DER ZUFALL IN DER MATHEMATIK

Fachbereichsarbeit aus Mathematik

Vorgelegt von
Victoria **T**arisai **T**iki, Klasse VIIIB

Eingereicht bei
Mag. **C**hrista **D**ell'mour

Bundesgymnasium Horn, Puechhaimgasse 21, 3580 Horn
Februar 2014, Schuljahr 2013/14

1. VORWORT

Das Thema „Zufall“ war in meinen Augen nicht zuletzt hinsichtlich physikalischer Belange und deren Bedeutung für den Menschen stets interessant. Nach eingehenderer Beschäftigung mit dieser Materie, entschloss ich mich Ende der 7. Klasse dazu, diese als Gegenstand meiner Fachbereichsarbeit, die ich schon seit Beginn der Oberstufe zu verfassen gedachte, zu wählen.

Der ursprüngliche Fokus auf die Physik der Zufallszahlen musste bereits im Zuge der Vorbereitungen erweitert und abgeändert werden, um auch die von der Mathematik dominierten Bereiche der Zufallszahlengenerierung einzubeziehen. Dadurch wurde ein Wechsel vom anfänglich vorgesehenen Fachbereich Physik zur Mathematik notwendig. Meine ursprüngliche Idee war nicht wesentlich über eine umfassende Beleuchtung aller gängigen Methoden der Zufallszahlengenerierung hinauszugehen, doch schon nach kürzester Zeit erwies sich, dass in deren Kontext auch Anwendungen und Gütekriterien von Bedeutung sind. Besonders letztere brachte mich häufig an die Grenzen meiner Kenntnisse, weswegen die Auseinandersetzung mit der vorliegenden Fachbereichsarbeit auch vom eigenständigen Erlernen eigentlicher mathematischer und informationstechnologischer Grundkompetenzen geprägt war. Ich nehme an, dass dieses, wie auch das generelle Befassen mit der wissenschaftlichen Methode, für meine weitere Laufbahn als Naturwissenschaftlerin von großer Bedeutung sein wird. Daher halte ich rückblickend den Entschluss diese Arbeit geschrieben zu haben, aus einer Vielzahl an Gründen für einen guten.

Ich danke allen, die zur Entstehung dieser Arbeit beigetragen haben. Dazu zählen besonders Mag. Christa Dell’mour, Betreuerin dieser Arbeit, sowie meine Familie und Freunde.

2. INHALTSVERZEICHNIS

3. EINLEITUNG

Werden hundert Menschen mit der Aufgabe konfrontiert eine Zufallszahl zwischen 1 und 10 niederzuschreiben, so wählt circa ein Drittel die Zahl 7, während dagegen nur zwei Personen an die 10 denken.[1] Dieses leicht durchzuführende Experiment beweist, dass sich der menschliche Geist nur bedingt zum Erzeugen einer zufälligen Ziffer eignet. Die vorliegende Arbeit behandelt unter anderem dieses psychologische Phänomen, aber auch den Zufall als solches, seine Unbestimmtheit sowie seine Berechenbarkeit.

Dazu wird eine Einführung in den allgemeinen Zufallsbegriff aus der Sicht diverser wissenschaftlicher Disziplinen, wie etwa der Psychologie und der Physik, geboten, um letztendlich die mathematische Bedeutung entsprechend definieren zu können. Fragen um den freien Willen oder die Vorbestimmtheit des Universums sind in diesem Zusammenhang zwei der bedeutendsten behandelten Thematiken. Bevor auf das Hauptmotiv dieser Fachbereichsarbeit eingegangen werden kann, ist es essentiell die Gütekriterien der einzelnen Zufallszahlengeneratoren zu beschreiben, da diese gleichsam das Fundament der Zufallszahlenerzeugung bilden. Die darauf folgenden Abschnitte befassen sich im Grunde mit Zufallszahlen und insbesondere mit den wichtigsten Methoden ihrer Generierung. Dies beinhaltet etwa deren Geschichte und Güte respektive Dienlichkeit in Bezug auf verschiedene Anwendungsgebiete.

[1] Siehe Anhang

4. EINE EINFÜHRUNG IN DEN ZUFALL

Der zunächst wissenschaftliche Begriff „Zufall“ ist längst in den Grundwortschatz der Umgangssprache übergegangen. Ob es sich um das Glückspiel, alltägliche Ereignisse oder lebensverändernde Entscheidungen handelt - der Faktor der Ungewissheit, eine bestimmte Zufallsvariable wird von Umsichtigen selten außer Acht gelassen. Allerdings muss hierbei der Tatsache Aufmerksamkeit geschenkt werden, dass der mathematische Zugang zu Wahrscheinlichkeiten und Zufallszahlen und jener anderer Wissenschaften, oder gar die der Prognosen aus der Alltagswissenschaft, entscheidend divergieren. Dem Nachkommenden liegt der Versuch zugrunde andere Disziplinen, die sich mit diesem Begriff auseinandersetzen, zu beleuchten, um den Zufall einerseits besser erfassen zu können und um andererseits zum Kern der mathematischen Definition und ihrer Bedeutung zu kommen.

4.1. BEGRIFFSDEFINITION

Zufälle werden gemeinhin über ihr unsicheres Verhältnis zu einer Ursache charakterisiert. Während die alltäglich gebräuchliche Deutung in dem betreffenden Begriff *„etwas, was man nicht vorausgesehen hat, was nicht beabsichtigt war, was unerwartet geschah“* [2] sieht, fällt die wissenschaftlichere Auslegung wesentlich komplexer aus. So schlug der österreichische Wissenschaftler Philipp Frank (1884 – 1966) schon im Jahre 1932 folgende Erklärung vor:

> *„Ein Zufall schlechthin, also gewissermaßen ein absoluter Zufall wäre dann ein Ereignis, das in bezug [sic!] auf alle Kausalgesetze ein Zufall ist, das also nirgends als Glied einer Kette auftritt“* [3]

Die Idee, dass ein zufälliges Ereignis nicht in eine Kausalkette einzuordnen ist, wird üblicherweise von der Wissenschaft als Definition anerkannt. Allerdings erweist es sich häufig als nicht eindeutig festlegbar, ob oder wie ein Ereignis in solch eine Kette passt, eine Problematik, die auch noch bei der Generierung von Zufallszahlen ein Kernpunkt sein wird.

4.2. DIE PHYSIK UND DER ZUFALL

Eine gravierende Bedeutung und vielfältige Anwendungsmöglichkeiten findet der Zufall respektive Zufallszahlen in der Physik. Die zentrale Frage stellt hierbei jene um den Determinismus dar, die im diametralen Verhältnis zwischen der Chaostheorie beziehungsweise der klassischen Mechanik und der Quantenphysik kulminiert.

[2] (Zufall: Duden, 2013)
[3] (Frank, 1932), Seite 156, 157

4.2.1. DETERMINISMUS

Die klassische Mechanik nach Isaac Newton (1643–1727) und Marquis Laplace (1749 – 1827) geht davon aus, dass die Teilchen des Universums einfachen Bewegungsgesetzen folgen und sich von jedem Ereignis eine Kausalkette zurück zu einem Ursprung finden lässt. Ähnlich setzt auch die Chaostheorie voraus, dass es einen Anfangszustand unendlich genau zu kennen gilt, um künftige exakt vorherzusagen zu können. In diesem deterministischen Konstrukt gibt es keinen Raum für Zufälle.[4] Theoretisch könnten leistungsfähige Computer nach dieser Annahme die Zukunft berechnen, vorausgesetzt es bestünde die Möglichkeit sie mit allen Informationen über den Anfangszustand des Universums zu speisen.

4.2.2. UNSCHÄRFE

Das deterministische Weltbild wurde im 20. Jahrhunderts maßgeblich durch die Quantenphysik erschüttert. Die Theorie hinter dem berühmten Ausspruch „*Gott würfelt nicht*“[5], der dem Physiker Albert Einstein (1879–1955) zugeschrieben wird, konnte 1927 von Werner Heisenberg (1901–1976) durch seine Formulierung der Unschärferelation (siehe *Formel 1*) falsifiziert werden. Der Determinismus basiert auf der Idee, dass die Bedingungen eines Ausgangszustands zu kennen sind, um künftige Zustände vorhersagen und somit die Variable des Zufalls aus der Gleichung streichen zu können. Durch die Unschärferelation wird eine solche präzise Aufnahme des Ausgangszustandes zu einer quantenphysikalischen Unmöglichkeit. Die berühmte Formel impliziert, dass es für einen objektiven Beobachter unmöglich ist, zur gleichen Zeit sowohl den Ort als auch die Geschwindigkeit

$$\Delta x \Delta p \sim h$$

Formel 1: Unschärferelation nach Werner Heisenberg. das Produkt aus der Unbestimmtheit des Ortes (Δx) und jener des Impulses (Δp) ist größer als das Planck'sche Wirkungsquantum (h)

eines Teilchens zu kennen.[6] Oder, in anderen Worten: „*Je genauer man die Position eines Teilchens misst, desto weniger genau kann man seine Geschwindigkeit bestimmen und umgekehrt.*“[7]

Durch diese Unsicherheit wird das wichtige Kriterium der deterministischen Physik, dass ein Anfangszustand in seiner Gänze zu verstehen ist, um eine bestimmte Prognose darauf basieren zu können, nicht erfüllt.

In makroskopischen Systemen kann die Unschärfe vernachlässigt werden, da es durch die Wechselwirkung der Teilchen mit anderen Systemen zu einer Dekohärenz, einem Verlust der Überlagerung mehrerer Zustände, kommt. In diesem Zusammenhang wird auch vom Kollaps der

[4] Vgl. (Al-Khalili, 1962)Seite 54
[5] (Einstein, 1926)
[6] Vgl. (Fischer, 2010) 171 f
[7] (Hawking) Seite 113

Wellenfunktion auf eine Wirklichkeit, die sich wiederum wie in der klassischen, deterministischen Physik verhält, gesprochen.[8]

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = i\hbar\frac{\partial\Psi}{\partial t}$$

Formel 2: Die Schrödinger-Gleichung nach Erwin Schrödinger (1887-1961) bestimmt die Änderungsrate der Wellenfunktion.
$\hbar$ bezeichnet die Planck'sche Konstante, m die Masse des beschriebenen Teilchens, ∇^2 den sogenannten Napla-Quadrat-Operator (er beschreibt, wie sich die Wellenfunktion ψ von einem Ort zum anderen ändert), V die auf das Teilchen einwirkenden Kräfte, i die imaginäre Zahl ($\sqrt{-1}$) und $\frac{\partial\psi}{\partial t}$ die Änderung von ψ in Abhängigkeit von der Zeit.[9]

Die Wellenfunktion ist die Lösung der Schrödinger-Gleichung und enthält *„alle über ein Quantensystem erreichbaren Informationen"*[10], indem sie die Wahrscheinlichkeiten der einzelnen Eigenschaften eines Systems angibt. Das bedeutet beispielsweise für ein Elektron, dass die Wellenfunktion den probabilistischen Raum seiner Aufenthaltsorte angibt, nicht aber eine eindeutige Position. Ähnlich verhält es sich auch mit der Geschwindigkeit des Teilchens.[11]

Die einzelnen Eigenschaften des Systems bleiben daher unsicher, hingegen ist die Wellenfunktion wohldefiniert. Wenn diese Wellenfunktion wie in der klassischen Physik als eindeutige Information behandelt wird, kann sie mithilfe der Schrödinger-Gleichung für einen beliebigen Zeitpunkt in der Zukunft berechnet werden. Es kann daher auch in der von Unschärfe gezeichneten Quantenphysik von einer Art Determinismus gesprochen werden.[12]

Ob das Universum vom Determinismus oder vielmehr von quantenphysikalischen Zufällen bestimmt wird, bleibt demnach selbst in der Physik eine Streitfrage.

4.2.3. CHAOS

Es gibt neben Determinismus und Unschärfe einen dritten in diesem Kontext stehenden Begriff, der wie ein Bindeglied zwischen den vorhergehenden fungiert.

Die Chaostheorie postuliert eine Art unscharfen Determinismus, indem sie zwar davon ausgeht, dass ein bestimmtes System klaren Regeln, die durchaus völlig im Bereich unseres physikalischen Verständnisses liegen, folgt, aber dennoch zu chaotisch ist, als dass wir seine künftigen Zustände berechnen könnten. Kleine Messungenauigkeiten und unbekannte Größen führen schlussendlich dazu, dass sich das System schon binnen kurzer Zeit unserem Vorhersagevermögen entzieht.[13] Klassisches Beispiel hierfür ist das Wetter: Da wir die komplexen Anfangszustände der gesamten Erdatmosphäre nicht exakt kennen, lässt sich immer nur eine Approximation der derzeitigen

[8] Vgl. (Müller, 2006) Seite 3f
[9] Vgl. (Al-Khalili, 1962) Seite 63
[10] (Al-Khalili, 1962)
[11] Vgl. (Al-Khalili, 1962) Seite 64f
[12] Vgl. (Hawking) Seite 116
[13] Vgl. (Letellier, 2010) Seite 24

Wetterlage ermitteln. Basierend auf dieser Approximation kann eine Wetterprognose für die Zukunft errechnet werden. Diese Vorhersage wiederum mag für kurze Zeit, etwa zwei Tage oder eine Woche, ausreichend realitätsnah sein, doch durch seine chaotische Natur sind Wetterprognosen für ein gesamtes System nahezu unmöglich. Durch kleine Ungenauigkeiten divergieren die tatsächlichen und die errechneten Werte erheblich voneinander.

Für den Menschen erscheinen solche Ereignisse dann so zufällig wie die Position eines Elektrons oder der Zerfall eines instabilen Isotops, tatsächlich folgt in der Chaostheorie dennoch Wirkung auf Ursache.

4.3. DER MENSCH UND DER ZUFALL

Von dieser eben erwähnten Streitfrage ist auch die Diskussion um die Existenz des freien Willens abhängig. Da menschliche Entscheidungen im Grunde auf in erster Instanz neurobiologischen und in weiterer auf physikalischen Prozessen beruhen, liegt der Schluss nahe, dass der menschliche Charakter letztlich ebenfalls den Gesetzen von Determinismus und Indeterminismus, von Vorherbestimmtheit und Zufall, zugrunde liegt. In diesem Fall scheint der Begriff „freier Wille“ lediglich als ein fiktives Konstrukt.

Neben der Philosophie befasst sich auch die Kognitionspsychologie mit dem Zufall und hier insbesondere mit dem subjektiven Bewerten zufälliger Ereignisse durch den Menschen, sowie Erwartungen an ebendiese. [14] Dass Menschen Wahrscheinlichkeiten und Zufälle nur bedingt bewerten oder gar vorhersagen können, beweist eine Beobachtung des deutschen Physikers Hans Reichenbach[15] (1891–1953), der erkannte, dass die menschliche Intuition besonders dann an ihre Grenzen stößt, wenn es darum geht die Frequenz aufeinanderfolgender identischer Charaktere zu bewerten[16] (siehe *Tabelle 1*).

Durch Münzwurf generierte Zufallszahlen	Von einem Menschen geschätzte Zufallszahlen
ZKZZKKKZZKKKKKKZKKKK	KZKKZZKZKZKKZZKKZKZK
KKZZKKKKKKZKKZKZKZKZ	ZKZKZKZKZKKKKZZKZKZK
KZZKKKKZZZZKKKKZZKKZ	ZZKKZZKKZKKKZKZZKZZK
ZKKKKKKKZKKKKKZZKKZZ	KZKKZKZKKZZKZKKKZZKZ
ZKZZKZZZZKKZKKKKKKKZ	ZZKKKZZKKKZZZKZKZZKK

Tabelle 1: Experiment zur Veranschaulichung von Reichenbachs Beobachtungen. Neben einer durch Münzwurf generierten Sequenz wurde ein Mensch gebeten das Ergebnis davon im Vorhinein zu schätzen (Die längste Folge aufeinanderfolgender, identer Ergebnisse ist jeweils hervorgehoben). Die Darstellung illustriert die Qualitätsunterschiede zwischen einem menschlichen und einem tatsächlichen Zufallszahlengenerator: Ein Mensch unterschätzt üblicherweise wie häufig idente Werte aufeinanderfolgen können.

[14] Vgl. (Wahrnehmung des Zufalls: wikipedia.org, 2013)
[15] Vgl. (Reichenbach, 1935), zitiert nach (Wahrnehmung des Zufalls: wikipedia.org, 2013)
[16] Vgl. (Wahrnehmung des Zufalls: wikipedia.org, 2013)

Der Zufall ist nicht nur in mathematischer Hinsicht interessant, wie die folgenden Kapitel suggerieren könnten, sondern auch ein bedeutender Aspekt in der Philosophie und der Geschichte. Der griechische Philosoph Demokrit (Ende 5. Jhdt. v. Chr.), berühmt für seine These des unteilbaren Teilchens (*atomos*), glaubte etwa nicht an den Zufall („*Die Menschen haben sich im Zufall ein Trugbild geschaffen, eine Ausrede für ihre eigene Torheit*“) [17] und lieferte dafür folgendes Gedankenexperiment:

Zwei Sklaven, die von ihren beiden Herren absichtlich zur selben Zeit an denselben Ort geschickt werden, mögen ihr Treffen für einen Zufall halten, unwissend um die Tatsache, dass ihr Zusammenstoß genau geplant war.[18] Analog dazu erkläre sich auch die Menschheit in Wirklichkeit genau berechenbare Ereignisse nur aus Unwissenheit um die wahren Umstände als durch den Zufall bestimmt. Dementsprechend wäre Demokrit aus heutiger Sicht ein Vertreter des Determinismus.

Demokrit war der Ansicht, dass ein objektiver Zufall nicht existiere[19], Ereignisse wie ein ungeplantes Treffen mögen nur dann zufällig erscheinen, wenn ein Beobachter oder eine Beobachterin nicht über die vollständige Bandbreite der Informationen, die zu diesem Ereignis geführt haben, Bescheid weiß. Der Philosoph Epikur (4. bis 3. Jhdt. v. Chr.) widersprach Demokrit, indem er zum Ausdruck brachte, dass „*der Zufall […] objektiv [ist], es ist die eigentliche Natur der Erscheinungen*“[20]. Folglich wäre Epikur mit seiner Einstellung, dass es sehr wohl einen objektiven Zufall gäbe, ein Gegner des Determinismus.

4.4. Die Mathematik und der Zufall

Die moderne Mathematik scheint durch das Bestimmte bedingt zu sein: Die Lösung einer Gleichung hängt nicht vom Zufall ab, Konstanten haben immer einen klar definierten Wert. Selbst π, die wohl berühmteste transzendente Zahl, ändert nicht zufällig ihre unendlichen Nachkommastellen.

All dies wirft die Frage auf, inwiefern diese wissenschaftliche Disziplin und der Begriff Zufall in Beziehung zueinander stehen. Der zentrale Begriff in diesem Zusammenhang ist der der Zufallszahl. Dieser ist jedoch irreführend, da er suggeriert, dass eine einzelne Zahl etwas Zufälliges haben könnte. Die Zahl 5 ist nicht „zufälliger“ als die Zahl 3456 oder vice versa. Eine Zufallszahl oder auch Zufallsvariable hat nicht per se diese Eigenschaft, sie ergibt sich ausschließlich daraus, dass sie Element einer Zufallssequenz, einer Folge von Zahlen in einem bestimmten Bereich, ist, wobei Position und Verteilung aller möglichen Zahlen vom Zufall bestimmt sind. Eine Zufallszahl ist daher nichts ohne die Zufallssequenz, der sie angehört.

[17] Zitiert nach (Demokrit: zitate.eu, 2013)
[18] Vgl. (Roney-Dougal, du Sautoy, & Gowers, 2011) ab Minute 5:20
[19] Vgl. (Hromkovic, 2008) Seite 188
[20] Zitiert nach (Hromkovic, 2008) Seite 188

Die Aufgabe der Mathematik liegt nun darin, derartige Sequenzen von Zufallszahlen, respektive Pseudozufallszahlen, für verschiedene Anwendungsbereiche (siehe *7 Anwendungen der* Zufallszahlengeneratoren) zu erzeugen. Die Formalwissenschaften, in deren Zusammenhang Wahrscheinlichkeiten und Zufälle der wesentliche Teil der Stochastik sind, stehen hier aber vor einem Problem, denn wie schon zu Beginn dieses Abschnittes geäußert, ist die Mathematik keine Wissenschaft des Unbestimmten.

Die durch Algorithmen erzeugten Sequenzen von Zufallszahlen sind daher nicht wirklich zufällig, sie erscheinen so lediglich Beobachtern oder Beobachterinnen, die nicht über die Mittel verfügen, die Sequenz nach allen Gütekriterien zu überprüfen. [21] Anhand der Veröffentlichung der Rand Corporation eines Buches, das eine Million Zufallszahlen enthält,[22] lässt sich nachweisen, dass Zufallszahlen nicht bloß eine mathematische Spielerei darstellen, sondern tatsächlich genügend

Zeile Nr.	Spalte Nr. 1–5	6–10	11–15	16–20	21–25	26–30	31–35	36–40	41–45	46–50
0	10097	32533	76520	13586	34673	54876	80959	09117	39292	74945
1	37542	04805	64894	74296	24805	24037	20636	10402	00822	91665
2	08422	68953	19645	09303	23209	02560	15953	34764	35080	33606
3	99019	02529	09376	70615	38311	31165	88676	74397	04436	27659
4	12807	99970	80157	36147	64032	36653	98951	16877	12171	76833
5	66065	74717	34072	76850	36697	36170	65813	39885	11199	29170
6	31060	10805	45571	82406	35303	42614	86799	07439	23403	09732
7	85269	77602	02051	65692	68665	74818	73053	85247	18623	88579
8	63573	32135	05325	47048	90553	57548	28468	28709	83491	25624
9	73796	45753	03529	64778	35808	34282	60935	20344	35273	88435
10	98520	17767	14905	68607	22109	40558	60970	93433	50500	73998
11	11805	05431	39808	27732	50725	68248	29405	24201	52775	67851
12	83452	99634	06288	98083	13746	70078	18475	40610	68711	77817
13	88685	40200	86507	58401	36766	67951	90364	76493	29609	11062
14	99594	67348	87517	64969	91826	08928	93785	61368	23478	34113
15	65481	17674	17468	50950	58047	76974	73039	57186	40218	16544
16	80124	35635	17727	08015	45318	22374	21115	78253	14385	53763
17	74350	99817	77402	77214	43236	00210	45521	64237	96286	02655
18	69916	26803	66252	29148	36936	87203	76621	13990	94400	56418
19	09893	20505	14225	68514	46427	56788	96297	78822	54382	14598
20	91499	14523	68479	27686	46162	83554	94750	89923	37089	20048
21	80336	94598	26940	36858	70297	34135	53140	33340	42050	82341
22	44104	81949	85157	47954	32979	26575	57600	40881	22222	06413
23	12550	73742	11100	02040	12860	74697	96644	89439	28707	25815
24	63606	49329	16505	34484	40219	52563	43651	77082	07207	31790

Abbildung 1: Ausschnitt aus dem Buch „A Million Random Digits with 100 000 Normal Deviates“, eine Tabelle mit Zufallsziffern zwischen 0 und 9

Anwendungsgebiete, die den Nutzen einer solchen Publikation rechtfertigen, finden. Jede Seite des Buches besteht im Wesentlichen aus Zahlen zwischen 0 und 9, die in einer riesigen Tabelle angeordnet sind[23] (siehe *Abbildung 1*). Die Zufallszahlengenerierung hat eine lange Geschichte, wobei sich etliche frühe Methoden nicht in die beiden heutigen bekannten, im Folgekapitel beschriebenen, Kategorien einteilen lassen.

So erstellte der englische Statistiker L.H.C. Tippet (1902 - 1985) im Jahre 1927 eine Liste von 41.600 Zufallszahlen, indem er die mittleren Stellen der gemessenen Gemeindeflächen Englands zusammenstellte und veröffentlichte.[24]

[21] Vgl. (Kütting, 1999) Seite 145
[22] Vgl. (TheRandCorporation, 1955)
[23] Vgl. (Kütting, 1999) Seite 143
[24] Vgl. (Peterson, 1998) Seite 170

4.4.1. Deterministische und physikalische Zufallszahlen

Numerische, respektive Pseudozufallszahlen sind deterministisch, das heißt sie werden von mathematischen Algorithmen hervorgebracht[25] und können üblicherweise jederzeit von einem Beobachter unter Anwendung derselben Formeln vorhergesagt oder reproduziert werden.[26] Da durch mathematische Formeln erzeugte Zufallszahlen nicht als „echt" gelten, laufen die Sequenzen gewöhnlich zunächst statistische Tests durch, um sie auf die Anwendbarkeit in speziellen Situationen testen zu können (siehe *5 Gütekriterien und -tests für Zufallszahlengeneratoren*). Die Qualität von Pseudozufallszahlen variiert je nach Methodik sehr stark, daher sind für bestimmte Anwendungsgebiete nur wenige Zufallszahlengeneratoren zweckdienlich. Ein Vorteil dieser Methode ist ihr simpler und schneller Gebrauch.[27]

Tatsächliche, „echte" Zufallszahlen entstehen durch physikalische Prozesse, wie etwa durch das Werfen eines Würfels oder das Rauschen der Atmosphäre im Radiofrequenzbereich.[28] Einige physikalische Methoden der Zufallszahlengewinnung sind allerdings in ihrer tatsächlichen Zufälligkeit umstritten. So wäre es beispielsweise möglich, das Ergebnis eines Münzwurfes mittels außerordentlich leistungsfähiger Rechengeräte und unter Kenntnis aller relevanten Anfangszustände des Würfels inklusive seiner Umgebung, zu berechnen.[29] Wie im Kapitel *Die Physik und der Zufall* dargestellt sind einige Vorgänge, die einem Beobachter zunächst zufällig erscheinen, dennoch den Regeln des Determinismus unterworfen. Davon sind auch physikalische „Zufallszahlengeneratoren" wie das Roulettespiel oder die Lottozahlenziehung betroffen.[30]

Physikalische Zufallszahlen erfüllen so gut wie alle statistischen Qualitätskriterien und haben daher theoretisch ein breites Anwendungsfeld, ihre langsame und komplexe Generierung schmälert allerdings die Häufigkeit ihrer Anwendung. Abgesehen davon sind die physikalischen Zufallsprozesse nicht auf exakt derselben Weise wiederholbar, was die dadurch entstandenen „echten" Zufallszahlensequenzen für etliche Disziplinen, wie etwa der Kryptographie, unbrauchbar macht.

4.4.2. Gleichverteilte und andersartig verteilte Zufallszahlen

Innerhalb der numerischen RNG (=random number generators), auch PRNG (=pseudorandom number generator), erfolgt eine weitere Differenzierung zwischen gleichverteilten Zufallszahlen und solchen mit andersartiger Distribution, wie etwa die Exponentialverteilung oder die Gauß-Verteilung (auch Normalverteilung). In den folgenden Kapiteln, die sich mit den numerischen Zufallszahlengeneratoren befassen, wird besonders auf gleichverteilte Zufallszahlen eingegangen, da

[25] Vgl. (Kütting, 1999) Seite 145
[26] Vgl. (Peterson, 1998) Seite 169
[27] Vgl. (Mordasini & Klahr, 2013)
[28] Vgl. (Mordasini & Klahr, 2013)
[29] Vgl. (Pseudozufall: wikipedia.org, 2013)
[30] Vgl. (Hromkovic, 2008) Seite 188f

daraus beliebige andere Verteilungen erzeugbar sind.[31] Allerdings offenbart besonders die Natur zahlreiche andere Verteilungen, wie etwa die Normalverteilung oder Gauß-Verteilung, die besonders treffend die Zerfallsrate eines radioaktiven Elements innerhalb eines kurzen Intervalls bezeichnet. [32] An dieser Stelle soll angemerkt werden, dass, obwohl man annehmen könnte, nur ein Zufallsexperiment in dem jeder Wert mit derselben Wahrscheinlichkeit auftritt (Wie das Werfen eines einzelnen Würfels), wäre tatsächlich zufällig, auch Variablen einer Normalverteilung zufällig sein können.

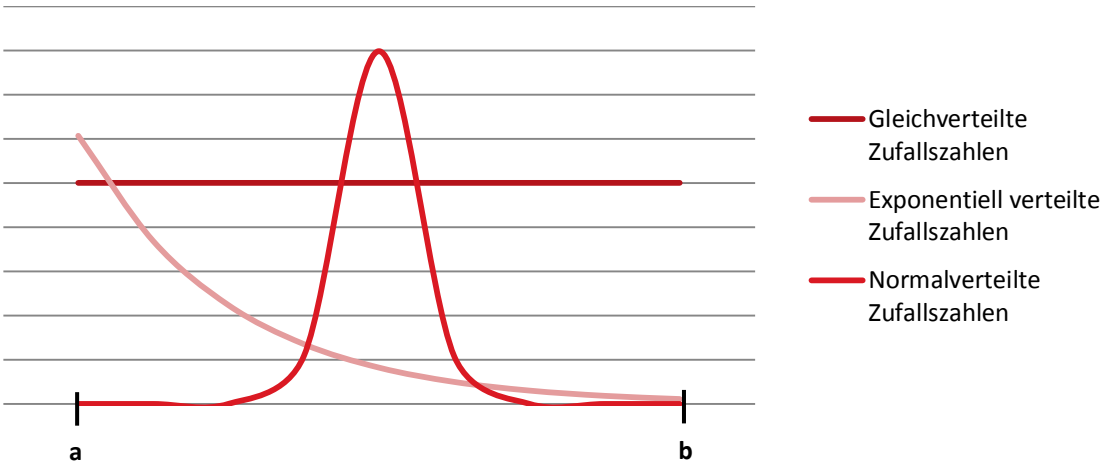

Abbildung 2: unterschiedliche Verteilungen von Zufallszahlen im Intervall (a,b)

Werden beispielsweise zwei Würfeln geworfen, so muss festgestellt werden, dass die Wahrscheinlichkeitsfunktion der dabei entstehenden Würfelpaare wie in Tabelle 2 einer Normalverteilung folgt (siehe *5.4.1.1 Der Chi-Quadrat-Test*). Diese Eigenschaft bedeutet dabei nicht, dass das Werfen mit zwei Würfeln nicht zufällig wäre, denn auch wenn jedes Augenpaar eine unterschiedliche Wahrscheinlichkeit hat, so lässt sich im Vorhinein nicht determinieren, welches schließlich obenauf landet. [33]

[31] Vgl. (Mordasini & Klahr, 2013)
[32] Vgl. (Clewett & Numberphile, 2013)
[33] Vgl. (Clewett & Numberphile, 2013)

5. GÜTEKRITERIEN UND -TESTS FÜR ZUFALLSZAHLENGENERATOREN

Unterschiedliche Anwendungsgebiete der Zufallszahlen stellen differenzierte Anforderungen an die Güte der jeweiligen Generatoren. Durch einen Algorithmus entstandene Zahlen sind nie wirklich zufällig, können aber ähnliche Eigenschaften wie tatsächliche Zufallszahlen haben. Die im Folgenden dargestellten Kriterien sind besonders in Bezug auf deterministische RNG wichtig, da diese im Gegensatz zu physikalischen Zufallszahlengeneratoren keine „echten“ und daher potentiell fehlerhafte Zufallssequenzen liefern. Beim Bewerten der Ergebnisse dieser Tests ist allerdings Vorsicht geboten, nicht immer, wenn eine Sequenz einen bestimmen Test nicht besteht, ist der Rückschluss, dass auch der dazugehörige RNG unbrauchbar ist, gerechtfertigt. [34] Andererseits wäre das ständige Bestehen eines (vor allem empirischen) Tests, Grund zum Verdacht der RNG könnte fehlerhaft sein.[35]

Die in *6 Zufallszahlengeneratoren* beschriebenen RNGs und PRNGs werden anhand der hier beschriebenen Gütekriterien bewertet. An dieser Stelle soll angemerkt werden, dass die beschriebenen Tests die tatsächliche Zufälligkeit einer Zahlensequenz nicht verifizieren oder falsifizieren, sondern nur Wahrscheinlichkeiten und grobe Anhaltspunkte angeben können.

5.1. PRAKTISCHE ANFORDERUNGEN

Drei praktische Anforderungen an RNG, ob deterministisch oder physikalisch, sind von Bedeutung:

Effizienz: RNGs deren Algorithmus sehr komplex und daher intensiv zu berechnen ist, werden wegen ihrer mangelnden Effizient selten, beziehungsweise nur eingeschränkt genutzt. [36]

Wiederholbarkeit: Besonders für Simulationen (siehe *7 Anwendungen der* Zufallszahlengeneratoren) müssen die Zufallszahlensequenzen leicht reproduzierbar sein; derselbe „seed“, also Anfangswert sollte immer dieselbe Sequenz erzeugen. [37]

Tragbarkeit: Tragbarkeit meint in diesem Zusammenhang, dass der entsprechende RNG relativ einfach auf diverser Hardware und Betriebssystemen zu implementieren ist. [38]

5.2. UNVORHERSEHBARKEIT

Dieses Kriterium wird tatsächlich nur von physikalischen Zufallszahlengeneratoren erfüllt, da nur diese effektiv unvorhersehbare Sequenzen erzeugen können. Deterministische RNGs werden über bestimmte Algorithmen, deren Wesen eine Unvorhersehbarkeit per Begriffsbestimmung unmöglich

[34] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[35] Vgl. (Peterson, 1998) Seite 180
[36] Vgl. (Random Number Generators: PJM's Sparse Home Page, 2013)
[37] Vgl. (Random Number Generators: PJM's Sparse Home Page, 2013)
[38] Vgl. (Random Number Generators: PJM's Sparse Home Page, 2013)

macht, definiert. Jede Person könnte, unter Anwendung desselben Algorithmus und „seeds“, ohne Probleme dieselbe Sequenz erzeugen und den Algorithmus gegebenenfalls nach Analyse der durch ihn entstehenden Sequenz auch ohne Hintergrundinformationen über die Parameter nachvollziehen. [39] Besonders im Bereich des Glücksspiels oder der Kryptographie ist die Ungelegenheit dieser Eigenschaft verständlich. In der Praxis ist es allerdings unwahrscheinlich, dass ein/e unwissender/unwissende Beobachter/in derartige Sequenzen analysieren kann, deshalb sollte darauf geachtet werden, Parameter zu wählen, die die Sequenz besonders komplex und den Algorithmus daher möglichst nicht nachvollziehbar erscheinen lassen.

a)	1, 4, 10, 22, 46	b)	3, 7, 29, 4, 68

Sequenz 1: Der Algorithmus hinter Sequenz 1a) ist sehr leicht nachzuvollziehen, die Ergebnisse sind vorhersagbar, hingegen ist der Algorithmus hinter Sequenz 1b komplexer

5.3. Periodenlänge

Ein PRNG wird von einem beliebigen „seed“ (zu Deutsch: Samen) gestartet und produziert unter der Voraussetzung der Verwendung desselben „seeds“ immer dieselbe Sequenz. Die Periodenlänge eines PRNG wird über das Maximum aller dieser Anfangszustände innerhalb der Länge einer wiederholungsfreien Sequenz von Zufallszahlen definiert[40] (siehe *Sequenz 2*).

