

Aufgabe 8

a)

Die Wahrscheinlichkeit, dass x einen Wert zwischen $\frac{1}{3}$ und $\frac{1}{2}$ annimmt ist gegeben durch

$$\int_{\frac{1}{3}}^{\frac{1}{2}} f(x) dx = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}. \quad (1)$$

b)

Die Wahrscheinlichkeit, dass irgendein exakter Wert angenommen wird ist 0. Dies ist einerseits klar, weil mit unendlich verschiedenen möglichen Werten sonst nie eine endliche Gesamtwahrscheinlichkeit erreicht werden könnte. Außerdem ist natürlich die integrierte Wahrscheinlichkeitsdichte

$$\int_a^a f(x) dx = 0. \quad (2)$$

c)

Da die verwendete Gleitkommazahldarstellung nicht genau angegeben ist, nehmen wir hierfür an, dass alle 23 Mantissenbits für die Darstellung von Zahlen zwischen 0 und 1 genutzt werden können und dass $\frac{1}{2}$ exakt darstellbar ist. In diesem Fall gibt es also insgesamt 2^{23} darstellbare Zahlen zwischen 0 und 1, und eine davon ist $\frac{1}{2}$ daher ist die Wahrscheinlichkeit genau $\frac{1}{2}$ (oder irgendeine andere der exakt darstellbaren Zahlen) zu ziehen 2^{-23} .

d)

Da $\frac{2}{3}$ eine Zahl mit unendlich vielen, periodischen Nachkommastellen ist, ist sie unabhängig von der gewählten Gleitkommazahldarstellung nie exakt darstellbar und daher ist die Wahrscheinlichkeit diese oder irgendeine andere nicht exakt darstellbare Zahl zu ziehen 0.

Aufgabe 9

a)

Gegeben ist der Linear-kongruenten Zufallszahlengenerator

$$x_n = (ax_{n-1} + b) \bmod m \quad (3)$$

mit $b = 3$ und $m = 1024$. Damit ist sofort klar, dass die maximale auftretende Periodenlänge nicht größer als $m = 1024$ sein kann. Wir bestimmen die Periodenlänge für alle a

von 0 bis 1024. Das Ergebnis ist in Abbildung 1 dargestellt. Zusätzlich ist in Abildung 2 der Bereich für a von 0 bis 32 höher aufgelöst dargestellt. Für die erzeugten Abbildungen wurde der Startwert $x_0 = 1$ verwendet, das Ergebnis ist unseren Tests nach jedoch unabhängig vom Startwert. Die maximale Periodenlänge ist wie erwartet $2^{10} = 1024$.

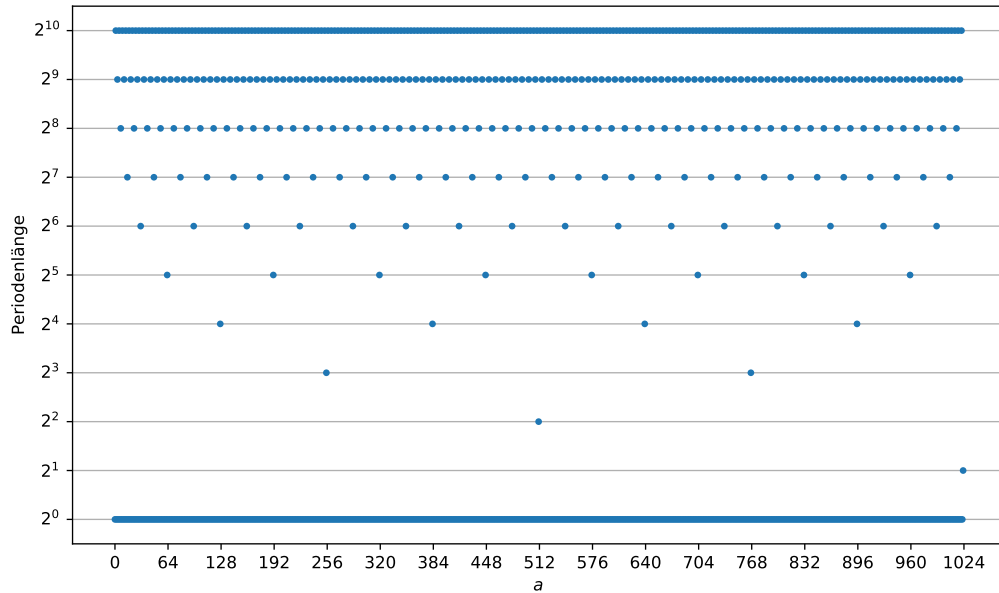


Abbildung 1: Periodenlänge für alle a von 0 bis 1024.

