

KOMMS Reports Nr. 8 (2019)

Reports zur Mathematischen Modellierung
in MINT-Projekten in der Schule



Numerische Simulationen mit Tabellenkalkulationsprogrammen

Patrick Capraro



Zusammenfassung:

Im folgenden wird an einigen Beispielen gezeigt, wie man numerische Simulationen in Tabellenkalkulationsprogrammen (hier speziell in Excel) erzeugen kann. Diese können beispielsweise im Kontext von mathematischer Modellierung verwendet werden.

Die Beispiele umfassen:

- Ein Modell zur Ausbreitung von Krankheiten
- Die Flugkurve eines Fußballs unter Berücksichtigung von Luftreibung
- Eine Monte-Carlo-Simulation zur experimentellen Bestimmung von π
- Eine Monte-Carlo-Simulation eines gemischten Kartenstapels
- Modellierung von Benzinpreisen mit einem Preistrend und Rauschen

Inhaltsverzeichnis

| | |
|--|-----------|
| 1 Grundlagen | 2 |
| 1.1 Formeln anwenden und auf Zellen verweisen | 2 |
| 1.2 Automatisches Weiterzählen der Zellenindizes | 2 |
| 1.3 Konstanten | 3 |
| 2 Zufallszahlen | 4 |
| 2.1 Zufallszahlen generieren und auf Bereiche anpassen | 4 |
| 2.1.1 Ganze Zahlen erzeugen | 4 |
| 2.2 IF und ELSE | 5 |
| 2.3 Zufallszahlen mit Wahrscheinlichkeitsdichte | 5 |
| 2.3.1 Binomialverteilung | 5 |
| 2.3.2 Normalverteilung | 5 |
| 3 Rekursive Folgen, Anfangswertprobleme und Schrittverfahren | 6 |
| 3.1 Krankheitsausbreitung | 6 |
| 3.1.1 Einfache Grundannahmen | 6 |
| 3.1.2 Genesung und Immunität | 6 |
| 3.2 Erweiterungen des Modells | 8 |
| 3.3 Schräger Wurf mit Luftreibung | 9 |
| 3.3.1 Schrittverfahren | 9 |
| 3.3.2 Ein mathematisches Modell für den fliegenden Fußball | 9 |
| 3.3.3 Ergänzung: Nulldurchgang der y -Koordinate | 10 |
| 4 Monte-Carlo-Simulationen | 12 |
| 4.1 Bestimmung der Zahl π | 12 |
| 4.2 Kartenmischen | 12 |
| 5 Deterministische Daten mit Rauschen | 14 |
| 5.1 Modellierung von Benzinpreisen | 14 |

1 Grundlagen

1.1 Formeln anwenden und auf Zellen verweisen

Um den Wert einer Zelle durch eine Formel berechnen lassen, wird zunächst ein Gleichheitszeichen eingegeben, gefolgt von den Größen, die berechnet werden:

$$=3,1415/2$$

Man kann dabei auch, statt Zahlen einzugeben, auf andere Zellen verweisen:

$$=(A2*B2)/2+1$$

Schließlich lassen sich auch eine Vielzahl von mathematischen Funktionen verwenden:

$$=WURZEL(2)*SIN(3,1415/2)$$

Zum Auswählen der passenden Funktion finden Sie bei den meisten Programmen Auswahllisten in den entsprechenden Menüs. Aber auch eine Onlinesuche kann zielführend sein, um den passenden Funktionsnamen zu finden.

Achtung mit den Funktionsparametern

Manche Funktionen benötigen mehrere Parameter, auch wenn man das nicht unbedingt vermuten würde. Die Funktion, die Zahlen rundet, nimmt beispielsweise 2 Parameter: Die Zahl, die gerundet werden soll, und die Anzahl der Nachkommastellen, z.B.

$$=RUNDEN(5,73;1)$$

Die Parameter werden durch einen Strichpunkt getrennt (um Verwechslungen mit dem Dezimalkomma zu vermeiden).

1.2 Automatisches Weiterzählen der Zellenindizes

Zelleninhalte – insbesondere auch die eingebetteten Formeln – lassen sich bequem auf andere Zellen in der selben Zeile oder Spalte übertragen: einfach eine Zelle auswählen und dann auf die rechte untere Ecke klicken. Hält man die Ecke geklickt und zieht den Mauszeiger über die Zeile bzw. Spalte, dann werden die Inhalte auf die anderen Zellen übertragen (siehe Abbildung 1).

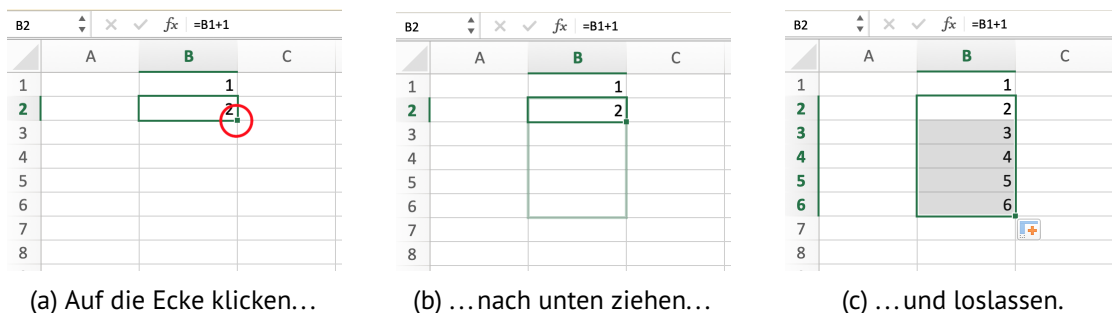


Abbildung 1: Zelleninhalt auf andere Zellen in der gleichen Spalte übertragen.

Das Praktische dabei: wenn wir in einer Formel auf Zellen verweisen, werden deren Indizes automatisch weitergezählt. Verweisen wir auf die Zelle A3 und kopieren den Inhalt in die nächste Zelle in der selben Spalte, dann wird dort auf A4 verwiesen. Kopieren wir den Inhalt

auf die nächste Zelle in der selben Zeile, dann wird dort auf B3 verwiesen. Damit lassen sich rekursiv definierte Größen spielend leicht erstellen.

