

Wie beeinflusst der Goldene Schnitt die Ziehungswahrscheinlichkeit unter zwei Fibonacci-Zahlen im Lotto 6 aus 49?

Shirin Shahidi, Matrikelnummer 455612

2019-07-20

Inhaltsverzeichnis

1. Einleitung	2
1.1. Forschungsfrage	2
1.2. Ziel der Studie/ Analyse	3
2. Methodik	3
2.1. Studiendesign	3
2.2. Angewandte statistische Methoden	3
3. Ergebnisse	4
3.1 Kennzahlen der Stichprobe	4
3.2 explorative Statistik	4
3.3 deskriptive Statistik	5
3.4 Inferenzstatistik	5
4. Diskussion	5
5. Schlussfolgerungen	5
Literaturverzeichnis	6

1. Einleitung

“Das, wobei unsere Berechnungen versagen, nennen wir Zufall.” - Albert Einstein (zitate.net, 2019, o.S.). Im deutschen Lotto 6 aus 49 lassen sich Erfolgswahrscheinlichkeiten errechnen, nicht jedoch künftige Ziehungen vorhersagen. Tendenzen, dass spezifische Ereignisse eintreten, können auf Basis bereits vergangener Ziehungen statistisch ermittelt werden (vgl. Ute Sproesser et al., 2014, S.33).

“Etwas hat Gestalt, weil es auf Zahlen beruht[...]” (Jonak, 2018, S.136). Die menschliche Gestaltwahrnehmung bevorzugt Vollkommenheit, Symmetrie und bestimmte Proportionen wie etwa den Goldenen-Schnitt. Diese Wahrnehmungspräferenzen von Gestalten lassen sich auch im Tierreich beobachten, bei dem Gestaltwahrnehmung instinktiv zu einer Reaktion führt(vgl. Oerter, 2014, S.252).

Auf Basis dieser Tatsachen stellt sich die Frage, ob die Gestalt der Lottoziehungen bestimmte Proportionen bevorzugt oder ob Ziehungen bestimmter Zahlenkombinationen rein zufällig erfolgen. Konzentriert man sich auf die Proportion des Goldenen Schnittes, so wird dieser gemäß aktuellem Forschungsstand durch die Formel $\frac{\sqrt{5}+1}{2} \sim 1,618$ errechnet (vgl. Kohn, 2019, S.160). Leonardo of Pisa, auch unter dem Namen Fibonacci bekannt, veröffentlichte 1202 das Werk *Liber Abbaci*(vgl. Hannah, 2007, o.S.). In diesem führt er zur Fibonacci-Sequenz hin, welche im Tierreich die Vermehrung von Hasen in Zahlen beschreibt(vgl. Fibonacci, 1987, S.18). Im Lotto 6 aus 49 sind die ersten acht Fibonacci-Zahlen (1,2,3,5,8,13,21,34) enthalten(vgl. Kohn, 2019, S.159). Dividiert man zwei aufeinander folgende Fibonacci-Zahlen, so konvergieren diese gegen den Goldene-Schnitt(vgl. Knebl & Springer Fachmedien Wiesbaden GmbH, 2019, S.29).

Um zur Forschungsfrage hinführen zu können, wird im Folgenden auf die beliebteste Lotterie namens Lotto 6 aus 49 unter den Einwohnern Deutschlands eingegangen(vgl. Lotto.de, 2019, o.S.). Spielberechtigt sind dabei erwachsene Personen ab 18 Jahren, da gemäß Jugendschutzgesetz § 6 Spielhallen, Glücksspiele Minderjährigen das Glücksspiel mit Chance auf hohe Geldbeträge untersagt ist(vgl. JuSchG, 2019, o.S.). Ziel eines Lottospiels ist es 6 zufällig gezogene Gewinnzahlen aus 1-49 vorgegeben Ziffern vorherzusagen. Dabei wird ein Lotto-Schein mit 12 möglichen Tippfeldern und jeweils 49 Zahlen ausgefüllt. Pro Tippfeld werden 6 Zahlen angekreuzt. Jeder Lotto-Schein enthält eine Spielscheinnnummer, deren letzte Ziffer die Superzahl darstellt. Diese Superzahl wird neben den 6 Gewinnzahlen zufällig aus den Ziffern 0-9 ermittelt. Es müssen minimal 2 Ziffern und eine Superzahl korrekt prophezeit werden um einen Gewinn zu erzielen. Der maximale Gewinn stellt 6 richtige Gewinnzahlen und die korrekte Superzahl dar. Der halbierte Spieleinsatz stellt die Gewinnsumme dar, welche sich auf 9 Gewinnklassen aufteilen lässt. Zusätzlich können die Lotterien Spiel 77 und SUPER 6 sowie die GlücksSpirale im Lotto-Schein ausgewählt werden. Ein Tippfeld im Lotto 6 aus 49 kostet 1,0€ zzgl. Bearbeitungsgebühren. Die Ziehungen finden wöchentlich mittwochs um 18:25 Uhr und samstags um 19:25 Uhr unter www.lotto.de statt(vgl. Lotto.de, 2019, o.S.). Die Wahrscheinlichkeit, zwei konkrete Zahlen im Lotto zu ziehen beträgt $1 : ((\frac{49}{6}) * (\frac{48}{5})) \sim 0.01276$ (vgl. Lottozahlen.com, 2019, o.S.).

