3. Simulation von Zufallsexperimenten mit Tabellenkalkulationsprogrammen

3.1 Erzeugen von Zufallszahlen

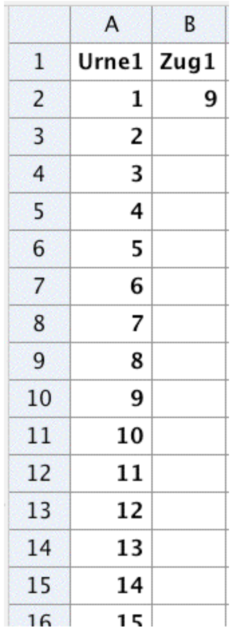
(Wir beschreiben Beispiele in der Software „Geogebra“)
In Tabellenkalkulationsprogrammen wie Excel, OpenOffice Calc oder dem Tabellenteil von GeoGebra arbeitet man vorwiegend mit Zahlen, z.T. können auch Texte in den Zellen stehen. Daher sind die Ergebnisse von Zufallsexperimenten in Rechenblättern üblicherweise Zahlen. Für manche Experimente muss man also die Ergebnisse in Zahlen übersetzen. Bei einer umkehrbar eindeutigen Zuordnung sind aber letztlich alle Experimente in Tabellenkalkulationsprogrammen nachbildbar.
In allen Tabellenkalkulationsprogrammen gibt es einen Befehl, mit dem man eine mehrstellige Zahl aus dem Intervall [0,1) erzeugen kann. (1 selbst ist nicht erreichbar) Diese Zufallszahlen sind gleich verteilt. D.h. hier, dass man gleich große Teilintervalle von [0,1) mit gleicher Wahrscheinlichkeit trifft. In GeoGebra heißt dieser Befehl „random()“. Es ist eine numerische Funktion, daher stehen die Parameter in runden Klammern, die hier allerdings leer bleiben, da kein Parameter erforderlich ist. Die angezeigte Stellenzahl hängt von der Einstellung im Programm ab, intern ist es eine 20-stellige Zahl.

	A
1	0.914816140859245
2	=random()

Die Abbildung zeigt in Zelle A1 das Ergebnis (Einstellung auf 15 Stellen) und in Zelle A2, wie der Aufruf der Funktion aussieht.
Sehr oft braucht man als Ergebnis eine ganzzahlige Zufallszahl aus einem begrenzten Intervall, wobei alle Zahlen als Elementarereignisse gleich wahrscheinlich sein sollen. So ist z.B. ein Münzwurf durch die Wahl von 0 oder 1 nachbildbar, ein Würfelwurf erzeugt eine Zahl von 1 bis 6. Dafür gibt es in den Tabellenkalkulationsprogrammen oftmals einen speziellen Befehl. In GeoGebra heißt er „Zufallszahl“ und ist ein GeoGebra-Befehl. Daher stehen die Parameter in eckigen Klammern und es müssen eine obere und untere Grenze für das Intervall angegeben werden, aus dem die Zufallszahlen gewählt werden.

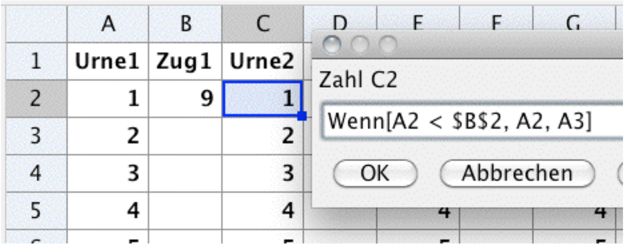
	A
1	=Zufallszahl[1,6
2	1
3	2
4	2
5	1
6	2
7	4
8	6
9	4
10	2
11	3
12	4
13	5
14	1
15	1