Automatisches Weiterzählen verhindern

Wenn wir das automatische Weiterzählen verhindern wollen, dann müssen wir beim Verweis auf eine Zelle ein \$ verwenden: =D\$4 verhindert, dass die Zeilen weitergezählt werden. Mit =\$D4 wird verhindert, dass die Spalten weitergezählt werden. Mit =\$D\$4 wird immer auf die selbe Zelle verwiesen.

Beispiel: Fibonacci-Zahlen

Mit dieser Methode können wir mit wenigen Eingaben die Fibonacci-Folge erzeugen. Wir brauchen zunächst 3 Eingaben: In den ersten beiden Zellen zwei Einsen als Startwerte und danach die Rekursionsformel, die wir mit dem obigen Trick kopieren.

| | A | B |
|----|-------------|--------|
| 1 | Folgenindex | Wert |
| 2 | 1 | 1 |
| 3 | 2 | 1 |
| 4 | 3 | =B2+B3 |
| 5 | 4 | |
| 6 | 5 | |
| 7 | 6 | |
| 8 | 7 | |
| 9 | 8 | |
| 10 | 9 | |
| 11 | 10 | |
| 12 | 11 | |
| 13 | 12 | |

(a) Startwerte und Rekursionsformel...

| | A | B |
|----|-------------|------|
| 1 | Folgenindex | Wert |
| 2 | 1 | 1 |
| 3 | 2 | 1 |
| 4 | 3 | 2 |
| 5 | 4 | 3 |
| 6 | 5 | 5 |
| 7 | 6 | 8 |
| 8 | 7 | 13 |
| 9 | 8 | 21 |
| 10 | 9 | 34 |
| 11 | 10 | 55 |
| 12 | 11 | 89 |
| 13 | 12 | 144 |

(b) ...und fertig!

Abbildung 2: Eine rekursive Folge erzeugen: Die Fibonaccifolge.

1.3 Konstanten

Nützlich ist es auch, Konstanten festzulegen. Dabei kann für den Wert einer Zelle ein individueller Name gewählt werden (z.B. ϵ). In einer Formel kann man dann, statt auf den Zellenindex, auf die Konstante verweisen. Das spart ggf. Tipparbeit und sorgt dafür, dass die Formeln leichter lesbar und verständlich werden.

| | A | B |
|----|-----------------|---|
| 1 | Wachstumsfaktor | |
| 2 | 5 | |
| 3 | | |
| 4 | | |
| 5 | | |
| 6 | | |
| 7 | | |
| 8 | | |
| 9 | | |
| 10 | | |

(a) Namen einer Konstanten festlegen...

| | A | B |
|----|-----------------|---|
| 1 | Wachstumsfaktor | |
| 2 | | 5 |
| 3 | | |
| 4 | Zeitschritt | |
| 5 | | 1 |
| 6 | | 2 |
| 7 | | 3 |
| 8 | | 4 |
| 9 | | 5 |
| 10 | | 6 |

(b) ... und einfach in die Formel einbinden.

Abbildung 3: Wenn man Konstanten an einer Stelle definiert und später darauf referenziert, lassen sich Parameter später leichter anpassen.

2 Zufallszahlen

2.1 Zufallszahlen generieren und auf Bereiche anpassen

Die einfachste Möglichkeit, Zufallszahlen zu generieren, verwendet die Funktion

`=ZUFALLSZAHL()`

Dabei wird eine Kommazahl im Intervall $[0,1]$ erzeugt. Selbstverständlich lässt sich diese Methode nutzen, um Zufallszahlen aus einem beliebigen Intervall $[a, b]$ zu erhalten. Man muss lediglich das Intervall strecken und verschieben:

`=ZUFALLSZAHL() * (b-a) + a`

Unter Umständen ist es erforderlich, die Zahlen auf eine bestimmte Anzahl an Nachkommastellen zu runden. Hier aber Vorsicht: Die Funktion

`=RUNDEN(ZUFALLSZAHL() * 5; 0) + 1`

erzeugt Zufallszahlen von 1 bis 6, ist aber nicht geeignet, um einen Würfel zu simulieren. Nur die Werte im Intervall $[0, \frac{1}{2})$ werden zur Null abgerundet und erzeugen (wegen +1) eine 1 als Ergebnis, während wir eine 2 als Ergebnis erhalten, wenn der Zufallszahlengenerator die Ergebnisse $[\frac{1}{2}, \frac{3}{2})$ erzeugt. Die Zahlen sind also nicht gleich wahrscheinlich. Hier kann man allerdings auch folgendermaßen verfahren:

`=AUFRUNDEN(ZUFALLSZAHL() * 6; 0) + 1`

2.1.1 Ganze Zahlen erzeugen

Will man gleichverteilte ganze Zahlen erzeugen, dann bietet sich eher folgende Funktion an:

`=ZUFALLSBEREICH(a; b)`

Diese Erzeugt ganze Zahlen im Intervall $[a, b]$.

2.2 IF und ELSE

Manchmal will man zufällige Effekte verwenden, die sich nicht ohne weiteres durch Zufallszahlen aus einem bestimmten Intervall darstellen lassen, z.B.

- Wechselndes Vorzeichen in einer Formel (d.h. ein Faktor, der entweder 1 oder -1 ist)
- Eine Zufallsvariable, die entweder den Wert 0 oder 500 annimmt
- Eine Größe, die von der Beschaffenheit anderer Größen abhängt

Manchmal lassen sich solche Dinge auf den obigen Fall zurückführen, z.B. wird

=ZUFALLSBEREICH(0;1)*2-1

entweder 1 oder -1. Ganz ähnlich ist

=ZUFALLSBEREICH(0;1)*500

entweder 0 oder 500. Manchmal ist es aber einfacher oder unumgänglich, die WENN-Funktion zu verwenden. Programmieraffine kennen das im Zusammenhang mit IF-ELSE-Verzweigungen

Die WENN-Funktion hat 3 Argumente: eine logische Aussage A (das, was geprüft wird), einen Wert x , falls die Aussage A wahr ist, und einen Wert y , falls die Aussage falsch ist. Z.B. ist

=WENN(ZUFALLSBEREICH(0;1)=1;500;0)

auch ein Beispiel für eine Zufallsvariable, die die Werte 0 und 500 annehmen kann. Außerdem kann man bei der zu prüfenden Bedingung auch auf andere Zellen verweisen. Beispielsweise

WENN(A3>5;1;-1)

Wenn also der Wert in der Zelle A3 größer als 5 ist, dann erhalten wir den Wert 1, andernfalls -1. Damit könnten wir z.B. das Vorzeichen in einer Formel abhängig von einem Wert in einer anderen Zelle machen.

