Mathe Ausarbeitung

über

Geburtstagsproblem Teil II

für den Kurs

Angewandte Informatik

an der

DHBW Mosbach

von

Tim Hönnige, Matthias JooSS und Liam Friedrich

Fällig am	31.03.2021
Kurs	INF19A
Dozent	David Weniger

Zusammenfassung

Angenommen wir betrachten das Sommermärchen 2006. Hierbei durfte jede Mannschaft 23 Spieler für ihren Kader nominieren. Der Kader der deutschen FuSSballnationalmannschaft für dieses WM-Turnier bestand unter anderem aus Mike Hanke und Christoph Metzelder. Obwohl die Beiden weder für den gleichen Verein spielten noch die gleiche Position im Feld innehattet, haben sie dennoch etwas gemeinsam: ihren Geburtstag! Ein Zufall ist dies schon, aber ist es auch ein seltenes Ereignis? Wie oft kommt so etwas vor? Um dieser Frage auf den Grund zu gehen, eignet sich ein Blick auf das sogenannte Geburtstagsproblem, das auch unter dem Begriff Geburtstagsparadoxon zu finden ist. Seinen Namen erhielt es aufgrund der Tatsache, dass viele Menschen es für eine scheinbar unsinnige Behauptung hielten, da sie nur ihr näheres Umfeld betrachteten und eine mögliche Ansammlung von 23 beliebigen, sich zufällig an einem Platz befindenden Menschen, völlig auSSer Acht lieSSen. Das Geburtstagsparadoxon gibt eine Antwort auf die Frage, wie viele Personen in einem Raum sein müssen, damit eine bestimmte Wahrscheinlichkeit besteht, dass mindestens zwei Personen den gleichen Geburtstag haben. Somit generiert es eine Wahrscheinlichkeit, dass unter k zufällig gewählten Personen mindestens zwei am selben Tag Geburtstag haben. In der folgenden Ausarbeitung wird das Geburtstagsparadoxon zunächst hergeleitet. Daraufhin folgt eine Darstellung der Näherung. Im Anschluss wird die Quantile beleuchtet und der Erwartungswert angegeben. Zum Schluss erfolgt ein Fazit.

Inhaltsverzeichnis

1 Berechnung

Das Geburtstagsproblem generiert eine Wahrscheinlichkeit das unter k zufällig gewählten Personen mindestens zwei am selben Tag Geburtstag haben. Es ist eine Abwandlung des Paradoxons der ersten Kollision. [1] Bei $k = 23$ ist die Wahrscheinlichkeit das zwei Personen bereits am gleichen Tag Geburtstag haben bei über 50%. Das ist gerade für stochastisch weniger bewanderte Personen erstaunlich, denn diese vermuten eine sehr viel geringere Wahrscheinlichkeit, da sie das Problem auf ihr Umfeld assoziieren.

Zuerst bestimmen wir n. Das ist die Wahrscheinlichkeit das eine Person an einem bestimmten Tag der Jahren Geburtstag hat. Das gewählte Jahr besitzt die Eigenschaften unserer Annahmen.

Daraus ergibt sich:

$$n = \frac{1}{365} \tag{1.1}$$

Um auszurechnen, wie viele Personen sich in einem Raum befinden müssen, so dass die Wahrscheinlichkeit, dass mindestens zwei Personen am selben Tag Geburtstag haben bei 50% oder mehr liegt verwenden wir das Gegenereignis. Wir berechnen die Wahrscheinlichkeit, dass alle Personen im Raum an verschiedenen Tagen Geburtstag haben und nähern uns den von oben den 50%

Für 2 Personen:

$$\frac{365}{365} * \frac{364}{365} = 0,997 \tag{1.2}$$

Die erste Person kann aus 365 Tagen wählen ohne das es zu einer Kollision kommt, für die 2. Person bleiben 364 Tage

Für 3 Personen:

$$\frac{365}{365} * \frac{364}{365} * \frac{363}{365} = 0,991 \tag{1.3}$$

Dies wird weitergeführt, bis die Wahrscheinlichkeit für das Gegenereignis bei etwa 50% liegt, somit liegt dann auch das Ereignis, dass mindestens Zwei Personen am selben Tag Geburtstag haben bei etwa 50%. Dieser Fall Tritt bei einer Personen Zahl von 23 ein.

