



universität  
wien

# DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

## Monte-Carlo Simulation Eine Simulation von Zufallszahlen im Mathematikunterricht

verfasst von / submitted by

Carolin Schauburger

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

A 190 406 299

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Lehramtsstudium UF Mathematik UF Psychologie/Philosophie

Betreut von / Supervisor:

ao. Univ.-Prof. Mag. Dr. Peter Raith



## Eidesstattliche Erklärung

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

---

Ort, Datum

---

Unterschrift



# Danksagung

Ich möchte mich besonders bei meinem Betreuer ao. Univ.-Prof. Mag. Dr. Peter Raith bedanken, der mich während meiner Diplomarbeit beraten und umfangreich unterstützt hat und mit seinen Anregungen eine mehr als hilfreiche Begleitung war.

Weiters möchte ich mich bei meiner Familie und meinen Freunden bedanken, die mich durch mein Studium begleitet und mich bei der Entstehung dieser Arbeit moralisch unterstützt haben.

Der größte Dank gilt allerdings meinen Eltern, die mich in schwierigen Situationen stets daran erinnert haben, mein Ziel nicht aus den Augen zu verlieren, was mich immer wieder bestärkt und motiviert hat. Außerdem sind sie mir bei der Abfassung der Diplomarbeit stets mit konstruktiven Beiträgen und Änderungsvorschlägen beigestanden. Danke, dass ihr immer für mich da seid.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>9</b>
<b>2</b>	<b>Geschichtlicher Hintergrund</b>	<b>10</b>
<b>3</b>	<b>Zufallsexperimente und Ereignisse</b>	<b>12</b>
<b>4</b>	<b>Zufallsvariablen</b>	<b>15</b>
4.1	Diskrete Zufallsvariablen . . . . .	16
4.1.1	Wahrscheinlichkeitsverteilungen diskreter Zufallsvariablen . .	17
4.2	Stetige Zufallsvariablen . . . . .	20
4.2.1	Wahrscheinlichkeitsverteilungen stetiger Zufallsvariablen . .	20
<b>5</b>	<b>Ausgewählte Wahrscheinlichkeitsverteilungen</b>	<b>24</b>
5.1	Normalverteilung . . . . .	24
5.2	Logarithmische Normalverteilung . . . . .	27
5.3	Binomialverteilung . . . . .	28
5.4	Poissonverteilung . . . . .	32
5.5	Anpassung von Zufallsvariablen an Verteilungsfunktionen . . . . .	33
5.5.1	Addiert oder multipliziert die Natur? . . . . .	34
<b>6</b>	<b>Zufallszahlen und deren Generatoren</b>	<b>38</b>
6.1	Zufallszahlen . . . . .	38
6.2	Zufallszahlengeneratoren . . . . .	38
6.2.1	Echte Zufallszahlengeneratoren . . . . .	39
6.2.2	Pseudozufallszahlengeneratoren . . . . .	40
6.2.3	Startwertproblem . . . . .	42
<b>7</b>	<b>Tests von Zufallszahlen</b>	<b>43</b>
7.1	Anpassungstests . . . . .	43
7.2	Tests für die Unabhängigkeit . . . . .	44
7.2.1	Runttest . . . . .	44
7.2.2	Pokertest . . . . .	45
7.2.3	Lückentest . . . . .	46
<b>8</b>	<b>Transformation von Zufallszahlen</b>	<b>47</b>
8.1	Direkte Transformation . . . . .	47
8.2	Inversionsmethode . . . . .	47
8.3	Verwerfungsmethode . . . . .	49
<b>9</b>	<b>Simulation von Zufallszahlen im Unterricht</b>	<b>51</b>

9.1	Der Modellbildungsprozess . . . . .	51
9.2	Modellieren im Unterricht . . . . .	52
9.3	Anwendungen im Unterricht . . . . .	54
9.3.1	Kombinatorik am Beispiel des Rebhuhnspiels . . . . .	55
9.3.2	Abschätzung der Zahl $\pi$ . . . . .	60
9.3.3	Random Walk: Das Herumirren in einer Stadt . . . . .	65
9.3.4	Vergleich von empirischen Daten mit der Normalverteilung .	70
9.3.5	Ableitung der Exponentialfunktion aus empirischen Daten .	72
<b>10</b>	<b>Zusammenfassung</b>	<b>76</b>
<b>11</b>	<b>Abstract</b>	<b>77</b>



# 1 Einleitung

Die meisten Glücks- und Gesellschaftsspiele haben eine Zufallskomponente um den Spielverlauf abwechslungsreich und interessant zu gestalten. Zu den ältesten Spielen zählt der Münzwurf und das Würfelspiel. Beide Spiele sind reine Glücksspiele mit einer Zufallsvariable. Entspricht diese Zufallsvariable einer Zahl, wird von einer Zufallszahl gesprochen. Die Verwendung von Zufallszahlen in der Naturwissenschaft und der Statistik wird mit dem Begriff Monte-Carlo verknüpft. Zufallszahlen werden in den Naturwissenschaften genutzt, um Zufallsvariablen mit vorgegebener Eigenschaft zu erzeugen. Mit Hilfe der Monte-Carlo Simulation können Prozesse simuliert werden. Es handelt sich um eine Methode der Stochastik, bei der eine große Zahl von Zufallsexperimenten simuliert werden.

Ziel dieser Arbeit ist es, verschiedene Beispiele im Unterricht so aufzubereiten, dass ein mathematisch komplexes Themengebiet im Schulunterricht gelehrt werden kann.

Den Anfang der Arbeit bildet ein Überblick über den geschichtlichen Hintergrund der Monte-Carlo Simulation. Im anschließenden Kapitel werden relevante Aspekte der Wahrscheinlichkeitstheorie behandelt. Dabei werden wesentliche Begriffe und Definitionen angeführt und erklärt. In Abschnitt vier werden neben den diskreten und stetigen Zufallsvariablen auch die dazu passenden Wahrscheinlichkeitsverteilungen erläutert. Im nächsten Abschnitt werden ausgewählte Wahrscheinlichkeitsverteilungen, wie zum Beispiel die Normal- oder die Poissonverteilung, behandelt. In Abschnitt sechs werden die Begriffe Zufallszahl und Pseudozufallszahl erklärt. Außerdem werden neben physikalischen Verfahren auch mathematische Methoden beschrieben, um Zufallszahlen beziehungsweise Pseudozufallszahlen zu erzeugen. Im folgenden Abschnitt werden unterschiedliche Tests von Zufallszahlen, mit deren Hilfe es möglich ist, die theoretischen Anforderungen zu überprüfen, bearbeitet. In Abschnitt acht werden verschiedenen Transformationen von Zufallszahlen behandelt. Diese beschäftigen sich damit, wie gleichverteilte Zufallszahlen in beliebig verteilte Zufallszahlen übergeführt werden können.

Im letzten Abschnitt wird die Simulation von Zufallszahlen im Unterricht, sowie Anwendungsbeispiele im Unterricht behandelt. Dabei werden verschiedene Fragestellungen mit unterschiedlichen Schwierigkeitsgraden aufbauend erklärt. Solch ein Anwendungsbeispiel gliedert sich in folgende drei Bereiche: Problemstellung, Theorie und Durchführung des Experiments. Bei der Durchführung wird einerseits das Experiment in der Klasse beschrieben und andererseits die exemplarische Auswertung mit Hilfe eines Tabellenkalkulationsprogramms dargestellt.

## 2 Geschichtlicher Hintergrund

Die Monte-Carlo Simulation wurde schon angewendet, bevor die theoretischen Grundlagen dazu untersucht wurden. Enrico Fermi arbeitete um 1930 mit Zufallszahlen um die Eigenschaften des, aus theoretischen Überlegungen vorhergesagten, Neutrons zu untersuchen. Damals wurden die Zufallszahlen einer sogenannten Zufallszahlentabelle entnommen [1]. Fermi war seit 1927 Professor für Physik in Rom, beschäftigte sich mit Elektrodynamik und entwickelte eine Statistik über das Verhalten von Elektronen. 1938 erhielt er den Nobelpreis für Physik für die Identifizierung neuer radioaktiver Elemente. Nach seiner Auswanderung nach Amerika arbeitete Fermi am Manhattan-Projekt mit [2], der Entwicklung und Bau der Atombombe. Dabei wurde ab 1940 die Monte-Carlo Methode intensiv eingesetzt und es konnten komplexe Kernreaktionen berechnet werden. Durch die Entwicklung von Computern konnte auch die Monte-Carlo Methode reifen und präzisiert werden [3, S. 23].

Stanislaw Ulam arbeitete bei dem Manhattan-Projekt mit. Er studierte Mathematik, war Professor an der University of Wisconsin und beschäftigte sich hauptsächlich mit Logik, Topologie und Maßtheorie [4]. Er fragte sich, wie groß die Wahrscheinlichkeit ist, Solitaire mit 52 Karten zu gewinnen. Er kam auf die Idee, alle möglichen Kombinationen um das Spiel zu gewinnen durchzugehen. Doch bald stieß er auf das Problem, dass es zu viele Möglichkeiten gibt, um diese durchspielen zu können. Daraufhin spielte er 100 zufällig sich ergebende Spiele und schrieb sich die gewonnenen Ergebnisse auf. Anhand dieser Idee bemerkte er, dass er ein sehr komplexes und fast unlösbares Problem in ein einfaches umgewandelt hatte [5].

Diese Methode, von etwas Kompliziertem auf etwas einfaches zu schließen, faszinierte auch John von Neumann. Neumann war ein angesehener Mathematiker und studierte auch Industrie- Chemie. Er war an den unterschiedlichsten Universitäten in Europa und den USA tätig. Neumann befasste sich mit den unterschiedlichsten Bereichen der Mathematik, unter anderem mit der Spieltheorie, Mengenlehre und Quantenmechanik [6]. Ulam und Neumann arbeiteten gemeinsam am Manhattan-Projekt [7].

Auch Nicolas Metropolis arbeitete am Manhattan-Projekt mit. Er war Professor für Physik in Chicago und war an der Entwicklung der frühesten Computer beteiligt. 1949 erfand er gemeinsam mit Ulam die Monte-Carlo Simulation, die numerische Berechnung der Zustandsgleichung für Materie unter externen Bedingungen. 1953 wurde die Monte-Carlo Simulation veröffentlicht- eine Variante davon wird auch Metropolisalgorithmus genannt.

Der Name selbst wurde von Metropolis geprägt, da Ulam einen Onkel hatte, der

sich immer Geld von Verwandten ausborgte, um nach "Monte-Carlo zu gehen". Dies ist natürlich eine Anspielung auf das Casino in Monte-Carlo [45].

### 3 Zufallsexperimente und Ereignisse

Unter einem Zufallsexperiment wird in der Wahrscheinlichkeitstheorie ein Versuch verstanden, der unter genau festgelegten Bedingungen durchgeführt wird und einen zufälligen Ausgang hat. Solch ein Versuch ist einmalig und unwiederholbar. Am Beispiel eines Münzwurfes sind diese Bedingungen ersichtlich, der Ausgang beim Werfen ist nicht vorhersehbar und der Vorgang kann beliebig oft bei gleichen Bedingungen wiederholt werden. Weiters muss jedes Zufallsexperiment zumindest zwei Ausgangsmöglichkeiten bei jedem Versuch haben.

Ein weiteres Beispiel für solch ein Zufallsexperiment, ist das Werfen eines Würfels. Auch hier treffen die drei Bedingungen zu. Erstens das Ergebnis ist nicht vorhersehbar, zweitens ist es beliebig oft wiederholbar und drittens gibt es sechs mögliche Ausgänge.

Die möglichen Ausgänge eines Zufallsexperiment werden mit dem griechischen Buchstaben  $\Omega$  beschrieben. Diese Menge wird Ergebnismenge genannt und ist wie folgt definiert:

$$\Omega = \{\omega_1, \omega_2, \omega_3, \dots\} \quad (1)$$

Ist  $\Omega$  jedoch endlich, setzt sich die Ergebnismenge folgendermaßen zusammen,  $\Omega = \{\omega_1, \omega_2, \omega_3, \dots, \omega_n\}$ .

Von dem oben genannten Würfelbeispiel ist die Ergebnismenge endlich. Sie enthält nur die sechs möglichen Ausgänge,  $\Omega = \{1, 2, 3, 4, 5, 6\}$ , die eines Münzwurfes nur zwei  $\Omega = \{Kopf, Zahl\}$ .

Weiters wird eine Menge  $E$  als Ereignismenge definiert. Darunter wird eine Teilmengen der Ergebnismenge verstanden. Wird hier noch einmal das Würfelbeispiel betrachtet, wird als Ergebnismenge alle möglichen Ergebnisse betrachtet. Im Gegenzug dazu können auch nur die Ergebnisse betrachtet werden bei denen eine Quadratzahl herauskommt. In dem Fall würde die Ereignismenge  $E$  wie folgt aussehen,  $E = \{1, 4\}$ . Dies ist nun eine Teilmenge von  $\Omega$ . Neben der Ereignismenge wird auch sehr oft den Begriff der Potenzmenge  $\mathcal{P}(\Omega)$  verwendet. Unter einer Potenzmenge wird ein Mengensystem verstanden. Dabei werden aus der Potenzmenge  $\mathcal{P}(\Omega)$  neue Mengen gebildet, die aus Teilmengen  $U$  von  $\Omega$  besteht. Dies sieht wie folgt aus:

$$\mathcal{P}(\Omega) := \{U | U \subseteq \Omega\} \quad (2)$$

Von einem sicheren Ereignis wird gesprochen, wenn die Ereignismenge  $E$  der Ergebnismenge entspricht,  $E = \Omega$ . Liegt jedoch eine leere Menge als Ergebnismenge vor, wird diese als unmögliches Ereignis bezeichnet.

Unter dem Begriff der  $\sigma$ -Algebra wird ein Mengensystem verstanden, wenn  $\Omega \neq \emptyset$  und eine Familie  $\mathcal{A}$  von Teilmengen von  $\Omega$  ist und gilt:

1.  $\Omega \in \mathcal{A}$
2. Falls  $A \in \mathcal{A}$ , dann ist  $\Omega \setminus A$ , dass bedeutet, wenn  $A \in \mathcal{A}$ , dann ist auch das Komplement enthalten.
3. Falls  $(A_j)_{j \in \mathbb{N}} \in \mathcal{A}$ , dann ist  $\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{A}$

Das Wahrscheinlichkeitsmaß  $P$  und der Wahrscheinlichkeitsraum sind weitere wesentliche Begriffe. Sei  $\Omega \neq \emptyset$  und  $\mathcal{A}$  eine  $\sigma$ -Algebra auf  $\Omega$ . Eine Funktion  $P : \mathcal{A} \mapsto \mathbb{R}$  heißt Wahrscheinlichkeitsmaß, falls folgendes gilt:

1.  $\forall A \in \mathcal{A} : 0 \leq P(A) \leq 1$
2.  $P(\Omega) = 1$
3. Falls  $(A_j)_{j \in \mathbb{N}} \in \mathcal{A}$  paarweise disjunkt sind, dann gilt  $P(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} P(A_j)$ .

Dann wird  $(\Omega, \mathcal{A}, P)$  einen Wahrscheinlichkeitsraum genannt [8].

Abschließend noch eine klassische Definition der Wahrscheinlichkeit nach Laplace. Bei der Definition ist ein Ereignis  $E$  eines Zufallsexperiment mit endlicher Ergebnismenge und das diese Ereignisse gleichwahrscheinlich sind gegeben. Die Anzahl der Elemente  $E$  stellt die Anzahl der günstigen Fälle  $N_E$  dar und die Anzahl der Elemente  $\Omega$  die möglichen Fälle  $N_\Omega$ . Für das Eintreten von  $E$  gilt folgende Wahrscheinlichkeit [28, S. 21]:

$$P(E) = \frac{|E|}{|\Omega|} \tag{3}$$

$$P(E) = \frac{N_E}{N_\Omega} \tag{4}$$

Folgende Eigenschaften der Laplace- Wahrscheinlichkeit gelten, wenn eine endliche Ergebnismenge vorliegt [20, S. 140]

1.  $0 \leq P(E) \leq 1 \quad \forall E \subseteq \Omega$ .
2.  $P(\Omega) = 1; P(\emptyset) = 0$ .

3.  $P(\bar{E}) = 1 - P(E) \quad \forall E \subseteq \Omega.$
4.  $P(E_1 \cup E_2) = P(E_1) + P(E_2) \quad E_1, E_2 \subseteq \Omega \quad E_1 \cap E_2 = \emptyset.$
5.  $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) \quad \forall E_1, E_2 \subseteq \Omega.$
6. Speziell gilt  $P(E) = \sum_{\omega \in E} P(\omega) \quad P(\Omega) = \sum_{\omega \in \Omega} = \sum_{i=1}^n P(\omega) = 1.$

## 4 Zufallsvariablen

Unter einer Zufallsvariable wird eine Größe verstanden, deren Wert vom Zufall abhängt. Dabei wird jedem möglichen Ergebnis eines Zufallsexperiment eine Größe zugeordnet [38].

Die Zufallsvariable  $X$  ordnet jedem möglichen Ergebnis eines Zufallsexperiment eine reelle Zahl zu. Hierbei werden vor der Durchführung des Experiments spezielle Ereignisse festgelegt.

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum. Dann ist eine Funktion  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable, falls  $\{\omega \in \Omega : X(\omega) \leq a\} \in \mathcal{A}$  für alle  $a \in \mathbb{R}$  gilt, dann ist die Funktion  $X : \Omega \rightarrow \mathbb{R}$ , folgendermaßen definiert:

$$X : \Omega \rightarrow \mathbb{R}, \omega \mapsto X(\omega). \quad (5)$$

Das Wahrscheinlichkeitsmaß wird nun auch von der Zufallsvariable  $X$  übernommen. Weiters muss diese Funktion  $X$  bestimmte Bedingungen erfüllen und steht im engen Zusammenhang mit der Verteilungsfunktion.

Bei der Verteilungsfunktion gibt es folgende Voraussetzungen. Sei  $X$  eine Zufallsvariable, dann nennt man die Funktion  $F(x)$ , die jedem  $x$  die Wahrscheinlichkeit  $P(X \leq x)$  zuordnet, die Verteilung von  $X$ . Es gilt also:

$$F(x) = P(X \leq x). \quad (6)$$

Die Verteilungsfunktion hat folgende Eigenschaften [9, S. 1]:

1.  $F$  ist monoton steigend in  $\mathbb{R}$ .
2.  $F(b) - F(a) = P(a < X \leq b)$  für alle  $a, b \in \mathbb{R}$  mit  $a \leq b$ .
3.  $P(X > a) = 1 - F(a)$  für alle  $a \in \mathbb{R}$ .
4.  $P(X = x_i) = F(x_i) - F(x_{i-1})$  falls  $X(\Omega) = x_1, x_2, \dots, x_n$  und  $x_1 < x_2 < \dots < x_n$ .

[20, S. 233]

Eine Zufallsvariable  $X$  wird endliche Zufallsvariable, falls sie nur endlich viele Werte annimmt. Mit  $\alpha_j$  werden die angenommenen Werte bezeichnet also  $X(\Omega) = \{\alpha_j : j \in J\}$ , wobei  $J$  eine endliche Menge ist.

Im folgenden werden noch Begriffe wie Erwartungswert, Varianz und Standardabweichung erläutert.

Sei ein Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  und  $X$  eine endliche Zufallsvariable auf  $\Omega$  gegeben, dann lautet der Erwartungswert,

$$E(X) := \sum_{j \in J} P(X = \alpha_j) * \alpha_j. \quad (7)$$

Der Erwartungswert ist der Mittelwert welche die Zufallsvariable annimmt, wenn eine Vielzahl von Realisierungen vorliegen.

Die Varianz  $V(X)$  ist die Abweichung von einer Zufallsvariable vom Erwartungswert  $E(X)$  und ist wie folgt definiert:

$$V(X) := \sum_{j \in J} P(X = \alpha_j) * (\alpha_j - E(X))^2. \quad (8)$$

Die Standardabweichung  $\sigma(X)$  ergibt sich aus der Varianz durch

$$\sigma(X) = \sqrt{V(X)}. \quad (9)$$

## 4.1 Diskrete Zufallsvariablen

Unter einer diskreten Zufallsvariable wird eine Zufallsvariable verstanden, die nur endlich viele oder abzählbar viele Werte annehmen kann. Dies ist mit einem Zähl-experiment vergleichbar. Die Verteilungsfunktion wird durch das Addieren der einzelnen Wahrscheinlichkeiten gebildet und kann in Form einer Treppenfunktion abgebildet werden.

Eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  wird Treppenfunktion genannt, wenn es endlich viele Stellen gibt mit  $a = x_0 < x_1 < x_2 < \dots < x_n = b$ , sodass  $f$  auf allen offenen Teilintervallen  $]x_{k-1}, x_k[$  dazwischen konstant ist. Eine solche Punktmenge

$$Z = (x_0, x_1, \dots, x_n) \quad \text{mit} \quad a = x_0 < x_1 < x_2 < \dots < x_n = b \quad (10)$$

wird als Zerlegung des Intervalls  $[a, b]$  bezeichnet. Die Teilpunkte werden als  $x_k$  und die Teilintervalle als  $]x_{k-1}, x_k[$  von  $Z$  definiert. Ist das Intervall  $[a, b]$  einpunktig, also gleich dem Intervall  $[a, a]$ , so ist nur die leere Zerlegung möglich. Jede Funktion auf  $[a, a]$  mit reellen Werten gilt als Treppenfunktion.

Neben dem Münzwurf und dem Würfel ist das Schießen auf eine Scheibe ein weiteres Beispiel für eine diskrete Zufallsvariable. Dabei schießt ein Schütze auf eine



Ringscheibe mit zehn gleichgroßen Ringen, so können als mögliche Trefferzahlen die Zahlen  $1, 2, 3, \dots, 10$  angenommen werden. Dabei wird von einer diskreten Zufallsvariable gesprochen, da jeder dieser Werte die gleiche Wahrscheinlichkeit hat mit,  $P_1 = P_2 = P_3 = P_4 = P_5 = P_6 = P_7 = P_8 = P_9 = P_{10} = \frac{1}{10}$  annimmt. Aber nur wenn die Flächen der Ringe gleich groß sind und alle Schüsse die Scheibe treffen [21, S. 225].

#### 4.1.1 Wahrscheinlichkeitsverteilungen diskreter Zufallsvariablen

Gegeben ist eine diskrete Zufallsvariable  $X$ , welche die Werte  $x_1, x_2, \dots$  mit von Null verschiedenen Wahrscheinlichkeiten  $P(X = x_i) = P(x_i)$  annehmen kann. Die Funktion  $f(x_i) = P(X = x_i)$ , die jedem  $x_i$  die Wahrscheinlichkeit  $f(x_i)$  zuordnet, heißt Wahrscheinlichkeitsfunktion [28, S. 47].

Mit dieser Definition sind zwei Eigenschaften verbunden:

1.  $0 \leq f(x_i) \leq 1$ .
2.  $\sum_i f(x_i) = 1$ .

Sollte die Zufallsvariable nur endlich viele Werte annehmen, so kann die Wahrscheinlichkeitsfunktion mit Hilfe einer Tabelle [28, S. 47] dargestellt werden:

$x_i$	$x_1$	$x_2$	$\dots$	$x_n$
$f(x_i)$	$f(x_1)$	$f(x_2)$	$\dots$	$f(x_n)$

Die Verteilungsfunktion  $F$  gibt auch hier wieder eine Wahrscheinlichkeit an, dass die Zufallsvariable  $X$  höchstens den Wert  $x$  annimmt. Dafür ist ein Wahrscheinlichkeitsraum  $(\Omega, P)$  und eine Zufallsvariable  $X$  gegeben. Dann wird diese Funktion wie folgt definiert [20, S. 232]:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i). \quad (11)$$

[28, S. 49]

Auch hierbei sind folgende Eigenschaften zu beachten [28, S. 49]:

1.  $F$  ist monoton steigend in  $\mathbb{R}$ .  
Da die Verteilungsfunktion durch Addieren der Wahrscheinlichkeiten entsteht, und diese niemals negativ sind, muss die Funktion monoton steigend sein.