2.3 Zufallszahlen mit Wahrscheinlichkeitsdichte

Bisher hatten wir es nur mit Zufallszahlen zu tun, die gleichverteilt waren (außer in dem Beispiel, in dem ungleichmäßig gerundet wurde). Um andere Wahrscheinlichkeitsverteilungen zu nutzen, muss man ggf. etwas kreativ sein, falls diese nicht automatisch bereitgestellt werden.

2.3.1 Binomialverteilung

Eine $B(n, p, k)$ -verteilte Zufallsvariable erzeugt man mit dem Befehl

=BINOM. VERT(n;p;k;FALSE)

Mit

=BINOM. VERT(n;p;k;TRUE)

erhält man die entsprechende kumulierte Verteilung.

2.3.2 Normalverteilung

Um eine Normalverteilte Zufallsvariable zu erhalten, nutzen wir die Funktion NORM.INV, also die Inverse der Verteilungsfunktion zur Normalverteilung. Dazu erzeugen wir eine

Zufallszahl im Intervall $[0,1]$ und berechnen dann den inversen Funktionswert:

$=\text{NORM.INV}(\text{ZUFALLSZAHL}(); m; s)$

Hier sind m der Erwartungswert und s die Standardabweichung der Verteilung.

3 Rekursive Folgen, Anfangswertprobleme und Schrittverfahren

Wir haben bereits gesehen, dass sich Tabellenkalkulation sehr gut eignet, um rekursiv Daten zu generieren. Damit lassen sich eine Vielzahl von Problemen lösen, die nach dem Schema eines Anfangswertproblems gestrickt sind. Beispiele dafür sind **Wachstumsprozesse** (Populationmodelle, Krankheitsausbreitung, Kapitalerträge, ...) oder aber Differentialgleichungen, die mit sogenannten **Schrittverfahren** gelöst werden können.

3.1 Krankheitsausbreitung

Die Ausbreitung einer Krankheit ist ein einfaches Beispiel für ein alltägliches Phänomen, das sich mit simplen mathematischen Mitteln modellieren lässt. Man kann dabei mit schlichten Annahmen beginnen und das Modell schrittweise um weitere Elemente ergänzen.

3.1.1 Einfache Grundannahmen

Ein einfaches Modell könnte darin bestehen, dass man eine Gesamtpopulation festlegt (z.B. eine Großstadt mit 100 000 Personen). Dann braucht man einen Startwert, also die Anzahl der Kranken, und eine Regel, die besagt, wie viele Leute sich jeden Tag zusätzlich infizieren. Man könnte mit folgendem Modell einsteigen:

- Gesamtpopulation: 100 000 Personen
- Tag 1: 5 sind krank
- Jeden Tag stecken 50% der Kranken eine Person an

Dazu brauchen wir eine Spalte, in der wir in der ersten Zelle den Startwert eintragen. In der zweiten Zelle steht dann die Funktion

$=B5+B5*0,5$

wobei der Wachstumsfaktor 0,5 der Einfachheit halber in eine Konstante gepackt werden kann. Da der Faktor keine ganze Zahl ist, erhalten wir als Zwischenergebnisse ebenfalls wieder rationale Zahlen. Je nach dem, wie das Modell interpretiert werden soll, kann es hier sinnvoll sein, die Zahlen zu runden. Wir passen also die Funktion entsprechend an (siehe Abbildung 4):

$=\text{RUNDEN}(B5+B5*0,5; 0)$

Dieses Modell führt zu einem exponentiellen Wachstum, so dass schon nach wenigen Tagen (in diesem Fall 20) die Zahl der Kranken so groß ist, dass sie die Gesamtpopulation übersteigt. Es ist also notwendig, das Modell zu überarbeiten.

3.1.2 Genesung und Immunität

Um folgende Ideen könnte man das Modell erweitern:

- Kranke werden irgendwann wieder gesund

RUNDEN \downarrow \times \checkmark f_x $=\text{RUNDEN}(B5+B5*f;0)$

| | A | B | C | D |
|----|------------------|-----------------------------|-----------------|---|
| 1 | Gesamtpopulation | | Wachstumsfaktor | |
| 2 | 100000 | | 0,5 | |
| 3 | | | | |
| 4 | Tag | krank | | |
| 5 | 1 | 5 | | |
| 6 | 2 | $=\text{RUNDEN}(B5+B5*f;0)$ | | |
| 7 | 3 | 12 | | |
| 8 | 4 | 18 | | |
| 9 | 5 | 27 | | |
| 10 | 6 | 41 | | |
| 11 | 7 | 62 | | |
| 12 | 8 | 93 | | |
| 13 | 9 | 140 | | |
| 14 | 10 | 210 | | |

Abbildung 4: Krankheitsausbreitung

- Manche Leute sind immun gegen die Krankheit
- Wenn sich Leute neu infizieren, dann sollte deren Anzahl nicht nur von der Anzahl der Kranken, sondern auch von der Anzahl der Gesunden abhängen

Die Umsetzung dieser Gedanken kann auf unterschiedliche Weise erfolgen. Hier ein Vorschlag:

- Zu Beginn sind 2000 Menschen immun gegen die Krankheit
- Kranke werden nach 5 Tagen wieder gesund und sind fortan immun
- Jeden Tag stecken 50% der Kranken eine Person an, wobei wir diesen Faktor zusätzlich mit dem Prozentsatz der Leute verrechnen müssen, die anfällig (also weder krank noch immun) sind

Wir ergänzen das alte Modell in zwei Schritten: zunächst erweitern wir die Tabelle um zwei Spalten: Leute die **krank** bzw. **immun** sind. Die Anzahl der Neuinfizierten wird um einen Faktor ergänzt, der sowohl die Gesamtpopulation, als auch die Summe den Leuten beinhaltet, die krank oder immun sind (siehe Abbildung 5).

Anschließend findet der Prozess der Genesung statt: Ab Tag 6 ziehen wir von der Anzahl der Kranken noch die Anzahl der Leute ab, die wieder gesund werden: Dies sind die Leute, die sich vor 5 Tagen infiziert haben (die also 5 Zeilen weiter oben bei der Anzahl der neuinfizierten aufgelistet waren). Diese Leute werden nicht einfach nur gesund, sondern immun.