$$\frac{365}{365} * \frac{364}{365} * \frac{363}{365} * ... * \frac{343}{365} = 0,493 \tag{1.4}$$

Die Wahrscheinlichkeit für keine Kollision liegt bei 0,493, somit liegt die Wahrscheinlichkeit für eine Kollision bei 0,507

$$1 - 0,492 = 0,507 \tag{1.5}$$

2 Annahmen

Um das Geburtstagsparadoxon als vereinfachtes statistisches Beispiel zu verwenden, werden folgende drei Annahmen getroffen:

- Jedes Jahr hat einheitlich 365 Tage, ohne Berücksichtigung des Schaltjahres.
  - Das Schaltjahr findet nur alle vier Jahre Anwendung und wäre im Experiment nicht so einfach zu berücksichtigen. Diese Annahme hat natürlich aber auch eine geringfügig gröSSere Wahrscheinlichkeit zu Folge.
- Jeder der 365 Tage eines Jahres ist als Geburtstag gleichwahrscheinlich.
  - Das Jahr erfüllt in unserer Variante die Anforderungen an ein La Place Experiment.
  - In der Realität ist das nicht so, denn hier gibt es eine Häufung an Geburten nach bestimmten Events. Ein Beispiel ist der Valentinstag.
- Das Auswahlkriterium der Testpersonen erfolgt hinsichtlich ihres Geburtstages.
- Das Ereignis "mindestens ein doppelter Geburtstag ist schwierig zu berechnen, da es eine groSSe Menge an Teilereignissen einschlieSSt (u.a. drei Personen haben den gleichen Geburtstag oder zwei Paare haben denselben Geburtstag). Daher wird in dieser Ausarbeitung das Gegenereignis kein doppelter Geburtstag angenommen.
- Als kleinste natürliche Zahl wird die 1 angenommen.

Auf diesen Annahmen beruhen alle unsere Berechnungen in dieser Ausarbeitung.

3 Gleichverteilung:

Sei eine endliche Ergebnismenge und $p : \Omega \to [0, 1]$ eine Wahrscheinlichkeitsfunktion. Wenn $p(x) = \frac{1}{|\Omega|}$ gilt, dann spricht man von Gleichverteilung, genauer: die Wahrscheinlichkeitsfunktion p ist gleichverteilt. Alle möglichen Ausgänge des Zufallsexperiments treten mit der gleichen Wahrscheinlichkeit auf.

4 La Place:

Ein Versuch heiSSt dann La Place Versuch, wenn keines der Elementarereignisse mit einer gröSSeren Häufigkeit auftritt als ein anderes. Somit ist die Wahrscheinlichkeit aller möglichen Ergebnisse gleich.

$$P(x) = \frac{1}{\Omega}$$

Bei Laplace Versuchen wird die Wahrscheinlichkeit $P(E)$ eines Ereignisses E wie folgt berechnet:

$$P(E) = \frac{AnzahlderzumEreignisEgehrendenErgebnisse}{AnzahlallermglichenErgebnisse} = \frac{IEI}{ISI}$$

Beispiel:

- Werfen einer fairen Münze ($\Omega = Kopf, Zahl$)

- N-maliges Werfen eines fairen Würfels ($\Omega = (1, ..., 6)^n$)

5 GauSSsche Summenformel:

Mithilfe der GauSSschen Summenformel lässt sich die Summe beliebig vieler ($= n$) natürlicher Zahlen berechnen. Hierbei addiert man alle natürlichen Zahlen von 1 bis zu der gewählten Grenze n.

$$0 + 1 + 2 + 3 + 4 + ... + n = \sum_{k=0}^{n} k = \frac{n(n+1)}{2} = \frac{n^2 + n}{2}$$

6 Herleitungen der verwendeten Formeln

Zur Modellierung des Geburtstagsproblems betrachten wir die Zufallsvariable:

$$X_n := \text{Zeitpunkt der ersten Kollision bei n Personen mit rein zufällig gewählten Geburtstagen} \tag{6.1}$$

Da zumindest zwei Personen vorhanden sein müssen damit es zu einer Kollision kommt, ist der minimale Wert 2. Höchstens sind es $n + 1$ Personen. Somit nimmt X_n die Werte $2, 3, \cdots, (n + 1)$ an und es gilt:

$$\mathbb{P}(X_n \geq k + 1) = \frac{n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot (n-k+1)}{n^k} \tag{6.2}$$

für jedes $k = 1, 2, \cdots, n + 1$. Durch die Annahme der gleichen Verteilung der Zufallsereignisse (Laplace-Modell), ergibt der Zähler von (6.2) die Anzahl der günstigen Fälle an.