Das Verhalten lässt sich sehr gut durch die Regeln für gute linear-kongruente Generatoren [1, S. 7] erklären: Die Regeln 1 und 2 sind erfüllt. Die Regel 3 ist für alle ungeraden Zahlen erfüllt, da die Primfaktorzerlegung von m durch 2^{10} gegeben ist. Wie gut zu erkennen ist, ist die Periodenlänge für alle geraden a minimal eins, was die Regel bestätigt. Für alle a , für die gilt, dass $a - 1$ außerdem durch 4 teilbar ist (Regel 4) ist die Periodenlänge maximal. Leider haben wir gerade keine Zeit mehr uns länger Gedanken darüber zu machen, welchen Regeln die übrigen Werte folgen, aber offenbar bildet sich bei jedem $a + 1 = 2^n$ eine neue Ebene mit der Periodenlänge 2^{11-n} .

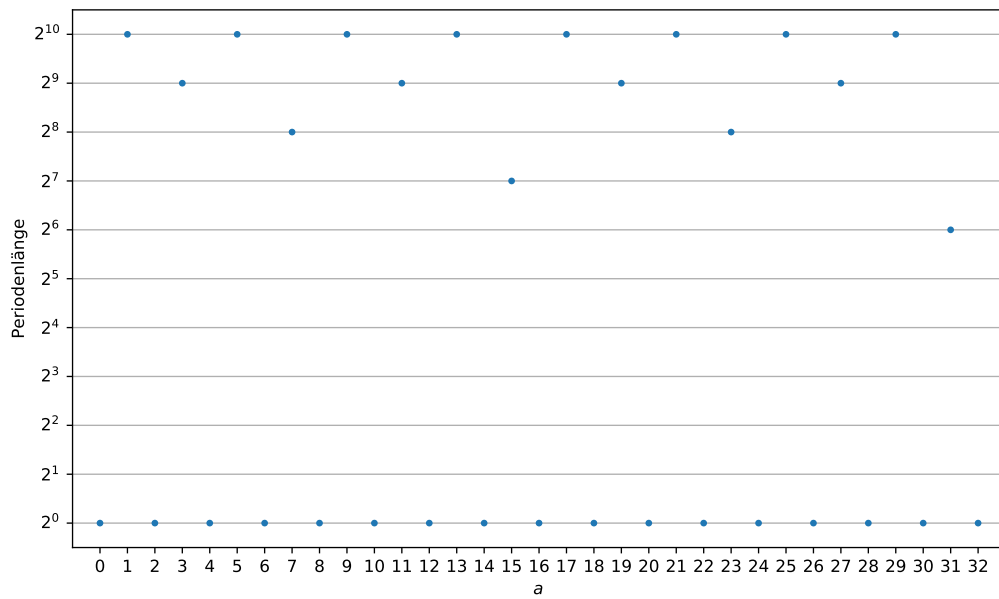


Abbildung 2: Periodenlänge für alle a von 0 bis 32.

c)

Das geforderte Histogramm ist durch Abbildung 3 gegeben. Das Ergebnis wurde für viele verschiedene Startwerte berechnet und scheint unabhängig von diesem zu sein. Es entspricht nicht den Anforderungen an einen guten Zufallszahlengenerator, da offenbar keine echte Gleichverteilung vorliegt, sondern manche Zahlen signifikant wahrscheinlicher sind als andere.