Üblicherweise sind die Grenzen ganze Zahlen, die als Zufallsergebnis auch eintreten können. So ist „Zufallszahl[1,6]“ der Aufruf, um einen Würfel zu simulieren. (siehe Abbildung, der Befehl wurde in den Zellen A2 bis A15 ausgeführt).
„Zufallszahl“ akzeptiert auch nichtganze Zahlen, die Ergebnisse sind dann aber auch nichtganzzahlig und liegen z.T. außerhalb des Intervalls. Man sollte also diesen Fall vermeiden.
Das Ziehen von Zahlen aus einer Urne lässt sich entsprechend mit dem Befehl „Zufallszahl“ simulieren, wenn das Ziehen mit Zurücklegen erfolgt. Für das Ziehen ohne Zurücklegen muss man etwas mehr Aufwand betreiben.
Wir wollen das Ziehen ohne Zurücklegen simulieren für die Ziehung der Lottozahlen, also 6 Zahlen aus den Zahlen 1 bis 49.



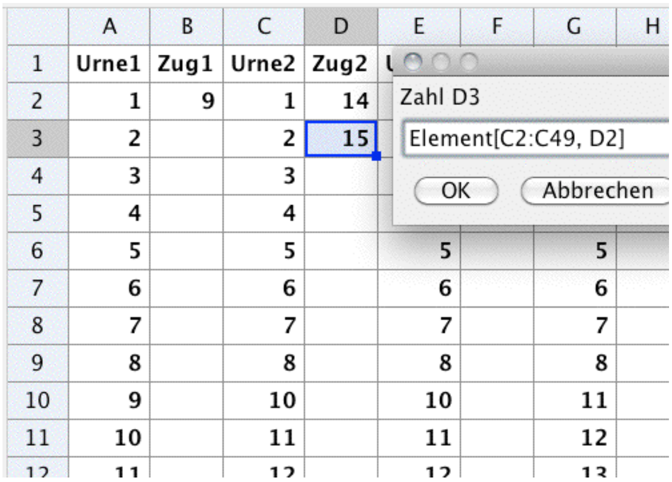
	A	B
1	Urne1	Zug1
2	1	9
3	2	
4	3	
5	4	
6	5	
7	6	
8	7	
9	8	
10	9	
11	10	
12	11	
13	12	
14	13	
15	14	

Dazu geben wir in der Tabelle zunächst in die Spalte A die Zahlen von 1 bis 49 ein (neben der Zeile 1, in der wir die Spalten bezeichnen, werden also die Zeilen 2 bis 50 belegt). In Zelle B2 lassen wir uns mit „=Zufallszahl[1,49]“ eine Zahl auswählen.
Nun müssen wir dafür sorgen, dass eine neue Urne (Spalte) gebildet wird, in der die gezogene 9 nicht mehr vorkommt. Dazu verwenden wir den „Wenn“-Befehl, mit dem wir eine bedingte Zuordnung für eine Zelle treffen können.



	A	B	C
1	Urne1	Zug1	Urne2
2	1	9	1
3	2		2
4	3		3
5	4		4

Die linke Abbildung zeigt die Formel für die Zelle C2. Ist die Zahl in A2 kleiner als die gezogene Zahl in B2, so wird die Zahl in A2 übernommen, ansonsten wird die nächste Zelle A3 genommen. Diese Formel ziehen wir runter bis zur Zeile 49. Bei B2 ist durch die $-Zeichen dafür gesorgt, dass immer mit dieser Zelle verglichen wird (absoluter Zellbezug).



	A	B	C	D	E	F	G
1	Urne1	Zug1	Urne2	Zug2			
2	1	9	1	14			
3	2		2	15			
4	3		3				
5	4		4				
6	5		5		5		5
7	6		6		6		6
8	7		7		7		7
9	8		8		8		8
10	9		10		10		11
11	10		11		11		12