Aus 6.2 folgt durch Verwendung des Gegenereignisses:

$$\mathbb{P}(X_n \geq k) = 1 - \prod_{j=1}^{k-1} (1 - \frac{j}{n}) \tag{6.3}$$

$$\mathbb{P}(Xn \leq 1) = 0 \tag{6.4}$$

Da bei einer einzigen Person $k = 1$ keine Kollision auftreten kann ist die Wahrscheinlichkeit für dieses Ereignis 0 (6.4). Deshalb ist der Wertebereich für k mit $k = [2; n + 1]$ angegeben.

In der Abbildung 3 ist die Wahrscheinlichkeit $P(X_n \leq k)$ durch eine Funktion von k mit dem Parameter $n = 365$ dargestellt. Unterschiedliche Zahlenwerte für n sind in den Tabellen im Absatz konkrete Zahlen aufgeführt. Für das Ereignis $X_n \leq 23$ ist die Wahrscheinlichkeit bereits höher als 50%.

Auf den ersten Blick scheint überraschen das bei $8,5 \cdot 10^{58}$ möglichen Kombinationen (365^{23}) der Geburtstagsverteilung bei 23 Personen. Die Wahrscheinlichkeit eines doppelten Geburtstags schon über 50% liegt. Die Erklärung hierfür ist das wir auf irgendeine und nicht auf eine bestimmte Kollision warten.

6.1 Vereinfachung der Wahrscheinlichkeitsverteilung

Die Vereinfachung der Wahrscheinlichkeitsverteilung baut auf der Ungleichung 6.5 auf.

$$1 - x \leq e^{-x} (x \in \mathbb{R}) \tag{6.5}$$

Diese lässt sich mit der Taylor-Entwicklung von e^x beweisen:

$$1 + x \leq 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = e^x. \tag{6.6}$$

Die Taylor Entwicklung beginnt bereits mit 1+x, somit ist sie mindestens genauso groSS wie die linke Seite der Gleichung und kommt als mögliche Näherung infrage [2, 560ff].

Unter Verwendung der Ungleichung 6.5, ist es uns möglich die unter 6.3 angegebene Funktion so weit zu vereinfachen, das kein Produkt- oder Summenzeichen mehr vorhanden ist, dadurch sind weitere Berechnungen einfacher zu realisieren.

$$\mathbb{P}(X_n \leq k) \approx 1 - \prod_{j=1}^{k-1} 1 - \frac{j}{n} \geq 1 - \exp(-\sum_{j=1}^{k-1} \frac{j}{n}) \tag{6.7}$$

$$\approx 1 - \exp(-\frac{k(k+1)}{2n}) \tag{6.8}$$

Bei der ersten Umformung in 6.7 wird das Produkt zu einer Summe im Exponenten von e, aufgrund der allgemeinen Potenzgesetze $a^r \cdot a^s = a^{r+s}$ [2, S. 267]. In der nächsten Umformung wird die im ersten Schritt geschaffene Summe mithilfe der GauSSschen Summenformel ersetzt [3, 9ff].

Somit ergibt sich zur Berechnung der Wahrscheinlichkeit die allgemeine Formel:

$$\mathbb{P}(X \leq k) \approx 1 - \exp(-\frac{k(k+1)}{2n}) \tag{6.9}$$

6.2 Wahrscheinlichkeitsdichte

In der Stochastik beschreibt die Wahrscheinlichkeitsdichte ist eine spezielle reellwertige Funktion zur Konstruktion von Wahrscheinlichkeitsverteilungen.

Im Unterschied zu Wahrscheinlichkeiten kann die Wahrscheinlichkeitsdichte auch gröSSere Werte als 1 annehmen. Dabei wird nicht der Funktionswert, sondern die Fläche unterm Funktionsgraphen berechnet, also das Integral.