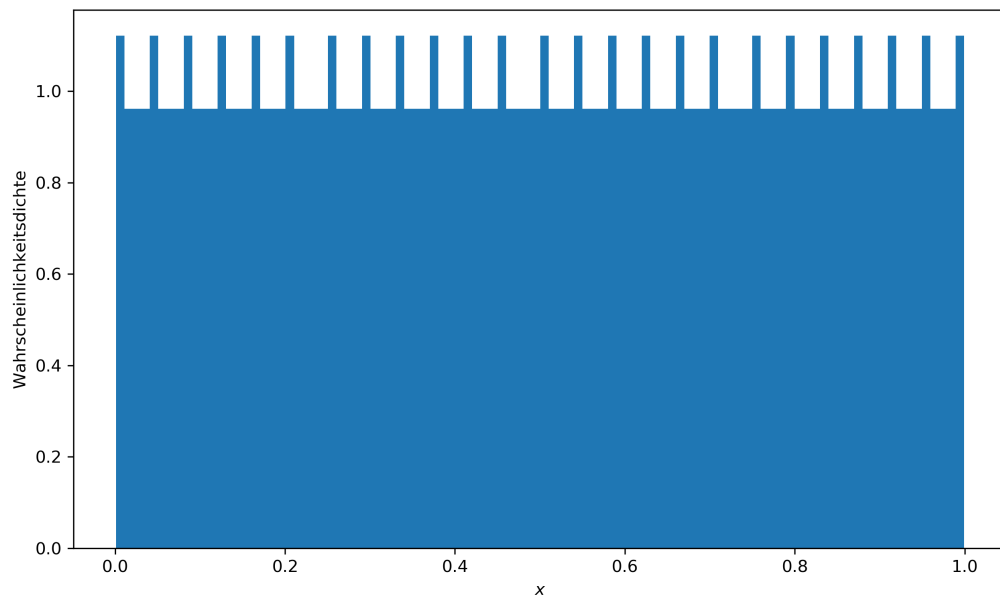


Abbildung 3: Histogramm für 10000 nach 3 erzeugte Zufallszahlen.

d)

Die geforderten Streudiagramme sind durch die Abbildungen 4 und 5 gegeben. Es ist klar zu erkennen, dass die Zahlen von einem schlechten Zufallszahlengenerator generiert wurden, da sich Ebenen bilden. Es besteht also eine vorhersagbare Korrelation zwischen aufeinanderfolgenden Zufallszahlen.

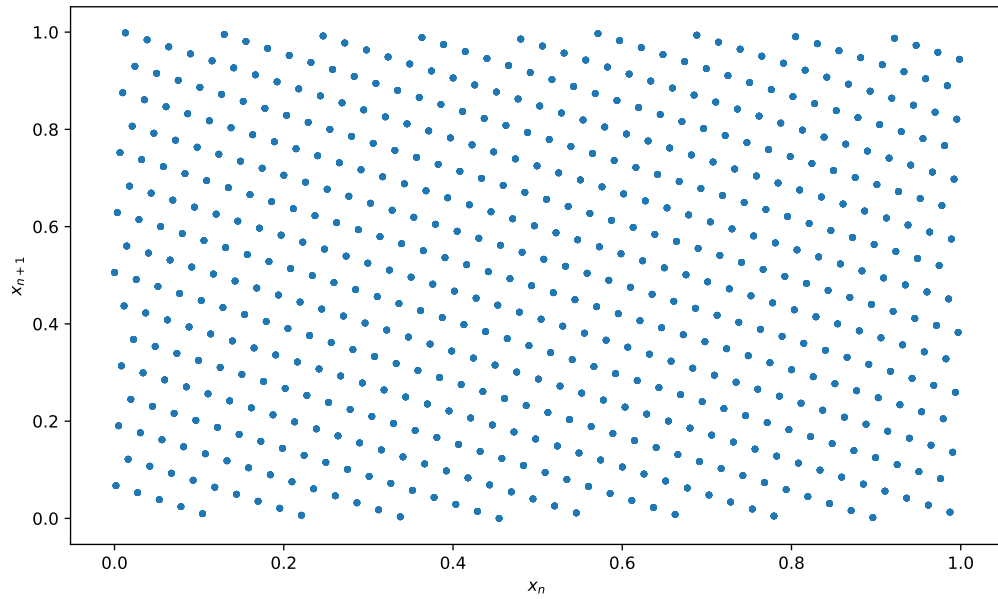


Abbildung 4: Zweidimensionales Streudiagramm für den Linear-kongruenten Zufallszahlengenerator.

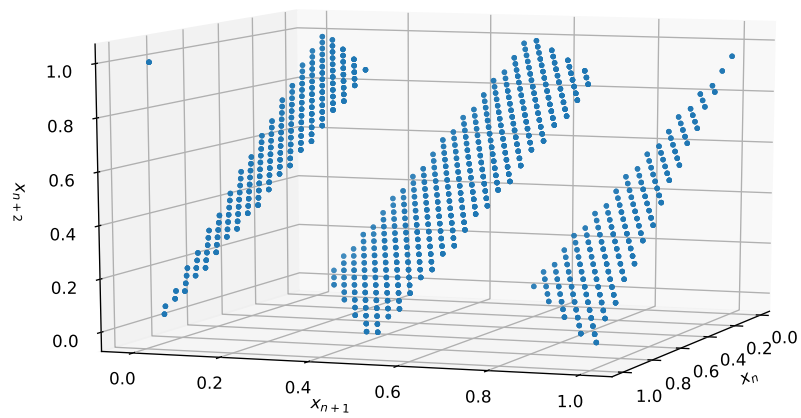


Abbildung 5: Dreidimensionales Streudiagramm für den Linear-kongruenten Zufallszahlengenerator.