Verfolgt man im Ergebnis die Spalte C bis zur Zeile 10, so erkennt man, dass in der Tat die 9 ausgelassen wurde.
Im nächsten Schritt wählen wir nun aus der neuen Liste in Spalte C wieder zufällig eine Zahl aus. Das machen wir in Zelle D2 mit dem Befehl „=Zufallszahl[1,48]“. Die obere Grenze ist nun 48, da ja nur noch 48 Zahlen zur Verfügung stellen. Diese Zahl ist nun aber nicht direkt die gezogene Zahl, denn hier könnte ja zufällig die gezogene 9 wieder gewählt werden. Die Zufallszahl in D2 gibt an, welche Zahl aus dem Bereich C2 bis C49 ausgewählt werden soll. Dabei kommt die zuerst gewählte Zahl ja nicht mehr vor. Dieses geschieht mit dem Befehl „=Element[C2:C49,D2]“[3]. Somit steht in D3 die nächste gezogene Zahl und es ist sichergestellt, dass es nicht zu einer Wiederholung kommen kann.
Die Tabelle wird nun in entsprechender Weise vervollständigt, bis aus Urne 6 die letzte Ziehung erfolgt. Zur besseren Übersicht fassen wir in Spalte M noch einmal die sechs gezogenen Zahlen zusammen.

[3] „Element“ ist eigentlich ein Befehl für Listen. Wir wollen aber hier nicht in das Arbeiten mit Listen einsteigen, sondern bewusst die Tabelle benutzen. Man kann in passenden Befehlen stets Listen durch Tabellenbereiche ersetzen.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Urne1	Zug1	Urne2	Zug2	Urne3	Zug3	Urne4	Zug4	Urne5	Zug5	Urne6	Zug6	Züge
2	1	4	1	6	1	6	1	41	1	32	1	15	4
3	2		2	7	2	8	2	44	2	35	2	18	7
4	3		3		3		3		3		3		8
5	4		5		5		5		5		5		44
6	5		6		6		6		6		6		35
7	6		7		8		9		9		9		18
8	7		8		9		10		10		10		
9	8		9		10		11		11		11		

In der fertigen (nach unten abgeschnittenen) Tabelle (Abbildung oben) kann man im Zug 3 sehr schön sehen, wie das Vermeiden von Wiederholungen funktioniert. Es wird durch „Zufallszahl“ zwar ein weiteres Mal die 6 gewählt, was aber bedeutet, die sechste Zahl aus Urne 3 zu wählen, welches die 8 ist.

Hat man in einer Tabelle Zufallszahlen erzeugt, möchte man zur Wiederholung des in der Tabelle modellierten Zufallsexperiments die Zufallszahlen neu erzeugen lassen. In Excel ist das mit dem Durchrechnen der Tabelle verbunden. Immer dann, wenn die Tabelle durchgerechnet wird, werden auch neue Zufallszahlen erzeugt. In der Regel geschieht das, wenn man in einer Zelle einen Eintrag verändert. In GeoGebra werden beim Verändern einer Zelle nur diejenigen aktualisiert, die von der veränderten Zelle logisch abhängen. In allen anderen Zellen bleiben die Werte unverändert, auch Zufallswerte.

Will man alle Zufallswerte einer Tabelle neu erzeugen, gibt es in GeoGebra den speziellen Befehl „cmd R“ (oder „str R“).

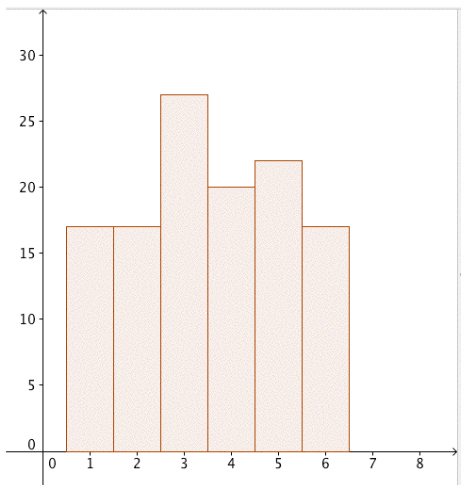
3.2 Auswerten von Zufallszahllisten

	A	B	C
1	4	1	17
2	5	2	17
3	6	3	27
4	3	4	20
5	5	5	22
6	2	6	17
7	1		
8	5		

Als Demonstrationsbeispiel erzeugen wir eine Liste von 120 Würfelwürfen. Dazu schreiben wir in Zelle A1 „=Zufallszahl[1,6]“ und ziehen dann dieses bis in Zelle A120. Von diesem Zellbereich wollen wir nun feststellen, wie oft die einzelnen Ergebnisse aufgetreten sind.