1.1. Forschungsfrage

Betrachtet man die bereits genannten Fibonacci Zahlen im Lotto 6 aus 49 und die Konvergenz gegen den Goldenen Schnitt, lässt sich die Forschungsfrage wie folgt ableiten:

Wie beeinflusst der Goldene Schnitt die Ziehungswahrscheinlichkeit unter zwei Fibonacci-Zahlen im Lotto 6 aus 49?

In einer ersten Hypothese soll geprüft werden, ob eine Fibonacci Zahl unter allen Fibonacci-Zahlen im Lotto 6 aus 49 am häufigsten mit ihrer Goldenen-Schnitt Zahl gezogen wird. Im Anschluss daran soll geprüft werden, ob eine Fibonacci Zahl unter allen Fibonacci-Zahlen im Lotto 6 aus 49 am häufigsten mittwochs mit ihrer Goldenen-Schnitt Zahl gezogen wird. Dieser Tag wird mit den Samstagsziehungen und den davon abweichenden Wochentagen verglichen. Die Hypothesen lassen sich wie folgt formulieren:

H1: Fibonacci-Zahlen werden unter allen Fibonacci-Zahlen im Lotto 6 aus 49 am häufigsten mit ihrer Goldenen-Schnitt Zahl gezogen

H2: Fibonacci-Zahlen werden unter allen Fibonacci-Zahlen im Lotto 6 aus 49 am häufigsten mittwochs mit ihrer Goldenen-Schnitt Zahl gezogen.

1.2. Ziel der Studie/ Analyse

Das Untersuchungsziel der Studie ist deskriptiv ausgerichtet mit der Anforderung an exakte Ergebnisse und eine repräsentative Stichprobe sowie der Beschreibung der Grundgesamtheit mit ihren wesentlichen Merkmalen.(vgl. Kuß, 2012, S.15) Ziel der Studie soll sein, herauszufinden, ob Fibonacci-Zahlenpaare, die dividiert den Goldenen-Schnitt ergeben im Lotto 6 aus 49 häufiger zusammen gezogen werden als Fibonacci-Zahlenpaare ohne den Goldenen-Schnitt Proportionen. Hierzu sollen nur die Gewinnzahlen ohne Superzahl in Betracht gezogen werden, da die Superzahlziehung eine abweichende Wahrscheinlichkeit von $\frac{1}{10}$ aufweist. Es werden die Häufigkeiten der Fibonacci-Zahlen, sowie jeglicher Fibonaccipaar-Ziehungen unter allen Lotto-Ziehungen von 1955 bis Ende Juni 2019 betrachtet. Sollte sich herausstellen, dass Zahlenpaare aus dieser Auswahl am häufigsten zusammen auftreten, so werden diese mit Hilfe der Inferenzstatistik genauer betrachtet um einen Einfluss des Goldenen-Schnittes auf diese Zahlenpaare ausfindig machen zu können.

2. Methodik

Um die Methodik festlegen zu können, wird auf das Studiendesign, sowie die Datenerhebung und die angewandten statistischen Methoden genauer Bezug genommen (vgl. Kuß, 2012, S.13).