(39, 84, 13, 4, 99, 29, 36, 54, 39, 84, 13, 4, 99, 29, 36, 54,

39, 84, 13, 4, 99, 29, 36, 54, 39, 84, 13, 4, 99, 29, 36, 54)

Sequenz 2: Die Periodenlänge dieser (außerordentlich kurzen) Sequenz ist 8, da 8 Elemente vorkommen, bevor sich die Abfolge wiederholt.

Die Periode sollte besonders im Vergleich zum Wertebereich, den sie durchläuft, lang sein (Der Mersenne-Twister, auf den an späterer Stelle noch genauer eingegangen wird, liefert hier eine besonders lange Sequenzlänge) [41]. In ihrer einfachsten Implementation durchlaufen lineare Kongruenzgeneratoren im besten Fall alle möglichen Zahlen in einem bestimmten Wertebereich genau einmal. Bei „schlechten“ Parametern wiederholt sich die Sequenz schon viel früher.[42]

Dass eine lange Periode zu bewirken nicht einfach ist, zeigt das Beispiel des renommierten Informatikers der Stanford University Donald Knuth (*1938), der eine komplizierte Methode in dreizehn Schritten entwickelte, um *„ein beinahe unendliches Angebot von unglaublich zufälligen Zahlen“* [43] zu erhalten. Entgegen der Erwartungen des Wissenschaftlers, erwies sich dieser

[39] Vgl. (Peterson, 1998) Seite 169
[40] Vgl. (Pseudorandom number generator: wikipedia.org, 2013)
[41] Vgl. (Güte eines Pseudozufallszahlengenerators: wikipedia.org, 2013)
[42] Vgl. (Mordasini & Klahr, 2013) Folie 6f
[43] (Knuth, 1969) Seite 4f, Zitat im Original: „almost an infinite supply of unbelievably random numbers“

Algorithmus als unbrauchbar: Die Sequenz blieb bereits beim ersten Durchlauf bei der Zahl 6065038420 hängen und wiederholte sich bei einem anderen schon nach 7401 Werten.[44]

5.3.1. SATZ VON KNUTH

Die von dem Informatiker Donald E. Knuth formulierten Anforderungen an die Parameter Modulus (m) und Multiplier (a) ermöglichen eine maximale Länge der Zufallszahlensequenz. Sie gelten insbesondere für lineare Kongruenzgeneratoren, die in *6.1 Deterministische Zufallszahlen* näher beschrieben werden. An dieser Stelle soll lediglich gesagt werden, dass lineare Kongruenzgeneratoren nach dem Algorithmus

$$I_{j+1} = aI_j + c \ (mod\ m)$$

Formel 3: Formel für den linearen Kongruenzgenerator

gebildet werden. Für Knuth ist hier die Wahl des Modulus besonders bedeutend, da dieser den Wertebereich, den die Elemente der Sequenz durchlaufen können, darstellt. Es sollte deshalb eine große Zahl für diese Variable gewählt werden[45], aber auch folgende Punkte, die sich auf die Parameter m, a und c beziehen, dürfen, um das Ziel einer maximalen Periodenlänge zu erreichen, nicht außer Acht gelassen werden:

1. Das Inkrement c soll zum Modulus m teilerfremd sein
2. a-1 ist ein Vielfaches von jedem Primfaktor, der m teilt
3. a-1 ist ein Vielfaches von 4, wenn m ein Vielfaches von 4 ist[46]

5.4. VERTEILUNG DER ZUFALLSZAHLEN

Ein Anspruch, der an die Eigenschaften eines PRNGs gestellt werden darf, ist jener, dass die Verteilung der von ihm erzeugten Zufallszahlen der erwarteten Verteilung folgt. Für eine Gleichverteilung bedeutet dies, dass alle Werte mit derselben Häufigkeit in der Sequenz auftreten sollten, wobei kleine Abweichungen immer im Bereich des Möglichen und Wahrscheinlichen liegen.

5.4.1. GRUNDLEGENDE VERTEILUNGSTESTS

Die nachkommenden beiden Verteilungstests sind dank ihrer Universalität und ihrer Einfachheit integraler Bestandteil zahlreicher anderer Gütetests für Zufallszahlengeneratoren, aber auch per se wichtige Kriterien für die Güte eines PRNG.

5.4.1.1. DER CHI-QUADRAT-TEST

Der χ^2-Test ist die Erfindung des englischen Mathematikers Karl Pearson (1857 – 1936) aus dem Jahre 1900.[47]

[44] Vgl. (Peterson, 1998)Seite 173
[45] Vgl. (Knuth, 1969) Seite 11
[46] Vgl. (Knuth, 1969) Seite 16

Die Idee hinter dieser Methode besteht im Wesentlichen darin, die Verteilung der durch einen PRNG entstandenen Zufallszahlen in Bezug auf eine Idealverteilung [48], beziehungsweise die Alternativhypothese $F(x) \neq F_0(x)$ gegen die Nullhypothese $H_0 : F(x) = F_0(x)$ zu prüfen[49]. Dazu wird der Bereich m, in dem die Zufallszahlen liegen, in k gleich große (daher m/k große) Kategorien aufgeteilt und die Wahrscheinlichkeit, dass die Zufallszahlen in den jeweiligen Kategorien liegen, ermittelt. Anschließend werden jene Wahrscheinlichkeiten verglichen mit den tatsächlichen Verteilungen. [50] Bei einem guten Zufallszahlengenerator treten bestimmte Verteilungen beziehungsweise Häufigkeiten in den k Kategorien auf. Dieses Verhalten lässt sich gut angesichts eines Beispiels [51] explizieren: Werden zwei Würfel n Mal geworfen, so sehen die jeweiligen Wahrscheinlichkeiten für die Augensummen (s) wie folgt aus:

Augensumme s	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Wahrscheinlichkeit p_s	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Tabelle 2

Wird dieses Experiment dann etwa 180 Mal durchgeführt, so mag das Ergebnis folgendermaßen aussehen:

Augensumme s	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Wahrscheinlichkeit np_s	5	10	15	20	25	30	25	20	15	10	5
Tatsächliche Verteilung Y_s	4	5	19	25	25	35	24	22	11	7	3

Tabelle 3

Es zeigt sich, dass die erwartete Verteilung stark von der tatsächlichen differiert. Diese Tatsache allein ist dennoch noch nicht Grund zur Annahme, dass der Zufallszahlengenerator (hier das Würfelspiel) fehlerhaft ist. Gewissheit darüber liefert hingegen der χ^2-Test in Kombination mit mehrfachen Durchläufen des Experiments. Berechnet wird in diesem Zusammenhang die Varianz V, die die Streuung der Werte vom erwarteten Ergebnis angibt. Für die erwarteten und tatsächlichen Augensummen lautet sie:

$$V = (Y_2 - np_2)^2 + (Y_3 - np_3)^2 + \cdots + (Y_{11} - np_{11})^2 + (Y_{12} - np_{12})^2$$

[47] Vgl. (informatik: uni-hamburg, 2013) Seite 30
[48] Vgl. (Sedgewick, 1992) Seite 585
[49] Vgl. (Sachs, 1984) Seite 251
[50] Vgl. (Knuth, 1969) Seite 39
[51] Parallel zum Beispiel aus (Knuth, 1969) Seite 39f

beziehungsweise

$$V = \sum_{s=2}^{12} (Y_s - np_s)^2$$

V sollte also möglichst gering sein, denn das bedeutet, dass die tatsächlichen Werte nicht stark von den erwarteten abweichen. Tritt hingegen der Fall ein, dass eine Sequenz etwa nur 3er oder nur 10er aufweist, so bedeutet auch das nicht, dass der entsprechende Zufallszahlengenerator unbrauchbar ist. Theoretisch sind alle diese „falsch“ erscheinenden Sequenzen möglich, es ist dabei aber unwahrscheinlich, dass sie mit einer Häufigkeit auftreten, die ein perfekter Zufallszahlengenerator so gut wie niemals zulässt.[52]

Die obenan stehende Formel ist unvollständig, denn sie geht davon aus, dass alle Abweichungen vom Standardwert $(Y_s - np_s)^2$ mit dergleichen Gewichtung ins Ergebnis einfließen. Tatsächlich ist die Streuung um die Augensumme 7 $(Y_7 - np_7)^2$ wesentlich wichtiger als beispielweise die um die Augensumme 2 $(Y_2 - np_2)^2$, da die Wahrscheinlichkeit eine 7 zu würfeln höher ist (siehe *Tabelle 2*). Die Formel muss demnach angepasst werden[53]:

$$\chi^2 = \frac{(Y_2 - np_2)^2}{np_2} + \frac{(Y_3 - np_3)^2}{np_3} + \cdots + \frac{(Y_{11} - np_{11})^2}{np_{11}} + \frac{(Y_{12} - np_{12})^2}{np_{12}}$$

Oder, allgemein formuliert:

$$\chi^2 = \sum_{1=s}^{k} \frac{(Y_s - np_s)^2}{np_s}$$

Formel 4: Formel des Chi-Quadrat-Tests

Setzt man die Werte des Zufallsexperimentes aus Tabelle 3 ein,

$$\chi^2 = \frac{(4 - 180 \times \frac{1}{36})^2}{180 \times \frac{1}{36}} + \frac{(5 - 180 \times \frac{1}{18})^2}{180 \times \frac{1}{18}} + \cdots + \frac{\left(7 - 180 \times \frac{1}{18}\right)^2}{180 \times \frac{1}{18}} + \frac{\left(3 - 180 \times \frac{1}{36}\right)^2}{180 \times \frac{1}{36}} = 8{,}8\dot{6}$$

so erhält man den Wert $8{,}8\dot{6}$. Die Formel allein liefert noch keine genauen Auskünfte darüber, ob diese Verteilung der Idealverteilung entspricht oder nicht. Die Summe sollte allerdings relativ nahe an der Anzahl an Kategorien k, in diesem Fall also 11 sein.[54] 8,86 scheint zwar nahe an 11 zu sein, dennoch sind Begriffe „nahe“ oder „weit entfernt“ nicht unbedingt von mathematischer Signifikanz. Daher wird anhand einer Tabelle, die Wahrscheinlichkeit, dass eine Sequenz echter Zufallszahlen ebenfalls die zu testende Verteilung aufweist, bestimmt (siehe *Tabelle 3*).

[52] Vgl. (Knuth, 1969) Seite 40
[53] Vgl. (Knuth, 1969) Seite 40
[54] Vgl. (Sedgewick, 1992)Seite 586

Die „Degrees of Freedom“ (Freiheitsgraden) *v* entsprechen dem Wert von $k - 1$. Für das Beispiel aus Tabelle 3 ergäbe dies ein *v* von 10, da $k = 11$. Die reellen Zahlen in den Zeilen links von v und den Spalten unter p entsprechen den χ^2 -Werten beziehungsweise dem V. Da ein beliebiges χ^2 in der Tabelle eine angenäherte Wahrscheinlichkeit von ≥p hat[55], lässt sich problemlos erörtern, mit welcher Häufigkeit ein „echter“ Zufallszahlengenerator die errechnete Varianz eines PRNG aufweist.

v	p								
	0,995	0,99	0,9	0,75	0,5	0,25	0,1	0,05	0,01
1	3,93E-05	0,0002	0,0158	0,1015	0,4549	1,3233	2,7055	3,8415	6,6349
2	0,0100	0,0201	0,2107	0,5754	1,3863	2,7726	4,6052	5,9915	9,2103
3	0,0717	0,1148	0,5844	1,2125	2,3660	4,1083	6,2514	7,8147	11,3449
4	0,2070	0,2971	1,0636	1,9226	3,3567	5,3853	7,7794	9,4877	13,2767
5	0,4117	0,5543	1,6103	2,6746	4,3515	6,6257	9,2364	11,0705	15,0863
6	0,6757	0,8721	2,2041	3,4546	5,3481	7,8408	10,6446	12,5916	16,8119
7	0,9893	1,2390	2,8331	4,2549	6,3458	9,0371	12,0170	14,0671	18,4753
8	1,3444	1,6465	3,4895	5,0706	7,3441	10,2189	13,3616	15,5073	20,0902
9	1,7349	2,0879	4,1682	5,8988	8,3428	11,3888	14,6837	16,9190	21,6660
10	2,1559	2,5582	4,8652	6,7372	9,3418	12,5489	15,9872	18,3070	23,2093
11	2,6032	3,0535	5,5778	7,5841	10,3410	13,7007	17,2750	19,6751	24,7250
12	3,0738	3,5706	6,3038	8,4384	11,3403	14,8454	18,5493	21,0261	26,2170
13	3,5650	4,1069	7,0415	9,2991	12,3398	15,9839	19,8119	22,3620	27,6882
14	4,0747	4,6604	7,7895	10,1653	13,3393	17,1169	21,0641	23,6848	29,1412
15	4,6009	5,2293	8,5468	11,0365	14,3389	18,2451	22,3071	24,9958	30,5779
16	5,1422	5,8122	9,3122	11,9122	15,3385	19,3689	23,5418	26,2962	31,9999
17	5,6972	6,4078	10,0852	12,7919	16,3382	20,4887	24,7690	27,5871	33,4087
18	6,2648	7,0149	10,8649	13,6753	17,3379	21,6049	25,9894	28,8693	34,8053
19	6,8440	7,6327	11,6509	14,5620	18,3377	22,7178	27,2036	30,1435	36,1909
20	7,4338	8,2604	12,4426	15,4518	19,3374	23,8277	28,4120	31,4104	37,5662

Tabelle 4: Die X^2-Verteilungstabelle zeigt Wahrscheinlichkeiten (p) abhängig von den Freiheitsgraden (d) und von X^2

Das hier angeführte Beispiel liefert ein χ^2 von $8{,}8\dot{6}$, wodurch sich dank der Tabelle in Tabelle 4 ein p-Wert im Bereich zwischen 0,75 und 0,5 ermitteln lässt. Das bedeutet, dass eine Sequenz von „physikalischen“ Zufallszahlen in ungefähr 50% respektive 75% der Fälle ein χ^2, das größer ist als $\sim 9{,}342$, respektive $\sim 6{,}737$, liefert. Exakter berechnet und mathematisch ausgedrückt hieße dies, dass

$$\mathrm{P}(\chi 2 \geq 8{,}86) = 5\dot{4}{,}48\%$$

[55] Vgl. (Knuth, 1969) Seite 41

Wird, im Gegensatz zur Verteilung in der Tabelle 3, 180 Mal hintereinander eine 7 geworfen, ergibt das eine Quantität χ^2 von 900. Anhand einer ausreichend weitführenden Tabelle lässt sich bestimmen, dass χ^2 in diesem Fall nur in $6{,}3678\times 10^{-185}$% der Fälle größer als oder gleich 900 ist. Dieser Test sollte mehrmals durchgeführt, da er in etwa jedes zehnte Mal ein falsches Ergebnis liefert.[56] Der Informatiker Marsaglia warnte selbst vor einer voreiligen Interpretation jener p-Werte, die integraler Bestandteil einiger von ihm eingeführter Zufallszahlentests sind: *„Such p's happen among the hundreds that DIEHARD produces, even with good RNG's. So keep in mind that ‚p happens'"*[57]

5.4.1.2. KOLMOGOROW-SMIRNOW-TEST

Der Kolmogorow-Smirnow-Test (auch KS-Test), benannt nach zwei russischen Statistikern[58], verhält sich ähnlich dem χ^2-Test von Pearson. Der Unterschied liegt im Groben darin, dass sich der KS-Test auch dann zum Vergleichen erwarteter mit tatsächlichen Verteilungen eignet, wenn nur ein kleines n vorliegt[59] (n sollte für den χ^2-Test groß genug sein, dass die Wahrscheinlichkeit jedes Ereignisses

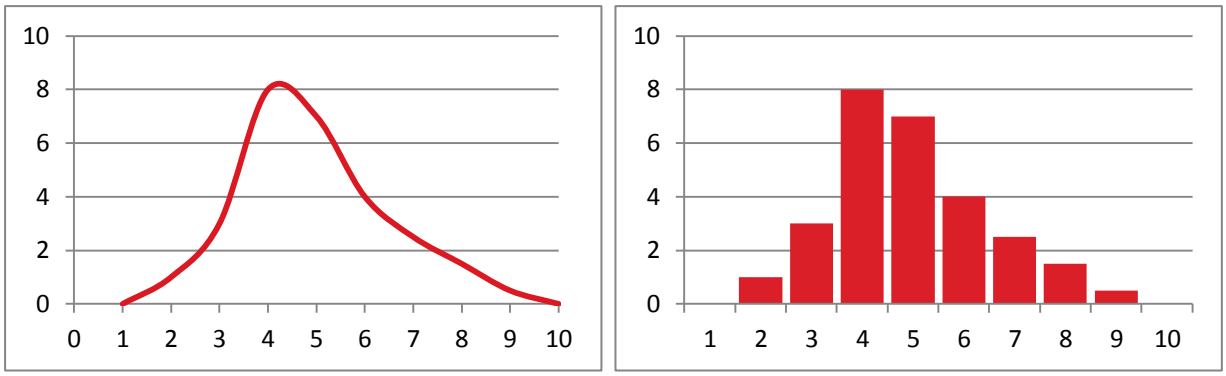

Abbildung 3: Der erste Graph zeigt eine kontinuierliche Wahrscheinlichkeitsfunktion, auf die der Kolmogorow-Smirnow-Test angewendet werden könnte.
Der zweite Graph stellt eine diskrete (sprunghafte) Funktion, wie das Ergebnis eines Würfelspiels, das durch den Chi-Quadrat-Test bewertet werden könnte, dar.

mindestens bei 5 liegt[60]) und gleichermaßen darin, dass er auch bei kontinuierlichen Verteilungen anwendbar ist, wohingegen der χ^2-Test in diesen Fällen lediglich eine Approximation darstellt[61] (siehe *Abbildung 3*). Diese Eigenschaft ist besonders dann von Bedeutung, wenn es sich um Sequenzen von reellen Zufallszahlen handelt, da diese unendlich viele Werte innerhalb des Wertebereichs annehmen können.

[56] Vgl. (Sedgewick, 1992) Seite 586
[57] „Solche ps passieren unter den hunderten, die Diehard produziert, selbst mit guten RNGs. Daher bedenken Sie, dass ‚p passiert'." Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[58] Vgl. (Plate J., 1993) Seite 307
[59] Vgl. (Sachs, 1984) Seite 256
[60] Vgl. (Knuth, 1969) Seite 42
[61] Vgl. (Kolmogorov-Smirnov Test: cs.indiana.edu, 2013)

Um den KS-Test für eine Zufallszahlensequenz zu implementieren, werden ihre Werte zunächst der Größe nach sortiert[62], um, vergleichbar mit dem χ^2-Test, unter Hinzuziehung der erwarteten Wahrscheinlichkeit für einen bestimmten Wert $F_e(x)$ und der tatsächlichen $F_b(x)$, die beiden Werte K^+ und K^- zu berechnen. K^+ bezeichnet hierbei den Wert der größten Abweichung der Sequenz von Zufallszahlen F_b von der Folge der erwarteten Werte F_e, wenn $F_b(x) > F_e(x)$ gilt. Entsprechend bezeichnet K^- die maximale Abweichung, wenn $F_b(x) < F_e(x)$ (siehe *Abbildung 4*) [63]. Die jeweiligen Formeln hierfür lauten:

$$K^+ = \sqrt{n}\ \max_{-\infty<x<\infty}(F_b(x) - F_e(x))$$
$$K^- = \sqrt{n}\ \max_{-\infty<x<\infty}(F_e(x) - F_b(x))$$

Formel 5

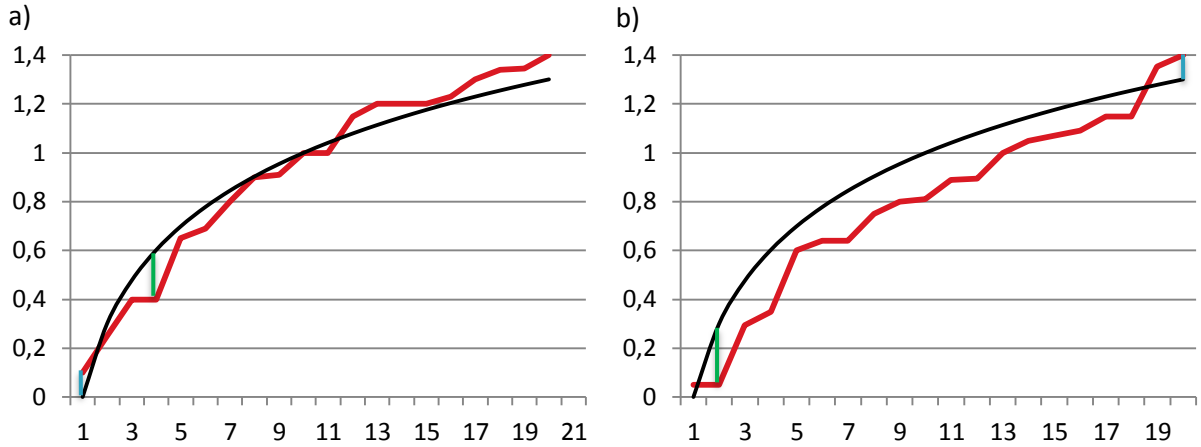

Abbildung 4: Beispiele für Verteilungen (rot), die mal über, mal unter der erwarteten (schwarz) liegen. Die blaue Linie in der Grafik indiziert jeweils $\frac{K^+}{\sqrt{n}}$, den Wert der größten Abweichung, wenn $F_b(x) > F_e(x)$, die grüne $\frac{K^-}{\sqrt{n}}$ für $F_b(x) < F_e(x)$.

Für die beiden Funktionen in Abbildung 4 ergäbe sich durch diese Methode ein K^+ von 0,4472 und ein K^- von 0,9036 für Beispiel a), beziehungsweise ein K^+ von 0,4426 und ein K^- von 1,1272 für Beispiel b).

Parallel zum χ^2-Test kann auch für die K^+ und K^- anhand einer entsprechenden Tabelle (siehe

	$p = 1\%$	$p = 5\%$	$p = 25\%$	$p = 50\%$	$p = 75\%$	$p = 95\%$	$p = 99\%$
$n = 1$	0.01000	0.05000	0.2500	0.5000	0.7500	0.9500	0.9900
$n = 2$	0.01400	0.06749	0.2929	0.5176	0.7071	1.0980	1.2728
$n = 3$	0.01699	0.07919	0.3112	0.5147	0.7539	1.1017	1.3589
$n = 4$	0.01943	0.08789	0.3202	0.5110	0.7642	1.1304	1.3777
$n = 5$	0.02152	0.09471	0.3249	0.5245	0.7674	1.1392	1.4024
$n = 6$	0.02336	0.1002	0.3272	0.5319	0.7703	1.1463	1.4144

Abbildung 5: Wahrscheinlichkeitspunkte für die Verteilungen für K^+ UND K^-

[62] Vgl. (Kolmogorov-Smirnov Test: cs.indiana.edu, 2013)
[63] Vgl. (Knuth, 1969) Seite 47

Abbildung 5) determiniert werden, ob sie signifikant von den erwarteten abweichen.[64] Die Wahrscheinlichkeit, dass K^+ 0,7703 oder weniger ist, beträgt demnach, bei einem Wert von 6 als Stichprobenumfang, 75%.

5.5. UNABHÄNGIGKEIT AUFEINANDERFOLGENDER ZAHLEN

Anders als bei physikalischen Zufallszahlengeneratoren, sind die durch Algorithmen generierten Zufallszahlen einer bestimmten Sequenz voneinander abhängig, da der „output" einer Iteration als „seed" für das nächste Element der Sequenz verwendet wird. Das Kriterium der völligen Unabhängigkeit wird hier folglich nicht erfüllt, allerdings kann der Anspruch gestellt werden, dass die einzelnen Glieder der Kette scheinbar unabhängig voneinander sind. Für einen Menschen, der dreimal hintereinander eine 2 würfelt, mag es unwahrscheinlich erscheinen, dies ein viertes Mal zu tun. Tatsächlich sind die einzelnen Würfe voneinander unabhängig:
Einen Wert zu erhalten macht es nicht unwahrscheinlicher ihn ein zweites Mal zu würfeln.[65]

5.5.1. VERHALTEN IM N-DIMENSIONALEN RAUM

5.5.1.1. HYPEREBENEN-VERHALTEN

Die eben beschriebene Abhängigkeit sich folgender Zufallszahlen manifestiert sich gemäß des Satzes von Marsaglia in sogenannten Hyperebenen. Der US-amerikanische Mathematiker und Informatiker George Marsaglia (1924 - 2011) beschrieb 1968, dass die ansonsten brauchbaren linearen Kongruenzgeneratoren (siehe *6.1 deterministische Zufallszahlengeneratoren*), gewisse „Defekte" aufweisen, die nicht durch das Anpassen der Parameter beseitigt werden können:

> *"If n-tuples (u_1, u_2,. . . u_n), (u_2,u_3,....,u_{n+1}),... of uniform variates produced by the generator are viewed as points in the unit cube of n dimensions, then all the points will be found to lie in a relatively small number of parallel hyperplanes. Furthermore, there are many systems of parallel hyperplanes which contain all of the points.* [66]

> *„Wenn n-Tupel*[67] *(u_1, u_2,. . . u_n), (u_2,u_3,....,u_{n+1}),… von durch den Generator erzeugten, gleichverteilten Zufallszahlen als Punkte in einem Einheitswürfel von n-Dimensionen gesehen werden, dann werden alle Punkte in einer relativ kleinen Anzahl von parallelen Hyperebenen liegen. Weiters gibt es viele Systeme paralleler Hyperebenen, die alle Punkte enthalten."*

Der Defekt der linearen Kongruenzgeneratoren liegt dementsprechend im Bilden jener Hyperebenen im Raum (idealerweise ein Einheitswürfel) (siehe *Abbildung 6*), ein Verhalten, das nicht demjenigen tatsächlich zufälliger Sequenzen entspricht. Die Parameter des linearen Kongruenzgenerators sollten

[64] Vgl. (Knuth, 1969) Seite 48
[65] Vgl. (analysis: random.org, 2013)
[66] (Marsaglia, Random numbers fall mainly in the planes, 1968)
[67] Ein n-Tupel ist eine Anordnung von n mathematischen Objekten. Ein 3-Tupel (Tripel) wäre etwa (0,53,1).

daher so gewählt werden, dass möglichst viele Hyperebenen den Einheitswürfel durchschneiden. Diese Hyperebenen liegen auf *n-1*-dimensionalen Ebenen[68], während die höchste Anzahl an möglichen Hyperebenen bei $(n!\, m)^{\frac{1}{n}}$ liegt (wobei *m* die maximale Sequenzlänge bezeichnet).[69]

a) b)

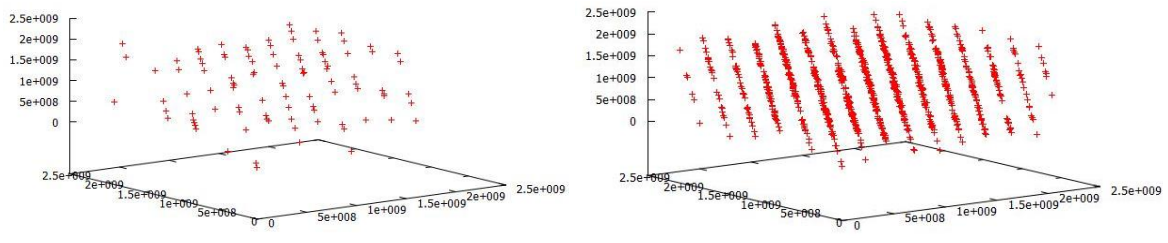

Abbildung 6: Durch den linearen Kongruenzgenerator RANDU erzeugte und in den Raum geplottete Punkte. Die Hyperebenen sind deutlich sichtbar [n für a) = 300; n für b) = 3000]

5.5.1.2. DER SPEKTRALTEST

Die Eigenschaft Hyperebenen zu bilden wird im Zuge des zwar komplexen, allerdings nicht minder bedeutsamen Spektraltests, der für lineare Kongruenzgeneratoren konzipiert wurde, zu einer vollständigen mathematischen Methode. Die Relevanz dieser Vorgehensweise liegt darin, dass die Inadequanz etlicher, heute als schlecht bekannter Zufallszahlengeneratoren, die andere Tests ohne Probleme bestehen, durch den Spektraltest dargelegt werden konnte.[70] So zum Beispiel wurde der berüchtigte Generator RANDU den Anforderungen an Zufallszahlen im 3-dimensionalen Bereich nicht gerecht, wie sich erst durch diesen Test herausstellte.[71]

Die Idee hinter diesem Test, der von Donald Knuth als der „*bei Weitem stärkste bekannte Test*“[72] bezeichnet wurde, stammt von R. R. Coveyou (1915 – 1996) und R. D.MacPherson (1944), die die Zufallszahlen noch als Quellen t-dimensionaler Wellen (Daher auch der Name „Spektraltest“ für Wellenfrequenzen oder Punkte im Spektrum) interpretierten, bevor andere Statistiker später die praktikablere Visualisierung anhand Punkte im t-dimensionalen Raum vorschlugen.[73].

5.5.1.3. BILDEN VON MUSTERN IM 3-DIMENSIONALEN RAUM

Parallel zur Methode von Marsaglia, die für lineare Kongruenzgeneratoren konzipiert wurde, können auch für andere PRNG aufeinanderfolgende 3-Tupel aus einer Sequenz von Zufallszahlen gebildet und in einem 3-dimensionalen Raum dargestellt werden. Um eventuell auftretende Strukturen noch besser zu visualisieren, kann auch die Farbe der Punkte von diesen drei sukzessiven Zahlen abhängig

[68] Vgl. (Mordasini & Klahr, 2013)
[69] Vgl. (Marsaglia, Random numbers fall mainly in the planes, 1968)
[70] Vgl. (Knuth, 1969) Seite 89
[71] Vgl. (Spectral Test: wikipedia.com, 2013)
[72] Im Original: „by far the most powerful test known“ (Knuth, 1969) Seite 89
[73] Vgl. (Knuth, 1969) Seite 110

gemacht werden. Gute PRNG beziehungsweise RNG bilden bei dieser Methode keine sichtbaren Muster und Strukturen (Abbildung 7).[74]

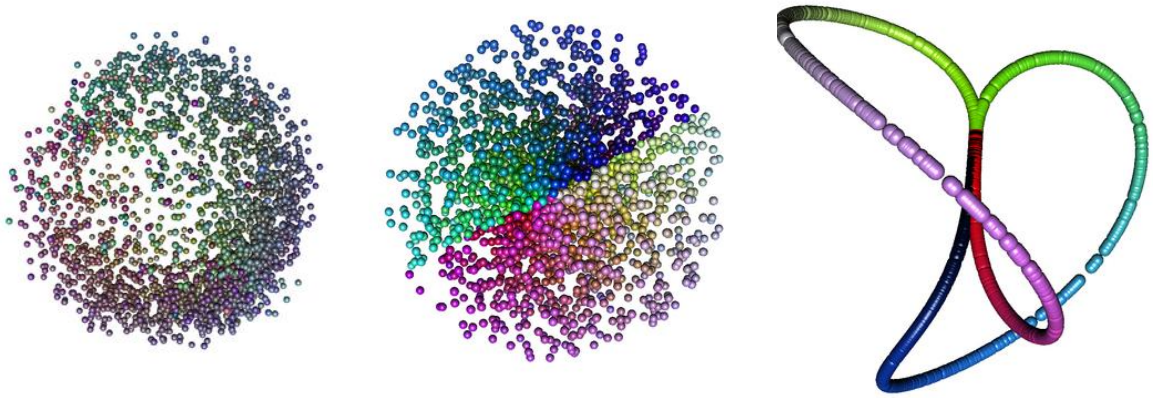

Abbildung 7: In den dreidimensionalen Raum geplottete 3-Tupel, gebildet aus einer Sequenz von Zufallszahlen. Die erste Grafik visualisiert die gleichmäßige Verteilung eines guten PRNG (hier *runif()* von R), die zweite zeigt schon eine klarere Aufteilung der Tupel (erzeugt durch den PRNG *rand* von Excel), während die dritte deutliche Muster bildet und daher einem eher schlechten PRNG (hier ein auf einer logistischen Gleichung basierender Generator) zuzuschreiben ist.

5.6. Eindeutig empirische Tests

Die im Vorhergehenden vorgestellten Gütetests waren partiell theoretischer Natur, das heißt, dass sie auf einem bestimmten, von anderen RNGs unabhängigen, Gedanken oder Anspruch, wie „die Periodenlänge sollte möglichst lang sein“ oder „die einzelnen Zufallszahlen sollen unabhängig von einander sein“ basieren. Manche schlechten RNGs passieren jedoch auch die theoretischeren Prüfmethoden, wie den Spektraltest, deshalb werden in der Praxis auch eindeutig empirische Tests herangezogen, um alle zu erfüllenden Anforderungen abzudecken und zu prüfen. Hauptmerkmal der Methoden dieser Kategorie ist, dass ihre Ergebnisse zunächst mit denen eines perfekten RNGs verglichen werden müssen und nicht per se aussagekräftig sind.

Im Grunde steckt in beinahe jeder Prüfmethode, so auch im χ^2-Test, etwas Empirie, da eine getestete Sequenz mit einer hundertprozentigen Sicherheit nicht tatsächlich zufällig wäre, sollte sie einen derartigen Test nicht bestehen. Eine solche Sicherheit kann allerdings praktisch nie gegeben sein oder angenommen werden.

Da das Erzeugen von Sequenzen von Zufallszahlen ein bedeutender Bestandteil der Mathematik ist, befassen sich seit dem Auftreten der deterministischen RNG einige Mathematiker/innen, Informatiker/innen und Statistiker/innen mit der Problematik des Testens von Zufallszahlengeneratoren. Zwei der bekanntesten, im Nachkommenden vorgestellten empirischen Testsammlungen stammen etwa von Donald Knuth und George Marsaglia. Sie testen insbesondere die Unabhängigkeit aufeinanderfolgender Zahlen, beziehungsweise deren Verteilung, zwei Merkmale

[74] Vgl. (Peterson, 1998) Color Plate 8

auf die schon in *5.4 Verteilung der Zufallszahlen* und 5.5 *Unabhängigkeit aufeinanderfolgender Zahlen* eingegangen wurde.