2.  $F$  ist rechtsseitig stetig.

Rechtsseitig stetig heißt, dass die Funktion an einer Stelle  $x$  einen Sprung hat, sodass sie den darüber liegenden Wert, als den Wert der oberen Sprunggrenze annimmt. Dies ergibt sich aus der Definition  $P(X \leq x)$ , denn würde dies durch  $P(X < x)$  ersetzt werden, dann wäre die Verteilungsfunktion linksseitig stetig.

3.  $\lim_{x \rightarrow -\infty} F(x) = 0$ .

Hierbei geht die Verteilungsfunktion gegen Null, wenn  $x$  immer kleiner wird. Im Gegensatz dazu steht die Eigenschaft vier.

4.  $\lim_{x \rightarrow \infty} F(x) = 1$

Der oben genannte Münzwurf wird der Gleichverteilung zugeordnet und gehört zu den diskreten Wahrscheinlichkeitsfunktionen. Wie schon bei den anderen Funktionen ist auch hier bei der Wahrscheinlichkeitsfunktion, siehe Abbildung 1, eine Zufallsvariable  $X$ , die die Werte  $x_i$  annimmt gegeben. Dafür gilt:

$$f(x_i) = \frac{1}{n} \quad i = 1, 2, \dots, n. \quad (12)$$

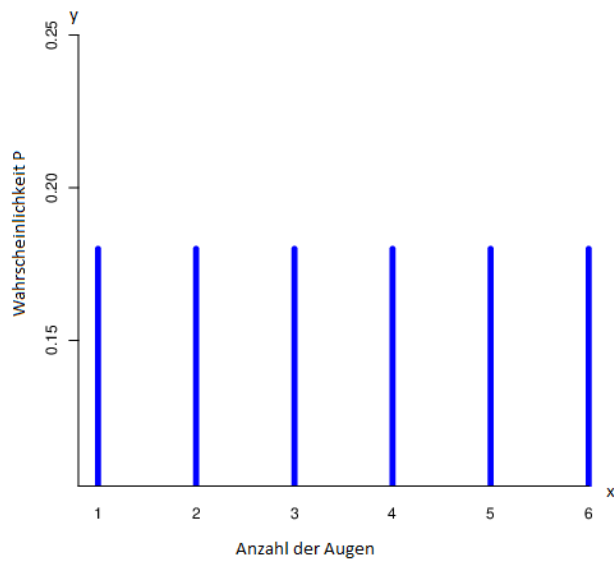
Bei der Verteilungsfunktion, siehe Abbildung 2, werden drei verschiedene Fälle unterschieden [28, S. 77] :

$$F(x) = \begin{cases} 0 & \text{für } x < x_1. \\ \frac{i}{n} & \text{für } x_i \leq x < x_{i+1} \quad \text{für } i = 1, \dots, n-1. \\ 1 & \text{für } x_n \leq x. \end{cases} \quad (13)$$

Das einfachste Beispiel der diskreten Gleichverteilung ist das Würfeln mit einem Würfel.

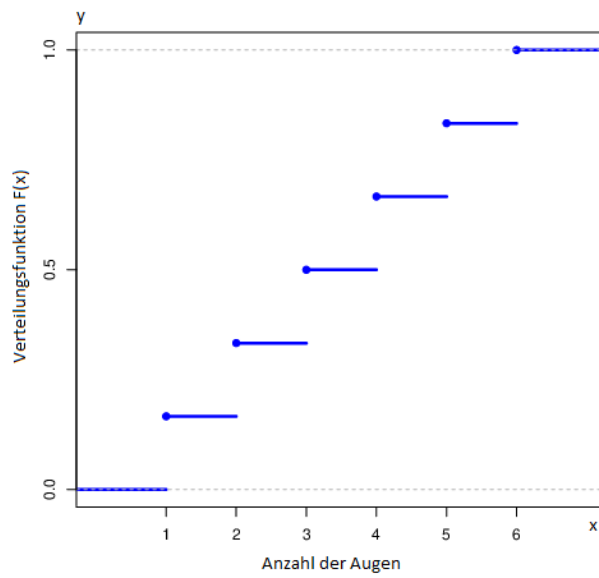
$$f(x_i) = \frac{1}{6} \quad \text{mit } i = 1, 2, \dots, 6. \quad (14)$$

Abbildung 1 zeigt die Wahrscheinlichkeitsfunktion und Abbildung 2 die Verteilungsfunktion.



[46]

Abbildung 1: Wahrscheinlichkeitsfunktion einer diskreten Gleichverteilung mit sechs möglichen Ausgängen (z.B. Würfelwurf)



[46]

Abbildung 2: Verteilungsfunktion einer diskreten Gleichverteilung am Beispiel des Würfelwurfs

## 4.2 Stetige Zufallsvariablen

Unter einer stetigen Zufallsvariable versteht man eine Zufallsvariable die nur überabzählbar viele Werte annehmen kann.

### 4.2.1 Wahrscheinlichkeitsverteilungen stetiger Zufallsvariablen

Im Gegensatz zur Wahrscheinlichkeitsfunktion einer diskreten Zufallsvariable wird bei stetigen Zufallsvariablen von einer Dichtefunktion gesprochen.

Eine Zufallsvariable  $X$  heißt stetige Zufallsvariable, falls es eine integrierbare Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  gibt, sodass für jedes abgeschlossene Intervall  $[a, b]$  Folgendes gilt [20, S. 283]:

$$P(a \leq X \leq b) = \int_a^b f(x)dx. \quad (15)$$

Eigenschaften dieser Funktion sind:

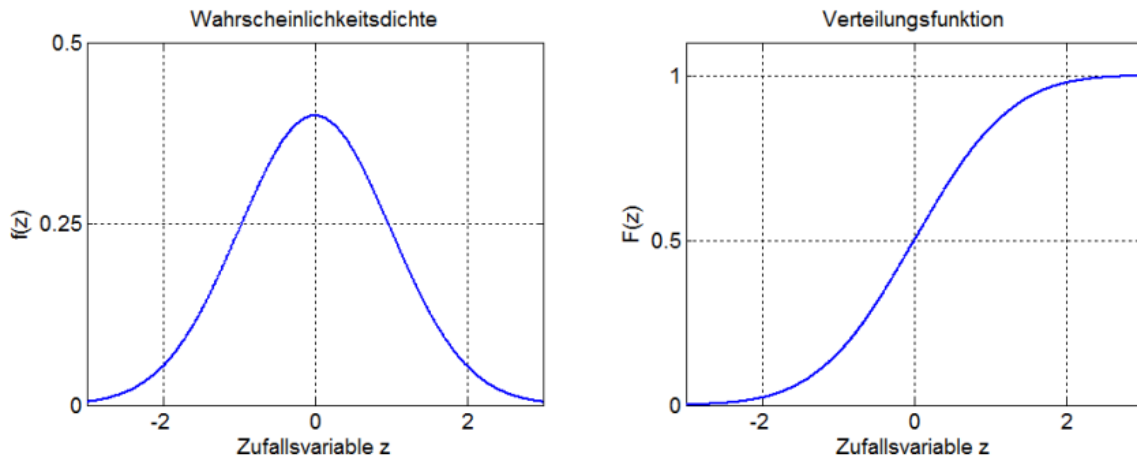
1.  $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$ .
2.  $P(X = x) = 0$ , das bedeutet, dass  $f(x)$  keine Wahrscheinlichkeit ist und dass die Dichtefunktion auch Werte von  $f(x) > 1$  annehmen kann.
3.  $f(x) \geq 0$ , bedeutet die Nichtnegativität der Funktion.
4. Normierungseigenschaften:  $\int_{-\infty}^{\infty} f(x)dx = 1$ , das bedeutet, dass die Gesamtfläche zwischen der x-Achse und der Dichtefunktion  $f(x)$  gleich 1 ist.

Die Dichtefunktion gibt an, mit welcher Wahrscheinlichkeit die Zufallsvariable  $X$  einen Wert in einem gegebenen Intervall annimmt. Dies wird mit Hilfe von einer Fläche mit den jeweiligen Integrationsgrenzen markiert.

Auch bei der stetigen Zufallsvariable entsteht eine Verteilungsfunktion. In dem Fall werden die Wahrscheinlichkeiten integriert und werden in Form einer monoton steigender Funktion abgebildet [28, S. 54]:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(\xi)d\xi \quad (16)$$

Die Eigenschaften dieser Verteilungsfunktion stimmen mit der Verteilungsfunktion der diskreten Zufallsvariable überein. In Abbildung 3 kann der Unterschied zwischen Verteilungsfunktion und Dichtefunktion erkannt werden.



[47]

Abbildung 3: Dichtefunktion  $f$  vs. Verteilungsfunktion  $F$  einer standardisierten Normalverteilung  $\mathcal{N}(0; 1)$  mit dem Mittelwert Null und der Standardabweichung Eins.

Rechnerisch kann die Dichtefunktion berechnet werden, indem die erste Ableitung der Verteilungsfunktion gebildet wird. Somit ist die Verteilungsfunktion eine Stammfunktion der Dichtefunktion. Mathematisch wird dieser Zusammenhang wie folgt dargestellt [28, S. 56]:

$$P(a < X \leq b) = \int_a^b f(x)dx = F(b) - F(a). \quad (17)$$

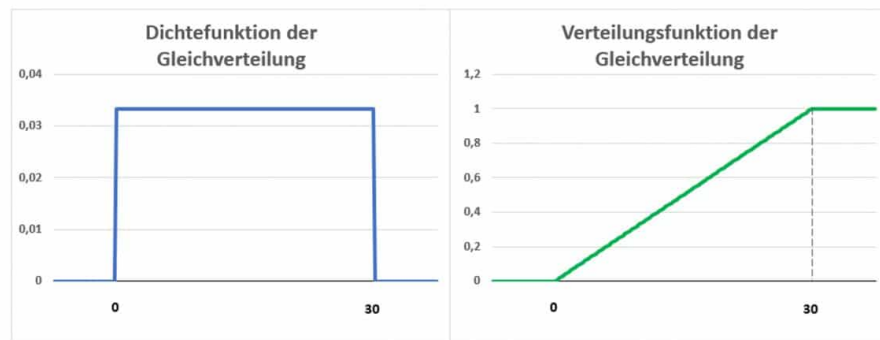
Nun zur Gleichverteilung, diese ist durch die Dichtefunktion mit  $a < b$  definiert:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a < x < b. \\ 0 & \text{sonst.} \end{cases} \quad (18)$$

Um von der Dichtefunktion die Verteilungsfunktion zu erhalten, ist folgender Rechenweg notwendig:

$$F(x) = \int_{-\infty}^x f(t) \, dt = \int_a^x \frac{1}{b-a} \, dt = \left| \frac{t}{b-a} \right|_a^x = \frac{x}{b-a} - \frac{a}{b-a} = \frac{x-a}{b-a}. \quad (19)$$

Diese Verteilung schwankt nur zwischen den zwei Werten  $a$  und  $b$ , siehe Abbildung 4, da sie alle Werte innerhalb dieses Intervalls gleich wahrscheinlich annimmt.



[48]

Abbildung 4: Dichte- und Verteilungsfunktion von einer stetigen Gleichverteilung im Intervall  $[0; 30]$

Die Verteilungsfunktion lässt sich in drei verschiedenen Fällen unterscheiden [12, S. 29]:

$$F(x) = \begin{cases} 0 & \text{für } x < a. \\ \frac{x-a}{b-a} & \text{für } a \leq x < b. \\ 1 & \text{für } b \leq x. \end{cases} \quad (20)$$

Der Erwartungswert von  $X$  und die Varianz von  $X$  können wie folgt hergeleitet werden:

$$E(X) = \int_a^b \frac{1}{b-a} x \, dx = \frac{1}{b-a} \left| \frac{1}{2} x^2 \right|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2} \quad (21)$$

Ebenso die Varianz mit  $V(X) = \frac{(b-a)^2}{12}$  kann mit Hilfe der Dichtefunktion hergeleitet werden:

$$V(X) = \int_a^b \frac{1}{b-a} x^2 \, dx - \left(\frac{b+a}{2}\right)^2 = \frac{1}{b-a} \left[ \frac{1}{3} x^3 \right]_a^b - \left(\frac{b+a}{2}\right)^2 \quad (22)$$

$$= \frac{b^3 - a^3}{3(b-a)} - \frac{(b+a)^2}{4} = \frac{(b-a)^2}{12} \quad (23)$$

## 5 Ausgewählte Wahrscheinlichkeitsverteilungen

Im folgenden Kapitel werden die wichtigsten stetigen Wahrscheinlichkeitsverteilungen beschrieben. Diese sind notwendig um empirische Daten möglichst gut beschreiben zu können. Anbei werden die Gleichverteilung, Normalverteilung und die logarithmische Normalverteilung genauer erklärt.

### 5.1 Normalverteilung

Die Normalverteilung oder auch Gauß'sche Verteilung genannt, ist eine der am häufigst vorkommensten Wahrscheinlichkeitsverteilung. Der Name stammt von Carl Friedrich Gauß auf den die Normalverteilung zurück geht, mit den beiden Parametern  $\mu$  und  $\sigma$ . Dabei ist  $\mu$  der Erwartungswert und  $\sigma$  die Standardabweichung [8]. In der folgenden wird die Dichte der Normalverteilung definiert:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{(x - \mu)^2}{2\sigma^2} \quad (24)$$

Solch eine Verteilung wird oft als  $\mathcal{N}(\mu, \sigma)$  dargestellt. Der Erwartungswert einer normalverteilten Zufallsvariable  $X$  ist definiert als  $E(X) = \mu$  und die Varianz  $V(X) = \sigma^2$ . Diese beiden Werte können direkt aus der Definition abgelesen werden. Wie schon vorhin erwähnt, hat die Dichtefunktion der Normalverteilung eine Gestalt wie eine Glocke, siehe Abbildung 5.

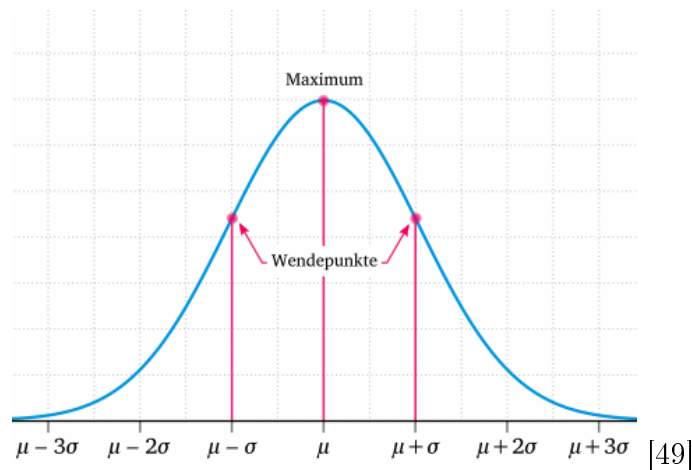


Abbildung 5: Dichtefunktion einer Normalverteilung



Graphisch lässt sich auch sehr gut die Veränderung von  $\sigma$  bei gleichbleibenden  $\mu$  darstellen. Wird der Wert der Standardabweichung kleiner, so wird der Graph um den Erwartungswert  $\mu$  steiler, siehe Abbildung 6.

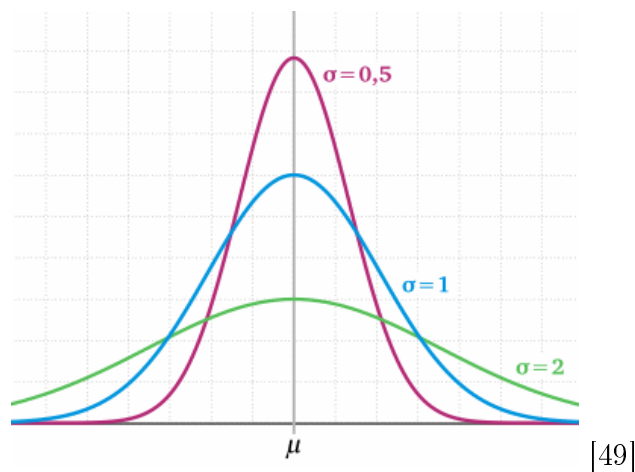


Abbildung 6: Dichtefunktion einer Normalverteilung  $\mathcal{N}(\mu; \sigma)$  mit  $\sigma = 1$ ,  $\sigma = 2$  und  $\sigma = 3$ .

Umgekehrt ist die Veränderung vom Erwartungswert  $E$  bei gleichbleibenden  $\sigma$ . Dabei verschiebt sich der Graph um  $\mu$  Einheiten nach links beziehungsweise nach rechts, siehe Abbildung 7 [28, S. 94ff.].

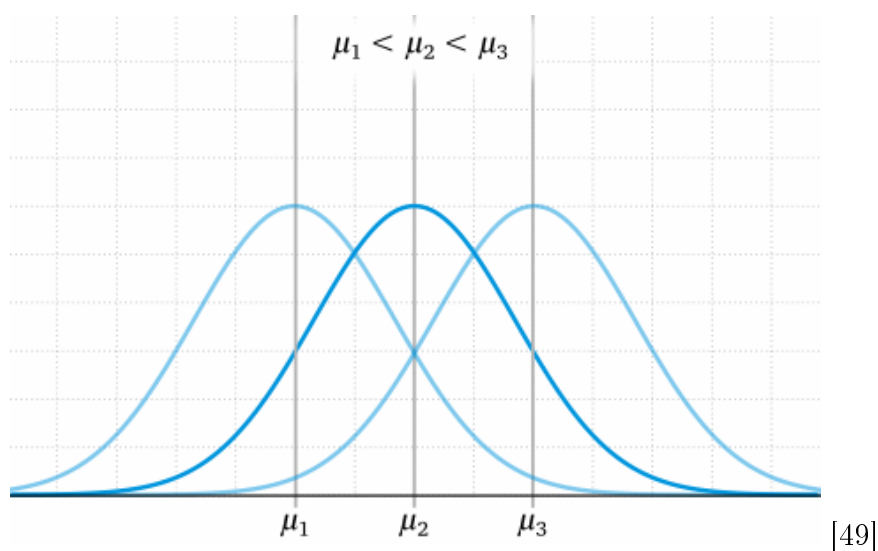


Abbildung 7: Dichtefunktion einer Normalverteilung  $\mathcal{N}(\mu; \sigma)$  mit konstanter Standardabweichung und drei Mittelwerten  $\mu_1 < \mu_2 < \mu_3$

Mittelwert    =Median    =Modus    = $\mu$   
 Varianz        = $\sigma^2$

Auf Grund der beiden Grafiken lässt sich schon erkennen, dass die Normalverteilung achsensymmetrisch bezüglich des Erwartungswertes  $\mu$  ist. Die Verteilung hat zwei Wendepunkte bei  $x = \mu \pm \sigma$ . Bei jeder Normalverteilung kann das Maximum bei  $\mu$  abgelesen werden. Eine weitere Eigenschaft die sich anhand der Grafik erkennen lässt ist, dass sich sobald die x- Werte gegen  $+\infty$  oder  $-\infty$  streben, sich asymptotisch der x- Achse nähern.

Sobald der Erwartungswert den Wert Null und die Standardabweichung den Wert Eins annimmt, wird von der Standardnormalverteilung  $\mathcal{N}(0; 1)$  gesprochen. Es sind  $\mu, \sigma \in \mathbb{R}$  definiert, wobei  $\sigma > 0$  sein muss. Dann heißt eine stetige Zufallsvariable  $X$  normalverteilt mit den Parametern  $\mu$  und  $\sigma$ , falls  $\forall -\infty \leq a < b \leq \infty$  gilt [12, S. 25]:

$$P(a \leq X \leq b) = \int_a^b f(t) \, dt = \int_a^b \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(t-\mu)^2}{2 \cdot \sigma^2}} dt \quad (25)$$

Die Dichtefunktion dieser Verteilung kann auch hier wieder direkt aus der Gleichung abgelesen werden, denn wenn die Verteilung  $F$  bekannt ist, wird die Dichte mit Hilfe der Integration  $f = F'$  berechnet.  $F$  wird hier als die Stammfunktion bezeichnet. Daraus lässt sich die Verteilungsfunktion ableiten:

$$F(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \int_{-\infty}^x e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}. \quad (26)$$

Jede Verteilung hat besondere Kenngrößen. Diese werden im Folgenden aufgelistet [32, S. 83]:

Wie bei Abbildung 8 zu sehen ist, gibt es bei der Normalverteilung bestimmte Intervalle in die die Werte hineinfallen. Ein Drittel (circa 34, 13%) werden im Bereich zwischen Mittelwert  $\mu$  und  $\mu + \sigma$  erwartet. Wird auch die negative Seite bezüglich  $\mu$  betrachtet, führt dies zur Verdoppelung des Wahrscheinlichkeitswertes also 68, 27%. Daher entspricht das Intervall  $(\mu \pm \sigma)$  circa zwei Drittel der Zufallsvariablen. Für  $(\mu \pm 2\sigma)$  beträgt die Wahrscheinlichkeit, dass ein Merkmal in diesen Bereich fällt 95, 45%. Das letzte Intervall  $(\mu \pm 3\sigma)$  wird mit 99, 73% beschrieben [32, S. 85].

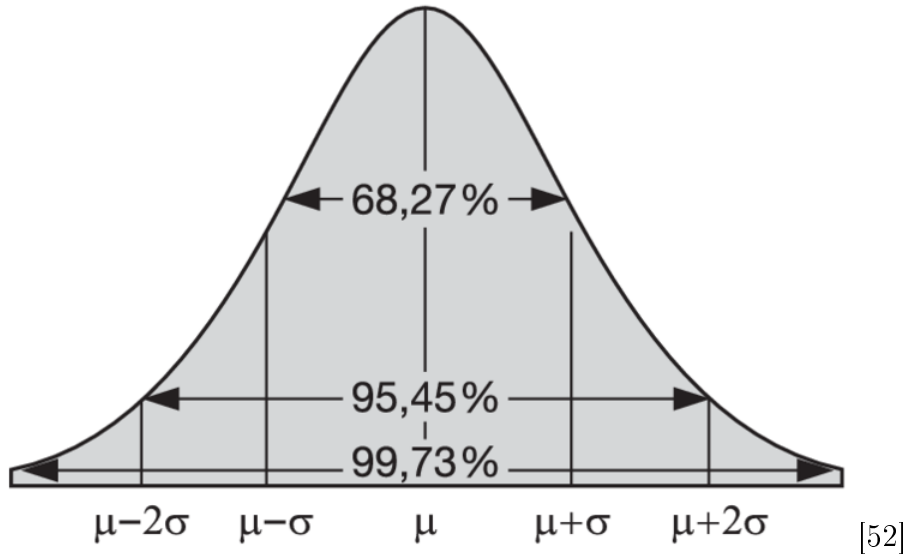


Abbildung 8: Dichteverteilung einer Normalverteilung  $\mathcal{N}(\mu; \sigma)$  und dem relativen Anteil der Zufallsvariable innerhalb der drei Intervalle  $(\mu \pm \sigma)$ ,  $(\mu \pm 2\sigma)$  und  $(\mu \pm 3\sigma)$

## 5.2 Logarithmische Normalverteilung

Die logarithmische Normalverteilung spiegelt sich bei einigen alltäglichen Vorgängen wieder. Denn die Merkmale dieser Verteilung haben einen unteren Grenzwert im beobachteten Wertebereich. Beispiele sind der Niederschlag, Windgeschwindigkeit, Erdbeben oder das auftreten von Gewittern. Treffen die oben genannten Merkmale zu, muss mit einer positiven schiefen Verteilung gerechnet werden, der logarithmischen Normalverteilung.

