabhängigen und identisch alternativverteilten Zufallsgrößen X_i auffassbar

$$S_n = \sum_{i=1}^n X_i$$

und asymptotisch normalverteilt mit den Parametern $\mu = E(S_n) = E(\sum_{i=1}^n X_i) = n \cdot p$ und $\sigma^2 = V(S_n) = V(\sum_{i=1}^n X_i) = n \cdot p \cdot (1 - p)$ ist.

$$S_n = \sum_{i=1}^n X_i \sim_{appr.} N(E(S_n), V(S_n))$$

Der zentrale Grenzwertsatz stellt eine Verallgemeinerung dieser Aussagen dar, denn er besagt, dass Summen beliebig, aber unabhängig und identisch, verteilter Zufallsvariablen X_i asymptotisch normalverteilt sind. Lässt man daher im *Grenzverteilungssatz von de Moivre* die Einschränkung weg, dass S_n binomialverteilt ist und betrachtet man stattdessen die Summe von beliebig aber unabhängig und identisch verteilten Zufallsgrößen $\sum_{i=1}^n X_i$, erhält man folgenden Satz:

Satz 6.2.1 (Zentraler Grenzwertsatz) Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge unabhängiger, identisch verteilter Zufallsgrößen mit $E(X_i) = \mu$ und $V(X_i) = \sigma^2$, so gilt:

$$\lim_{n \rightarrow \infty} P \left(a \leq X^{(n)*} = \frac{\sum_{i=1}^n X_i - n \cdot \mu}{\sqrt{n} \cdot \sigma} \leq b \right) = \int_a^b \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt$$

In Worten: Normierte Summen von unabhängigen Zufallsgrößen sind asymptotisch standardnormalverteilt.

$$X^{(n)*} = \frac{\sum_{i=1}^n X_i - n \cdot \mu}{\sqrt{n} \cdot \sigma} \sim_{approx.} N(0, 1)$$

Macht man die Standardisierung rückgängig, erhält man in Analogie zur Schulbuchvariante des Grenzwertsatzes von de Moivre (Satz 6.1.2)

$$X^{(n)} = \sum_{i=1}^n X_i \sim_{approx.} N(n\mu, n\sigma^2) \quad (6.1)$$

Da man den Ausdruck

$$X^{(n)*} = \frac{\sum_{i=1}^n X_i - n \cdot \mu}{\sqrt{n} \cdot \sigma}$$

zu folgendem Ausdruck

$$\bar{X}^{(n)*} = \frac{\frac{1}{n} \sum_{i=1}^n X_i - \mu}{\frac{\sigma}{\sqrt{n}}}$$

umformen kann, ist auch das arithmetische Mittel $\bar{X}^{(n)}$ einer Folge $(X_n)_{n \in \mathbb{N}}$ näherungsweise normalverteilt.

$$\bar{X}^{(n)} = \frac{1}{n} \sum_{i=1}^n X_i \sim_{approx.} N\left(\mu, \frac{\sigma^2}{n}\right) \quad (6.2)$$

Anstelle eines formalen Beweises wollen wir uns diesen Sachverhalt nun mittels Simulation veranschaulichen:

Beispiel 1 - Stetige Gleichverteilung

Betrachten wir nun exemplarisch eine Folge $(X_n)_{n \in \mathbb{N}}$ stetig in $[0,1]$ gleichverteilter Zufallsvariablen X_i .

$$X_i \sim S_{0,1} \text{ mit } \mu = E(X_i) = \frac{1}{2} \text{ und } \sigma^2 = V(X_i) = \frac{1}{12}$$

Das arithmetische Mittel $\bar{X}^{(n)}$ der Summe einer Folge $X^{(n)} = \sum_{i=1}^n X_i$ ist nach (6.2) asymptotisch normalverteilt

$$\bar{X}^{(n)} \sim_{approx.} N\left(\frac{1}{2}, \frac{1}{12n}\right)$$

Erzeugen wir nun unterschiedlich lange Folgen von auf $[0,1]$ stetig gleichverteilten Zufallszahlen $(X_n)_{n \in \mathbb{N}}$, ermitteln die arithmetischen Mittel $\bar{X}^{(n)}$ der Folgen und stellen die Häufigkeiten dieser mithilfe eines Histogramms dar. (vgl. Abbildung 6.2, S. 63)

Wir beginnen mit 1000 Folgen (X_n) der Länge 1, also 1000 einzelnen Zufallszahlen.¹ Die Häufigkeit der einzelnen Zahlen stellen wir als Histogramm dar und legen die Näherung durch die Normalverteilung $N(\frac{1}{2}, \frac{1}{12})$ darüber. Wie in Abbildung 6.2, Bild 1, zu erkennen, ist die Näherung nicht besonders gut.

Betrachtet man allerdings längere Folgen, erkennt man bereits ab der Folgenlänge 2 einen annähernd glockenförmigen Verlauf des Histogramms. Umso größer n ist, desto besser schmiegt sich das Histogramm unter die Normalverteilung.

¹Da wir nur einzelne Werte betrachten, fällt die Berechnung des arithmetischen Mittels weg.

Beispiel 2 - Exponentialverteilung

Allgemein ist die Summe von Zufallsvariablen beliebiger Verteilungen, die eine positive und endliche Varianz besitzen, wie z.B. die Exponentialverteilung, näherungsweise normalverteilt.

Betrachten wir nun eine Folge $(X_n)_{n \in \mathbb{N}}$ exponentialverteilter Zufallsvariablen X_i mit Parameter $\lambda = 1$.

$$X_i \sim Ex_\lambda \text{ mit } \mu = E(X_i) = \frac{1}{\lambda} \text{ und } \sigma^2 = V(X_i) = \frac{1}{\lambda^2}$$

Das arithmetische Mittel $\bar{X}^{(n)}$ der Summe einer Folge $X^{(n)} = \sum_{i=1}^n X_i$ ist nach (6.2) asymptotisch normalverteilt

$$\bar{X}^{(n)} \sim_{approx.} N\left(\frac{1}{\lambda}, \frac{1}{n \cdot \lambda^2}\right)$$

Gehen wir nun wie im vorherigen Beispiel vor, erhalten wir eine graphische Veranschaulichung des *zentralen Grenzwertsatzes* für die Summe exponentialverteilter Zufallsvariablen. (vgl. Abbildung 6.3, S. 64)

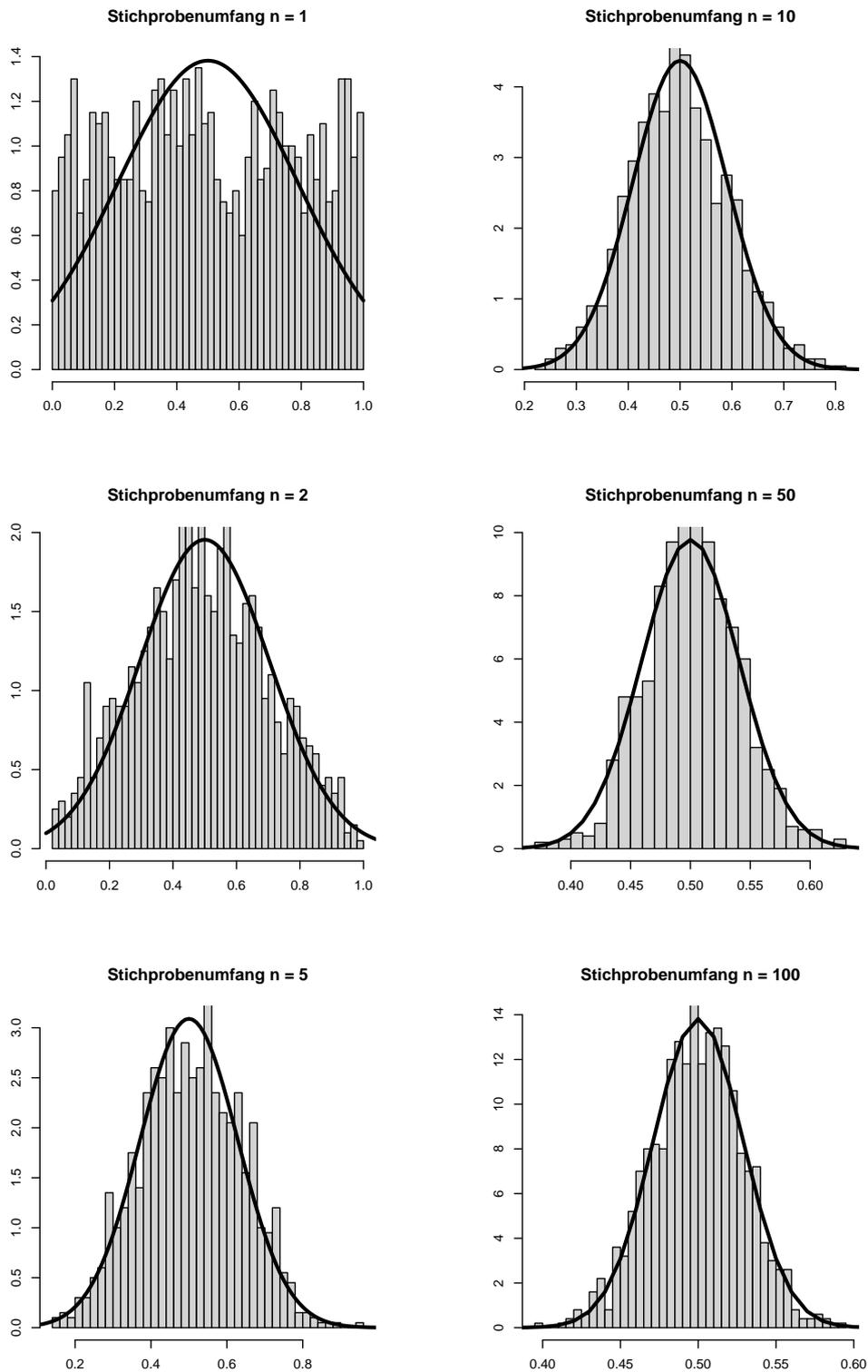


Abbildung 6.2.: Arithmetische Mittel von n stetig in $[0,1]$ gleichverteilten Zahlen

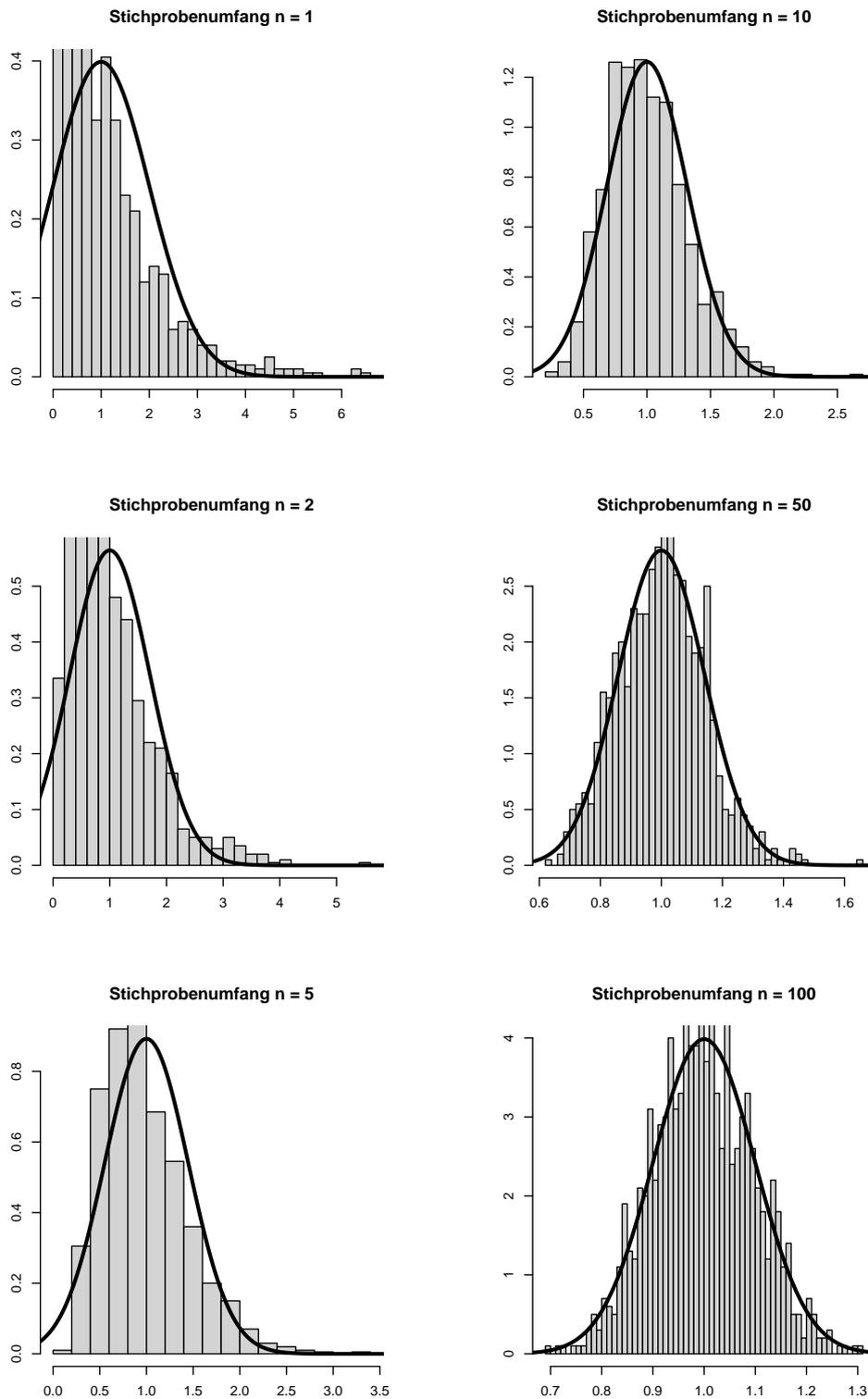


Abbildung 6.3.: Arithmetische Mittel von n exponentialverteilten Zufallszahlen mit Parameter $\lambda = 1$

Kapitel 7.

Subjektive Wahrscheinlichkeiten und Simulation

Bei einmalig stattfindenden Veranstaltung, wie sie im Sport oft vorkommen, ist ein mehrmaliges Durchführen des Experiments nicht möglich und die Wahrscheinlichkeiten werden nach subjektiven Einschätzungen gegeben. Wir wollen im Folgenden auf die Einschätzungen der Buchmacher und Buchmacherinnen vertrauen und versuchen mithilfe dieser die voraussichtliche Teilnahme einer Mannschaft bei einer Sportveranstaltung vorauszusagen.

7.1. Fußball WM-Qualifikation

Aufgabenstellung Wird das Österreichische Fußballnationalteam bei der Weltmeisterschaft 2014 mitspielen? Mit welcher Wahrscheinlichkeit belegt es welchen Platz in der Gruppenphase der Qualifikation zur Weltmeisterschaft?

Vorüberlegungen

Das Österreichische Fußballnationalteam (O) muss sich gegen die Mannschaften aus Deutschland (D), Färöer (F), Irland (I), Kasachstan (K) und Schweden (S) durchsetzen, um bei der Weltmeisterschaft 2014 mitzuspielen. Jedes Team spielt gegen jedes andere zweimal (einmal zuhause und einmal auswärts) und wer am Ende am meisten Punkte hat, qualifiziert sich für das Turnier. Neben dieser Qualifikationsgruppe gibt es noch acht weitere. Unter den besten acht Zweitplatzierten aller neun Gruppen werden weitere vier Startplätze für das Turnier ausgespielt. Wie könnte man nun vorgehen, um die Platzierung der österreichischen Mannschaft vorauszusagen?

Begeben wir uns zuerst in die Welt der Buchmacher und Buchmacherinnen. Diese legen nach subjektivem Ermessen und Befragungen Wettquoten für die einzelnen Spiele fest. Beispielsweise hat der Wettanbieter *tipico* für das Spiel Irland gegen Österreich am 26.03.2013 in Dublin folgende Wettquoten q_i für die möglichen Ereignisse i , 1 (Sieg der Heimmannschaft), 0 (Unentschieden) und 2 (Sieg der Gastmannschaft), festgelegt:

q_1	q_0	q_2
2.2	3.3	3.5

Setzt man auf einen Sieg Irlands, zahlt einem *tipico*, im Fall eines Sieges, das 2,2-fache. Bei einem Sieg von Österreich erhält man für einen richtigen Tipp sogar das 3,5-fache und

bei einem richtigen Tipp auf Unentschieden das 3,3-fache. Doch wie kommen diese Quoten zustande?

Betrachtet man ein Fußballspiel als ein Zufallsexperiment, so kann man den möglichen Ereignissen (0, 1 und 2) bestimmte (subjektive) Wahrscheinlichkeiten p_i ($i = 0, 1, 2$) zuordnen. Der Kehrwert dieser Wahrscheinlichkeiten p_i bildet dann die Quote q_i und umgekehrt

$$q_i = \frac{1}{p_i} \quad \Leftrightarrow \quad p_i = \frac{1}{q_i}$$

Wir können daher aus den Quoten die zugrundeliegenden subjektiven Wahrscheinlichkeiten ermitteln.

p_1	p_0	p_2	$\sum_i p_i$
0.45	0.3	0.29	1.04

In der Regel ist die Summe der ermittelten Häufigkeiten größer als 1. Das ist bewusst von den Wettanbietern so gewählt, damit sie einen Gewinn erzielen: Denn höhere Wahrscheinlichkeiten bedeuten geringere Quoten!

Wir normieren daher die Wahrscheinlichkeiten p_i , indem wir sie durch die Summe $\sum_i p_i$ dividieren.

\hat{p}_1	\hat{p}_0	\hat{p}_2	$\sum_i \hat{p}_i$
0.44	0.29	0.27	1

Die Wahrscheinlichkeiten, die ich auf diese Weise berechnet habe, sind in den Tabellen auf Seite 113 zu finden.¹ Doch was können wir nun mit diesen Wahrscheinlichkeiten anfangen?