5.6.1. DIEHARD TESTS

Kurz vor seiner Pensionierung von seinem Lehrposten an der Florida State University, vergab der Informatiker George Marsaglia hochwertige Sequenzen von Zufallszahlen, die er auf CD-ROMs brannte. Die „*Marsaglia Random Number CD-Rom*“[75] enthält 4,8 Milliarden Zufallsbits in 60 Dateien. Diese zufälligen Bits wurden durch das Kombinieren dreier Quellen von elektronischem Rauschen mit dem Output vom neuesten und besten deterministischen RNG (entwickelt von Marsaglia und Zaman) erzeugt. Weiters fügte Marsaglia digitale Tonspuren von Rap- und klassischer Musik, sowie durch Bilder erzeugte Bits in die Sequenz ein. Um die von ihm kreierten Zufallszahlen auf ihre Güte zu prüfen, erfand der Informatiker, zusätzlich zu den bis dahin bestehenden traditionellen Tests, die meist keine Garantie auf brauchbare Zufälligkeit bieten, etliche neue Zufallszahlentests, die unter dem Begriff „Diehard Tests“ zusammengefasst werden.[76]

Marsaglias Tests sind empirischer Natur, das bedeutet, dass die theoretische Grundlage für die Bewertung des PRNGs vernachlässigt wurde, und stattdessen die Ergebnisse eines dieser Tests mit denen eines „perfekten“ RNGs verglichen werden.

Da die der Diehardbatterie entstammenden Tests in Abschnitt *6. Zufallszahlengeneratoren* als Mittel zur Bestimmung der Güte der jeweiligen, dort vorgestellten Generatoren dienen, soll an dieser Stelle besonders umfassend auf diese eingegangen werden.

5.6.1.1. GEBURTSTAGSABSTÄNDE

Der erste dieser Tests ist angelehnt an das sogenannte Geburtstagsproblem, das die 50-prozentige Wahrscheinlichkeit beschreibt, dass sich in einer Gruppe von 23 Menschen zumindest zwei einen Geburtstag (unabhängig vom Geburtsjahr) teilen[77] und integriert mit dem χ^2-Test und dem Kolmogorow-Smirnow-Test, zwei der wichtigsten Gütetests für PRNG.

Marsaglia interpretierte für diesen Test die Größe n als Anzahl an Tagen in einem Jahr, m als zufällig gewählte Tage (Geburtstage), und j als Anzahl von Abständen, die öfter als einmal vorkommen.[78]

Allgemein lässt sich sagen, dass die Elemente m einer Sequenz von Zufallszahlen mit einer Länge n aufsteigend sortiert werden. Nach der Bestimmung der Differenzen $(Y_n - Y_{n-1};\ Y_{n-1} - Y_{n-2};\ \dots\ ;\ Y_3 - Y_2; Y_2 - Y_1)$ einander nachfolgender Zahlen, wird die Anzahl an Paaren solcher

[75] Abrufbar auf: (Marsaglia, CDRom: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[76] Vgl. (Peterson, 1998) Seite 178 ff
[77] Vgl. (Peterson, 1998) Seite 89
[78] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

Geburtstage mit einem Abstand von mehr als einem Tag [79] ermittelt. Dieses $j = (Anzahl(Vorkommnissen(\Delta Y > 1) > 1)$ ist Poisson-verteilt, mit

$$\mu = \lambda = \frac{m^3}{4n}$$

Formel 6

(siehe *Abbildung 8*). Wird beispielsweise ein Jahr, bestehend aus fünfzehn Tagen (n), gewählt und

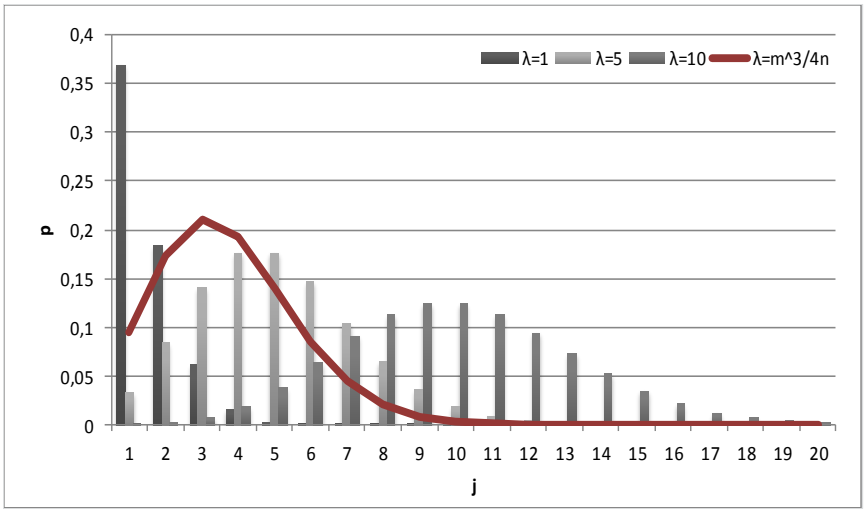

Abbildung 8: Poisson-Verteilung mit unterschiedlichen λ, wobei m (Geburtstage)= 20 und n (Tage im Jahr)=365

generiert weiters der zu testenden Zufallszahlengenerator die vier (m) Zahlen 3, 6, 10 und 14, so sieht die Aufteilung wie folgt aus:

|…1…|…2…|…✕…|…4…|…5…|…✕…|…7…|…8…|…9…|…✕…|…11…|…12…|…13…| …✕…| …15…|

Die Folge der Abstände lautet demnach (6-3; 10-6; 14-10) oder (3; 4; 4). Obwohl zwei unterschiedliche Abstände (3 und 4) vorkommen, ist nur der Abstand mit dem Wert 4 von Bedeutung für die Statistik, da dieser mehr als einmal vorkommt. J ist daher in diesem Fall 1.

In der Praxis werden viel größere Werte für m und n eingesetzt, George Marsaglia sieht beispielsweise ein n von 2^{24} und ein m von 2^9 vor. Der weitere Testvorgang besteht aus zwei Teilen, wobei der erste die Poisson-verteilten j mehrere Male einem χ^2-Test unterzieht und der zweite einen KS-Test auf die dadurch entstehenden p-Werte anwendet.[80]

Werden zwei RNGs mit unterschiedlicher Qualität, wie etwa der KISS-Generator und ein Fibonacci-Generator (siehe *6.1 Deterministische Zufallszahlengeneratoren*), getestet, so mögen die erwarteten und schlussendlich ermittelten Werte wie in den folgenden beiden Tabellen aussehen.

[79] Vgl. (Birthday-spacings: software.intel-Website, 2013)
[80] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

Doppelt vorkommende Abstände (j)	Beobachteter Wert	Erwarteter Wert
0	66	67.668
1	156	135.335
2	129	135.335
3	88	90.224
4	40	45.112
5	17	18.045
6 oder mehr	4	8.282
p-Wert (mit v=6) ≈ 0,620230		

Tabelle 6: Verhalten eines guten RNG (KISS) im „Geburtstagsabstände-Test“

Doppelt vorkommende Abstände (j)	Beobachteter Wert	Erwarteter Wert
0	17	67.668
1	59	135.335
2	114	135.335
3	95	90.224
4	90	45.112
5	62	18.045
6 oder mehr	63	8.282
p-Wert (mit v=6) ≈ 1,000000		

Tabelle 5: Verhalten eines schlechten RNG (Fibonacci) im „Geburtstagsabstände-Test“

5.6.1.2. AFFENTEST (ORIGINAL: MONKEY TEST) ODER „BITSTREAM“-TEST

1913 beschrieb der französische Mathematiker Émile Borel (1871 – 1956) in seiner Abhandlung *„La mécanique statique et l'irréversibilité“* ein Gedankenexperiment, das später namensgebend für einen von Marsaglias Diehard Tests wurde: Eine Million speziell dafür dressierter Affen werden dazu angewiesen zehn Stunden pro Tag zufällig auf die Tasten einer Schreibmaschine zu schlagen. Nach einem Jahr sollen die zusammengeführten Texte der Affen exakte Kopien aller naturwissenschaftlichen Bücher der größten Bibliotheken der Welt enthalten. [81] Die Wahrscheinlichkeit dieses Ereignisses ist so gering, dass es beinahe als unmöglich gewertet werden kann. Allein die Wahrscheinlichkeit, dass diese Affen in einem vergleichbaren Experiment im Laufe eines Jahres diesen hier vorliegenden Absatz getippt hätten beläuft sich auf circa $6{,}6638 \times 10^{-154}$% (Equation 1).

$$\binom{n}{k} p^k (1-p)^{n-k} = \binom{72000}{1} (9.255235 \times 10^{-161})^1 (1 - 9.255235 \times 10^{-161})^{72000-1} = \\ = 6{,}6638 \times 10^{-156}$$

Equation 1: Die Binomialverteilung liefert die Wahrscheinlichkeit, dass die Affen innerhalb von 72000 Versuchen (n) versuchen per Zufall einmal (k) jenen Text zu schreiben, den erfolgreich einzutippen beim ersten Versuch mit einer Wahrscheinlichkeit von 9.255235E-161 (p) gelingt. Dies gilt für den Fall, dass der Text aus circa 90 Zeichen besteht, wobei angenommen wird, dass es drei Minuten braucht diesen einzutippen.

(1,0,0,1,1,0,1,0,1,0,0,1,0,0,1,1,1,1,1,0,1,1,1,1,0,0,1,1,0,0,1,0,1,0,
1,1,0,0,1,1,1,1,0,1,1,1,1,0,0,1,1,1,0,1,0,0,0,0,0,0,1,1,0,0,0,1,1,1
,0,1,0,0,0,1,1,0,0,0,0,0,1,1,0,1,0,0,0,0,0,1,0,1,0,1,1,0,1,1,1,0,1)

Sequenz 3: 100 Zufallsbits; grau hinterlegt ist das erste Wort, bestehend aus den Bits b_1=1, b_2=0, ... b_{20}=0. Die rote Schrift bedeutet das zweite Wort, bestehend aus den Bits *b_2=0, b_3=0 ...b_{21}=1*.

[81] Vgl. (Borel, 1913)

Parallel zu Borels Gedankenexperiment befasst sich auch Marsaglias Testmethode für die Güte von PRNG mit dem scheinbar arbiträren Generieren von Worten aus einer Sequenz von Zufallszahlen. Eine Sequenz von Zufallsbits (wie Sequenz 3) mit den Elementen b_1, b_2, ... wird als Abfolge von aus zwei „Buchstaben“ (0 und 1) bestehenden Wörtern interpretiert. Diese „Wörter“ überlappen sich und bestehen aus 20 solcher „Buchstaben“, das erste „Wort“ wird entsprechend von den Bits $b_1,b_2...b_{20}$, das zweite von den Bits $b_2,b_3...b_{21}$ konstituiert, und so weiter, bis die gesamte Sequenz in

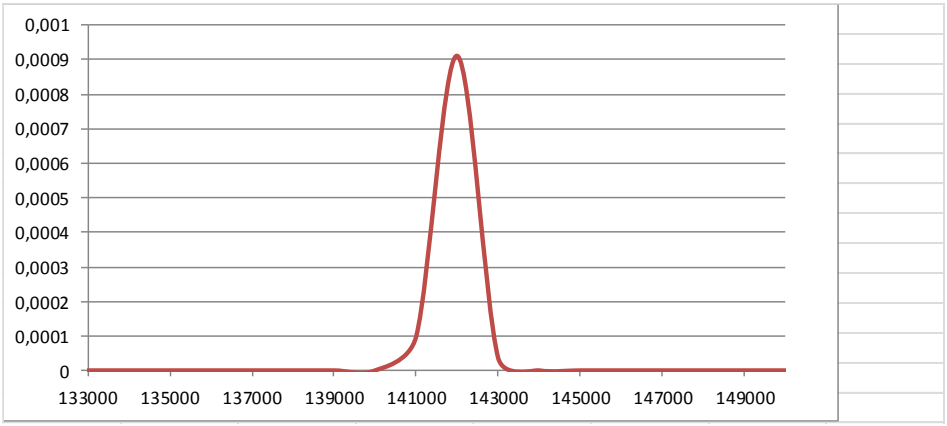

Abbildung 9: Ideale Verteilung der j des Affentests (μ=141 909 Sigma σ=428).

derartige Wörter umgewandelt ist. [82]

Es gibt 2^{20} verschiedene Möglichkeiten die Bits 0 und 1 innerhalb eines 20 Bits langen Wortes anzuordnen. Unter diesem Wissen werden für den Affentest die Kombinationen gefunden, die in der Sequenz „fehlen“, die Anzahl der fehlenden Kombinationen wird mit dem Buchstaben „j“ bezeichnet. Statistische Tests zeigen, dass für eine 2097171 Bits lange Sequenz von tatsächlich zufälligen Zahlen j normalverteilt wie Abbildung 9 sein soll.[83]

5.6.1.3. OPERM5[84]

Der OPERM5 Test unterteilt ganze Zahlen einer Zufallssequenz in überlappende Gruppen von jeweils fünf ganzen Zahlen. Daher ergeben sich für die einzelnen Gruppen $5! = 120$ verschiedene Ordnungs-möglichkeiten, beziehungsweise Zustände. Infolgedessen wird die Häufigkeit jeder dieser Zustände ermittelt und ein χ^2-Test auf die Häufigkeiten angewendet.

[82] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[83] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[84] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

5.6.1.4. RÄNGE VON MATRIZEN

Marsaglia beschrieb im Rahmen des Matrizentests drei geringfügig unterschiedliche Varianten. Ihnen allen ist gemeinsam, dass eine gewisse Anzahl von Bits einer Zufallszahlensequenz als Element einer $m \times n$ binären Matrix betrachtet wird. Daraufhin wird der Zeilenrang dieser Matrix bestimmt.[85]

Der Zeilenrang einer Matrix kann zwischen 0 und m liegen und hängt im Wesentlichen von der Anzahl an Dimensionen einer Zeile ab. Der Rang beziehungsweise die Dimension wird in folgenden Fällen verringert:

1. Die Zeile enthält nur „0“en
2. Die Zeile ist identisch mit, proportional zu oder eine lineare Kombination einer anderen[86]

Je nach Variation des Tests ergeben sich unterschiedliche Wahrscheinlichkeiten für die Spaltendimensionen. Wird eine 31×31 Matrix erzeugt, so kann die Spaltendimension theoretisch zwischen 0 und 31 liegen, wird einen Wert von 28 jedoch selten unterschreiten. In diesem Fall wird ein χ^2-Test für die Spaltendimensionen 31,30,29 und ≤28 durchgeführt und ein entsprechender p-Wert errechnet.[87] Die erwarteten Werte liefern die in Tabelle 7 angeführten Wahrscheinlichkeiten:

Spaltenrang	Wahrscheinlichkeit in einem perfekten RNG
31	0.289
30	0.578
29	0.128
≤28	0.005

Tabelle 7

Die beiden Variationen neben der, die eine 31×31 Matrix erzeugt, behandeln einerseits die Ränge einer 32×32 Matrix, mit einem Mittelwert μ von 31 und andererseits die einer 6×8 Matrix, mit einem Mittelwert μ von 6.[88]

5.6.1.5. ZÄHLE DIE 1EN

Der nächste aus Marsaglias Diehard Testsammlung stammende Gütetest befasst sich mit der

(1,0,0,1,1,0,1,0, 1,0,1,1,0,0,1,1, 1,1,1,0,1,1,1,1,
0,0,1,0,0,1,0,1, 0,1,1,0,0,1,1,1, 1,0,1,1,1,1,0,0,
1,1,1,0,1,0,0,0, 0,0,0,1,1,0,0,0, 1,1,1,0,1,0,0,0)

Sequenz 4: 72 Zufallsbits, hervorgehoben ist der erste „Buchstabe“, unterstrichen das erste „Wort“ der Sequenz

[85] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[86] Vgl. (Matrix-Rank: easycalculation.com, 2013)
[87] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[88] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

Häufigkeit eines bestimmten Elements in einer Sequenz von Bits und heißt entsprechend „Zähle-die-1en"-Test. Hier wird die Sequenz in Bytes zu jeweils 8 Bits aufgeteilt, in denen folglich jeweils 0 bis 8 Mal die Ziffer „1" vorkommen kann. Die Wahrscheinlichkeiten für diese Häufigkeiten (siehe *Tabelle 8*) bilden gewissermaßen eine Glockenkurve. Nun werden die einzelnen Bytes als Buchstaben interpretiert, die von der Anzahl an 1en in ihrer 8-gliedrigen Sequenz von Bits determiniert werden[89] (siehe *Tabelle 8*).

Anzahl an 1en	Wahrscheinlichkeit	Entsprechender Buchstabe	Wahrscheinlichkeit für den entsprechenden Buchstaben
0	0,00390625	A	0,14453125
1	0,03125	A	0,14453125
2	0,109375	A	0,14453125
3	0,21875	B	0,21875
4	0,2734375	C	0,2734375
5	0,21875	D	0,21875
6	0,109375	E	0,14453125
7	0,03125	E	0,14453125
8	0,00390625	E	0,14453125

Tabelle 8

Die Sequenz 4 besteht aus 72 Zufallsbits, respektive 12 Bytes. Das erste Byte (**1**,0,0**,1,1**,0,**1**,0) besteht aus insgesamt vier 1en und wird daher als C interpretiert. Marsaglias Test besagt weiters, dass fünf dieser Buchstaben zu einem ganzen Wort zusammengefasst werden sollen. Das erste Wort der Sequenz 3 ist folglich CDEBD (**1**,0,0,**1**,**1**,0,**1**,0, **1**,0,**1**,**1**,0,0,**1**,**1**, **1**,**1**,**1**,0,**1**,**1**,**1**,**1**, 0,0,**1**,0,0,**1**,0,**1**, 0,**1**,**1**,0,0,**1**,**1**,**1**).

Es gibt 5^5 verschiedene Möglichkeiten für derartige Wörter. Für eine Sequenz von 256 000 überlappenden „Wörtern" aus 5 Buchstaben, wie sie in der Diehardbatterie zum Einsatz kommt, wird die Frequenz jedes einzelnen Wortes ermittelt. Ein χ^2-Test liefert p-Werte, die die Interpretation der Ergebnisse ermöglichen.[90]

5.6.1.6. OPSO, OQSO UND DNA[91]

Diese drei Tests verhalten sich insofern ähnlich zum vorhergehenden, als dass auch hier gewisse Kombinationen von Bits als Buchstaben interpretiert werden, um daraus entstehende, fehlende „Wörter" zu ermitteln.

[89] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[90] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[91] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

OPSO steht für *Overlapping-Pairs-Sparse-Occupancy (Überlappende-Paare-spärliche-Belegung)*. Anders als der Zähle-die-1en-Test, der Wörter aus fünf Buchstaben betrachtet, verwendet OPSO durch einen Zufallszahlengenerator generierte Wörter, die aus zwei Buchstaben (eines Alphabets von 1024 Elementen) bestehen. Der Test generiert 2^{21} solcher überlappender Wörter und bestimmt die Anzahl an „fehlenden" Worten. Da die dadurch entstehende Kurve einer Normalverteilung von μ 141 909 und σ von 290 folgen sollte, kann die Abweichung von der idealen Verteilung und darauf hin ein p-Wert (Wahrscheinlichkeit, dass ein perfekter Zufallsgenerator dieselbe Verteilung bildet) bestimmt werden.

OQSO steht für *Overlapping-Quadruples-Sparse-Occupancy (Überlappende-Quadrupel-spärliche-Belegung)*. Hier werden Wörter aus vier Buchstaben und einem Alphabet von 32 Elementen betrachtet. Erneut werden 2^{21} solcher Wörter generiert und durch den Test die Anzahl an „fehlenden" Worten festgelegt. Die dadurch entstehende Kurve sollte einer Normalverteilung von μ 141 909 und σ von 295 folgen.

Der DNA-Test geht von einem Alphabet aus 4 Elementen beziehungsweise den Buchstaben C, G, A und T aus, die von zwei bestimmten Bits der Sequenz von Zufallszahlen determiniert werden. Generiert ein RNG Wörter aus jeweils zehn derartigen Buchstaben, so gibt es 2^{20} Möglichkeiten für diese. Bei einer Sequenz von 2^{21} einander überlappender Wörter sollte der Mittelwert μ der fehlenden Wörter bei 141 909 und die Standardabweichung σ bei 339 liegen.

5.6.1.7. PARKPLATZTEST

Der χ^2 Test zum Determinieren der Güte eine Zufallszahlensequenz ist nur dann sinnvoll, wenn n, die Länge der Sequenz, groß genug ist, dass jeder Wert in den jeweiligen Kategorien theoretisch zumindest 5-mal vorkommt.[92] Der Parkplatztest, in Donald Knuths Testsammlung noch als Kollisionstest bezeichnet, ist auch noch dann gültig, wenn die Anzahl der Kategorien viel größer ist als n, also die Anzahl an generierten Zahlen.[93]

Gegeben ist ein Quadrat mit der Länge 100, in das nach vom Zufallsgenerator festgelegten Koordinaten „Autos geparkt" werden, die, als Kreise betrachtet, einen Radius von 1 haben (Abbildung 10). Wenn eines dieser Autos in ein Einheitsquadrat, wo schon ein anderes steht, zu stellen versucht wird, führt dies zu einem „crash".[94]

Donald E. Knuth schlug in seinem Werk „*The Art of Computer Programming*" einen sehr ähnlichen Gütetest vor, wobei bei ihm Bälle, die in Urnen geworfen werden sollen für die Autos, die auf freie

[92] Vgl. (Knuth, 1969) Seite 42
[93] Vgl. (Knuth, 1969) 68
[94] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

beziehungsweise besetzte Parkplätze gestellt werden, stehen. Beim Kollisionstest, wie auch beim Parkplatztest, wird die Anzahl an Parkplätzen, respektive Urnen weitaus höher sein als die Anzahl an Autos beziehungsweise Bällen.[95]

Wenn die Anzahl an Versuchen diese Autos zu parken als n, und die Anzahl an tatsächlich erfolgreichen Autos als k gesehen wird, dann sollte die entstehende Kurve für k derer von echten Zufallszahlengeneratoren ähnlich sein. Dies bedeutet, dass, bei einem n von 12 000, k

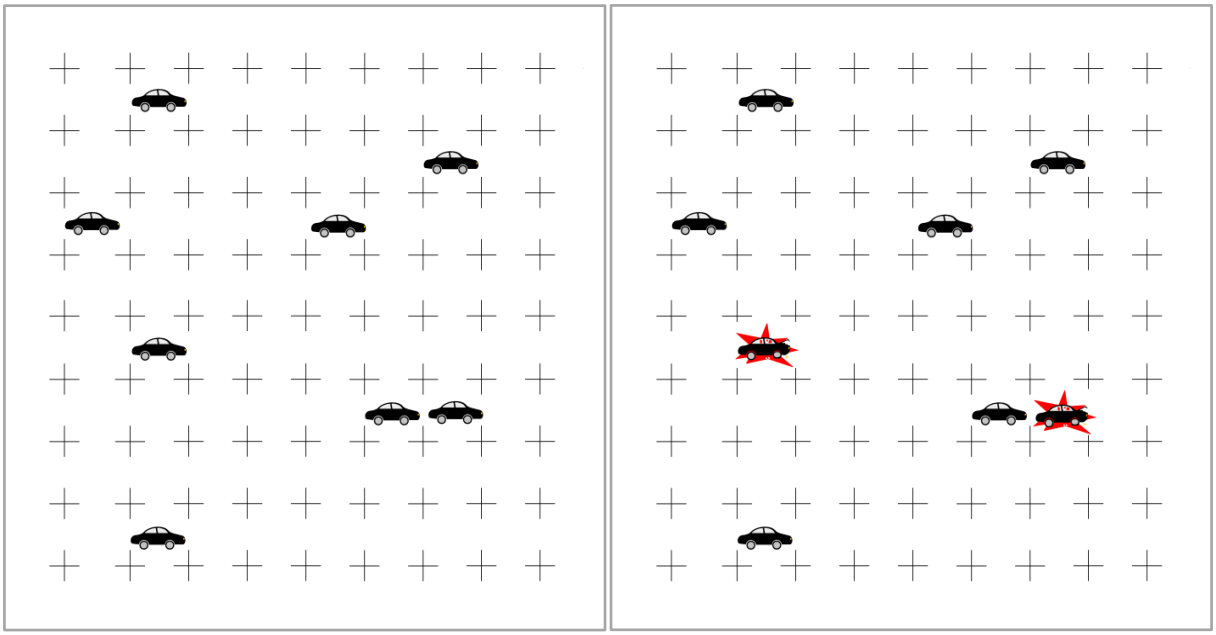

Abbildung 10: Geparkte Autos in einem Quadrat mit Länge 100. Das erste Quadrat zeigt 7 erfolgreich geparkte Autos (n=8, k=8), das zweite weist zusätzlich zwei Zusammenstöße vor (n=10, k=8); eine Auto wurde dorthin zu stellen versucht, wo schon ein anderes stand

durchschnittlich bei 3523 (mit einer Standardabweichung σ 21,9) liegen und nahe an einer Normalverteilung sein sollte. Durch die Anwendung eines Kolmogorow-Smirnow-Tests lässt sich feststellen, ob die für den PRNG errechneten Werte dem eines echten RNG entsprechen.[96]

5.6.1.8. MINIMUMDISTANZTEST

Hier werden 8000 Punkte in ein Quadrat mit einer Seitenlänge von 10.000 geplottet, wobei die Koordinaten der Punkte vom getesteten Zufallszahlengenerator determiniert werden. Zwischen den $\frac{n^2-n}{2}$ Punktepaaren wird die kleinste Distanz d bestimmt. Sind die Punkte wirklich unabhängig voneinander, dann sollte d^2 nahe an einer exponentiellen Verteilung mit einem Mittelwert von 995 liegen. Ein KS-Test liefert genauere Auskünfte über die Wahrscheinlichkeit, dass die von dem zu

[95] Vgl. (Knuth, 1969) Seite 69

[96] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

testenden PRNG erzeugten Verteilungen für die Minimumdistanz auch von einem „echten“ Zufallszahlengenerator stammen könnten. [97]

5.6.1.9. ZUFÄLLIGE-KUGELN-TEST

Der Zufällige-Kugeln-Test ist gewissermaßen das „Übersetzen“ des Minimumdistanz-tests in eine höhere Dimension:

In einem Würfel mit einer Seitenlänge von 1000 werden 4000 zufällige Punkte geplottet, wobei die Koordinaten der Punkte vom getesteten Zufallszahlengenerator determiniert werden. An jeden Punkt wird der Mittelpunkt einer Sphäre gesetzt, deren Radius groß genug sein soll, dass er den nächsten Punkt erreicht. Das Volumen der kleinsten derartigen Sphäre sollte nahe an einer exponentiellen Verteilung mit einem Mittelwert von $\frac{120\pi}{3}$ sein, der Radius ist daher exponentiell mit einem Mittelwert von 30 (siehe *Abbildung 11*) verteilt. Der Test wird 20 Male wiederholt und wie beim Minimumdistanztest werden die Ergebnisse mittels eines KS-Tests zusammengefasst.[98]

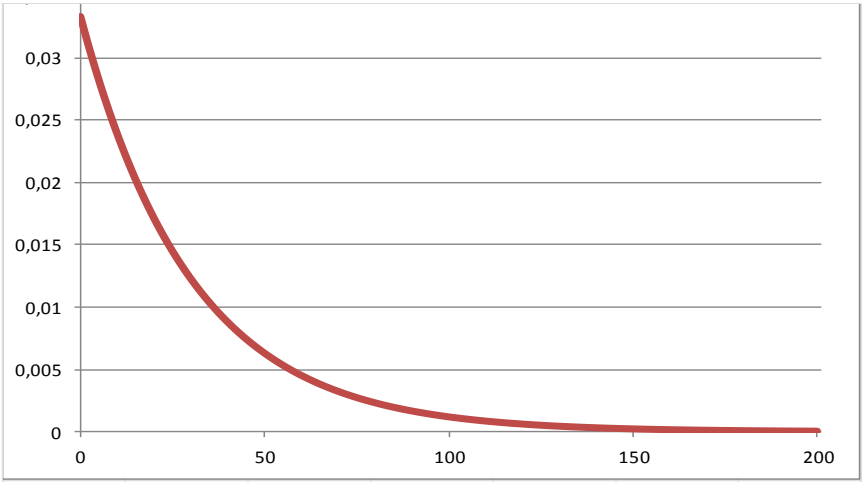

Abbildung 11: Ideale Verteilung der Radien für den Zufällige-Kugeln-Test

5.6.1.10. „SQUEEZE“-TEST

Beginnend mit der Zahl 2^{31} wird eine Zahl solange mit einer zufälligen reellen im Intervall [0,1) multipliziert, bis das Produkt der Rechnung 1 ergibt. Die Formel hierzu lautet

$$k_{n+1} = ceil(k_n \times U)$$

wobei U die reelle Zufallszahl im Intervall [0,1) bedeutet und die Abrundungsfunktion *ceil(x)* die nächsthöhere ganze Zahl zur reellen Zahl, die bei der Multiplikation $k \times U$ entsteht, wiedergibt. J beschreibt die Anzahl an Iterationsverfahren, die nötig sind, um k zu 1 zu reduzieren (siehe Beispiel in Tabelle 9). Schließlich wird ein χ^2-Test auf die dadurch ermittelten J eingesetzt. [99]

[97] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[98] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[99] Vgl. ebd.

Anzahl der Iterationen (J)	$ceil(k_n \times U)$	k_{n+1}
1	ceil(2^{31}×0,27591642)	592525994
2	ceil(592525994×0,91324721)	541122713
3	ceil(541122713×0,26907047)	145600141
4	ceil(145600141×0,35470658)	51645328
5	ceil(51645328×0,29070892)	15013758
6	ceil(15013758×0,08123641)	1219664
7	ceil(1219664×0,44763117)	545960
8	ceil(545960×0,64273248)	350907
9	ceil(350907×0,86459341)	303392
10	ceil(303392×0,73081335)	221723
11	ceil(221723×0,37314823)	82736
12	ceil(82736×0,8183026)	67704
13	ceil(67704×0,32596247)	22069
14	ceil(22069×0,73535518)	16229
15	ceil(16229×0,18738962)	3042
16	ceil(3042×0,82255788)	2503
17	ceil(2503×0,38265519)	958
18	ceil(958×0,06489883)	63
19	ceil(63×0,46946908)	30
20	ceil(30×0,96936825)	30
21	ceil(30×0,22470317)	7
22	ceil(7×0,78735285)	6
23	ceil(6×0,03317527)	**1**

Tabelle 9: Durchgeführter „Squeeze"-Test. Nach 23 Iterationen wurde k zu 1 reduziert (j=23)

5.6.1.11. ÜBERLAPPENDE-SUMMEN-TEST

Für den überlappende-Summen-Test (Original: „Overlapping sums test") wird eine Sequenz aus reellen Zufallszahlen U(1),U(2),... erzeugt, wobei alle U im Intervall [0,1) liegen. Anschließend an die Bildung überlappender Summen S(1)=U(1)+...+U(100), S2=U(2)+...+U(101),... aus dieser Sequenz, wird ein KS-Test auf die unterschiedlichen S angewandt. Der Test liefert 10 p-Werte, die wiederum als Input für einen weiteren KS-Test dienen.[100]

5.6.1.12. LÄUFE-TEST

Im Zuge des Läufe-Tests wird eine Sequenz von Zufallszahlen im Intervall [0,1) erzeugt. Daraufhin werden „Rauf-Läufe" („*runs up*") und „Runter-Läufe" („*runs down*") bestimmt.[101]

Die Sequenz

(0,365 **0,356** **0,968** **0,341** **0,347** **0,454** **0,902** **0,961** 0,667 0097)

beispielsweise beginnt mit einen „run down" (grau unterlegt) von der Länge 2, wird durch einen „run up" (unterstrichen und fett formatiert) und anschließend einem „run down" der Länge 2 fortgesetzt, schließlich von einem „run-up" der Länge 5 weitergeführt und einem „run down" der Länge 3 abgeschlossen.

[100] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

[101] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

Donald Knuth rät von einem direkten χ^2-Test der Unabhängigkeit auf die „runs down“ und „runs up“ ab, da die Läufe in ihrer Natur nicht unabhängig voneinander sind und beispielsweise ein langer Lauf meist von einem kurzen nachgefolgt wird und vice versa.[102]

Stattdessen wendet Marsaglia eine kompliziertere Methodik über eine Kovarianzmatrix für die „Rauf-Läufe“ und „Runter-Läufe“ an, die letzten Endes ebenfalls p-Werte, auf die ein KS-Test anwendet werden kann, liefert.[103]

5.6.1.13. „CRAPS“-TEST[104]

Craps ist ein im englischsprachigen Raum beliebtes Würfelspiel, bei dem im Grunde die Augensummen eines Würfelpaares über Gewinn oder Niederlage entscheiden. Im Zuge dieses Zufallszahlentests werden 200 000 „Spiele“ davon gespielt, also das Ergebnis der Würfe durch den RNG simuliert, um die „Anzahl an Gewinnen“ p und die der „Würfe“ t, die nötig sind um die jeweiligen Spiele zu beenden, zu errechnen.

Die Anzahl der Siege p sollte nahe an einer Normalverteilung mit bestimmten Parametern für den Mittelwert und die Standardabweichung liegen.

Für den χ^2-Test werden alle t (Würfe, die nötig sind um die jeweiligen Spiele zu beenden), für die ≥21 gilt in einer Kategorie zusammengefasst und die Wahrscheinlichkeit des jeweiligen Ergebnisses berechnet.

5.6.2. METHODEN NACH DONALD KNUTH

1962 begann Donald Knuth, damals noch Student am „Case Institute of Technology“ nach Anfragen eines von Knuths Begabung für die Informatik und Mathematik in Kenntnis gesetzten Verlages damit, ein Buch über Compiler zu verfassen. Aus einem Projekt, dass ursprünglich nicht mehr als eine Ausgabe umfassen sollte, wurde das 7-bändige Standardwerk der Informatik *„The Art of Computer Programming“*, von dem gegenwärtig noch drei Bände in Planung sind.[105] Der zweite davon enthält den Abschnitt über Zufallszahlen, der als Rahmen für Knuths Vorschläge für Zufallszahlentests, von denen allerdings ein Großteil auf anderen weithin bekannten Tests basiert, dient.

5.6.2.1. FREQUENZ- ODER GLEICHVERTEILUNGSTEST

Um die Gleichverteilung im Intervall [0,1) zu testen, soll nach Knuth entweder ein χ^2-Test (siehe *5.4.1.1*), beziehungsweise ein Kolmogorow-Smirnow-Test (siehe *5.4.1.2*), [106] je nach zur Verfügung stehendem Wertebereich und der Menge an generierten Zufallszahlen, angewendet werden.

[102] Vgl. (Knuth, 1969) Seite 65
[103] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[104] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[105] Vgl. (The Art of Computer Programming: wikipedia, 2013)
[106] Vgl. (Knuth, 1969)Seite 59

5.6.2.2. SERIENTEST

Nicht nur einzelne Zahlen, auch Zahlenpaare sollen gleichmäßig und unabhängig voneinander verteilt sein.