Die Dichte kann mit zwei verschiedenen Herangehensweisen konstruiert werden: Durch eine Funktion, die aus der Wahrscheinlichkeitsverteilung generiert wird, oder durch die Ableitung der Wahrscheinlichkeitsverteilung. Es unterscheidet sich nur die Herangehensweise [1, 560ff].

Im Weiteren wird nur noch auf den Fall eingegangen, die in dem die Dichte aus der Wahrscheinlichkeitsverteilung abgeleitet wird [4, 22ff].

Allgemein ist die Wahrscheinlichkeitsdichte dann folgendermaSSen definiert:

$$\mathbb{P}(]-\infty, a]) = \int_{-\infty}^{a} f(x)\,\mathrm{d}x \tag{6.10}$$

bzw.

$$\mathbb{P}(X \leq a) = \int_{-\infty}^{a} f(x)\,\mathrm{d}x \tag{6.11}$$

Für das Geburtstagsproblem ist die Wahrscheinlichkeitsverteilung aus Formel 6.9 als gegeben anzusehen. Es lässt sich aus der Definition für die Dichte folgende Gleichung ableiten:

$$\mathbb{P}(X \leq k) \approx \int_{-\infty}^{k} f(k)dk \tag{6.12}$$

$$\implies 1 - \exp(-\frac{k(k+1)}{2n}) \approx \int_{-\infty}^{k} f(k)dk \tag{6.13}$$

Durch Ableiten der Gleichung ergibt sich eine Definition für die Funktion f(k):

$$\implies 1 - \exp(-\frac{k(k+1)}{2n}) \approx \int_{-\infty}^{k} f(k)dk \quad | \quad \frac{d}{dk} \tag{6.14}$$

$$\implies \frac{(k-1)}{n} \cdot \exp(-\frac{k(k+1)}{2n}) \approx f(k) \tag{6.15}$$

Somit ist die Dichte für das Geburtstagsproblem folgendermaSSen definiert:

$$\mathbb{P}(X = k) \approx \frac{(k-1)}{n} \cdot \exp(-\frac{k(k+1)}{2n}) \tag{6.16}$$

6.3 Erwartungswert

Der Erwartungswert für das Geburtstagsproblem lässt sich mithilfe der Weibull-Verteilung konstruieren. In dieser Ausarbeitung wird nicht näher auf die Weibull-Verteilung oder die verwendete Gamma-Funktion eingegangen, diese werden als vorausgesetzt angesehen [5].

Für die um 1 nach rechts verschobene Weibull-Verteilung werden folgende Parameter verwendet. $k = 2$, dadurch ergibt sich eine Rayleigh-Verteilung und $\lambda = \frac{1}{\sqrt{2n}}$. $\mathcal{T}$ bezeichnet die Gammafunktion [5].

Somit ergibt sich die Funktion:

$$\mathbb{E}(X-1) = \mathbb{E}(X) - 1 \approx \sqrt{2n} \cdot \mathcal{T}(1 + \frac{1}{2}) = \sqrt{2n} \cdot \frac{\sqrt{\pi}}{2} = \sqrt{\frac{\pi \cdot n}{2}} \tag{6.17}$$

$$\mathbb{E}(X) \approx \sqrt{\frac{\pi \cdot n}{2}} + 1 \tag{6.18}$$

für den Erwartungswert.

6.4 Varianz

Die Varianz wird so wie der Erwartungswert auch mithilfe der Weibull-Verteilung bestimmt. Die Parameter sind hierbei dieselben wie schon für den Erwartungswert verwendet wurden [5].

$$\begin{align} Var(X) &\approx \frac{1}{\lambda^2} \cdot [\mathcal{T}(1 + \frac{2}{k}) - \mathcal{T}^2(1 + \frac{1}{k})] \tag{6.19} \\ &\approx 2n \cdot [\mathcal{T}(1 + \frac{2}{2}) - \mathcal{T}^2(1 + \frac{1}{2})] \tag{6.20} \\ &\approx 2n \cdot [1 - \frac{\pi}{4}] \tag{6.21} \end{align}$$

Die Definition der Varianz der Weibull-Verteilung ist in 6.19 gegeben [5].