Je nachdem, wie man die Werte in die Tabelle einsetzt, kann es hier nützlich sein, den Tag 0 einzuführen, an dem noch niemand krank ist, wo es aber 5 Neuinfizierte gibt.

| RUNDEN | | | |
|--------|------------------|-----------------|-------|
| A | B | C | D |
| 1 | Gesamtpopulation | Wachstumsfaktor | |
| 2 | 100000 | 0,5 | |
| 3 | | | |
| 4 | Tag | krank | immun |
| 5 | 1 | 5 | 2000 |
| 6 | 2 | | |
| 7 | 3 | | |
| 8 | 4 | | |
| 9 | 5 | | |
| 10 | 6 | | |
| 11 | 7 | | |
| 12 | 8 | | |
| 13 | 9 | | |
| 14 | 10 | | |

| RUNDEN | | | |
|--------|------------------|-----------------|-------|
| A | B | C | D |
| 1 | Gesamtpopulation | Wachstumsfaktor | |
| 2 | 100000 | 0,5 | |
| 3 | | | |
| 4 | Tag | krank | immun |
| 5 | 1 | 5 | 2000 |
| 6 | 2 | | |
| 7 | 3 | 10 | 2000 |
| 8 | 4 | 15 | 2000 |
| 9 | 5 | 22 | 2000 |
| 10 | 6 | 33 | 2000 |
| 11 | 7 | 49 | 2000 |
| 12 | 8 | 73 | 2000 |
| 13 | 9 | 109 | 2000 |
| 14 | 10 | 162 | 2000 |

Abbildung 5: Die Zahl der neu infizierten wird durch einen Faktor ergänzt, der Wert in eine eigene Spalte geschrieben. Genesung findet noch nicht statt.

| RUNDEN | | | |
|--------|------------------|-----------------|-------|
| A | B | C | D |
| 1 | Gesamtpopulation | Wachstumsfaktor | |
| 2 | 100000 | 0,5 | |
| 3 | | | |
| 4 | Tag | krank | immun |
| 5 | 0 | 0 | 2000 |
| 6 | 1 | 5 | 2000 |
| 7 | 2 | 7 | 2000 |
| 8 | 3 | 10 | 2000 |
| 9 | 4 | 15 | 2000 |
| 10 | 5 | 22 | 2000 |
| 11 | 6 | | |
| 12 | 7 | 42 | 2000 |
| 13 | 8 | 63 | 2000 |
| 14 | 9 | 94 | 2000 |
| 15 | 10 | 140 | 2000 |

| RUNDEN | | | |
|--------|------------------|-----------------|-------|
| A | B | C | D |
| 1 | Gesamtpopulation | Wachstumsfaktor | |
| 2 | 100000 | 0,5 | |
| 3 | | | |
| 4 | Tag | krank | immun |
| 5 | 0 | 0 | 2000 |
| 6 | 1 | 5 | 2000 |
| 7 | 2 | 7 | 2000 |
| 8 | 3 | 10 | 2000 |
| 9 | 4 | 15 | 2000 |
| 10 | 5 | 22 | 2000 |
| 11 | 6 | 28 | 2007 |
| 12 | 7 | 42 | 2010 |
| 13 | 8 | 63 | 2015 |
| 14 | 9 | 94 | 2022 |
| 15 | 10 | 140 | 2033 |

Abbildung 6: Kranke werden nach 5 Tagen wieder gesund

3.2 Erweiterungen des Modells

Obwohl dieses Modell schon recht realistische Werte liefert, gibt es dennoch einige Probleme, die auftreten können. Möglicherweise sind einige dieser Annahmen auch unrealistisch. Hier ein paar Anregungen zum Weiterarbeiten:

- Durch das Einbeziehen der Frage, ob neuinfizierte vielleicht immun (oder bereits erkrankt) sind, hat man im Prinzip ein logistisches Wachstum, welches verhindern soll, dass die Zahl der Kranken die Gesamtpopulation übersteigt. Dennoch kann genau das passieren, da das Problem nicht kontinuierlich, sondern diskret formuliert ist. Unter Umständen muss man mit Funktionen wie MAX und MIN Maximal- und Minimalwerte für die Größen festlegen.
- Sobald die Krankheitswelle vorüber ist, hat man je nach Startwerten manchmal den Effekt, dass es über viele Tage hinweg gleich viele (aber sehr wenige) Neuerkrankungen gibt. In diesem Fall ergibt sich ein zeitweise stationärer Zustand. Das liegt daran, dass sich der Quotient aus Anfälligen zur Gesamtpopulation und auch die Gesamtzahl der Kranken nur sehr schwach ändert. Die Änderung spielt nach dem Runden auf ganze Zahlen keine Rolle mehr.
- Das Modell ist deterministisch. Unter Umständen könnte es interessanter (vielleicht

auch realistischer) sein, einige der Größen Zufallsabhängig zu machen. So könnte die Zahl der Neuinfizierten eine Zufallsvariable sein und auch die Dauer der Krankheit müsste nicht auf 5 Tage beschränkt sein.

3.3 Schräger Wurf mit Luftreibung

Als Beispiel eines Schrittverfahrens betrachten wir ein Problem, das in der vereinfachten Form im Physik- und Mathematikunterricht sehr geläufig ist: Der schräge Wurf wird gerne als Paradebeispiel für quadratische Funktionen verwendet. Dabei ist diese Sichtweise nur akzeptabel, wenn andere physikalische Effekte (wie etwa Reibung) vernachlässigbar klein sind. Im Falle eines fliegenden Fußballs ist das aber kaum noch haltbar, da das Gewicht des Balls in Relation zur Größe problematisch ist.

Will man einen schrägen Wurf mit Luftreibung mathematisch korrekt angehen, befindet man sich bereits im Themenfeld der Differentialgleichungen (die Beschleunigung ist abhängig von der momentanen Geschwindigkeit, d.h. die erste und zweite Ableitung sind miteinander gekoppelt). Damit ist die Aufgabe für die Schule völlig ungeeignet (sieht man einmal von Schülern ab, die sich kurz vor dem Abitur befinden und sich mit Differentialgleichungen beschäftigen haben).

Wenn man sie jedoch mit numerischen Methoden löst, dann lässt sich die Aufgabe bereits in der Mittelstufe behandeln.