Eine beliebige Sequenz ganzer Zahlen wird für $0 \leq j < n$ in die Zahlenpaare $(Y_{2j}, Y_{2j+1}) = (q, r)$ aufgeteilt und ein χ^2-Test wird für die unterschiedlichen Kategorien beziehungsweise Arten von Paarungen (d^2, wobei d den Wertebereich bezeichnet) angewandt. Der Test kann auch zu einer Serie von Tripel, Quadrupel und n-Tupel erweitert werden.[107]

Der Wertebereich im Beispiel aus Sequenz 5 liegt im Intervall [1,9], es gibt daher 9^2 verschiedene Anordnungsmöglichkeiten. Das erste Paar heißt $(Y_{2j}, Y_{2j+1}) = (5,2)$, das zweite $(Y_{2j+2}, Y_{2j+3}) = (7,7)$ etc., wobei jede Paarung eine Wahrscheinlichkeit von $\frac{1}{9^2}$ hat.

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

Sequenz 5

Inwiefern dieser Test als eine wichtige Ergänzung zum Gleichverteilungstest fungiert, beschreibt Donald E Knuth sehr treffend, wenn er schreibt: *„The sun comes up just about as often as it goes down, in the long run, but this doesn't make its motion random"*[108]

Die Sequenz (0,1,0,1,0,1,0,1,0,1,0,1,..) mag den herkömmlichen Gleichverteilungstest ohne Probleme bestehen, da jedes Element, 0 beziehungsweise 1, mit exakt der erwarteten Häufigkeit auftreten, allerdings bestätigt nur der Serientest den Verdacht, der sich einem/einer menschlichen Beobachter/in sofort auftut, dass die Sequenz tatsächlich nicht zufällig ist.

Bei idealer Verteilung sollten die Paare (0,0), (0,1), (1,0) und (1,1) mit derselben Häufigkeit in der Folge auftreten[109], dennoch ist dies im hier vorliegenden Beispiel, dessen Versinnbildlichung in gewisser Hinsicht die Bewegung der Sonne im Himmel repräsentieren soll, nicht der Fall.

5.6.2.3. LÜCKENTEST

Der Lückentest wird verwendet um die Länge der „Lücke" zwischen zwei Vorkommnissen gewisser, in einem festgelegten [α,β]-Intervall von Werten liegender Zahlen zu ermitteln (siehe *Abbildung 12*). [110]

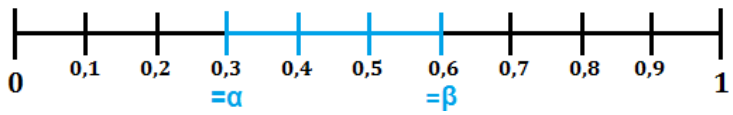

Abbildung 12: Das betreffende Intervall heißt hier [0,3=α,β=0,6]. Wenn eine Zufallszahl der zu testenden Sequenz in dieses fällt, so wird sie U_j genannt, die nächste in das Intervall fallende Zahl ist U_{j+r+1}. Die Anzahl aller Werte die zwischen den Vorkommnissen von U_j und U_{j+r+1} liegen ist r, die Länge der Lücke.

[107] Vgl. (Knuth, 1969) Seite 60
[108] Die Sonne geht so oft auf, wie sie untergeht, doch schlussendlich macht das ihre Bewegung nicht zufällig. (Knuth, 1969)Seite 60
[109] Vgl. (pRNGs: math.umn.ed, 2013) Seite 10
[110] Vgl. (Knuth, 1969) Seite 60

Knuth sieht vor, diesen Test auf eine Sequenz von reellen Zufallszahlen zwischen 0 und 1 (daher gilt $0 \leq \alpha < \beta \leq 1$)[111] anzuwenden, dennoch kann er, wenn es die Praxis erfordert, auch auf ganze Zahlen ausgeweitet werden. In diesem Falle fällt die Voraussetzung, zunächst ein Intervall an Zahlen zu wählen, weg (siehe *Abbildung 13*).

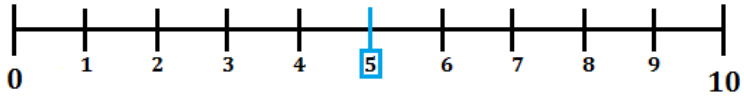

Abbildung 13: Bei ganzen Zahlen ist kein Intervall nötig. In diesem Fall ist r, die Länge der Lücke, die Anzahl an Werten, die zwischen allen Vorkommnissen der Zahl 5 liegen.

Auch aus der Sequenz 5 lässt sich beispielsweise eine Folge von der Länge der Lücken zwischen Auftreten eines bestimmten Werts bilden. Wird dieser Wert als 5 festgelegt, so sieht die „Lückensequenz" folgendermaßen aus: 6, 18, 15.

Auf die Ermittlung der Lücken folgt ein χ^2-Test auf die „beobachteten" Längen der Lücken im Vergleich zu den erwarteten. Für die Bestimmung der unterschiedlichen Klassen k sollten mehrere Lücken mit schwächerer Wahrscheinlichkeit zu einer zusammengefasst werden[112], vorausgesetzt die erwarteten Häufigkeiten folgen keiner Gleichverteilung.

5.6.2.4. POKERTEST[113]

Ein klassischer Pokertest zerlegt eine Sequenz von Zufallszahlen in fünf direkt aufeinanderfolgende ganze Zahlen, die 7 unterschiedliche Muster (abcde, aabcd, aabbc, aaabc, aaabb, aaaab, aaaaa) bilden können und wendet dann den χ^2-Test auf die Anzahl der tatsächlich erhaltenen Muster im Vergleich mit den erwarteten an. Von Donald Knuth wird in diesem Zusammenhang eine simplere und vor allem in Programmiersprachen besser zu implementierende Version dieses Tests vorgeschlagen. Es können anstatt festgelegter fünf, nun nach eventueller Anpassung k aufeinanderfolgende Zahlen zu n Gruppen zusammengefasst werden, wobei die Muster beziehungsweise Kategorien aussehen wie folgt:

5 unterschiedlich	5♣ 2♦ 3♥ 6♥ 8♠	All different
4 unterschiedlich	5♣ 5♦ 3♥ 6♥ 8♠	One pair
3 unterschiedlich	5♣ 5♦ 3♥ 3♠ 8♥ oder 5♣ 5♦ 5♥ 6♥ 8♠	Two pairs oder Three of a kind
2 unterschiedlich	5♣ 5♦ 5♥ 8♣ 8♠ oder 5♣ 5♦ 5♥ 5♠ 8♠	Full House oder Four of a kind
1 unterschiedlich	5♣ 5♦ 5♥ 5♠ 5◊	Five of a kind

Tabelle 10: Ein sogenanntes „Five of a kind" ist im Poker technisch nicht möglich (außer es wird geschummelt) und daher hier mit dem zusätzlichen Symbol ◊ dargestellt.

[111] Vgl. (Knuth, 1969) Seite 60
[112] Vgl. (Köchel & Flohrer, 1995)
[113] Vgl. (Knuth, 1969) Seite 62

Ähnlich dem traditionellen Pokertest wird auch bei Knuths Version die Anzahl an k-Tupel verschiedener Zustände r erfasst und ein χ^2-Test auf r angewendet.

5.6.2.5. COUPONSAMMLERTEST

Das Coupon Collector's-Problem, zu Deutsch auch das Sammelbilderproblem, beschreibt wie viele Sammelbilder einer festgelegten Serie es im Mittel zu kaufen gelte, um die vollständige Serie zu erhalten.[114]

Der Couponsammler-Test läuft analog dazu ab, indem er innerhalb einer Sequenz Y_1, Y_2 … von ganzen Zufallszahlen die Anzahl an Elementen, die es braucht um eine vollständige Serie $Y_{j+1}, Y_{j+2} \ldots Y_{j+r}$, zu bilden (wobei eine Serie aus allen ganzen Zahlen im Intervall [0;d) besteht), ermittelt.[115]

Wird beispielsweise angenommen, dass d den Wert 5 annimmt, so besteht eine vollständige Serie aus den Zahlen 0, 1, 2, 3 und 4. Nach Analyse der Sequenz (2, 4, 1, 2, 0, 0, 3, 4, 0, 1, 2, 0, 2, 4, 2, 0, 2, 2, 4, 0, 1, 0, 2, 2, 1, 2, 1, …) lässt sich feststellen, dass die Serie erst nach dem 7. Wert (3) vollendet ist, folglich beträgt deren Länge 7. Ein χ^2-Test auf die Länge der einzelnen vollendeten Serien liefert Information darüber, ob die zu testende Sequenz ähnliche Eigenschaften wie eine tatsächlich zufällige aufweist.[116]

5.6.2.6. PERMUTATIONSTEST

Dieser Test weißt starke Parallelen zu Marsaglias OPERM5 aus der Diehard Testbatterie auf. Einer der größten Unterschiede besteht dabei darin, dass die jeweiligen Permutationen bei Knuth einander nicht überlappen sondern aneinandergereiht werden. Der Test ist nur unter der Annahme fundiert, dass innerhalb eines t-Tupels niemals dieselben Elemente vorkommen.[117] Dadurch wird bei Marsaglia ein neuer Zustand durch den Wert an der Stelle t, t+1, t+2, … t+(n-t) abgeschlossen, bei Knuth durch jene an der Stelle t, 2t, … $\frac{n}{t}$ t.

5 2 7 1 ***6 1 5 4*** ***7***	52716 15479 86534 52617 ***93476***
Überlappende Permutationen	aufeinanderfolgende Permutationen
OPERM5 – George Marsaglia (t ist immer 5)	**Permutationstest – Donald Knuth (t ist hier 5, zum Vergleich mit OPERM5)**

Tabelle 11: fünf 5-Tupel in OPERM5 und dem Permutationstest

Für den anschließenden χ^2-Test wird t!, die Anzahl an Anordnungsmöglichkeiten, mit k und die Wahrscheinlichkeit, dass ein beliebiger Wert in einer beliebigen Kategorie liegt mit $\frac{1}{t!}$ gleichgesetzt.[118]

[114] Vgl. (Sammelbilderproblem: Wikipedia.org, 2013)
[115] Vgl. (Knuth, 1969)Seite 63
[116] Vgl. (Knuth, 1969) Seite 62
[117] Vgl. (Knuth, 1969)Seite 64
[118] Vgl. (Knuth, 1969)Seite 64

5.6.2.7. LAUFTEST

Der Lauf-Test ist im Grunde äquivalent zu Marsaglias Test, beschrieben *in 5.6.1.12.*

8	3	1	6	3	9	$=t_1$	$\max(t_1) = 9$
5	8	7	6	3	1	$=t_2$	$\max(t_2) = 8$
9	7	9	4	4	5	$=t_3$	$\max(t_3) = 9$
4	4	4	2	1	7	$=t_4$	$\max(t_4) = 7$
2	1	9	2	0	0	$=t_5$	$\max(t_5) = 9$

Sequenz 6

5.6.2.8. MAXIMUM-AUS-T-TEST

Der Maximum-aus-t-test sieht vor, für einen bestimmten Bereich t einer Zufallszahlensequenz den Maximalwert $\max(U_{tj}, U_{tj+1} \dots U_{tj+(t-1)})$ zu berechnen. Nach der Ermittlung der größten Werte der jeweiligen Bereiche $t_1; t_2 \dots$ wird der Kolmogorow-Smirnow-Test auf die Folge der jeweiligen Maximalwerte $V_1, V_2 \dots V_{n-1}$ angewendet.[119]

Wird beispielsweise die Sequenz 6 in 5 Bereiche t eingeteilt, so lässt sich die Sequenz der Maximalwerte (9, 8, 9, 7, 9) determinieren.

5.6.2.9. KOLLISIONSTEST

Der Kollisions-Test ist im Grunde äquivalent zu Marsaglias Parkplatz-Test, beschrieben in 5.6.1.7.

5.7. ÜBER DIE INTERPRETATION VON P-WERTEN

Im nächsten Kapitel sollen die Tests der Diehard-Batterie auf die bekanntesten Zufallszahlengeneratoren angewandt werden. Dazu gehört zunächst allerdings ein tieferes Verständnis der p-Werte, da die meisten Tests einen solchen als Resultat ausgeben.

Anders als in der allgemeinen Statistik bedeuten p-Werte nahe 0 und 1 (wie 0,0012 oder 0,9983) in der Bewertung der Resultate dieser Tests nicht unbedingt, dass der zugehörige RNG den Test nicht bestanden hat.[120] Es müssen daher mehrere p-Werte betrachtet und in den größeren Kontext gestellt werden. Wenn ein Test aber ausschließlich die Werte 1,000000 oder 0,000000 ausgibt, so kann der dazugehörende Zufallszahlengenerator durchaus als defekt bezeichnet werden.

[119] Vgl. (Knuth, 1969)Seite 68

[120] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

6. ZUFALLSZAHLENGENERATOREN

6.1. DETERMINISTISCHE ZUFALLSZAHLENGENERATOREN

Deterministische, numerische oder auch Pseudozufallszahlengeneratoren (PRNGs) erzeugen Zufallszahlen mittels eines Algorithmus, dessen Existenz per Definition Unvorhersagbarkeit und daher Zufälligkeit ausschließt. Im Allgemeinen wird für die Umsetzung derartiger Generatoren ein „seed", also ein Anfangswert verwendet, um eine Sequenz von Zufallszahlen zu erzeugen.[121]

Generelle Vorteile[122]	Generelle Nachteile
Einfache und schnelle Erzeugung, genügt den praktischen Anforderungen	Nicht tatsächlich zufällig
wiederhohlbar (Wichtig für die Reproduktion von Simulationen und Versuchen)	Auch Unbefugte können unter Wissen der genauen Parameter die Sequenz nachvollziehen
	Bildet häufig Muster und unerwünschte Regelmäßigkeiten
	Bei schlechten Parametern ungenügend lange Sequenzen[123]

Tabelle 12: Generelle Vorteile und Nachteile deterministischer RNGs

Daher sind aufeinanderfolgende Elemente der Zufallssequenz voneinander abhängig und nicht zufällig, können allerdings bei guten Parametern so erscheinen. Das Ziel der meisten Algorithmen zur Erzeugung von Zufallszahlen ist also eine scheinbare, also theoretische sowie empirische Tests bestehende Zufälligkeit. Im Folgenden werden einige PRNGs vorgestellt, die diesen Anforderungen auf unterschiedlich zufriedenstellende Weise gerecht werden.

Die meisten dieser Generatoren bedienen sich einer sogenannten Modulooperation in der Form $15\ mod\ 9 = 6$. Der Modulooperator mod gibt den Rest einer Division an. So ist beispielsweise 6 der Rest, der bei der Division 15 durch 9 entsteht.

[121] Vgl. (Mordasini & Klahr, 2013)
[122] Vgl. (Mordasini & Klahr, 2013)
[123] Vgl. (Mordasini & Klahr, 2013)

6.1.1. DEZIMALENTWICKLUNGEN IRRATIONALER ZAHLEN

14159265358979323846264338327950288419716939937510
58209749445923078164062862089986280348253421170679
82148086513282306647093844609550582231725359408128
48111745028410270193852110555964462294895493038196

Sequenz 7: Die 200 ersten Nachkommastellen von π

6.1.1.1. METHODE

Die dieser Gruppe angehörenden RNGs basieren auf einer einzelnen Zahl, wie π oder $\sqrt{2}$,als Zufallszahlensequenz.[124] Dazu werden die Nachkommastellen $s_1, s_2, s_3, s_4, \ldots$ dieser irrationalen Zahlen $s, s_1 s_2 s_3 s_4 \ldots$ als Zufallsziffern zwischen 0 und 9 betrachtet.

Sequenz 7 mag zwar zunächst zufällig erscheinen, dennoch ist sie von einer strengen inneren Ordnung bestimmt, jedes weitere Element kann theoretisch vorhergesagt werden[125], vorausgesetzt die Formel zur Errechnung dieser Zahl, oder andere Annäherungsmethoden sind bekannt und besonders leistungsfähige Rechner kommen zum Einsatz. Diese Eigenschaft trifft allerdings auch auf jeden anderen deterministischen RNG zu und darf daher per se nicht als Grund, diese Methode zu verwerfen, gesehen werden. Deshalb gilt es noch andere Kriterien im Betracht irrationaler Zufallszahlengeneratoren zu überprüfen.

Um eine irrationale Zahl als verlässlichen RNG gebrauchen zu können, muss nachgewissen werden, dass es sich bei der vorliegenden um eine sogenannte „normale" Zahl handelt.

Eine reelle Zahl $s, s_1 s_2 s_3 s_4 \ldots$ heißt normal, wenn alle Nachkommastellen $s_1, s_2, s_3, s_4, \ldots$ mit dergleichen Häufigkeit auftreten[126]. Im Falle von Pi bedeutet dies eine Wahrscheinlichkeit von $\frac{1}{10}$ pro möglicher Ziffer. Weiters umschließt die Definition all jene reelle Zahlen, deren mögliche Ziffernblöcke alle mit jeweils derselben Wahrscheinlichkeit vorkommen.[127] Erneut auf die Kreiszahl bezogen heißt dies, dass jedes Ziffernpaar s_n, s_{n+1} , beziehungsweise $s_{n+1}, s_{(n+1)+1}$ mit einer Wahrscheinlichkeit oder Häufigkeit von $\frac{1}{10^2}$ beziehungsweise $\frac{1}{10^3}$ auftritt.

Von der Kreiszahl π darf angenommen werden, dass sie normal ist, ein endgültiger Beweis dafür kann für diese Behauptung trotz ihrem Bestehen aller bekannten Tests auf Normalität nicht geliefert werden.[128] Derzeit sind 12 Trillionen Nachkommastellen der Zahl bekannt (Die Errechnung all dieser Stellen dauerte circa ein Jahr)[129], dennoch kann nicht ohne weiteres ausgeschlossen werden, dass sich nach der 13 Trillionsten Nachkommastelle plötzlich nur noch die Zahl 6 oder das Zahlenpaar 25 wiederholt.

[124] Vgl. (Softwaretechnische Realisierungen: wikipedia.org, 2013)
[125] Vgl. (Milavec, Zufall, Ordnung und Chaos am Computer, 1995) Seite 2
[126] Vgl. (Aistleitner, 2006) Seite 5
[127] Vgl. (Aistleitner, 2006) Seite 5
[128] Vgl. (Milavec, Zufall, Ordnung und Chaos am Computer, 1995) Seite 2
[129] Vgl. (Yee & Kondo, 2013)

6.1.1.2. Gütekriterien

6.1.1.2.1. Praktische Anforderungen

Die approximierende Berechnung der Nachkommastellen transzendenter, irrationaler Zahlen ist äußerst komplex, demzufolge wird in der Praxis meist auf einfachere Methoden zurückgegriffen.

6.1.1.2.2. Periodenlänge

Durch ihre aperiodische Natur lassen sich mit dieser Art von PRNGs, vorausgesetzt die Zufallszahlen leiten sich tatsächlich von normalen Zahlen ab, unendlich lange Sequenzen von Zufallszahlen bilden.

6.1.1.2.3. Verteilung und Unabhängigkeit

Sollte es sich bei den irrationalen Zahlen tatsächlich auch um normale handeln, so sind die Zahlen gleichverteilt, mit einer jeweiligen Häufigkeit von $\frac{1}{10}$ und erscheinen statistisch unabhängig voneinander.

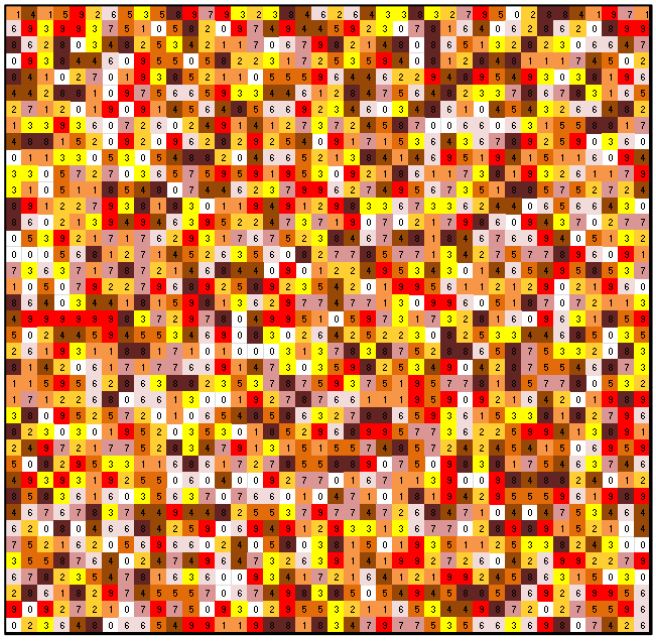

Abbildung 14: Visuelle Darstellung der Nachkommastellen von Pi. Es sind keine bestimmten Muster, Abhängigkeiten oder gleichmäßige Verteilung erkennbar

6.1.2. Mittelquadratmethode

735352, 742563, 399808, 846436, 453902

Sequenz 8: 5 durch die Mittelquadratmethode generierte Zufallszahlen

6.1.2.1. Methode

John von Neumann (1903-1957) schlug 1946 die sogenannte Mittelquadratmethode (oder „Mitte der Quadratzahl“) als Möglichkeit zur Zufallszahlengenerierung vor. Dabei wird eine n-Ziffern lange Zahl (zunächst der „seed“) quadriert und die mittleren n-Ziffern der dadurch entstehenden Zahl (also die

erste Zufallszahl) ermittelt. Die daraus resultierende n-stellige Zahl dient anschließend wiederum als „seed“ und der Vorgang wird für die gewünschte Länge der Sequenz fortgesetzt.[130]

Mit seed 735352 ließe sich entsprechend eine Sequenz wie in Tabelle 13 bilden:

Seed (6-stellig)	**735352**
$seed^2$	540**742563**904
	↓
1. Zufallszahl	742563
$(1.\ Zufallszahl)^2$	551**399808**969
	↓
2. Zufallszahl	399808
$(2.\ Zufallszahl)^2$	159**846436**864
	↓
3. Zufallszahl	846436
$(3.\ Zufallszahl)^2$	716**453902**096
	↓
4. Zufallszahl	453902
$(4.\ Zufallszahl)^2$	206**027025**604
	↓
5. Zufallszahl	027025

Tabelle 13: Die Mittelquadratmethode

6.1.2.2. GÜTEKRITERIEN

6.1.2.2.1. PRAKTISCHE ANFORDERUNGEN

Wie ein Großteil der deterministischen RNGs ist auch die Erzeugung von Zufallszahlen mittels der Mittelquadratmethode einfach und schnell.

6.1.2.2.2. PERIODENLÄNGE

Die kurze Periodenlänge ist wohl der Hauptgrund, aus dem diese Technik heute lediglich von historischer Bedeutung ist.[131] Nach Analysen des von Neumann geschaffenen RNGs lässt sich errechnen, dass die Sequenz mit einem seed von 0 bis 9999 nach durchschnittlich 45,8 Iterationen in einem regelmäßigen Zyklus endet.[132] In anderen Worten ist die durchschnittliche Sequenzlänge 45,8; viel zu kurz für einen Großteil der Anwendungsgebiete für Zufallszahlengeneratoren. Die Sequenz in TABELLE 13 endet schon nach der 4. Zufallszahl, da sie dann auf eine (n-1)-stellige Zahl „abstürzt“.

In einzelnen Fällen lässt sich nicht eine einzige Zufallszahl generieren, da das Mittelquadrat eines bestimmten seeds (wie 3792) denselben Wert zurückgibt (3792^2=14**3792**64).[133]

6.1.2.2.3. EMPIRISCHE TESTS

Die Sequenzen sind generell bei Weitem zu kurz, um die empirischen Tests von beispielsweise Marsaglia oder Knuth darauf anwenden zu können.

[130] Vgl. (Milavec, Zufall, Ordnung und Chaos am Computer, 1995) Seite 5
[131] Vgl. (Mordasini & Klahr, 2013)
[132] Vgl. (Milavec, Zufall, Ordnung und Chaos am Computer, 1995) Seite 5
[133] Vgl. (Milavec, Zufall, Ordnung und Chaos am Computer, 1995) Seite 5

6.1.3. DER ALLGEMEINE LINEARE KONGRUENZGENERATOR (LCG)

148	45	38	154	222	35	259	30	176
355	153	128	100	177	62	88	210	68
265	207	173	73	348	269	196	300	14
220	234	2	253	240	179	250	345	374
4	54	110	343	186	134	277	174	167
283	351	164	1	159	305	97	282	257
229	306	191	217	339	197			

Sequenz 9: 60 durch einen linearen Kongruenzgenerator entstandene Zufallszahlen m=387; a=868;c=65

Lineare Kongruenzgeneratoren zählen dank ihrer Einfachheit zu den wichtigsten deterministischen RNGs.

Sie liefern schnell Zufallszahlen und verbrauchen eine minimale Größe im Cache eines Taschenrechners oder PCs. Ihre Güte hängt im Wesentlichen von den jeweiligen verwendeten Werten für die zugrundeliegenden Parameter in Formel 7 ab.[134] Die Methode der linearen Kongruenzgeneratoren stammt vom amerikanischen Mathematiker Derrick Henry Lehmer (1905-1991).[135]

6.1.3.1. METHODE

Der Algorithmus hinter dem linearen Kongruenzgenerator liefert positive ganze Zahlen[136] und lautet:

$$I_{j+1} = aI_j + c \ (mod\ m)$$

Formel 7: Formel für den linearen Kongruenzgenerator

Wobei m den Modulus (m>0), a den Multiplier (0<a<m) und c den Inkrement ($0 \leq c < m$) bezeichnet.[137] I_j beschreibt den Startwert oder „seed“ der Sequenz, wohingegen der auf den Startwert folgende Wert I_{j+1} genannt wird.

6.1.3.2. GÜTEKRITERIEN

6.1.3.2.1. PRAKTISCHE KRITERIEN

Die weite Verbreitung dieser Art von RNGs begründet sich in der Einfachheit und Geschwindigkeit ihrer Implementation. Genügen auch weniger zuverlässige RNGs den jeweiligen Anforderungen, so können auch sehr schlichte Parameter verwendet werden.

Wie bereits zuvor angemerkt, verbraucht ein RNG grundsätzlich eine minimale Größe im Cache[138] und wird daher häufig auf Taschenrechnern und PCs implementiert.

[134] Vgl. (Mordasini & Klahr, 2013)
[135] Vgl. (Lehmer, 1949) Seiten 141-146
[136] Vgl. (Dagpunar, 1988) Seite 19
[137] Vgl. (Mordasini & Klahr, 2013)
[138] Vgl. (Mordasini & Klahr, 2013)

6.1.3.2.2. PERIODENLÄNGE

Die Periodenlänge eines RNGs hängt stark von den verwendeten Werten für die Parameter ab. Die von Informatiker Donald E. Knuth formulierten Anforderungen an die Parameter Modulus (m) und Multiplier (a) ermöglichen eine maximale Länge der Zufallszahlensequenz. Für Knuth ist hier die Wahl des Modulus besonders bedeutend, da dieser den Wertebereich, den die Elemente der Sequenz durchlaufen können, darstellt. Es sollte deshalb eine große Zahl für diese Variable gewählt werden[139], aber auch folgende Punkte, die sich auf die Parameter m, a und c beziehen, dürfen für das Ziel einer maximalen Periodenlänge nicht außer Acht gelassen werden:

1. Das Inkrement c soll zum Modulus m teilerfremd sein
2. a-1 ist ein Vielfaches von jedem Primfaktor, der m teilt
3. a-1 ist ein Vielfaches von 4, wenn m ein Vielfaches von 4 ist[140]

Gute Parameter liefert das Werk *„The Art of Scientific Computing“* aus der Buchserie *„Numerical Recipes“* der Autoren William H. Press und Saul A. Teukolsky. Sie schlagen für die Parameter diese Werte vor[141]:

Multiplier a	Modulo m	Inkrement c
1664525	2^{32}	1013904223

Tabelle 14

6.1.3.2.3. VERTEILUNG UND UNABHÄNGIGKEIT

Lineare Kongruenzgeneratoren weisen eine hohe sequentielle Korrelation auf, die sich in sogenannten Hyperebenen (siehe *5.5.1.1 Hyperebenen-Verhalten*), manifestiert.[142] Das heißt, dass die Zufallszahlen, besonders bei einer leichtfertigen Wahl der Parameter, in hohem Maße voneinander abhängig erscheinen. Wird dieser Test auf eine 15.000 Zahlen lange, durch einen LCG erzeugte Sequenz (mit den vorgeschlagenen Parametern von Press und Teukolsky) angewandt, so

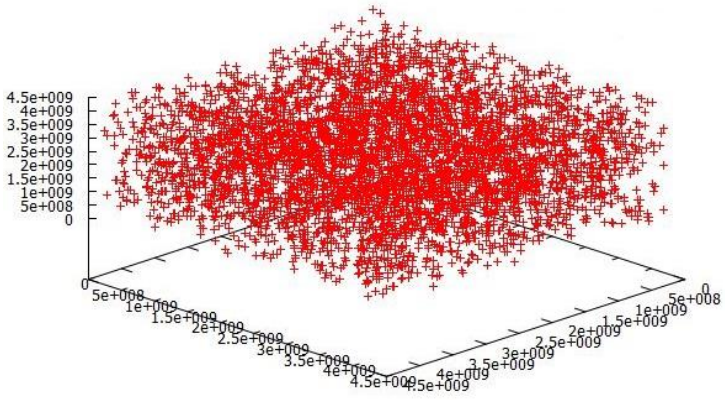

Abbildung 15: In den Raum geplottete 3-Tupel, erzeugt durch einen linearen Kongruenzgenerator. n=15.000 a=1.664.525 m=2^{32} c=1.013.904.223

[139] Vgl. (Knuth, 1969) Seite 11
[140] Vgl. (Knuth, 1969) Seite 16
[141] Vgl. (Press & Teukolsky, 1992), zitiert nach (Linear congruentia generator - Parameters in common use: wikipedia.org, 2013)
[142] Vgl. (Mordasini & Klahr, 2013)

mag die entstehende Grafik der in Abbildung 15 gleichen.

Die Variablen sind offenbar so gut gewählt, dass Marsaglias Methode keine sichtbaren Hyperebenen bildet. Es gibt allerdings auch andere Parameter, die deutliche Hyperebenen bilden, wie hier an späterer Stelle noch gezeigt werden soll.

Ist die Periodenlänge nicht so lang wie der Wertebereich, was bei schlechten Parametern durchaus zu bedenken ist, so ergibt sich keine Häufigkeit von $\frac{1}{m}$ für jeden möglichen Wert.

6.1.3.2.4. Empirische Tests

Der lineare Kongruenzgenerator, wenngleich zweckdienlich für eine Vielzahl an Anwendungen, passiert empirische Tests, wie jene der Diehard-Sammlung oft nur auf unbefriedigende Weise. Eine durch einen LCG erzeugte und von Diehard getestete Sequenz mag Ergebnisse wie in Tabelle 15 liefern.

Test	p-Werte des ersten Tests (meist Chi-Quadrat-Test, selten auch Kolmogoroff-Smirnow-Test)					p-Werte des zweiten Test (Kolmogoroff-Smirnow-Test)
Geburtstagsabstände	0,815381	0,584275	0,450692			1,000000
	0,251101	0,639621	0,249072			
	0,964143	1,000000	1,000000			
Affentest	20 Mal 1,00000					kein KS-Test ist hier vorgesehen
OPERM	0,989455		0,991610			kein KS-Test ist hier vorgesehen
Ränge von Matrizen (31x31)	1,000000					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (32x32)	1,000000					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (6x8)	0,559911	0,407869	0,191633	0,077537	0,260300	0, 999405
	0,224860	0,460637	0,491865	0,010131	0,224307	
	0,598960	0,701048	0,822454	0,217878	0,773705	
	0,909089	0,738260	0,598277	0,659908	0,335548	
	0,382177	0,189177	0,949383	1,00000	1,000000	
Parkplatztest	0,807188	0,218799	0,218799	0,590298	0,554479	0,227256
	0,625377	0,899470	0,692266	0,842447	0,100530	
OPSO	20 Mal 1,00000					kein KS-Test ist hier vorgesehen
OQSO	20 Mal 1,00000					kein KS-Test ist hier vorgesehen
DNA	20 Mal 1,00000					kein KS-Test ist hier vorgesehen
Zähle die 1en (1)	1,00000		1,00000			kein KS-Test ist hier vorgesehen
Zähle die 1en (2)	0,750958	0,222288	0,390153	0,045203	0,432683	kein KS-Test ist hier vorgesehen
	0,131185	0,999006	0,049539	0,731208	0,153910	
	0,930091	0,390874	0,131390	0,753024	0,943672	
	0,502724	0,293915	0,742763	0,950126	0,158213	
	0,704941	0,044028	1,000000	1,000000	1,000000	
Minimumdistanztest	-					0,999803
Zufällige-Kugeln-Test	0,93969	0,36933	0,77607	0,24554	0,62774	0,686994
	0,87634	0,90171	0,91568	0,55636	0,98973	
	0,74064	0,70925	0,91968	0,35546	0,25534	
	0,11686	0,29578	0,07195	0,63635	0,25551	
Squeeze-test	1,000000					kein KS-Test ist hier vorgesehen
Überlappende-Summen-Test	0,959197	0,715695	0,576400	0,149183	0,026835	0,111437
	0,532157	0,065138	0,808407	0,255416	0,745975	
Läufe-test	-					„Raufläufe“ „Runterläufe“
						0,102111 0,885028
						0,822715 0,275828
Craps-test	Anzahl an Gewinnen		Anzahl an Würfe/Spiel			kein KS-Test ist hier vorgesehen
	0,740433		0,998219			

Tabelle 15: Exemplarisches Verhalten eines LCG (a=1664525 m=2^{32} c=1013904223) in der Diehard-Testsammlung. Kritische Werte (sehr nahe an 0 oder 1) sind hervorgehoben

6.1.4. DER MULTIPLIKATIVE LINEARE KONGRUENZGENERATOR

83	62	23	227	53	338	38	89	239	20	332	248
92	134	212	191	152	356	182	80	167	218	368	149
74	377	221	263	341	320	281	98	311	209	296	347
110	278	203	119	350	5	83	62	23	227	53	338
38	89	239	20	332	248	92	134	212	191	152	356

Sequenz 10: 60 durch einen multiplikativen linearen Kongruenzgenerator entstandene Zufallszahlen. m= 387; a=868

6.1.4.1. METHODE

Der multiplikative lineare Kongruenzgenerator, nach seinen Entwicklern auch Park-Miller PRNG, ist ein Spezialfall des allgemeinen, für den $c = 0$ gilt.[143] Daher lautet die geänderte Formel:

$$I_{j+1} = aI_j \ (mod\ m)$$

Formel 8: Formel für den multiplikativen linearen Kongruenzgenerator

6.1.4.2. GÜTEKRITERIEN

Allgemein müssen die Werte für die Parameter des multiplikativen linearen Kongruenzgenerators sehr vorsichtig gewählt werden.[144]

6.1.4.2.1. PRAKTISCHE KRITERIEN

Diese Methode heißt „Minimal Standard“ (MINSTD) Generator,[145] da sie simpel und leicht zu implementieren ist. Zudem ist der Generator tragbar, da jeder wissenschaftliche Taschenrechner die Möglichkeit hat, den entsprechenden Algorithmus anzuwenden.