Eine analytische Behandlung der anfänglichen Fragestellung werden wir nicht so einfach erhalten. Wir können natürlich anhand der Wahrscheinlichkeiten die erwartete Punktezahl von Österreich ermitteln, erhalten dadurch aber keinen Vergleich zu den anderen Mannschaften. Beispielsweise könnten unabhängig von den Ergebnissen der Spiele der österreichischen Mannschaft die Spiele zwischen Schweden und Irland einen großen Einfluss auf die Platzierung des österreichischen Teams haben. Wir haben es somit mit einem sehr komplexen System zu tun und werden an einer einfachen analytischen Lösung des Problems scheitern, können uns allerdings mithilfe von Simulationen Abhilfe schaffen.

Lösung mithilfe von Simulation

Die experimentelle Vorgangsweise ist mathematisch relativ einfach. Wir simulieren jedes Spiel der Gruppenphase anhand der Wahrscheinlichkeiten² Bildlich kann man sich die Gewinnung eines Ergebnisses so vorstellen, dass man aus einer Urne mit Kugeln, die mit 0,

¹Als zugrundeliegende Quoten habe ich für bereits gespielte Partien jene von <http://www.fussballportal.de/wm-2014/wm-2014-qualifikation/europa/spielplan>, Stand 27.03.2013, verwendet und noch nicht gespielte Partien mithilfe dieser selbst geschätzt.

²Für schnell durchführbare Simulationen sind wir auf spezielle Statistiksoftware, wie z.B. R, angewiesen.

1 oder 2 beschriftet sind und jeweils mit den Häufigkeiten p_0 , p_1 bzw. p_2 vorkommen, einmal zieht.

Führt man diese Simulation für jede Begegnung durch, so erhält man ein ähnliches Ergebnis wie in Tabelle 7.1, wobei die Zeilen den Heimmannschaften und die Spalten den Auswärtsmannschaften entsprechen. Die sich daraus ergebende Punkteverteilung kann der Tabelle 7.2 entnommen werden.

	D	F	I	K	O	S
D		1	1	1	1	0
F	2		0	0	2	2
I	2	1		1	0	2
K	2	1	0		1	2
O	0	0	2	1		2
S	2	1	1	1	2	

Tabelle 7.1.: Ergebnisse bei einmaliger Simulation

	D	F	I	K	O	S
Punkte	29	6	18	11	15	25

Tabelle 7.2.: Punkte der Mannschaften bei einmaliger Simulation

Platzierung Österreichs

Österreich würde in diesem Fall den 4. Platz erreichen. Um aussagekräftige Prognosen abgeben zu können, wird man die Spiele der Gruppenphase öfter simulieren. Die absoluten und relativen Häufigkeiten der Platzierung der Österreichischen Fußballnationalmannschaft bei 1000 maliger Simulation der Gruppenphase sind der Tabelle 7.3 zu entnehmen. Vertraut man auf meine Einschätzungen und die der Buchmacher und Buchmacherinnen, kann man diese relativen Häufigkeiten als Schätzwerte für die Wahrscheinlichkeiten der einzelnen Ränge verwenden.

Für die WM-Teilnahme sind die Platzierungen eins und zwei relevant. Nehmen wir an, dass die Wahrscheinlichkeit, dass sich die österreichische Mannschaft als Zweitplatzierte noch für die Weltmeisterschaft qualifiziert, $1/2$ beträgt, können wir die Gesamtwahrscheinlichkeit einer Qualifikation angeben als

$$p_{\text{Teilnahme}} = p_{1.\text{Rang}} + \frac{1}{2} \cdot p_{2.\text{Rang}} = 0.042 + \frac{1}{2} \cdot 0.176 = 0.13$$

	abs.Hauf.	rel.Hauf
Rang 1	42	0.042
Rang 2	176	0.176
Rang 3	326	0.326
Rang 4	389	0.389
Rang 5	61	0.061
Rang 6	6	0.006

Tabelle 7.3.: Ränge Österreichs bei 1000 Simulationen

Ich schätze daher die Wahrscheinlichkeit einer WM-Teilnahme mit 13% ein.

Platzierung Deutschlands

Für die deutschen Leser und Leserinnen dieser Arbeit, wollen wir nun auch noch die Ränge der deutschen Nationalmannschaft analysieren. Die Simulationsergebnisse sind der

	abs.Hauf.	rel.Hauf
Rang 1	783	0.783
Rang 2	174	0.174
Rang 3	30	0.030
Rang 4	12	0.012
Rang 5	1	0.001
Rang 6	0	0.000

Tabelle 7.4.: Ränge Deutschland bei 1000 Simulationen

Tabelle 7.4 zu entnehmen. Dieselben Überlegungen wie zuvor gehen davon aus, dass das deutsche Nationalteam mit einer Wahrscheinlichkeit von 87% bei der Welmeisterschaft 2014 mitspielt.

Graphische Veranschaulichung der Platzierungen

Die absoluten Häufigkeiten der Ränge aller Mannschaften bei 1000 Simulationen sind in den Abbildungen 7.1 und 7.2 auf den Seiten 69 und 70 als Säulen- und Tortendiagramme dargestellt.

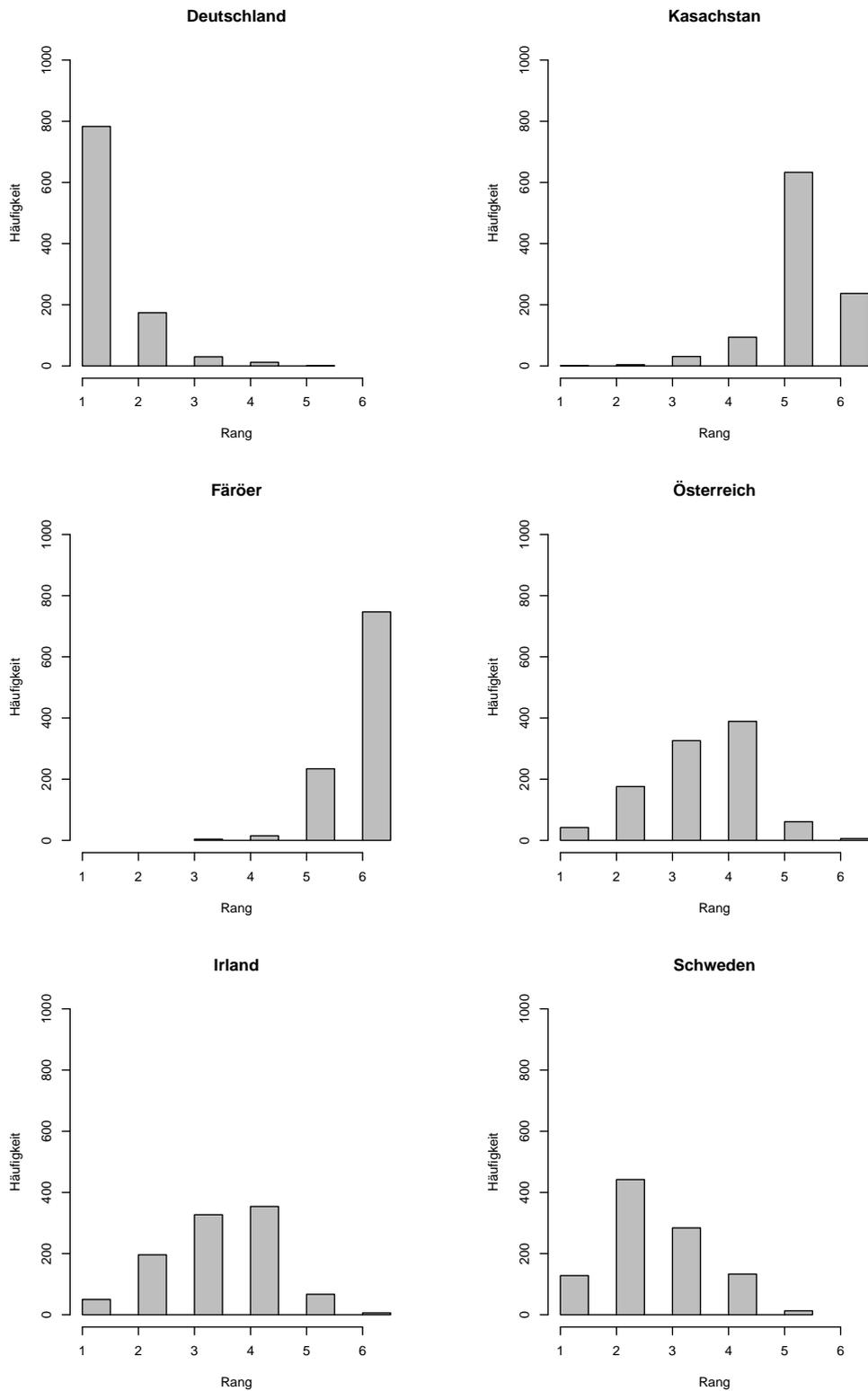


Abbildung 7.1.: Ränge aller Mannschaften bei 1000 Simulationen der Gruppenphase als Säulendiagramme

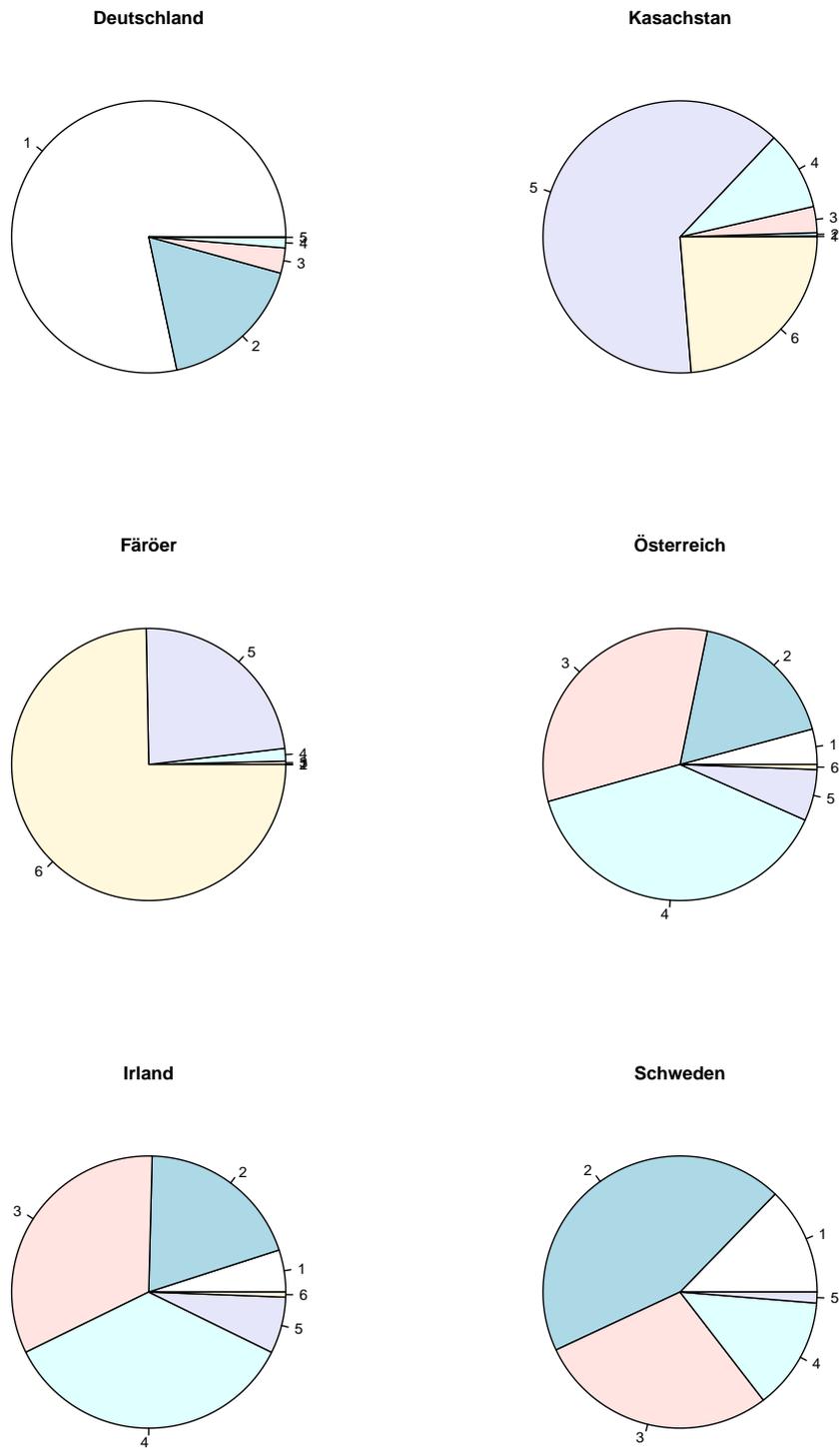


Abbildung 7.2.: Ränge aller Mannschaften bei 1000 Simulationen der Gruppenphase als Tortendiagramme

Kapitel 8.

Wartezeiten und Warteschlangen

8.1. Wartezeiten

8.1.1. Warten auf Augenzahl 6

Die meisten kennen vermutlich das Brettspiel „*Mensch ärgere dich nicht*“, bei welchem die Augenzahl 6 von großer Bedeutung ist. Erst wenn man diese würfelt, kann man seine Figuren ins Spiel bringen. Besonders ärgerlich ist es, wenn man keine Figuren mehr im Spiel hat und vergeblich auf die Augenzahl 6 wartet, um wieder aktiv mitspielen zu können. Obwohl man drei mal würfeln darf, schafft man es mitunter trotzdem nicht, den gewünschten Sechser zu würfeln.

Nun wollen wir uns die Frage stellen, ob es wirklich besonders viel Pech bedarf, um in eine solche Lage zu kommen und mit wie vielen Versuchen zu rechnen ist, bis man die Augenzahl 6 würfelt.

Lösung mithilfe von Simulation

Nähert man sich dieser Fragestellung mittels Simulation, d.h. würfelt man viele Male mit einem Würfel bis man einen Sechser erhält, notiert sich die Wartezeiten und berechnet das arithmetische Mittel dieser, so wird man feststellen, dass man im Durchschnitt in etwa sechs Versuche zu warten hat, bis man die 6 würfelt.

Analytische Lösung

Die analytische Lösung bereitet im Normalfall auch keine großen Schwierigkeiten, wenn man bereits weiß, dass es sich beim mehrmaligen Würfeln um unabhängige aufeinanderfolgende Alternativversuche handelt. Daher kann man die Wahrscheinlichkeit P für die Wartezeit $X = n$ als Produkt von $n - 1$ Misserfolgen mit den Wahrscheinlichkeiten $1 - p = 1 - \frac{1}{6}$ und schlussendlich einem Erfolg mit der Wahrscheinlichkeit $p = \frac{1}{6}$ angeben

$$P(X = n) = (1 - p)^{n-1} \cdot p = \left(1 - \frac{1}{6}\right)^{n-1} \cdot \frac{1}{6}$$

Die daraus resultierende Wahrscheinlichkeitsverteilung wird *Geometrische Verteilung* genannt und besitzt den Erwartungswert

$$E(X) = \frac{1}{p} = 6$$

Es ist daher zu erwarten, dass wir in etwa 6 Mal würfeln müssen, bis wir die Augenzahl 6 erhalten und ein Scheitern bei nur 3 Versuchen ist daher nichts Untypisches.

8.1.2. Sammlerproblem

In letzter Zeit werden Sammel-Aufkleber in vielen der bekannten Supermarktketten zu einem Einkauf geschenkt. Man erhält pro 10 bezahlten Euro ein Briefchen, das fünf der möglichen 200 Aufkleber beinhaltet. Nun stellen wir uns die Frage, wie lange man sammeln muss, d.h. wie viele Einkäufe zu 10 Euro man tätigen muss, bis man jeden Aufkleber besitzt. Wir werden sehen, dass eine analytische Lösung bereits um einiges anspruchsvoller als jene im vorherigen Beispiel ist.

Lösung mithilfe von Simulation

Betrachten wir vorerst das Kaufen einzelner Aufkleber und nehmen wir an, dass alle Aufkleber gleich häufig sind und daher jeder mit gleicher Wahrscheinlichkeit gezogen wird.

Dann können wir uns die Aufkleber als von 1 bis 200 durchnummerierte Zettel in einer Urne, aus welcher wir ziehen, vorstellen. Wir wollen wissen, wie lange es dauert bis wir alle Zettel einmal gezogen haben, wobei wir nach jeder Ziehung den Zettel wieder in die Urne zurücklegen.¹ Wir bezeichnen die Anzahl der Ziehungen mit der Zufallsgröße X .

Die Ergebnisse solcher Simulationen für Aufkleberanzahlen n von 6, 20, 50, 100 und 200 sind in folgender Tabelle zu finden.

	n=6	n=20	n=50	n=100	n=200
X	15	72	225	523	1170
X/n	2.43	3.59	4.49	5.23	5.85

Sammeln wir die Aufkleber, kommen wir nicht daran vorbei, mindestens n Aufkleber zu kaufen. In der Regel werden es aber mehr sein. Wir sind daher daran interessiert, welches Vielfache der Aufkleberanzahl n wir in etwa kaufen müssen?

In der obigen Tabelle sind auch diese Vielfachen $\frac{X}{n}$ dargestellt. Man kann erkennen, dass sie bei wachsendem n größer werden. Können wir eine Formel für diese Vielfache vermuten?

In Abbildung 8.1 wurden die Werte X/n in Abhängigkeit von n in ein Koordinatensystem aufgetragen. Wie schon im *Lotto-Kugel-Paradoxon* in Abschnitt 4.3 auf Seite 49 wird die Punktwolke am besten durch den natürlichen Logarithmus $\ln(n)$, in diesem Fall durch $f(x) = \ln(x) + 0.5$, approximiert.

¹Entspricht die Anzahl der Aufkleber der Flächenzahl eines Würfels (6, 8, 12 oder 20) kann man diese Problemstellung auch durch Würfeln simulieren.