3.3.1 Schrittverfahren

Der Grundgedanke von Schrittverfahren besteht darin, dass man die Zeit (oder eine andere Größe, von der die Funktionen abhängen) in kleine Schritte dt unterteilt. Entlang dieser kleinen Schritte nimmt man an, dass die Ableitungen der Funktion (d.h. hier die Geschwindigkeit, Beschleunigung, etc.) konstant sind. Man kann also mit einem Startwert (Zeit $t = 0$) starten und gemäß der physikalischen Gesetze die Funktionswerte der Ableitungen für den ersten Schritt $t = dt$ neu berechnen. Im zweiten Schritt berechnet man die Werte für $t = 2dt$ usw. Man macht dabei stets kleine Fehler, die jedoch vernachlässigbar sind, solange der Zeitschritt dt klein genug gewählt wurde¹.

3.3.2 Ein mathematisches Modell für den fliegenden Fußball

Zur Berechnung der Flugkurve soll folgendes Modell verwendet werden:

- Koordinatensystem: x -Achse: Vorwärtsrichtung, y -Achse: senkrecht nach oben
- Fußball wird gekickt und fliegt mit Anfangsgeschwindigkeit v_0 im Winkel α los
- Ball wird durch die Erdbeschleunigung g in negativer y -Richtung beschleunigt

¹Hier ist vorsicht geboten. Im Allgemeinen wird die Simulation besser (d.h. genauer), je kleiner der Zeitschritt gewählt wird. Dies geht auf Kosten der Rechenzeit (im Falle einer Tabellenkalkulation erkennt man die Kosten in Form der benötigten Tabelleneinträge). Allerdings gibt es eine untere Grenze für die Schrittweite. Wird diese Grenze unterschritten, dann wird die Genauigkeit des Ergebnisses wieder schlechter, obwohl der Rechenaufwand weiter wächst.

- Luftreibungskraft ist antiparallel zur Geschwindigkeit

$$F = 6\pi v\eta \quad (\text{Stokessche Reibung, nur bei laminarer Strömung})$$

$$F = C_w A \frac{1}{2} \rho v^2 \quad (\text{Reibungskraft für turbulente Strömungen})$$

- Weitere Faktoren, wie Eigenrotation des Balls, Wind, etc. werden vernachlässigt
- Masse: $m = 0.45 \text{ kg}$
- Radius: $r = 0.11 \text{ m}$
- Strömungswiderstandskoeffizient²: $C_w = 0.4$
- Luftdichte³: $1.29 \frac{\text{kg}}{\text{m}^3}$
- Reibungskraft für turbulente Strömungen:

$$F_w = C_w A \frac{1}{2} \rho v^2 = 0.009 \frac{\text{kg}}{\text{m}} v^2.$$

- Alle Einheiten werden in SI-Einheiten angegeben

Konstanten für die Tabellenkalkulation

- Abwurfwinkel: `alpha`
- Anfangsgeschwindigkeit: `v`
- Zeitschritt: `dt`

Startwerte

- x -Koordinate: 0
- y -Koordinate: 0
- Geschwindigkeit v_x in x -Richtung: `v*cos(alpha*3,1415/180)`
- Geschwindigkeit v_y in y -Richtung: `v*sin(alpha*3,1415/180)`

Zuwächse je Zeitschritt

- x -Koordinate: `vx * dt`
- y -Koordinate: `vy * dt`
- Geschwindigkeit in x -Richtung: `vx*ABS(vx)*0,009*dt/0,45`
- Geschwindigkeit in y -Richtung: `vy*ABS(vy)*0,009*dt/0,45-9,81*dt`

(hier sind mit v_x und v_y die jeweiligen Werte des vorherigen Zeitschritts gemeint, siehe Abbildung 7).

3.3.3 Ergänzung: Nulldurchgang der y -Koordinate

Sobald der Ball an Höhe verliert, wird er früher oder später wieder die Ausgangshöhe $y = 0$ erreichen und dann zwangsläufig negative y -Werte annehmen. Um das Modell an dieser Stelle trotzdem noch weiterzuspinnen, kann hier eine Reflexion eingebaut werden: Dazu müssen wir dieses Event lokalisieren und zu gegebener Zeit die Geschwindigkeit mit dem Wert -1 multiplizieren. Hier hilft die WENN-Funktion weiter. Man erhält ein charakteristisches Bild (Abbildung 8).

²Demtröder Experimentalphysik 1 (Mechanik und Wärme), Springer-Verlag, 3. Auflage, Abb. 8.36.

³<https://www.lernort-mint.de/alltagstoffe/stoffeigenschaften/luftdichte.html>.

ABS \updownarrow \times \checkmark f_x $=D4-0,009*D4*ABS(D4)*dt/0,45$

| | A | B | C | D | E | F |
|----|-------------|---|------|------|--------------------------------|-------|
| 1 | Fußball | | | v | alpha | dt |
| 2 | | | | | 20,00 | 30,00 |
| 3 | Zeitschritt | x | y | vx | vy | |
| 4 | | 1 | 0,00 | 0,00 | 17,32 | 10,00 |
| 5 | | 2 | 0,09 | 0,05 | $=D4-0,009*D4*ABS(D4)*dt/0,45$ | 9,94 |
| 6 | | 3 | 0,17 | 0,10 | 17,26 | 9,88 |
| 7 | | 4 | 0,26 | 0,15 | 17,23 | 9,82 |
| 8 | | 5 | 0,35 | 0,20 | 17,20 | 9,76 |
| 9 | | 6 | 0,43 | 0,25 | 17,17 | 9,71 |
| 10 | | 7 | 0,52 | 0,30 | 17,14 | 9,65 |
| 11 | | 8 | 0,60 | 0,34 | 17,11 | 9,59 |
| 12 | | 9 | 0,69 | 0,39 | 17,08 | 9,53 |

Abbildung 7: Ein schräg geschossener Fußball mit Luftwiderstand.