6.1.4.2.2. PERIODENLÄNGE

Neben anderen Vorgaben für Kongruenzgeneratoren soll das Inkrement c zum Modulus m teilerfremd sein, um die längste Periode zu ermöglichen.[146] Dieses Kriterium wird nicht erfüllt wenn $c = 0$, ergo lassen sich mit dem Park-Miller PRNG niemals m-lange Sequenzen erzeugen.[147]

6.1.4.2.3. VERTEILUNG UND UNABHÄNGIGKEIT

Der multiplikative lineare Kongruenzgenerator weist in vielerlei Hinsicht im Vergleich zum allgemeinen linearen erhebliche Mängel auf. Entsprechend müssen auch die durch ihn entstehenden Verteilungen und sukzessive Paare sorgfältiger analysiert werden. Auch die Parameter sollten vorsichtiger gewählt werden, um Hyperebenen und ungleichmäßige Verteilungen vorzubeugen.

[143] Vgl. (Park & Miller, 1988), zitiert nach (Mordasini & Klahr, 2013)
[144] Vgl. (Mordasini & Klahr, 2013)
[145] Vgl. (Mordasini & Klahr, 2013)
[146] Vgl. (Knuth, 1969) Seite 16
[147] Vgl. (Dagpunar, 1988) Seite 20

6.1.4.2.4. EMPIRISCHE TESTS

Dieser Kongruenzgenerator verhält sich in empirischen Tests ähnlich wie der allgemeine lineare, verdeutlicht durch das Beispiel in Tabelle 16.

Test	p-Werte des ersten Tests (meist Chi-Quadrat-Test, selten auch Kolmogoroff-Smirnow-Test)					p-Werte des zweiten Test (Kolmogoroff-Smirnow-Test)
Geburtstagsabstände	0,387836	0,869970	0,927293			0,999991
	0,724988	0,925504	0,941279			
	0,743748	0,786273	1,000000			
Affentest	20 Mal 1,000000					kein KS-Test ist hier vorgesehen
OPERM	0,370030		0,271489			kein KS-Test ist hier vorgesehen
Ränge von Matrizen (31x31)	0,991220					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (32x32)	1,000000					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (6x8)	0,398510	0,750517	0,559377	0,849908	0,484966	0,904848
	0,209402	0,317190	0,273200	0,910864	0,889815	
	0,313637	0,318794	0,481155	0,484266	0,747420	
	0,834539	0,807779	0,056888	0,267494	0,897087	
	0,141099	0,676595	0,381926	0,167115	1,000000	
Parkplatztest	0,427537	0,625377	0,276387	0,676028	0,092718	0,581005
	0,590298	0,692266	0,276387	0,218799	0,409702	
OPSO	1,0000	0,8951	0,8369	0,8402	0,5147	kein KS-Test ist hier vorgesehen
	0,0324	0,0169	0,0250	0,0043	0,8044	
	0,3442	0,1686	0,9458	0,4203	0,0416	
	0,5339	0,8887	0,9831	0,8291	0,9012	
	0,4379	0,3777	0,1102			
OQSO	1,0000	0,3244	0,8427	0,4124	0,4874	kein KS-Test ist hier vorgesehen
	0,8873	0,2125	0,3467	0,5562	0,4658	
	0,4456	0,4914	0,1278	0,2047	0,4617	
	0,0250	0,6917	0,7799	0,5293	0,2716	
	0,9363	0,0105	0,0297	0,1162	0,6374	
	0,0728	0,2807	0,8585			
DNA	1,0000	0,6598	0,2665	0,1212	0,3825	kein KS-Test ist hier vorgesehen
	0,1354	0,1800	0,6246	0,8076	0,5196	
	0,4926	0,4006	0,9402	0,7260	0,3984	
	0,7945	0,2466	0,7936	0,5791	0,5173	
	0,1731	0,0915	0,8180	0,1777	0,5837	
	0,8711	0,0380	0,5278	0,9555	0,2627	
	0,1348					
Zähle die 1en (1)	1,000000			1,000000		kein KS-Test ist hier vorgesehen
Zähle die 1en (2)	0,037343	0,050777	0,065245	0,134225	0,445672	kein KS-Test ist hier vorgesehen
	0,916983	0,204326	0,967082	0,895104	0,660246	
	0,671086	0,815013	0,125077	0,981846	0,453979	
	0,570722	0,380152	0,110568	0,374552	0,425104	
	0,509384	0,081113	0,950882	0,239285	1,000000	
Minimumdistanztest	-					,449972
Zufällige-Kugeln-Test	0,65297	0,49857	0,44732	0,99840	0,88264	0,944076
	0,55278	0,36241	0,69915	0,98853	0,48457	
	0,38750	0,49234	0,98150	0,15523	0,93187	
	0,68215	0,46182	0,80182	0,25351	0,64459	
Squeeze-test	0,647875					kein KS-Test ist hier vorgesehen
Überlappende-Summen-Test	0,822387	0,901573	0,634354	0,735625	0,062202	,067264
	0,269622	0,296079	0,879644	0,226955	0,355645	
Läufe-test	-					„Raufläufe“ „Runterläufe“
						0,051944 0,845152
						0,882697 0,662618
Craps-test	Anzahl an Gewinnen			Anzahl an Würfe/Spiel		kein KS-Test ist hier vorgesehen
	0,646663			0,801175		

Tabelle 16: Exemplarisches Verhalten eines MLCG (a= 48271 m= 2^{31-} 1) in der Diehard-Testsammlung. Kritische Werte (sehr nahe an 0 oder 1) sind hervorgehoben

6.1.4.3. RANDU

> *„Line 22 [≡ RANDU, Anm. der Verfasserin] is, regrettably, the generator that has actually been used on such machines in most of the world's scientific computing centers for about a decade; Its very name RANDU is enough to bring dismay into the eyes and stomachs of many computer scientists!"*[148]
>
> *„Line 22 [≡ RANDU, Anm. der Verfasserin] ist bedauerlicherweise der Generator der für circa eine Dekade in den meisten Zentren für Computerwissenschaften eingesetzt wurde; Allein der Name RANDU lässt Schrecken in die Glieder vieler Computerwissenschaftler fahren!"*

Wie Donald Knuth feststellt, ist der multiplikative lineare Kongruenzgenerator RANDU der wohl berüchtigtste Zufallszahlengenerator in der Geschichte der Stochastik.

Dieser PRNG wurde in den 60er Jahren auf IBM Mainframes installiert und erfuhr in den darauffolgenden Jahrzehnten weite Verbreitung. Als sich schließlich herausstellte, dass der ihm zugrundeliegende Algorithmus große Mängel aufweist, waren die Ergebnisse zahlreicher Forschungen, Studien und Experimente obsolet geworden.[149] Die ihnen zu Grunde liegende Annahme, die statistische Zufälligkeit des PRNGs, war schlichtweg falsch.

Die Formel für diesen Generator lautet:

$$I_{j+1} = 65539 I_j \ (mod\ 2^{31})$$

Formel 9: Formel für RANDU

Die Problematik bezieht sich im Wesentlichen auf die mangelnde Unabhängigkeit zwischen Elementen der Sequenz.[150] Dieser Defekt lässt sich durch das Plotten von 3-Tupel in den Raum darstellen:

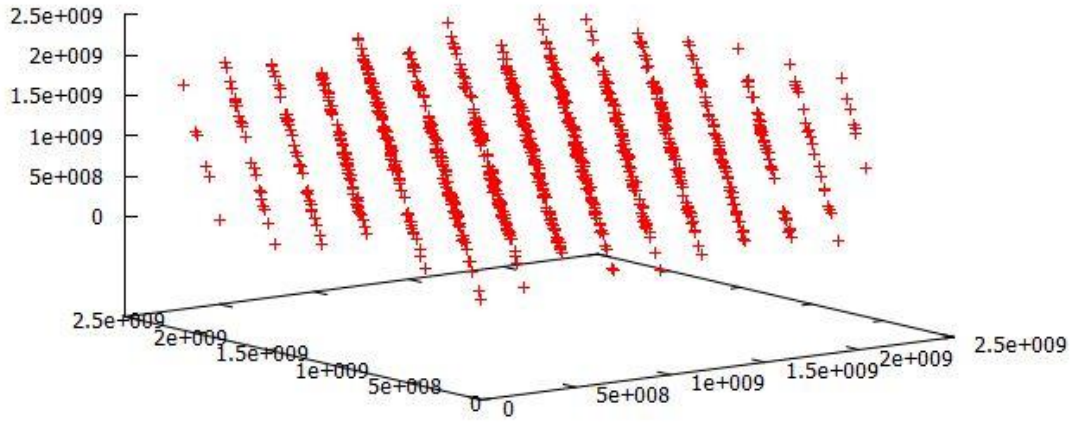

Abbildung 16: Marsaglias Methode, angewandt auf den multiplikativen linearen Kongruenzgenerator RANDU

Es ist deutlich sichtbar, dass die Tupel auf einer geringen Anzahl von Hyperebenen liegen und daher von einer starken inneren Ordnung definiert werden.

---

[148] (Knuth, 1969) Seite 104
[149] Vgl. (Mordasini & Klahr, 2013)
[150] Vgl. (Knuth, 1969) Seite 104

RANDU ist nur noch von historischer Bedeutung, von seinem Gebrauch kann jedenfalls unabhängig vom vorgesehenen Anwendungsgebiet der Zufallszahlen abgeraten werden.

6.1.4.4. VERWENDUNG

Obwohl ein multiplikativer Kongruenzgenerator Defekte aufweist, wird er häufig in Softwareprogrammen implementiert, besonders dann, wenn die Ansprüche nicht wissenschaftlicher oder sonstiger professioneller Natur sind.

Der Zufallszahlengenerator, der beispielsweise für die Taschenrechner der Marke „Texas Instruments“ verwendet wird, ist ein kombinierter multiplikativer linearer Kongruenzgenerator mit einer Periodenlänge von $2{,}30584 \times 10^{18}$.[151]

Das Tabellenkalkulationsprogramm Excel der Microsoft Corporation verwendet einen von B.A. Wichman und I.D. Hill spezifizierten multiplikativen Kongruenzgenerator und lässt ihn drei Zufallszahlen erzeugen, bevor er aus den dreien eine weitere Zufallszahl generiert. Dieser Generator erzeugt sehr zuverlässige Sequenzen, die die Diehard-Tests bestehen und eine Periodenlänge von etwa 10^{13} haben.[152]

6.1.5. DER GEMISCHT-LINEARE KONGRUENZGENERATOR

Der gemischt-lineare Kongruenzgenerator entspricht im Wesentlich dem allgemeinen, wenn $c \neq 0$[153] (anders als beim multiplikativen, für den $c = 0$ gilt).

Für die entsprechenden Informationen zu diesem Generator siehe *6.1.3 Der allgemeine lineare Kongruenzgenerator.*

6.1.6. DER ADDITIVE KONGRUENZGENERATOR

108, 119, 227, 346, 573, 919, 1492, 2411, 3903, 6314, 832, 7146, 7978, 5739, 4332, 686, 5018, 5704, 1337, 7041, 8378, 6034, 5027, 1676, 6703, 8379, 5697, 4691, 1003, 5694, 6697, 3006, 318, 3324, 3642, 6966, 1223, 8189, 27, 8216, 8243, 7074, 5932, 3621, 168, 3789, 3957, 7746, 2318, 679,

Sequenz 11: 50 durch einen additiven RNG generierte Zufallszahlen (in diesem Fall ein Fibonacci-Generator mit m=9385)

6.1.6.1. METHODE

Lineare Kongruenzgeneratoren ermöglichen lediglich eine Sequenzlänge von m, bevor sich die Folge wiederholt. Wenn I_{j+1} hingegen von mehreren vorausgehenden Elementen der Sequenz, wie etwa I_j und I_{j-1}, anstatt allein von I_j, abhängig gemacht wird, so kann die Periodenlänge erheblich erhöht werden.[154] Für den additiven Kongruenzgenerator gilt eine solche maximale Periode von $m^p - 1$,[155]

[151] Vgl. (Mangaldan, 2014)
[152] Vgl. (Beschreibung der Funktion ZUFALLSZAHL in Excel 2007 und Excel 2003: support.microsoft.com, 2008)
[153] Vgl. (Burns, 2004) Seite 4
[154] Vgl. (Milavec, Zufall, Ordnung und Chaos am Computer, 1995) Seite 7

eine Eigenschaft die einen derartigen RNG sehr günstig erscheinen lässt. Der darüber definierte allgemeine Kongruenzgenerator oder additive RNG heißt[156]:

$$I_j = (a_1 I_{j-1} + \cdots + a_p I_{j-p})(mod\ m)$$
$$p > 1, ap \neq 0$$

Formel 10: Formel für den additiven RNG

Im Nachfolgenden werden Abwandlungen oder Ausformungen dieses Generators vorgestellt.

6.1.6.2. DER FIBONACCI-GENERATOR

Der einfachste dieser Art von Zufallszahlengeneratoren ist der den 50er Jahren entstammende Fibonacci-Generator.[157] Er gilt für den Sonderfall, dass p=2 und a_1=a_2=1[158], der zugrundeliegende Algorithmus lautet demnach:

$$I_j = (I_{j-1} + I_{j-2})(mod\ m)$$

Formel 11: Formel für den Fibonacci Kongruenzgenerator

Der Generator heißt Fibonacci-Generator, da der Algorithmus dem hinter der Fibonacci-Folge (1,1,2,3,5,8,13,21,34,55,…n=(n-1)+(n-2)) sehr ähnlich ist.

Die maximale Sequenzlänge beträgt für den Fibonacci-Generator bis zu m^2-1. Allerdings weist dieser Zufallszahlengenerator Defekte auf, die sich mittels empirischer Tests, wie etwa dem „Gap-Test“ (siehe *5.6.2.3 Lückentest*), belegen lassen.[159] Die Methode ist zwar sehr simpel, aber scheint nicht wirklich zufällig und wurde daher von Donald Knuth als „schönes ‚schlechtes Beispiel‘ “[160] eines additiven RNGs bezeichnet.

6.1.6.3. DER VERZÖGERTE (LAGGED) FIBONACCI-GENERATOR

4647, 1231, 945, 5051, 1585, 1017, 5455, 1939, 1089, 5859, 2293, 1161, 6263, 2647, 1233, 6667, 3001, 1305, 7071, 3613, 956, 4855, 4310, 1039, 5320, 5007, 1122, 5785, 5704, 1205, 6250, 6401, 1288, 6715, 7098, 1371, 7180, 7795, 1454, 7645, 8492, 1537, 8110, 9447, 1199, 5955, 10487, 1293, 6481, 11527

Sequenz 12: 50 durch den Mitchell-Moore RNG generierte Zufallszahlen

Mitchell und Moore entwickelten daraufhin einen additiven Generator nach folgendem Algorithmus:

$$I_j = (I_{j-24} + I_{j-55})(mod\ m)$$

Formel 12: Formel für den additiven Kongruenzgenerator nach Mitchell und

---

[155] Vgl. (Linearer Kongruenzgenerator: wikipedia.org, 2013)
[156] Vgl. (Dagpunar, 1988) Seite 32
[157] Vgl. (Milavec, Zufall, Ordnung und Chaos am Computer, 1995) Seite 7
[158] Vgl. (Dagpunar, 1988) Seite 32
[159] Vgl. (Bert F. Green, 1959) Seiten 527 – 537 zitiert nach (Knuth, 1969) Seite 26
[160] Im Original: „nice ‚bad example‘“ (Knuth, 1969) Seite 26

Der Verwendung der Werte 24 und 55 liegt eine theoretische Basis zugrunde, denn diese ermöglicht eine vergleichbar hohe, Mindestperiodenlänge von 2^{55} - 1.[161]

Wird diese Idee verallgemeinert, sodass die beiden „seeds“ einer Zufallszahl unterschiedlich weit zurückliegen können (anstatt wie bei Mitchell und Moore auf 24. und 55. vorhergehenden Stelle), so resultiert daraus der Algorithmus für den verzögerten oder „lagged“ Fibonacci RNG[162]:

$$I_j = (I_{j-r} + I_{j-s})(mod\ m)$$

Formel 13: Formel für den verzögerten Fibonacci-Generator

Die Theorie die schon bei Mitchell und Moore wirksam wurde, gilt auch für den verzögerten Fibonacci-Generator. Sie manifestiert sich darin, dass nur gewisse Zahlenpaare verwendet werden können, um eine maximale Periode von $2^s - 1$ zu bewirken.[163]

6.1.6.4. GÜTEKRITERIEN

6.1.6.4.1. PRAKTISCHE ANFORDERUNGEN

Fibonacci-Generatoren sind relativ schnell, da sie im Gegensatz zu den linearen Kongruenzgeneratoren keine Multiplikation erfordert.[164] Generell sind RNGs dieser Art tragbar und die Ergebnisse wiederholbar.

6.1.6.4.2. PERIODENLÄNGE

Besonders im Vergleich zu den linearen Kongruenzgeneratoren ist die durch additive Kongruenzgeneratoren erzeugte Sequenz ausnehmend lang. Während die maximale Periode bei ersteren m beträgt, so kann sie bei letzteren eine Länge von $m^p - 1$ erreichen.

6.1.6.4.3. VERTEILUNG UND UNABHÄNGIGKEIT

Wie zufällig die Verteilung und die Unabhängigkeit aufeinanderfolgender Elemente einer beliebigen Sequenz erscheinen, hängt von den jeweiligen verwendeten Abwandlungen des additiven

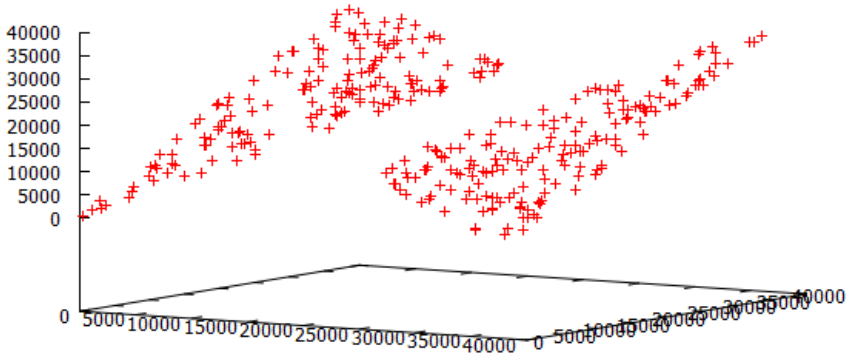

Abbildung 17: Hyperebenen des Fibonacci-Generators (m=38603, n=999)

[161] Vgl. (Knuth, 1969) Seiten 26f
[162] Vgl. (Linearer Kongruenzgenerator: wikipedia.org, 2013)
[163] Vgl. (Linearer Kongruenzgenerator: wikipedia.org, 2013)
[164] Vgl. (Dagpunar, 1988) Seite 7

Kongruenzgenerators ab. Die Fibonacci Methode beispielsweise verursacht bei sehr schlecht gewählten Parametern deutliche Hyperebenen (siehe *Abbildung 17*).

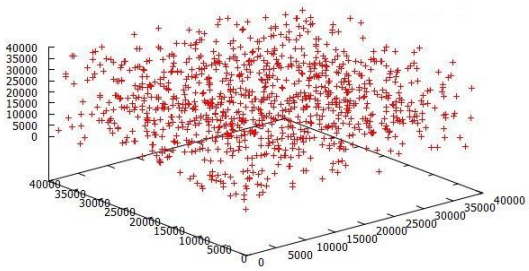

Abbildung 18: Die 3-Tupel eines additiven Kongruenz-generators ergeben keine sichtbaren Strukturen

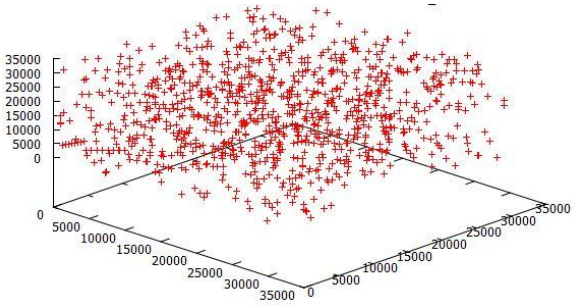

Abbildung 19: Die 3-Tupel des Mitchell- Moore-Generators

Eine Sequenz, erzeugt nach der Formel $I_j = \left(I_{j-1} + I_{j-2} + I_{j-3}\right)(mod\ m)$, mag beim Plotten ihrer 3-Tupel in den Raum eine Grafik wie in *Abbildung 18* erzeugen. Man erkennt, dass dieser nur geringfügig komplexere Generator die starken Muster bedeutend reduziert.

6.1.6.4.4. EMPIRISCHE TESTS

Die Tests der Diehard-Batterie wurden für Tabelle 17 auf den von Mitchell und Moore definierten, verzögerten Fibonacci-Generator angewandt. In diesem Fall besteht die vom Generator erzeugte Sequenz von Zufallszahlen alle Tests, bis auf den Geburtstagsabstände-Test, problemlos.

Test	p-Werte des ersten Tests (meist Chi-Quadrat-Test, selten auch Kolmogoroff-Smirnow-Test)					p-Werte des zweiten Test (Kolmogoroff-Smirnow-Test)
Geburtstagsabstände	1,000000		1,000000		1,000000	1,000000
	1,000000		1,000000		1,000000	
	1,000000		1,000000		1,000000	
Affentest	0,95675	0,78968	0,79170	0,04311	0,00411	kein KS-Test ist hier vorgesehen
	0,39109	0,78833	0,31904	0,84116	0,48385	
	0,18274	0,36183	0,77314	0,51181	0,71225	
	0,88129	0,46154	0,48851	0,17783	0,16768	
OPERM	0,926364			0,829432		kein KS-Test ist hier vorgesehen
Ränge von Matrizen (31x31)	0,320860					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (32x32)	0,341326					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (6x8)	0,914805	0,846804	0,465689	0,980056	0,246127	0,173739
	0,299733	0,398817	0,520520	0,864121	0,881407	
	0,763233	0,376630	0,716774	0,641740	0,661042	
	0,992302	0,017465	0,491517	0,329217	0,201657	
	0,481204	0,326268	0,554985	0,143167	0,187813	
Parkplatztest	0,180558	0,340551	0,625377	0,753306	0,518210	0,227380
	0,108811	0,625377	0,291865	0,767486	0,767486	
OPSO	0,4163	0,8238	0,0689	0,6345	0,3278	kein KS-Test ist hier vorgesehen
	0,6837	0,3106	0,7067	0,3118	0,5923	
	0,5949	0,8887	0,3416	0,1462	0,2176	
	0,0216	0,1686	0,8729	0,9878	0,8900	
	0,7736	0,4789	0,2961			
OQSO	0,4739	0,7946	0,9525	0,6637	0,1343	kein KS-Test ist hier vorgesehen
	0,3330	0,1568	0,9175	0,1617	0,6563	
	0,6686	0,6038	0,8515	0,4124	0,1264	
	0,6078	0,3567	0,8674	0,9144	0,6784	
	0,4456	0,3810	0,7128	0,8854	0,6450	

	0,7140	0,0521	0,2638			
DNA	0,1831	0,2126	0,8748	0,3602	0,5477	kein KS-Test ist hier vorgesehen
	0,5384	0,1001	0,5929	0,3171	0,8302	
	0,8635	0,4867	0,8980	0,2803	0,8642	
	0,8784	0,5137	0,5384	0,9692	0,5489	
	0,4109	0,3680	0,8084	0,2627	0,4914	
	0,4761	0,1731	0,3814	0,1862	0,2169	
	0,0800					
Zähle die 1en (1)	0,846357			0,842735		kein KS-Test ist hier vorgesehen
Zähle die 1en (2)	0,874037	0,921162	0,305811	0,450949	0,410465	kein KS-Test ist hier vorgesehen
	0,439567	0,275317	0,148134	0,015481	0,837264	
	0,128125	0,772802	0,387999	0,485018	0,782483	
	0,836493	0,015994	0,016701	0,592295	0,674035	
	0,470833	0,427934	0,833569	0,923177	0,347268	
Minimumdistanztest	-					0,573004
Zufällige-Kugeln-Test	0,30367	0,78501	0,08064	0,99726	0,30620	0,468207
	0,43010	0,66130	0,57239	0,78243	0,40207	
	0,22982	0,91910	0,42573	0,49482	0,27077	
	0,33955	0,42913	0,35100	0,92156	0,76511	
Squeeze-test	0,121207					kein KS-Test ist hier vorgesehen
Überlappende-Summen-Test	0,069837	0,492373	0,517531	0,164855	0,329915	0,054398
	0,908260	0,700711	0,718067	0,924793	0,603278	
Läufe-test	-					„Raufläufe“ „Runterläufe“
						0,584768 0,785331
						0,263216 0,775167
Craps-test	Anzahl an Gewinnen			Anzahl an Würfe/Spiel		kein KS-Test ist hier vorgesehen
	0,,585238			0, 657645		

Tabelle 17: Exemplarisches Verhalten eines verzögerten Fibonacci-Generators (r=24; s=55) in der Diehard-Testsammlung. Kritische Werte (sehr nahe an 0 oder 1) sind hervorgehoben.

6.1.7. DER INVERSE KONGRUENZGENERATOR (ICG)

1	0	3
2	4	(1)

Sequenz 13: 5 durch einen inversen Kongruenz-generator generierte Zufallszahlen (m=5, a=2, c=3)

Neben linearen wird mit dem inversen, nonlinearen KG von Eichenauer und Lehn (1986)[165] eine weitere Art von Kongruenzgeneratoren vorgestellt. Der wesentliche Vorteil eines inversen Kongruenzgenerators im Vergleich zu den anderen RNGs seiner Kategorie, besteht in ihrem Vermeiden von Hyperebenenstrukturen. Jede Hyperebene in einem n-dimensionalen Raum enthält maximal d Punkte, weshalb die Verwendung dieses Generators adäquat für diverse Simulationen ist.[166]

6.1.7.1. METHODE

Diese Art von Zufallszahlengeneratoren wird über die Formel

$$I_j = \left(a\overline{I_{j-1}} + c\ \right)(mod\ m)$$

Formel 14: Formel für den inversen Kongruenzgenerator

definiert.

[165] Vgl. (Hellekalek, 2014)
[166] Vgl. (Eichenauer-Herrmann J. , 1991) Seite 297, bzw. 1

Nimmt I_{j-1} den Wert 0 an, so gilt $\overline{I_{j-1}} = 0$. Für $I_{j-1} \neq 0$ fällt die Bestimmung seines sogenannten multiplikativ-inversen Elements $\overline{I_{j-1}}$ komplexer aus:

Generell gilt: $\left(\overline{I_{j-1}} \times I_{j-1}\right)(mod\ m) = 1(mod\ m)$ [167] Das bedeutet, dass $\overline{I_{j-1}}$ mit I_{j-1} multipliziert und durch m dividiert, denselben Rest wie 1 durch m haben muss. Ist m beispielsweise 5, so wäre das multiplikativ-inverse Element zu 3 (innerhalb des Wertebereichs m) entweder die Zahl 2, da $(2 \times 3)(mod\ 5) = 1(mod\ 5)$ oder auch $\frac{1}{3}$, da auch $\left(\frac{1}{3} \times 3\right)(mod\ 5) = 1(mod\ 5)$. Im zweiten Fall liegt das multiplikativ-inverse Element $\overline{I_{j-1}}$ nicht im Zahlenraum der ganzen Zahlen $\mathbb{Z}$, sondern in $\mathbb{R}$. Im Zuge der Zufallszahlengenerierung soll I_j aber immer eine ganze Zahl sein, darum wird das in $\mathbb{R}$ liegende multiplikativ-inverse Element hier übergangen.

6.1.7.2. Der explizit inverse Kongruenzgenerator (EICG)

Der explizit inverse Kongruenzgenerator kann als der „der lässige Bruder des ICG“[168] bezeichnet werden und gleicht seinem „großen Bruder“ in den wesentlichen Eigenschaften. Der Hauptunterschied liegt darin, dass er nicht über einen rekursiven Algorithmus definiert wird:[169]

$$I_j = \left(\overline{a(J - J_0) + c}\right)(mod\ m)$$

Formel 15: Formel für den explizit inversen Kongruenzgenerator

6.1.7.3. Gütekriterien

6.1.7.3.1. Praktische Anforderungen

Die „Invertierung“ der Zufallszahl erhöht die Berechnungskomplexität im Vergleich zu einem linearen Kongruenzgenerator erheblich. Im ungünstigsten Fall ist die Zeiteinbuße um den Faktor 3 höher als beim allgemeinen LCG.[170]

6.1.7.3.2. Periodenlänge

Die maximale Sequenzlänge dieses Generators beträgt m, allerdings nur dann, wenn er bestimmten Anforderungen gerecht wird.[171]

6.1.7.3.3. Verteilung und Unabhängigkeit

Wie schon zu Beginn dieses Abschnittes erörtert, besteht der wesentliche Vorteil der ICGs im Vergleich zu den LCGs darin, dass sie keine Hyperebenen bilden.

Die Unabhängigkeit der Elemente der Sequenz ist von den gewählten Parametern unabhängig, wenn also zudem a und c so gewählt wurden, dass die maximale Sequenzlänge gewährleistet ist, so sollten laut Wahler, Rose und Schömig zuverlässige Sequenzen generiert werden können.[172] Wird der

[167] Vgl. (Wahler, Rose, & Schömig, 1997) Seite 8
[168] Im Original: „‘easy-going brother‘ of the ICG“ (Hellekalek, 2014)
[169] Vgl. (Wahler, Rose, & Schömig, 1997) Seite 10f
[170] Vgl. (Wahler, Rose, & Schömig, 1997) Seite 8
[171] Vgl. (Wahler, Rose, & Schömig, 1997) Seite 9
[172] Vgl. (Wahler, Rose, & Schömig, 1997) Seite 9

Kolmogorow-Smirnow-Test auf eine durch den ICG generierte Sequenz von Zufallszahlen angewandt, so besteht diese ihn ohne große Probleme.[173]

6.1.7.3.4. Empirische Tests

Laut dem Informatiker George Marsaglia liegt der Wert dieser Art von Generatoren hauptsächlich darin, dass deren theoretische Basis interessant ist. Der inverse Kongruenzgenerator besteht einen Großteil der Diehard-Tests nicht und weist in diesem Gebiet ähnliche oder sogar schlechtere Eigenschaften auf als die herkömmlichen LCG.[174]

6.1.8. Marsaglia-Zaman-Generatoren

Im Jahre 1991 stellten George Marsaglia und Arif Zaman von der State University eine neue Art von Zufallszahlengeneratoren vor, zu denen die Methoden „Add-with-carry“ und „Subtract-with-borrow“ zählen. Der Vorteil beider Generatoren liegt in ihren langen Perioden und der nachweisbaren Gleichverteilung der Zufallszahlen. [175] Ihr Prinzip wird im Folgenden kurz vorgestellt, da ihr charakterisierendes Element ebenfalls essentiell für einen der bedeutendsten existierenden PRNGs (den Multiply-with-carry) ist.

6.1.8.1. Add with carry (AWC)

Dieser Generator weist gewisse Ähnlichkeiten zum verzögerten Fibonacci-Generator auf, allerdings sieht der AWC das Designieren eines sogenannten „carry bit“, das sich bei jeder Iteration ändert und entweder einen Wert von 0 oder 1 annimmt, vor. [176] Die entsprechende Formel lautet:

$$I_j = \left(I_{j-r} + I_{j-s} + carry\right)(mod\ m)$$

Formel 16: Formel für den Add-with-carry-Generator

Für den Wert eines „carry bits“ gilt:

$$Wenn\ I_{j-r} + I_{j-s} + carry_{\ 1} < b \quad dann \quad carry_2 = 0$$
$$Wenn\ I_{j-r} + I_{j-s} + carry_{\ 1} \geq b \quad dann \quad carry_2 = 1$$

Formel 17: Formel für den carry des Add-with-carry Generators

Für den AWC wird gewöhnlich ein b von 10 gewählt. [177] Wird r ein Wert von 1 und s ein Wert von 2 zugeschrieben, so kann die offensichtliche Relation zu einem herkömmlichen Fibonacci-Generator erkannt werden.

Die maximale Periodenlänge dieses Generators ist $b^r + b^s - 2$.[178]

[173] Vgl. (Eichenauer-Herrmann J. , 1994) Seite 783 und (Niederreiter, 1992)
[174] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[175] Vgl. (Marsaglia & Zaman, A New Class of Random Number Generators, 1991) Seite 462
[176] Vgl. (Marsaglia & Zaman, A New Class of Random Number Generators, 1991) Seite 464f
[177] Vgl. (Marsaglia & Zaman, A New Class of Random Number Generators, 1991) Seite 464f
[178] Vgl. (Tezuka, L'Ecuyer, & Couture, 1995) Seite 2

6.1.8.2. SUBTRACT WITH BORROW (SWB)

Dieser Generator basiert praktisch auf der Umkehrung der Rechenregeln des Add-with-carry.

Die entsprechende Formel lautet entsprechend:

$$I_j = (I_{j-r} - I_{j-s} - carry)(mod\ m)$$

Formel 18: Formel für den Subtract-with-borrow-Generator

Für den Wert eines „carry bits“ gilt:

$$Wenn\ I_{j-r} - I_{j-s} - carry_{\ 1} < 0 \quad dann \quad carry_2 = 0$$
$$Wenn\ I_{i-r} - I_{i-s} - carry_{\ 1} \geq 0 \quad dann \quad carry_2 = 1$$

Formel 19: Formel für den carry des Subtract-with-borrow-Generators

Für den SWB wird gewöhnlich ein b von 10 gewählt.[179] Wird r ein Wert von 1 und s ein Wert von 2 zugeschrieben, so kann die offensichtliche Relation zu einem herkömmlichen Fibonacci-Generator erkannt werden.