Die Dichtefunktion sind wie folgt definiert:

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \cdot \frac{1}{x} \cdot \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right) & x > 0. \\ 0 & x \leq 0. \end{cases} \quad (27)$$

Die dazu passende Verteilungsfunktion lässt sich hier wieder mit dem Integral berechnen. Auch hier werden zwei verschiedene Fälle unterschieden:

$$F(x) = \begin{cases} 0 & x \leq 0. \\ \frac{1}{\sqrt{2\pi}\sigma} \int_0^x \frac{1}{t} e^{-\frac{(\ln(t)-\mu)^2}{2\sigma^2}} dt & x > 0. \end{cases} \quad (28)$$

Erwartungswert und Varianz lassen sich aus der Verteilungsfunktion herleiten und es gilt:  $E(X) = \exp^{\mu + \frac{\sigma^2}{2}}$ . Mit Hilfe des Erwartungswertes lässt sich die Varianz herleiten:  $V(X) = \exp^{2\mu + \sigma^2} \cdot (\exp^{\sigma^2} - 1)$ .

Die Kenngrößen der logarithmischen Normalverteilung werden im folgenden noch einmal zusammengefasst:

Mittelwert	$\mu = e^{\mu + \frac{\sigma^2}{2}}$
Median	$\mu_+ = e^\mu$
Modus	$\mu_\wedge = e^{\mu - \sigma^2}$
Varianz	$\sigma^2 = e^{2\mu + \sigma^2} \cdot (e^{\sigma^2} - 1)$

An dem unten angeführten Beispiel, siehe Abbildung 9, kann der enge Zusammenhang zwischen der Normal- und logarithmische Normalverteilung dargestellt werden. Diese Funktion ist positiv schief. Wenn die Zufallsvariable logarithmisch transformiert wird, dann ergibt sich daraus eine Normalverteilung [?, S. 27].

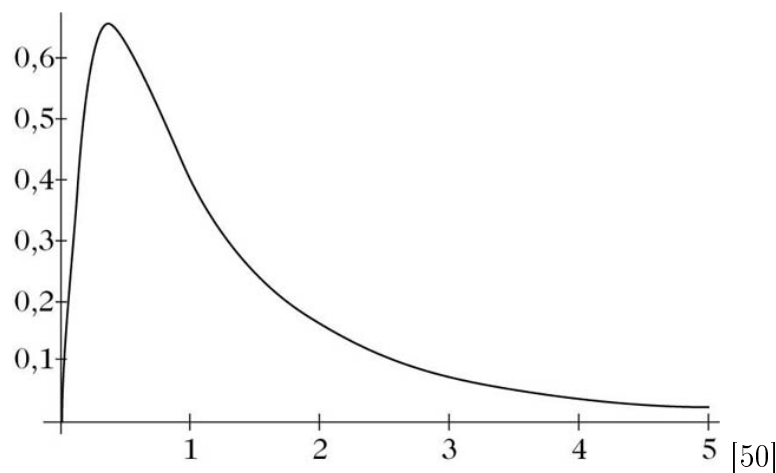


Abbildung 9: Dichteverteilung einer logarithmischen Normalverteilung  $\mathcal{L}(0; 1)$ .

### 5.3 Binomialverteilung

Die Binomialverteilung basiert auf dem Bernoulli- Experiment. Dabei wird ein Experiment mit zwei möglichen Ereignissen  $A$  und  $\bar{A}$ , wobei die Wahrscheinlichkeiten  $P(A) = p$  und  $P(\bar{A}) = 1 - P(A) = q$  gelten betrachtet. Wenn dieses Zufallsexperiment  $n$ -mal wiederholt wird, wird von einem Bernoulli- Experiment gesprochen. Wenn  $A$  aber genau  $x$ -mal auftritt, ohne Berücksichtigung auf die Anordnung und ohne Wiederholung, dann ist das ein Problem der Kombinatorik [28, S. 81].

Dieses Beispiel ist ein Grundaufgabe aus der Kombinatorik  $k$  Kugeln auf  $n$  Plätze zu verteilen. Dies kann wie folgt geschrieben werden

$$\binom{n}{k}. \quad (29)$$

Dabei wird  $k$ -mal das Ereignis  $A$  bei  $n$  Versuchen eintreten. Wird dies nun auf das Ereignis genau  $k$ -mal Eins gespiegelt, tritt mit Hilfe der Pfad-Multiplikation folgende Formel zum Vorschein:

$$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}. \quad (30)$$

Das Ergebnis dieser Formel ist die Wahrscheinlichkeitsfunktion der Binomialverteilung  $B(n, p)$ . Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $X$  eine Zufallsvariable, die die Werte  $0, 1, \dots, n$  annimmt. Die Dichtfunktion  $f(k)$  mit  $X$  heißt binomialverteilt mit den Parametern  $n$  und  $p$ , wenn es ein  $p \in [0; 1]$  gibt sodass,

$$P_X(k) = P(X = k) = f(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad \text{für } k = 0, 1, \dots, n. \quad (31)$$

Die Verteilungsfunktion lautet [32, S. 77]:

$$F(k) = \sum_{i=1}^k \binom{n}{i} \cdot p^i \cdot q^{n-i}. \quad (32)$$

Auch hier gibt es wieder die beiden Kerngrößen Erwartungswert und Standardabweichung. Beide können mit Hilfe des Bernoulli- Experimentes dargestellt werden.

Dabei werden mehrerer dieser Experimente additiv verknüpft, sodass  $X = \sum_{i=1}^n X_i$  entspricht. Daher gilt  $P(X_i = 1) = p$  und  $P(X_i = 0) = 1 - p = q$ . Daraus lässt sich dann der Erwartungswert  $E(X_i) = 0 \cdot (1 - p) + 1 \cdot p = p$  mit Hilfe folgenden Satzes,

$$P(X = x) = P(x) = \sum_{\omega \in \Omega} P(\omega) \quad (33)$$

und

$$\begin{aligned}
E(X + Y) &= \sum_{\omega \in \Omega} P(\omega) \cdot (X(\omega) + Y(\omega)) = \\
&= \sum_{\omega \in \Omega} P(\omega) \cdot X(\omega) + \sum_{\omega \in \Omega} P(\omega) \cdot Y(\omega) = \\
&= E(X) + E(Y).
\end{aligned}$$

mit

$$E(X) = \sum_{i=1}^n E(X_i) = n \cdot p \quad (34)$$

herleiten.

Die Varianz kann mit dem selben Schema hergeleitet werden. Auch hier gilt folgendes  $V(X_i) = (0 - p)^2 \cdot (1 - p) + (1 - p)^2 \cdot p = p \cdot (1 - p)$ , sowie folgende Sätze sind gegeben:

$$\begin{aligned}
V(X + Y) &= E((X + Y)^2) - (E(X + Y))^2 = \\
&= E(X^2 + 2XY + Y^2) - (E(X) + E(Y))^2 = \\
&= E(X^2) - E(X)^2 + E(Y^2) - E(Y)^2 + 2E(XY) - 2E(X) \cdot E(Y) = \\
&= V(X) + V(Y) + 2(E(XY) - E(X) \cdot E(Y))
\end{aligned}$$

und

$$\begin{aligned}
E(X \cdot Y) &= \sum_{z \in XY(\Omega)} P(XY = z) \cdot z = \\
&= \sum_{z \in XY(\Omega)} \left( \sum_{x \in X(\Omega), y \in Y(\Omega), x \cdot y = z} P(X = x \wedge Y = y) \right) \cdot z = \\
&= \sum_{x \in X(\Omega), y \in Y(\Omega)} x \cdot y \cdot P(X = x \wedge Y = y) = \\
&= \sum_{x \in X(\Omega), y \in Y(\Omega)} x \cdot y \cdot P(X = x) \cdot P(Y = y) = \\
&= \sum_{x \in X(\Omega)} x \cdot P(X = x) \cdot \sum_{y \in Y(\Omega)} y \cdot P(Y = y) = \\
&= E(X) \cdot E(Y)
\end{aligned}$$

Aus diesen beiden Sätzen folgt direkt der hier wesentliche Satz,

$$V(X + Y) = V(X) + V(Y), \quad (35)$$

da auf Grund der Unabhängigkeit der Term verschwindet. Daraus folgt dann die Varianz der Binomialverteilung [20, S. 247ff.],

$$V(X) = \sum_{i=1}^n V(X_i) = n \cdot p \cdot (1 - p). \quad (36)$$

Weitere Kenngrößen sind [32, S. 79]:

Mittelwert	$\mu = n \cdot p$
Median	$\mu_+ = \begin{cases} n \cdot p, & \text{falls } n \cdot p \text{ ganzzahlig,} \\ \text{andernfalls je nach Assymetrie der } n \cdot p \text{ nächstliegende Zahl.} \end{cases}$
Modus	$\mu_\wedge = \begin{cases} (n + 1) \cdot p, & \text{falls } (n + 1) \cdot p \text{ ganzzahlig,} \\ \text{andernfalls größte ganze Zahl mit } z < (n + 1) \cdot p \text{ nächstliegende Zahl.} \end{cases}$
Varianz	$\sigma^2 = n \cdot p \cdot q$

Zuletzt noch eine wichtige Eigenschaft der Binomialverteilung, die Reproduktivität. Dabei sind zwei unabhängige Zufallsvariablen  $X$  und  $Y$  gegeben, für die gilt  $X + Y$  ist binomialverteilt mit den Parametern  $n + m$  und  $p$ . Dabei ist zu beachten, dass die beiden ursprünglichen Binomialverteilungen dieselbe Wahrscheinlichkeit haben müssen.

Es gibt für die Approximation der Binomialverteilung zwei verschiedene Möglichkeiten. Dies hängt einerseits von der Anzahl der Wiederholungen des Zufallsexperiment und andererseits von der Wahrscheinlichkeit  $p$ , dass ein Ereignis eintritt ab.

Die folgenden Anwendungen sind Faustregeln, da es sich um keine mathematische Begründungen handelt. Sei  $n \cdot p \leq 10$  und  $n \geq 1500 \cdot p$  eine  $B(n; p)$ - verteilte Zufallsvariable, dann wird von einer näherungsweise Poissonverteilung mit dem Parameter  $\mu = np$  gesprochen. Dazu gegenüber steht die Approximation bezüglich der Normalverteilung. Dafür gilt,  $n \cdot p \cdot q > 9$ , damit eine  $B(n; p)$ - verteilte Zufallsvariable mit den Parametern  $\mu = np$  und  $\sigma^2 = npq$  näherungsweise normalverteilt ist [28, S. 84].

## 5.4 Poissonverteilung

Die letzte Verteilung lässt sich auf das vorhin erwähnte Bernoulli- Experiment zurückführen. Jedoch wird hier bei der Poissonverteilung eine sehr kleine Erfolgswahrscheinlichkeit angenommen. Das heißt die Wahrscheinlichkeit  $p$  für das Eintreten des Ereignisses  $A$  ist sehr klein und die Anzahl  $n$  der Ausführungen des Experiments sehr groß.

Die Poissonverteilung geht auf Simeon Denis Poisson zurück. 1837 hat er diese in seinem Werk hergeleitet, da er bemerkt hat, dass die Konvergenz der Binomialverteilung gegen die Normalverteilung sehr schlecht ist. vor allem wenn die Erfolgswahrscheinlichkeit und die Gegenwahrscheinlichkeit sehr unterschiedlich sind.

Wird nun von dem Bernoulli- Experiment ausgegangen und  $p \rightarrow 0$  und  $n \rightarrow \infty$  gilt, sodass der Erwartungswert  $E(X) = \mu np$  konstant bleibt, gilt folgendes:

Eine Zufallsvariable  $X$  mit den Werten  $\mathbb{N}_0$  heißt Poissonverteilt mit dem Parameter  $\lambda > 0$ , wenn gilt:

$$P(X = x) = f(x) = \frac{e^{-\lambda}}{x!} \cdot \lambda^x. \quad (37)$$

Die Verteilungsfunktion  $F(k)$  ist wie folgt definiert:

$$F(x) = e^{-\lambda} \cdot \sum_{i=0}^x \frac{\lambda^i}{i!} \quad (38)$$

Für den Erwartungswert und die Varianz erhält man:

$$E(X) = \lambda \quad \text{und} \quad V(X) = \lambda. \quad (39)$$

Wie schon bei den vorigen Verteilungen, werden auch hier die wichtigsten Kerngrößen zusammengefasst [32, S. 81]:

Mittelwert	$\mu = \lambda$
Median	$\mu_+ = \begin{cases} \frac{1}{\sqrt{\lambda}} & \text{falls } \lambda \text{ ganzzahlig, andernfalls je nach Schiefe} \\ \text{die } \lambda \text{ nächstliegende ganze Zahl.} \end{cases}$
Modus	$\mu_\wedge = \begin{cases} \lambda \quad \text{und} \quad (\lambda - 1) & \text{falls } \lambda \text{ ganzzahlig} \\ \text{andernfalls größte ganze Zahl mit } z < \lambda. \end{cases}$
Varianz	$\sigma^2 = \lambda$

Damit nachvollzogen werden kann was es bedeutet, dass  $p \rightarrow 0$  und  $n \rightarrow \infty$  gehen, einige Beispiele:



1. Anzahl von Druckfehlern pro Seite.
2. Schadenshäufigkeit bei Versicherungen.
3. Eintreten eines Impfschadens.

[20, S. 261ff.]

## 5.5 Anpassung von Zufallsvariablen an Verteilungsfunktionen

In der Realität entspricht die Verteilung einer realen Zufallsvariable oftmals nicht den oben angeführten theoretischen Verteilungen. Daher werden diese an theoretische Verteilungen angepasst. Diese unterschiedlichen Transformationen werden im Folgenden erklärt.

Mit Hilfe der linearen Transformation einer Normalverteilung kann jede beliebige  $\mathcal{N}(\mu; \sigma)$ - verteilte Zufallsvariable in eine  $\mathcal{N}(0; 1)$ - verteilte Zufallsvariable  $Y$  transformiert werden.

Sei  $X$  eine  $\mathcal{N}(\mu; \sigma)$ - verteilte Zufallsvariable, dann gilt

$$Y = aX + b, \quad a, b \in \mathbb{R}, \quad \text{eine } \mathcal{N}(a\mu + b; |a|\sigma) - \text{verteilte Zufallsvariable} \quad (40)$$

Werden  $a$  und  $b$  geeignet gewählt, dann gilt für den Erwartungswert  $E$  und die Varianz  $V$ :

$$E(Z) = 0 = a\mu + b \text{ und } V(Z) = 1 = a^2\sigma^2. \text{ Daraus lässt sich folgern: } a = \frac{1}{\sigma} \text{ und } b = -\frac{\mu}{\sigma}.$$

Mit Hilfe dieser Informationen kann nun eine  $\mathcal{N}(\mu; \sigma)$  -verteilte Zufallsvariable standardisiert werden:

$$Z = \frac{1}{\sigma}X - \frac{\mu}{\sigma} = \frac{X - \mu}{\sigma} \quad (41)$$

und hat somit eine  $\mathcal{N}(0; 1)$ - verteilte Zufallsvariable [28, S. 96].

Eine weitere Transformation ist die Box-Cox Transformation, welche von Box und Cox 1964 [35, S. 46] entwickelt wurde. Da viele statistische Verfahren normalverteilte Daten voraussetzen, gibt es eine weitere Möglichkeit, diese Daten zu einer Normalverteilung zu transformieren. Bei der vorhin erwähnten lineare Transformation, wird lediglich nur ein Faktor addiert beziehungsweise multipliziert. Dabei

wird aber das Aussehen der Verteilung nicht verändert. Bei der Box-Cox Transformation passt sich die Verteilung auch graphisch an. Eigentlich ist es nur ein Nebeneffekt, dass Daten normalverteilt werden. Denn eigentlich sucht die Box-Cox Transformation nach der kleinsten Standardabweichung [33].

Sei  $X$  eine Zufallsvariable mit  $X = x_0, x_1, \dots, x_{n-1}$  und  $i = 1, 2, \dots, N - 1$  dann gilt,

$$x_i(\lambda) = \begin{cases} \frac{x_i^\lambda - 1}{\lambda} & \text{für } \lambda \neq 0. \\ \log(x_i) & \text{für } \lambda = 0. \end{cases} \quad (42)$$

Diese Transformation besitzt nur einen Parameter,  $\lambda$ . Wenn  $\lambda$  gleich Null ist, erfolgt die logarithmische Umwandlung der Eingangsfolge, bei jedem anderen Wert erfolgt dies schrittweise. Ist  $\lambda$  Eins, so bleibt die Verteilungsgesetzmäßigkeit der Ausgangsfolge unverändert [34].

Die letzte Transformation wird als  $\log(x + C)$ - Transformation beschrieben. Wie schon erwähnt, wird die logarithmische Transformation dafür verwendet, um die Schiefe und die Wölbung von der original Verteilung zu reduzieren. Jedoch nimmt diese nicht die Form der Gauß'schen Glocke an. Zur weiteren Verbesserung wird eine Konstante  $C$  hinzugefügt. Wenn jedoch eine positive Schiefe nach der logarithmischen Transformation bleibt, wird die Konstante  $C$  negativ gewählt. Im Gegensatz dazu ist  $C$  positiv, wenn die Schiefe negativ ist.

Diese Transformation ist eine Verallgemeinerung der Box-Cox- Transformation. Dabei gibt es statt einem Parameter nun zwei:

$$x_i(\lambda, C) = \begin{cases} \frac{(x_i + C)^\lambda - 1}{\lambda} & \text{für } \lambda \neq 0. \\ \log(x_i + C) & \text{für } \lambda = 0. \end{cases} \quad (43)$$

### 5.5.1 Addiert oder multipliziert die Natur?

Im Folgenden wird das Problem analysiert, wann die Normalverteilung und wann die logarithmische Normalverteilung angewendet werden soll. Hierbei geht es um das Problem, ob Zufallszahlen mit beliebigen Verteilungsfunktionen additiv oder multiplikativ verknüpft werden. Vor allem in der Natur tritt dieses Problem oft auf.

Ein Beispiel um dies deutlicher zu machen: In einem Stall gibt es unterschiedlich große Mastschweine. Um die Wachstumszunahme darzustellen, kann der Höhenunterschied in Zentimeter oder der Unterschied vom Ausgangspunkt in Prozent

dargestellt werden. Bei der Addition würde dies wie folgt aussehen:  $h \rightarrow h + x$ .  $h$  ist die Größe bevor des Schweins und  $x$  die Wachstumszunahme. Am Mastanfang hat das Schwein die Größe  $h$  und  $h + x$  am Mastende. Im Fall der Multiplikation hat es folgende Gestalt:  $h \rightarrow h \cdot (1 + \frac{y}{100})$ . In dem Fall kommt noch  $y$  als Prozentsatz dazu.

Vor allem in der Physik, Biologie und Chemie stehen die multiplikativen Verknüpfungen im Vordergrund. Wie schon vorhin erwähnt an Hand des Beispiels mit dem Mastschwein, beruhen auch das Wachstum und die Vermehrung auf multiplikativen Verknüpfungen. So können empirische Daten besser bezeichnet und beschrieben werden [13, S. 191]. Da hauptsächlich multiplikative Verknüpfungen in der Realität vorkommen, sind die vorhin Transformationen unumgänglich.

Der zentrale Grenzwertsatz ist ein Hauptsatz in der theoretischen Statistik. Seien  $X = X_1 + X_2 + \dots + X_n$  die Summe von  $n$  unabhängigen und identischen verteilten Zufallsvariablen  $X_i$  mit positiver Varianz. Weiter sei  $T = \frac{Z - \mu_Z}{\sigma_Z}$  die standardisierte Zufallsvariable zu  $Z$  mit  $\mu_T = 0, \sigma_T = 1$ . Dann gilt  $\lim_{n \rightarrow \infty} P(T \leq x) = f(x)$

Die Zufallsvariable  $Z_1$  wird der Erwartungswert  $\mu_1$  und die Standardabweichung  $\sigma^2$  zugeordnet. Wegen der Unabhängigkeit der  $Z_i$  gilt  $\mu_Z = E(Z_1 + Z_2 + \dots + Z_n) = E(Z_1) + E(Z_2) + \dots + E(Z_n) = n\mu_1$  und analog  $\sigma = \sqrt{n} \cdot \sigma_1$ . Das Verhalten der Summe hängt also asymptotisch nur vom Erwartungswert und der Varianz der  $Z_i$ , nicht aber von der speziellen Form ihrer Verteilung ab [20, S. 288].

Ist die Summe von  $n$  unabhängigen stochastischen Zufallsvariablen einigermaßen normalverteilt, wird von dem zentralen Grenzwertsatz gesprochen [13, S. 191].

Im Gegenzug dazu steht die Grenzverteilung als das Produkt von unabhängigen Zufallsvariablen. Hierbei handelt es sich um die logarithmische oder multiplikative Normalverteilung. Um die beiden Verteilungen auch empirisch erstellen zu können, gibt es ein physikalisches Simulationsmodell. Dies geht auf den Naturforscher Sir Francis Galton zurück. Galton wurde 1822 in Birmingham geboren. Er war nicht nur Physiker, sondern auch Geologe und Meteorologe, forschte über Afrika und die Vererbungslehre. Das Galtonbrett geht auf ihn zurück [13, S. 192].

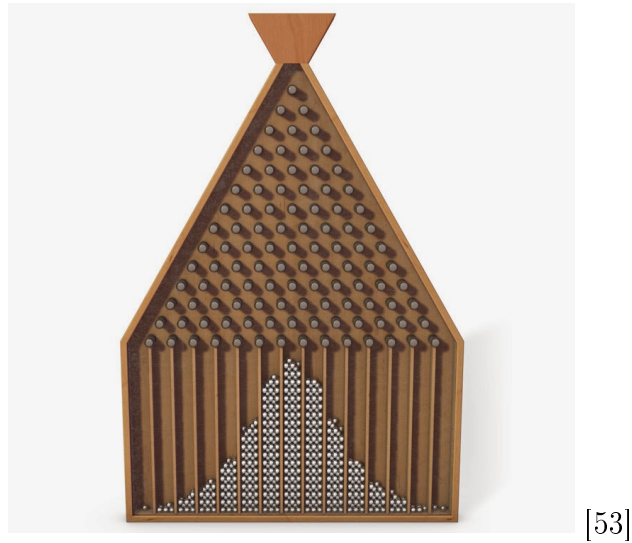
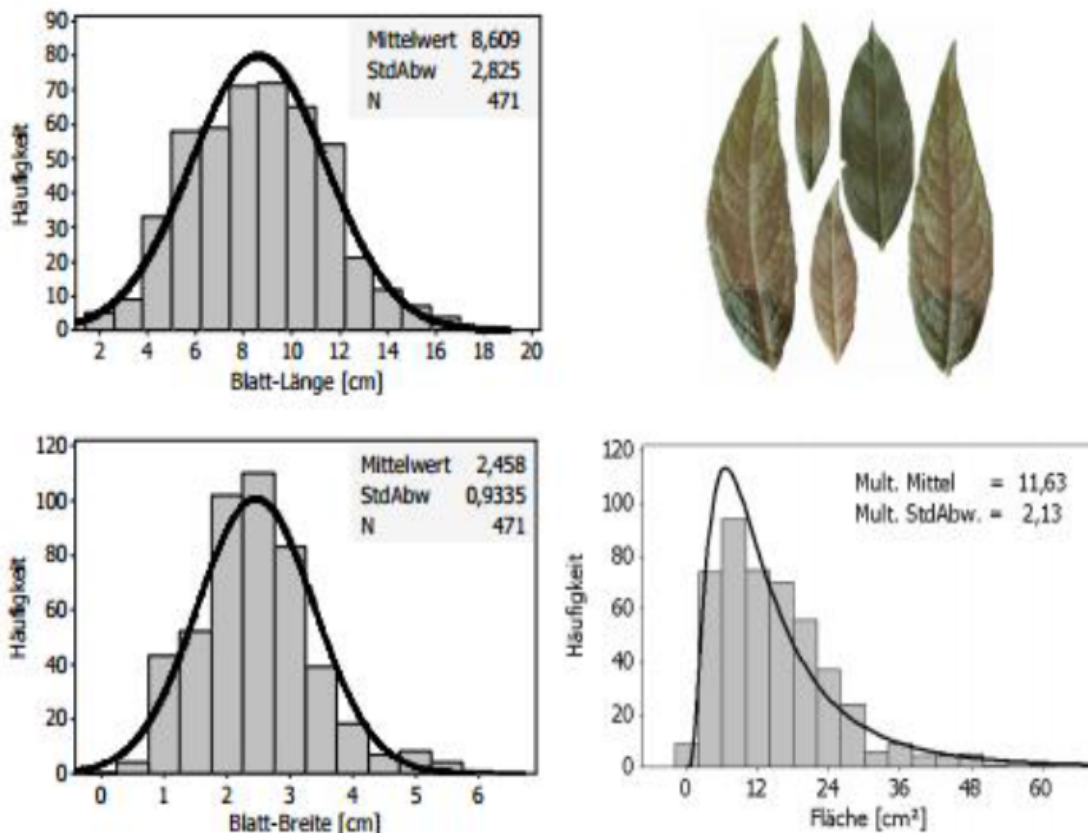


Abbildung 10: Galtonbrett

Dieses Brett ist mit Nägeln versehen, welche einen konstanten horizontalen Abstand haben und in mehreren Reihen versetzt zueinander angeordnet sind. Diese Anordnung hat auch Ähnlichkeiten zu einem Pascal'schen Dreieck. Nun werden nacheinander Kugeln los gelassen. Dabei hat die Kugel bei jedem Nagel die Möglichkeit entweder nach links oder rechts zu rollen. Die Wahrscheinlichkeit beträgt für beide Seiten 0,5. Am Ende des Brettes befinden sich Behälter um die Kugeln aufzufangen. Dies kann auch mit einem gewöhnlichen Baumdiagramm verglichen werden. Die Anzahl der Kugeln in den Behältern entspricht der Wahrscheinlichkeit der einzelnen Wege, siehe Abbildung 10. Bei dieser additiven Verknüpfung sammeln sich die meisten Kugeln in den mittleren Behältern und werden am Rand immer weniger, dies entspricht einer Normalverteilung. Dabei wird von einer Gauß-Normalverteilung mit  $\mathcal{N}(\mu, \sigma)$  gesprochen. Bei der Variante mit der Multiplikation entsteht eine Logarithmische- Normalverteilung mit  $\mathcal{L}(\mu, \sigma)$ . Hier ist kein großer Unterschied zur Normalverteilung, da bei der Multiplikation von Logarithmen die einzelnen Logarithmen addiert werden [15].