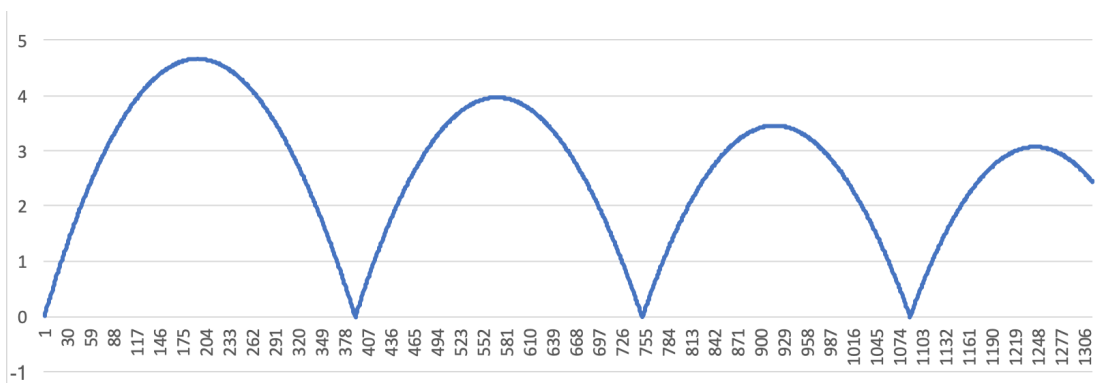


Abbildung 8: Der Fußball wird am Boden reflektiert und hüpft.

4 Monte-Carlo-Simulationen

Eine **Monte-Carlo-Simulation** ist ein Verfahren, das eine gesuchte Größe mit Hilfe von relativen Häufigkeiten bestimmen kann. Die Grundidee ist, dass ein Zufallsexperiment sehr häufig wiederholt wird. Das **Gesetz der großen Zahlen** besagt dann, dass die relative Häufigkeit eines bestimmten Ereignisses (stochastisch) gegen die tatsächliche Wahrscheinlichkeit des Ereignisses konvergiert.

Monte-Carlo-Simulationen können verwendet werden, um Wahrscheinlichkeitswerte für sehr komplexe Zufallsexperimente zu bestimmen, die sich nur noch sehr schwer exakt (z.B. durch Kombinatorik) berechnen lassen. Man kann sie jedoch auch verwenden, um Fragen zu beantworten, die gar nicht aus dem Bereich der Wahrscheinlichkeitsrechnung stammen.

4.1 Bestimmung der Zahl π

Ein sehr bekanntes Beispiel für eine Monte-Carlo-Simulation ist das Buffonsche Nadelproblem: Durch ein Zufallsexperiment (eine Nadel fällt zufällig auf ein Schachbrettmuster), lässt sich die Kreiszahl π bestimmen (wie oft schneidet die Nadel eine Kante des Schachbrettmusters).

Wir wollen hier ein einfacheres Experiment machen, das jedoch zum selben Ergebnis führt: Man lässt Punkte zufällig in das Quadrat $[-1, 1] \times [-1, 1]$ fallen. Wir zählen alle Punkte, die einen Abstand von höchstens 1 zum Ursprung haben (d.h. sie liegen im Inneren des Einheitskreises). Das Verhältnis $\frac{p}{n}$ von positiven Ereignissen p zur Gesamtzahl n der Durchführungen sollte gegen das Verhältnis $\frac{Q}{K}$ der Flächeninhalte vom Kreis $K = \pi$ zum Quadrat $Q = 4$ konvergieren. Multiplizieren wir die statistisch bestimmte relative Häufigkeit mit 4, dann erhalten wir einen Schätzwert für π .

| ZUFALLSZ... $f_x = \text{ZUFALLSZAHL}()*2-1$ | | | | SUMME $f_x = \text{WENN}(A2^2+B2^2<1;1;0)$ | | | |
|--|------------|--------------------|---|--|------------|---------------|------------------------|
| | A | B | C | | A | B | C |
| 1 | x | y | | 1 | x | y | im Kreis? |
| 2 | -0,7843121 | =ZUFALLSZAHL()*2-1 | | 2 | 0,0920109 | -0,5187393 | =WENN(A2^2+B2^2<1;1;0) |
| 3 | 0,58235549 | 0,898081908 | | 3 | 0,04033504 | -0,2183869 | 1 |
| 4 | 0,30227395 | -0,298761937 | | 4 | -0,8341546 | -0,1548372 | 1 |
| 5 | -0,6048125 | 0,575145894 | | 5 | 0,09153215 | -0,6443968 | 1 |
| 6 | -0,122887 | 0,673580501 | | 6 | 0,56932733 | 0,77218883 | 1 |
| 7 | -0,5151111 | 0,040108919 | | 7 | -0,9165143 | -0,6346994 | 0 |
| 8 | 0,77078409 | -0,869213408 | | 8 | 0,94377891 | -0,5086972 | 0 |
| 9 | 0,45801799 | 0,280323624 | | 9 | 0,81688749 | -0,126284 | 1 |
| 10 | 0,54254789 | 0,003502485 | | 10 | -0,8624179 | -0,9761111 | 0 |
| 11 | | | | 11 | | | |
| 12 | | | | 12 | | rel. Häufigke | 0,66666667 |

Abbildung 9: Wir erzeugen Zufallszahlen im Intervall $[-1,1]$ und testen, ob die Punkte im Kreis liegen. Die relative Häufigkeit wird mit $\text{SUMME}(C2:C10)/9$ berechnet.

4.2 Kartenmischen

Im vorherigen Beispiel haben wir gesehen, wie wir das Monte-Carlo-Verfahren benutzen können, um Probleme zu lösen, die eigentlich nichts mit Stochastik zu tun haben. Jetzt

wollen wir der folgenden Frage nachgehen: Wie wahrscheinlich ist es, dass in einem gemischten Kartenstapel (32 Karten, französisches Blatt) genau ein Paar Ass hintereinander liegt.

Dazu stellen wir uns vor, dass wir die Positionen der vier Ass im Kartenstapel kennen müssen. Andere Informationen sind irrelevant. Die Positionen werden durch Zahlen von 1 bis 32 beschrieben. Dann bestimmen wir die Differenzen der Positionen aller möglichen Paare dieser vier Ass. Ist irgendwo der Betrag der Differenz 1, dann liegen die Ass hintereinander.