6.5 Quantile

Ein Quantil ist ein LagemaSS in der Statistik. Den meisten ist der Median bekannt, dabei handelt es sich um das 50% oder $\frac{1}{2}$-Quantil. Es lassen sich aber auch beliebige Quantile zwischen 0 und 1 bestimmen. Allgemein sind Quantile Schwellenwerte. Werden die gegebenen Daten nach ihrer Wertigkeit sortiert, ist ein bestimmter Anteil kleiner als das Quantil [1, S. 32, 35, 37].

Gegeben sei eine beliebige Zufallsvariable X. Dann ist x_p das p-Quantil von X, wenn gilt:

$$\mathbb{P}(X \leq x_p) \geq p \tag{6.22}$$

und

$$\mathbb{P}(x_p \leq X) \geq 1 - p \tag{6.23}$$

Im Folgenden wird beschrieben wie aus dieser Definition eine Funktion konstruiert werden kann, mit der sich die Quantile für das Geburtstagsparadoxon bestimmen lassen.

Für das $\frac{1}{2}$-Quantil (Median):

$$\begin{aligned}
&\mathbb{P}(X \leq k) = 1 - \mathbb{P}(X > k) \geq 1 - \exp(-\frac{k(k-1)}{n}) \stackrel{!}{=} \frac{1}{2} \\
&\Leftrightarrow ln(\frac{1}{2}) = \frac{k(k-1)}{2n} \\
&\Leftrightarrow -2n \cdot ln(2) = k(k-1) \\
&\Leftrightarrow 2n \cdot ln(2) = k(k-1) \\
&\Leftrightarrow 2n \cdot ln(2) = k^2 - k \\
&\Leftrightarrow -k^2 + k + 2n \cdot ln(2) = 0 \\
&\Leftrightarrow k^2 - k - 2n \cdot ln(2) = 0 \\
&\Leftrightarrow k = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2n \cdot ln(2)}
\end{aligned}$$

Daraus lässt sich dann folgendes Ableiten:

$$Q_{\frac{1}{2}}(X) \leq \left(\frac{1}{2} + \sqrt{\frac{1}{4} + 2n \cdot ln(2)}\right) \leq \left(1 + \sqrt{2n \cdot ln(2)}\right) \tag{6.24}$$

Das Quantil befindet sich somit in den Grenzen der quadratischen Funktion:

$$k = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2n \cdot ln(2)} \tag{6.25}$$

Das spezielle $\frac{1}{2}$-Quantil lässt sich auch allgemein bestimmen, sodass der Schwellwert p ein Parameter der Funktion ist:

$$\begin{aligned}
&1 - \exp(-\frac{k(k-1)}{n}) \stackrel{!}{=} p \\
&\Leftrightarrow ln(1-p) = \frac{k(k-1)}{2n} \\
&\Leftrightarrow 2n \cdot ln(1-p) = k(k-1) \\
&\Leftrightarrow 2n \cdot ln(1-p) = k^2 - k \\
&\Leftrightarrow -k^2 + k + 2n \cdot ln(1-p) = 0 \\
&\Leftrightarrow k^2 - k - 2n \cdot ln(1-p) = 0 \\
&\Leftrightarrow k = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2n \cdot ln(1-p)}
\end{aligned}$$

Daraus lässt sich wie bei der speziellen Lösung, folgende Aussage ableiten:

$$Q_p(X) \leq \left(\frac{1}{2} + \sqrt{\frac{1}{4} + 2n \cdot ln(1-p)}\right) \leq \left(1 + \sqrt{2n \cdot ln(1-p)}\right) \tag{6.26}$$

Somit lassen sich die Quantile durch folgende Funktionen approximieren:

$$Q_p(X) \approx \frac{1}{2} + \sqrt{\frac{1}{4} + 2n \cdot ln(1-p)} \tag{6.27}$$

$$Q_p(X) \approx 1 + \sqrt{2n \cdot ln(1-p)} \tag{6.28}$$

7 Näherung

Das offenbar spektakuläre Ereignis besteht scheinbar darin, dass bei der schrittweisen rein zufälligen Belegung von $n = 365$ unterschiedlichen Tagen die erste Kollision bereits nach dem 23. Tag stattfand, d.h. ein bereits durch einen Geburtstag belegter Tag wurde erneut belegt. Intuitiv würde man erwarten, dass der Zeitpunkt dieser ersten Kollision viel später liegt. Würden Sie z. B. bei n= 1000 Fächern darauf wetten, dass die erste Kollision spätestens nach 50 Versuchen auftrat?
Zur Modellierung des Kollisionsphänomens betrachten wir die Zufallsvariable

$X_n =$ Zeitpunkt der ersten Kollision beim sukzessivenrein zufälligen Besetzen von n Tagen.