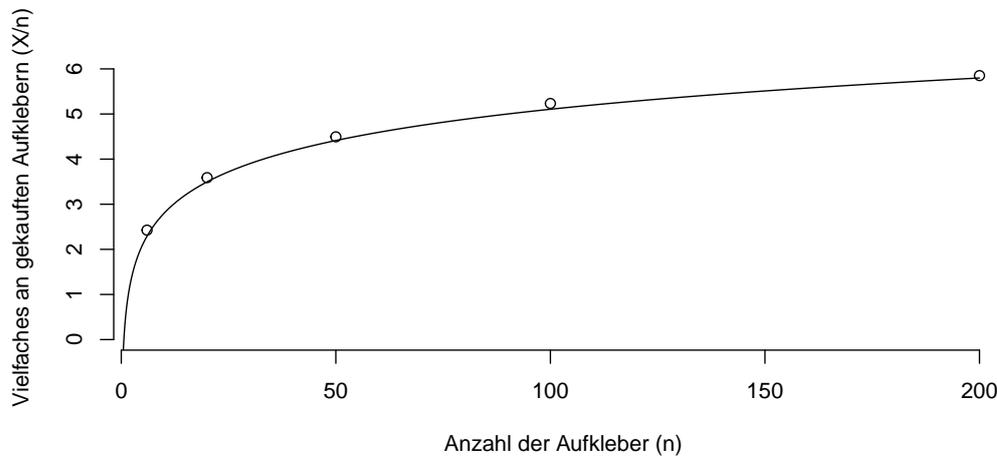


Abbildung 8.1.: Vielfaches an Sammelbildern (X/n) in Abhängigkeit von der Gesamtmenge n an Sammelbildern inklusive Näherung durch die Funktion $f(n) = \ln(n) + 0.5$.

Aufgrund unserer Simulationen vermuten wir daher, dass wir bei n verschiedenen Aufklebern in etwa $\frac{X}{n} \approx \ln(n)$ so viele Aufkleber kaufen müssen, bis wir alle besitzen. Wir müssen daher in etwa

$$E(X) = n \cdot \ln(n)$$

Käufe abwarten, bis unsere Sammlung vollständig ist. Da man bei jedem Einkauf ein Briefchen mit 5 Bildern erhält, sind damit Z 10-Euro-Einkäufe zu erwarten

$$E(Z) = \frac{n \cdot \ln(n)}{5}$$

Das würde bei $n = 200$ bedeuten, dass man in etwa 212 10-Euro-Einkäufe tätigen muss, um das Sammelalbum zu vervollständigen. Das entspricht einer Ausgabe von mehr als 2120 Euro.

Analytische Lösung

Versuchen wir nun die aufgestellte Formel zu bestätigen. Wenn wir der Reihe nach Aufkleber kaufen, hängt die Wahrscheinlichkeit, einen neuen zu erhalten, nur davon ab, wie viele Aufkleber wir aktuell besitzen und nicht davon wie viele wir zu einem anderen, vorausgegangen Zeitpunkt besessen haben. Wir nennen diese Eigenschaft *Gedächtnislosigkeit* und können das Problem daher mittels einer *Markov*-Kette beschreiben. Betrachten wir zu Beginn den Fall mit 6 Aufklebern, der in Abbildung 8.2 dargestellt ist. Die Knoten $i.Z$ kennzeichnen den Zustand, in dem wir i verschiedene Aufkleber besitzen und die Werte über den Pfeilen geben die Wahrscheinlichkeit für den Übergang zwischen zwei Zuständen beim Kauf eines neuen Aufklebers an.

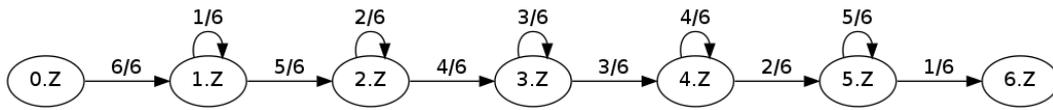


Abbildung 8.2.: Markov-Kette zur Sammlung von 6 Aufklebern

Besitze ich noch keinen Aufkleber (0.Z) ist es so gut wie sicher ($p=6/6$), dass ich in den Zustand (1.Z) gelange, d.h. einen neuen Aufkleber erhalte. Umso mehr verschiedene Sticker ich besitze, desto unwahrscheinlicher wird es, einen neuen Sticker zu erhalten.

Die Wartezeit X_i , um vom Zustand $i - 1$ in den Zustand i zu gelangen, ist wiederum geometrisch verteilt:

$$X_i \sim G_{p_i} \quad \text{mit} \quad p_i = \frac{6 - i + 1}{6} \quad i = 0, 1, \dots, 6$$

Da die einzelnen Wartezeiten X_i in etwa

$$E(X_i) = \frac{1}{p_i} = \frac{6}{6 - i + 1}$$

betragen, ist insgesamt mit einer Wartezeit $X = \sum_{i=1}^6 X_i$ von

$$E(X) = E\left(\sum_{i=1}^6 X_i\right) = \sum_{i=1}^6 E(X_i) = \sum_{i=1}^6 \frac{6}{6 - i + 1} = 6 \cdot \sum_{i=1}^6 \frac{1}{6 - i + 1} = 6 \cdot \sum_{i=1}^6 \frac{1}{i}$$

zu rechnen. Wie bereits beim *Lotto-Kugel-Paradoxon* können wir die im Ausdruck enthaltene Summe durch den Logarithmus approximieren:

$$E(X) = 6 \cdot \sum_{i=1}^6 \frac{1}{i} \approx 6 \cdot \int_1^6 \frac{1}{x} dx = 6 \cdot \ln(6)$$

Würden wir die selben Überlegungen für die Aufkleberanzahl n durchführen, erhielten wir dasselbe Ergebnis. Daher erhalten wir durch Ersetzen von 6 mit n in der Formel den Erwartungswert der Wartezeit $X^{(n)}$ für den allgemeinen Fall:

$$E(X) = n \cdot \sum_{i=1}^n \frac{1}{i} \approx n \cdot \ln(n)$$

8.2. Warteschlangen

8.2.1. Handy-Reperatur

Bevor wir uns genauer mit Warteschlangen beschäftigen wollen, betrachten wir ein einführendes Beispiel, um ein erstes Gefühl für zufällige Ereignisse zu erhalten.

Beispiel Das Geschäft eines großen Mobilfunkbetreibers will für seine Kunden folgende Dienstleistung zur Verfügung stellen:

Wenn man im Laufe des Vormittags sein kaputtes Handy im Geschäft abgibt, kann man es am Abend in funktionstüchtigem Zustand wieder abholen.

Die Geschäftsleitung ist daher auf der Suche nach einem Dienstleister, der die Telefone abholt, repariert und am selben Tag wieder zurückbringt.

Wir wollen einen solchen Dienst anbieten. Es wird uns mitgeteilt, dass mit ungefähr 50 Reparaturen pro Woche zu rechnen ist, die sich zufällig auf die 6 Werktage aufteilen.

Wir errechnen aus den 50 Reparaturen pro Woche durchschnittlich $50/6 \approx 8$ Reparaturen pro Arbeitstag und nehmen daher an, dass unsere zwei Angestellten, die jeweils 5 Reparaturen pro Nachmittag durchführen können, ausreichend sind. Liegen wir in unserer Annahme richtig?

Lösung mithilfe von Simulation

Sehen wir uns nun an, an wie vielen Tagen wir die Mobiltelefone nicht am selben Tag reparieren können, d.h. mehr als 10 Mobiltelefone bei uns eintreffen.

Zur Simulation würfeln wir 50 Mal und fassen die Augenzahlen als Wochentage auf. Die Häufigkeiten der gewürfelten Augenzahlen, d.h. die Anzahl an Mobiltelefonen, die am jeweiligen Tag zur Reparatur gebracht werden, ergeben sich in unserem Fall wie folgt:

	Mo	Di	Mi	Do	Fr	Sa
Anzahl an kaputten Mobiltelefonen	5	7	11	8	12	7

Wir erkennen, dass unsere Anzahl an Angestellten nicht ausreichen wird, da wir an zwei Tagen die an uns gestellte Aufgabe nicht bewältigen können und das, obwohl pro Tag mehr Telefone repariert werden können als im Schnitt erwartet werden. Wir tun uns anscheinend schwer, zufällige Ereignisse einzuschätzen und stellen fest, dass Zufälligkeit eben nicht annähernde Gleichverteilung bedeutet.

8.2.2. Definition einer Warteschlange

Im Alltag begegnen wir oft *Warteschlangen*: Sei es, dass wir einen Fahrschein an einem Schalter kaufen, uns in der Kantine anstellen oder in der Praxis eines Arztes im Warteraum Platz nehmen. Warteschlangen müssen aber nicht aus Menschen bestehen: Flugzeuge warten auf eine Landeerlaubnis, Akten in einem Amt auf ihre Bearbeitung und Programme am Computer auf ihre Ausführung. Versuchen wir nun **Warteschlangensysteme**² allgemein zu beschreiben.

²auch: *Bedienungssysteme*

Ein Bedienungssystem besteht, wie in folgender Abbildung dargestellt, aus einer **Förderquelle**, einem **Warterraum** und einem **Bedienungsknoten**.



Die Förderquelle beschreibt das Ankommen von neuen Forderungen, den sogenannten Ankunftsprozess. Zur mathematischen Modellierung des Ankunftsprozesses werden meist die **Zwischenankunftszeiten**, d.h die Zeiten zwischen zwei Forderungen, herangezogen.

Kommen neue Forderungen hinzu, müssen diese, sofern noch andere Forderungen vorhanden sind, auf ihre Bedienung warten. Einerseits unterscheidet man zwischen Warterräumen mit *endlicher* und *unendlicher Kapazität* und andererseits zwischen verschiedenen **Wartedisziplinen**. Beispiele für solche Wartedisziplinen sind folgende:

FIFO	<i>(first in - first out)</i>	Die Forderungen werden in der Reihenfolge ihres Eintreffens bedient. (Beispiel: Anstellen an einer Kassa.)
LIFO	<i>(last in - first out)</i>	Die zuletzt eingetroffene Forderung wird zuerst bedient. (Beispiel: Abarbeiten eintreffender Akten in einem Amt.)
PRI	<i>(priority)</i>	Den Forderungen werden bestimmte Prioritäten zugewiesen und sie entsprechend dieser abgearbeitet. (Beispiel: Landeerlaubnis für Flugzeuge richtet sich nach der Tankfüllung.)

Bei den folgenden Beispielen kommt immer das *FIFO*-Prinzip zur Anwendung.

Zu guter Letzt werden die Forderungen auch abgearbeitet. Für die Modellierung ist dabei vor allem die **Bedienzeit**, d.h. die Dauer der Abarbeitung, von großer Bedeutung.

8.2.3. Warteschlange an einer Kassa mittels Würfeln

Zwischenankunftszeiten deutlich größer als die Bedienzeit

Stellen wir uns nun eine Kassa in einem Lebensmittelgeschäft vor. Die Wartedisziplin folgt dem *FIFO*-Prinzip und wir nehmen an, dass die *Zwischenankunftszeiten* und *Bedienzeiten* in etwa 3,5 sowie 1,75 min. betragen. Die erwartete *Zwischenankunftszeit* ist *deutlich größer* als die erwartete *Bedienzeit*. Wie können wir nun zufällige Zeiten simulieren, die

diese Erwartungswerte besitzen?

Am einfachsten können wir das mithilfe eines Würfels bewerkstelligen. Die Augenzahlen stellen die Zwischenankunftszeiten X und die Hälfte der Augenzahlen die Bedienzeiten Y dar. Wir nehmen also an, dass die Zeiten diskret gleichverteilt auf $M_X = \{1, 2, 3, 4, 5, 6\}$ bzw. $M_Y = \{\frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, 3\}$ sind.

Wir werden daher in der Regel doppelt so schnell bedient, wie jemand ankommt. Intuitiv würden wir wiederum annehmen, dass gar keine Warteschlangen entstehen werden. Liegen wir diesmal richtig?

Zuerst legen wir einen Zeitraum fest, den wir betrachten wollen. Wir wählen eine Zeitspanne von 30 min. Nun erzeugen wir so lange durch Würfeln zufällige Zwischenankunftszeiten Z , bis ihre Summe größer als 30 ist. Die Teilsummen dieser Zeiten bilden die absoluten Ankunftszeiten A . Für jede ankommende Forderung erzeugen wir durch erneutes Würfeln zufällige Bedienzeiten B . Als Ergebnis erhalten wir folgende Tabelle, in der die Beschriftungszeile die i -te im System ankommende Person bezeichnet.

	1	2	3	4	5	6	7	8	9
Z	0	6	2	3	4	6	2	6	6
A	0	6	8	11	15	21	23	29	35
B	2.0	2.0	0.5	1.0	1.0	2.5	1.5	2.5	1.5

Der Tabelle ist zu entnehmen, dass zum Zeitpunkt $A = 0$ die erste Person ankommt und innerhalb von $B = 2$ Minuten bedient wird. Die nächste Person erreicht zum Zeitpunkt $A = 6$ das System und wird wiederum nach $B = 2$ abgefertigt.

Diese Werte kann man nun in ein Koordinatensystem eintragen und erhält den Verlauf der Warteschlangenlänge (y -Wert) über die Zeit (x -Wert): Trägt man die Werte von A an der x -Achse auf, wächst an diesen Stellen die Warteschlange um den Wert 1 und zählt man zu den Stellen von A die Werte B hinzu, wird an die Warteschlange an diesen Stellen um den Wert 1 kleiner.

Eine solche grafische Darstellung für eine Simulation der Warteschlangenlänge über den Zeitraum einer Stunde ist in Abbildung 8.3 dargestellt und die ersten 15 Funktionswerte der entstandenen Treppenfunktion sind der folgenden Tabelle zu entnehmen.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Zeit	0	0	2	2	5	6	8	9	11	14	14	18	19	20	22
Personen	0	1	0	1	0	1	2	1	0	1	0	1	0	1	0

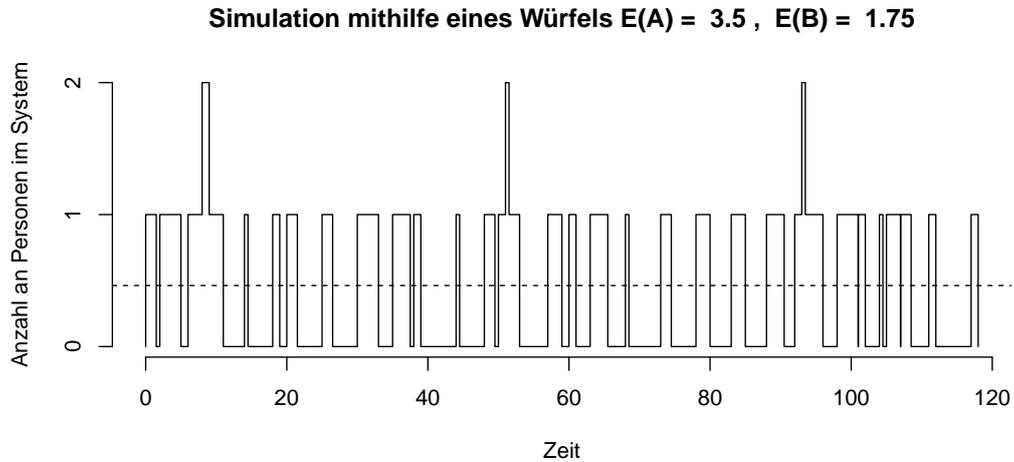


Abbildung 8.3.: Anzahl an Personen im System (y) in Abhängigkeit von der Zeit (x)

Wir erkennen wiederum, dass drei Mal eine Warteschlange entsteht. Können wir mit dem Verhalten des Warteschlangensystems zufrieden sein?

Interessierende Größen eines Warteschlangensystems

Als Betreiber oder Betreiberin eines Bediensystems könnten die *Länge der Warteschlange* und die *Anzahl an Personen im System*, die *durchschnittliche Wartezeit* bzw. *Gesamtaufenthaltsdauer* und die *Auslastung der Angestellten* von Interesse sein.

Warteschlangenlänge und Personen im System Umso länger die Warteschlange ist, desto eher werden wir dazu verleitet sein, uns nicht anzustellen.

Die maximale Anzahl an Personen im System ist aus der Abbildung 8.3 leicht abzulesen: Es befinden sich maximal drei Personen im System, das einer Warteschlangenlänge von 2 Personen entspricht. Um die durchschnittliche Länge zu erhalten müssen wir das Mittel aus allen möglichen Längen $l = 0, \dots, n$ (wobei n die maximal auftretende Länge bezeichnet), die mit der Dauer ihres Auftretens gewichtet sind, berechnen. Die durchschnittliche Länge L_q beträgt in diesem Fall

$$L_q \approx 0.02$$

Die durchschnittliche Anzahl von Personen im System L ist in Abbildung 8.3 strichliert aufgetragen und beträgt in etwa

$$L \approx 0.46$$

Ohne Meinungsumfragen durchführen zu müssen sind diese Werte aus der Sicht eines Kunden oder einer Kundin sehr gut.

Durchschnittliche Wartezeit und Gesamtaufenthaltsdauer Eine weitere aus unserer Sicht als Käufer oder Käuferin bedeutende Kenngröße ist die Zeit, die wir uns durch-

schnittlich im System befinden, d.h. die Zeit zwischen der Ankunft und dem Verlassen des Bediensystems. Wir erhalten folgende durchschnittliche Gesamtaufenthaltsdauer

$$W \approx 1.68$$

Die mittlere Wartedauer in der Warteschlange W_q beträgt

$$W_q \approx 0.06$$

Auslastung eines Angestellten Die Unternehmensführung wird daran interessiert sein, ihre Angestellten möglichst auszulasten. Wie groß ist die Zeit, in denen ein Bediener bzw. eine Bedienerin beschäftigungslos ist? Wir berechnen den Anteil der Zeit, in der die Angestellten unbeschäftigt sind, an der Gesamtzeit und erhalten

$$A \approx 0.54$$

Als Angestellter oder Angestellte ist man daher (in unserem Fall) mehr als die Hälfte der Arbeitszeit beschäftigungslos.