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
|-----|----|----|----|----|------|------|------|------|-----|-----|-----|-----|-----|-----|--------------|
| 1 | Z1 | Z2 | Z3 | Z4 | ass1 | ass2 | ass3 | ass4 | 1+2 | 1+3 | 1+4 | 2+3 | 2+4 | 3+4 | nur ein Paar |
| 2 | 20 | 16 | 25 | 20 | 20 | 16 | 27 | 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 30 | 19 | 25 | 7 | 30 | 19 | 26 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 1 | 9 | 21 | 15 | 1 | 10 | 23 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 23 | 2 | 19 | 29 | 23 | 2 | 20 | 32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 10 | 26 | 4 | 1 | 10 | 27 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 2 | 10 | 8 | 15 | 2 | 11 | 9 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 12 | 27 | 12 | 18 | 12 | 28 | 13 | 20 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| 9 | 10 | 28 | 26 | 10 | 10 | 29 | 27 | 11 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 10 | 25 | 31 | 19 | 10 | 25 | 32 | 19 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 28 | 21 | 19 | 3 | 28 | 21 | 19 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 32 | 23 | 18 | 16 | 32 | 23 | 18 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 15 | 10 | 2 | 18 | 15 | 10 | 2 | 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 4 | 25 | 7 | 3 | 4 | 26 | 8 | 3 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 15 | 25 | 27 | 15 | 16 | 25 | 28 | 15 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 22 | 21 | 10 | 17 | 22 | 21 | 10 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 1 | 9 | 21 | 15 | 1 | 10 | 23 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | 23 | 2 | 19 | 29 | 23 | 2 | 20 | 32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | 10 | 26 | 4 | 1 | 10 | 27 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20 | 2 | 10 | 8 | 15 | 2 | 11 | 9 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21 | 12 | 27 | 12 | 18 | 12 | 28 | 13 | 20 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| 22 | 10 | 28 | 26 | 10 | 10 | 29 | 27 | 11 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 23 | 25 | 31 | 19 | 10 | 25 | 32 | 19 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 24 | 28 | 21 | 19 | 3 | 28 | 21 | 19 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 25 | 32 | 23 | 18 | 16 | 32 | 23 | 18 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 26 | 15 | 10 | 2 | 18 | 15 | 10 | 2 | 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 27 | 4 | 25 | 7 | 3 | 4 | 26 | 8 | 3 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 28 | 25 | 27 | 15 | 16 | 25 | 28 | 15 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 29 | 12 | 1 | 16 | 23 | 12 | 1 | 18 | 26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 290 | 12 | 1 | 16 | 23 | 12 | 1 | 18 | 26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 291 | 26 | 11 | 27 | 18 | 26 | 11 | 29 | 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 292 | 13 | 29 | 7 | 15 | 13 | 30 | 7 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 293 | 1 | 25 | 29 | 6 | 1 | 26 | 31 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 294 | 8 | 28 | 17 | 9 | 8 | 29 | 18 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 295 | 25 | 3 | 19 | 14 | 25 | 3 | 20 | 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 296 | 5 | 9 | 14 | 29 | 5 | 10 | 16 | 32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 297 | 28 | 2 | 5 | 21 | 28 | 2 | 6 | 23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 298 | 8 | 12 | 24 | 1 | 8 | 13 | 26 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 299 | 22 | 10 | 27 | 22 | 22 | 10 | 29 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 300 | 21 | 21 | 18 | 23 | 21 | 22 | 18 | 26 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 301 | 25 | 9 | 28 | 1 | 25 | 9 | 30 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 302 | | | | | | | | | | | | | | | |
| 303 | | | | | | | | | | | | | | | 0,33 |

Abbildung 10: Monte-Carlo-Simulation eines gemischten Kartenstapels

Die fertige Simulation ist in Abbildung 10 abgebildet. Zunächst erzeugen wir in den Spalten A bis D vier Zufallsvariablen. Die Ereignisräume sind natürlich unterschiedlich, da für das erste Ass noch 32 Möglichkeiten zur Verfügung stehen, beim vierten nur noch 29. Wir verwenden hier die Formeln

=ZUFALLSBEREICH(1;32)

für das erste Ass (Spalte A). Der Bereich wird dann in jedem Schritt etwas kleiner, bis schließlich für das letzte Ass (Spalte D)

=ZUFALLSBEREICH(1;29)

verwendet wird. In den Spalten E bis H werden dann diese Zufallswerte in Positionen umgerechnet. Hier ist folgendes zu beachten: Ist der Zufallswert einer Karte größer oder gleich wie der Zufallswert einer vorhergehenden Karte, dann muss der Wert um 1 vergrößert werden⁴. Wir realisieren das mit folgenden Formeln für die Spalten F,G und H:

⁴Die Interpretation ist folgende: Wenn bereits eine Karte auf den Platz 10 gelöst wurde, dann steht diese Position nicht mehr zur Verfügung. Wird jetzt nochmal die 10 gelöst, dann wird die darauf folgende Position

=B2+WENN(B2>=A2;1;0)

=C2+WENN(C2>=A2;1;0)+WENN(C2>=B2;1;0)

=D2+WENN(D2>=A2;1;0)+WENN(D2>=B2;1;0)+WENN(D2>=C2;1;0)

Anschließend vergleichen wir in den Spalten I bis N paarweise die Positionen derASSE mit Formeln, die folgendermaßen aussehen:

=WENN(ABS(E2-F2)=1;1;0)

Eine 1 bedeutet dabei, dass die beiden verglichenenASSE benachbart sind. In der letzten Spalte werden dann die Spalten I bis N addiert. Das Event wird nur dann gezählt, wenn die Summe gleich 1 ist (wenn es also ein Paar gibt, aber nicht mehrere).

=WENN(SUMME(I2:N2)=1;1;0)

Schließlich können wir dann die relative Häufigkeit der positiven Events in Spalte O berechnen, um einen Schätzwert für die tatsächliche Wahrscheinlichkeit zu erhalten.

5 Deterministische Daten mit Rauschen

Beim Arbeiten mit Daten hat man oft den folgenden Fall: Man weiß bereits im Voraus, dass die Daten einem gewissen Muster folgen (man spricht von einem Trend). Dennoch sind die Werte dem Zufall unterworfen oder die exakte Berechnung ist sehr schwierig, so dass man keine völlig genaue Vorhersage treffen kann. In diesem Fall ist es in der mathematischen Modellierung üblich, den Trend durch eine recht einfache Funktion darzustellen und diese mit statistischen Schwankungen zu stören: man spricht davon, dass man **Rauschen** addiert.

5.1 Modellierung von Benzinpreisen

Ein Beispiel für dieses Phänomen aus dem Alltag sind Benzinpreise. Diese schwanken recht deutlich von Tag zu Tag. Dennoch kann man übers Jahr hinweg einen typischen Trend feststellen: in manchen Jahreszeiten ist der Benzinpreis niedriger.