6.1.8.3. GÜTEKRITERIEN

Diese beiden RNGs sind dem verzögerten Fibonacci-Generator sehr ähnlich. Sie haben beide lange Perioden und schneiden auch bei den empirischen Tests (siehe *Tabelle 18*) ähnlich gut ab.[180]

Test	p-Werte des ersten Tests (meist Chi-Quadrat-Test, selten auch Kolmogoroff-Smirnow-Test)					p-Werte des zweiten Test (Kolmogoroff-Smirnow-Test)
Geburtstagsabstände	1,000000		1,000000		1,000000	1,000000
	1,000000		1,000000		1,000000	
	1,000000		1,000000		1,000000	
Affentest	0,88496	0,53971	0,15399	0,84564	0,57841	kein KS-Test ist hier vorgesehen
	0,89883	0,15622	0,13540	0,24509	0,98797	
	0,82839	0,92148	0,31904	0,22359	0,04551	
	0,00250	0,87078	0,32490	0,33591	0,87611	
OPERM	0,879219			0,455769		kein KS-Test ist hier vorgesehen
Ränge von Matrizen (31x31)	0,978126					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (32x32)	0,644719					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (6x8)	0,518613	0,183674	0,403555	0,794729	0,971652	0,369660
	0,384395	0,666352	0,615107	0,272532	0,826568	
	0,463099	0,609373	0,987806	0,134022	0,245597	
	0,841385	0,606482	0,178175	0,417754	0,698498	
	0,322465	0,254503	0,798523	0,304256	0,477356	
Parkplatztest	0,659449	0,340551	0,500000	0,323972	0,723613	0,519287
	0,463618	0,276387	0,607947	0,708135	0,934075	
OPSO	0,7412	0,4002	0,2866	0,1893	0,8184	kein KS-Test ist hier vorgesehen
	0,8938	0,3790	0,5161	0,5869	0,3316	
	0,2855	0,9466	0,7566	0,7673	0,4625	
	0,6676	0,7967	0,1584	0,3896	0,0386	
	0,8101	0,4163	0,4203			
OQSO	0,7508	0,7550	0,6760	0,7888	0,9755	kein KS-Test ist hier vorgesehen

[179] Vgl. (Marsaglia & Zaman, A New Class of Random Number Generators, 1991) Seite 464f

[180] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

	0,0027	0,3087	0,7789	0,8419	0,9269	
	0,2594	0,6117	0,4631	0,0105	0,6550	
	0,9728	0,8546	0,5198	0,1727	0,9264	
	0,8232	0,7023	0,6284	0,5090	0,5881	
	0,1780	0,6893	0,2594			
DNA	0,5196	0,0143	0,5582	0,0761	0,4667	kein KS-Test ist hier vorgesehen
	0,9660	0,8964	0,2862	0,0661	0,5020	
	0,5255	0,4890	0,5524	0,3547	0,4667	
	0,7782	0,0413	0,5675	0,5090	0,3438	
	0,6402	0,1910	0,7348	0,0684	0,2588	
	0,2466	0,9673	0,4973	0,7556	0,2763	
	0,5126					
Zähle die 1en (1)	0,418371			0,877437		kein KS-Test ist hier vorgesehen
Zähle die 1en (2)	0,275464	0,063417	0,500799	0,420622	0,840136	kein KS-Test ist hier vorgesehen
	0,820103	0,737408	0,569578	0,937045	0,261199	
	0,804347	0,633297	0,360645	0,350522	0,481673	
	0,088323	0,785861	0,064731	0,956001	0,256884	
	0,097499	0,075377	0,782542	0,987213	0,956963	
Minimumdistanztest	-					0,766313
Zufällige-Kugeln-Test	0,77387	0,68886	0,86249	0,78470	0,56102	0,549946
	0,64310	0,26014	0,02239	0,98281	0,04814	
	0,92615	0,17022	0,86948	0,07564	0,35608	
	0,89095	0,70097	0,04224	0,44811	0,68439	
Squeeze-test	0,895423					kein KS-Test ist hier vorgesehen
Überlappende-Summen-Test	0,547240	0,903548	0,247380	0,192504	0,629629	0,133208
	0,317846	0,740587	0,131306	0,738208	0,258321	
Läufe-test	-					„Raufläufe“ „Runterläufe“
						0,993104 0,102433
						0,075308 0,669737
Craps-test	Anzahl an Gewinnen			Anzahl an Würfe/Spiel		kein KS-Test ist hier vorgesehen
	0,638317			0,134943		

Tabelle 18: Exemplarisches Verhalten eines Subtract-with-borrow-Generators (s=24; r=37 m=2^{32}) in der Diehard-Testsammlung. Kritische Werte (sehr nahe an 0 oder 1) sind hervorgehoben.

6.1.9. MULTIPLY WITH CARRY (MWC)

Im Mai 2003 entwickelte George Marsaglia, in Anlehnung an seinen zuvor mit Zaman beschriebenen Add-with-carry PRNG, einen neuen Pseudozufallszahlengenerator, den Multiply-with-carry (MWC).[181]

6.1.9.1. METHODE

Die einfache Formel hinter Marsaglias MWC lautet:

$$I_j = \left(aI_{j-1} + carry_{j-1}\right)(mod\ m)$$

Formel 20: Formel des MWC-Generators

Für den Carry gilt:

$$carry_j = \left\lfloor \frac{\left(aI_{j-1} + carry_{j-1}\right)}{b} \right\rfloor$$

Formel 21: Formel für den carry des MWC-Generators. $\lfloor x \rfloor$ oder auch floor(x) bezeichnet die Gaußklammer, sie gibt die größte ganze Zahl, die kleiner oder gleich x ist, an (siehe *5.6.1.10*)

[181] Vgl. (Marsaglia, Random Number Generators, 2003) Seiten 2-13

6.1.9.2. GÜTEKRITERIEN

6.1.9.2.1. PRAKTISCHE KRITERIEN

Der MWC benötigt Zahlen, die durch physikalische Zufallszahlengeneratoren erzeugt wurden, als „seeds“, da er ansonsten erst nach einer etwaigen „Warmlaufphase“ gleichverteilte Zahlen generiert.[182] Dieser PRNG ist verhältnismäßig schnell[183] und seine Ergebnisse sind, wie jene anderer PRNGs, wiederholbar.

6.1.9.2.2. PERIODENLÄNGE

Der große Vorteil dieser Art von Zufallszahlengeneratoren liegt in ihrer, besonders im Vergleich zu den linearen Kongruenzgeneratoren (die eine maximalen Periodenlänge vom nur m aufweisen) extrem langen Sequenzlänge.[184]

6.1.9.2.3. VERTEILUNG UND UNABHÄNGIGKEIT

Nach der „Warmlaufphase“ erzeugt dieser PRNG gleichverteilte Zufallszahlen,[185] die eine niedrige serielle Korrelation aufweisen. Er stellt auch insofern eine Verbesserung der LCG dar, als dass er keine Hyperebenen bildet.

6.1.9.2.4. EMPIRISCHE TESTS

George Marsaglia selbst war davon überzeugt, dass der MWC den linearen Kongruenzgenerator als beliebtesten und am häufigsten eingesetzten PRNG ersetzen würde. Diese Annahme begründet er nicht zuletzt darin, dass der Multiply-with-carry sich ausgezeichnet in den empirischen Tests verhält.[186]

Test	p-Werte des ersten Tests (meist Chi-Quadrat-Test, selten auch Kolmogoroff-Smirnow-Test)					p-Werte des zweiten Test (Kolmogoroff-Smirnow-Test)
Geburtstagsabstände	0,305990		0,841998		0,702975	0,803954
	0,401961		0,711431		0,516435	
	0,557861		0,973123		0,838907	
Affentest	0,803954	0,90331	0,85489	0,86015	0,33250	kein KS-Test ist hier vorgesehen
	0,19732	0,15344	0,17420	0,44765	0,72953	
	0,34705	0,71781	0,99203	0,67791	0,10679	
	0,42373	0,40915	0,56189	0,89842	0,75663	
OPERM	0,015984			0,046516		kein KS-Test ist hier vorgesehen
Ränge von Matrizen (32x32)	0,931982					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (31x31)	0,760710					kein KS-Test ist hier vorgesehen
Ränge von Matrizen (6x8)	0,671600	0,857414	0,463783	0,725270	0,132573	0,872124
	0,190799	0,358332	0,356497	0,192235	0,159959	
	0,242293	0,631777	0,095338	0,504848	0,602928	
	0,155751	0,328030	0,228766	0,188069	0,688158	
	0,076182	0,010591	0,579102	0,849488	0,756067	

[182] Vgl. (Multiply with carry: wikipedia.org, 2013)
[183] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)
[184] Vgl. (Multiply with carry: wikipedia.org, 2013)
[185] Vgl. (Multiply with carry: wikipedia.org, 2013)
[186] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

Parkplatztest	0,982156 0,168804	0,157553 0,891189	0,108811 0,642555	0,500000 0,554479	0,975204 0,392053	0,427810
OPSO	0,6563 0,2514 0,9653 0,4069 0,3480	0,4803 0,8678 0,5298 0,3634 0,4461	0,1266 0,0594 0,9180 0,9048 0,2985	0,7390 0,2046 0,5707 0,6689	0,6069 0,3989 0,8976 0,2395	kein KS-Test ist hier vorgesehen
OQSO	0,2561 0,0832 0,4671 0,7966 0,7561 0,8267	0,6284 0,8854 0,3442 0,8546 0,3504 0,6051	0,3875 0,9420 0,1888 0,6323 0,8327 0,3914	0,1175 0,6905 0,7421 0,1710 0,2086	0,2358 0,9384 0,4071 0,2518 0,7321	kein KS-Test ist hier vorgesehen
DNA	0,0357 0,7141 0,4996 0,3938 0,5396 0,3825 0,8052	0,2512 0,6674 0,3702 0,8432 0,3384 0,1136	0,5872 0,2793 0,1563 0,7894 0,6780 0,4363	0,2033 0,7319 0,5872 0,7620 0,7434 0,2301	0,9374 0,0056 0,9539 0,4340 0,1831 0,6346	kein KS-Test ist hier vorgesehen
Zähle die 1en (1)	0,685650			0,397062		kein KS-Test ist hier vorgesehen
Zähle die 1en (2)	0,205138 0,963766 0,683016 0,058914 0,634241	0,336875 0,438389 0,707221 0,230738 0,015270	0,865703 0,678198 0,749498 0,565538 0,293666	0,648975 0,808719 0,455330 0,493188 0,541272	0,570880 0,729105 0,818628 0,700176 0,120794	kein KS-Test ist hier vorgesehen
Minimumdistanztest	-					0,050850
Zufällige-Kugeln-Test	0,61841 0,31171 0,29292 0,06068	0,49120 0,73327 0,11213 0,17614	0,16368 0,21711 0,25286 0,42042	0,73597 0,03821 0,05802 0,35531	0,88528 0,45380 0,58878 0,46551	0,921109
Squeeze-test	0,093860					kein KS-Test ist hier vorgesehen
Überlappende-Summen-Test	0,915753 0,414646	0,941622 0,103141	0,017796 0,823972	0,618955 0,752187	0,345700 0,106864	0,211705
Läufe-test	-					„Raufläufe“ „Runterläufe“ 0,496844 0,578668 0,537900 0,350337
Craps-test	Anzahl an Gewinnen 0,826358			Anzahl an Würfe/Spiel 0,594892		kein KS-Test ist hier vorgesehen

Tabelle 19: Exemplarisches Verhalten eines Multiply-with-carry-Generators (a=1683268614; m=2^{32}) in der Diehard-Testsammlung. Kritische Werte (sehr nahe an 0 oder 1) sind hervorgehoben.

6.1.10. DER MERSENNE-TWISTER

Der Mersenne-Twister wurde 1997 von Makoto Matsumoto und Takuji Nishimura vorgestellt.[187] Mittlerweile existieren drei Generationen dieses RNGs, der TinyMT (2001), der MTGP (2009) und der SFMT (2007), wobei sich letzterer in seiner höheren Geschwindigkeit, besseren Gleichverteilung und einer kürzeren „Warmlaufphase“ vom originalen Mersenne-Twister unterscheidet.[188]

6.1.10.1. DER NAME

Ursprünglich war vorgesehen, den Mersenne-Twister „Primitive Twisted Generalized Feedback Shift Register Sequence“ zu taufen, was Donald Knuth mit „the name is mouthful“[189] kommentiert haben soll. Makoto Matsumoto schlug daraufhin den Namen „Mersenne Twister“ vor, da sich dieser PRNG

[187] Vgl. (Matsumoto & Nishimura, Mersenne twister. A 623-dimensionally equidistributed uniform pseudorandom number generator, 1998) Seite 3-30
[188] Vgl. (Mersenne Twister Home Page, 2014)
[189] „Der Name ist allerhand“

der Mersenne-Primzahlen bedient und wie eine „Achterbahn“, also „ziemlich schnell, einfach zu merken und auszusprechen“, ist. Zudem sind die Initialen der Entwickler des „MT“ in diesem Namen versteckt.[190]

6.1.10.2. METHODE[191]

Für die Umsetzung des Mersenne-Twisters müssen zuerst die Startwerte I_1 bis I_n (n=624) mit einem anderen RNG generiert werden. Der darauffolgende Algorithmus heißt:

$$h := I_{j-n} - I_{j-n} mod 2^{31} + I_{j-n+1} mod\ 2^{31}$$
$$I_j := I_{j-227} \oplus \left\lfloor \frac{h}{2} \right\rfloor \oplus ((h\ mod\ 2) \times 9908b0df_{hex}$$

Formel 22: Definition des Mersenne-Twister. $\lfloor x \rfloor$ ist das Symbol für die Gaußklammer, $\oplus$ bezeichnet die Kontravalenz (auch XOR-Verknüpfung), $_{hex}$ steht für hexadezimal.

6.1.10.3. GÜTEKRITERIEN

Der Mersenne-Twister wurde mit dem Hintergedanken kreiert, einen PRNG zu schaffen, der frei von den Defekten herkömmlicher Zufallszahlengeneratoren ist.[192]

6.1.10.3.1. PRAKTISCHE ANFORDERUNGEN

Wie bereits angedeutet, muss vor Initialisierung des Mersenne-Twisters erst eine große Anzahl an Zufallszahlen mittels eines anderen PRNGs oder RNGs erzeugt werden. Dies ist sicherlich ein Faktor, der vor allem die ersten Werte der generierten Zufallszahlensequenz unsicher machen kann, vor allem wenn die „seeds“ keinem verlässlichen RNG entstammen.

Die Generierung der Zufallszahlen ist verhältnismäßig schnell und belegt keine beträchtliche Größe im Cache.[193]

6.1.10.3.2. PERIODENLÄNGE

Matsumotos und Nishimuras RNG hat eine bestechend lange Periodenlänge von 2^{19937}-1. [194]

6.1.10.3.3. VERTEILUNG UND UNABHÄNGIGKEIT

Der Mersenne-Twister liefert, wie alle anderen durch lineare Rekursion erzeugten Zufallszahlensequenzen, keinen Garant für kryptographische Sicherheit, ist dafür aber für Simulationen gut geeignet.[195] Die entstehenden Zufallszahlen sind gleichverteilt und weisen eine sehr niedrige sequentielle Korrelation auf.[196]

[190] Vgl. (Matsumoto & Nishimura, The origin of the name MT: Mersenne Twister Home Page, 2014)
[191] Vgl. (Mersenne-Twister: wikipedia.org, 2013)
[192] Vgl. (Mersenne Twister Home Page, 2014)
[193] Vgl. (Mersenne Twister Home Page, 2014)
[194] Vgl. (Mersenne Twister Home Page, 2014)
[195] Vgl. (Mersenne Twister Home Page, 2014)
[196] Vgl. (Mersenne-Twister: wikipedia.org, 2013)

6.1.11. ANDERE METHODEN ZUR PSEUDOZUFALLSZAHLENGENERIERUNG

6.1.11.1. WELL[197]

Das Akronym WELL steht für einen von Francois Panneton und Pierre L'Ecuyer an der Universität von Montréal entwickelten Zufallszahlengenerator mit Namen „Well Equidistributed Long-period Linear". Diese Klasse an Zufallszahlengeneratoren ist zwar vergleichsweise komplex, doch im Vergleich zum Mersenne-Twister und anderen PRNGs mit langen Perioden, nicht weniger schnell. Zudem weisen die erzeugten Zufallszahlen eine bessere Gleichverteilung (Equidistribution) als die meisten anderen PRNGs mit vergleichbarer Sequenzlänge auf.

Die Mängel eines WELL-Generators verdeutlichen sich, wenn dieser von statistischen Tests beleuchtet wird. Wenngleich er etliche dieser Tests problemlos passiert, scheitert der WELL-Generator an Tests, die lineare Abhängigkeiten in langen Zufallszahlensequenzen prüfen (wie der Matrizen-Test).

Die Schnelligkeit dieser Generatoren und die Länge der durch sie erzeugten Sequenzen gewährleisten eine große Sachdienlichkeit und Sicherheit für eine Vielzahl an Simulationen.

6.1.11.2. PRIMZAHLEN

Primzahlen können, ähnlich wie die Kreiszahl Pi, als Quelle von zwar deterministischen, aber dennoch aperiodischen Zufallszahlensequenzen gesehen werden.

Dazu wird nicht der Wert der Primzahl per se, sondern die Abstände zwischen zwei Primzahlen betrachtet. Die Verteilung dieser Zahlen, die nur durch sich selbst und 1 teilbar sind, wirkt, obwohl sie deterministisch ist (es gibt feste Kriterien, über die eine Primzahl definiert werden kann), zufällig.[198] Zwar werden die Abstände mit wachsenden Stellen immer größer, doch wie bereits im Eingangskapitel beschrieben, bedeuten Distributionen, die von der Gleichverteilung abweichen, nicht, dass die Zahlen nicht zufällig wären. Aus den ersten 45 Primzahlen lassen sich durch diese Methode nun Zufallszahlen generieren:

Primzahlen 2-181	Differenzen zwischen den Primzahlen 2-181
2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167, 173, 179, 181	3-2, 5-3, 7-5, 11-7, 13-11, 17-13, 19-17, 23-19, 29-23, 31-29, 37-31, 41-37, 43-41, 47-43, 53-47, 59-53, 61-59, 67-61, 71-67, 73-71, 79-73, 83-79, 89-83, 97-89, 101-97, 103-101, 107-103, 109-107, 113-109, 127-113, 131-127, 137-131, 139-137, 149-139, 151-149, 157-151, 163-157, 167-163, 173-167, 179-173, 181-179

1, 2, 2, 4, 2, 4, 2, 4, 6, 2, 6, 4, 2, 4, 6, 6, 2, 6, 4, 2, 6, 4, 6, 8, 4, 2, 4, 2, 6, 14, 4, 6, 2, 10, 2, 6, 4, 4, 6, 6, 2

Sequenz 14: Durch Ermittlung von Primzahlendifferenzen erzeugte Zufallszahlen

Sequenz 14, eine durch Primzahlendifferenzen gebildete Sequenz von Pseudozufallszahlen, weist erhebliche Mängel auf. So sind beispielsweise alle Primzahlen ab 2 ungerade, weshalb die

[197] Vgl. (Panneton & L'Ecuyer) Seite 10ff

[198] Vgl. (Roney-Dougal, du Sautoy, & Gowers, 2011) circa ab Minute 15

numerischen Abstände zwischen den Zahlen, also die Zufallszahlen, ab dem ersten Element allesamt gerade sind.

6.1.11.3. XORSHIFT

Xorshift ist der Name einer Klasse von Zufallszahlengeneratoren, die schnell, einfach und leicht als Teiloperation in andere PRNGs integrierbar sind. Ihre Periodenlängen betragen 2^k-1, wobei k die Werte 32, 64, 96, 128, 160 und 192 annehmen kann. [199] Xorshift-Generatoren passieren laut ihrem Entwickler Marsaglia die gängigen Tests auf Zufälligkeit, so auch die Diehard-Batterie, wohingegen Panneton und L'Ecuyer in *„On the Xorshift Random Number Generators*" veranschaulichten, dass PRNGs mit sehr wenigen und simplen Xorshift-Operationen teilweise auch einfache statistische Tests nicht bestehen.[200] Vorsicht ob leichtfertiger Implementation ist also dennoch gefragt.

6.1.11.4. KISS[201]

Der KISS-Generator ist ein von George Marsaglia entwickelter PRNG, dessen Name für „Keep It Simple, Stupid" („Halte es einfach, Dummkopf") steht. Diese Methode kombiniert mit dem linearen Kongruenzgenerator, Xorshift und einem abgewandelten Multiply-With-Carry-Generator, drei schnelle und einfach zu programmierende Generatoren. Die Periodenlänge des Generators beträgt $>2^{127}$. Marsaglia empfiehlt KISS besonders an Assembler-Programmierer/innen, da neben seiner Einfachheit auch der Umstand für ihn spricht, dass er alle statistischen Tests zu bestehen scheint.

[199] Vgl. (Marsaglia, Xorshift RNGs, 2003) Seite 1ff
[200] Vgl. (L'Ecuyer, 2005) Seite 1
[201] Vgl. (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

6.2. PHYSIKALISCHE ZUFALLSZAHLENGENERATOREN

Wie schon im Einführungskapitel beschrieben, können nur durch physikalische Prozesse generierte Zufallszahlen tatsächlich als solche bezeichnet werden. Nach dem heutigen physikalischen Weltbild sind etwa radioaktive Zerfälle oder der exakte Aufenthaltsort eines Elektrons oder Photons, vom Zufall bestimmt, daher eignen sich diese Vorgänge gut zum Erzeugen „echter" Zufallszahlensequenzen. Es gibt für jedes Zufallsexperiment eine Vielfalt an Methoden davon Zufallszahlen abzuleiten, ob durch die numerische Abweichung von der erwarteten Verteilung oder durch Messung der Zeitintervalle zwischen zwei zufällig auftretenden Ereignissen.

Generelle Vorteile	Generelle Nachteile
Meist tatsächlich zufällig oder zumindest chaotisch, also sehr schwer zu berechnen	Komplizierte und langwierige Erzeugung
Sehr sicher und zuverlässig	Nicht wiederholbar
Beliebig lange Sequenzen möglich	Schwer zu kontrollierende Verfälschungen durch Messungen[202]
Ist das Zufallsexperiment wirklich zufällig und die Messung korrekt, so besteht die Sequenz die Tests auf Verteilung und Unabhängigkeit	Meist nicht gleichverteilt
Besteht die empirischen Tests (Da sich empirische Tests an den Eigenschaften physikalischer RNGs orientieren)	

Tabelle 20: Generelle Vorteile und Nachteile physikalischer RNGs

[202] Vgl. (Mordasini & Klahr, 2013)

6.2.1. DER MÜNZWURF, WÜRFEL UND LOTTOZAHLEN

5, 4, 2, 3, 3, 5, 2, 1, 4, 1, 5, 5, 4, 2, 2, 2, 2, 4, 4, 3, 4, 5, 1, 3, 6,
2, 5, 1, 2, 2, 5, 5, 4, 5, 5, 1, 5, 1, 3, 5, 5, 3, 3, 2, 6, 6, 6, 2, 3, 2,
2, 3, 3, 2, 6, 2, 1, 3, 2, 2, 4, 6, 3, 5, 6, 2, 3, 5, 5, 5, 3, 4, 6, 3, 3,
5, 1, 3, 3, 1, 5, 6, 4, 3, 3, 2, 6, 5, 5, 3, 4, 4, 4, 5, 5, 4, 4, 5, 2, 2,
5, 2, 1, 6, 2, 5, 5, 3, 2, 4, 3, 1, 3, 5, 1, 4, 6, 3, 1, 6, 1, 4, 5, 2, 1,
5, 3, 3, 4, 1, 3, 1, 1, 3, 2, 6, 1, 4, 4, 2, 4, 2, 5, 2, 2, 4, 5, 1, 1, 6,
5, 2, 1, 5, 4, 5, 5, 3, 6, 3, 3, 2, 5, 1, 6, 5, 5, 5, 1, 5, 6, 2, 2, 5, 5,
5, 6, 1, 1, 2, 1, 5, 1, 6, 2, 3, 4, 6, 2, 3, 6, 2, 3, 1, 5, 5, 2, 6, 6, 3

Sequenz 15: 200 durch einen Würfel generierte Zufallszahlen
n= 200; m=6; F(x)=idealerweise ⅙ (gleichverteilt)

6.2.1.1. DIE ZUFÄLLIGKEIT EINES MÜNZWURFS

Es gibt auch physikalische Zufallszahlengeneratoren, deren Status fragwürdig ist. Dazu gehören unter anderem Würfelspiele, Münzwürfe oder die Lottozahlenziehungen, da deren Ergebnisse streng genommen berechnet werden könnten, vorausgesetzt der genaue Anfangszustand dieser Systeme ist bekannt.

In den 1970ern gelang es einer Gruppe von Wissenschaftlern an der University of California at Santa Cruz ein Computerprogramm zu schreiben, das Resultate eines Rouletterads analysieren und aufgrund der dadurch gewonnen Daten mit ziemlicher Genauigkeit vorhersagen konnte, welche Zahlen Gewinn bringen würden. [203] Zudem ist auch bekannt, dass langjährige Croupiers sehr wohl in einem Maße mit den Besonderheiten ihres Rouletterads vertraut sind, welches das Bedenken zulässt, der Ausgang des „Zufallsexperimentes Roulette“ könne mehr vom Croupier als vom Zufall determiniert sein.[204]

Auch ein Würfel ist nicht perfekt, denn die kleinsten Fehler in der Produktion (etwa ein minimaler Gewichtsunterschied auf einer der Seiten) bewirkt eine Verfälschung der Wahrscheinlichkeiten. Ein idealer, perfekter Würfel, dessen Seiten jeweils eine exakte Wahrscheinlichkeit von ⅙ haben obenauf zu landen, wird Laplace’scher Würfel genannt.

Selbst der Münzwurf, mehrfacher Streitschlichter und Spielentscheider, ist nicht perfekt. Besonders, wenn man die Münze zuerst auf den Boden fallen lässt, kann die Wahrscheinlichkeit beträchtlich manipuliert werden. Der Wurf eines US-Penny beispielsweise liefert in nur 30% der Fälle Kopf, wenn die Münze auf den Boden aufkommt.[205]

6.2.1.2. GÜTEKRITERIEN

Wie auch für die deterministischen Zufallszahlengeneratoren können die im Abschnitt *5. Gütekriterien und -tests für Zufallszahlengeneratoren* beschriebenen Tests auch auf die

[203] Vgl. (Peterson, 1998) Seite 182
[204] Vgl. (Peterson, 1998) Seite 182
[205] Vgl. (Peterson, 1998) Seite 4f

physikalischen RNGs angewandt werden, wobei sich im Zuge dieser Analyse einige Redundanzen ergeben werden.

6.2.1.2.1. Praktische Anforderungen

Physikalische Prozesse wie das Würfelspiel oder die Lottozahlenziehung sind im Allgemeinen sehr impraktikabel, da sie sehr viel Zeit in Anspruch nehmen und sich kaum auf Hardwareelementen implementieren lassen. Für einzelne Anwendungen fällt dieser Aufwand allerdings nicht ins Gewicht, da etwa Glücksspiele keine langen Sequenzen von Zufallszahlen benötigen, sondern das dahinführende Zufallsexperiment (das Drehen des Rouletterads, das Rollen der Lottokugeln) per se zum Zelebrieren des Spiels gehört.

Das Generieren einer bestimmten Sequenz physikalischer Zufallszahlen ist nicht wiederholbar. Selbst für sehr kurze Sequenzen erkennt man, dass es sehr unwahrscheinlich ist zweimal hintereinander dieselbe zu erhalten. Daher sind diese RNGs etwa für die Kryptographie ungeeignet.

6.2.1.2.2. Periodenlänge

Diese Zufallsexperimente können bis in die Unendlichkeit fortgesetzt werden, allerdings wird diese theoretisch außerordentlich lange Periodenlänge wegen der langen Generierungszeit und des unverhältnismäßigen Aufwands selten ausgenutzt.

6.2.1.2.3. Verteilung und Unabhängigkeit

Ein Laplace'scher Würfel, also ein Würfel dessen Seiten jeweils eine Wahrscheinlichkeit von ⅙ haben obenauf zu landen, liefert Verteilungen, die sowohl den χ^2- als auch den Kolmogorow-Smirnow-Test bestehen (Tatsächlich würde für ein Würfelspiel oder ein Münzwurf eher der χ^2-Test angewandt, da es sich bei der erwarteten Verteilung um eine diskrete Funktion handelt).

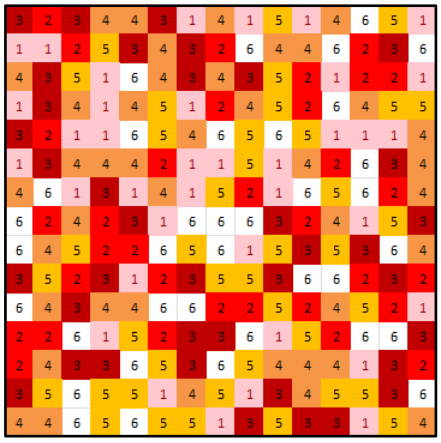

Abbildung 20: Visuelle Darstellung der Ergebnisse eines Würfelspiels. Es sind keine Muster zu erkennen (n=225)

Für eine große Anzahl an Versuchen n zeigt sich auch in empirischen Tests, dass die Kopf:Zahl-Ratio ziemlich genau bei 1:1 liegt. Der französische Naturalist Georges-Louis Leclerc (1707-1788) warf eine Münze beispielsweise 4.040 Male und erhielt 2 048 Mal Kopf (50,69%), während der Englische

Mathematiker Karl Pearson (1857–1936) 24 000 Mal eine Münze warf und 12.012 Mal Kopf erhielt (50,05%)[206]. Ähnlich wie mit der Verteilung verhält es sich mit den diversen Unabhängigkeitstests. Das Problem voneinander abhängiger Sequenzelemente in der Zufallszahlengenerierung entsteht in erster Linie dadurch, dass der „output“ des einen Rechenvorganges als „seed“ für den darauffolgenden verwendet wird. Da sich physikalische Zufallszahlen einer grundsätzlich anderen Methodik bedienen, laufen diese RNGs eher weniger Gefahr voneinander abhängige Zahlen zu generieren.

6.2.1.2.4. EMPIRISCHE TESTS

Die hier vorgestellten PRNGs empirischen Tests zu unterziehen, ist deshalb impraktikabel, da diese Tests lange Sequenzen von bis zu einer Millionen Zufallszahlen benötigen, welche zu generieren wiederum kompliziert ist. Es darf hierbei ohnehin angenommen werden, dass diese Sequenzen Tests, wie etwa die der Diehard-Sammlung, bestünden, da sich empirische Tests hauptsächlich am Verhalten „echter“ RNGs orientieren.

6.2.2. ON-CHIP ODER HARDWARE GENERATOREN

```
10111010010100101110100111101001000001000111010110100101101
10010100100011001000100010010111101000111010010000110001000 1
10111000001111111000000000011110010101000101000001111010110 1
11011110000000011001001100111101000110001100010100001010111 1
00100011110001100001111001001101011001000110101001101110000 0
10010001000100100010100011101100001000010111001011010000011 0
01101000010101100010110011010010011001010110101100110111101 1
10100101101000101011000110001100111000011001001101111001010 0
```

Sequenz 16: 480 durch On-chip-Generatoren erzeugte Zufallsbits

6.2.2.1. METHODE

Für Hardware RNGs wird eine Sequenz von Zufallszahlen durch ein mikroskopisches Zufallsexperiment, das auch innerhalb einer elektronischen Schaltung durchgeführt werden kann, erzeugt.

6.2.2.1.1. THERMISCHES RAUSCHEN

Der On-Chip Random Number Generator nach Kinniment und Chester macht sich das interne Rauschen in elektronischen Schaltungen zu Nutze, um daraus Zufallszahlen zu erzeugen.[207] Für gewöhnlich wird dazu ein thermisches Rauschen verwendet, das schließlich weiterverarbeitet wird, um den störender Effekt äußerer Einflüsse zu minimieren.[208] Der Begriff „Rauschen“ bezeichnet im allgemeinen Gebrauch einen unerwünschten Schall.[209] Analog dazu beschreibt er in der Physik eine Störgröße, die als Überlagerung vieler Schwingungen oder Wellen mit unterschiedlicher Frequenz

[206] Vgl. (Peterson, 1998) Seite 4
[207] Vgl. (Kinniment & Chester, 2002) Seite 595 ff
[208] Vgl. (Jun & Kocher)
[209] Vgl. (Rechner-Rauschen: sengpielaudio.com, 2013)

und Amplitude beziehungsweise Wellenlänge gesehen werden kann.[210] Im Speziellen wird thermisches Rauschen durch unkontrollierte thermische Bewegung der Elektronen in einem Leiter erzeugt.[211]

Da dieses Rauschen zwar einer bestimmen Verteilung folgt, aber von zufälligen Parametern bedingt wird, eignet es sich gut zum Erzeugen von Sequenzen von Zufallszahlen. Dazu wird die Schaltung, die das thermische Rauschen erzeugt, in die Hardware, die daraus Zufallszahlen generiert, integriert.

Abbildung 21: Mögliche Schaltung für einen On-chip Zufallszahlengenerator (hier als USB-Stick realisiert)

6.2.2.1.2. QUANTENMECHANISCHES RAUSCHEN

Wie im Eingangskapitel beschrieben, entziehen sich quantenphysikalische Prozesse großteils den Regeln der deterministischen Physik und eignen sich daher relativ gut, um davon „echte" Zufallszahlen abzuleiten. In diesem Zusammenhang ist, analog zum im letzten Abschnitt beschriebenen thermischen Rauschen, das sogenannte Schrotrauschen von eminenter Bedeutung.

Schrotrauschen ist eine Art elektronischen Rauschens, das durch energietragende Teilchen (wie Elektronen in einem Schaltkreis oder Photonen in einem optischen Bauelement), die statistische Fluktuationen in Messungen bewirken, entsteht.[212] Diese Fluktuationen folgen zufälligen Parametern und können deshalb genutzt werden, daraus Zufallszahlen zu erzeugen.

Die emittierten Photonen sind Poisson-verteilt und je nach Intensität der Quelle erhält man unterschiedliche Mittelwerte für die Fluktuationen, beziehungsweise ein unterschiedliches Mittel an emittierten Photonen.[213] Man kann hier ähnlich der Zufallszahlengewinnungsmethode für radioaktive Zerfallsprozesse (siehe *6.2.4.1 Methode*) Zeiten zwischen zwei Fluktuationshochs messen, um daraus Zufallsbits zu erzeugen oder die zufällige numerische Abweichung von der erwarteten Verteilung nutzen, um eine Sequenz von Zufallszahlen zu erzeugen.

[210] Vgl. (Rauschen: wikipedia.org, 2013)
[211] Vgl. (Rechner-Rauschen: sengpielaudio.com, 2013)
[212] Vgl. (Shot noise: princeton.edu, 2013)
[213] Vgl. (Shot noise: princeton.edu, 2013)

6.2.2.2. GÜTEKRITERIEN

6.2.2.2.1. PRAKTISCHE ANFORDERUNGEN

Die Ergebnisse dieses Generators sind nicht wiederholbar, allerdings lassen sie sich im Gegensatz zu anderen physikalischen RNGs mühelos realisieren und oft vergleichsweise schnell umsetzen.

Dennoch gibt es ein beschränktes Maß an Bits, die pro Sekunde erzeugt werden können, daher werden die durch Hardware RNGs erzeugten Zahlen meist nur als „seeds“ für schnellere PRNGs verwendet.[214]

Hingegen kann das Ergebnis leicht manipuliert werden, indem etwa die Verteilung des Rauschens durch elektromagnetische Interferenz verändert wird. Aus diesem Grund sollte der Output eines On-Chip-Generators regelmäßig überprüft werden.[215]

6.2.2.2.2. PERIODENLÄNGE

Wie alle anderen Zufallszahlengeneratoren haben auch Hardware RNGs eine unendlich lange Periode.