e)

Die geforderten Streudiagramme sind durch die Abbildungen 6 und 7 gegeben. Es ist klar zu erkennen, dass die Zahlen von einem guten Zufallszahlengenerator generiert wurden, da sich keine erkennbaren Muster bilden. Es besteht also keine vorhersagbare Korrelation zwischen aufeinanderfolgenden Zufallszahlen.

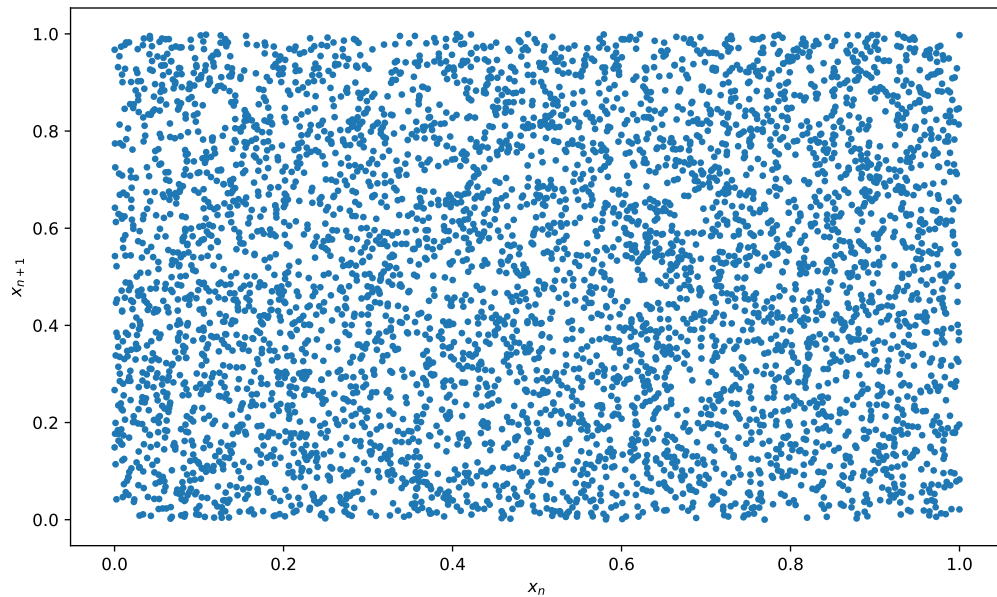


Abbildung 6: Zweidimensionales Streudiagramm für den NUMPY Zufallszahlengenerator.

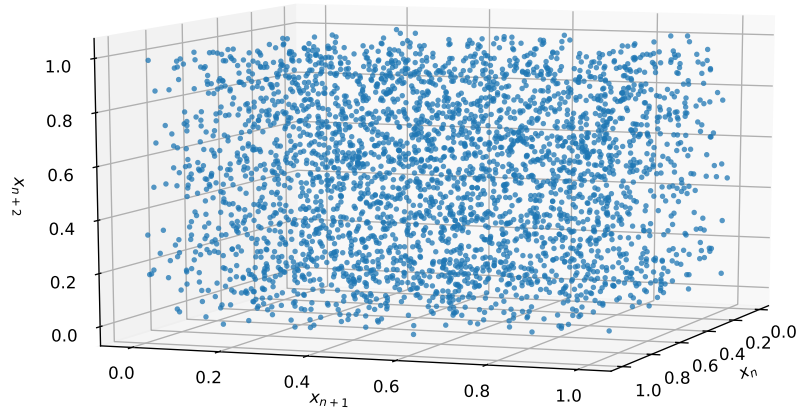


Abbildung 7: Dreidimensionales Streudiagramm für den Numpy Zufallszahlengenerator.

f)

Die vom Zufallszahlengenerator generierten Zahlen bestehen für alle Startwerte, die ungerade Vielfache von 8 sind, zu 0,16 % aus dem Wert $\frac{1}{2}$. Für alle anderen Startwerte wird nie der Wert $\frac{1}{2}$ generiert.