Bedien- und Zwischenankunftszeiten kommen einander näher

Sehen wir uns nun an, was passiert, wenn die erwarteten *Zwischenankunftszeiten* A und die *Bedienzeiten* B einander *näher kommen*. Wir betrachten nun drei Szenarien von verschiedenen Zeiten A und B , welche wiederum als diskret gleichverteilt angenommen und mithilfe eines Würfels mit den Augensummen 1 bis 6 erzeugt werden.

	A	B	$E(A)$	$E(B)$
Szenario 1	Augensumme	Augensumme/2	3.5	1.75
Szenario 2	Augensumme	$(3 \cdot \text{Augensumme})/4$	3.5	2.625
Szenario 3	Augensumme	$(7 \cdot \text{Augensumme})/8$	3.5	3.0625

Die interessierenden Größen der Simulationen sind in folgender Tabelle aufgelistet

Größe	Symbol	Szenario 1	Szenario 2	Szenario 3
<i>mittlere Länge der Warteschlange</i>	L_q	0.05 P.	0.17 P.	0.67 P.
<i>mittlere Anzahl an Personen im System</i>	L	0.54 P.	0.88 P.	1.52 P.
<i>mittlere Wartezeit</i>	W_q	0.18 min	0.61 min	2.33 min
<i>mittlere Aufenthaltsdauer</i>	W	1.87 min	3.14 min	5.31 min
<i>Anteil der unbeschäftigten Zeit</i>	A	0.51	0.29	0.15

Wie es nicht anders zu erwarten war und auch in der Abbildung 8.4 auf Seite 87 am Ende des Kapitels zu erkennen ist, werden die Wartezeiten der Kunden mit kürzeren Ankunftszeiten immer länger und die Angestellten immer stärker ausgelastet.

Die Frage, die wir uns allerdings stellen müssen, ist, ob unser verwendetes Modell mit den diskret gleichverteilten Zwischenankunftszeiten wirklich der Realität entspricht. Das Personen immer zu ganzen Minuten ankommen und Angestellte immer halbe Minuten brauchen klingt nicht sehr plausibel. Wir werden daher im folgenden Abschnitt ein besseres Modell verwenden.

8.2.4. Warteschlange an einer Kassa als M|M|1-System

Im vorherigen Beispiel haben wir das Bediensystem durch die drei Größen *Zwischenankunftszeit*, *Bedienzeit* und *Warteschlangendisziplin* beschrieben. Meist verwendet man zur Charakterisierung von Warteschlangensystemen die fünf Größen A|B|X|Y|Z:

A	<i>Verteilung der Zwischenankunftszeiten</i>	z.B. Exponentialverteilung (M)
B	<i>Verteilung der Bedienzeiten</i>	z.B. Exponentialverteilung (M)
X	<i>Anzahl der parallelen Schalter</i>	1, 2, ..., ∞
Y	<i>Beschränkung der Systemkapazität</i>	1, 2, ..., ∞
Z	<i>Warteschlangendisziplin</i>	FIFO, LIFO, PRI

Das Symbol M , mit welchem die *Exponentialverteilung* bezeichnet wird, steht für die **Markoveigenschaft**³.

Bezogen auf unser Warteschlangensystem bedeutet diese Eigenschaft, dass sich das Bediensystem nicht merkt, wie lange es bereits auf die Ankunft einer neuen Person wartet. Die Wahrscheinlichkeit, dass, nachdem es bereits eine Zeitspanne t_0 gewartet hat, in der folgenden Zeitspanne der Länge t eine Person ankommt, ist gleich groß wie jene, von Anfang an t lang zu warten. Mathematisch formuliert

$$P(T \leq t_0 + t | T \geq t_0) = P(T \leq t)$$

Durch diese Eigenschaft wird ein maximaler Grad an Zufälligkeit erreicht. Die einzige stetige Verteilung, die diese Eigenschaft erfüllt, ist die (negative) *Exponentialverteilung*

$$\begin{aligned} P(T \leq t_0 + t | T \geq t_0) &= \\ &= \frac{P(T \leq t_0 + t, T \geq t_0)}{P(T \geq t_0)} = \frac{\int_{t_0}^{t_0+t} \lambda e^{-\lambda t} dt}{\int_{t_0}^{\infty} \lambda e^{-\lambda t} dt} = \frac{e^{-\lambda t_0} - e^{-\lambda(t_0+t)}}{e^{-\lambda t_0}} = 1 - e^{-\lambda t} = \end{aligned}$$

³auch: *Gedächtnislosigkeit*

$$= P(T \leq t)$$

Unter Annahme, dass die Zwischenankunftszeiten unabhängig exponentialverteilt mit Parameter λ sind, ist die Anzahl der Ankünfte N_t in einem Zeitintervall der Länge t poissonverteilt mit Parameter λt

$$P(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

und die Familie $(N_t, t \geq 0)$ nennt man einen **Poissonprozess**.

Wir betrachten nun das spezielle Warteschlangensystem $M|M|1|\infty|FIFO$ für die Warteschlange im Supermarkt. Wählt man die Systemkapazität, d.h. die Warteraumgröße, unbeschränkt ($Y = \infty$) und die Warteschlangendisziplin als FIFO, wie wir es tun, spricht man meist kurz von einem $M|M|1$ -System.

Lösung mithilfe von Simulation

Wir wählen die Zwischenankunftszeiten als exponentialverteilt mit dem Parameter $\lambda = \frac{20}{h}$, d.h. es ist mit 20 Ankünften pro Stunde⁴ zu rechnen. Die Bedienzeiten nehmen wir auch als exponentialverteilt mit dem Parameter $\mu = \frac{40}{h}$ an. Nun benötigen wir nur mehr exponentialverteilte Zufallszahlen, um das Beispiel für unterschiedliche Zeiträume $[0, t_{\text{Ende}}]$ (in Minuten) simulieren zu können.

Erzeugung von exponentialverteilten Zufallszahlen⁵

Wie bereits in Abschnitt 1.8 auf Seite 20 beschrieben, kann man mithilfe der *Inversionsmethode* Zufallszahlen beliebiger Verteilungen simulieren. Wenn X eine auf $[0,1]$ stetig gleichverteilte Zufallsvariable ist, erhält man mithilfe folgender Transformation eine exponentialverteilte Zufallsvariable Z :

$$Z = F^{-1}(X) = -\frac{1}{\lambda} \cdot \ln(1 - X)$$

Wir erzeugen nun eine Reihe von Zufallszahlen aus dem Intervall $[0,1]$ und wenden auf diese die Funktion F^{-1} an, um exponentialverteilte Zufallszahlen zu erhalten:

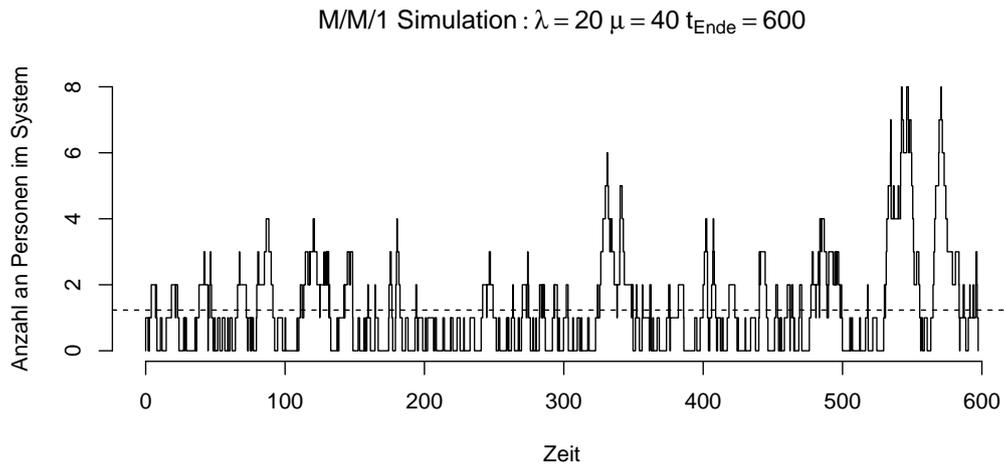
	Zahl 1	Zahl 2	Zahl 3	Zahl 4	Zahl 5	Zahl 6	Zahl 7
X	0.89670	0.26551	0.37212	0.57285	0.90821	0.20168	0.89839
Z	0.11350	0.01543	0.02327	0.04253	0.11941	0.01126	0.11433

⁴Das entspricht einer erwarteten Zwischenankunftszeitdauer von 3 min (vgl. vorheriges Beispiel)

⁵Verwendet man das Statistikprogramm R, kann man mithilfe des Befehls `rexp`, ohne Transformationen durchführen zu müssen, exponentialverteilte Zufallszahlen erzeugen.

Feste Öffnungszeiten

Nehmen wir an, dass das Geschäft von 8.00 bis 18.00 Uhr geöffnet hat, interessiert uns der Warteschlangenverlauf während der Öffnungszeiten, wie er in folgender Abbildung dargestellt ist.⁶



Wir erkennen, dass, obwohl μ doppelt so groß wie λ ist, sehr häufig Warteschlangen auftreten und sogar solche der Länge 7 möglich sind. Wie ändert sich der Verlauf der Warteschlange, wenn mehr Leute ankommen?

Variation von λ

Betrachten wir zum Vergleich die Fälle, dass mehr Leute ankommen, d.h. λ größer wird. In Abbildung 8.5 auf Seite 88 sind für $\mu = 40$ und $\lambda = 20, 30, 40, 50$ die Entwicklungen der Warteschlangen dargestellt und in Tabelle 8.1 die berechneten Kenngrößen angegeben.

Wir erkennen, dass im Falle $\lambda < \mu$ bei unseren Simulationen die Warteschlange ein stabiles Verhalten aufweist und immer wieder abgebaut wird. Sobald λ jedoch gleich μ wird, so legen es uns die Ergebnisse der Simulationen nahe, scheint die Warteschlange immer größer zu werden, obwohl wir das vorab nicht vermutet hätten. Wird λ sogar größer als μ , explodiert die Warteschlange regelrecht und wir müssen im Durchschnitt mehr als $1\frac{1}{2}$ Stunden in einer 60 Personen langen Warteschlange warten, um bedient zu werden.

Rund um die Uhr geöffnet

Bis jetzt haben wir immer den Fall einer Warteschlange an einer Kassa eines Geschäfts betrachtet, welches nur einen bestimmten Zeitraum geöffnet hat. Daher ist nur der Verlauf

⁶In der Realität wird es Zeiten geben, in denen mehr Leute ankommen und andere in denen weniger ankommen. Hier wird der Einfachheit halber angenommen, dass zu jeder Tageszeit in etwa gleich viele Leute ankommen.

Größe	Symbol	$\lambda = 20$	$\lambda = 30$	$\lambda = 40$	$\lambda = 50$
mittlere Länge der Warteschlange	L_q	0.63 P.	2.34 P.	13.29 P.	60.96 P.
mittlere Anzahl an Personen im System	L	1.23 P.	3.14 P.	14.24 P.	61.96 P.
mittlere Wartezeit	W_q	1.66 min	4.48 min	21.11 min	92.16 min
mittlere Aufenthaltsdauer	W	3.25 min	6.03 min	22.61 min	93.67 min
Anteil der unbeschäftigten Zeit	A	0.39	0.19	0.05	0.01

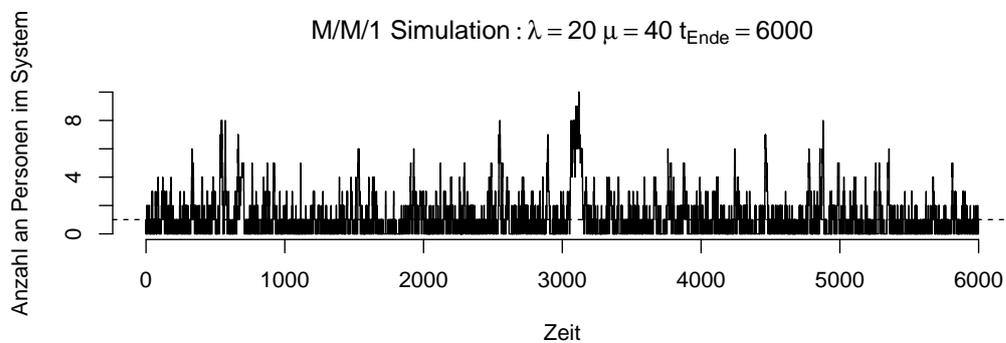
Tabelle 8.1.: Kenngrößen von Warteschlangen für verschiedene λ und μ

der Warteschlangenlänge über diese Zeitspanne von Interesse und nach dem Schließen des Geschäfts, hat der Bediener oder die Bedienerin noch die Möglichkeit, die Warteschlange abzubauen.

Betrachten wir nun den Fall, dass ohne Unterbrechung Forderungen ankommen und bedient werden, wie beispielsweise bei einem Warenlager. In diesem kommen pausenlos Waren an, welche anschließend von den Angestellten an die richtige Stelle weitergeleitet werden.

Wir sind in diesem Fall an dem Verhalten der Warteschlange auf lange Sicht interessiert. Wir wollen wissen, wie sich das System verhält, nachdem eine gewisse Vorlaufzeit vergangen ist und es sich eingeschwungen hat. Dann können wir nämlich langfristige Voraussagen über das Verhalten des Bediensystems treffen.

Aufgrund der folgenden Abbildung können wir vermuten, dass bei $\lambda = 20$ und $\mu = 40$ das **System** auf lange Sicht **stabil** ist, da die Anzahl der Personen im System beschränkt bleibt.

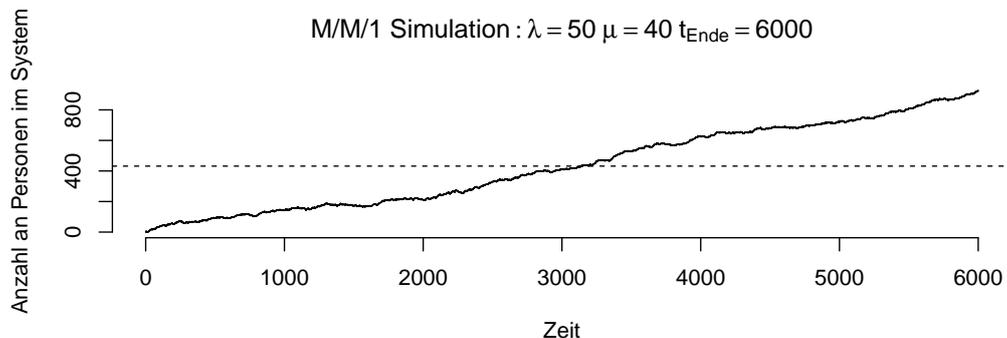


Größe	Symbol	t_{Ende} 600	t_{Ende} 3000	t_{Ende} 6000	lange Sicht
mittlere Länge der Warteschlange	L_q	0.63 P.	0.48 P.	0.51 P.	≈ 0.5 P.
mittlere Anzahl an Personen im System	L	1.23 P.	0.99 P.	1 P.	≈ 1 P.
mittlere Wartezeit	W_q	1.66 min	1.48 min	1.54 min	≈ 1.5 min
mittlere Aufenthaltsdauer	W	3.25 min	3.06 min	3.05 min	≈ 3 min
Anteil der unbeschäftigten Zeit	A	0.39	0.49	0.5	≈ 0.5

Tabelle 8.2.: Kenngrößen der Warteschlangen für $\lambda = 20$ und $\mu = 40$

Betrachten wir die Warteschlange für immer längere Zeitintervalle und berechnen wir für jeden dieser Abschnitte die Kenngrößen (vgl. Tabelle 8.2). Es macht den Anschein, dass sich die Kenngrößen auf lange Sicht in der Nähe von bestimmten Werten befinden. Ist diese Vermutung richtig, ist bei einer beliebigen Ankunftszeit in das eingeschwungene System mit einer Wartezeit von 1,5 Minuten zu rechnen.

Betrachten wir hingegen den Fall $\lambda > \mu$, legt unsere Simulation nahe, dass das System nicht stabil ist und die Warteschlange immer länger wird.



Wenn unsere Annahme, dass die Kenngrößen auf lange Sicht bei $\lambda < \mu$ in der Nähe von bestimmten Werten liegen, richtig ist, dann müsste es doch eine Möglichkeit geben, die Kenngrößen ohne den Umweg der Simulation zu berechnen?

Analytische Lösung

Auf lange Sicht könnte man vermuten, dass sich die Größen in der Nähe von bestimmten Werten befinden. Wir wollen nun allgemeine Aussagen darüber treffen, wie die Warte-

schlange in etwa aussieht, wenn wir sie zu einem beliebigen Zeitpunkt beobachten.