Um dies noch in einen praktischen Kontext zu bringen, können Blätter, siehe Abbildung 11 betrachtet werden. Die Länge und die Breite sind additive Verknüpfungen und passen zu Gauß- Normalverteilung. Die Fläche ist eine multiplikative Verknüpfung und so der logarithmischen Normalverteilung zugeordnet.



[13]

Abbildung 11: Entstehung einer logarithmischen Normalverteilung  $\mathcal{L}(\mu; \sigma)$  aus den beiden Normalverteilungen der Blattlänge  $L$  (links oben), der Blattbreite  $B$  (links unten) und der Blattfläche  $F$  (rechts unten)

Weitere Beispiele aus der Wissenschaft bei der die Log- Normalverteilung verwendet wird, ist die Humanmedizin. Dabei kann die latente Periode von Infektionskrankheiten von der Infektion bis zu den ersten Symptomen mit dieser besonderen Verteilung dargestellt werden. Auch im Bereich der Umwelt, spiegelt sich diese Verteilung wieder. Ein Beispiel ist die Verteilung von chemischen und organischen Teilchen in der Luft [29, S. 346].

## 6 Zufallszahlen und deren Generatoren

### 6.1 Zufallszahlen

Unter Zufallszahlen wird das Ergebnis von speziellen Zufallsexperimenten verstanden. Solche Zufallszahlen werden vorallem in der Statistik verwendet. Weiters liegen den Zufallszahlen verschiedene Verteilungen zu Grunde.

Eine einzelne Zahl wird streng genommen nicht als Zufallszahl definiert, da über diese nicht genau gesagt werden kann, ob zum Beispiel die Zahl fünf zufällig gewählt wurde oder ob es das Ergebnis einer Berechnung ist. Daher spricht Knuth [10] von einer "Folge von unabhängigen Zufallszahlen zu einer bestimmten Verteilung". Daraus ergibt sich, dass jede Zahl zufällig gewählt wird und nicht in einem Zusammenhang zu der Folge steht. Weiters sollte die Zahl in einem bestimmten Wertebereich einer vorgegebenen Wahrscheinlichkeit liegen.

Solche Folgen von Zufallszahlen können mit Hilfe eines Algorithmus erstellt werden. Dabei wird von sogenannten Pseudozufallszahlen gesprochen. Damit nicht ewig viele Tabellen mit Zufallszahlen gebraucht werden, wird vor allem auf Zufallszahlen zurückgegriffen, die durch Algorithmen hergestellt werden können. In der gegebenen Folge sollen keine Muster, keine Reihenfolgen oder Strukturen zuerkennen sein [10, S. 46].

Bei Zufallszahlen wird zwischen den diskreten und kontinuierlichen Zufallszahlen unterschieden. Bei den kontinuierlichen Zufallszahlen spricht man auch von den Standardzufallszahlen, die im Intervall  $[0, 1]$  gleichverteilt sind. Zufallszahlen können beliebigen Verteilungen genügen. Wobei in den meisten Fällen eine Gleichverteilung vorliegt.

### 6.2 Zufallszahlengeneratoren

Es werden zwischen Echten- und Pseudozufallszahlengeneratoren unterschieden. Echte Zufallszahlengeneratoren nutzen physikalische Erscheinungen und ihre Messungen zur Generierung von Folgen von Zufallszahlen.

Verfahren zur Generierung von Folgen von Zufallszahlen werden als Zufallszahlengeneratoren bezeichnet. Die ersten Methoden gehen auf Hammersley und Handscob zurück, die mit Hilfe von Beobachtung Zufallszahlen erzeugt haben. Unter der Beobachtungsmethode wird vor allem das Würfeln und Ziehen von Zahlen aus einer Urne, sowie das Drehen beim Roulette verstanden. Um aber auch später auf diese Zufallszahlen zurückgreifen zu können, werden diese in Form einer Tabelle

dargestellt. Jedoch ist bei dieser Art von Erzeugung mit einem sehr großen Aufwand zu rechnen. Daher wurde sehr rasch begonnen, den Computer als Hilfsmittel einzusetzen. Im Gegenzug dazu, gibt es auch die Pseudozufallszahlengeneratoren, welche mit deterministischen Algorithmen Zufallszahlen entwickeln [18, S.23].

Eine wesentliche Eigenschaft eines Zufallsgenerator ist die Korrelationsfreiheit. Unter der Eigenschaft der Korrelationsfreiheit wird folgendes verstanden, "[lat.], das Vorliegen eines irgendwie gearteten Zusammenhanges zwischen zwei oder mehreren Zufallszahlen und seine mathemat. Darstellung."

In diesem Fall sind es statistische Zusammenhänge zwischen aufeinanderfolgenden Zufallszahlen.

Folgende Bedingungen sollten Generator erfüllen, die Standardzufallszahlen generieren:

1. Gleichverteilung: Eine Zufallsfolge sollte der Gleichverteilung zu Grunde liegen. Daher ist eine Zufallszahl immer an einen Bereich gebunden in dem Fall ist es das Intervall  $[0, 1]$ .
2. Vorhersagbarkeit: Sobald eine Zahl erzeugt wurde, sollte die darauf folgende Zahl nicht errahnt werden können. Aus diesem Grund sollte das Konstruktionsprinzip so gestaltet sein, dass der Algorithmus nicht gleich ersichtlich ist.
3. Periodenlänge: Die Periodenlänge an Zufallszahlen muss zumindest so groß sein wie Zufallszahlen benötigt werden, damit sich die Zahlen auf keinen Fall wiederholen. Außerdem sollte die Degeneration vermieden werden. Dabei wiederholt sich dieselbe Zahl immer wieder. Dies ist ein besonderer Fall der auf keinen Fall eintreffen darf. [18, S. 28].

### 6.2.1 Echte Zufallszahlengeneratoren

Echte Zufallszahlengeneratoren lassen sich auf physisikalische Prozesse zurückführen.

Den Beginn haben Kendall und Babington-Smith [18, S. 23] 1938 gemacht. Sie haben eine Scheibe in Sektoren geteilt und diese mit den Ziffern Null bis Neun beschriftet. In kurzen Abständen wurde dann immer ein Sektor beleuchtet und die dazugehörige Zahl aufgeschrieben. So konnten sie über 100000 Zahlen zufällig erzeugen. All diese Zahlen wurden in Tabellen festgehalten mussten dort abgespeichert und anschließend in einen Computer eingelesen werden. Diese Methode ist Zeitaufwendig. Um diesen Vorgang etwas allgemeiner zu beschreiben, wird ange-

nommen, dass der Generator ganzzahlige Werte zwischen  $r_{min}$  und  $r_{max}$  annimmt, also

$$r \in \{r_{min}, r_{min} + 1, \dots, r_{max}\} \quad (44)$$

Durch wiederholtes Anwenden liefert diese Funktion eine Sequenz von Zufallszahlen [11, S. 58].

In den 1940er und 50er Jahren wurde vor allem mit numerischen und arithmetischen Verfahren gearbeitet, um an Zufallszahlen zu kommen. Der erste bekannte echte Zufallszahlengenerator ERNIE wurde von Sidney Broadhurst, Tommy Flowers und Harry Fenson entwickelt. ERNIE ist die Abkürzung für Electronic Random Number Indicator Equipment. Dabei wurden mit Leuchtstoffröhren Geräuschimpulse erzeugt, welche in einem konstanten Zeitintervall Zufallszahlen gezählt und dargestellt haben. Mit dieser Methode wurden bis zu 50 Zufallszahlen erzeugt. ERNIE wurde vom Britischen Postministerium zur Ziehung der Gewinner des "premium bond saving lottery" verwendet. The Rand Corporation hat 1955 ein Buch veröffentlicht, mit circa einer Million Zufallszahlen. Diese wurden mit Hilfe eines Impulsgebers, der circa 100000 Impulse pro Minute erzeugt hat entwickelt. 1983 wurde durch das Zählen vom Gammaquanten eine weitere Methode von Miyake entwickelt, um eine große Zahl von Zufallszahlen zu erhalten [18, S. 23ff.].

Diese Art von Zufallszahlengeneratoren werden auch physikalische Generatoren genannt. Da zum Beispiel der Zerfall radioaktiven Materials, Schwankungen in elektrischen Netzen oder thermisches Rauschen gemessen wird [10, S. 46].

Ein großer Vorteil dieser Generatoren ist, dass all die Zahlenfolgen nicht vorhersehbar sind und daher von einer zufälligen Folge ausgegangen werden kann. Jedoch gibt es auch einige Gründe warum echte Zufallszahlengeneratoren heutzutage kaum noch in Verwendung sind. Einer davon ist, dass all die Zusatzgeräte gewartet werden müssen, um zu prüfen ob die erzeugte Zufallsfolge der gewünschten Verteilung entspricht. Weiters ist bei Generatoren die Reproduzierbarkeit ein großer Vorteil. Doch dies ist hier nicht möglich, da nie die selben Zahlenfolge auftreten wird [18, S.25ff.].

### 6.2.2 Pseudozufallszahlengeneratoren

Pseudozufallszahlengeneratoren erzeugen Zufallszahlen nach einem bestimmten Algorithmus. Es wird eine Zahlenfolge erzeugt, die zwar den Anschein macht zufällig zu sein, es aber nicht ist. Da sie von einem Algorithmus berechnet wird .



Bei Pseudozufallszahlengeneratoren kommt es vor allem auf die Komplexität des Algorithmus an, da ein enger Zusammenhang zwischen Rechenzeit und Algorithmus besteht. Je geringer die Rechenzeit, desto einfacher ist der Algorithmus aufgebaut und umso schneller wird eine Zufallszahl produziert. Ein anderes positives Argument ist die Stabilität der Pseudozufallszahlengeneratoren. Denn im Vergleich zu physikalischen Generatoren können bei diesem keine Störungsquellen, wie zum Beispiel fehlende Kabel, auftreten [18, S. 24]. Ein weiterer Vorteil ist die Reproduzierbarkeit. Dabei ist es möglich, dass eine schon erzeugte Folge von Zufallszahl noch einmal reproduziert werden kann. Um Fehler so leichter finden und Simulationen besser miteinander vergleichen zu können. Dies ist möglich, indem von den identischen Anfangsbedingungen ausgegangen wird [10, S. 46]. Dies wird oftmals als Seeding bezeichnet und ermöglicht mit dem selben Anfangswert idente Zufallszahlen zu erzeugen.

Einer der ersten Algorithmen war die von Neumann genannte "middle-square-Methode" bei der die vorherigen Zufallszahlen quadriert wurden und die mittleren Ziffern als neue Zufallszahl auftreten. Zu diesem Generator gibt es etliche Studien, die sehr bald bewiesen haben, dass diese Methode nicht weiterführend ist. Denn wenn einmal die Zahl Null vorkommt, wiederholt sich diese Zahl unendlich oft. Wenn die beiden letzten signifikanten Ziffern Null sind, konvergiert die Folge gegen Null oder es entsteht eine sehr kurze Sequenz. Ein weiteres Problem ist, dass es durchaus Zahlen gibt, die sich selbst reproduzieren. Ein Beispiel ist die Zahl 3792. Wird die Zahl 3792 quadriert, ergibt das 14379264. Werden von dieser Zahl die mittleren Ziffern genommen, ergibt das 3792, was dazu führt, dass die selbe Zufallszahl wie am Beginn erscheint.

Ein weiterer Generator namens Tausworthe, welcher 1965 entwickelt wurde, stützt sich auf binäre Zahlenfolgen. Auch der Generalized Feedback Shift Register Generator, auch GFSR abgekürzt, basiert auf diesem Prinzip. Im Jahr 1989 wurde bewiesen, dass diese beiden Generatoren äquivalent sind [18, S. 25].

Der Lagged-Fibonacci-Generator, basiert wie der Name schon vermuten lässt auf der Fibonacci-Folge. Anfangs hatte der Generator mit der lang andauernden Rechenzeit Schwierigkeiten und war daher nicht sehr oft in Verwendung.

Auch einige Generatoren, die auf der Basis des linearen Kongruenzgeneratoren basieren, wurden vor allem von Lehmer 1951 und von Eichenauer und Lehn 1986 geprägt. Weitere Methoden zur Erzeugung von Zufallszahlen sind die quadratischen Generatoren und die Kombination mehrerer Zufallszahlengeneratoren [18, S. 24f].

### 6.2.3 Startwertproblem

Bei echten Zufallszahlengeneratoren ist es nicht notwendig an einer bestimmten Stelle der Tabelle oder Folge zu beginnen, solange die gleiche Verteilung vorliegt.

Bei Pseudozufallszahlengeneratoren ist der Algorithmus determiniert, was heißt, dass er für die gleichen Eingabewerte immer die selben Ausgangswerte liefert. Wenn er alle Zufallszahlen durchlaufen hat, wiederholt sich die Folge von vorne. Bei dieser Wiederholung wird von einer Periode gesprochen.

Definition einer Periode laut Härtel [18, S. 26]:

Die Periode eines Pseudozufallszahlengenerators ist der Zyklus der sich immer wiederholenden Zahlenfolge. Dieser Zyklus tritt entweder sofort oder erst nach einer gewissen Anzahl von Werten (der sogenannten Vorperiode) auf. Die Anzahl in diesem Zyklus wird Periodenlänge genannt.

Die Periodenlänge bei Pseudozufallszahlen ist vom Startwert abhängig, daher der Name Startwertproblem [18, S. 26].

## 7 Tests von Zufallszahlen

Bei Tests von Zufallszahlen wird die Gleichverteilung sowie die Unabhängigkeit von Zufallszahlen überprüft. Vor allem für die Unabhängigkeit müssen mehrere Tests durchgeführt werden, um auf ein aussagekräftiges Ergebnis schließen zu können.

Zuerst noch eine Definition, die sehr wesentlich für die Überprüfung der Gleichverteilung ist. Unter einer Nullhypothese wird eine Annahme oder Aussage verstanden, die erst durch ihre Falsifizierbarkeit verworfen werden kann.

Für die Gleichverteilung gilt folgende Nullhypothese beziehungsweise Alternativhypothese:

$$\begin{array}{l} H_0^a : U_i \sim U[0, 1] \\ \text{versus} \\ H_1^a : U_i \not\sim U[0, 1]. \end{array} \quad (45)$$

$H_0$  bedeutet in diesem, dass die Zahlen im Intervall  $[0, 1]$  gleichverteilt sind.

Für die Unabhängigkeit gilt folgende Nullhypothese:

$$\begin{array}{l} H_0^b : U_i \text{ unabhängig von } U_j \quad \forall i \geq j \\ \text{versus} \\ H_1^b : U_i \text{ nicht unabhängig von } U_j \quad \forall i \geq j \end{array} \quad (46)$$

Hier wird von der Nullhypothese gesprochen, wenn die aufeinanderfolgenden Zahlen voneinander unabhängig sind. Auch hier gilt, ein Test reicht um die Hypothese verwerfen zu können. Jedoch kann es auch öfters passieren, dass mehrere Tests notwendig sind um die Nullhypothese auch wirklich bestätigen zu können.

### 7.1 Anpassungstests

Um zu überprüfen, ob empirisch gewonnene Zufallszahlen einer vorgegebenen Verteilung entsprechen, kann ein entsprechender Anpassungstest gewählt werden. Dabei wird eine Testgröße anhand der Zufallszahl berechnet und mit einem Referenzwert verglichen. Die Testgröße wird durch den Vergleich zwischen der empirischen Verteilung und der vorgegebenen theoretischen Verteilung berechnet. Beispiele für solche stochastischen Tests sind der Kolmogorov- Smirnov- Test und der  $\chi^2$ - Test. Beide Tests sind für beliebige Verteilungen anwendbar [18, S. 29ff.].

Länge Run 1: 1  
Länge Run 2: 3  
Länge Run 3: 2  
Länge Run 4: 4  
Länge Run 5: 2

## 7.2 Tests für die Unabhängigkeit

Dabei werden die einzelnen Zahlen einer Folge von Zufallszahlen nach ihrer Unabhängigkeit überprüft. Es kann nicht direkt geprüft werden, ob die Zahlen unabhängig sind. Daher kann nur gesagt werden ob ein Generator hinreichend gut oder schlecht ist.

### 7.2.1 Runtest

Ein Runtest wird angewendet um die Zufälligkeit einer Folge zu bestimmen. Bei diesen Tests wird die Länge monotoner Teilfolgen ermittelt und untersucht, ob diese monoton fallend oder steigend sind. Als einen Run, oder im deutschen auch als Iteration bekannt, wird eine monotone Folge von Zahlen bezeichnet.

Nun ein Beispiel zu einem Run- Test [16, S. 643].

Ausgegangen wird von einer Folge von Zufallszahlen:

2 1 2 3 2 4 1 7 8 9 0

Da bei diesem Test die Monotonie ein wesentliches Merkmal spielt, wird nun die Folge nach ihren Runs untersucht. In dem Fall können die Runs folgendermaßen unterteilt werden. Dabei beginnt ein neuer Run sobald die folgende Zahl kleiner als die vorangegangene Zahl ist.

Run 1: 2  
Run 2: 1 2 3  
Run 3: 2 4  
Run 4: 1 7 8 9  
Run 5: 0

Die Länge dieser Runs hängt nun immer von der Anzahl der Zahlen zusammen, die in einem Run vorkommen.

Bei solch einer Folge wird von einer Iteration nach oben oder einem Runs- Up gesprochen, da folgende Eigenschaft gilt:  $U_j > U_{j+1}$ . Im Gegensatz dazu gibt es

auch die monoton fallenden Teilfolgen, welche die analoge Eigenschaft  $U_j < U_{j+1}$  besitzen und Iteration nach unten oder Runs- Down genannt werden [18, S. 32].

Es ist hier nicht möglich mit der Anzahl der Iterationen nach oben beziehungsweise nach unten einen  $\chi^2$ - Test durchzuführen, da aufeinanderfolgende Iterationen nicht voneinander unabhängig sind. Daher wird die unmittelbar auf eine Iteration folgende Realisation  $U_{i+1}$  verworfen und die nächste Iteration wird erst mit  $U_{i+2}$  begonnen. Erst dann kann ein  $\chi^2$ - Test angewendet werden.

### 7.2.2 Pokertest

Der Pokertest überprüft die Unabhängigkeit einer Zahlenfolge. Dabei werden Zufallszahlen in Fünfer Gruppen eingeteilt, wobei die Zahlen in aufeinanderfolgender Reihenfolge abgebildet werden. Anhand dieser Aufteilung kann nun beobachtet werden, welche der folgenden sieben Muster auftreten. Diese Muster wurden aus den Bezeichnungen von Poker übernommen, daher der Name Pokertest.

alle verschieden:	abcde
ein Paar:	aabcd
zwei Paare:	aabbc
Full House:	aaabb
drei Gleiche:	aaabc
vier Gleiche:	aaaab
fünf Gleiche:	aaaaa

Anschließend kann ein  $\chi^2$ - Test mit sechs Freiheitsgraden angewendet werden. Es gelten folgende Wahrscheinlichkeiten, welche auch als Erwartungswerte definiert werden:

$P(abcde)=$	0,3024
$P(aabcd)=$	0,504
$P(aabbc)=$	0,108
$P(aaabb)=$	0,009
$P(aaabc)=$	0,072
$P(aaaab)=$	0,0045
$P(aaaaa)=$	0,0001

Ein Problem ergibt sich, dass sehr viele Beobachtungswerte benötigt werden, damit die Bedingung erfüllt wird, dass die Anzahl in jeder Klasse zumindest größer

gleich fünf ist. Aus der theoretischen Häufigkeit für fünf gleiche Zufallszahlen mit  $P(aaaaa) = 0,0001$  ergibt sich, dass zumindest  $n \geq 50000$  Zufallszahlen erzeugt werden müssen, um die Beobachtung dieser Bedingung zu erfüllen.

Es gibt noch weitere Methoden für den Pokertest, indem weniger Realisationen angenommen werden. Also in dem Fall statt fünf Gruppen nur vier oder drei Gruppen. Jedoch verändert sich der Beobachtungswert bei vier Gruppen auf  $n \geq 1839$ , da zwei Gruppen zu einer gemacht werden. Bei drei Gruppen verringert sich dieser Wert wieder auf  $n \geq 741$ , da wieder die beiden letzten Gruppen zu einer vereinigt werden [18, S. 34]. Mit dem  $\chi^2$ -Test kann nun die theoretische Häufigkeit mit der empirischen Häufigkeit der Zufallszahlen-Folge verglichen werden. Wenn die Summe des  $\chi^2$  Werte aller Klassen kleiner als der kritische Wert ist, dann ist die Bedingung erfüllt, dass das Auftreten von Wiederholungen von Zufallszahlen der theoretischen Wahrscheinlichkeitsverteilung entspricht.

### 7.2.3 Lückentest

Auch dieser Test bietet die Möglichkeit Zahlenfolgen auf ihre Unabhängigkeit zu überprüfen. Dabei wird beobachtet wie lange die Lücke zwischen zwei Erfolgen ist. Es wird untersucht, ob dieser Wert in dem kritischen Bereich  $0 \leq \alpha < \beta \leq 1$  auftritt. Dabei werden unter  $\alpha$  und  $\beta$  reelle Zahlen verstanden, die folgende Eigenschaft haben:  $\alpha < \beta$ . Eine Lücke ist die Anzahl von Zufallszahlen zwischen zwei Erfolgen. Die Lücke, die zwischen den beiden Bereichen auftritt, wird Gap genannt. Daher auch der Namen Gaptest oder Lückentest [18, S. 33].

Anschließend kann man mit einem  $\chi^2$ -Test überprüfen, ob die beobachtete Anzahl von der zu erwartenden Anzahl der Lücken zu stark abweicht [42, S. 29].

## 8 Transformation von Zufallszahlen

Da Zufallszahlengeneratoren üblicherweise Zufallszahlen liefern die gleichverteilt sind, sind Verfahren notwendig um Zufallszahlen zu transformieren die auch andere Verteilungsfunktionen generieren. Das können sowohl diskrete als auch stetige Verteilungsfunktionen sein.