6.2.2.2.3. VERTEILUNG UND UNABHÄNGIGKEIT

Die Wahrscheinlichkeiten für die jeweiligen Ausprägungen des thermischen Rauschens folgen einer Normalverteilung[216], genauer gesagt einer kontinuierlichen Normalverteilung. Der χ^2-Test könnte hier zwar angewandt werden, wäre aber nur eine Approximation, daher scheint der Kolmogorow-Smirnow-Test in diesem Fall geeigneter.

6.2.2.2.4. EMPIRISCHE TESTS

Da die Verteilung der Zufallsbits leicht zu manipulieren ist, sollte sie regelmäßig überprüft werden. Liegt keine physikalische Manipulation vor, so besteht die erzeugte Sequenz die üblichen empirischen Tests.

6.2.3. HINTERGRUNDRAUSCHEN DER ATMOSPHÄRE

18, 13, 78, 96, 10, 43, 58, 2, 64, 11, 27, 44, 24, 93, 25, 41, 66, 94, 77, 26,
99, 68, 49, 47, 8, 70, 67, 75, 21, 42, 80, 38, 86, 9, 17, 39, 15, 79, 89, 54,
30, 84, 53, 61, 4, 71, 69, 95, 32, 73, 97, 7, 50, 65, 20, 55, 22, 85, 33, 82,
48, 76, 28, 83, 1, 56, 3, 46, 60, 31, 92, 88, 29, 23, 5, 14, 19, 90, 16, 51,
81, 34, 98, 57, 6, 91, 74, 52, 72, 36, 45, 100, 87, 35, 63, 59, 37, 62, 40, 12

Sequenz 17: 100 durch atmosphärisches Hintergrundrauschen entstandene Zufallszahlen (n= 100 m=100 p(x)=$\frac{1}{100}$, also gleichverteilt)

6.2.3.1. METHODE

Nicht nur elektronische Elemente oder quantenphysikalische Zustände können „rauschen“, auch das zufällige Rauschen der Radiowellen in der Erdatmosphäre kann dazu verwendet werden Zufallszahlen

[214] Vgl. (Hardware Random number generator: wikipedia.org, 2013)
[215] Vgl. (Kinniment & Chester, 2002) Seite 595
[216] Vgl. (Kinniment & Chester, 2002) Seite 595

zu erzeugen.[217] Dafür werden etwa kleine (zufällige) Veränderungen in der Amplitude des atmosphärischen Rauschens genutzt.[218] Ein ähnliches Phänomen wie atmosphärisches Rauschen ist auch aus dem Alltag bekannt: Das Rauschen, das auf einem Fernsehbildschirm (ABBILDUNG 22), ohne bestimmtes Signal als Input, erscheint, ist zu einem sehr kleinen Prozentsatz Resultat einer Art „Störung“ durch Mikrowellen aus dem Weltall (die sogenannte Hintergrundstrahlung). Hauptsächlich

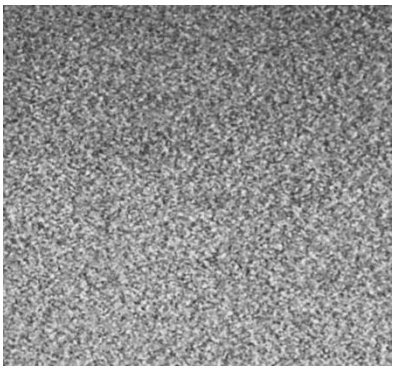

Abbildung 22: Ein rauschender Fernsehbildschirm ist eine Visualisierung des (atmosphärischen und thermischen) Phänomens Rauschen

besteht dieser Effekt jedoch aufgrund thermischen Rauschens innerhalb der Schaltung des Fernsehers selbst.[219]

Das atmosphärische Rauschen ist zwar chaotisch, aber gleichermaßen deterministisch, das heißt, dass, obwohl das System sehr komplex ist, theoretisch voraussehend berechnet werden könnte, welche Amplituden die Frequenz liefert, auch wenn dazu die Geschwindigkeit und Position jedes einzelnen Moleküls in der Atmosphäre bekannt sein müsste. Auf diesen Umstand stützen Kritiker dieser Methode ihren Einwand, durch atmosphärisches Rauschen erzeugte Zufallszahlen seien nicht tatsächlich zufällig.[220]

Abbildung 23: Früher Aufbau des Zufallszahlengenerators der Website RANDOM.org. Zu sehen sind das Radio, als Empfänger der Wellen, links davon ein Sun SPARCstation-Arbeitsplatzrechner mit dem Betriebssystem Solaris und ganz links im Bild eine 500 MB-Festplatte zum Speichern der erzeugten Bits.

[217] Vgl. (Mordasini & Klahr, 2013)
[218] Vgl. (Haahr, 2013)
[219] Vgl. (Rothermel, 2002)
[220] Vgl. (Haahr, 2013)

Eine sowohl für pädagogische als auch für professionelle Zwecke nutzbare Quelle von durch atmosphärisches Rauschen erzeugten Zufallszahlen ist die Website RANDOM.org, die 1997 als einfaches Projekt vierer Studenten begann und heute, nach zahlreichen Umgestaltungen, als einer der vertrauensvollsten Online-Anbieter „echter" Zufallszahlen gilt. Als Empfänger für die Radiowellen wird ein einfaches Radiogerät (siehe *Abbildung 23*) verwendet.[221]

6.2.3.2. Gütekriterien

6.2.3.2.1. Praktische Anforderungen

Wie schon Abbildung 23 suggeriert, erfordert die Implementierung dieser Methode einen komplexen Aufbau und ist daher sehr impraktikabel. Besonders wenn die verwendeten Hardwareelemente nicht dem neuesten Stand der Leistungsfähigkeit entsprechen, ergeben sich äußerst lange Berechnungszeiten für die Zufallszahlen. Die Sequenzen sind nicht wiederholbar und die Generatoren nur dann tragbar, wenn man Radiowellenempfangsgerät und entsprechende Software immer bei sich hat.

6.2.3.2.2. Periodenlänge

Durch diese Methode lassen sich endlos lange Perioden von sich nicht wiederholenden Zufallszahlen erzeugen.

6.2.3.2.3. Verteilung und Unabhängigkeit

Die einzelnen Elemente dieser Zufallszahlentests sind unabhängig voneinander, eine Eigenschaft die wegen der Kontinuität der Wahrscheinlichkeitsfunktion am bestem mit dem Kolmogorow-Smirnow-Test nachgewiesen (und mittels des χ^2–Tests approximiert) werden kann.

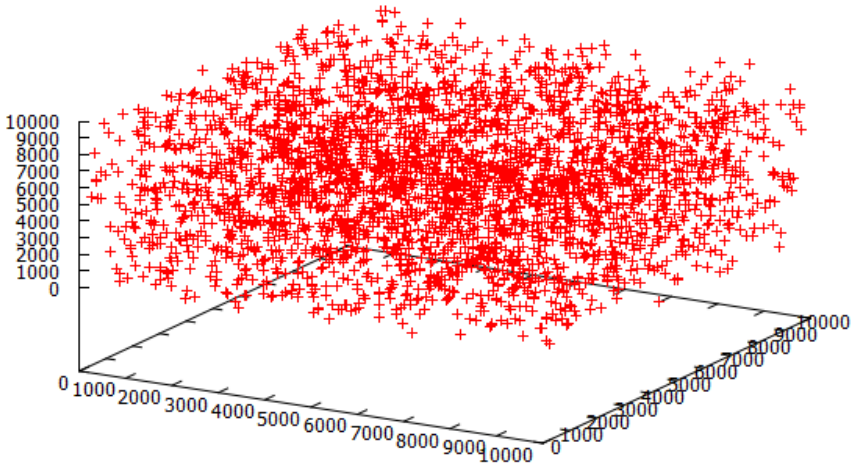

Abbildung 24: Das Plotten von 3-Tupel in den Raum zeigt, dass die durch atmosphärisches Rauschen generierten Zufallszahlen voneinander unabhängig sind. n=9999

[221] (History: random.org, 2013)

6.2.3.2.4. EMPIRISCHE TESTS

Wie die meisten anderen physikalischen RNGs besteht auch der Generator, der sich dieser Methodik bedient, die gängigen empirischen Tests.

6.2.4. RADIOAKTIVE ZERFALLSPROZESSE

255, 152, 201, 61, 129, 145, 222, 73, 30, 54, 19, 218, 60, 169, 240, 236, 46, 55, 193, 39, 73, 195, 252, 40, 192, 127, 247, 132, 53, 23, 208, 132, 48, 240, 32, 171, 40, 159, 89, 209, 126, 66, 185, 141, 238, 21, 248, 91, 254, 51, 249, 26, 89, 192, 142, 195, 85, 230, 17, 25, 148, 83, 72, 73, 30, 91, 157, 118, 73, 28, 199, 209, 86, 10, 211, 32, 179, 238, 161, 89, 166, 162, 141, 100, 19,135, 227, 241, 48, 210, 252, 75, 244, 34, 241, 44, 215, 234, 238, 164

Sequenz 18: 100 durch radioaktive Zerfallsprozesse erzeugte Zufallszahlen

Anders als beispielsweise ein Münzwurf, der streng genommen deterministischen Charakteristika folgt, ist ihre Unberechenbarkeit ein fester Bestandteil der Quantenphysik. Während die klassische Mechanik immer denselben Verlauf eines Systems mit selben Ausgangszustand beschreibt, können in der Quantenphysik idente Anfangsstadien zu unterschiedlichen Ergebnissen führen. [222] Es reicht also nicht den Anfangszustand eines Systems exakt zu kennen, um künftige Ereignisse zu berechnen.

In Bezug auf radioaktive Elemente bedeutet das Folgendes: Wird angenommen, es stünden eine Million radioaktiver Atomkerne zur Verfügung, so kann mithilfe der Quantenphysik die sogenannte Halbwertszeit, also die Zeit, die es braucht, bis die Hälfte der Atomkerne zerfallen ist, ermittelt werden. Diese gibt allerdings keine Auskunft über die genaue Zerfallszeit jedes einzelnen Kernes und ist daher nur ein statistisches Mittel, geschaffen für eine große Anzahl an Atomen.[223] Die genauen Vorgänge in jedem einzelnen Atomkern sind unbestimmt und passieren zufällig. Der kanadische Wissenschaftsjournalist Ivars Peterson drückte diesen Sachverhalt derart aus:

> *„The theory of quantum mechanics […] posits that the instant at which an atom decays occurs by chance. […] There's no warning sign or yellow light signalling the approaching event"*[224]
>
> *"Die Theorie der Quantenmechanik […] postuliert, dass der Augenblick des atomaren Zerfalls zufällig stattfindet. […] Es gibt kein Warnsignal oder ein gelbes Licht, welches das nahende Ereignis ankündigt.*

Wenn instabile Isotope (Isotope sind Atome, deren Neutronenzahl höher ist als die der Protonen. Daher sind diese zumeist instabil und radioaktiv) spontan zu stabileren zerfallen, so strahlen sie ein Wirkungsteilchen[225], im Falle von etwa Uran ein Alphateilchen, aus, das mit einem Geiger-Müller-

[222] Vgl. (Al-Khalili, 1962) Seite 59
[223] Vgl. (Al-Khalili, 1962) Seite 59
[224] (Peterson, 1998) Seite 181
[225] Vgl. (Breider, 1995) Seite 43

Zählrohr gemessen werden kann.[226] Durch die auf diese Weise durchgeführte Messung des Zufallsexperimentes lassen sich Zufallszahlen generieren.

6.2.4.1. METHODE

Es gibt, wie auch für die anderen hier vorgestellten Zufallsexperimente, mehrere Methoden aus ihnen Sequenzen von Zufallszahlen zu generieren.

Eines dieser Verfahren beinhaltet eine atomare Quelle von Radioaktivität, wie etwa das instabile Element Strontium 90, das zufällig Elektronen emittiert, während sich Neutronen in Protonen verwandeln. Für ein bestimmtes Zeitintervall kann etwa durch einen Geiger-Zähler die Anzahl der durch diese Weise emittierten Wirkungsteilchen gemessen werden. Da die Anzahl an abgestrahlten Elektronen einer Normalverteilung folgt, kann die numerische Abweichung einer beliebigen Messung von dem Idealwert bestimmt und davon eine Zufallszahl abgeleitet werden. Für jedes Mal, dass der gemessene Wert genau dem erwarteten entspricht, die Abweichung also 0 ist, erhält man auch die Zufallszahl 0. Diesen Wert zu erhalten ist zwar allgemein am wahrscheinlichsten, dennoch ist er zufällig.[227]

Eine andere Methode sieht hingegen vor, die Zeit zwischen radioaktiven Zerfällen zu messen, um die Zeiten zweier unmittelbar aufeinanderfolgender Messungen zu vergleichen. Sollte das erste gemessene Zeitintervall länger sein als das zweite, so nimmt das resultierende Zufallsbit den Wert 0 an, während eine 1 generiert wird, wenn das zweite Intervall länger ist.[228] Auf diese Weise wird eine Sequenz von Nullen und Einsen erzeugt.

6.2.4.2. GÜTEKRITERIEN

6.2.4.2.1. PRAKTISCHE ANFORDERUNGEN

Die Resultate dieser Zufallsexperimente sind nicht wiederholbar, insofern eignen sie sich nicht für viele praktische Anwendungsgebiete, wie diverse Simulationen. Zudem ist die Implementierung dieser Methode äußerst komplex und steht auch nur einem begrenzten Teil der Bevölkerung zur Verfügung, weshalb sie die Anforderung „Tragbarkeit“ nur in ungenügendem Maße erfüllt. Radioaktive Zerfallsprozesse zu messen ist daher eine sehr impraktikable Vorgehensweise, doch in heiklen Fällen, die „echte“ Zufälligkeit erfordern, erweist sich diese Methode, zumindest nach dem heutigen Stand der Wissenschaft, als die adäquateste.

6.2.4.2.2. PERIODENLÄNGE

Durch diese Methode physikalischer Zufallszahlengewinnung ist die Erzeugung endlos langer Perioden von sich nicht wiederholenden Zufallszahlen möglich.

[226] Vgl. (Peterson, 1998) Seite 181
[227] Vgl. (Clewett & Numberphile, 2013)
[228] Vgl. (Peterson, 1998) Seite 181

6.2.4.2.3. VERTEILUNG UND UNABHÄNGIGKEIT

Die Experimente, die schließlich zu den Zufallszahlen beziehungsweise –bits führen, folgen „echt“ zufälligen Parametern. Dies gilt jedoch nur für die Experimente selbst (in diesem Fall für die Anzahl an emittierten Elektronen oder die verstrichene Zeit zwischen zwei Zerfällen) und nicht zwangsläufig für die Zählungen, denn diese können durch die Messinstrumente verfälscht werden:
Beispielsweise könnte die Zeit, die benötigt wird, um den Zähler zurückzusetzen, von der vorherigen Messung abhängen, weswegen die darauf folgende Messung geringe Abhängigkeiten von der vorherigen aufweisen wird[229], obwohl die verstrichenen Zeiten selbst unabhängig voneinander sind.

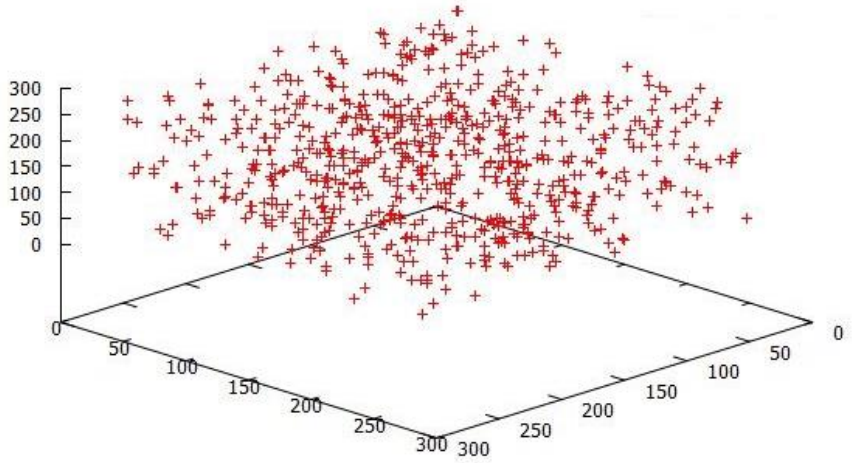

Abbildung 25: Das Plotten von 3-Tupel in den Raum zeigt, dass die durch radioaktive Zerfälle generierten Zufallszahlen (hier nach der Methode der aufeinanderfolgenden Zeitintervalle) voneinander unabhängig sind. n=2046

6.2.4.2.4. EMPIRISCHE TESTS

Eine Vielzahl empirischer Tests richtet sich nach den Eigenschaften, die von physikalischen RNGs wie diesem vorgegeben werden. Daher scheint es nur logisch, dass physikalische RNGs diese Gütetests, die ja praktisch auf sie zugeschnitten sind, bestehen.

6.2.5. ANDERE PHYSIKALISCHE METHODEN

6.2.5.1. LAVALAMPEN

1996 kamen drei Arbeiter an der Computerfirma Silicon Graphics auf die Idee, Lavalampen als Zufallszahlengeneratoren zu zweckentfremden und meldeten das Patent ihres Produkts mit dem Namen *Lavarand* an.[230] Eine Lavalampe ist ein chaotisches System[231], das heißt sie ist zwar sehr komplex, aber dennoch deterministisch. Aus diesem Grund, der praktischen Unvorhersehbarkeit der Lampen, eignen sie sich ebenfalls zur Erzeugung von Zufallszahlen.

[229] Vgl. (Peterson, 1998) Seite 181
[230] Vgl. (McNichol, 2014)
[231] Vgl. (everything2.com, 2014) im Original von (lavarand.org, 2000)

Abbildung 26: Lavalampen

Um die Zufallszahlen erzeugen zu können, wird zuerst eine Kamera, die in regelmäßigen Abständen Fotos schießt, vor sechs Lavalampen platziert. Ein spezieller Algorithmus verwandelt daraufhin die digitalen Informationen der Bilder in Zufallszahlen, die dann als „seed" für einen PRNG dienen. [232]

6.3. ZUSAMMENFASSUNG

Tabelle 21 liefert einen Überblick über die im Vorhergehenden vorgestellten Zufallszahlengeneratoren samt ihrer unterschiedlichen Gütemerkmale.

Name	Formel/Methode	Vorteile	Nachteile
Deterministische RNGs			
Dezimalentwicklungen irrationaler Zahlen	Nachkommastellen einer irrationalen Zahl (bei π entstünde dadurch die Sequenz (1,4,1,5,9,2,…))	• gute Verteilung und Unabhängigkeit (falls normal) • unendlich lange Sequenz (falls normal)	• leicht vorhersehbar (nicht für die Kryptografie geeignet) • Berechnung (ab einer gewissen Stelle) sehr komplex
Mittelquadratmethode	Eine Zufallszahl mit n Stellen wird quadriert. Die n mittleren Stellen fungieren als die nächste Zufallszahl der Sequenz	• Sehr simpel • Historisch interessant	• Sehr kurze Sequenzen • „abstürzen" auf niedrige Werte • Schlechte Verteilung • Hohe sequentielle Korrelation
Der multiplikative lineare Kongruenzgenerator	$I_{j+1} = aI_j \ (mod\ m)$	• Minimale Größe im Cache • Simpel, einfach und schnell • Relativ zuverlässig bei guten Parametern	• Hohe sequentielle Korrelation (→ Hyperebenen) • Noch kürzere Perioden als beim gemischt-linearen KG • Passieren viele empirische Tests nicht
Der gemischt-lineare Kongruenzgenerator	$I_{j+1} = aI_j + c \ (mod\ m)$	• Minimale Größe im Cache • Simpel, einfach und schnell • Relativ zuverlässig bei guten Parametern	• Hohe sequentielle Korrelation (→ Hyperebenen) • Bei schlechten Parametern: kurze Perioden, schlechte Verteilungen • Passieren viele empirische Tests nicht

[232] Vgl. (everything2.com, 2014) im Original von (lavarand.org, 2000)

Der additive RNG	$(a_1 I_{j-1} + \cdots + a_p I_{j-p}) (mod\ m) = Ij$	• Schnell und einfach (besonders Fibonacci) • Lange Perioden von bis zu m^2-1 • Relativ zuverlässig bei guten Parametern (verzögerter Fibonacci)	• Bei schlechten Parametern (Fibonacci) deutliche Hyperebnen • Schlechtes Verhalten in empirischen Tests
Der inverse Kongruenzgenerator	$I_j = (a\overline{I_{j-1}} + c)(mod\ m)$	• Bei guten Parametern Sequenzlänge m • Keine Hyperebenen	• Langsamer als ein LCG • Schlechtes Verhalten in empirischen Tests
Marsaglia-Zaman-Generatoren	$(I_{j-r} + I_{j-s} + carry) (mod\ m) = I_j$ Oder $(I_{j-r} - I_{j-s} - carry) (mod\ m) = I_j$	• Relativ gutes Verhalten in empirischen Tests • Schnell und einfach • Lange Perioden	• Eventuelle Warmlaufphasen
Multiply with carry	$I_j = (aI_{j-1} + carry_{j-1}) (mod\ m)$	• Sehr lange Periode • Geringe sequentielle Korrelation • Schnell (nach Warmlaufphase) • Gutes Verhalten in empirischen Tests	• Bei schlechten seeds lange Warmlaufphase • braucht physikalische RNGs für seed (kompliziert)
Der Mersenne-Twister	$I_j \coloneqq I_{j-227} \oplus \left\lfloor \frac{h}{2} \right\rfloor \oplus ((h\ mod\ 2) \times 9908b0df_{hex}$	• Gutes Verhalten in empirischen Tests • Keine allzu große Größe im Cache • Extrem lange Periode (2^{19937}-1) • Niedrige sequentielle Korrelation	• Bei schlechten seeds lange Warmlaufphase (kürzer als beim MWC) • braucht physikalische RNGs für seed (kompliziert)
Physikalische RNGs			
Der Münzwurf	Münzwurf, Würfel, Lottozahlenziehung, Rouletterad etc.	• theoretisch: niedrige sequentielle Korrelation und perfekte Verteilung • unendlich lange Perioden • nicht vorhersehbar (chaotisch)	• praktisch: kleine Produktionsfehler im Würfel (oder Münze etc.) verfälschen die Verteilung & Unabhängigkeit • nicht wiederholbar • sehr aufwändig
Hardware Generatoren	Verschiedenartiges Rauschen in Schaltungen wird von diesen als Quelle von Zufallszahlen verwendet	• schnell im Vergleich zu anderen physikalischen RNGs • nicht vorhersehbar (chaotisch) • (bei regelmäßiger Überprüfung) sehr zuverlässig • Unendlich lange Perioden	• Zufallsbits folgen der Normalverteilung • Sequenzen müssen regelmäßig überprüft werden • Anfällig auf äußere Störungen • Nicht wiederholbar
Hintergrundrauschen der Atmosphäre	Zufälliges Rauschen in der Atmosphäre wird als Quelle von Zufallszahlen angesehen	• nicht vorhersehbar (chaotisch) • sehr zuverlässig • Unendlich lange Perioden	• Sehr langsam • Sehr aufwändig
Radioaktive Zerfallsprozesse	Zufällige Zerfälle in Atomkernen werden als Quelle von Zufallszahlen angesehen	• nicht vorhersehbar (vielleicht sogar völlig indeterministisch) • sehr zuverlässig • Unendlich lange Perioden	• Sehr langsam • Sehr aufwändig

Tabelle 21: Zusammenfassung der in dieser Arbeit beschriebenen PRNGs und RNGs

7. ANWENDUNGEN DER ZUFALLSZAHLENGENERATOREN

Die letzten Kapitel umfassen ausführliche Darstellungen der Gütekriterien und Methoden deterministischer sowie physikalischer Zufallszahlengeneratoren. Die Beschäftigung mit dieser schier unerschöpflichen Thematik wirft letzten Endes die unerlässliche Frage auf: Wozu der Aufwand?
Tatsächlich kennen Zufallszahlen zahlreiche Anwendungsmöglichkeiten in den unterschiedlichsten Bereichen, von denen an dieser Stelle einige beschrieben werden.

7.1. IN DER UNTERHALTUNG

Sowohl deterministische als auch physikalische Zufallszahlen werden im Glücksspiel und Casinos eingesetzt.
Die Ermittlung der Lottozahlen erfolgt zumeist durch das randomisierte Auswählen von Lottokugeln. Diese Methode ist physikalisch, da kein Algorithmus eingesetzt wird, um die Zufallszahlen zu ermitteln. Dennoch ist sie nicht tatsächlich zufällig, sondern nur chaotisch. Durch diese Methodik kann das Resultat weder von den Lottobetreiber/inne/n manipuliert, noch von den Spieler/inne/n im Vorhinein durch Analyse eines Algorithmus vorausberechnet werden.
Diese Technik hat allerdings ihre Tücken; Im Zuge einer Lottozahlenziehung des deutschen Senders ZDF im April 2013 passierte eine Panne, die die gezogenen Zahlen ungültig machte. Zwei Kugeln blieben am sogenannten Schlitten (Abbildung 27) hängen, sodass nur 47 anstatt der vorgesehenen 49 Kugeln in die Trommel fielen. Die Ziehung war daher nicht „6 aus 49“, sondern „6 aus 47“, nicht alle Zahlen hätten mit derselben Wahrscheinlichkeit gezogen werden können. Die Ziehung musste entsprechend wiederholt werden.[233] Es kann dagegen argumentiert werden, dass derartige Fehler

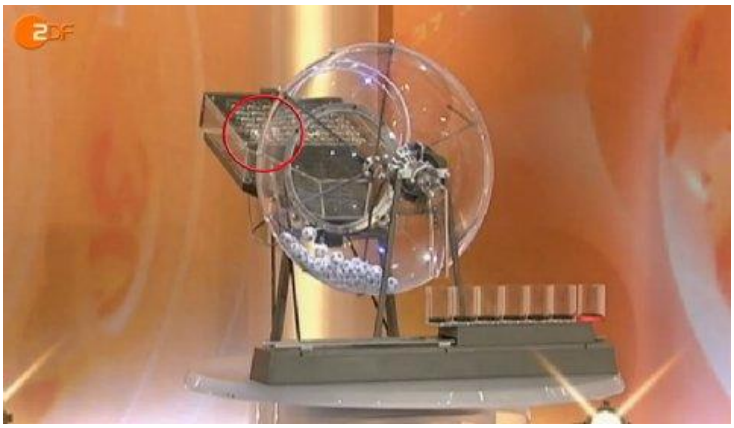

Abbildung 27: Eine sogenannte Trommel für die Lottozahlenziehung. Dahinter ist der Schlitten, mit der Stelle an der die Kugeln hängenblieben (rot) zu sehen

nur äußerst selten passieren, ein Umstand der die Verwendung dieser physikalischen Methode, im Vergleich zu deterministischen RNGs durchaus rechtfertigt.

[233] Vgl. (Lotto am Mittwoch: sueddeutsche, 2013)

Generell gehört das Rollen der Lottokugeln per se zum Zelebrieren des Glücksspiels und kontribuiert zu seinem Unterhaltungswert. Aus diesem Grund fällt die lange Generierungszeit und die Impraktikabilität der Methode kaum ins Gewicht und erhöht im Gegenteil die Spannung auf das Ergebnis noch mehr.

Das Rouletterad und das Würfelspiel sind gleichermaßen Methoden, deren Zuverlässigkeit sich Glücksspielinstitutionen bedienen. Diese können allerdings mitunter vom Croupier zum Vorteil des Casinos verfälscht werden, wie bereits unter *6.2.1. Der Münzwurf, Würfel und Lottozahlen* erörtert wurde.

Glücksspielautomaten oder sogenannte „einarmige Banditen" waren bis vor wenigen Dekaden aufgrund etlicher Zahnräder, Federn, Hebel und Gewichte, rein mechanischer Natur. Die Automaten waren allerdings besonders bei schlechter Kalibrierung sehr anfällig für Defekte und mechanische Manipulation (ein/e Spieler/in konnte etwa die Gewinnwahrscheinlichkeit durch das Ziehen des Griffs im richtigen Moment erhöhen).[234] In modernen Glücksspielsautomaten, die ausschließlich auf Algorithmen, also deterministischen RNGs, basieren, sind die Wahrscheinlichkeiten pro mögliche Symbolkombination nicht ident, wie man für PRNGs mit gleichverteilten Output annehmen könnte. Casinos beeinflussen die Wahrscheinlichkeiten in ihrem Interesse, sodass der Jackpot weit weniger wahrscheinlich ist als jede andere Kombination.[235]

Allerdings sind auch diese Generatoren mitunter nicht vor Manipulation gefeit: Ein besonders aufmerksamer Spieler erkannte, dass ein gewisses Casino ihre Glücksspielautomaten über Nacht zurückstellte, sodass sie ihre Sequenzen jeden Tag wiederholten. Mit diesem Wissen gelang er zu einem Gewinn von umgerechnet etwa 450 000€.[236]

7.2. In der Kryptographie

Die Kryptographie, also die Verschlüsselung von Nachrichten, hat sich besonders im Zusammenhang mit der Informationstechnologie und dem wachsenden Spionageproblem zu einem die Computerwissenschaften dominierenden Themengebiet entwickelt.

Es gibt unterschiedliche Methoden einen Klartext b mithilfe von Zufallszahlen in einen Chiffriertext c umzuwandeln. Der sogenannte One Time Pad (OTP), ein bereits 1917 von G. Vernam entwickeltes Chiffriersystem, beispielsweise verschlüsselt den Klartext mithilfe eines Zufallsbits r nach der Methode $b \rightarrow c = b \oplus r$.[237] Der/Die Empfänger/in der verschlüsselten Nachricht kann diese, soweit er/sie die verwendete Sequenz von Zufallszahlen kennt, unter Anwendung der umgekehrten

[234] Vgl. (Peterson, 1998) Seite 168
[235] Vgl. (Peterson, 1998) Seite 169
[236] Vgl. (Peterson, 1998) Seite 183f
[237] Vgl. (Meier, 2009)

Verschlüsselungsmethode wieder dechiffrieren. Folgendes Beispiel soll diese Methode veranschaulichen[238]:

Eine beliebige, zu sendende Nachricht wird zunächst in eine Sequenz S_1 von binären Zahlen umgeschrieben. Anschließend wird eine ebenso lange Sequenz S_2 von binären Zufallszahlen generiert und nach den Regeln der binären modularen Arithmetik zur bereits bestehenden S_1 addiert. Diese Regeln lassen sich in ihrer einfachsten Form wie folgt zusammenfassen:

0+0=0 1+1=0
0+1=1 1+0=1

Die gesamte binäre Nachricht wird auf diese Weise umgewandelt und dem/der Empfänger/in gesendet:

Nachricht	**Z**	**U**	**F**	**A**	**L**	**L**
In binären Zahlen (S_1)	01011010	01010101	01000110	01000001	01001100	01001100
binäre Zufallszahlen (S_2)	00010111	01111111	10101011	10101111	00110010	01001011
Verschlüsselte Nachricht ($S_1+S_2=S_3$)	01001101	00101010	11101101	11101110	01111110	00000111
Entschlüsselte Nachricht ($S_3-S_2=S_1$)	01011010	01010101	01000110	01000001	01001100	01001100

Tabelle 22: Ver- und Entschlüsselung einer Nachricht

Jede Methode hat allerdings ihre Schwachstellen, denn ihre Zuverlässigkeit ist vom verwendeten Verschlüsselungsalgorithmus und Zufallszahlengenerator, wie auch von anderen Parametern abhängig.

7.3. In der Wissenschaft

Bis Mitte des 20. Jahrhunderts fehlte es den Naturwissenschaften, insbesondere der Physik an einer Methode komplexe Zustände, deren Ausgangszustände nicht exakt bekannt sind, zu simulieren.

Daraufhin schlugen die beiden amerikanischen Mathematiker Stanislaw Ulam und John von Neumann 1947 eine spezielle Methode vor, um diesem Problem entgegenzuwirken, indem die Wahrscheinlichkeitslehre in physikalische Berechnungen integriert wird. Diese Technik, derer statistische Stichprobenentnahmen und Zufallszahlen zugrunde liegen, wird Monte-Carlo-Methode genannt. Der Name rührt von der Gewohnheit eines Onkels von Stanislaw Ulam her, sich Geld von seinen Angehörigen zu borgen, um es im monegassischen Casino Monte Carlo am Rouletterad zu verspielen.[239]

[238] Vgl. Original beschrieben in: (Peterson, 1998) Seite 205
[239] Vgl. (Peterson, 1998) Seite 171f

Die Monte-Carlo-Methode zur Simulation von Zufallsprozessen wurde zunächst für den Bau der Atombombe eingesetzt, um die Wechselwirkungen zwischen Neutronen und Materie vorherzusagen und lässt sich heute auf zahlreiche naturwissenschaftliche, technische und medizinische Probleme anwenden.[240]

Die Physik hat hierbei ihren alleinigen Anspruch auf diese Simulationsmethode längst verloren: Sie wird von Fachleuten in Branchen wie zum Beispiel Finanz- und Projektmanagement, Energiewirtschaft, Forschung und Entwicklung, Versicherung, Öl- und Gaswirtschaft, Transport und Umwelttechnik zur Analyse diversester Situationen eingesetzt. [241]

Es lässt sich schon aufgrund der zahlreichen Anwendungsgebiete konkludieren, dass es zahlreiche differenzierte Ausformungen und entsprechend keine einheitliche Methode für derartige Simulationen gibt, allerdings lässt sich ein gemeinsamer Nenner aller Techniken bilden:

1. Zunächst wird das Gebiet möglicher Inputs definiert
2. Daraufhin werden nach dem Zufallsprinzip Inputs aus einer Wahrscheinlichkeitsverteilung generiert
3. Die jeweiligen Inputs werden deterministischen Berechnungen unterzogen
4. Schließlich werden die Ergebnisse zusammengefasst[242]

Ein populäres Exempel für die Anwendungsmöglichkeiten der Monte-Carlo-Methode ist die relativ simple Annäherung der Kreiszahl π. Die dieser Approximation zugrundeliegende Vorgehensweise sieht als Basis einen in einem Einheitsquadrat liegenden Viertelkreis vor (siehe *Abbildung 29*).