Die Anzahl an Personen im System N_t im Laufe der Zeit t kann als ein stochastischer Prozess $\{N_t : t \geq 0\}$ angesehen werden. Wir bezeichnen mit $p_n(t)$ die Wahrscheinlichkeit, dass sich zum Zeitpunkt t genau n Personen im System befinden. Nun wollen wir wissen, wie wahrscheinlich bestimmte Anzahlen im System sind, nachdem sich das System eingependelt hat und suchen die Wahrscheinlichkeiten, dass sich auf lange Sicht n Personen im System befinden:

$$\lim_{t \rightarrow \infty} p_n(t) = p_n$$

Diese Grenzverteilung wird als **stationäre Verteilung** oder **stabiler Zustand des Wartesystems** (p_n) bezeichnet. Aufgrund der Eigenschaften der Poissonverteilung⁷ können wir die *Kolmogorov-Vorwärts-Differentialgleichungen* für p_n aufstellen:

$$\frac{dp_n}{dt} = -(\lambda + \mu) \cdot p_n(t) + \mu \cdot p_{n+1}(t) + \lambda \cdot p_{n-1}(t) \quad n \geq 0 \quad (8.1)$$

$$\frac{dp_0(t)}{dt} = -\lambda \cdot p_0(t) + \mu \cdot p_1(t) \quad (8.2)$$

Da wir an der stationären Lösung von (p_n) interessiert sind und sich in diesem Fall die Wahrscheinlichkeiten nicht mehr ändern, setzen wir die Ableitungen in (8.1) und (8.2) gleich Null und erhalten folgende Differenzgleichungen:

$$p_{n+1} = \frac{\lambda + \mu}{\mu} \cdot p_n - \frac{\lambda}{\mu} \cdot p_{n-1} \quad n \geq 0 \quad (8.3)$$

$$p_1 = \frac{\lambda}{\mu} \cdot p_0 \quad (8.4)$$

Durch Einsetzen des Startwerts (8.4) in die rekursive Formel (8.3), erhält man folgende Ausdrücke zur expliziten Berechnung der p_n :

$$p_1 = \left(\frac{\lambda}{\mu}\right)^1 \cdot p_0, \quad p_2 = \left(\frac{\lambda}{\mu}\right)^2 \cdot p_0, \quad \dots, \quad p_n = \left(\frac{\lambda}{\mu}\right)^n \cdot p_0 \quad (8.5)$$

Sei

$$\rho = \frac{\lambda}{\mu}$$

der *Benutzungsfaktor*. Dann erhalten wir, da die Summe aller Wahrscheinlichkeiten gleich dem Wert 1 sein muss:

$$1 = \sum_{n=0}^{\infty} p_n = \sum_{n=0}^{\infty} \rho^n p_0 \quad \Leftrightarrow \quad p_0 = \frac{1}{\sum_{n=0}^{\infty} \rho^n}$$

Für den Fall dass $\rho = \frac{\lambda}{\mu} < 1$, d.h. $\lambda < \mu$, können wir der geometrischen Reihe im Nenner einen Wert zuschreiben, da sie konvergiert. Andernfalls, $\frac{\lambda}{\mu} \geq 1$, gibt es keine

⁷Zur Nachlese der theoretischen Hintergründe sei auf Heller et. al. [10] verwiesen.

stabile Verteilung, wie wir es bereits zuvor anhand der Simulationsergebnisse vermutet haben. Die Warteschlange wird immer länger. Im ersten Fall gilt daher

$$p_0 = \frac{1}{\frac{1}{1-\rho}} = 1 - \rho$$

und somit insgesamt

$$\boxed{p_n = \rho^n(1 - \rho)} \quad \text{wenn } \lambda < \mu \quad (8.6)$$

Mithilfe dieser Verteilung kann man Formeln für die interessierenden Größen einer Warteschlange herleiten. An dieser Stelle wird auf die Herleitung mit einem Verweis auf Heller et. al. [10] verzichtet. Die Formel für die Auslastung A als Anteil der beschäftigungslosen Zeit, welche in der angegebenen Literatur nicht behandelt wird, ist gleich der Wahrscheinlichkeit, dass sich niemand im System befindet

$$A = p_0 = 1 - \rho$$

Die restlichen Formeln sind in der folgenden Tabelle aufgelistet

Größe	Formel
<i>mittlere Länge der Warteschlange</i>	$L_q = \frac{\lambda^2}{\mu(\mu-\lambda)}$
<i>mittlere Anzahl an Personen im System</i>	$L = \frac{\lambda}{\mu-\lambda}$
<i>mittlere Wartezeit</i>	$W_q = \frac{\lambda}{\mu(\mu-\lambda)}$
<i>mittlere Aufenthaltsdauer</i>	$W = \frac{1}{\mu-\lambda}$
<i>Anteil der unbeschäftigten Zeit</i>	$A = 1 - \frac{\lambda}{\mu}$

Setzt man die Werte $\lambda = 20$ und $\mu = 40$ aus dem vorherigen Beispiel ein, erhält man genau die damals vermuteten Werte.

	L_q	L	W_q	W	A
$\lambda = 20, \mu = 40$	1	1	1.5	3	0.5

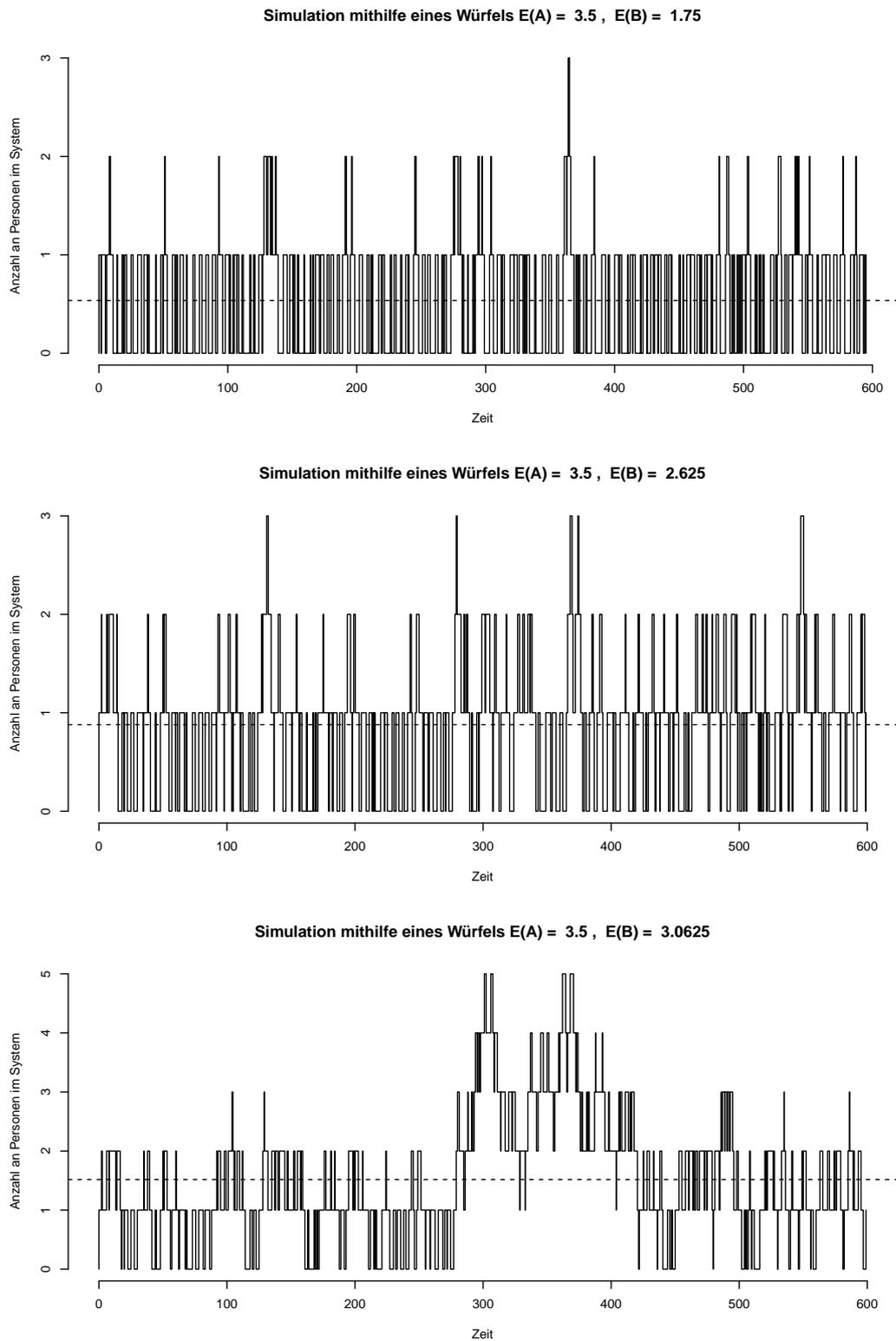


Abbildung 8.4.: Anzahl an Personen im System bei diskret gleichverteilten Ankunfts- und Bedienzeiten

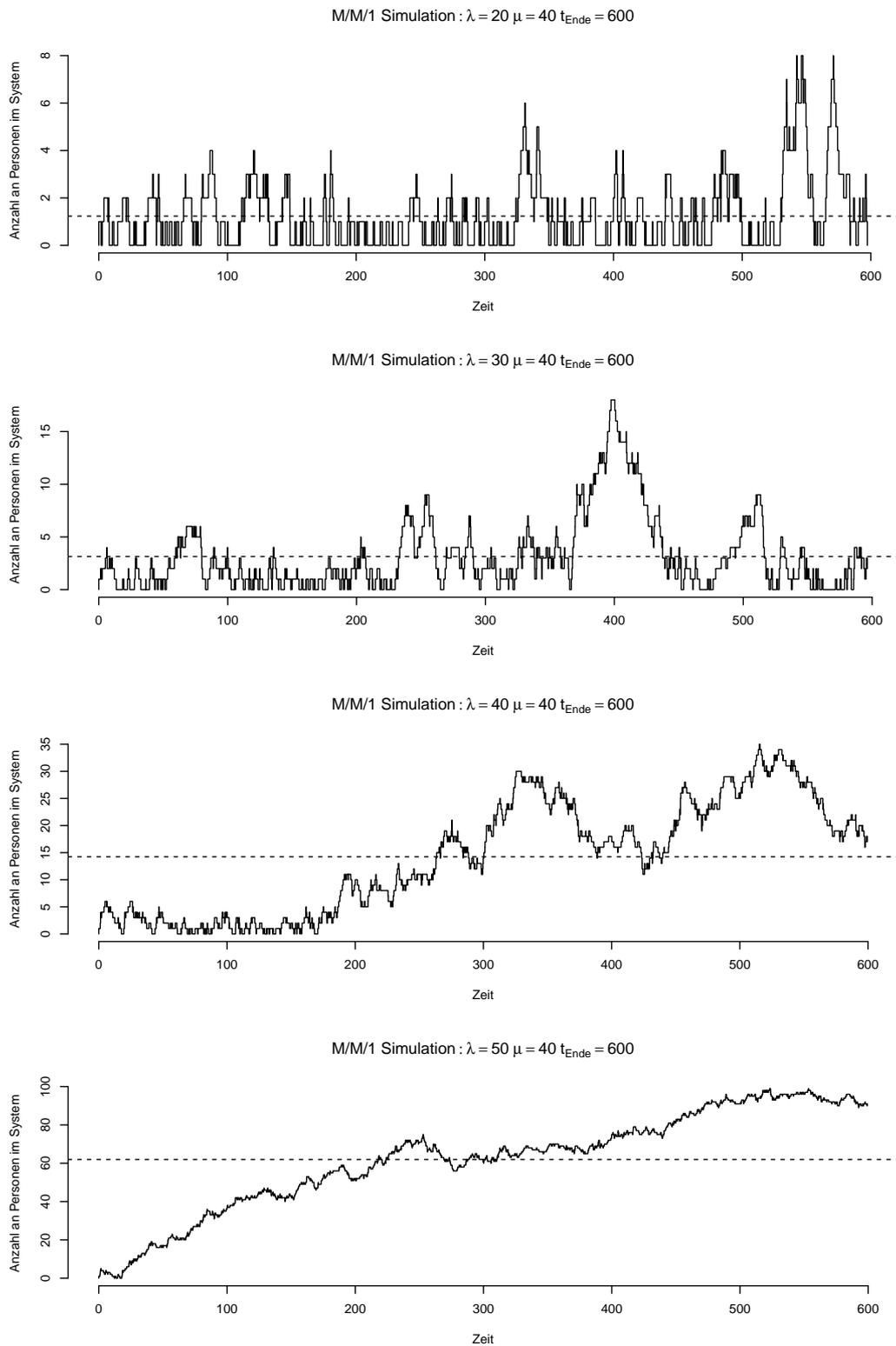


Abbildung 8.5.: Anzahl an Personen im System M/M/1

Kapitel 9.

Schließende Statistik

Dieses Kapitel widmet sich den Monte-Carlo-Methoden in der schließenden Statistik. Zuerst wird eine frequentistische Interpretation von Konfidenzintervallen gegeben und anschließend statistische Tests vorgestellt, die Monte-Carlo-Methoden verwenden.

9.1. Konfidenzintervalle

Mithilfe einer Erhebung in einer Maturaklasse haben wir bestimmte Daten (Größe, Gewicht, Blutgruppe, Geschlecht) ermittelt. Das Ergebnis dieser Erhebung ist auf Seite 114 zu finden. Wir wollen nun mithilfe dieser Stichprobe allgemeine Aussagen treffen.

Im Folgenden werden wir die Variable *Größe* genauer analysieren, da wir daran interessiert sind, wie groß der durchschnittliche Gymnasiast ist. Wir bezeichnen die durchschnittliche Körpergröße aller Gymnasiasten mit μ .

Die Berechnung des Mittelwerts \bar{x} der *Größen* von allen männlichen Schülern unserer Stichprobe ergibt den Wert

$$\bar{x} \approx 179.29$$

und könnte eine erste Schätzung für die Durchschnittsgröße sein. Man nennt \bar{x} einen Punktschätzer. Das Problem bei der Schätzung von \bar{x} ist allerdings, dass wir den eigentlichen Wert μ in der Regel nicht treffen werden. Wir sind daher daran interessiert, einen Bereich anzugeben, in dem sich die Durchschnittsgröße wahrscheinlich befindet.

Wir ermitteln ausgehend von unserer Stichprobe ein sogenanntes Konfidenzintervall, in dem sich die eigentliche Durchschnittsgröße μ mit bestimmter Wahrscheinlichkeit γ befindet.

Nehmen wir an, dass die *Größen* normalverteilt sind, können wir die Grenzen eines Konfidenzintervalls aufgrund unbekannter Varianz σ^2 und kleinem n ($n = 28 \leq 50$) folgendermaßen berechnen

$$\bar{x} \pm t_{n-1, \frac{1+\gamma}{2}} \frac{s}{\sqrt{n}}$$

wobei $t_{n-1, \frac{1+\gamma}{2}}$ dem $\frac{1+\gamma}{2}$ -Fraktile der t -Verteilung mit $n - 1$ Freiheitsgraden und

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

der Stichprobenstandardabweichung entspricht.

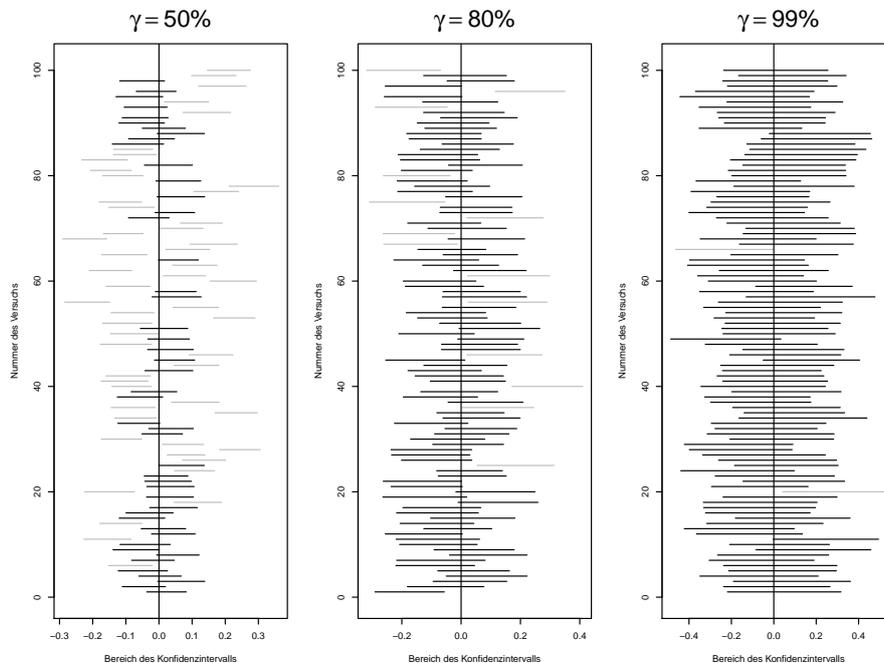
Wir wählen $\gamma = 99\%$, berechnen die benötigten Werte $s \approx 6.62$, $\bar{x} \approx 179.29$, $n = 28$ sowie $t_{27, 0.995} \approx 2.77$ und erhalten das folgende Konfidenzintervall:

$$I = [175.82, 182.75]$$

Zu 99% liegt der Mittelwert μ in dem von uns angegebenen Intervall. Können wir das festgelegte Sicherheitsniveau γ von 99% auch frequentistisch interpretieren?

Angenommen wir haben die Möglichkeit, die Erhebung in 100 Maturaklassen durchzuführen. Berechnen wir für jede dieser Stichproben ein Konfidenzintervall, so wird sich ein ähnliches Bild wie in der Abbildung auf der nächsten Seite ergeben. Kennen wir den wahren Wert μ , dann wird dieser voraussichtlich von einem Anteil γ aller Konfidenzintervalle überdeckt. Die Konfidenzintervalle, die den wahren Mittelwert μ überdecken, wurden schwarz eingezeichnet, die anderen grau.

Wählen wir beispielsweise $\gamma = 0.5 = 50\%$, dann überdecken voraussichtlich 50 der 100 Konfidenzintervalle den Wert μ . Bei $\gamma = 0.8 = 80\%$ werden in etwa 80 Konfidenzintervalle μ überdecken und bei $\gamma = 0.99 = 99\%$ ist sogar damit zu rechnen, dass nur ein einziges Intervall μ nicht überdeckt.



9.2. Statistische Tests und Resampling

Statistische Tests wurden bereits in Abschnitt 1.7 anhand des χ^2 -Tests dargestellt. Im Falle des χ^2 -Tests berechnen wir unter Verwendung unserer Nullhypothese H_0 die Teststatistik χ^2 und überprüfen, ob sie im Konfidenzintervall der χ_{k-1}^2 -Verteilung¹ mit Überdeckungswahrscheinlichkeit $\gamma = 1 - \alpha$ liegt.