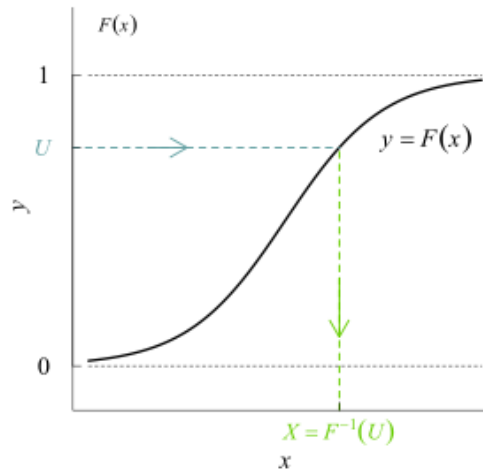
### 8.1 Direkte Transformation

Diese Methode realisiert eine Verteilungsfunktion mit Hilfe von Standardzufallszahlen im Intervall  $[0; 1]$ . Es ist offensichtlich das stetige Zufallszahlen die im Intervall  $[0; 1]$  liegen auf beliebige Intervallgrenzen transformiert werden können. Durch Zusammensetzen verschiedener Realisationen von Standardzufallszahlen kann eine direkte Transformation durchgeführt werden. Dabei werden die Verteilungen entweder multipliziert oder addiert, damit eine neue oder die gewünschte Verteilungsfunktion gebildet werden kann. Zum Beispiel bei der Poissonverteilung wird diese Methode angewendet, indem Standardzufallszahlen mit einander multipliziert werden [18, S. 125].

### 8.2 Inversionsmethode

Wie der Name schon sagt, spielt hier das Inverse eine wesentliche Rolle. Nur wenige Verteilungen können für diese Methode verwendet werden, da die Berechnung sehr kompliziert ist. Beispiele dafür sind die Exponentialverteilung so wie die Cauchyverteilung. Hierbei werden Pseudozufallszahlen erzeugt, die als künstliche herbeigeführte Realisationen statistische Zufallsvariablen nachstellen. Dabei werden aus gleichverteilten Zufallsvariablen andere Wahrscheinlichkeitsverteilungen erzeugt. Diese Folgen werden verwendet um das Verhalten komplexer Systeme die man mit Funktionen analysieren kann, zu untersuchen.

Für die Erzeugung muss die Inverse oder auch Quantilfunktion  $F^{-1}$  zu der Verteilungsfunktion  $F$  bekannt sein. Durch die Berechnung von  $x = F^{-1}(u)$  kann eine Realisation  $x$  der Wahrscheinlichkeitsverteilung  $F(x)$  erzeugt werden. Unter  $u$  wird eine Standardzufallszahl welche auf dem Intervall  $[0, 1]$  gleichverteilt ist. Die Inversefunktion  $F^{-1}$  ist immer auf dem Wertebereich von  $F$  definiert. Die Funktion  $F$  ist definiert auf dem Intervall  $[0, 1]$  mit  $u \in [0, 1]$ .



[51]

Abbildung 12: Inversionsmethode mit dem Inversen  $F^{-1}$  der Verteilungsfunktion  $F$

Im folgenden Beweis lässt sich leicht zeigen, dass  $x$  eine Realisation der Zufallsvariable  $X$  mit der Verteilung  $F$  ist:

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x). \quad (47)$$

Auf Grund der Gleichverteilung von  $U$  auf  $[0, 1]$  und  $0 \leq F(x) \leq 1$  ist die letzte Gleichung bewiesen. Bei diskreten Verteilungen führt die Inversionsmethode zu Problemen, da die Verteilungsfunktion Sprungstellen besitzt [18, S. 126].

In Abbildung 13 ist von einigen Dichtefunktionen die Verteilungsfunktion  $F$  und ihre Umkehrfunktion  $F^{-1}$  zusammengefasst.



Dichtefunktion	Verteilungsfunktion	Quantilfunktion	Bereich
$f(x) = x \cdot \text{const.}$	$F(x) = x^2 \cdot \text{const.}$	$F^{-1}(U) = a \cdot \sqrt{U}$	$0 \leq U \leq 1, \quad 0 \leq x \leq a$
$f(x) = x^n \cdot \text{const.}$	$F(x) = x^{n+1} \cdot \text{const.}$	$F^{-1}(U) = a \cdot \sqrt[n+1]{U}$	$0 \leq U \leq 1, \quad 0 < x \leq a, \quad n > -1$
$f(x) = \frac{\text{const.}}{x^n}$	$F(x) = \frac{\text{const.}}{x^{n-1}}$	$F^{-1}(U) = \frac{a}{\sqrt[n-1]{U}}$	$0 < U \leq 1, \quad a \leq x < \infty, \quad n > 1$
$f(x) = e^{-\lambda x} \cdot \lambda$	$F(x) = 1 - e^{-\lambda x}$	$F^{-1}(U) = -\frac{1}{\lambda} \ln(U)$	$0 < U \leq 1, \quad 0 \leq x < \infty$
$f(x) = \sin(x)$	$F(x) = 1 - \cos(x)$	$F^{-1}(U) = \arccos(U)$	$0 \leq U \leq 1, \quad 0 \leq x \leq \pi/2$
$f(x) = \cos(x)$	$F(x) = \sin(x)$	$F^{-1}(U) = \arcsin(U)$	$0 \leq U \leq 1, \quad 0 \leq x \leq \pi/2$
$f(x) = \frac{1}{\pi(1+x^2)}$	$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan(x)$	$F^{-1}(U) = -\cot(\pi U)$	$0 < U < 1, \quad -\infty < x < \infty$

[54]

Abbildung 13: Einige Dichtefunktionen mit ihrer Verteilungsfunktion  $F$  und ihre Umkehrfunktion  $F^{-1}$

### 8.3 Verwerfungsmethode

Die Verwerfungsmethode geht auf John von Neumann zurück. Dabei werden Zufallszahlen zu einer vorgegeben Verteilung erzeugt. Seien  $F$  und  $G$  Verteilungsfunktionen mit den dazu passenden Dichtefunktionen  $f(x)$  und  $g(x)$ . Weiters muss ein  $k \in \mathbb{R}$  mit  $f(x) \leq k \cdot g(x)$  mit  $\forall x, k > 0$ . Die Methode funktioniert wie folgt: Zuerst werden Zufallszahl  $u_i$  und eine Standardzufallszahlen  $v_i$  der zu  $G$  gehörigen Verteilungsfunktion erzeugt. Gilt  $j := \inf\{n \geq 1 \mid k \cdot u_n \cdot g(v_n) < f(v_n)\}$ , als Zufallszahl  $x := v_j$  der Verteilungsfunktion  $F$ . Dies bedeutet, dass auf einen Treffer, der unterhalb von  $f(x)$  liegt, gewartet wird. Andernfalls wird die Zufallszahl verworfen.

Diese Methode kann graphisch sehr gut veranschaulicht werden, siehe Abbildung 14. Dabei werden zwischen der x-Achse und dem Graph von  $k \cdot g(x)$  Zufallszahlen, also Zufallspunkte verteilt. Als x-Koordinate des Punktes  $i$  wird die G-verteilte Zufallszahl  $v_i$  genommen. Für die y-Koordinate wird die standardverteilte Zufallszahl  $u_i$  angenommen. Liegt der Punkt  $(v_i \mid u_i)$  oberhalb des Graphen von  $f(x)$ , werden die Punkte verworfen. Die x-Koordinaten der restlichen Punkte sind dann nach der Dichtefunktion  $f(x)$  verteilt. Damit nun eine Zufallszahl nach dieser Verteilung erzeugt wird, werden so lange Zufallszahlen erzeugt bis eine unterhalb von  $f(x)$ , indem Fall wie der Punkt  $C$ , liegt. [36].

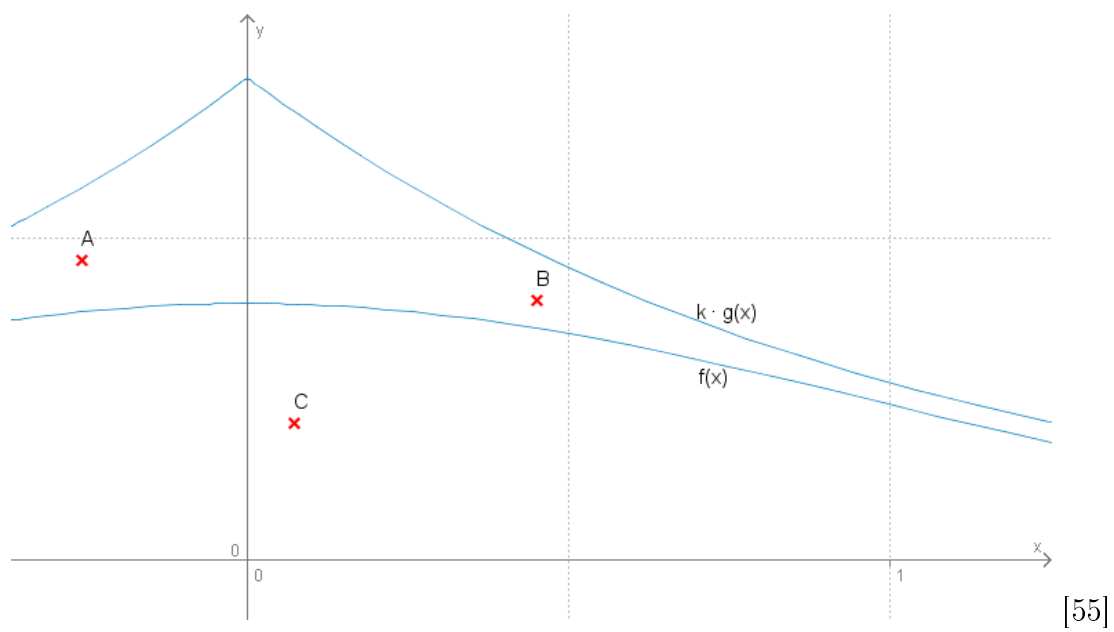


Abbildung 14: Graphische Darstellung der Verwerfungsmethode mit dem Graphen  $f(x)$  und  $k \cdot g(x)$

## 9 Simulation von Zufallszahlen im Unterricht

### 9.1 Der Modellbildungsprozess

Albert Einstein sagte: Insofern sich Sätze der Mathematik auf die Wirklichkeit beziehen, sind sie nicht sicher, und insofern sie sicher sind, beziehen sie sich nicht auf die Wirklichkeit. [22, S. 121]

Der Modellbildungsprozess hat sein Ursprungsproblem nicht in der Mathematik, sondern aus der Realität, vor allem in den Bereichen Wirtschaft und Sozialwissenschaften, sowie Umwelt- und Verkehrsfragen. Dabei ist der Ausgangspunkt des Problems offen. Jedoch müssen Strukturen so weit erkennbar sein, dass mathematische Formulierungen und Vorgehensweisen ermöglicht werden können. Dieser Prozess gliedert sich in fünf Schritte. Als Schritt Null wird die Formulierung des Problems bezeichnet. Erst dann startet der Prozess.

**Schritt 1:** Die Grundaufgabe dieses Arbeitsschrittes ist die Schaffung eines Systems. Dabei müssen zuerst alle Bedingungen, Voraussetzungen und Einflussgrößen gesammelt und strukturiert werden. Ein System besteht aus einem oder mehreren strukturellen Elementen, deren Zustände von anderen Elementen oder sich selbst abhängen und die Zustände anderer Elemente beeinflussen. Ein System hat auch immer einen Zweck. Außerdem hat es immer eine Systemgrenze, die es von seiner Systemumwelt trennt. Diese Grenze ist nur dort eindeutig, wo kein Austausch von Stoffen oder Energie besteht [39, S. 10ff.]. Damit nun mit diesem System sinnvoll weiter gearbeitet werden kann, müssen oft Vereinfachungen und Idealisierungen gefunden werden, da es für die Lösung eines Problems nicht unbedingt notwendig ist alle Gegebenheiten zu berücksichtigen. Diese Entscheidungen sind oft subjektiv beeinflusst [22, S. 121].

**Schritt 2:** Hierbei steht die Mathematisierung, also der Übergang des Systems in ein formales, mathematisches Modell mit Hilfe von Mengen, Graphen und Funktionen, des Realmodells, im Vordergrund. Bei diesem Schritt wird deutlich, dass sich zwar das mathematische Modell und das System weitgehend entsprechen, jedoch die Übereinstimmung zwischen Ursprungsproblem und mathematischem Modell vom zwischengeschalteten System abhängen [22, S. 122].

**Schritt 3:** Dieser Schritt bereitet am wenigsten Schwierigkeiten, denn es wird nur an einer mathematischen Lösung gearbeitet. Da das mathematische Modell schon in Schritt zwei ausgearbeitet wurde, ist noch ein innermathematischer Prozess am Laufen. Es gibt zwei verschiedene Lösungswege: Entweder es wurde schon eine Lösung gefunden oder das Modell verlangt, dass noch weitere vereinfachende Maßnahmen getroffen werden, damit eine passende Lösung entsteht.

**Schritt 4:** Ein sehr wichtiger Schritt im Modellbildungsprozess ist die Interpretation der mathematischen Lösung und Validierung des Modells. Bei der Interpretation wird das mathematische Ergebnis in die Realität rückübersetzt. Die Validierung ist der nächste Schritt. Hierbei wird überprüft, ob das Ergebnis auch wirklich eine sinnvolle Lösung des Ursprungsproblem darstellt. Dabei wird auch beobachtet, ob trotz der Vereinfachung des Ursprungsproblem alle relevanten Faktoren übereinstimmen [22, S. 123]. Ob das Modell richtig ist, lässt sich nicht allgemein beweisen. Für die vorhin genannten Faktoren gibt es vier wesentliche Aspekte:

**Strukturgültigkeit:** Für die Strukturgültigkeit muss gezeigt werden, dass die Struktur des Modells der Struktur des Ursprungsproblem entspricht.

**Verhaltensgültigkeit:** Dabei muss gezeigt werden, dass alle Anfangsbedingungen und Umwelteinflüsse, die im Ursprungsproblem vorkommen auch auf das Modell die gleichen dynamischen Auswirkungen haben.

**Empirische Gültigkeit:** Hierbei müssen numerische Ergebnisse des Modells mit den empirischen Daten des Ursprungsproblem weitgehend übereinstimmen.

**Anwendungsgültigkeit:** Bei der Anwendungsgültigkeit muss gezeigt werden, dass die Modelldarstellung dem Zweck, der in Schritt 1 definiert wurde, entspricht.

**Schritt 5:** Der letzte Schritt, die Veränderung des Modells, steht in einem sehr engen Zusammenhang mit Schritt vier. Stimmen die Ergebnisse in Schritt vier nicht mit dem realen Modell überein, muss eine neue mathematische Herangehensweise gefunden werden und somit ein neues Modell aufgestellt werden [39, S. 14.] .

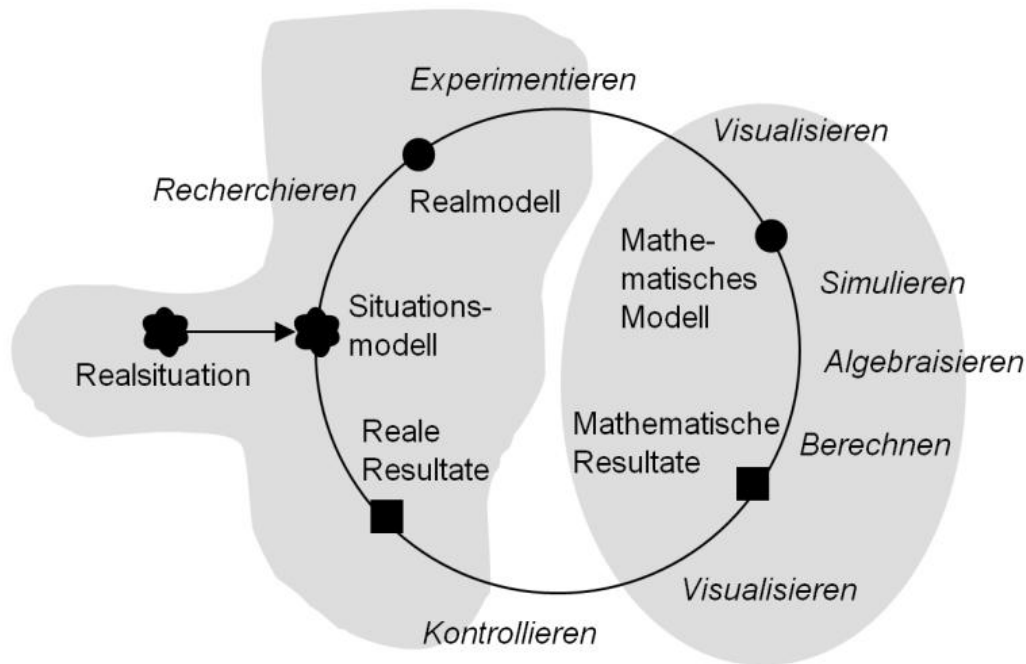
An diesem Punkt würde sich dann der Prozess mit den fünf Schritten wiederholen, bis das gewünschte Modell mit den richtigen Ergebnissen realisiert wurde [22, S. 121ff.].

Eine Grafik nach Blum/Leiß, siehe Abbildung 15, stellt die einzelnen Schritte des Modellbildungsprozesses dar.

## 9.2 Modellieren im Unterricht

In diesem Abschnitt soll der Weg vom abstrakten Modellbildungsprozess zum Modellieren in der Schule beleuchtet werden. Dabei werden vor allem der Prozess der Entwicklung, der Darstellung und der Verwendung von Modellen, um reale Situationen zu analysieren, Phänomene zu verstehen und Entscheidungen zu treffen, behandelt.

Ein wichtiger Zugang ist, dass Schülerinnen und Schüler nicht nur mathematische Zusammenhänge lernen, sondern ihren eigenen Weg zur Lösung von realistischen



[22]

Abbildung 15: Modellbildungsprozess

Problemen entwickeln sollen. Außerdem soll nicht, wie im traditionellen Unterricht, eine bestimmte Lösung für ein Problem gefunden werden, sondern die Schülerinnen und Schüler sollen es schaffen eine Lösung oder ein Modell zu entwickeln, dass sich vielleicht sogar auf andere Problemsituationen verallgemeinern lässt.

Damit so ein modellorientierter Unterricht auch wirklich umsetzbar ist, hat Lesh folgende sechs Prinzipien entwickelt:

1. Das Realitätsprinzip. Kann dieses Problem im Alltag wirklich Auftreten? Können die Schülerinnen und Schüler, auf Grund ihrer Erfahrung und ihres Wissens dieses Problem wirklich verstehen?
2. Das Prinzip der Modellkonstruktion. Können die Schülerinnen und Schüler wirklich erkennen, dass es sich um die Erarbeitung eines Modells handelt?
3. Das Prinzip der Modelldokumentation. Dabei soll sicher gestellt werden, dass die Schülerinnen und Schüler ihre Antworten, Muster oder Ideen darlegen können.
4. Das Prinzip der Eigenevaluation. Sind die Schülerinnen und Schüler in der Lage ihre möglichen Lösungen richtig einzuschätzen?
5. Das Prinzip der Modellverallgemeinerung. Kann das Modell auch auf weitere

Probleme angewandt werden?

6. Das Prinzip des effektiven Lernprototyps. Ist die Ausgangssituation so einfach, dass trotzdem komplexere Modelle entwickelt werden können?

Während dieses Prozesses ist es notwendig, dass die Lehrperson über die Vorkenntnisse und Erfahrungen der Schülerinnen und Schüler Bescheid weiß, damit eine wissensgerechte Erarbeitung eines Problems stattfinden kann. Weiters sollte der Lehrperson die Ergebnisse und Lösungen bekannt sein, damit auf etwaige Schwierigkeiten oder Missverständnisse eingegangen werden kann. Jedoch sollte darauf geachtet werden, dass trotz der Unterstützung die Autonomie und Selbstevaluation nicht gehindert wird. Zuletzt sollen die Schülerinnen und Schülern im Erkennen der Verallgemeinerung ihres Modells sowie der Anwendung auf ähnliche Ausgangssituationen unterstützt werden.

All diese Prinzipien und Vorgehensweise sind notwendig um diesen Prozess fördern, begleiten und unterstützen zu können [30, S. 83ff.].

### 9.3 Anwendungen im Unterricht

Eine Möglichkeit den Modellbildungsprozess anzuwenden ist die Monte-Carlo Simulation. Dies ist eine Methode um mathematische, physikalische aber auch statistische Probleme mit Hilfe von Modellen zu lösen. Dabei liegt das Augenmerk nicht nur auf einem bestimmtem Algorithmus, sondern auf,

eine Gruppe von numerischen Methoden, die Zufallszahlen zur approximativen Lösung oder zur Simulation verschiedener Prozesse einsetzen [23].

Dabei muss auch die Rechenzeit des Computers berücksichtigt werden, denn je länger diese ist, umso besser ist die Lösung. Daher ist es sinnvoll, bei jedem Beispiel abzuschätzen, welche die wichtigere Komponente von beiden ist [19].

Die Monte-Carlo Simulation kann nicht nur in der höheren Mathematik angewendet werden, sondern auch im Alltag. Um diesen Vorgang den Schülerinnen und Schülern besser näher zu bringen, werden im folgenden Abschnitt Beispiele beschrieben, analysiert und erklärt. Dies wird anhand von einem Kalkulationsprogramm veranschaulicht, da Schülerinnen und Schüler in diesem Bereich bereits ein gewisses Vorwissen haben sollten, welches anhand von Bildungsstandards festgelegt ist.

Unter Bildungsstandards werden konkret formulierte Lernergebnisse verstanden, welche die Zielsetzung des Lehrplans konkretisieren. Es sind Kompetenzen die

Schülerinnen und Schüler erwerben sollen. Dadurch kann auch ein Lehrer leichter der Ist- und Sollzustand bestimmt werden [24].

In den folgenden Beispielen wird die Monte-Carlo Simulation gemeinsam mit den Schülerinnen und Schülern erarbeitet. Dabei wird immer nach dem selben Prinzip vorgegangen. Es gibt eine Problemstellung, diese wird mit Hilfe der Theorie erläutert und anschließend wird ein Anwendungsbeispiel mit Hilfe einer Simulation aufbereitet. Dieses kann dann im Unterricht eingesetzt werden. Dabei können Kalkulationsprogramme wie Excel oder Libre-Office-Calc verwendet werden.

### 9.3.1 Kombinatorik am Beispiel des Rebhuhnspiels

#### Problemstellung

Am Beispiel des Rebhuhnschießen wird ein Experiment aus der Kombinatorik durchgeführt. Eine Gruppe von Jägern steht die gleiche Anzahl an Rebhühnern gegenüber. Wenn jeder Jäger einen Schuss hat der immer trifft und sich jeder zufällig ein Rebhuhn aussucht, stellt sich in dem Experiment die Frage, wie viele Rebhühner getroffen werden. Dieses Experiment wird mehrfach durchgeführt und die mittlere Zahl der Treffer mit dem Ergebnis aus der Kombinatorik verglichen.