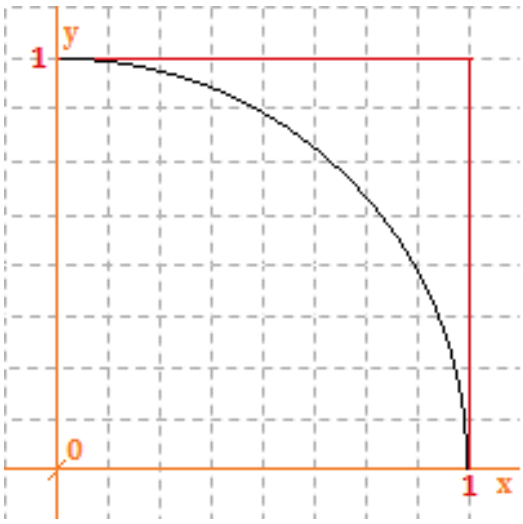

Abbildung 29

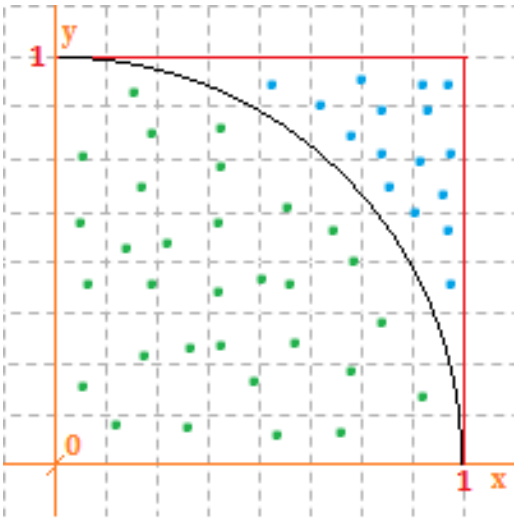

Abbildung 28

[240] Vgl. (Monte-Carlo-Simulation: exp.univie.ac.at, 2014)
[241] Vgl. (Monte Carlo-Simulation: Palisade, 2014)
[242] Vgl. (Monte-Carlo-Simulation: wikipedia.org, 2013)

Ein Zufallszahlengenerator bestimmt daraufhin die Koordinaten der Punkte, die in dieses Einheitsquadrat geplottet werden sollen, also Werte im Intervall [0,1] (siehe *Abbildung 29*). Anschließend muss die Anzahl der Punkte P_0, die im Viertelkreis liegen, sowie diejenige der Punkte P_1, die ins kleinere Feld außerhalb des Kreises geplottet wurden, ermittelt werden. [243]Die finale Approximation der Kreiszahl geschieht schließlich durch die Formel:

$$\pi \approx 4 \frac{P_0}{P_0 + P_1}$$

Formel 23

Werden beispielsweise durch die *Zufallszahl()*-Funktion 1000 Zufallszahlen im Tabellenkalkulationsprogramm Excel erzeugt, so mag die Anzahl an im Viertelkreis liegenden Zahlen 785 betragen, was laut Formel 23 einen angenäherten Wert für Pi von *3,136* liefert.

$$\pi \approx 4 \frac{784}{1000} = 3{,}136$$

Bedenkt man, dass die Kreiszahl mit einer Genauigkeit von drei Dezimalen *3,141* lautet, so ist dies eine verhältnismäßig gute Annäherung. Der Mittelwert mehrerer solcher Iterationen ergibt schlussendlich eine genauere Approximation.

Die Monte-Carlo-Methode kann auch dazu genutzt werden, einen Zufallszahlengenerator auf seine Güte zu prüfen. Dazu wird die Genauigkeit, mit der sich die Kreiszahl mithilfe des fraglichen PRNG approximieren lässt, getestet. Mithilfe von 500.000 Byte an Zufallszahlen, generiert durch radioaktiven Zerfall, ließ sich Pi auf diese Weise auf einen Wert von 3,143580574 annähern.[244]

7.4. In der Statistik[245]

Anlässlich des Wahlkampfes zwischen dem Demokraten Franklin D. Roosevelt und dem Republikaner Alf Landon um die US-amerikanische Präsidentschaft im Jahre 1936 wurde die Bevölkerung der Vereinigten Staaten per Telefoninterview befragt, welchem Kandidaten sie ihre Stimme gäben. Die auf den Befragungen gestützten Prognosen sahen einen erdrutschartigen Sieg für Landon vor, tatsächlich gewann aber Roosevelt die Wahl mit einem beträchtlichen Vorsprung. Die falsche Vorhersage rührt daher, dass die Statistiker und Statistikerinnen ihre Stichprobenelemente nicht zufällig aus der Gesamtbevölkerung ausgewählt, sondern ausschließlich die Teile der Gesellschaft, die im Besitz eines Telefons waren, interviewt hatten. Im frühen zwanzigsten Jahrhundert gehörten diese Menschen einer bestimmten finanziellen und sozialen Schicht wohlhabender Bürger und Bürgerinnen, an, die sich erfahrungsgemäß eher mit der republikanischen Partei identifizieren.

[243] Vgl. Methodik beschrieben auf (Wie berechnet man Pi nach der Monte-Carlo-Methode ?, 2014)
[244] Vgl. (Walker, 2008)
[245] Vgl. (Roney-Dougal, du Sautoy, & Gowers, 2011) ab Minute 25

Nicht nur zuverlässige Prognosen um den Ausgang politischer Wahlen, sondern auch aussagekräftige Testungen in der Medizin werden von einer intelligenten Auswahl der Stichproben gewährleistet. Wird ein bestimmtes Medikament beispielsweise nur an Männern mittleren Alters getestet, so haben die gewonnenen Daten keine Aussagekraft über die Auswirkungen des Mittels auf jüngere Menschen oder Frauen.

Es ist daher in der Statistik von eminenter Bedeutung, Stichproben ohne jegliche Struktur, also allein mithilfe von Zufallszahlen, bestimmen zu können.

8. ANHANG

Der Mensch und die Zufallszahlengenerierung

Die Daten, auf die sich die Einleitung dieser Arbeit bezieht, entstammen einer größtenteils an den Gymnasien der Stadt Horn durchgeführten Umfrage. Unter anderen wurden 170 Schülerinnen und Schüler, sowie Personen des Lehrerkörpers gebeten, eine Zufallszahl zwischen 1 und 10 niederzuschreiben. Die wesentlichen Ergebnisse dieser Erhebung sind in Abbildung 30 zusammengefasst.

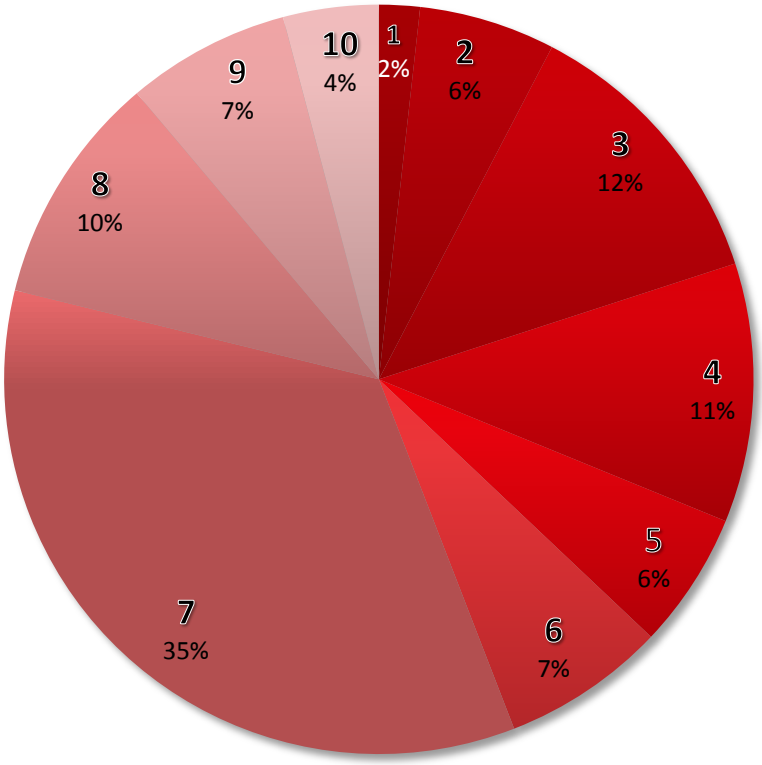

Abbildung 30: Die Zahl 7 ist mit großem Vorsprung die am häufigsten gewählte Zahl, wohingegen die Zahl 1 nur mit zwei Prozent und die Zahl 10 nur mit vier Prozent vertreten ist.

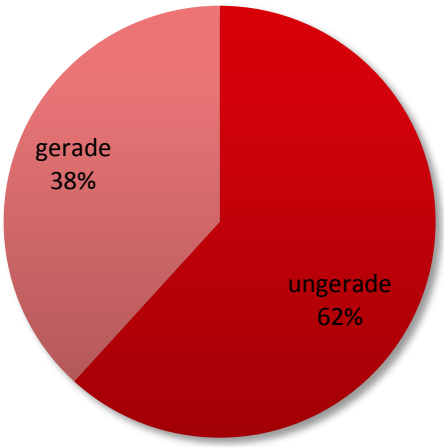

Abbildung 31: Häufigkeit der geraden/ungeraden Zahlen

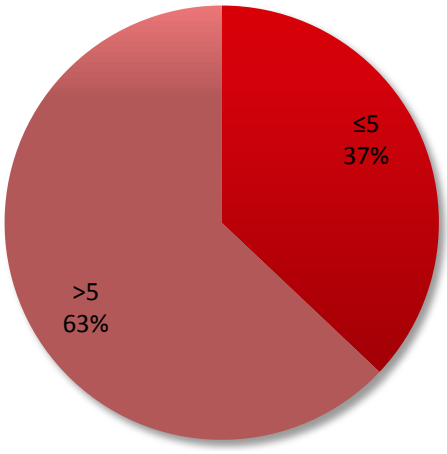

Abbildung 32: Häufigkeit der Zahlen nach Größe

Zwei weitere Grafiken (Abbildung 31 und 32) illustrieren die Tendenzen der Befragten zur Wahl von ungeraden Zahlen und zu Zahlen >5. Wie lassen sich nun diese Präferenzen, die offensichtlich wider einen objektiven Zufall sprechen, begründen?

Das menschliche Gehirn ist, evolutionsgeschichtlich bedingt, darin geübt, bestimmte Muster und Regelmäßigkeiten in seiner Umwelt zu erfassen. Wann immer eine Person derartige Strukturen zu erkennen meint, so ordnet sie sie in bestimmte Kategorien ein, denen folglich keine Zufälligkeit zuzusprechen ist. Paradoxerweise ist es gerade diese Eigenschaft, die es dem Menschen verwehrt (zumindest ohne Hinzuziehung wissenschaftlicher Methoden) wahre Zufälligkeit als solche zu erkennen oder sie gar selbst zu erzeugen. Erfahrungsgemäß ist eine Verteilung oder eine Zahl laut menschlichem Verstand nicht zufällig, wenn sie bestimmte Eigenschaften zu haben scheint. Wahre Zufälle folgen dieser Logik aber nicht: Ob eine Zahl beispielsweise ungerade ist oder nicht, macht in der Wahrscheinlichkeitsrechnung nicht zwingend einen Unterschied.

Hieraus lässt sich schließen, dass ein Großteil der befragten Personen hinter den von ihnen nicht gewählten Zahlen gewisse „nicht zufällige“ Eigenschaften vermuteten:

9. Die Randwerte: Die Randwerte werden am seltensten gewählt. Sie sind schon in der Aufgabenstellung vorgegeben, also determiniert worden und scheinen somit nicht zufällig. Zudem haben sie sehr klare Eigenschaften (der kleinste und der größte mögliche Wert). Übrig bleiben schließlich die Werte

   | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

10. Gerade Zahlen: Gerade Zahlen haben mit ihrer Teilbarkeit durch 2 eine gemeinsame klar definierte Eigenschaft. Zudem wirken sie auf die menschliche Ratio fast schon zu „makellos“ um zufällig zu sein. Ungerade Zahlen scheinen hingegen keinerlei Struktur zu besitzen, sie wirken auf die Psyche „krumm“ und „uneben“, in anderen Worten: zufällig. Übrig bleiben schließlich die Werte:

    | 3 | | 5 | | 7 | | 9 |

11. Die Tendenz zu Primzahlen: Primzahlen, wie 3 und 7, die sich durch keine andere Zahl als durch sich selbst und 1 teilen lassen, wirken deshalb auf viele Menschen zufällig, da sie ein tatsächliches, statistisches Merkmal von Zufälligkeit zu erfüllen scheinen: Die Unabhängigkeit. Während gerade Zahlen von der Zahl 2 abhängig wirken, oder die Zahl 9 von der 3, so sehen Primzahlen unabhängig aus, es lässt sich keine „(Kausal)kette“ zu einer anderen Zahl bilden.
12. 9: Diese Zahl hat eine klar definierte Eigenschaft, sie ist durch 3 teilbar, also keine Primzahl, und liegt zudem sehr nahe am Randwert. Übrig bleiben:

    | 3 | | 5 | | 7 | | |

13. 5: Diese Zahl scheint wenig zufällig, da sie sich, gemeinsam mit der 6, in der Mitte des möglichen Bereiches befindet und somit wieder eine klare Eigenschaft hat. Übrig bleiben die Werte:

| 3 | | | | 7 | | |

14. Zahlen ≤5: 3 ist der zweithäufigste Wert in der Umfrage, dennoch wurde er weniger als halb so oft wie die Zahl 7 gewählt. Das liegt wohl einerseits daran, dass Zahlen über 5 generell attraktiver auf den Menschen wirken als (zu) niedrige Werte und andererseits daran, dass 7 weiter vom Randwert entfernt ist als drei. Übrig bleibt daher in über einem Drittel der Fälle:

| | | | | 7 | | |

9. NACHWORT

Abschließend, bleibt zu sagen, dass die vorhergehenden Seiten sowohl einen Überblick, also auch tiefgehendere Einsicht in die Thematik der Zufallszahlengenerierung liefern, wenngleich diese eine Materie ist, die sicherlich Stoff für eine Vielzahl an zusätzlichen Kapiteln bereitstellen könnte.
Über all den mathematischen Algorithmen und Formeln, informationstechnologischen Gütetests und p-Werten darf nicht vergessen werden, dass Zufallszahlen eine unanfechtbare Relation zur Realität jedes einzelnen Individuums haben. Durch unser Verständnis eines mathematischen Determinismus, der es uns ermöglicht pseudozufällige Sequenzen von Zahlen zu erzeugen, aber auch eines Determinismus, der dem subjektiven und möglicherweise auch dem objektiven Zufall zugrunde liegt, erhält die Wissenschaft erst die Möglichkeit existentielle Fragen zu stellen und zu beantworten.
Aus diesem Grund sind die in dieser Arbeit beschriebenen Themen nicht nur eine Spielerei der Mathematik, sondern auch für zahlreiche anderen Disziplinen, wie auch das Selbstverständnis des Menschen von hoher Signifikanz.

10. ENGLISH ABSTRACT

This paper's underlying idea is to examine the topic of randomness in general and random number generation in particular. In that context, a thorough overview of all common types of random number generators, alongside methods of testing their quality is given. Additionally, also applications of random numbers are discussed.

The reader is confronted with existential or interesting questions like "Is there a free will?" or "Can I really know what is random and what not?" especially in the opening chapter of this paper and will get some scientific insight on those issues.

The following sections are of a more technological and mathematical nature: The random number generators in questions include, amongst others, various linear congruential generators, as they are the most popular kind of deterministic random number generators, the additive RNG, the Multiply-with-carry and the Mersenne Twister. Furthermore, this paper also analyses random number generators based on occurrences in the world of physics and everyday-life, like radiation, thermal noise and even lava-lamps. This paper is finally completed with a discourse on the applications of random numbers.

11. QUELLENVERZEICHNIS

11.1. PRINTMEDIEN

Abramowitz, M., & Stegun, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* Wahington, D.C: Dover Publications.

Aistleitner, C. (2006). *Normale Zahlen.* Technischen Universität Wien.

Al-Khalili, J. (1962). *Quantum: moderne Physik zum Staunen.*

Bert F. Green, J. K. (1959). *Journal of the ACM (Association for Computing Machinery) 6.*

Borel, É. (1913). La mécanique statique et l'irréversibilité. *Journal de Physique*(3, Numéro 1), S. 189-196.

Breider, H. (1995). *Über Zufall und Wahrscheinlichkeit: Sternschnuppen - schwarze Löcher - Seifenblasen.* Frankfurt am Main: Haag und Herchen.

Burns, P. (2004). *Linear, Congruential Random Number Generators.*

Dagpunar, J. (1988). *Principles of Random Variate Generation.* Oxford: Claredon Press.

Eichenauer-Herrmann, J. (Ausgabe 56 Nummer 193. Januar 1991). Inversive congruential pseudorandom numbers avoid the planes. *Mathematics of Computation - American Mathematical Society*, S. 297-301.

Eichenauer-Herrmann, J. (1994). Improved lower bounds for the discrepancy on inversive congruential pseudorandom numbers. *in: Mathematical Computation Volume 62*.

Einstein, A. (4. Dezember 1926). Brief an Max Born oder Niels Bohr.

Fischer, E. P. (2010). *Die Hintertreppezum Quantensprung.* München: F.A. Herbig Verlagsbuchhandlung, München.

Frank, P. (1932). *Das Kausalgesetz und seine Grenzen.* Wien, Österreich: Springer.

Hawking, S. (kein Datum). *Das Universum in der Nussschale.* Deutscher Taschenbuch Verlag.

History: random.org. (30. Dezember 2013). Von random.org: http://www.random.org/history/ abgerufen

Hromkovic, J. (2008). *Sieben Wunder der Informatik.* Deutschland: Springer.

Jun, B., & Kocher, P. (kein Datum). The INTEL Random Number Generator. Cryptography Research Inc.

Kinniment, D., & Chester, E. (24.–26.. September 2002). Design of an On-Chip Random Number Generator using Metastability. Florenz, Italien at the 28th European Solid-State Circuits Conference: Proceedings of the 28th European Solid-State Circuits Conference.

Knuth, D. (1969). *The Art of Computer Programming.* Reading, Mass.: Addison-Wesley.

Kütting, H. (1999). *Elementare Stochastik.* Berlin - Heidelberg: Spektrum Akademischer Verlag.

L'Ecuyer, F. P. (Oktober 2005). On the Xorshift Random Number Generators. *im Rahmen von: ACM Transactions on Modeling and Computer Simulation*. Université de Montréal.

Lehmer, D. (1949). Mathematical methods in large-scale computing units. In *Proceedings of a Second Symposium on Large-Scale Digital Calculating Machinery* (S. 141–146). Harvard University: Harvard University.

Letellier, C. (2010). Chaos unter Kontrolle. *Spektrum der Wissenschaft Spezial - Zufall und Chaos*(1/2010), S. 24-31.

Marsaglia, G. (24. Juni 1968). *Random numbers fall mainly in the planes*. Washington.

Marsaglia, G. (Mai 2003). Random Number Generators. *Journal of Modern Applied Statistical Methods: Vol 2. No 1,2-13*.

Marsaglia, G. (Juli 2003). Xorshift RNGs. *in: Journal of Statistical Software 8 (14)*. Florida State University.

Marsaglia, G., & Zaman, A. (1991). A New Class of Random Number Generators. *in: The Annals of Applied Probability, Vol 1, No. 3*. The Florida State University.

Matsumoto, M., & Nishimura, T. (1998). Mersenne twister. A 623-dimensionally equidistributed uniform pseudorandom number generator. *In: ACM Transactions on Modeling and Computer Simulation. 8*. Keio University/Max-Planck Institut für Mathematik.

Milavec, T. J. (1995). *Zufall, Ordnung und Chaos am Computer.* Universität Wien.

Müller, R. (2006). Dekohärenz (vom Erscheinen der klassischen Welt). Sektion Physik der Ludwig-Maximilians-Universität München.

Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte-Carlo-Methode.* Philadelphia: SIAM.

Panneton, F., & L'Ecuyer, P. (kein Datum). Improved Long-Period Generators Based on Linear Recurrences Modulo 2. Université de Montréal.

Park, S. K., & Miller, K. W. (1988). *Random Number Generators: Good Ones Are Hard To Find.* Communications of the ACM.

Peterson, I. (1998). *The Jungles Of Randomness.* New York: John Wiley & Sons Inc.

Plate J., E. (1993). *Statistik und angewandte Wahrscheinlichkeitslehre für Bauingenieuere.* Berlin: Ernst, Verlag für Architektur u. techn. Wissenschaften.

Press, W. H., & Teukolsky, S. A. (1992). *Numerical Recipes* (2. Ausg., Bd. The Art of Scientific Computing). Cambridge: Cambridge University Press.

Reichenbach, H. (1935). *Wahrscheinlichkeitslehre. Eine Untersuchung über die logischen und mathematischen Grundlagen der Wahrscheinlichkeitsrechnung.* Leiden: Sijthoff.

Sachs, L. (1984). *Angewandte Statistik.* Berlin, Heidelberg: Springer.

Sedgewick, R. (1992). *Alorithmen.* Deutschland: Addison-Wesley.

Tezuka, S., L'Ecuyer, P., & Couture, R. (27. November 1995). On the lattice structure of the Add-with-carry and Subtract-with-borrow random number generators. Tokyo Research Laboratory (Tezuka) & Université de Montréal (L'Ecuyer) & Université Laval (Couture).

TheRandCorporation. (1955). *A Million Random Digits with 100 000 Normal Deviates.* Glencoe Illinois: Free press.

Wahler, S., Rose, O., & Schömig, A. (September 1997). Implementierung und Test neuartiger Zufallszahlengeneratoren. Würzburg, Deutschland: Institut für Informatik, Universität Würzburg.

11.2. Elektronische Quellen

Kolmogorov-Smirnov Test: cs.indiana.edu. (24. 12 2013). Von cs.indiana.edu: http://www.cs.indiana.edu/~kapadia/project2/node14.html abgerufen

Mersenne Twister Home Page. (5. Januar 2014). Von Mersenne Twister Home Page: http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html abgerufen

31x31 Binary Matrices Test: software.intel. (21. 12 2013). Von software.intel.com: http://software.intel.com/sites/products/documentation/doclib/mkl_sa/111/vslnotes/index.htm#8_3_4_Rank_of_31x31_Binary_Matrices_Test.htm abgerufen

analysis: random.org. (29. Dezember 2013). Von random.org: http://www.random.org/analysis/ abgerufen

Beschreibung der Funktion ZUFALLSZAHL in Excel 2007 und Excel 2003: support.microsoft.com. (7.. Januar 2008). Von support.microsoft.com: http://support.microsoft.com/kb/828795/de abgerufen

Birthday-spacings: software.intel-Website. (19. 12 2013). Von software.intel-Website: http://software.intel.com/sites/products/documentation/hpc/mkl/vslnotes/8_3_2_Birthday_Spacing_Test.htm abgerufen

Clewett, J., & Numberphile), (. d. (10. April 2013). *Random Numbers - Numberphile: Youtube.com*. Von youtube.com/user/numberphile: http://www.youtube.com/watch?v=SxP30euw3-0 abgerufen

Demokrit: zitate.eu. (27. 12 2013). Von zitate.eu: http://www.zitate.eu/de/autor/739?page=2 abgerufen

everything2.com. (1. Januar 2014). Von lavarand: everything2.com: http://everything2.com/title/lavarand abgerufen

Güte eines Pseudozufallszahlengenerators: wikipedia.org. (6. November 2013). Von wikipedia.org: http://de.wikipedia.org/wiki/Zufallszahlengenerator#G.C3.BCte_eines_Pseudozufallszahlengenerators abgerufen

Haahr, M. (30. Dezember 2013). *Randomness: random.org*. Von random.org: http://www.random.org/randomness/ abgerufen

Hardware Random number generator: wikipedia.org. (22. Dezember 2013). Von wikipedia.org: http://en.wikipedia.org/wiki/Hardware_random_number_generator abgerufen

Hellekalek, P. (4. Januar 2014). *Inverse Generators*. Von Inverse Pseudorandom Number Generators: Concepts Results and Links: http://random.mat.sbg.ac.at/generators/wsc95/inversive/node2.html#SECTION00020000000000000000 abgerufen

History: random.org. (30. Dezember 2013). Von random.org: http://www.random.org/history/ abgerufen

informatik: uni-hamburg. (18. Dezember 2013). Von Universität Hamburg-Website: http://www.informatik.uni-hamburg.de/TKRN/world/abro/NMI/kapitel07.pdf abgerufen

Köchel, & Flohrer. (29. März 1995). *Der Gap-Test: tu-chemnitz.de*. Von tu-chemnitz.de: http://www-user.tu-chemnitz.de/~jflo/Simulation/ZZ/gaptest.html abgerufen

lavarand.org. (20. Juni 2000). Von lavarand.org: http://lavarand.sgi.com/ abgerufen

Linear congruentia generator - Parameters in common use: wikipedia.org. (29. November 2013). Von wikipedia.org: http://en.wikipedia.org/wiki/Linear_congruential_generator#Parameters_in_common_use abgerufen

Linearer Kongruenzgenerator: wikipedia.org. (25. Dezember 2013). Von wikipedia.org: http://de.wikipedia.org/wiki/Linearer_Kongruenzgenerator#Fibonacci-Generator abgerufen

Lotto am Mittwoch: sueddeutsche. (3. April 2013). Von sueddeutsche-Website: http://www.sueddeutsche.de/panorama/lotto-am-mittwoch-panne-bei-ziehung-zahlen-ungueltig-1.1639793 abgerufen

Mangaldan, J. (5. Januar 2014). *On emulating the Texas Instruments random number generator: The Pleasure of Figuring Things Out*. Von The Pleasure of Figuring Things Out/Mangaldan-Website: http://tpfto.wordpress.com/2012/02/12/on-emulating-the-texas-instruments-random-number-generator/ abgerufen

Marsaglia, G. (11. 12 2013). *CDRom: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness*. Von http://www.stat.fsu.edu/pub/diehard/cdrom/ abgerufen

Marsaglia, G. (11. 12 2013). *Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness*. Von http://www.stat.fsu.edu/pub/diehard/cdrom/source/tests.txt abgerufen

Matrix-Rank: easycalculation.com. (21. 12 2013). Von easycalculation.com: http://easycalculation.com/matrix/matrix-rank.php abgerufen

Matsumoto, M., & Nishimura, T. (03. Februar 2014). *The origin of the name MT: Mersenne Twister Home Page*. Von Mersenne Twister Home Page: http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/ename.html abgerufen

McNichol, T. (1. 1 2014). *Totally Random: wired.com*. Von wired.com: http://www.wired.com/wired/archive/11.08/random.html abgerufen

Meier, W. (4. 11. 2009). Zufallszahlen in der Kryptographie. Uni Basel. Von http://web.fhnw.ch/personenseiten/marcel.steiner/Weiterbildung/Kolloquium/Zufallszahlen%20in%20der%20Kryptographie%20(W.Meier,%2004.11.2009).pdf abgerufen

Mersenne-Twister: wikipedia.org. (20. Dezember 2013). Von wikipedia.org: http://de.wikipedia.org/wiki/Mersenne-Twister abgerufen

Monte Carlo-Simulation: Palisade. (6. Januar 2014). Von Palisade-Website: http://www.palisade.com/risk/de/monte_carlo_simulation.asp abgerufen

Monte-Carlo-Simulation: exp.univie.ac.at. (06. Januar 2014). Von exp.univie.ac.at: http://www.exp.univie.ac.at/sc/sim/sim.pdf abgerufen

Monte-Carlo-Simulation: wikipedia.org. (13. Dezember 2013). Von wikipedia.org: http://en.wikipedia.org/wiki/Monte_Carlo_simulation abgerufen

Mordasini, C., & Klahr, H. (3. Januar 2013). *Mordasini: UKnum: Max-Planck-Institut für Astronomie.* Von Max-Planck-Institut für Astronomie-Website: http://www.mpia-hd.mpg.de/~mordasini/UKNUM/randomnumbers.pdf abgerufen

Multiply with carry: wikipedia.org. (16. November 2013). Von wikipedia.org: http://de.wikipedia.org/wiki/Multiply-with-carry abgerufen

pRNGs: math.umn.ed. (24. 12 2013). Von math.umn.ed: http://www.math.umn.edu/~garrett/students/reu/pRNGs.pdf abgerufen

Pseudorandom number generator: wikipedia.org. (5. December 2013). Von wikipedia.org: http://en.wikipedia.org/wiki/Pseudorandom_number_generator#Periodicity abgerufen

Pseudozufall: wikipedia.org. (15. August 2013). Von Wikipedia.org: http://de.wikipedia.org/wiki/Pseudozufall abgerufen

Random Number Generators: PJM's Sparse Home Page. (27. Dezember 2013). Von PJM's Sparse Home Page: http://www.paulm.org/random.html abgerufen

Rauschen: wikipedia.org. (20. November 2013). Von wikipedia.org: http://de.wikipedia.org/wiki/Rauschen_(Physik) abgerufen

Rechner-Rauschen: sengpielaudio.com. (30. 12 2013). Von sengpielaudio.com: http://www.sengpielaudio.com/Rechner-rauschen.htm abgerufen

Reichenbach, H. (1935). *Wahrscheinlichkeitslehre. Eine Untersuchung über die logischen und mathematischen Grundlagen der Wahrscheinlichkeitsrechnung.* Leiden: Sijthoff.

Roney-Dougal, C., du Sautoy, M., & Gowers, T. (13. Januar 2011). Random and Pseudorandom. Großbritannien: BBC Radio 4.

Sammelbilderproblem: Wikipedia.org. (17. November 2013). Von Wikipedia.org: http://de.wikipedia.org/wiki/Sammelbilderproblem abgerufen

Shot noise: princeton.edu. (30. Dezember 2013). Von princeton.edu: http://www.princeton.edu/~achaney/tmve/wiki100k/docs/Shot_noise.html abgerufen

Softwaretechnische Realisierungen: wikipedia.org. (14. Dezember 2013). Von wikipedia.org: http://de.wikipedia.org/wiki/Zufallszahlengenerator#Softwaretechnische_Realisierungen abgerufen

Spectral Test: wikipedia.com. (25. 12 2013). Von wikipedia.com: http://en.wikipedia.org/wiki/Spectral_test abgerufen

The Art of Computer Programming: wikipedia. (19. December 2013). Von wikipedia.com: http://en.wikipedia.org/wiki/The_Art_of_Computer_Programming abgerufen

Wahrnehmung des Zufalls: wikipedia.org. (12. August 2013). Von wikipedia.org: http://de.wikipedia.org/wiki/Zufall#Wahrnehmung_des_Zufalls abgerufen

Walker, J. (28. Januar 2008). *ENT (A Pseudorandom Number Sequence Test Program): fourmilab*. Von fourmilab: http://www.fourmilab.ch/random/ abgerufen

Wie berechnet man Pi nach der Monte-Carlo-Methode ? (3. Februar 2014). Von ekg-lemgo.de: http://www.ekg-lemgo.de/html/unterricht/faecher/diff-inf-mathe/kortemeier/seite2/screen2.html abgerufen

Yee, A. J., & Kondo, S. (28. Dezember 2013). *Round 2... 10 Trillion Digits of Pi: numberworld.org*. Von numberworld.org: http://www.numberworld.org/misc_runs/pi-10t/details.html abgerufen

Zufall: Duden. (10. Dezember 2013). Von Duden-Website: http://www.duden.de/rechtschreibung/Zufall abgerufen

12. ABBILDUNGSVERZEICHNIS

12.1. ABBILDUNGEN NACH SEITE

12.2. QUELLEN DER ABBILDUNGEN

Deckblatt: Eigenes Layout, Bild: *Cristian Ilies Vasile.* „Flow of life flow of pi“ (Created with Circos) Mit freundlicher Genehmigung von Christian Vasile

Abbildung 1: Entnommen: *Kütting, Herbert.* Elementare Stochastik, Heidelberg, Berlin - Spektrum Akad. Verl., 1999 (dort wiederum zitiert nach: *W.A. Wallis/H.V. Roberts*. Methoden der Statistik, Rowohlt 1977, S.523)

Abbildung 2: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 3: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 4: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 5: Entnommen von: *Knuth, D.* (1969). The Art of ComputerProgramming. Reading, Mass.: Addison-Wesley. Table 2 auf Seite 48

Abbildung 6: Eigene Darstellung (Programm: gnuplot [Version: 4.6] [12. März 2012])

Abbildung 7: Entnommen von URL: http://casoilresource.lawr.ucdavis.edu/drupal/book/export/html/371 [07.02.2014]

Abbildung 8: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 9: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 10: Eigene Darstellung

Abbildung 11: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 12: Eigene Darstellung

Abbildung 13: Eigene Darstellung

Abbildung 14: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 15: Eigene Darstellung (Programm: gnuplot [Version: 4.6] [12. März 2012])

Abbildung 16: Eigene Darstellung (Programm: gnuplot [Version: 4.6] [12. März 2012])

Abbildung 17: Eigene Darstellung (Programm: gnuplot [Version: 4.6] [12. März 2012])

Abbildung 18: Eigene Darstellung (Programm: gnuplot [Version: 4.6] [12. März 2012])

Abbildung 19: Eigene Darstellung (Programm: gnuplot [Version: 4.6] [12. März 2012])

Abbildung 20: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 21: (leicht nachbearbeitet) Entnommen von: *Simtec Electronics*. Entropy-key-Website. URL: http://www.entropykey.co.uk/ [07.02.2014]

Abbildung 22: Eigene Fotografie

Abbildung 23: Entnommen von: *random.org.* random.org-website (The History of RANDOM.ORG). URL: http://www.random.org/history/ [07.02.14]

Abbildung 24: Eigene Darstellung (Programm: gnuplot [Version: 4.6] [12. März 2012])

Abbildung 25: Eigene Darstellung (Programm: gnuplot [Version: 4.6] [12. März 2012])

Abbildung 26: Entnommen von: *Ian White*. Wired-website (Totally Random). URL: http://www.wired.com/wired/archive/11.08/random.html [07.02.14]

Abbildung 27: Original von: *ZDF (*Lottozahlenziehung April 2013), Entnommen von: URL: http://www.berliner-kurier.de/panorama/ziehung-ungueltig--lotto-panne--hier-stecken-die-kugeln-fest,7169224,22276712.html [07.02.2014]

Abbildung 28: Eigene Darstellung

Abbildung 29: Eigene Darstellung

Abbildung 30: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 31: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

Abbildung 32: Eigene Darstellung (Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010])

13. TABELLENVERZEICHNIS

13.1. TABELLEN NACH SEITE

13.2. QUELLEN DER TABELLEN

Tabelle 1: Eigene Darstellung

Tabelle 2: (Knuth, 1969) Seite 39

Tabelle 3: Werte stammen von eigenem Experiment, nach der Vorlage von (Knuth, 1969) Seite 39

Tabelle 4: Eigene Darstellung, (verwendetes Programm: *Windows*. Microsoft Excel 2010 [Version 14.0] [15. Mai 2010]), nach der Tabelle in: (Abramowitz & Stegun, 1964) Tabelle 26.8

Tabelle 5: Eigene Darstellung (verwendetes Programm: *Marsaglia*. DIEHARD.exe [29.12.1995])

Tabelle 6: Eigene Darstellung (verwendetes Programm: *Marsaglia*. DIEHARD.exe [29.12.1995])

Tabelle 7: Eigene Darstellung, nach Informationen von: (31x31 Binary Matrices Test: software.intel, 2013)

Tabelle 8: Eigene Darstellung, nach Informationen von: (Marsaglia, Tests: The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 2013)

14. FORMEL- UND SEQUENZVERZEICHNIS

14.1. FORMELN

14.2. SEQUENZEN

π