Sehen wir uns hierzu nochmals die Dichtefunktion der χ_3^2 -Verteilung in Abbildung 9.1 an. Liegt der von uns ermittelte Wert χ^2 innerhalb des Konfidenzintervalls, nehmen wir die Hypothese an (oberes Bild). Man kann diesen Sachverhalt aber auch anders formulieren: Ist die graue Fläche unter der Funktion ab unserem Wert χ^2 größer als α , nehmen wir die Hypothese an. Diese Fläche bezeichnet man als **p-Wert** (unteres Bild). Der p-Wert entspricht der Wahrscheinlichkeit, dass die Werte der Teststatistik T , die χ_{k-1}^2 -verteilt sind, größer als der von uns ermittelte Wert χ^2 sind:

$$\text{p-Wert} = P(T \geq \chi^2)$$

Ist der p-Wert kleiner als die von uns vorgegebene Schranke α , werden wir die Nullhypothese verwerfen. Kennen wir die Verteilung der Teststatistik, können wir den p-Wert mithilfe der Verteilungsfunktion berechnen.

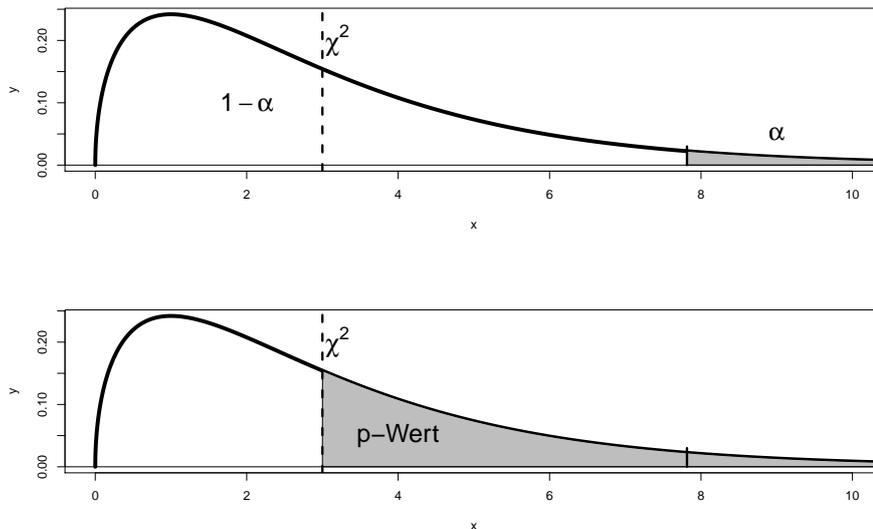


Abbildung 9.1.: Dichtefunktion der χ_3^2 -Verteilung

9.2.1. Test des Mittelwerts: t-Test vs. Bootstrapping

Kehren wir zum Beispiel der Körpergrößen zurück. Ein Schüler der Klasse ist der Meinung, dass er und seine Mitschüler besonders groß sind. Er hat gelesen, dass die Körpergröße

¹ $k - 1$ entspricht dabei der Anzahl an Freiheitsgraden

eines Mannes normalverteilt mit dem Mittelwert 179 cm ist und hält es nicht für möglich, dass die Knaben der Klasse so klein sind und ihre Größen dieser Verteilung entstammen. Formuliert man diese Fragestellung in einem statistischen Test, soll

$$H_0 : \mu \leq 179 \quad \text{gegen} \quad H_1 : \mu > 179$$

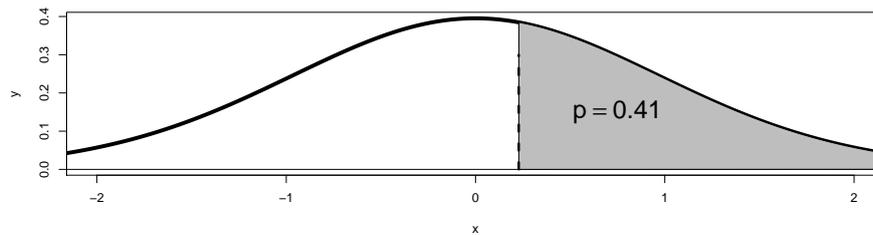
getestet werden. Wie kann man dabei vorgehen?

Eine erste Variante wäre der sogenannte **t-Test**, der die Normalverteilung der Grundgesamtheit, die in diesem Fall gegeben ist, voraussetzt. Man berechnet die Teststatistik

$$T_0 = \sqrt{n} \cdot \frac{\bar{x} - \mu_0}{s} \approx 0.23,$$

die t-verteilt mit $n - 1$ Freiheitsgraden ist. Da man die Verteilung der Teststatistik kennt, kann man mithilfe der Verteilungsfunktion der t-Verteilung den p-Wert berechnen:

$$p = P(T \geq T_0) \approx 0.41$$



Bei einem Signifikanzniveau α von beispielsweise 0.05 können wir die Nullhypothese nicht verwerfen, da $p > 0.05$ gilt.

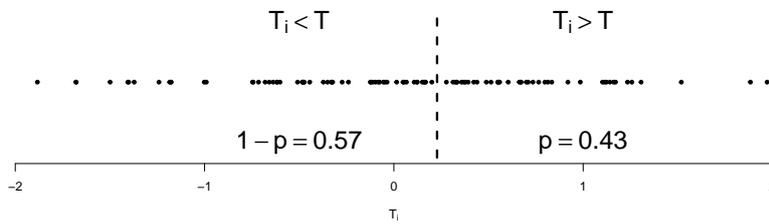
Wir konnten den p-Wert einfach berechnen, da wir die Verteilung der Teststatistik kannten. Ist das nicht der Fall, könnte man den p-Wert auch mithilfe von **Sampling** schätzen.

Wir ermitteln unseren p-Wert, indem wir viele Teststatistiken T_i ($i > 0$) berechnen und mit unserer Teststatistik T_0 vergleichen. Doch woher bekommen wir die Daten, um das zu tun?

Unsere Nullhypothese lautet, dass die Körpergrößen X normalverteilt mit dem Mittelwert $\mu = 179$ sind. Nehmen wir die Varianz σ^2 als die empirische Varianz $s^2 = 43.84$ der Stichprobe an, haben wir die Verteilung der Größen komplett beschrieben und können die Daten 1000 Mal simulieren. Wir erzeugen Stichproben der Länge $n = 28$, die dieser Verteilung entsprechen und berechnen für jede Stichprobe anschließend die Teststatistik T .

Nun können wir den p-Wert schätzen, indem wir den relativen Anteil aller Teststatistiken T_i ($i > 0$), die größer als der zuvor ermittelte Wert T_0 sind, ermitteln:

$$p = |\{T_i : T_i > T_0\}| \approx 0.43$$



Diese Variante des *Sampling* nennt man **Bootstrapping**.

9.2.2. Vergleich von Mittelwerten: t-Test vs. Permutationstest

Da wir die Körpergrößen eigentlich für zwei unterschiedliche Gruppen, für männliche (M) Schüler und weibliche (W) Schülerinnen, erhoben haben, wollen wir nun diese beiden Gruppen vergleichen. Dazu betrachten wir die Mittelwerte der beiden Beobachtungen (Größen der Knaben: $x_1, \dots, x_{m=28}$, Größen der Mädchen: $y_1, \dots, y_{n=36}$) aus unterschiedlichen Grundgesamtheiten und testen sie mithilfe des **Zwei-Stichproben-t-Tests** auf signifikante Unterschiede.

$$H_0 : \mu_x = \mu_y \quad \text{gegen} \quad \mu_x \neq \mu_y$$

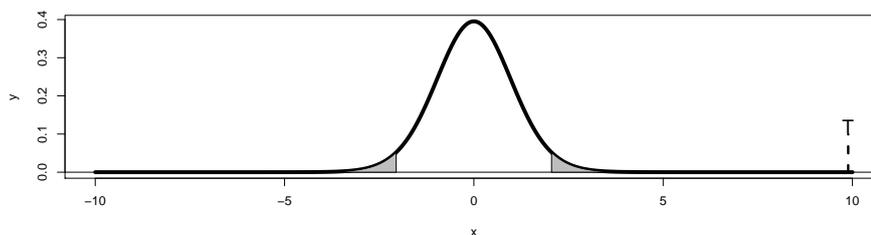
Wir nehmen wiederum eine Normalverteilung der Daten an. Da wir aber nicht wissen, ob die Varianzen der beiden Grundgesamtheiten gleich sind, verwenden wir die Variante des *Welch-Tests* mit der Teststatistik

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{m} + \frac{s_y^2}{n}}} \approx 9.89$$

die t -verteilt mit $d = \min(m - 1, n - 1)$ Freiheitsgraden ist und s_x, s_y den Stichprobenstandardabweichungen der Körpergrößen der Jungen sowie Mädchen entsprechen. Da es sich in diesem Fall um einen zweiseitigen Test handelt, liegt der Annahmebereich in der Mitte und es gibt einen linken und einen rechten Verwerfungsbereich und wir verwerfen die Nullhypothese, wenn

$$|T| > t_{d, 1 - \frac{\alpha}{2}}$$

wobei t das $1 - \frac{\alpha}{2}$ -Fraktile der t -Verteilung mit d Freiheitsgraden ist.



Wir erhalten als Wert für unsere Teststatistik $T = 9.89$ und erkennen, dass dieser deutlich im grauen Verwerfungsbereich für $\alpha = 0.05$ liegt. Der p-Wert beträgt nur $8e-13$. Es ist daher sehr unwahrscheinlich, dass die Körpergrößen der Jungen und Mädchen derselben Verteilung entsprechen.

Dasselbe Ergebnis können wir wiederum mittels Sampling erhalten. Betrachten wir hierzu den sogenannten **Permutationstest**. Da dieser Test keine bestimmte Verteilung voraussetzt, ist diese Variante auf beliebig verteilte Daten anwendbar. Daraus ergibt sich auch der Vorteil gegenüber dem t-Test, welcher normalverteilte Daten benötigt.

Sehen wir uns vor der Testdurchführung einen Ausschnitt der erhobenen Daten an.

	1	2	3	4	5	6	7	8	9	10
Größe	170	180	173	160	168	162	161	162	159	179
Geschlecht	M	M	W	W	W	W	W	W	W	M

Um die Körpergrößen der Jungen und Mädchen zu vergleichen, berechnen wir den Unterschied der Mittelwerte

$$T = \bar{x} - \bar{y} = 14.34$$

und verwenden diesen Wert als Teststatistik.

Unter der Annahme, dass beide Stichproben derselben Grundgesamtheit entstammen und somit der Faktor *Geschlecht* unerheblich ist, können wir die Werte des Faktors beliebig untereinander vertauschen (permutieren), ohne dass sich die Daten signifikant ändern. (vgl. Tabelle 9.1)

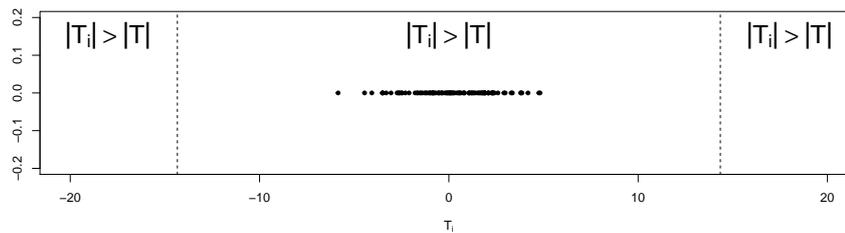
	1	2	3	4	5	6	7	8	9	10
Größe	170	180	173	160	168	162	161	162	159	179
Geschlecht	M	M	W	W	W	W	W	W	W	M
Geschlecht*1	M	W	W	W	M	W	W	W	M	W
Geschlecht*2	W	W	M	M	W	W	M	W	W	W
Geschlecht*3	W	M	W	W	W	W	W	W	M	M
Geschlecht*4	M	M	W	W	M	W	W	W	W	W
Geschlecht*5	W	W	W	M	W	W	M	M	W	W

Tabelle 9.1.: Permutationen des Faktors *Geschlecht*

Für jede Permutation $Gechlecht^*i$ des Faktors $Geschlecht$ berechnen wir die Teststatistik T_i . Den p-Wert schätzen wir als Anteil aller Teststatistiken T_i , die dem Betrag nach größer als das anfangs ermittelte T sind. Bei 1000 simulierten Permutationen erhalten wir folgenden p-Wert:

$$p = |\{T_i : |T_i| > |T|\}| \approx 0$$

Wie in der folgenden Grafik zu erkennen ist, ist kein einziges T_i größer als unser anfangs ermitteltes T und somit der p-Wert gleich 0.



Der eben durchgeführte Test wird als **Permutationstest** bezeichnet. Im Gegensatz zum *Bootstrapping*, bei dem man zufällige Daten anhand von Verteilungsannahmen, wie im vorherigen Beispiel (*parametrisches Bootstrap*), oder mithilfe der empirischen Verteilungsfunktion der Stichprobe (*nicht-parametrisches Bootstrap*) generiert, verwendet man beim *Permutationstest* ausschließlich die erhobenen Werte und würfelt diese durcheinander. Man kann sich daher *Bootstrapping* als Ziehung von Elementen aus der Grundgesamtheit *mit Zurücklegen* und den Permutationstest als Ziehung *ohne Zurücklegen* vorstellen.

Vor- und Nachteile der Resampling-Verfahren

Die Vorteile dieser sogenannten *Resampling*-Verfahren sind, dass keine Verteilungsannahmen getroffen werden müssen. Beim *Permutationstest* haben wir keinerlei Annahmen über die Verteilung der Körpergrößen getroffen. Dieser Test ist daher für beliebige Verteilungen anwendbar, während der t-Test normalverteilte Daten voraussetzt. Beim *Bootstrap*-Beispiel aus Unterabschnitt 9.2.1 haben wir keine Verteilungsannahmen bezüglich der Teststatistik getroffen. Der Nachteil der Verfahren ist, dass sie mitunter sehr rechenintensiv sind und Informationen verschwenden, sofern die Verteilung der Daten bekannt ist.

Kapitel 10.

Flächenberechnungen mit Monte-Carlo-Methoden

Im Folgenden werden wir sehen, dass Flächeninhalte, bei denen der Zufall eigentlich keine Rolle spielt, mittels Monte-Carlo-Methoden näherungsweise ermittelbar sind.

Die einführenden Abschnitte zur Treffermethode (10.1, 10.2) und zur Mittelwertmethode (10.5) stellen Verfahren zur Flächenberechnung mit Monte-Carlo-Methoden vor und zum Verständnis dieser Algorithmen sind keine Vorkenntnisse vonnöten. Man könnte sie daher bereits vor der oder als Ergänzung zur Berechnung von Integralen behandeln.

In Abschnitt 10.3 werden die in der Treffermethode verwendeten relativen Häufigkeiten mithilfe der Binomialverteilung modelliert, um theoretische Überlegungen zu den Beobachtungen aus den vorherigen Abschnitten anstellen zu können. Für das Verständnis dieses Abschnitts sind die im Schulunterricht vermittelten Grundlagen, meiner Meinung nach, ausreichend.

Die Abschnitte 10.4 und 10.6 sowie die Zusammenführung der Erkenntnisse dieser in Abschnitt 10.7 beschäftigen sich mit der Güte der Verfahren. Diese Überlegungen bilden die Grundlage dafür, dass die Verfahren „gute“ Ergebnisse liefern. Zum Verständnis dieser sind allerdings die Gütekriterien der Parameterschätzung aus Unterabschnitt 1.9, S. 24, vonnöten. Daher sehe ich diesen Teil als notwendige Grundlage für den Lehrenden, allerdings nicht notwendigerweise als Pflichtaufgabe jedes Lernenden, sondern, im Zuge eines leistungsdifferenzierten Unterrichts, als Erweiterungsaufgabe für Interessierte.

10.1. Schätzen von π

Voraussetzungen Flächeninhalt eines Quadrats bestimmen können

Durch Abzählen

Betrachten wir einen Viertelkreis mit Radius 1, der in einem Quadrat mit der Seitenlänge 1 eingebettet ist und werfen wir zufällig Punkte in dieses Quadrat - siehe Abbildung 10.1.

Da wir die Punkte gleichmäßig verteilen, können wir die Wahrscheinlichkeit, dass die Punkte unter dem Kreis landen, mithilfe geometrischer Überlegungen bestimmen. Die

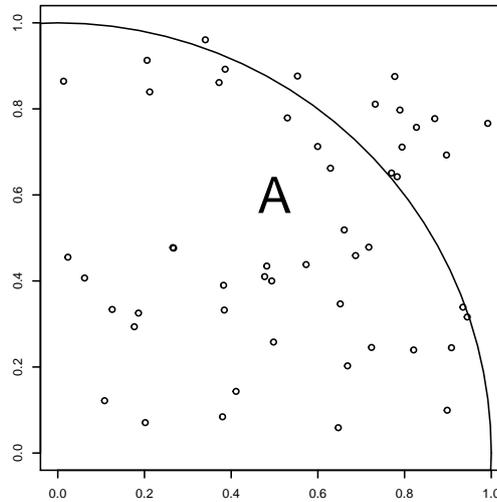


Abbildung 10.1.: Punktwolke mit 50 Punkten

Wahrscheinlichkeit, dass ein Punkt im Quadrat landet, ist natürlich 1 und entspricht daher der Fläche des Quadrats. Die Wahrscheinlichkeit, dass der Punkt unter dem Kreis landet, entspricht der Teilfläche A unter dem Kreis, also einem Wert kleiner als 1.