Aufgabe 10

Für die Gleichverteilung nutzen wir in der gesamten Aufgabe den Zufallszahlengenerator von Numpy. Nach Teilaufgabe a) verwenden wir diesen auch für die Generation von gleichverteilten Zufallszahlen mit anderen Grenzen als 0 und 1, obwohl dies streng genommen nach Aufgabenstellung nicht erlaubt ist, es macht aber keinen Unterschied und ist so vermutlich etwas schneller in der Ausführung. Es könnte auch überall im Code einfach die Numpy Funktion durch unsere selbst implementierte Variante ersetzt werden.

Wir gehen davon aus, dass in den Teilaufgaben b) bis d) mit f gewünschte Dichtefunktionen, und keine kumulierten Wahrscheinlichkeitsfunktionen gemeint sind.

a)

Gleichverteilte Zahlen zwischen x_{\min} und x_{\max} lassen sich einfach generieren, indem die zwischen 0 und 1 gleichverteilten Zahlen z nach

$$x = (x_{\max} - x_{\min})z + x_{\min} \quad (4)$$

transformiert werden. Ein Histogramm einer auf diese Weise mit den Grenzen $x_{\min} = 2$ und $x_{\max} = 5$ generierten Verteilung ist zur Verifikation in Abbildung 8 dargestellt.

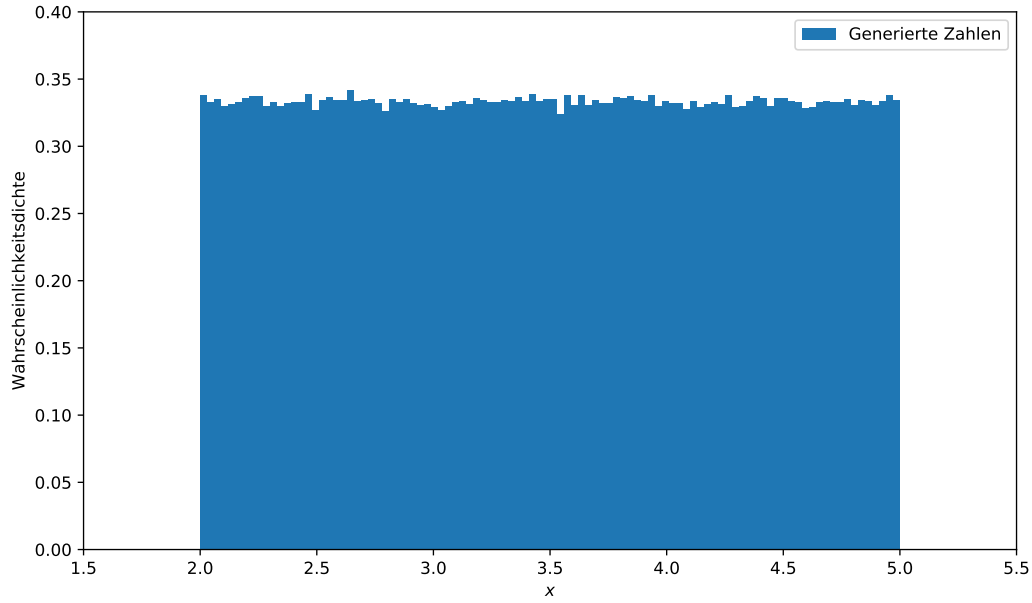


Abbildung 8: Histogramm der transformierten Gleichverteilung mit den Grenzen $x_{\min} = 2$ und $x_{\max} = 5$.

b)

Die gewünschte Dichtefunktion

$$f(t) = N \exp\left(-\frac{t}{\tau}\right), \quad t \in [0, \infty) \quad (5)$$

lässt sich integrieren. Daher kann zur Erzeugung dieser Verteilung die Transformationsmethode verwendet werden. Für die Verteilungsfunktion ergibt sich

$$F(t) = \int_0^t N \exp\left(-\frac{x}{\tau}\right) dx = 1 - \frac{\exp\left(-\frac{t}{\tau}\right)}{\tau}, \quad (6)$$

mit der Normierungskonstante $N = \tau^{-1}$. Diese kann nun invertiert werden:

$$z = F(x) = 1 - \frac{\exp\left(-\frac{x}{\tau}\right)}{\tau} \quad (7)$$

$$\Leftrightarrow (1 - z)\tau = \exp\left(-\frac{x}{\tau}\right) \quad (8)$$

$$\Rightarrow x = -\tau \log((1 - z)\tau), \quad z \in \left[1 - \frac{1}{\tau}, 1\right]. \quad (9)$$

Wenn z gleichverteilte Zufallszahlen mit den angegebenen Grenzen sind, dann sind die so transformierten x mit der gewünschten Dichte f verteilt. z kann hier natürlich wieder gemäß (4) erzeugt werden, diese Transformation ließe sich auch direkt hier einsetzen. Zur Verifikation erstellen wir auf diese Weise ein Histogramm mit $\tau = 1$, dieses ist in Abbildung 9 dargestellt.