2.1. Studiendesign

Das Studiendesign wird definiert durch die Art der Datenerhebung und Ziehung der Stichprobe. Die vorliegende Studie basiert auf den Sekundärforschungen der Westdeutsche Lotterie GmbH & Co. oHG, sowie den Datenerhebungen unter www.lottozahlenonline.de durch Gustav Zygmund. (vgl. Kuß, 2012, S.15)

2.1.1 Datenerhebung

Da es sich um die Beschreibung der Veränderungen der Lotto-Ziehungen im Zeitablauf von 1955 bis 2019 handelt, wurden die in der Studie verwendeten Sekundärdaten in Form von Längsschnitt-Untersuchungen als Beobachtung erhoben. Da das Forschungsproblem auf Basis von Sekundärdaten bearbeitbar ist, wird in dieser Studie auf die Erhebung von Primärdaten verzichtet. Die bereits vorhandenen Daten werden für die Studie neu aufbereitet und analysiert. (vgl. Kuß, 2012, S.16)

2.1.2 Stichprobe

Die erhobenen Sekundärdaten stellen alle Ziehungen seit Anbeginn des Lottos 6 aus 49 von 1955 bis zum Zeitpunkt der Studienanalyse tabellarisch dar. Werden potenzielle Ungenauigkeiten wie etwa Tipp- oder Eingabefehler in den bereits vorhandenen Sekundärdaten vernachlässigt, kann von einer Totalerhebung der Daten ausgegangen werden. Dies führt dazu, dass Stichprobenfehler die Aussagekraft nicht einschränken können.(vgl. Kuß, 2012, S.43 f.)

2.2 Angewandte statistische Methoden

Um die Daten hinsichtlich ihrer Eignung für statistische Methoden zu überprüfen, wird eine explorative Datenanalyse zur Verdichtung der Daten durchgeführt. Da es sich um Häufigkeitsbeobachtungen der Fibonacci-Zahlenpaare in der Stichprobe handelt, werden Schätzverfahren angewandt, um Annahmen über entsprechende Wahrscheinlichkeitsverteilungen in der Grundgesamtheit treffen zu können (vgl. Kuß, 2012, S.193 f.). Eine Annahme oder Ablehnung der Hypothesen erfolgt mittels der Zweiseitigen Fragestellung (vgl. Hellbrück, 2009, S.73–76). Um den Datensatz auf Linearität zu überprüfen wird anschließend eine lineare Regression durchgeführt (vgl. Hellbrück, 2009, S.250). Für die Ergebnisse der statistischen Methoden wird die Programmiersprache R verwendet. Diese ermöglicht den Einsatz moderner Verfahren der Datenanalyse und ist reproduzierbar, automatisierbar, quelloffen, kostenlos und grafisch abbildbar. (vgl. Sauer, 2018, S.16 f.) Für den Datensatz wurden die Ergebnisse der Lotto-Ziehungen aus den Quellen Westdeutsche Lotterie GmbH & Co. oH

3. Ergebnisse

Um zu den Ergebnissen der Studie hinführen zu können werden die bereits genannten statistischen Methoden im Folgenden anhand von R-Syntax-Abschnitten (sogenannten Chunks) erklärt, sowie grafisch veranschaulicht (vgl. Sauer, 2018, S.26).

3.1 Kennzahlen der Stichprobe

Um einen Überblick über die Kennzahlen der Stichprobe zu erhalten, wird die Bibliothek “mosaic” geladen und der R-Befehl “inspect” ausgeführt:

```
##
## categorical variables:
##      name  class levels    n missing
## 1 Wochentag factor      6 4295      0
##                                distribution
## 1 Samstag (65.3%), Mittwoch (22.6%) ...
##
## quantitative variables:
##      name  class min Q1 median Q3 max    mean    sd    n missing
## 1 GewZahl_1 integer  1 12    25 37  49 24.64331 14.31717 4295      0
## 2 GewZahl_2 integer  1 12    24 37  49 24.64983 14.25556 4295      0
## 3 GewZahl_3 integer  1 13    25 37  49 25.11665 13.86886 4295      0
## 4 GewZahl_4 integer  1 13    25 37  49 25.06682 13.98750 4295      0
## 5 GewZahl_5 integer  1 13    26 38  49 25.45285 14.06999 4295      0
## 6 GewZahl_6 integer  1 13    26 38  49 25.46123 14.44079 4295      0
```