Wir können diese Überlegungen nun dazu verwenden, um die Fläche unter dem Kreis zu schätzen: Die Anzahl der Punkte X unter dem Kreis im Verhältnis zu allen Punkten N beträgt in etwa dem Flächeninhalt des Kreises. Wir schätzen den Flächeninhalt A daher mit:

$$\tilde{A} = \frac{X}{N}$$

In unserem Fall können wir \tilde{A} mittels Abzählen¹ bestimmen:

$$\tilde{A} = \frac{50 - 12}{50} = \frac{38}{50} = 0.76$$

Multiplizieren wir den Wert \tilde{A} mit 4, erhalten wir eine Schätzung für den Flächeninhalt des Einheitskreises, also π .

$$\tilde{\pi} = 4 \cdot \tilde{A} = 4 \cdot 0.76 = 3.04$$

Der von uns gemacht Fehler beträgt in etwa:

$$|\pi - \tilde{\pi}| \approx |3.14 - 3.04| = 0.1$$

¹ Es geht schneller, die Punkte außerhalb des Viertelkreises abzuzählen und diesen Wert von allen Punkten abzuziehen.

Mithilfe eines Computerprogramms

Wünschenswert wäre eine Schätzung, die genauer als die vorherige ist. Diese können wir in der Regel aufgrund des *Empirischen Gesetzes der Großen Zahlen* durch eine größere Anzahl an Punkten erreichen. Da wir in diesem Fall mit dem Abzählen überfordert wären, verwendet man am besten ein Computerprogramm. Die Anzahl der Punkte unter dem Viertelkreis erhalten wir mithilfe der Überprüfung von

$$x^2 + y^2 \leq 1$$

für jeden Punkt (x,y) .

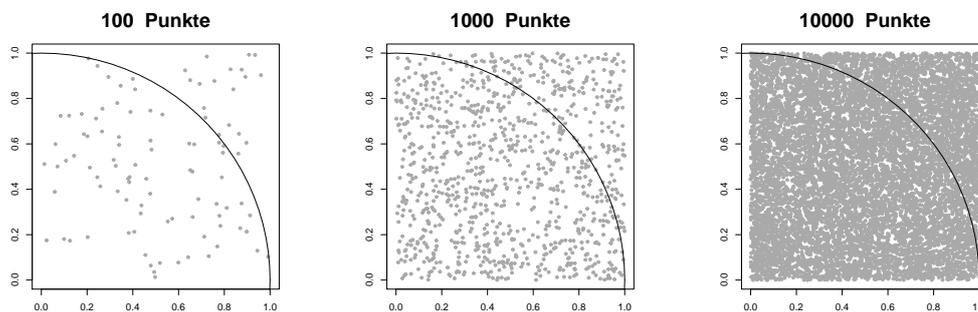


Abbildung 10.2.: Punktwolken mit 100, 1.000 und 10.000 Punkten

Die Schätzwerte für die Punktwolken in Abbildung 10.2 sind folgende:

$$\tilde{A}_1 = 0.81, \tilde{A}_2 = 0.791, \tilde{A}_3 = 0.7795$$

Daher ergeben sich die Schätzwerte für π als

$$\tilde{\pi}_1 = 3.24, \tilde{\pi}_2 = 3.164, \tilde{\pi}_3 = 3.118$$

und somit Abweichungen von

$$|\pi - \tilde{\pi}_1| \approx 0.098, |\pi - \tilde{\pi}_2| \approx 0.022, |\pi - \tilde{\pi}_3| \approx 0.024$$

Wie man sieht, kann eine höhere Anzahl an Versuchsdurchführungen, in diesem Fall Punkten, eine bessere Schätzung ergeben - sie muss es aber nicht.

Doch wie viele Versuche muss ich durchführen, damit ich ein gewisses Sicherheitsniveau erreiche? In Abschnitt 10.3 werden wir uns mit dieser Frage auseinandersetzen.

10.2. Treffermethode

Vorkenntnisse Fläche eines Rechtecks berechnen können

Das eben genannte Verfahren zum Schätzen von π kann allgemein zur näherungsweisen Berechnung von Flächen unter Kurven, d.h. Integralen, verwendet werden. Das allgemeine Verfahren nennt sich *Hit-or-Miss-Method* oder *Treffermethode*.

Betrachten wir eine Funktion $f(x)$. Suchen wir den Flächeninhalt unter der Kurve zwischen den Grenzen a und b , siehe Abbildung 10.3, dann können wir diesen erneut mit

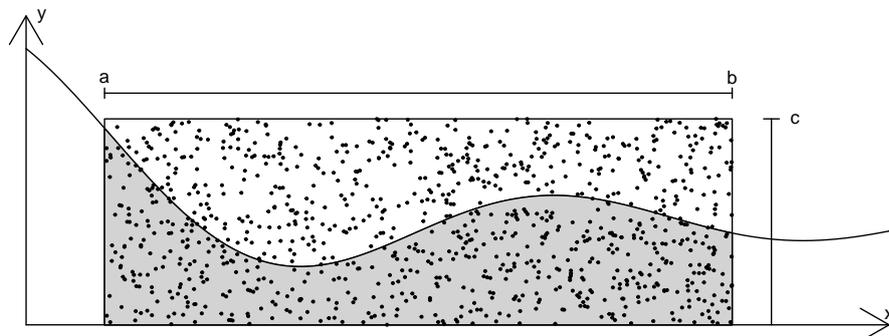


Abbildung 10.3.: Integral $\int_a^b f(x)dx = \int_2^{10} \left(\frac{2 \cdot \sin(x)}{x} + 1\right)dx$

Monte-Carlo-Methoden ermitteln. Wir umschreiben einen Teil des Funktionsgraphen mit einem Rechteck, welches durch a und b begrenzt ist und die Höhe c , wobei c größer als jeder Funktionswert in $[a, b]$ ist, aufweist. Nun werfen wir wiederum zufällige Punkte in dieses Rechteck und ermitteln die Anzahl an Punkten X unter der Kurve, indem wir für jeden Punkt (x, y) die Ungleichung $y \leq f(x)$ überprüfen. Diese Anzahl X dividieren wir durch die Anzahl aller Punkte N . Anschließend müssen wir dieses Verhältnis noch mit der Fläche des Rechtecks multiplizieren, um den geschätzten Flächeninhalt zu erhalten.

$$\tilde{A} = \frac{X}{N} \cdot (b - a) \cdot c$$

Beispiel 10.2.1 Das in Abbildung 10.3 dargestellte Integral hat den Integranden $f(x) = \frac{2 \cdot \sin(x)}{x} + 1$ und die Grenzen $a = 2$, $b = 10$ und $c = 2$. Von den 1000 zufällig erzeugten Punkten liegen 534 unterhalb der Kurve. Als geschätzten Flächeninhalt erhalten wir:

$$\tilde{A} = \frac{534}{1000} \cdot (10 - 2) \cdot 2 = 8.544$$

Da man das Integral nicht auf analytischem Weg lösen kann, greifen wir auf die numerische Berechnung von R zurück. R liefert:

8.105869 with absolute error < $9e-14$

Die Werte der beiden Verfahren unterscheiden sich somit um ungefähr 0.438.

Im Gegensatz zu unserem Monte-Carlo-Algorithmus gibt uns R zusätzlich zum Wert des Integrals eine obere Grenze für den gemachten Fehler des verwendeten Verfahrens an. Daher werden wir uns im nächsten Abschnitt damit beschäftigen, Aussagen über den Bereich, in dem sich unser Schätzwert mit großer Wahrscheinlichkeit befindet, zu machen.

10.3. Intervallschätzer und Genauigkeit der Treffermethode

Vorkenntnisse Erwartungswert, Varianz, Binomialverteilung, Normalverteilung, Grenzwertungssatz von de Moivre

Wir werden nun versuchen, anstelle eines einzelnen Wertes einen ganzen Bereich zu schätzen, in dem der gesuchte Wert liegt.

10.3.1. Konfidenzintervalle

Führen wir das Werfen von 1000 Punkten ($n = 1000$) wie in den Abbildungen 10.1 und 10.2 10000 mal durch und bestimmen wir für jede Versuchsdurchführung die Anzahl X an Punkten unterhalb des Viertelkreises. Tragen wir unsere Ergebnisse in Form eines Histo-

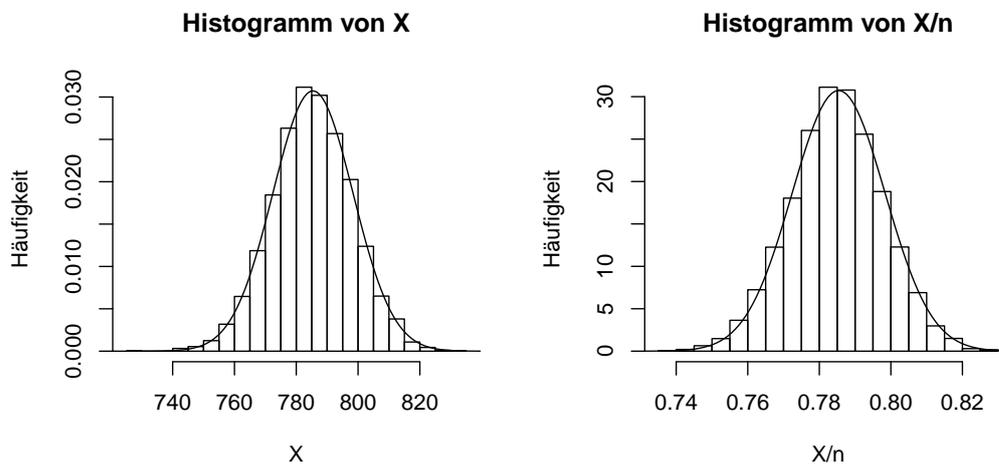


Abbildung 10.4.: Histogramm von X bzw. $\frac{X}{n}$ und Approximation durch Normalverteilung

gramms auf, siehe Abbildung 10.4, kann man bereits vermuten, dass X binomialverteilt ist. Dass dies auch wirklich der Fall ist, liegt auf der Hand. (Wir führen 1000 mal den Alternativversuch mit den zwei möglichen Ausgängen (1) Punkt liegt unter dem Viertelkreis

oder (0) er liegt oberhalb, mit den Wahrscheinlichkeiten $p(1) = \frac{\pi}{4}$ (Fläche unter Viertelkreis) und $p(0) = \frac{3\pi}{4}$ durch).

Für X gilt daher

$$X \sim B_{n,p}$$

mit

$$n = 1000, p = \frac{\pi}{4}, E(X) = n \cdot p, V(X) = n \cdot p \cdot (1 - p)$$

Uns interessiert allerdings der Wert

$$\tilde{p} = \frac{X}{n}$$

den wir als Schätzwert für p verwenden. Ausgehend von Erwartungswert und Varianz von X können wir dieselben Werte für $\frac{X}{n}$ berechnen

$$E\left(\frac{X}{n}\right) = \frac{E(X)}{n} = p, V\left(\frac{X}{n}\right) = \frac{V(X)}{n^2} = \frac{p \cdot (1 - p)}{n} \quad (10.1)$$

Die Verteilung von $\frac{X}{n}$ können wir mithilfe des *Zentralen Grenzwertsatzes* durch die Normalverteilung approximieren.

$$\frac{X}{n} \sim_{approx.} N\left(p, \frac{p \cdot (1 - p)}{n}\right)$$

Standardisieren wir $\frac{X}{n}$ zu

$$\frac{\frac{X}{n} - p}{\sqrt{\frac{p \cdot (1 - p)}{n}}} \sim N(0, 1)$$

dann gilt folgende Beziehung

$$P\left(u_{\frac{1-\gamma}{2}} \leq \frac{\frac{X}{n} - p}{\sqrt{\frac{p \cdot (1 - p)}{n}}} \leq u_{\frac{1+\gamma}{2}}\right) = \gamma$$

und wir erhalten durch Umformen und Einsetzen von \tilde{p} folgende Gleichung

$$P\left(p - u_{\frac{1+\gamma}{2}} \sqrt{\frac{p(1-p)}{n}} \leq \tilde{p} \leq p + u_{\frac{1+\gamma}{2}} \sqrt{\frac{p(1-p)}{n}}\right) = \gamma \quad (10.2)$$

Diese Gleichung besagt, dass der von uns geschätzte Wert \tilde{p} mit Wahrscheinlichkeit γ im angegebenen Bereich liegt, wenn wir n Versuche durchführen.

Konfidenzintervall für den Schätzwert von π

Wir wollen eine 95% sichere Aussage ($\gamma = 0.95$), in welchem Bereich sich \tilde{p} befindet und erhalten durch Einsetzen in obige Formel:

$$[0.76, 0.811]$$

Somit liegt unsere Schätzung von π mit 95 prozentiger Wahrscheinlichkeit in

$$[3.04, 3.244].$$

Konfidenzintervall für das Integral

In analoger Weise können wir das 95%-Konfidenzintervall für das Integral aus Abschnitt 10.2 durch Verwendung von Gleichung (10.2) angeben ($\tilde{p} = \frac{503}{1000}$ und $n = 1000$). Der Wert des Integrals liegt voraussichtlich im Intervall

$$[7.552, 8.544]$$

10.3.2. Erreichen einer gewissen Genauigkeit

Interessanter für uns ist, wie viele Versuchsdurchführungen, d.h. wie viele Punktwürfe, wir machen müssen, damit wir mit großer Sicherheit ein bestimmtes Genauigkeitsniveau erreichen.

Aus Gleichung (10.2) ergibt sich eine Länge l des Intervalls I von

$$l(I) = 2 \cdot u_{\frac{1+\gamma}{2}} \sqrt{\frac{p(1-p)}{n}} \quad (10.3)$$

Wollen wir den zu schätzenden Wert mit einer Sicherheit von γ auf ε genau bestimmen, so gilt es, folgende Gleichung nach n aufzulösen

$$2 \cdot u_{\frac{1+\gamma}{2}} \sqrt{\frac{p(1-p)}{n}} \leq 2 \cdot \varepsilon \quad (10.4)$$

Genauigkeit des Schätzwerts von π

Wir wollen π auf 0.01 genau schätzen und unsere Aussage soll eine Sicherheit von 99% aufweisen.

Da wir $\frac{\pi}{4}$ schätzen, berechnen wir zuerst das Konfidenzintervall für $\frac{\pi}{4}$ und erhalten durch Multiplikation mit 4 das Konfidenzintervall für π . Wollen wir daher π auf 0.01 genau schätzen, müssen wir $\frac{\pi}{4}$ auf $\frac{1}{4} \cdot \frac{1}{100}$ genau schätzen. Wir müssen daher

$$2 \cdot u_{0.995} \sqrt{\frac{p(1-p)}{n}} \leq 2 \cdot \frac{1}{100} \cdot \frac{1}{4}$$

gewährleisten. Durch Umformen erhalten wir

$$\left(400 \cdot u_{0.995} \cdot \sqrt{p \cdot (1-p)}\right)^2 \leq n$$

und somit als untere Grenze für n (aufgerundet)

$$178928 \leq n.$$

Wir müssen also mindestens 178928 Punkte auf das Quadrat werfen, um mit 99 prozentiger Wahrscheinlichkeit eine auf 0.01 genaue Schätzung zu erhalten.

Genauigkeit des Integrals

Zur Bestimmung der Genauigkeit des Integrals A müssen wir beachten, dass wir eigentlich $\frac{A}{16}$ schätzen und daher 16 mal so genau schätzen müssen. Für eine Schätzung von obiger Genauigkeit und dem selben Sicherheitsniveau erhalten wir als untere Grenze für die Anzahl der Versuchsdurchführungen n (aufgerundet)

$$4246181 \leq n$$

10.4. Güte der Treffermethode

Vorkenntnisse aufbauend auf Abschnitt 10.3 und zusätzlich Gütekriterien eines Punktschätzers (vgl. Seite 24)

Erwartungstreue

Wir haben bei der Treffermethode die Fläche unter einem Funktionsgraphen $f(x)$ mittels

$$\tilde{I} = \frac{X}{N} \cdot (b - a) \cdot c = \tilde{p} \cdot (b - a) \cdot c$$

geschätzt, wobei X die Anzahl der Punkte unter dem Graphen, N die Anzahl aller Punkte und $(b - a) \cdot c$ der Flächeninhalt des umgebenden Rechtecks ist.

Aufgrund der Gleichung (10.1) folgt, dass \tilde{I} ein *erwartungstreuer* Schätzer ist.

$$E(\tilde{I}) = E\left(\frac{X}{N} \cdot (b - a) \cdot c\right) = E\left(\frac{X}{N}\right) \cdot (b - a) \cdot c \stackrel{(10.1)}{=} p \cdot (b - a) \cdot c = I$$

Varianz und Konsistenz

Der Schätzer weist folgende Varianz auf

$$V(\tilde{I}) = V\left(\frac{X}{N} \cdot (b - a) \cdot c\right) = V\left(\frac{X}{N}\right) \cdot [(b - a) \cdot c]^2 \stackrel{(10.1)}{=} \frac{p \cdot (1 - p)}{n} \cdot [(b - a) \cdot c]^2 = \frac{I}{n} \cdot [c \cdot (b - a) - I] \quad (10.5)$$

und da die Varianz für wachsendes n gegen Null strebt, ist der Schätzer auch *konsistent* (Satz 1.9, Seite 24).

10.5. Mittelwertmethode

Vorkenntnisse Mittelwert bestimmen können

optional Mittelwertsatz der Integralrechnung

Betrachten wir das Integral

$$\int_a^b f(x) dx = \int_2^{10} \left(\frac{\sin(x)}{2} + \frac{3}{2} \right) dx$$

in Abbildung 10.5. Normalerweise geht man wie im linken Bild vor: Man berechnet die

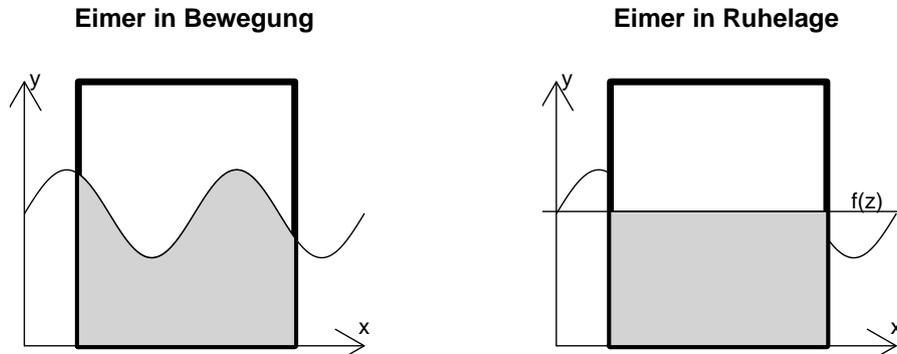


Abbildung 10.5.: Integral $\int_2^{10} (\frac{\sin(x)}{2} + \frac{3}{2}) dx$

graue Fläche unter dem Funktionsgraphen. Eine andere Möglichkeit wäre es, sich den unrandeten Bereich als Eimer und den Funktionsgraph als Oberfläche des darin enthaltenen Wassers vorzustellen. Aufgrund von Bewegungen des Eimers, erhält man eine wellenförmige Oberfläche, wie im linken Bild.