Da mindestens 2 und höchstens $n+1$ Versuche bis zur ersten Kollision nötig sind, nimmt X_n die Werte $2, 3, ..., n+1$ an, und es gilt

$$P(X_n \geq k+1) = \frac{n * (n-1) * (n-2) * ... * (n-k+1)}{n^k} = \frac{n^{\underline{k}}}{n^k}$$

für jedes $k = 1, 2, ..., n+1$.
Daraus folgt das Gegenereignis

$$P(X_n \leq k) = 1 - \prod_{j=1}^{k-1}(1 - \frac{j}{n})$$

$$(k = 2, 3, ..., n+1; P(X_n \leq 1) = 0)$$

Die Formel zeigt die Wahrscheinlichkeiten $P(X_n \leq k)$ als Funktion von k für den Fall $n = 365$. Für das Ereignis $X_n \leq 23$ gilt$P(X_n \leq 23) = 0,5073$

Es vermag vermutlich überraschend zu erscheinen, dass wir bei fast 365 möglichen Tagen im Jahr durchaus auf das Auftreten des doppelten Geburtstags nach höchstens 23 Personen wetten können. Doch der Grund hierfür ist, dass wir auf irgendeine beliebige und nicht nach einer bestimmten Kollision suchen. Der Zeitpunkt der ersten Kollision bei der rein zufälligen sukzessiven Besetzung von n Tagen ist von der GröSSenordnung $\sqrt{n}$.

Da beim Geburtstagsproblem $P(X365 \leq 23) = 0.507 > 1/2$ gilt, kann durchaus darauf gewettet werden, dass unter 23 aber auch mehr Personen mindestens zwei am gleichen

Tag Geburtstag haben. Bei $\sqrt{365 * 2 * ln(2)} = 22.49$ ist die Approximation schon für n = 365 sehr gut.

8 Konkrete Zahlen

Hier werden die hergeleiteten Formeln angewendet um konkrete Zahlen zu erhalten. Um einen MaSSstab zu sehen werden jeweils ein Erdenjahr (365 Tage) ein Marsjahr (780 Tage) und ein Jupiterjahr (4330 Tage) gezeigt.

8.1 Wahrscheinlichkeit der ersten Kollision

Mithilfe der Formel 8.1 wird die Wahrscheinlichkeiten dafür berechnet, dass $X \leq k$ ist. X ist hierbei der Zeitpunkt der 1. Kollision. Mit der Formel 8.2 wird die Wahrscheinlichkeit erechnet, dass die Kollision bei K stattfindet.

$$P(X \leq k) \approx 1 - e^{-\frac{k(k-1)}{2n}} \tag{8.1}$$

$$P(X = k) \approx \frac{(k-1)}{n} \cdot e^{-\frac{(k-1)^2}{2n}} \tag{8.2}$$

8.1.1 $P(X \leq k)$

Erde

k	$P(X_n) \leq k)$	k	$P(X_n) \leq k)$
0	0,001	40	0,876
10	0,105	50	0,963
20	0,390	60	0,992
30	0,684	70	0,999

Tabelle 1: Wahrscheinlichkeit für die 1. Kollision nach höhstens k Personen auf der Erde

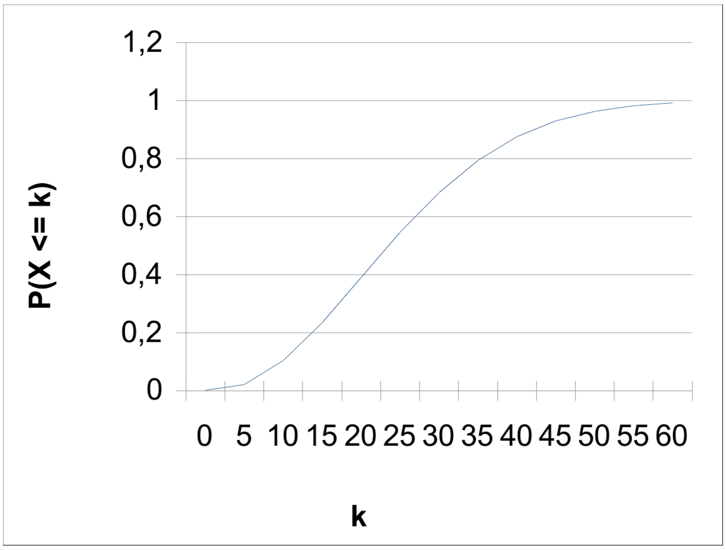

Abbildung 2: Wahrscheinlichkeit für den 1. gleichen Geburtstag nach höchstens k Ziehungen