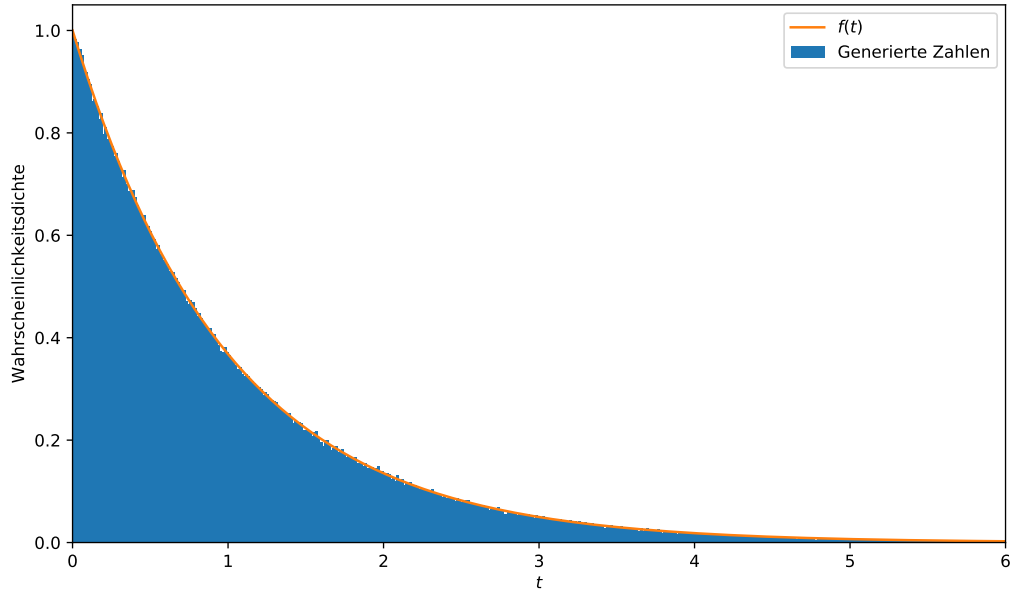


Abbildung 9: Histogramm der generierten Verteilung aus Aufgabe 10b).

c)

Die gewünschte Dichtefunktion

$$f(x) = Nx^{-n}, \quad x \in [x_{\min}, x_{\max}] \quad (10)$$

mit $n \geq 2$ lässt sich integrieren. Daher kann zur Erzeugung dieser Verteilung erneut die Transformationsmethode verwendet werden. Für die Verteilungsfunktion ergibt sich

$$F(x) = \int_{x_{\min}}^x f(x') dx' = \frac{N}{1-n} (x^{1-n} - x_{\min}^{1-n}), \quad x \in [x_{\min}, x_{\max}] \quad (11)$$

mit der Normierungskonstante

$$N = \frac{1-n}{x_{\max}^{1-n} - x_{\min}^{1-n}}. \quad (12)$$

Diese kann nun invertiert werden:

$$z = F(x) \quad (13)$$

$$\Rightarrow x = \left(z(x_{\max}^{1-n} - x_{\min}^{1-n}) + x_{\min}^{1-n} \right)^{\frac{1}{1-n}}, \quad z \in [0, 1]. \quad (14)$$

Wenn z gleichverteilte Zufallszahlen zwischen 0 und 1 sind, dann sind die so transformierten x mit der gewünschten Dichte f verteilt. Es ist zu erkennen, dass sich hier von selbst die Transformation (4) ergeben hat. Zur Verifikation erstellen wir auf diese Weise ein Histogramm mit $n = 2$ und den Grenzen $x_{\min} = 1$ und $x_{\max} = 5$, dieses ist in Abbildung 10 dargestellt.