#### Theorie

Dieses Beispiel kann mit Hilfe eines Urnenmodells mit Zurücklegen verglichen werden. Dabei werden aus  $n$ - Elementen  $k$ - Elemente gezogen. Die Reihenfolge spielt hier keine Rolle. Die Anzahl  $A$  der Kombinationen kann wie folgt berechnet werden,  $A = \binom{n+k-1}{n-1}$ . Weiters ist es notwendig die Wahrscheinlichkeit  $P(R)$  zu berechnen, dass eine Kugel nicht gezogen wird. Dafür wird die Definition von Laplace verwendet. Unter der Annahme von  $n$ - Elementen ergibt sich eine Wahrscheinlichkeit mit  $P(R = n - 1) = \frac{n-1}{n}$ . Unter der Annahme von  $n = 10$ , wird im folgenden mit der Gegenwahrscheinlichkeit gerechnet, so dass für die Überlebenswahrscheinlichkeit für jedes Huhn pro  $j$  Durchgang gilt:

$$P(X = n - 1) = \left(\frac{n-1}{n}\right)^j \quad (48)$$

$$P(X = 9) = \left(\frac{9}{10}\right)^{10} \quad (49)$$

$$P(X = 9) = 0,349 \quad (50)$$

Anschließend kann der Erwartungswert  $E$ , für die erlegten Hühner pro Versuch berechnet werden. Der Erwartungswert  $E$  mit einer diskreten Zufallsvariable  $X$ ,

die die Werte  $x_i$  mit den jeweiligen Wahrscheinlichkeiten  $p_i$  annimmt kann nun anschließend definiert werden:  $E(X) = \sum_i x_i \cdot p_i$ . Für das vorher genannte Beispiel, wird mit der Gegenwahrscheinlichkeit  $q_i = 1 - p_i$  gerechnet:

$$E(X) = n \cdot (1 - p) \quad (51)$$

$$E(X) = 10 \cdot (1 - 0,349) \quad (52)$$

$$E(X) = 6,51. \quad (53)$$

### Durchführung des Experiments

Das Experiment wird mit zehn Jägern und zehn Rebhühnern durchgeführt. Es fliegen zehn Rebhühner auf und die zehn Jäger suchen sich nun zufällig jeweils ein Rebhuhn aus. Die Treffsicherheit der Jäger beträgt 100%. Die Frage ist nun, wie viele Rebhühner werden im Durchschnitt bei einem Experiment getroffen.

Die Durchführung in der Klasse gliedert sich in drei Schritte. Der erste Schritt die Diskussion, der zweite die Modellannahme und der letzte Schritt die Erzeugung von Zufallszahlen sollen mit den Schülerinnen und Schülern durchgeführt werden.

Beim ersten Schritt sollen die Schülerinnen und Schüler miteinander diskutieren und eine mögliche Lösung aufschreiben. Jedoch sollte seitens der Lehrperson noch keine Lösung verlautbart werden. Gemeinsam mit der Klasse sollen nun Modelle erarbeitet werden, um folgenden Vorgang zu simulieren: Jeder der zehn Jäger wählt zufällig einen von den zehn Hühnern aus. Dabei ist vor allem auf die vereinfachte Darstellung zu achten, dass die Treffsicherheit 100% beträgt. Weiters wählen die Jäger ihr Huhn unabhängig voneinander aus und jedes Huhn hat die gleiche Chance getroffen zu werden.

Der zweite Schritt des Experiments ist die Erarbeitung von Modellen, bei denen mit Hilfe von Zufallszahlen dieser Modellprozess realisiert wird. In diesem Fall entspricht eine Zufallszahl dem Treffen eines der zehn Rebhühner. Folgende Modelle können bei der Realisierung helfen.

1. Jeder der zehn Jäger wählt eine Zahl aus dem Bereich  $\{0,1,2 \dots 9\}$
2. Zehnmaliges Ziehen mit Zurücklegen aus einer Urne mit Kugeln von Null bis Neun
3. Zehnerblock aus eine Liste von Zufallszahlen auswählen



Unter Betrachtung der vereinfachten Darstellung fallen aber nur die Modelle zwei und drei, da bei dem ersten Modell Zahlen aufgeschrieben werden, die einem zufällig einfallen. Diese Zahlen sind dann aber keine Zufallszahlen sondern Pseudozufallszahlen. Trotzdem sollen die Schülerinnen und Schüler wissen, dass dieses Modell angewendet werden kann. Das Ergebnis erfüllt jedoch nicht die mathematischen Bedingungen die an die Generierung von Zufallszahlen gestellt werden.

Der letzte Schritt dient dazu, dass alle Schülerinnen und Schüler mitarbeiten können. Dabei bietet es sich an mit dem dritten Modell zu arbeiten und ihnen so die Idee der Monte-Carlo Simulation näher zu bringen. In Zweierteams sollen die Schülerinnen und Schüler aus einer Tabelle mit Zufallszahlen, siehe Abbildung 16, jeweils Zehnerblöcke an Zufallszahlen herausschreiben, dies gibt die Anzahl der Treffer an. Die Tabelle mit Zufallszahlen kann der Lehrkörper den Schülerinnen und Schülern zur Verfügung stellen. Jedoch ist es hierbei sehr wichtig, dass jedes Team an einer anderen Stelle der Tabelle beginnt die Zahlen zu notieren. Da sich sonst die Kombination von Zufallszahlen wiederholen könnte. Die Ziffer Null entspricht dem Dieb Nummer zehn. Dabei entspricht ein Zehnerblock einem Spieldurchgang.

Anschließend erstellt jedes Zweierteam eine Tabelle damit die gesammelten Daten auch dokumentiert werden. In der Realität entspricht die erste Spalte, dass jeder Jäger eines von den Hühnern auswählt. In der nachfolgenden Spalte werden die abgeschossenen Hühner angeschrieben und in der letzten Spalte stellt wird die Anzahl der getroffenen Hühner dargestellt. Im Fall der Modelle werden die Spalten wie folgt beschriftet, die erste Spalte steht für den Zehnerblock aus der Zufallszahlentabelle, die zweite Spalte für die vorkommenden Zahlen und die letzte Spalte für die Anzahl der verschiedenen Zahlen.

In der Tabelle 1 ist exemplarisch die Auswertung von zehn Spieldurchgängen dargestellt. Der empirische Mittelwert ergibt eine Wahrscheinlichkeit von  $P(X = 10) = 0,740$ .

Anschließend sollen die Schülerinnen und Schüler immer zehn Zehnerblöcke zusammenfassen und den Mittelwert bilden. Nun können die Werte in der Klasse verglichen werden. Das Simulationsergebnis kann mit der Wahrscheinlichkeit dieses stochastischen Problems verglichen werden.

Da in der Klasse nun circa 100 Simulationen durchgeführt wurden, sollten die Schülerinnen und Schüler erkennen, dass je mehr Simulationen durchgeführt werden, desto besser ist die Approximation durch das Zufallsexperiment.

Mit Hilfe dieser Modellannahme kann den Schülerinnen und Schülern die Idee, einen stochastischen Prozess mit Zufallszahlen darzustellen und somit der Anfang

Zufallszahlen	Getroffene Rebhühner	Anzahl der getroffenen Rebhühner
1379562099	01235679	8
1762938157	12356789	8
9898708142	0124789	7
4418675409	01456789	8
0416345928	012345689	9
4518200338	0123458	7
7444836989	346789	6
1363954466	134569	6
2950268407	02456789	8
4397213984	1234789	7

Tabelle 1: Beispiel für die Auswertung vom Experiment Rebhuhnschießen in der Klasse

der Monte-Carlo Simulation, näher gebracht werden [27]. Bei diesem Einführungsbeispiel ist es gar nicht notwendig tiefer in die Materie einzudringen, da hier nur die Modellierung mit Zufallszahlen verstanden werden soll.

3	2	0	5	1	4	4	6	5	4	1	5	9	1	1	2	6	2	3	0	2	0	5	0	5	8
9	3	0	0	0	2	4	4	2	0	8	8	0	9	4	3	1	0	1	6	5	3	5	7	9	1
5	3	2	0	0	3	4	3	5	6	9	8	7	2	9	9	6	1	7	9	4	2	8	9	2	4
5	7	8	2	9	7	8	1	8	2	3	8	9	5	6	6	9	0	4	4	9	1	1	3	6	7
8	1	4	6	8	7	7	8	4	2	8	0	7	6	5	2	6	5	2	3	8	4	6	0	0	2
1	5	6	4	1	5	1	6	3	6	2	8	6	8	5	5	9	8	2	1	6	7	6	8	9	4
7	9	4	7	5	8	6	2	4	9	7	0	0	8	5	5	0	0	6	2	3	0	5	7	2	4
4	8	9	7	6	9	5	4	2	2	4	3	7	6	2	3	4	3	0	5	8	6	2	6	8	4
8	9	4	3	1	5	9	7	6	7	4	8	4	6	5	8	3	0	2	5	7	0	3	5	9	7
5	6	1	0	5	6	3	4	0	3	4	7	7	5	0	6	0	4	3	9	8	0	1	8	6	2
3	0	0	6	5	3	6	9	1	7	4	0	4	8	5	9	0	5	8	0	0	9	1	8	3	5
0	4	6	1	4	7	7	7	1	0	8	8	3	6	3	5	9	7	3	5	1	2	9	4	2	7
9	5	6	3	5	2	0	5	2	9	0	8	2	1	7	0	8	5	7	5	8	9	2	3	6	9
4	0	1	3	1	0	9	1	9	4	0	7	7	2	9	0	4	8	0	6	9	3	3	7	7	1
3	7	4	1	1	5	6	6	5	4	2	9	8	2	7	3	2	0	8	4	5	2	0	3	7	3
9	3	0	9	0	4	2	8	1	1	3	2	4	4	8	6	5	2	2	4	7	4	7	6	6	1
9	4	4	8	4	5	1	1	1	0	7	0	8	6	8	6	3	9	9	6	1	4	8	7	4	9
9	5	0	1	8	3	9	1	6	8	1	1	5	5	9	1	8	9	0	2	9	8	2	1	2	4
8	6	9	8	8	6	0	0	4	4	8	1	4	3	8	1	7	5	4	0	7	8	5	2	3	2
9	7	1	5	3	0	6	5	4	6	2	5	1	5	7	0	7	7	1	6	4	7	8	5	8	7
5	2	1	5	8	7	2	3	9	7	7	1	6	2	8	6	4	7	9	5	9	4	0	4	2	5
8	0	5	5	0	3	8	5	3	9	9	5	0	0	5	6	6	0	3	9	0	8	9	4	5	1
8	3	8	7	0	9	4	2	7	8	5	7	6	8	8	0	9	5	5	4	5	5	4	6	2	9
3	5	9	7	7	1	8	7	8	6	1	8	8	7	8	1	4	1	2	1	1	2	5	3	7	2
5	2	7	0	0	1	1	9	7	9	6	5	2	3	2	0	2	2	6	5	4	5	7	6	8	1
6	6	9	4	2	0	7	5	1	2	5	2	2	2	8	8	0	2	1	4	4	9	4	3	7	2
2	3	0	9	2	3	3	6	6	9	3	1	6	2	0	0	6	8	2	3	2	6	0	4	7	2
7	7	1	0	0	4	8	0	2	1	2	2	7	9	5	1	8	2	1	0	8	2	0	0	9	4
1	2	4	0	3	9	2	3	8	5	0	1	7	6	5	4	4	3	7	0	7	1	8	4	2	5
7	4	0	0	5	5	6	3	0	8	1	0	8	3	6	4	5	9	4	1	0	0	9	3	4	4
9	5	2	7	4	7	5	0	0	1	8	8	4	3	4	2	1	0	9	4	4	0	7	9	1	9
5	8	5	5	4	3	5	9	9	6	5	2	6	8	9	1	2	4	6	5	8	2	7	9	1	2
5	1	2	8	6	6	6	4	3	5	7	8	5	4	6	4	4	4	3	4	0	9	4	1	5	6
9	2	9	5	6	2	3	0	0	6	2	7	8	4	5	2	7	5	0	8	5	5	6	5	1	9
0	7	9	4	3	2	4	5	4	3	5	8	8	1	2	2	9	6	8	0	0	3	6	6	7	9
2	5	4	5	5	6	4	7	3	3	6	3	8	9	0	9	6	2	0	9	4	7	4	9	1	6
2	0	6	9	6	7	1	3	4	8	7	9	8	2	7	6	4	9	8	0	7	7	1	7	3	4
4	8	9	5	8	6	4	2	2	3	5	0	8	6	0	9	6	2	9	3	9	9	0	9	2	1
1	9	7	5	4	5	5	4	1	4	4	0	1	1	1	1	2	4	0	1	3	6	0	9	6	4
1	5	0	6	8	6	6	9	0	8	0	0	7	2	8	1	5	8	6	8	0	3	5	2	1	6
3	8	2	2	3	4	8	4	7	2	6	5	8	0	5	9	7	2	0	0	9	4	4	8	2	1
8	1	4	3	3	4	1	8	3	4	4	6	1	2	5	4	9	5	5	1	8	7	7	9	8	3
9	2	7	5	4	8	4	4	9	1	5	1	8	3	5	4	2	7	0	0	2	8	2	1	9	8
1	1	2	2	2	6	4	1	8	2	4	4	1	5	4	0	1	3	1	8	9	1	0	6	3	5
2	2	6	9	6	8	3	3	2	6	0	4	6	1	1	6	3	1	3	5	4	2	0	6	2	5
3	2	8	7	3	6	0	6	0	5	0	5	2	7	5	1	8	1	5	3	6	0	1	9	9	4
2	9	4	2	8	6	0	7	5	0	8	3	7	8	0	6	0	2	6	1	5	1	9	0	7	3
6	2	3	3	2	7	7	3	6	5	0	3	0	5	0	2	8	5	3	0	7	4	9	2	0	7
3	0	6	1	6	0	5	1	9	8	8	3	1	7	3	5	6	5	4	9	7	5	7	7	2	0
1	7	1	8	8	7	0	9	2	7	3	4	1	8	0	7	7	4	4	7	9	1	8	2	4	0

Abbildung 16: Arbeitsmaterial: Tabelle mit Zufallszahlen im Intervall  $[0; 9]$  für die Erarbeitung des Modells 3

### 9.3.2 Abschätzung der Zahl $\pi$

#### Problemstellung

Wie kann die Kreisfläche mit Hilfe der Fläche eines umschriebenen Quadrats berechnet werden. Das Verhältnis dieser beiden Flächen entspricht der Kreiszahl  $\pi$ . Die gleiche Bedingung ist auch erfüllt, für ein Quadrat in das ein Viertelkreis mit dem Radius der Quadratseite eingeschrieben ist.

#### Theorie

Für das Verhältnis von Quadrat und Viertelkreis wird angenommen, dass die Fläche des Viertelkreises mit  $A_V = \frac{1}{4} \cdot r^2 \pi$  beschrieben werden kann. Die Fläche des Quadrates wird mit  $A_Q$  mit  $A_Q = r^2$  definiert. Diese beiden Definitionen werden nun in einem Verhältnis dargestellt:

$$\frac{A_V}{A_Q} = \frac{\pi \cdot r^2}{4 \cdot r^2} \quad (54)$$

$$\frac{A_V}{A_Q} = \frac{1}{4} \cdot \pi \quad (55)$$

Da ein Näherungsverfahren für  $\pi$  gesucht ist, wird nun diese Formel nach  $\pi$  umgeformt und man erhält für das Verhältnis, folgende Umformung:

$$\pi = 4 \cdot \frac{A_V}{A_Q} \quad (56)$$

#### Durchführung des Experiments

Die Durchführung dieses Anwendungsbeispiels gliedert sich in zwei Arbeitsschritte. Im ersten Schritt wird ein Experiment mit Hilfe der Schülerinnen und Schüler durchgeführt. Beim zweiten Schritt wird das Beispiel mit Hilfe von einem Kalkulationsprogramm umgesetzt.

Beim ersten Schritt wird die Klasse auf mehrere Kleingruppen aufgeteilt, die jeweils unabhängig die Experimente durchführen. Dabei bekommt jede Kleingruppe einen quadratischen Karton und 40 kleine Nudeln. Auf den quadratischen Karton sollen nun die Schülerinnen und Schüler einen Viertelkreis in das Quadrat einzeichnen. Der Radius entspricht der Länge der Quadratseite. Anschließend ist es die Aufgabe der Schülerinnen und Schüler diese 40 kleine Nudeln zufällig auf den Karton zu

werfen. Nun werden alle Nudeln gezählt, die innerhalb beziehungsweise außerhalb des Viertelkreises liegen. Bei diesem Arbeitsschritt stellt sich jedoch ein Problem. Was ist wenn eine Nudel genau auf der Linie zum Liegen kommt. Hierbei sollte wie bei den Tennisregeln vorgegangen werden. Fällt eine Nudel auf die Linie, zählt sie zur Fläche des Viertelkreises.

Zeitgleich sollte mit den Schülerinnen und Schülern besprochen werden, ob das Werfen der Nudeln den Bedingungen einer Zufallszahl entspricht. Dabei können mögliche Fehler, wie Wegspringen der Nudeln oder Werfen der Nudeln in eine Richtung, diskutiert werden.

Als Beispiel werden nun die Werte einer Gruppe heran genommen. Dabei liegen zum Beispiel 31 der 40 Nudeln innerhalb und neun außerhalb des Viertelkreises. Die Gesamtanzahl der Nudeln werden mit  $N_0 = 40$  bezeichnet. Die 31 Nudeln die innerhalb liegen werden mit  $N = 31$  definiert. Diese beiden Zahlen sollen nun in einem Verhältnis, wie in der Theorie beschrieben angeschrieben werden. Dafür wird  $\kappa$  folgendermaßen definiert:

$$\kappa = 4 \cdot \frac{N}{N_0} \quad (57)$$

$$\kappa = 4 \cdot \frac{31}{40} \quad (58)$$

$$\kappa = 3,10 \quad (59)$$

Nun kann das empirische Näherungsverfahren mit der Zahl  $\kappa$  mit dem theoretischen Näherungsverfahren von  $\pi$  verglichen werden. Dabei ist  $N$  die Approximation für  $A_V$ . Werden nun die beiden Zahlen  $\kappa$  und  $\pi$  mit einander verglichen, sollte der Klasse die Ähnlichkeit zwischen Theorie und Experiment auffallen und das obwohl nur wenige Zufallszahlen in dem Fall Nudeln verwendet wurden. Auf Grund des ersten Beispiels mit den Rebhühnern sollten die Schülerinnen und Schüler nun wissen, je mehr Realisierungen beim Zufallsprozess vorgenommen werden, desto besser ist die Approximation.

Im Folgenden wird den Schülerinnen und Schülern die Theorie hinter diesem Experiment erklärt:

Als Ausgangsposition wird der 1. Quadrant mit Länge  $l = 1$  angenommen. Es wird ein Quadrat ebenfalls mit der Seitenlänge  $l = 1$  konstruiert. Nun wird in dieses Quadrat ein Viertelkreis mit Radius  $r = 1$  eingeschrieben (Abbildung 17).

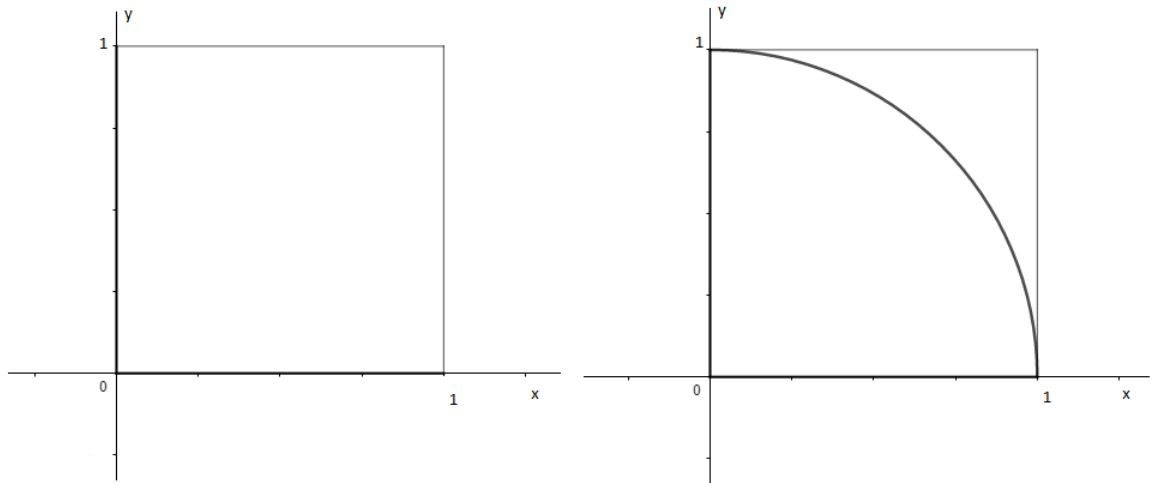


Abbildung 17: 1.Quadrant mit Quadrat mit Seitenlänge  $l = 1$  und Viertelkreis mit Radius  $r = 1$

Um nun die Verbindung zwischen Experiment und Theorie mathematisch darzulegen, es ist notwendig wichtige Variablen mit einander in Verbindung zu bringen. Dabei gilt folgendes,  $A_V \hat{=} N$  und  $A_Q \hat{=} N_0$ .

$$A_{\text{Quadrat}} = r^2 \quad (60)$$

$$A_{\text{Kreis}} = r^2 \cdot \pi \quad (61)$$

$$A_{\text{Viertelkreis}} = \frac{r^2 \cdot \pi}{4} \quad (62)$$

Da  $r = 1$  gilt, ist folgendes für den Flächeninhalt des Viertelkreises bekannt:

$$A_{\text{Viertelkreis}} = \frac{\pi}{4} \quad (63)$$

Anschließend wird das Quadrat mit dem eingeschriebenen Viertelkreis mit beliebigen Punkten "berechnet". Unter diesen beliebigen Punkten werden gleichverteilte Zufallszahlen verstanden. In diesem Fall sind es Zufallspaare, da sowohl der x-Achse als auch der y-Achse jeweils eine Zufallszahl zugeordnet wird. Diese Zufallszahlen müssen zwischen Null und Eins liegen, da der Radius  $r$  eins ist. Nun muss überprüft werden, ob diese Zufallspaare innerhalb oder außerhalb des Viertelkreises liegen. Da dieses Zufallspaar ein Punkt im Koordinatensystem darstellt gilt für

diesen Punkt,  $P = (ZZ_1|ZZ_2)$ . Unter  $ZZ_1$  und  $ZZ_2$  werden zwei verschiedene gleichverteilte Zufallszahlen im Intervall  $[0; 1]$  verstanden. Dafür lassen sich zwei Fälle unterscheiden.

Beim ersten Fall liegt der Punkt innerhalb des vorgegeben Radius, mit  $x^2 + y^2 \leq r^2$ . Ist der Wert jedoch größer, tritt der zweite Fall ein, dann liegt der Punkt nicht mehr im Viertelkreis, da  $x^2 + y^2 > r^2$ . Diese beiden Fälle lassen sich in der Abbildung 18 gut erkennen.

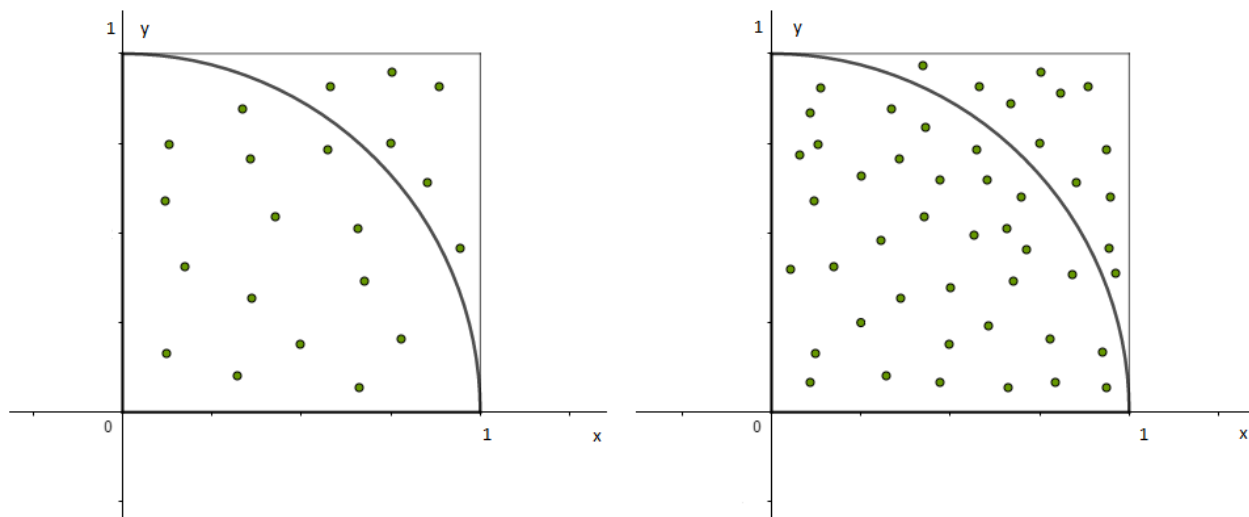


Abbildung 18: "Berechnung" mit Zufallszahlen

Wie schon beim Experiment werden alle Zufallszahlen die innerhalb des Viertelkreises liegen als  $N$  und die Summe aller Zufallszahlen als  $N_0$  bezeichnet. Die Anzahl der Punkte  $N_0$  und  $N$  werden als Approximation der beiden Flächen  $A_Q$  und  $A_V$  angenommen. Um nun die Punkte im Viertelkreis berechnen zu können, wird die Fläche des Viertelkreises mit der vorher genannten Formel  $A_V = \frac{r^2 \cdot \pi}{4} = 0,785$  gelöst. Daher ist nun bekannt, dass 78,5% der Punkte von  $N_0$  im Viertelkreis liegen. Dafür gilt folgende Formel:  $N = 0,785 \cdot N_0$ . Um eine Schätzung durchführen zu können, wird diese Formel umgeformt:  $\kappa = \frac{4N}{N_0}$ . So kann der enge Zusammenhang zwischen Theorie und Zufall dargestellt werden.