Die Stichprobe des Datensatzes Lotto 6 aus 49 von 1955 bis 2019 enthält 6 numerische und eine kategoriale Variable. Insgesamt enthält die Stichprobe 4295 Beobachtungen. Die numerischen Variablen stellen die Gewinnzahlen kategorisiert nach Ziehungsrunde dar. Das Minimum liegt bei der Ziffer 1 und das Maximum bei Ziffer 49. Die kategoriale Variable nennt den Wochentag, an dem die Beobachtungen aufgezeichnet wurden. Die Lagemaße und Streumaße sind unterschiedlich ausgeprägt, wie anhand der Werte von Quantil 1 und 3 sowie der Mittelwerte und dem Median zu erkennen ist. Die Standardabweichung fällt ebenfalls unterschiedlich aus. Die Variablenamen werden wie folgt ausgegeben: GewZahl_1, GewZahl_2, GewZahl_3, GewZahl_4, GewZahl_5, GewZahl_6, Wochentag. Insgesamt sind keine fehlenden Angaben vorhanden. Einen Überblick über die letzten 6 Beobachtungen des Datensatzes verschafft der Befehl “tail”:

```
##      GewZahl_1 GewZahl_2 GewZahl_3 GewZahl_4 GewZahl_5 GewZahl_6 Wochentag
## 4290         3         11         15         18         33         45   Samstag
## 4291         3         13         25         27         32         48   Mittwoch
## 4292         5          6         14         16         31         48   Samstag
## 4293         3         22         24         25         43         45   Mittwoch
## 4294         5         14         32         37         46         47   Samstag
## 4295        10         15         31         34         35         45   Mittwoch
```

3.2 explorative Statistik

Anhand der explorativen Statistik sollen Häufigkeitsverteilungen der Fibonacci-Zahlen betrachtet werden. Hierzu wird der Datensatz nach Fibonacci-Zahlen mittels R-Skript “_01_Sort_by_Fibo.R” sortiert und gefiltert. Die daraus gewonnen Daten werden auf Fibonacci-Zahlenpaare im Goldenen-Schnitt Verhältnis und deren Häufigkeitsverteilung überprüft und in eine Rangfolge gebracht.

```
##      1&13 1&34  1&8  1&5 1&21  1&3  1&2
## 1 2024 2053 2092 2109 2137 2140 2193

##      2&13  2&5 2&21  2&3  2&8 2&34  2&1
## 1 2010 2051 2078 2110 2127 2137 2179
```

```
## 1&13 1&34 1&8 1&5 1&21 1&3 1&2
## 1 2024 2053 2092 2109 2137 2140 2193
```

3.3 deskriptive Statistik

3.4 Inferenzstatistik

3.4.1 Bootstrap

3.4.2 Permutation

3.4.3 Lineare Regression

Gehen Sie hier näher auf die *verwendeten* Variablen ein. In R markdown Dateien können Sie einfach den R Code in Chunks einfügen (sogenannte “Code Chunks”). Entweder über das Menü **Insert -> R** oder über die Tastenkombination **strg+alt+i**.

Hier stellen und beschreiben Sie die Ergebnisse Ihrer kleinen Studie. Ihre Ergebnisse stellen Sie graphisch und/ oder tabellarisch dar. Ganz allgemein wird im Resultateteil NICHT interpretiert, sondern die Ergebnisse ausschließlich beschrieben.

Der Ergebnisteil kann zur besseren Orientierung und zum besseren Lesen weitere Gliederungspunkte enthalten (z.B. bei Hypothesenwechsel, für jedes Ergebnis o.ä.).

- die Abbildungen werden durchnummeriert (Abbildung 1 bis Abbildung xx) ebenso die Tabellen.
- Jede Abbildung erhält eine Bildunterschrift und jede Tabelle eine Beschreibung (über der Tabelle).
- die Bildunterschrift dürfen Sie mit R markdown machen oder aber in word! Eine Vorlage für eine Bildunterschrift ist im Template enthalten.

Beispiel für Ergebnisbeschreibung und einer Figure caption = Bildunterschrift:

Die Ergebnisse graphisch und tabellarisch dargestellt. Eine tabellarische Darstellung der Ergebnisse ist im Teil der deskriptiven Statistik oft sinnvoll. Sie entscheiden hier was eine sinnvolle Darstellung ist!

Beschreiben Sie das zentrale Ergebnis und Auffälligkeiten.

Folgendes sollte im Ergebnisteil enthalten/ ver- oder bearbeitet sein: 1. Kennzahlen 2. explorative / deskriptive Statistik 3. Inferenzstatistik

Beschreiben Sie Ihre Stichprobe und Ihre Variablen. Das ist hilfreich zum Verständnis und Nachvollziehbarkeit Ihrer Studie, der Datenerhebung und der Hypothesentestung.

Gehen Sie hier genauer auf die untersuchten Hypothesen und Modelle ein. Geben Sie zur Methodik 1-3 Literaturquellen an.