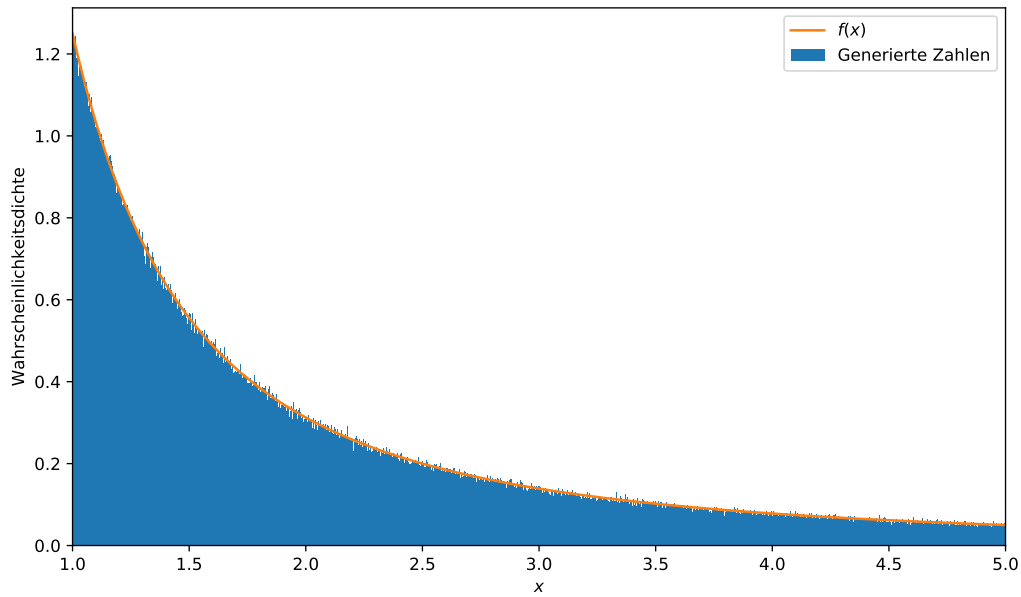


Abbildung 10: Histogramm der generierten Verteilung aus Aufgabe 10c).

d)

Die Wahrscheinlichkeitsdichte der CAUCHY-Verteilung

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}, \quad x \in (-\infty, \infty) \quad (15)$$

ist bereits normiert und lässt sich integrieren. Daher verwenden wir wieder die Transformationsmethode. Die Verteilungsfunktion lautet

$$F(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(x') dx' = \frac{\arctan x}{\pi} + \frac{1}{2}. \quad (16)$$

Invertieren der Verteilungsfunktion ergibt

$$F(x) = z \quad (17)$$

$$\Rightarrow x = -\cot(\pi z), \quad z \in [0, 1] \quad (18)$$

Wenn z gleichverteilte Zufallszahlen zwischen 0 und 1 sind, dann sind die so transformierten x CAUCHY-verteilt. Zur Verifikation erstellen wir auf diese Weise ein Histogramm, dieses ist in Abbildung 11 dargestellt.

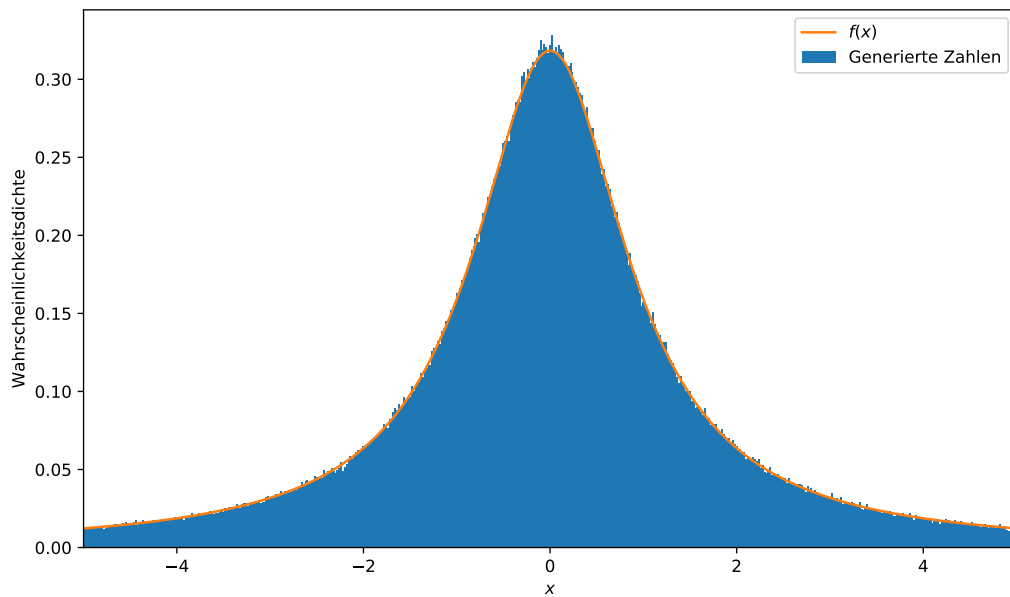


Abbildung 11: Histogramm der generierten Verteilung aus Aufgabe 10d).