Nach dem Experiment mit den Nudeln wird die Approximation der Kreiszahl  $\pi$  mit Hilfe von einem Kalkulationsprogramm noch einmal durchgeführt. Dabei werden gleichverteilte Zufallszahlen im Intervall  $[0; 1]$  in Form vom Koordinaten produziert. Jene Zahlenpaare  $(ZZ_1|ZZ_2)$  die außerhalb vom Kreisbogen liegen, fallen weg, da diese für die Simulation keine Bedeutung mehr haben. Jedoch werden die Koordinatenpaare, welche innerhalb vom Radius liegen, direkt in die Grafik übernommen.

Anhand dieses Beispiels wird zum ersten Mal eine Simulation mit einem Kalkulationsprogramm den Schülerinnen und Schüler näher gebracht und so wird auch ein wichtiges Verständnis für die Monte-Carlo Simulation geboten.



### 9.3.3 Random Walk: Das Herumirren in einer Stadt

#### Problemstellung

Der Random Walk ist die mathematische Beschreibung der Brownschen Molekularbewegung. Dabei bewegen sich Moleküle rein zufällig durch zahlreiche Zusammenstöße mit anderen Molekülen. Anhand eines Experimentes wird gezeigt, dass die zufällige Bewegung auf einem Gitternetz, mit dem Random Walk verglichen werden kann. Solche Zufallsbewegungen werden anschließend mit dem Computer simuliert [43].

#### Theorie

Das einfachste Modell des Random Walks reduziert sich auf eine Dimension mit einer fixen Schrittlänge  $a$ . Wobei mit gleicher Wahrscheinlichkeit ein Schritt nach links oder rechts gemacht wird. Dabei werden die einzelnen Schritte einer Folge von Zufallsvariablen  $d_1, d_2, d_3, \dots$  zugeordnet. Diese Folge entspricht den einzelnen Schritten, welche jeweils unabhängig die Wahrscheinlichkeit  $P(d) = \frac{1}{2}$  annehmen. Diese Zufallsbewegung entsteht durch die Aufsummierung dieser Zufallsvariablen.

Als Startwert wird  $x_0$  mit  $x_0 = 0$  angenommen. Unter  $x_0, x_1, x_2, \dots$  werden die einzelnen Positionen verstanden.

$$x_1 = \quad \quad \quad x_0 + d_1 = \quad \quad \quad d_1 \quad (64)$$

$$x_2 = \quad \quad \quad x_1 + d_2 = \quad \quad \quad d_1 + d_2 \quad (65)$$

$$x_3 = \quad \quad \quad x_2 + d_3 = \quad \quad \quad d_1 + d_2 + d_3 \quad (66)$$

Die Position nach  $n$ - Schritten, wird wie folgt berechnet:

$$x_n = d_1 + d_2 + \dots + d_n = \sum_{j=1}^n d_j \quad (67)$$

[44].

Dabei entspricht jedes  $x_n$  einer Funktion der Zufallsvariablen  $d_1, d_2, d_3, \dots$ .

#### Durchführung des Experiments

Das Experiment wird mit Hilfe des Random Walks durchgeführt. Dabei wird zuerst mit den Schülerinnen und Schüler das Problem eines desorientierten Fußgängers in einem Straßengitter erarbeitet und analysiert und anschließend wird mit Hilfe von einem Kalkulationsprogramm solch ein Random Walk simuliert. Abschließend wird

noch die Frage, wie weit bewegt sich im Mittel ein orientierungsloser Fußgänger vom Startpunkt weg, wenn dieser  $n$ - Schritte zurücklegt bearbeitet.

Die Durchführung des ersten Schrittes findet in Zweierteams statt. Jedes Team bekommt ein Stadtplan von Mannheim mit einem integrierten Koordinatensystem (siehe Abbildung 19) sowie einen Würfel. Anschließend wird der Klasse folgendes erklärt:



[56]

Abbildung 19: Stadtplan von Mannheim mit eingezeichneten Koordinatengitter sowie gelb markierten Startpunkt

Ein desorientierter Fußgänger möchte die Stadt Mannheim erkunden. Der Straßenplan dieser Stadt gleicht einem symmetrischen Straßengitter. Da dieser Fußgänger orientierungslos ist, wählt er an jeder Kreuzung rein zufällig und gleichwahrscheinlich einer der vier Richtungen aus. Startpunkt ist der gelbe Punkt, welcher den Nullpunkt des Koordinatengitters symbolisiert. Die rein zufällige Wahl der Richtung wird mit einem Würfel simuliert. Dabei entspricht die Zahl Eins einem Zug nach rechts bis zur nächsten Kreuzung, Zahl Zwei einem Zug nach links bis zur nächsten Kreuzung, Zahl Drei einem Zug nach oben bis zur nächsten Kreuzung und das Würfeln der Zahl Vier entspricht einem Zug nach unten bis zur nächsten Kreuzung. Die Zahlen Fünf und Sechs werden verworfen. Jedes Zweierteams soll circa zehn Simulationen vornehmen und diese Schritte beziehungsweise Züge entlang eines Häuserblocks einzeichnen. Abschließend können die einzelnen Stadtpläne, mit den jeweiligen Endpunkten, mit einander verglichen werden.

Im zweiten Schritt wird diese Simulation mit Hilfe von einem Kalkulationsprogramm durchgeführt. Dabei werden in zwei Spalten Zufallszahlen im Intervall  $[0; 3]$  in Form von Koordinaten produziert. Ist solch eine Zufallszahl gerade verläuft die Bewegung entlang der x-Achse, und umgekehrt entlang der y-Achse. Anschließend wird noch zwischen links und rechts unterschieden. Für einen Schritt auf der x-Achse werden die Zufallszahlen Null und Zwei angenommen. Wobei Null einem Schritt nach links, also  $0 \hat{=} -1$  und Zwei ein Schritt nach rechts mit  $2 \hat{=} 1$  bedeutet. Analog dazu die Schritte auf der y-Achse. Dabei entspricht  $1 \hat{=} -1$  und  $3 \hat{=} 1$ . Diese Koordinatenpunkte werden dann mit Hilfe eines Punktdiagramms graphisch dargestellt. In den folgenden Abbildungen werden einige Realisierungen vorgenommen. Dabei werden jeweils unterschiedliche Anzahlen an Schritten/Zügen vorgenommen. In Abbildung 20 werden 100 Züge mit  $n = 100$  realisiert. In den Abbildungen 21 werden 1000 Züge und in Abbildung 22 10000 Züge graphisch dargestellt.

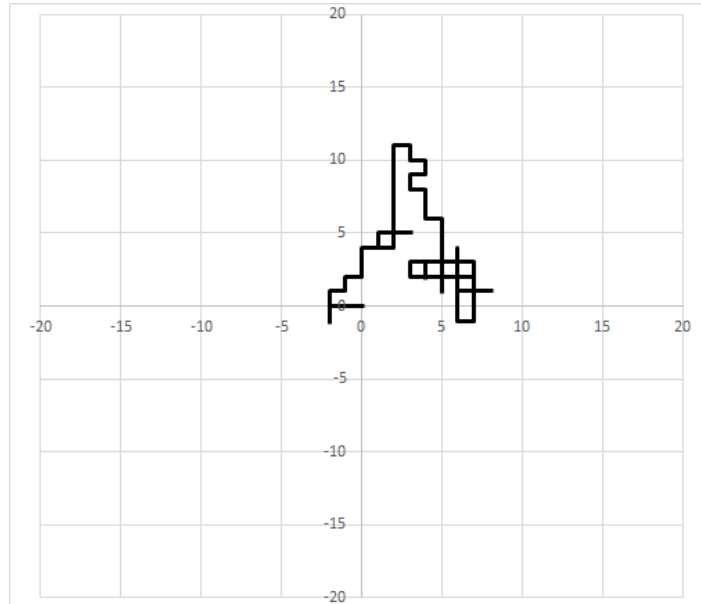


Abbildung 20: Graphische Darstellung eines Random Walks mit Hilfe eines Kalkulationsprogramms mit  $n = 100$

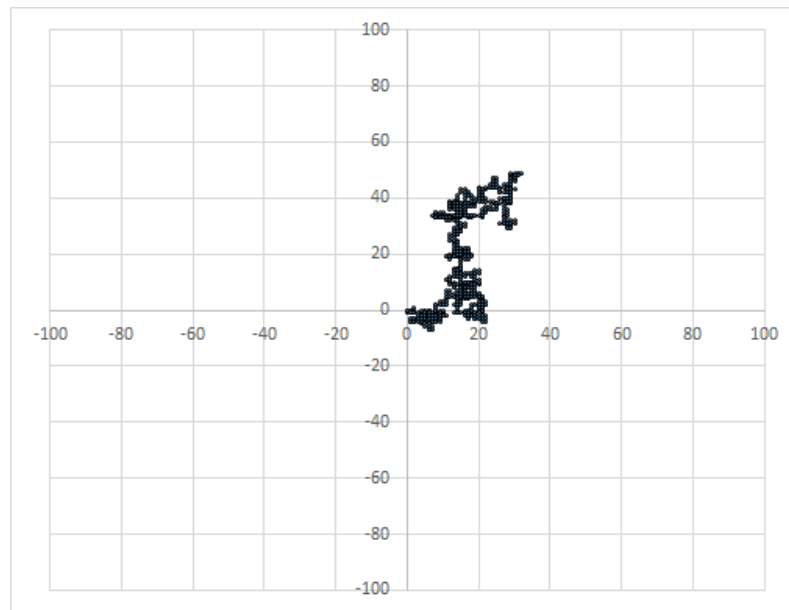


Abbildung 21: Graphische Darstellung eines Random Walks mit Hilfe eines Kalkulationsprogramms mit  $n = 1000$

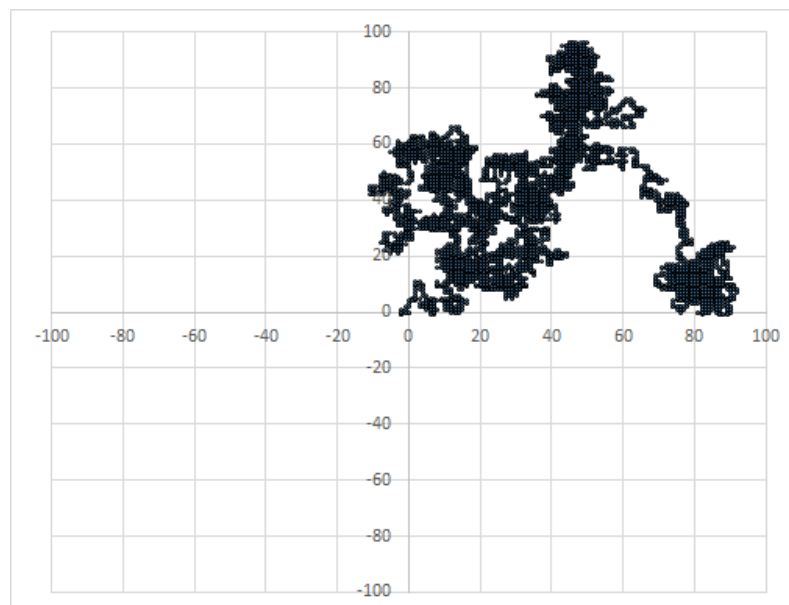


Abbildung 22: Graphische Darstellung eines Random Walks mit Hilfe eines Kalkulationsprogramms mit  $n = 10000$

Eine abschließende Frage kann nun noch gemeinsam mit den Schülerinnen und

Schülern bearbeitet werden. Der Fußgänger ist irgendwann so erschöpft, dass er sich an einer Hausmauer anlehnen muss. Wie weit bewegt sich im Mittel ein orientierungsloser Fußgänger vom Startpunkt weg, wenn dieser  $n$ - Wohnblöcke beziehungsweise Züge zurücklegt?

Dabei muss zwischen einer einfachen Zufallsbewegung mit  $n$ - Zügen (Random Walk) und einer zielgerichteten Bewegung mit  $n$ - Zügen unterschieden werden. Findet ein Random Walk statt, ist der Weg proportional zu  $\sqrt{n}$ . Im Gegensatz dazu, ist bei einer zielgerichteten Bewegung, der Weg in  $n$ - Zügen proportional zu dieser Anzahl von Schritten. Daher ist der Weg eines Random Walks deutlich geringer [43].

### 9.3.4 Vergleich von empirischen Daten mit der Normalverteilung

#### Problemstellung

Die Normalverteilung ist eine wesentliche, stetige Verteilung in der Statistik. Eine solche Verteilung ergibt sich aus der Summe von mehrerer Zufallsvariablen, die einer beliebigen Verteilung genügen. In einem Experiment wird gezeigt, dass die Summe aus sechs gleichverteilten Zufallsvariablen einer normalverteilten Zufallsvariable entspricht.

#### Theorie

Die Zufallsvariable  $Z_n = Z_1 + Z_2 + \dots + Z_n$  sei die Summe von  $n$  unabhängigen und identisch verteilten Zufallsvariablen  $Z_i$  [20, S. 288]. Der Erwartungswert von  $Z_n$  ist  $E(Z_n) = n\mu$  und die Varianz ist mit  $V(Z_n) = n\sigma^2$  definiert. Daraus wird die standardisierte Zufallsvariable mit  $S_n = \frac{Z_n - n\mu}{\sigma\sqrt{n}}$  gebildet. Nun konvergiert die Verteilungsfunktion  $S_n$  für  $n \rightarrow \infty$  punktweise gegen die Verteilungsfunktion der Standardnormalverteilung  $\mathcal{N}(0; 1)$  [40].

Dies besagt, dass die Summe unabhängiger Zufallsvariablen, die alle die gleiche Verteilung haben, näherungsweise normalverteilt sind. Je größer die Anzahl der Summanden, desto besser ist die Näherung.

#### Durchführung des Experiments

Im folgenden wird ein Experiment durchgeführt, welches einerseits sehr gut in der Klasse und andererseits auch mit einem Kalkulationsprogramm veranschaulicht werden kann. Diese zwei Schritte werden im folgenden analysiert. Zuerst wird die Idee mit den Schülerinnen und Schüler erarbeitet, dann werden Daten gesammelt und graphisch dargestellt. Im zweiten Schritt wird das Experiment mit Hilfe von Zufallszahlen simuliert und zuletzt mit der theoretischen Verteilung verglichen.

Bei der Durchführung des ersten Schritten werden Zweierteams innerhalb der Klasse gebildet. Dabei bekommt jedes Zweierteam sechs Würfel und soll diese sechs Würfel immer gemeinsam zehn Mal werfen. Nach jedem Wurf soll die Summe der Augenzahlen der sechs Würfel gebildet werden. Damit es für die Schülerinnen und Schüler übersichtlicher wird, sollten geeignete Intervalle herangezogen werden, damit eine übersichtliche Tabelle mit den Häufigkeitsverteilungen gestaltet werden kann. Nun können alle Ergebnisse an der Tafel gesammelt werden. Aus all diesen Werten kann nun das arithmetische Mittel  $\mu$  und die Standardabweichung  $\sigma$  gebildet werden. Die gesammelten Werte können anschließend graphisch dargestellt werden. Das Ergebnis eines solchen Experiments ist in Abbildung 23 dargestellt.

Nun erfolgt der Übergang zu Schritt zwei. Damit die Idee der Monte-Carlo Si-

mulation vertieft werden kann, wird das selbe Prozedere mit Hilfe des Computers durchgeführt. Bei diesem Beispiel muss den Schülerinnen und Schülern jedoch noch der Unterschied zwischen empirischer Verteilung und theoretischer Verteilung erklärt werden.

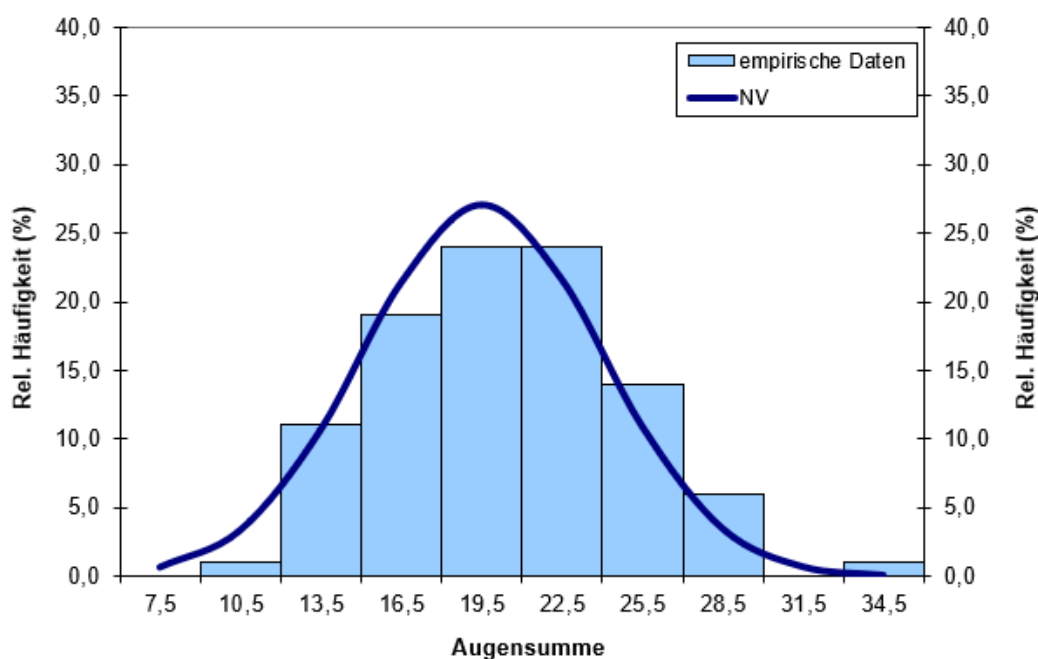


Abbildung 23: Vergleich der empirischen Daten mit der Normalverteilung, mit dem Mittelwert  $\bar{x} = 21$  und der Standardabweichung  $\sigma = 4,4$

Abschließend kann nun die Grafik, siehe Abbildung 23 mit den empirischen Werten und der Normalverteilung verglichen werden. Hierbei spielen bei der Normalverteilung der Mittelwert  $\bar{x} = 21$  und die Standardabweichung  $\sigma = 4,4$  eine wesentliche Rolle. Die Häufigkeiten der Normalverteilung berücksichtigt die Klassen in diesem Fall nicht.

Im nächsten Schritt wird die Verbindung zur Monte-Carlo Simulation erarbeitet. Die Würfel produzieren gleichverteilte Zufallszahlen, diese werden durch einen Zufallszahlengenerator ersetzt. Anhand solcher Beispiele kann gezeigt werden, dass komplexe Beispiele oftmals einfach umzusetzen sind und sie damit auch in der Klasse anwendbar sind.

### 9.3.5 Ableitung der Exponentialfunktion aus empirischen Daten

#### Problemstellung

Unter Radioaktivität versteht man den spontanen Zerfall instabiler Atomkerne. Der zeitliche Verlauf der verbleibenden Atomkerne  $N$  wird mit der Exponentialfunktion  $N = N_0 \cdot e^{-\lambda t}$  berechnet, wobei  $N_0$  die Anzahl der Atomkerne zum Zeitpunkt  $t = 0$  und  $\lambda$  die Zerfallskonstante ist. Anhand eines Experiments wird gezeigt, dass dieser physikalischer Vorgang mit Hilfe von Zufallszahlen approximiert werden kann.

#### Theorie

Beim radioaktiven Zerfall ist die Änderung (Abnahme)  $dN$  der Kerne im Zeitintervall  $dt$  durch die Beziehung  $dN = -\lambda \cdot N \cdot dt$  gegeben. Als Anfangsbedingung ist die Zahl der Kerne  $N_0$  zur Zeit  $t = 0$  gegeben.

Anhand der ursprünglichen Gleichung kann das Gesetz für den radioaktiven Zerfall hergeleitet werden [37].

$$dN = -\lambda \cdot N \cdot dt \quad (68)$$

$$\int_N dN = \int_t -\lambda \cdot N \cdot dt \quad (69)$$

$$\int_N dN = -\lambda \int_t N \cdot dt \quad (70)$$

$$\int_N \frac{dN}{N} = -\lambda \int_t dt \quad (71)$$

$$\int_N \frac{1}{N} \cdot dN = -\lambda \int_t 1 \cdot dt \quad (72)$$

Durch lösen dieses Integrals, gilt dann folgende Gleichung für die Anzahl der Atome beim radioaktiven Zerfall:

$$N(t) = N_0 \cdot e^{-\lambda t} \quad (73)$$

#### Durchführung des Experiments

Um zu zeigen wie die Monte-Carlo Simulation mit Hilfe einer Exponentialfunktion durchgeführt werden kann, zeigt das Beispiel des radioaktiven Zerfalls mit Hilfe von



Zufallszahlen. Dieses Experiment wird wieder in zwei Schritte gegliedert. Zuerst die Erarbeitung und Durchführung des Experiments in der Klasse und anschließend die Berechnung mit einem Kalkulationsprogramm.

Damit Schritt eins wieder mit den Schülerinnen und Schülern erarbeitet werden kann, sollte allen die Bedeutung und Rechenweise der Exponentialfunktion mit  $N(t) = N_0 \cdot e^{-\lambda t}$  klar sein. Anfangs müssen alle Schülerinnen und Schüler aufstehen und jeder bekommt einen Würfel. Nun soll mit Hilfe der Klasse ein Zerfall dargestellt werden. Jeder einzelne aus der Klasse kann als einzelnes Atom gesehen werden, die Gesamtanzahl der Schülerinnen und Schüler wird als  $N_0$  bezeichnet. Damit die Daten erfasst und dokumentiert werden, wird ein Koordinatensystem gezeichnet. Auf der y-Achse wird  $N_0$  als die Anzahl der Schülerinnen und Schüler und auf der x-Achse wird die Anzahl der Zeitschritte markiert. Nun stehen alle auf und würfeln all jene die die Augenzahl sechs gewürfelt haben, müssen sich hinsetzen. Das heißt sie sind zerfallen. Nun wird die Anzahl der übrigen Schülerinnen und Schüler nach dem ersten Zeitschritt in das Koordinatensystem eingetragen. Nun wird wieder gewürfelt. All jene mit einer Augensumme zwischen eins und fünf dürfen stehen bleibe

Wenn nach mehreren Zeitschritten nur mehr wenige Schülerinnen und Schüler ( $\sim 10$ ) übrig bleiben, dann sollte darauf hingewiesen werden, dass der weitere Zeitverlauf stark von der Exponentialfunktion abweichen kann, siehe Abbildung 24.