Wir illustrieren das Thema an folgender Fragestellung: Es soll eine Strategie entwickelt werden, wie man beim Tanken des Autos vorgehen soll. Das Ziel ist stets, bei einer Tankstelle zu tanken, die günstiger ist als andere und im Idealfall auch an einem Tag, an dem der Preis geringer ist als an den Tagen zuvor oder danach.

Wenn verschiedene Strategien ausgearbeitet wurden, könnte man diese an realen Daten testen. Die Daten zu erheben könnte jedoch sehr aufwändig sein, daher ist es eine interessante Alternative, die Daten selbst zu erzeugen. Das wollen wir hier tun.

Zunächst betrachten wir die durchschnittlichen monatlichen Benzinpreise im Verlauf eines Jahres (Abbildung 11). Für diese Kurve bestimmen wir eine Regressionskurve, die den Trend vorgibt.

Beispielsweise könnte der Verlauf als Summe zweier Sinusfunktionen dargestellt werden:

$$P(x) = 11 \sin\left((x - 8,5)\frac{\pi}{3}\right) + 4 \sin\left((x - 11,5)\frac{\pi}{6}\right) + \frac{2}{12}x + 149, \quad x \text{ in Monaten.}$$

Der lineare Anstieg, der dazuaddiert wurde, soll noch den jährlichen Preisanstieg darstellen, der hier mit 2 cent geschätzt wurde.

vergeben, nämlich Platz 11, ebenso müssen alle weiteren Positionen um 1 verschoben werden.

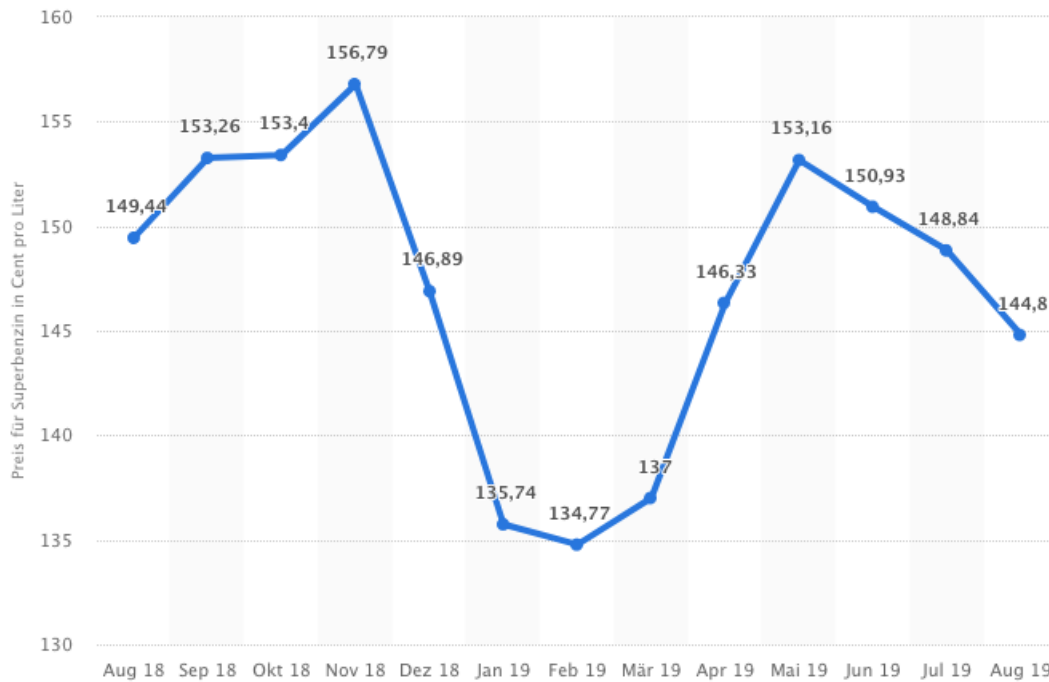


Abbildung 11: Daten für die Benzinpreise im monatlichen Durchschnitt von August 2018 bis August 2019 (Quelle: <https://de.statista.com/statistik/daten/studie/1690/umfrage/preis-fuer-einen-liter-superbenzin-monatsdurchschnittswerte/>).

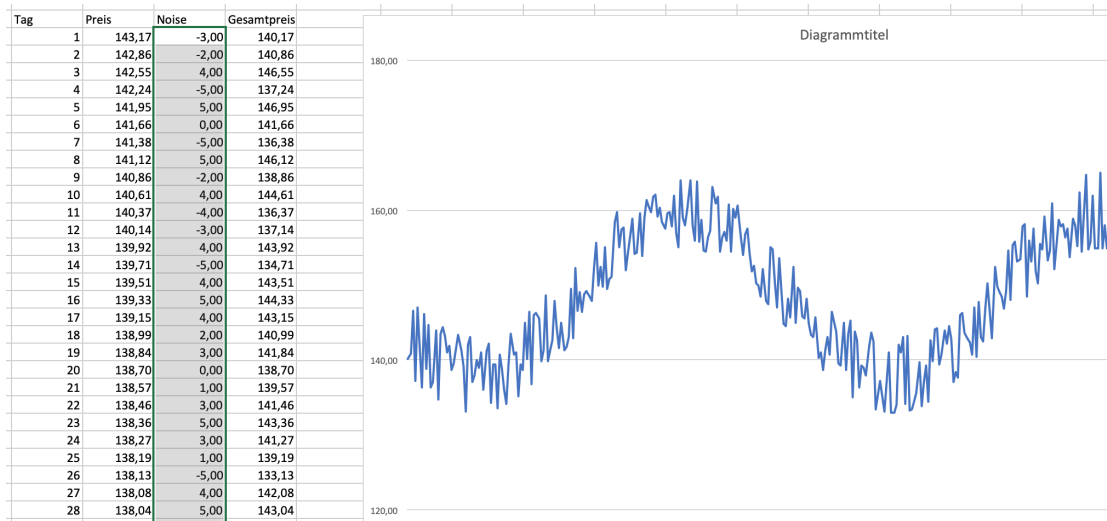


Abbildung 12: Benzinpreise: Trendkurve mit Rauschen.

Im nächsten Schritt können wir diese Funktion verwenden und jeden Tag einen gewissen Zufallswert hinzuaddieren oder abziehen (Abbildung 12). Dazu wird für jeden Tag im Jahr anhand der Formel für die Trendkurve der jeweilige Tagespreis berechnet. Anschließend addiert man einen Zufallswert (z.B. `ZUFALLSBEREICH(-5;5)`).