Die Idee ist nun, den Eimer in Ruhelage zu bringen und so lange zu warten, bis die Oberfläche wieder eine horizontale Linie bildet. In diesem Fall kann man den Flächeninhalt, der die Form eines Rechtecks besitzt, ganz einfach berechnen. Für diese Berechnung benötigen wir allerdings den Wasserpegel $f(z)$ in Ruhelage.

Dass ein Wert $z \in [a, b]$ existiert, sodass

$$\int_a^b f(x) dx = f(z) \cdot (b - a)$$

für $f : [a, b] \rightarrow \mathbb{R}$ gilt, besagt der **Mittelwertsatz der Integralrechnung**. Dieser gibt allerdings kein Verfahren an, um den Wert $f(z)$ zu berechnen. Wir werden daher im Folgenden versuchen, $f(z)$ mithilfe von Monte-Carlo-Methoden zu schätzen.

Wir erzeugen uns für diesen Zweck zufällige Werte $u_i \in [a, b]$ und berechnen die Funktionswerte $f(u_i)$. Das arithmetische Mittel dieser Funktionswerte

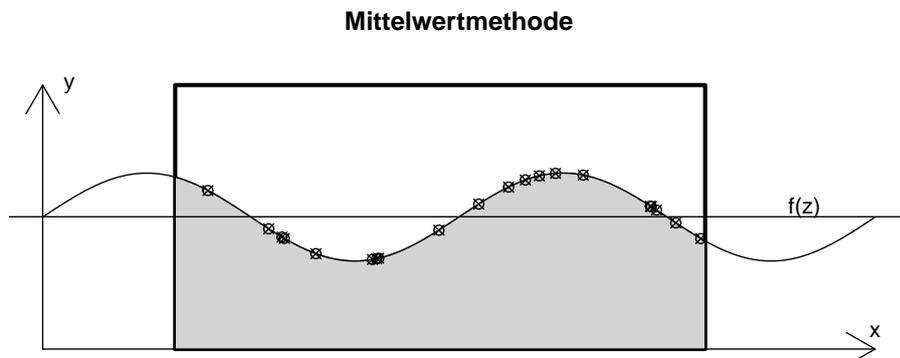
$$\widetilde{f(z)}_n = \frac{1}{n} \cdot \sum_{i=1}^n f(u_i)$$

verwenden wir als Schätzwert für $f(z)$. Multiplizieren wir $\widetilde{f(z)}_n$ mit $(b - a)$ erhalten wir einen Schätzwert für das Integral I :

$$\tilde{I}_n = \frac{(b - a)}{n} \cdot \sum_{i=1}^n f(u_i)$$

Diese Variante der Monte-Carlo-Integration nennt man **Sample-Mean Method** oder **Mittelwertmethode**.

Eine Durchführung des Verfahrens mit 20 Punkten ist in der folgenden Abbildung dargestellt.



Aus den 20 Punkten wurden folgende Schätzwerte ermittelt:

$$\widetilde{f(z)}_{20} = 1.5 \text{ sowie } \tilde{I}_{20} = 12$$

Der exakte Wert des Integrals beträgt:

$$\int_2^{10} \left(\frac{\sin(x)}{2} + \frac{3}{2} \right) dx = \left(-\frac{\cos(x)}{2} + \frac{3}{2} \cdot x \right) \Big|_2^{10} \approx 12.21$$

Verwendet man mehr Punkte zur Approximation, erhält man in der Regel bessere Näherungswerte

$$\tilde{I}_{10^2} \approx 12.135, \tilde{I}_{10^3} \approx 12.064, \tilde{I}_{10^4} \approx 12.202, \tilde{I}_{10^5} \approx 12.226, \tilde{I}_{10^6} \approx 12.216$$

10.6. Güte der Mittelwertmethode

Voraussetzungen Erwartungswert, Varianz, Gütekriterien eines Punktschätzers (Seite 24)

Erwartungstreue

Betrachten wir das Integral

$$I = \int_a^b f(x)dx = \int_a^b \frac{f(x)}{\frac{1}{b-a}} \cdot \frac{1}{b-a} dx$$

Da $\frac{1}{b-a}$ die Dichtefunktion der stetigen Gleichverteilung ist, handelt es sich bei obigem Ausdruck um den Erwartungswert

$$E\left(\frac{f(U)}{\frac{1}{b-a}}\right) = E(f(U)) \cdot (b-a)$$

wobei U eine auf $[a,b]$ stetig gleichverteilte Zufallsgröße ist. Weil das arithmetische Mittel ein erwartungstreuer Schätzer ist, folgt, dass auch unser Schätzer \tilde{I} erwartungstreu ist

$$\tilde{I} = \frac{1}{n} \sum_{i=1}^n (f(u_i)) \cdot (b-a)$$

Varianz und Konsistenz

Die Berechnung der Varianz, mittels Verschiebungssatz $V(X) = E(X^2) - E(X)^2$, ergibt

$$V(\tilde{I}) = \frac{1}{n} \cdot \left((b-a) \cdot \int_a^b f^2(x)dx - I^2 \right) \quad (10.6)$$

Es folgt wiederum die *Konsistenz* des Schätzers. (vgl. Satz 1.9, Seite 24)

10.7. Effizienz - Vergleich der beiden Verfahren

Voraussetzungen Conclusio aus den Abschnitten 10.6 und 10.4.

Wir haben nun zwei unterschiedliche Verfahren zur Ermittlung eines Integrals mit Monte-Carlo-Methoden besprochen. Der Schätzer des jeweiligen Verfahrens, der eine geringere Varianz aufweist, heißt *effizient*. Bezeichnen wir mit \tilde{I}_T den Schätzer der *Treffermethode* und mit \tilde{I}_M jenen der *Mittelwertmethode*. Subtrahieren wir (10.6) von (10.5), dann erhalten wir

$$V(\tilde{I}_T) - V(\tilde{I}_M) = \frac{1}{n} \cdot (b-a) \cdot \left(c \cdot I - \int_a^b f^2(x)dx \right)$$

Wegen

$$f(x) \leq c$$

gilt

$$c \cdot I - \int_a^b f^2(x)dx \geq 0$$

und ferner

$$V(\tilde{I}_T) - V(\tilde{I}_M) \geq 0.$$

Deswegen ist die *Mittelwertmethode* *effizienter* als die *Treffermethode*.

Literaturverzeichnis

- [1] ARENS, T., F. HETTLICH, C. KARPFINGER, U. KOCKELKORN, K. LICHTENEGGER und H. STACHEL: *Mathematik*. Spektrum Akademischer Verlag, Heidelberg, 2009.
- [2] BECKER, G., J. HENNING, V. LINDENAU, K.-D. MAI und M. SCHINDLER: *Anwendungsorientierter Mathematikunterricht*, Kap. Warteschlangen, S. 51–77. Julius Klinkhardt, Bad Heilbrunn/Obb., 1979.
- [3] BEHR, A.: *Einführung in die Statistik mit R*. Verlag Franz Vahlen, München, 2005.
- [4] BIEHLER, R., A. EICHLER, J. ENGEL und E. WARMUTH: *Leitidee Daten und Zufall für die Sekundarstufe II – Kompetenzprofile für die Bildungsstandards aus Sicht der Stochastik und ihrer Didaktik*. Paderborn [u.a.], 2010.
- [5] BÜCHTER, A. und H.-W. HENN: *Elementare Stochastik*. Springer, Berlin, Heidelberg, 2. Aufl., 2007.
- [6] CLAUS, H. J.: *Einführung in die Didaktik der Mathematik*. Wissenschaftliche Buchgesellschaft, Darmstadt, 2. Aufl., 1995.
- [7] DAHL, D. B.: *xtable: Export tables to LaTeX or HTML*, 2013. R package version 1.7-1.
- [8] GUNTHER, N.: *Simulating a Queue in R*. <http://perfdynamics.blogspot.co.at/2010/05/simulating-queue-in-r.html>, 2010. [Online; zugegriffen am 13. April 2013].
- [9] HAGMANN, S.: *Erzeugung, Tests und Anwendungen von Pseudo-Zufallszahlen*. Diplomarbeit, Technische Universität Wien, 2000.
- [10] HELLER, W.-D., H. LINDENBERG, M. NUSKE und K.-H. SCHRIEVER: *Stochastische Systeme*. Walter de Gruyter, Berlin, New York, 1978.
- [11] KOLONKO, M.: *Stochastische Simulation*. Vieweg+Teubner, Wiesbaden, 1. Aufl., 2008.
- [12] KRAMPE, H., J. KUBÁT und W. RUNGE: *Bedienungsmodelle*. Oldenbourg, München, Wien, 1973.
- [13] KÜTTING, H.: *Didaktik der Stochastik*. BI-Wiss.-Verl., Mannheim [u.a.], 1994.
- [14] KÜTTING, H. und M. SAUER: *Elementare Stochastik*. Spektrum Akademischer Verlag, Heidelberg, 3. Aufl., 2011.

- [15] LEMON, J.: *Plotrix: a package in the red light district of R*. R-News, 6(4):8–12, 2006.
- [16] LEUDERS, T. (Hrsg.): *Mathematik Didaktik*. Cornelsen Verlag Scriptor, Berlin, 2003.
- [17] LEUDERS, T.: *Perspektiven von Mathematikunterricht*, S. 15–58. In: [16], 2003.
- [18] MALLE, G., E. RAMHARTER, A. ULOVEC und S. KANDL: *Mathematik verstehen 8. öbv*, Wien, 2007.
- [19] MÜLLER-GRONBACH, T., E. NOVAK und K. RITTER: *Monte Carlo-Algorithmen*. Springer, Berlin, Heidelberg, 2012.
- [20] R DEVELOPMENT CORE TEAM: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011.
- [21] REICH, K.: (Hrsg.): *Methodenpool*. In: url: <http://methodenpool.uni-koeln.de>.
- [22] RIZZO, M. L.: *Statistical Computing with R*. Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton [u.a.], 2008.
- [23] RSTUDIO TEAM: *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA, 2012.
- [24] RUBINSTEIN, R. Y.: *Simulation and the Monte Carlo Method*. John Wiley & Sons, New York [u.a.], 1981.
- [25] SCHWARZ, A.: *Jackknife- und Bootstrap-Verfahren*. Diplomarbeit, Universität Wien, 1992.
- [26] STADLER, H.: *Paradoxien der Wahrscheinlichkeitsrechnung und Statistik, Teil 1*. Didaktik der Mathematik, 14(2):134–152, 1986.
- [27] TIETZE, U.-P., H. WOLPERS und M. KLIKA (Hrsg.): *Mathematikunterricht in der Sekundarstufe II. 3. Didaktik der Stochastik*. Vieweg, Braunschweig [u.a.], 2002.
- [28] TRAUERSTEIN, H.: *Zur Simulation mit Zufallsziffern im Mathematikunterricht der Sekundarstufe I*. Stochastik in der Schule, 10(2), 1990.
- [29] VENABLES, W. N. und B. D. RIPLEY: *Modern Applied Statistics with S*. Springer, New York, 4. Aufl., 2002.
- [30] WITTMANN, E.: *Grundfragen des Mathematikunterrichts*. Friedrich Vieweg & Sohn, Braunschweig, Wiesbaden, 6. Aufl., 1981.

Anhang

Simulationstafel zum Drei-Türen-Problem

1.	Tür mit Geschenk	Gewählte Tür	Gezeigte Tür
	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>
	Tür bei Wechsel	Wechsel	Kein Wechsel
	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>	gewonnen verloren <input type="radio"/> <input type="radio"/>	gewonnen verloren <input type="radio"/> <input type="radio"/>
2.	Tür mit Geschenk	Gewählte Tür	Gezeigte Tür
	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>
	Tür bei Wechsel	Wechsel	Kein Wechsel
	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>	gewonnen verloren <input type="radio"/> <input type="radio"/>	gewonnen verloren <input type="radio"/> <input type="radio"/>
3.	Tür mit Geschenk	Gewählte Tür	Gezeigte Tür
	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>
	Tür bei Wechsel	Wechsel	Kein Wechsel
	Tür 1 Tür 2 Tür 3 <input type="radio"/> <input type="radio"/> <input type="radio"/>	gewonnen verloren <input type="radio"/> <input type="radio"/>	gewonnen verloren <input type="radio"/> <input type="radio"/>

Zufallszahlen von 1 bis 365

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
1	292	16	116	164	339	69	307	184	243	282	308	47	17	244
2	103	202	116	103	271	155	331	102	119	295	92	35	321	260
3	226	316	196	175	220	281	43	106	249	66	324	175	211	25
4	271	319	185	80	149	231	362	45	255	246	132	166	271	242
5	182	254	96	183	1	320	338	85	252	108	74	37	231	363
6	256	266	319	110	138	126	247	192	302	21	273	19	69	69
7	115	236	31	276	173	135	204	325	119	188	25	35	28	112
8	297	232	107	249	119	257	129	22	137	330	123	39	37	158
9	216	81	166	258	267	80	302	143	305	329	248	99	216	160
10	128	338	217	7	124	43	130	258	86	328	55	333	191	258
11	85	138	276	206	74	243	313	213	285	268	121	113	281	335
12	327	154	9	150	62	265	294	18	287	232	109	67	201	308
13	58	227	259	85	256	181	361	230	166	60	157	183	65	313
14	293	159	129	213	365	328	211	192	303	310	240	10	199	136
15	68	59	211	169	306	364	156	345	296	171	102	331	71	354
16	272	49	25	274	338	64	232	350	237	210	130	149	5	338
17	144	245	28	166	273	139	210	38	105	32	64	31	256	229
18	171	287	319	201	160	1	23	259	151	116	62	277	224	40
19	97	294	223	115	64	132	129	338	267	131	41	134	289	136
20	52	76	306	77	210	96	287	169	1	273	181	134	60	95
21	95	94	179	108	10	87	80	187	256	187	118	301	93	27
22	180	90	209	21	350	331	288	182	24	39	318	108	100	139
23	199	100	292	67	40	5	308	241	259	160	206	139	353	16
24	177	219	315	288	209	237	185	230	71	152	295	201	250	67
25	217	347	158	6	285	196	34	44	267	289	227	184	219	38

Wahrscheinlichkeiten der möglichen Spielergebnisse

Gewinnwahrscheinlichkeit der Heimmannschaft

	D	F	I	K	O	S
D	0.00	0.94	0.69	0.91	0.76	0.75
F	0.02	0.00	0.05	0.27	0.08	0.04
I	0.13	0.86	0.00	0.63	0.44	0.20
K	0.04	0.55	0.18	0.00	0.17	0.06
O	0.11	0.84	0.27	0.77	0.00	0.29
S	0.14	0.88	0.64	0.87	0.46	0.00

Wahrscheinlichkeit für ein Unentschieden

	D	F	I	K	O	S
D	0.00	0.05	0.21	0.07	0.18	0.17
F	0.06	0.00	0.21	0.18	0.14	0.09
I	0.22	0.10	0.00	0.25	0.29	0.24
K	0.10	0.27	0.26	0.00	0.25	0.10
O	0.19	0.10	0.29	0.17	0.00	0.29
S	0.21	0.09	0.22	0.09	0.28	0.00

Gewinnwahrscheinlichkeit der Auswärtsmannschaft

	D	F	I	K	O	S
D	0.00	0.02	0.11	0.02	0.06	0.08
F	0.92	0.00	0.74	0.55	0.78	0.87
I	0.65	0.04	0.00	0.12	0.27	0.56
K	0.87	0.18	0.56	0.00	0.58	0.84
O	0.69	0.06	0.44	0.06	0.00	0.43
S	0.65	0.03	0.13	0.04	0.26	0.00

Statistische Erhebung beim Maturajahrgang eines Gymnasiums²

Größe	Gewicht	Blutgruppe	Geschlecht	Größe	Gewicht	Blutgruppe	Geschlecht
170	65	0	M	170	50	A	W
180	65	A	M	181	69	AB	M
173	59	0	W	182	74	0	M
160	48	B	W	165	53	AB	W
168	54	A	W	168	54	B	W
162	53	A	W	172	69	A	M
161	45	B	W	181	82	B	M
162	53	0	W	173	54	A	W
159	51	0	W	169	53	0	W
179	65	A	M	184	66	0	M
167	71	AB	M	194	86	A	M
180	68	A	M	183	73	B	M
168	55	A	W	172	65	A	M
164	53	0	W	183	78	0	M
168	69	B	W	165	56	0	W
160	45	A	W	167	60	A	M
169	55	A	W	164	53	B	W
159	50	AB	W	162	49	A	W
183	60	B	M	158	58	A	W
179	73	0	M	166	54	A	W
179	68	0	M	190	85	0	M
160	60	B	W	160	52	A	W
170	80	A	M	162	55	0	W
165	50	A	W	167	58	0	W
189	70	A	M	172	57	A	W
177	53	AB	M	165	55	AB	W
164	53	0	W	168	62	A	W
174	68	A	M	186	78	B	M
164	53	0	W	174	54	0	W
176	65	0	M	159	48	A	W
165	54	A	W	181	70	0	M
182	73	A	M	179	76	A	M

²Quelle: Vorlesungsunterlagen von Heinz Stadler zu „Stochastik für Lehramt 2“, WS 2012, Technische Universität Wien