Angenommen 24 Schülerinnen und Schüler nehmen an dem Experiment teil, siehe Abbildung 24. Dann ist  $N_0 = 24$ . Nach dem ersten Durchgang, nach einem Zeitschritt, stehen noch 21 Schülerinnen und Schüler und nach einem weiteren Durchgang stehen nur noch 18 Schülerinnen und Schüler. Dies Vorgang kann wie folgt beschrieben werden:

$$\delta N_0 = -\lambda N_0 \cdot \delta t \quad (74)$$

$$\delta N_1 = -\lambda(N_0 - \delta N_0)\delta t \quad (75)$$

$$\Rightarrow N_1 = N_0 - \delta N_0 \quad (76)$$

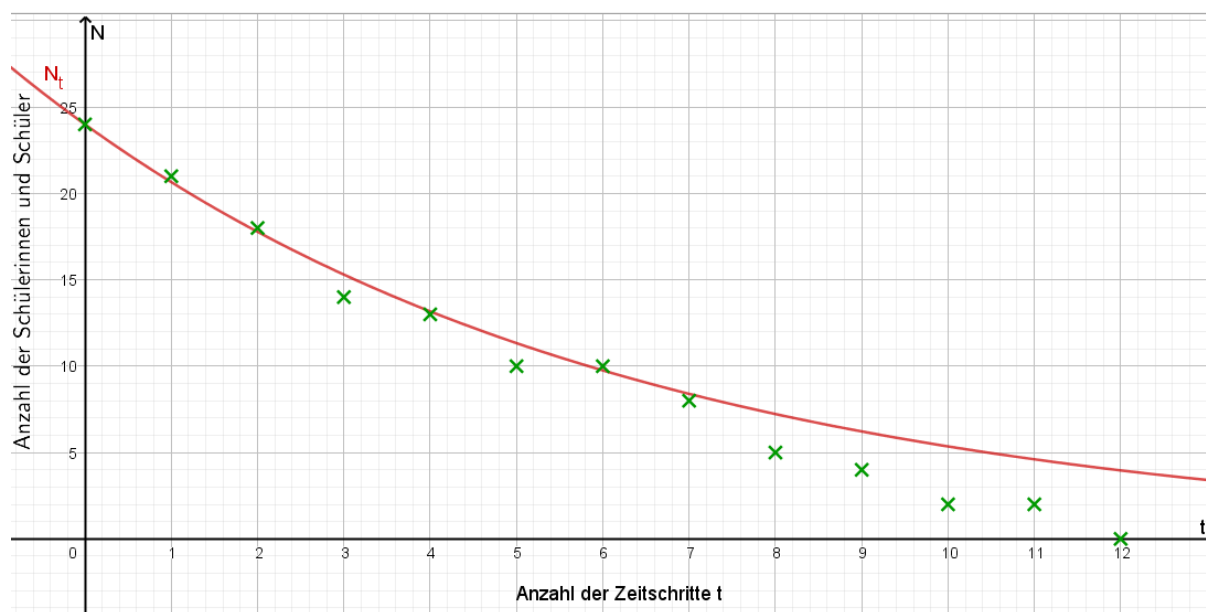


Abbildung 24: Beispiel für ein Experiment in der Klasse: Anzahl der verbleibenden Schülerinnen und Schüler  $N$  abhängig vom Zeitschritt  $t$

Nun werden die Ergebnisse der empirischen Verteilung mit der theoretischen Verteilung verglichen. Dabei wird, wie schon oben erwähnt, mit der Exponentialfunktion gearbeitet. Um mit der Funktionsgleichung  $N(t) = N_0 \cdot e^{-\lambda t}$  zu arbeiten, sind die Parameter  $N_0$ , als Gesamtzahl der Atome und  $\lambda$  als Zerfallskonstante anzunehmen. Beim Kalkulationsprogramm wird  $N_0 = 100$  und  $\lambda = \frac{1}{6}$  pro Zeitschritt angenommen. In dem Fall entscheidet nicht ein Sechser, ob ein Atom zerfällt, sondern die Zerfallskonstante. Hier werden, so wie bei der empirischen Vorgangsweise, auch einige Versuche als Wiederholungen nötig sein, um eine Verteilung zu bekommen. Unter den Bedingungen, dass eine produzierte Zufallszahl kleiner als 0,85, dann existiert diese weiterhin. Ist die Zufallszahl größer als 0,85, zerfällt diese und scheidet aus. Nun wird dieser Versuch bei allen Zufallszahlen, die noch existieren, wiederholt. Es wird eine neue Zufallszahl produziert und anschließend fällt wieder die Entscheidung zerfällt oder zerfällt nicht. Dies kann beliebig oft wiederholt werden. Damit auch hier den Schülerinnen und Schülern der Unterschied zwischen theoretischer und empirischer Verteilung vor Augen geführt werden kann, ist es hilfreich den Vorgang graphisch darstellen zu lassen (Abbildung 24).

In Abbildung 25 wird ein Vergleich eines Experiments (grüne Kreuze) mit Zufallszahlen und dem radioaktiven Zerfallsgesetz (rote Linie) für  $N_0 = 24$  und  $\lambda = \frac{1}{6}$  dargestellt. Das Experiment kann mit Würfeln durchgeführt werden.

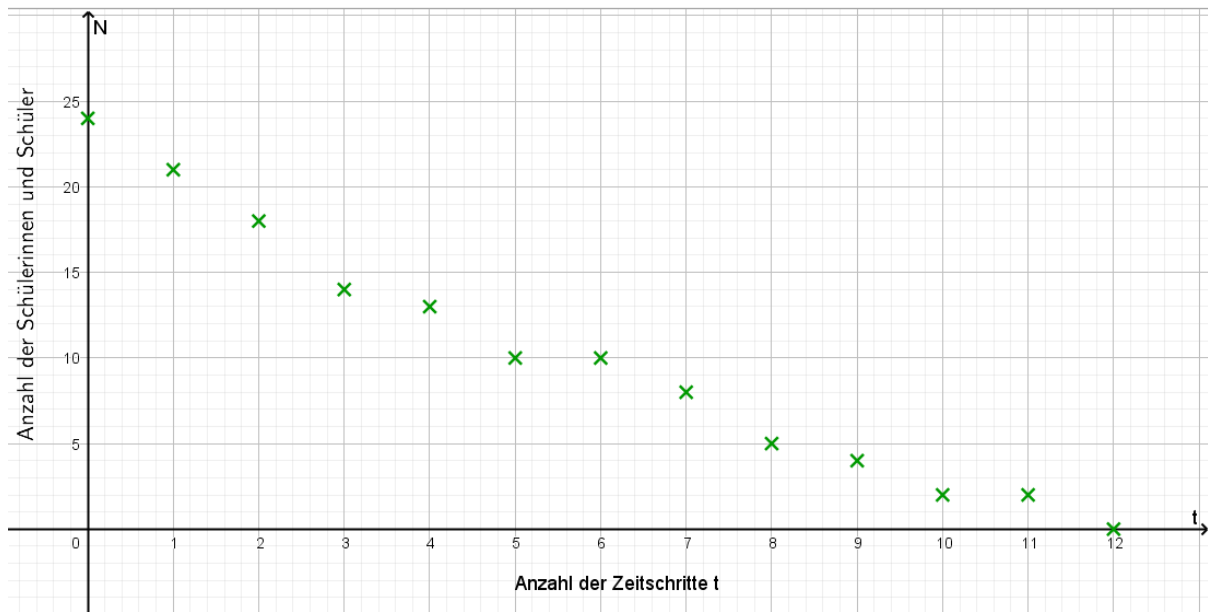


Abbildung 25: Darstellung eines Experiments (grüne Kreuze) mit Würfeln und dem radioaktiven Zerfallsgesetz (rote Linie) mit  $N_0 = 24$  und  $\lambda = \frac{1}{6}$

## 10 Zusammenfassung

Die Monte-Carlo Simulation ist nicht nur in der Mathematik, sondern auch in den Natur- und Wirtschaftswissenschaften weit verbreitet. Damit können einerseits stochastische Probleme (z.B. Value at Risk) und andererseits deterministische Fragestellungen (z.B. Strahlungstransfer durch die Atmosphäre) berechnet werden.

Anhand einiger einfacher stochastischer Experimente mit Münze und Würfel werden die Grundzüge der Monte-Carlo Simulation exemplarisch dargestellt. Dazu wird der notwendige Formelapparat und die theoretischen Grundlagen für den Lehrkörper behandelt. Darauf aufbauend werden die stochastischen Experimente erklärt und exemplarisch diskutiert.

In der Diplomarbeit werden ausgehend von den relevanten Begriffen der Wahrscheinlichkeitstheorie die diskreten und stetigen Zufallsvariablen und deren Verteilungen behandelt. Weiters werden noch ausgewählte Wahrscheinlichkeitsverteilungen erläutert. Außerdem wird der Begriff der Zufallszahl, sowie die Erzeugung von Zufallszahlen mit Hilfe von Generatoren behandelt. Neben physikalischen Verfahren werden auch mathematische Methoden beschrieben, um Zufallszahlen beziehungsweise Pseudozufallszahlen zu erzeugen. Darauf folgend werden unterschiedliche Tests von Zufallszahlen, mit deren Hilfe es möglich ist die theoretischen Anforderungen zu überprüfen, bearbeitet. Weiters werden verschiedenen Transformationen von Zufallszahlen behandelt. Diese beschäftigen sich damit, wie gleichverteilte Zufallszahlen in beliebig verteilte Zufallszahlen übergeführt werden können.

Im letzten Abschnitt wird die Simulation von Zufallszahlen im Unterricht, sowie Anwendungsbeispiele im Unterricht behandelt. Dabei werden verschiedene Fragestellungen mit unterschiedlichen Schwierigkeitsgraden aufbauend erklärt. Solch ein Anwendungsbeispiel gliedert sich in folgende drei Bereiche: Problemstellung, Theorie und Durchführung des Experiments. Bei der Durchführung wird einerseits das Experiment in der Klasse beschrieben und andererseits die exemplarische Auswertung mit Hilfe eines Tabellenkalkulationsprogramms dargestellt.

## 11 Abstract

The Monte-Carlo Simulation is not only prevalent in the world of mathematics, but it is also widespread in other natural sciences and economics. Stochastic problems (e.g. value-at-risk calculations) as well as deterministic questions (e.g. radiation through the atmosphere) can be therewith resolved.

The fundamentals of the Monte-Carlo Simulation are here illustrated through simple stochastic experiments with coins and dice. Underlying formulas and theoretical principles for the use by teachers and lecturers are covered and the stochastic experiments explained and discussed.

The relevant terms and concepts of probability theory, discrete and continuous random variables and their distribution are studied in this thesis, with selected probability distributions examined in more detail. The term of random variable and the generation of random variables are illustrated with the use of generators. The generation of random variables and pseudo-random variables are described through physics and mathematical methods. Various random variable tests validating theoretical requirements are subsequently examined. Different transformations of random variables explain how uniformly distributed random variables can be transferred into randomly distributed random variables.

The last section illustrates the simulation of random numbers as well as application examples and question formulations with various levels of difficulties used in teaching. Such application examples can be divided into three steps: formulation of the problem, theory and the realisation of the experiment. The realisation of the experiment in class is described and the exemplary interpretations with help of a spreadsheet program showcased.

## Literatur

- [1] <http://library.lanl.gov/cgi-bin/getfile?00326886.pdf>, 14.Jänner 2016
- [2] <https://de.wikipedia.org/wiki/Manhattan-Projekt>, 14.Jänner 2016
- [3] [http://www.nobelprize.org/nobel\\_prizes/physics/laureates/1938/fermi-bio.html](http://www.nobelprize.org/nobel_prizes/physics/laureates/1938/fermi-bio.html), 14.Jänner 2016
- [4] <http://www.atomicarchive.com/Bios/Ulam.shtml>, 14. Jänner 2016
- [5] <http://people.cs.ubc.ca/nando/papers/mlintro.pdf>, 15. Jänner 2016
- [6] <http://www.deutsche-biographie.de/pnd118770314.html>, 14. Jänner 2016
- [7] <http://www.mathepedia.de/Monte-Carlo-Methode.aspx>, 14. Jänner 2016
- [8] P. Raith, *Vorlesung Stochastik*, Universität Wien, 2013.
- [9] R. Kosfeld, *Statistik II*, Universität Kassel,
- [10] [https://www.uni-frankfurt.de/63287925/Einführung\\_in\\_die\\_Numerik](https://www.uni-frankfurt.de/63287925/Einführung_in_die_Numerik), Jänner 2016.
- [11] H. Hinrichsen, *Computational Physics*, Universität Würzburg, Februar 2014.
- [12] G. Tutz, *Kleines Kompendium nützlicher Verteilungen*, LMU München, Jänner 2008.
- [13] E. Limpert; G. Ohmayer; W.A. Stahel *Eine Grundfrage der Datenanalyse: Addiert oder multipliziert die Natur?*, Seminar for Statistics, ETH Zürich.
- [14] M. Hudec, *Statistik 2 für SoziologInnen Zentraler Grenzwertsatz*, Universität Wien.
- [15] [https://www.psychologie.hu-berlin.de/de/institut/kabinett/ausstellungen/galton\\_brett](https://www.psychologie.hu-berlin.de/de/institut/kabinett/ausstellungen/galton_brett). 12.März 2016.
- [16] <https://www.informatik.hu-berlin.de/de/forschung/gebiete/algorithmenII/Lehre/w> November 2019.
- [17] *Meyers Großes Taschenlexikon 20*, Bibliographisches Institut Mannheim, Wien, Zürich, 1981.
- [18] F. Härtel, *Zufallszahlen für Simulationsmodelle Vergleich und Bewertung verschiedener Zufallszahlengeneratoren*, Difo- Druck GmbH, Bamberg, 1994.
- [19] <http://ls4-www.cs.tu-dortmund.de/download/LehreMaterialien/MAO2011/Kap7.pdf> 18.Jänner 2016

- [20] A. Büchter; H.W. Henn, *Elementare Stochastik*- Eine Einführung in die Mathematik der Daten und des Zufalls, Springer- Verlag Berlin Heidelberg, 2005.
- [21] R. Walter, *Einführung in die Analysis 1*, Walter de Gruyter GmbH & Co.KG, Berlin, 2007.
- [22] U.P. Tietze; M. Klika; H. Wolpers, *Mathematikunterricht in der Sekundarstufe II*- Band 1: Fachdidaktische Grundfragen- Didaktik der Analysis, Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig / Wiesbaden, 1997.
- [23] [https : //itp.tugraz.at/MML/MonteCarlo/MCIntro.pdf](https://itp.tugraz.at/MML/MonteCarlo/MCIntro.pdf), März 2016.
- [24] [https : //www.bifie.at/bildungsstandards](https://www.bifie.at/bildungsstandards), März 2016.
- [25] [https : //www.bmbf.gv.at/schulen/unterricht/lp/lp\\_neu\\_a\\_h\\_s\\_14\\_1866.pdf?4dzgm2](https://www.bmbf.gv.at/schulen/unterricht/lp/lp_neu_a_h_s_14_1866.pdf?4dzgm2), April 2016.
- [26] [https : //www.bifie.at/node/81](https://www.bifie.at/node/81), April 2016.
- [27] [https : //www.stochastik - schule.de/sisonline/struktur/jahrgang10 - 90/heft2/1990 - 2Tra.pdf](https://www.stochastik-schule.de/sisonline/struktur/jahrgang10-90/heft2/1990-2Tra.pdf), September 2019.
- [28] J. Schwarze, *Grundlagen der Statistik II* Wahrscheinlichkeitsrechnung und induktive Statistik, Neue Wirtschafts-Briefe, Herne/Berlin, 2001.
- [29] E. Limpert; W.A. Stahel; M. Abbt, *Log-normal Distributions across the Sciences: Keys and Clues* On the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: That is the question, *Bioscience* 51, 2001.
- [30] R.B. Ferri; W. Blum, *Lehrerkompetenzen zum Unterrichten mathematischer Modellierung* Konzepte und Transfer, Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature, 2018.
- [31] K. Meyberg; P. Vachenauer, *Höhere Mathematik 1* Differential- und Integralrechnung, Vektor- und Matrizenrechnung, Springer-Verlag Berlin Heidelberg, 1997.
- [32] C.D. Schönwiese, *Praktische Statistik* für Meteorologen und Geowissenschaftler, Gebr. Borntraeger Verlagsbuchhandlung, Stuttgart/Berlin, 2013.
- [33] [https : //statistikguru.de/rechner/box - cox.html](https://statistikguru.de/rechner/box-cox.html), Dezember 2019.
- [34] [https : //www.mql5.com/de/articles/363](https://www.mql5.com/de/articles/363), Dezember 2019.

- [35] E.K. Harris; J.C. Boyd, *Statistical bases of reference values in laboratory medicine*, Marcel Dekker, New York, 1995.
- [36] <https://de.wikipedia.org/wiki/Verwerfungsmethode>, Dezember 2019.
- [37] <https://me-lrt.de/a22-differentialgleichung-radioaktiver-zerfall-zerfallskonstante>, Dezember 2019.
- [38] <https://de.wikipedia.org/wiki/Zufallsvariable>, Jänner 2020.
- [39] H. Bossel, *Systemdynamik Grundwissen, Methoden und BASIC-Programme zur Simulation dynamischer Prozesse*, Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig, 1987.
- [40] [https://de.wikipedia.org/wiki/Zentraler\\_Grenzwertsatz](https://de.wikipedia.org/wiki/Zentraler_Grenzwertsatz), Dezember 2019.
- [41] <https://de.wikipedia.org/wiki/Zufallszahlengenerator>, Jänner 2020.
- [42] K.H. Waldmann; W.E. Helm, *Simulation stochastischer Prozesse Eine anwendungsorientierte Einführung*, Springer-Verlag Berlin Heidelberg, 2016.
- [43] [http://www.langemathenacht.de/zufallskurven/Modul\\_zufallskurven\\_excel.pdf](http://www.langemathenacht.de/zufallskurven/Modul_zufallskurven_excel.pdf), Jänner 2020.
- [44] <https://homepage.univie.ac.at/franz.embacher/Lehre/aussermathAnw2012/RandomW>, Jänner 2020.
- [45] <https://de.wikipedia.org/wiki/Monte-Carlo-Simulation>, Jänner 2020.
- [46] [https://mars.wiwi.hu-berlin.de/mediawiki/mmstat3/index.php/Diskrete\\_Gleichverteilung](https://mars.wiwi.hu-berlin.de/mediawiki/mmstat3/index.php/Diskrete_Gleichverteilung), November 2019.
- [47] <https://www.eit.hs-karlsruhe.de/mesysto/teil-c-stochastische-signale/univariate-wahrscheinlichkeitstheorie/spezielle-stetige-verteilungen/normalverteilung.html>, November 2019.
- [48] <https://www.statistik-nachhilfe.de/ratgeber/statistik/wahrscheinlichkeitsrechnung-stochastik/wahrscheinlichkeitsverteilungen/stetige-verteilungen/stetige-gleichverteilung-rechteckverteilung>, November 2019.
- [49] <https://matheguru.com/stochastik/normalverteilung.html>, November 2019.
- [50] <https://www.spektrum.de/lexikon/mathematik/lognormalverteilung/6046>, November 2019.
- [51] <https://de.wikipedia.org/wiki/Inversionsmethode>, November 2019.



- [52] [https : //wirtschaftslexikon.gabler.de/definition/normalverteilung - 39769](https://wirtschaftslexikon.gabler.de/definition/normalverteilung-39769), Dezember 2019.
- [53] [http : //1.bp.blogspot.com/-73EcSyRdCcs/Uzqfkya1WII/AAAAAAAAAB8/87loD0JyIl8/s16](http://1.bp.blogspot.com/-73EcSyRdCcs/Uzqfkya1WII/AAAAAAAAAB8/87loD0JyIl8/s16) November 2019.
- [54] [https : //de.wikipedia.org/wiki/Inversionsmethode](https://de.wikipedia.org/wiki/Inversionsmethode), Dezember 2019.
- [55] [https : //de.wikipedia.org/wiki/Verwerfungsmethode](https://de.wikipedia.org/wiki/Verwerfungsmethode), Jänner 2020.
- [56] [https : //de.wikipedia.org/wiki/Schloss\\_Mannheim](https://de.wikipedia.org/wiki/Schloss_Mannheim), Jänner 2020.
- [57] [http : //www.langemathenacht.de/zufallskurven/Modulzufallskurven\\_excel.pdf](http://www.langemathenacht.de/zufallskurven/Modulzufallskurven_excel.pdf), Jänner 2020.

# Abbildungsverzeichnis

1	Wahrscheinlichkeitsfunktion einer diskreten Gleichverteilung mit sechs möglichen Ausgängen (z.B. Würfelwurf) . . . . .	19
2	Verteilungsfunktion einer diskreten Gleichverteilung am Beispiel des Würfelwurfs . . . . .	19
3	Dichtefunktion $f$ vs. Verteilungsfunktion $F$ einer standardisierten Normalverteilung $\mathcal{N}(0; 1)$ mit dem Mittelwert Null und der Standardabweichung Eins. . . . .	21
4	Dichte- und Verteilungsfunktion von einer stetigen Gleichverteilung im Intervall $[0; 30]$ . . . . .	22
5	Dichtefunktion einer Normalverteilung . . . . .	24
6	Dichtefunktion einer Normalverteilung $\mathcal{N}(\mu; \sigma)$ mit $\sigma = 1, \sigma = 2$ und $\sigma = 3$ . . . . .	25
7	Dichtefunktion einer Normalverteilung $\mathcal{N}(\mu; \sigma)$ mit konstanter Standardabweichung und drei Mittelwerten $\mu_1 < \mu_2 < \mu_3$ . . . . .	25
8	Dichteverteilung einer Normalverteilung $\mathcal{N}(\mu; \sigma)$ und dem relativen Anteil der Zufallsvariable innerhalb der drei Intervalle $(\mu \pm \sigma)$ , $(\mu \pm 2\sigma)$ und $(\mu \pm 3\sigma)$ . . . . .	27
9	Dichteverteilung einer logarithmischen Normalverteilung $\mathcal{L}(0; 1)$ . . . . .	28
10	Galtonbrett . . . . .	36
11	Entstehung einer logarithmischen Normalverteilung $\mathcal{L}(\mu; \sigma)$ aus den beiden Normalverteilungen der Blattlänge $L$ (links oben), der Blattbreite $B$ (links unten) und der Blattfläche $F$ (rechts unten) . . . . .	37
12	Inversionsmethode mit dem Inversen $F^{-1}$ der Verteilungsfunktion $F$ . . . . .	48
13	Einige Dichtefunktionen mit ihrer Verteilungsfunktion $F$ und ihrer Umkehrfunktion $F^{-1}$ . . . . .	49
14	Graphische Darstellung der Verwerfungsmethode mit dem Graphen $f(x)$ und $k \cdot g(x)$ . . . . .	50
15	Modellbildungsprozess . . . . .	53
16	Arbeitsmaterial: Tabelle mit Zufallszahlen im Intervall $[0; 9]$ für die Erarbeitung des Modells 3 . . . . .	59
17	1.Quadrant mit Quadrat mit Seitenlänge $l = 1$ und Viertelkreis mit Radius $r = 1$ . . . . .	62
18	"Berechnung" mit Zufallszahlen . . . . .	63
19	Stadtplan von Mannheim mit eingezeichneten Koordinatengitter sowie gelb markierten Startpunkt . . . . .	66
20	Graphische Darstellung eines Random Walks mit Hilfe eines Kalkulationsprogramms mit $n = 100$ . . . . .	67
21	Graphische Darstellung eines Random Walks mit Hilfe eines Kalkulationsprogramms mit $n = 1000$ . . . . .	68

22	Graphische Darstellung eines Random Walks mit Hilfe eines Kalkulationsprogramms mit $n = 10000$ . . . . .	68
23	Vergleich der empirischen Daten mit der Normalverteilung, mit dem Mittelwert $\bar{x} = 21$ und der Standardabweichung $\sigma = 4,4$ . . . . .	71
24	Beispiel für ein Experiment in der Klasse: Anzahl der verbleibenden Schülerinnen und Schüler $N$ abhängig vom Zeitschritt $t$ . . . . .	74
25	Darstellung eines Experiments (grüne Kreuze) mit Würfeln und dem radioaktiven Zerfallsgesetz (rote Linie) mit $N_0 = 24$ und $\lambda = \frac{1}{6}$ . . . . .	75