e)

Es gibt verschiedene Möglichkeiten Zufallszahlen nach einer empirischen Verteilung zu erzeugen. Am einfachsten wäre es vermutlich, die Rückweisungsmethode zu verwenden. Diese ist allerdings sehr ineffizient und deswegen langsam. Stattdessen berechnen wir aus den gegebenen Histogramm Daten die kumulierten Wahrscheinlichkeiten und interpolieren daraus die Inverse der Verteilungsfunktion. Anwenden dieser auf gleichverteilte Zufallszahlen liefert Zahlen, die nach der empirischen Verteilung verteilt sein sollten. Zur Verifikation erstellen wir erneut ein Histogramm, dieses ist zusammen mit dem ursprünglich gegebenen Histogramm in Abbildung 12 dargestellt.

Unsere Vorgehensweise scheint relativ gut zu funktionieren, allerdings scheint es so, als wären die Wahrscheinlichkeiten bei aufsteigenden Flanken etwas zu hoch, und bei abfallenden etwas zu niedrig. Dies kommt vermutlich dadurch zu Stande, dass die Stützstellen in den Mitten der Bins liegen und so beim Bilden der kumulierten Wahrscheinlichkeiten bei steigenden Flanken immer bereits ein halber Bin zu viel, und bei fallenden ein halber zu wenig in den Stützpunkt einfließt. Außerdem ist zu sagen, dass nur sehr wenig Daten gegeben sind, und die zu Grunde liegende Verteilung daher sowieso nur sehr ungenau bekannt ist.

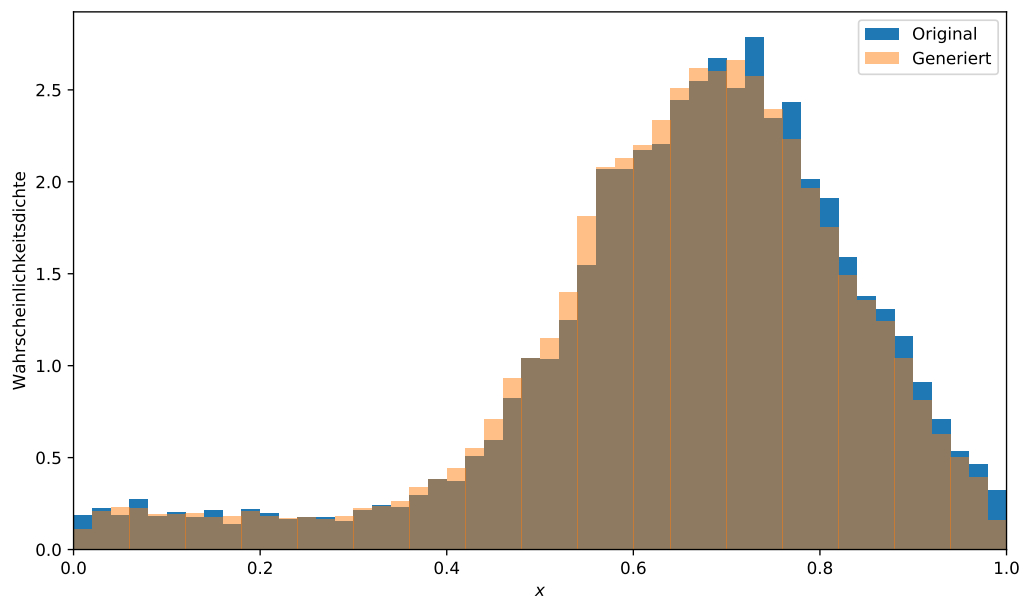


Abbildung 12: Darstellung des mittels interpolierter Verteilungsfunktion generierten Histogramms zusammen mit dem ursprünglichen Histogramm aus Aufgabe 10e)

Literatur

- [1] Wolfgang Rhode *Generation von Zufallszahlen* https://moodle.tu-dortmund.de/pluginfile.php/618120/mod_resource/content/1/Generation%20Zufallszahlen.pdf