Mars

k	$P(X_n \leq k)$	k	$P(X_n \leq k)$
0	0,001	60	0,893
15	0,118	75	0,970
30	0,417	90	0,994
45	0,711		

Tabelle 2: Wahrscheinlichkeit für die 1. Kollision nach höhstens k Personen auf dem Mars

Jupiter

k	$P(X_n \leq k)$	k	$P(X_n \leq k)$
0	0	120	0,805
30	0,093	150	0,923
60	0,331	180	0,975
90	0,599	210	0,994

Tabelle 3: Wahrscheinlichkeit für die 1. Kollision nach höhstens k Personen auf dem Jupiter

8.1.2 $P(X = k)$

Erde

k	$P(X_n = k)$	k	$P(X_n = k)$	k	$P(X_n = k)$
5	0,011	25	0,030	45	0,009
10	0,022	30	0,025	50	0,005
15	0,029	35	0,019	55	0,003
20	0,032	40	0,013	60	0,001

Tabelle 4: Wahrscheinlichkeit für die 1. Kollision bei der k. Person

Die höchste Wahrscheinlichkeit dafür, dass die 1. Kollision genau bei k auftritt liegt auf der Erde bei $\boldsymbol{\sim 20,105}$ (zu sehen auf Grafik 3), auf dem Mars bei $\boldsymbol{\sim 28,928}$ und auf dem Jupiter bei $\boldsymbol{\sim 66,803}$ also jeweils bei $\boldsymbol{1 + \sqrt{n}}$.

8.2 Erwartungswerte

Der Erwartungswert für die 1. Kollision wird über die Formel 8.3 errechnet.

$$E(X) \approx 1 + \sqrt{\frac{1}{2}\pi n} \tag{8.3}$$

Für die verschiedenen Planeten kommen dabei folgende Werte heraus:

Erde	24,945
Mars	36,003
Jupiter	83,472

8.3 Quantile

Die Quantile für die 1. Kollision werden über die Formel 8.4 berechnet.

$$q_\alpha \approx 1 + \sqrt{-n \cdot ln(1 - \alpha)} \tag{8.4}$$

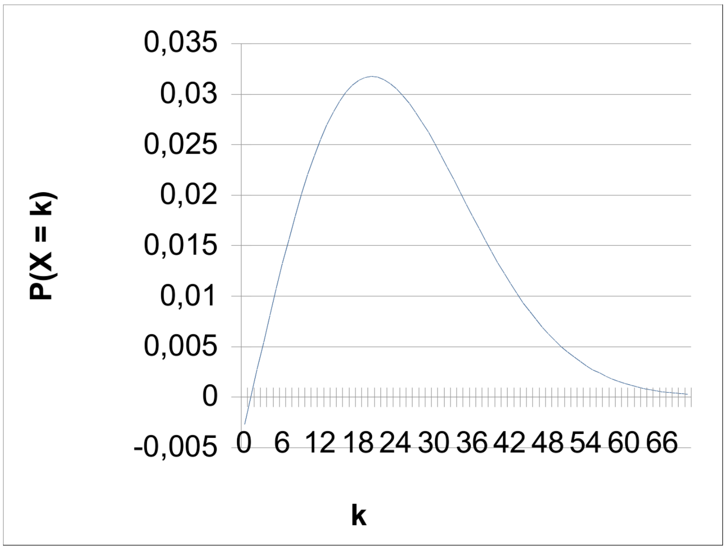

Abbildung 3: Wahrscheinlichkeit für die 1. Kollision bei der k. Person

α	q_α Erde	q_α Mars	q_α Jupiter
$\frac{1}{4}$	15,492	22,185	50,913
$\frac{1}{2}$	23,494	33,883	78,477
$\frac{3}{4}$	32,812	47,504	110,569