4. Diskussion

Für die Diskussion setzen Sie das Ergebnis in Kontext zu bereits publizierter Literatur! (hier ist das Lesen und die Angabe (CAVE: richtig zitieren) von Literatur notwendig) Mind. 3 Literaturquellen

5. Schlussfolgerungen

“Es steckt oft mehr Geist und Scharfsinn in einem Irrtum als in einer Entdeckung.” - Joseph Joubert

Fassen Sie hier kurz die zentralen Ergebnisse für Ihre Forschungsthematik zusammen. Gehen Sie auch auf die Grenzen Ihrer Analyse ein.

Literaturverzeichnis

Hier stehen die im Text verwendeten Quellen:

- Nachname Autor1, Anfangsbuchstabe Vorname Autor1, Nachname Autor2, Anfangsbuchstabe Vorname Autor2 1 & Nachname Autor3, Anfangsbuchstabe Vorname Autor3, ... (Jahr der Veröffentlichung). Titel des Beitrags. Journal, Volume, Issue, Seitenzahlen

Fibonacci, L.P. (1987) INTRODUCTION: A Brief Biography of Leonardo Pisano (Fibonacci) [1170-post 1240]. In: LEONARDO PISANO Fibonacci (Hrsg.). *The Book of Squares*. [Online]. San Diego, Academic Press. S. xv–xx. Available from: doi:10.1016/B978-0-08-088650-3.50005-0 [Zugegriffen: 15 Juli 2019].

Hannah, J. (2007) False position in Leonardo of Pisa's Liber Abbaci. *Historia Mathematica*. [Online] 34 (3), 306–332. Available from: doi:10.1016/j.hm.2006.10.004 [Zugegriffen: 15 Juli 2019].

Hellbrück, R.P. (2009) OCLC: 458747755. *Angewandte Statistik mit R: eine Einführung für Ökonomen und Sozialwissenschaftler*. Lehrbuch. 1. Aufl. Wiesbaden, Gabler.

Jonak, U. (2018) *Essays zur Architektur*. [Online]. Wiesbaden, Springer Fachmedien Wiesbaden. Available from: doi:10.1007/978-3-658-19129-0 [Zugegriffen: 8 Juli 2019].

JuSchG (2019) *JuSchG - Jugendschutzgesetz*. [Online]. 2019. Available from: <https://www.gesetze-im-internet.de/juschg/BJNR273000002.html> [Zugegriffen: 16 Juli 2019].

Knebl, H. & Springer Fachmedien Wiesbaden GmbH (2019) OCLC: 1098316960. *Algorithmen und Datenstrukturen Grundlagen und probabilistische Methoden für den Entwurf und die Analyse*.

Kohn, W. (2019) OCLC: 1102475595. *MATHEMATIK FR WIRTSCHAFTSINFORMATIKER: Grundlagen und anwendungen*. S.l., SPRINGER.

Kuß, A. (2012) OCLC: 795116334. *Marktforschung: Grundlagen der Datenerhebung und Datenanalyse*. Springer-Gabler-Lehrbuch. 4., überarb. Aufl. Wiesbaden, Springer Gabler.

Lotto.de (2019) *LOTTO 6aus49 Spielregeln LOTTO.de*. [Online]. 2019. Available from: <https://www.lotto.de/lotto-6aus49/spielregeln> [Zugegriffen: 16 Juli 2019].

Lottozahlen.com (2019) *6 aus 49 Wahrscheinlichkeit einfach erklärt. Lottozahlen*. [Online]. 2019. Available from: <https://www.lottozahlen.com/lotto-nachrichten/6-aus-49-wahrscheinlichkeit-einfach-erklart> [Zugegriffen: 16 Juli 2019].

Oerter, R. (2014) OCLC: 887522626. *Der Mensch, das wundersame Wesen: was Evolution, Kultur und Ontogenese aus uns machen*. Wiesbaden, Springer Spektrum.

Sauer, S. (2018) *Moderne Datenanalyse mit R: Daten einlesen, aufbereiten, visualisieren und modellieren*. FOM-Edition. 1. Auflage 2019. Wiesbaden, Springer Fachmedien Wiesbaden GmbH.

Ute Sproesser et al. (2014) *Daten, Zufall und der Rest der Welt: didaktische Perspektiven zur anwendungsbezogenen Mathematik*. Research. Silvia Wessolowski & Claudia Wörn (Hrsg.). Wiesbaden, Springer Spektrum.

zitate.net (2019) *Zufallzitate Top 20 Zitate und Sprüche über Zufälle ... Zitate.net*. [Online]. 2019. Available from: <http://zitate.net/zufall-zitate> [Zugegriffen: 14 Juli 2019].