Sind wir an genauen Zahlen interessiert, so verwenden wir den GeoGebra-Befehl „ZähleWenn[Bedingung, Tabellenbereich]“ (ACHTUNG: In Excel gibt es auch diesen Befehl, der dort aber „Zähle**n**wenn“ heißt). In unserem Fall lassen wir uns in sechs Zellen anzeigen, wie viele Einsen, Zweien, ..., Sechsen in der Liste der 120 Zufallszahlen auftreten. Für eine gewisse Übersichtlichkeit und Flexibilität schreiben wir die Zahlen 1 bis 6 in die Zellen B1 bis B6 und in Zelle C1 „=ZähleWenn[x==B1,$A$1:$A$120]“. Für den Vergleich wird die allgemeine Variable x verwendet. Das vergleichende Gleichheitszeichen ist das doppelte „==“, das beim nachträglichen Editieren als ein Gleichheitszeichen mit Fragezeichen dargestellt wird.

Den Inhalt von Zelle C1 ziehen wir nun runter bis C6 und erhalten so die gewünschte Übersicht über die Häufigkeit der einzelnen Ergebnisse. Das Bild zeigt, dass durchaus große Abweichungen vom theoretischen Wert 20 auftreten können.



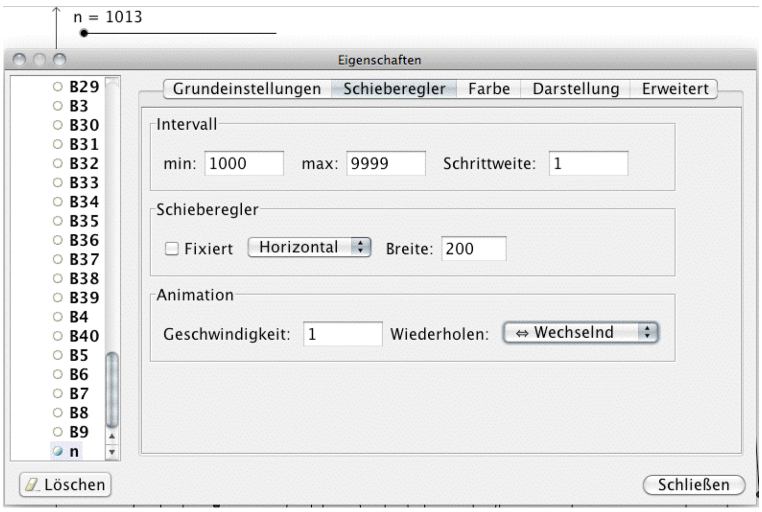
	A
1	4
2	5
3	6
4	3
5	5
6	2
7	1
8	5
9	6
10	1
11	2
12	1
13	1
14	5
15	2
16	3
17	2

Ist man eher an einer grafischen Darstellung interessiert, verwendet man den ausgesprochen komfortablen Befehl „Balkendiagramm“. (Hier zeigt sich z.B., dass GeoGebra eher an den Bedürfnissen der Mathematik ausgerichtet ist als die anderen Tabellenkalkulationsprogramme.) „Balkendiagramm“ kann mit verschiedenen Parametern aufgerufen werden, wir verwenden hier die Rohdaten und die Balkenbreite. Wir geben „Balkendiagramm[A1:A120,1]“ ein und erhalten im Grafikfenster das passende Diagramm. Mit „cmd R“/„str R“ kann man die 120 Würfelwürfe immer wieder neu durchführen. Das Diagramm wird sofort aktualisiert.

3.3 Visualisieren von Zufallsfolgen

Für das Testen und Visualisieren von Zufallsfolgen (Pseudozufallszahlen) stellt man die einzelnen Werte durch Punkte dar, die man zum leichteren Verfolgen der Abfolge mit Strecken verbindet. Das soll am Beispiel des vierstelligen Quadratsummenalgorithmus gezeigt werden (siehe auch Ü1).

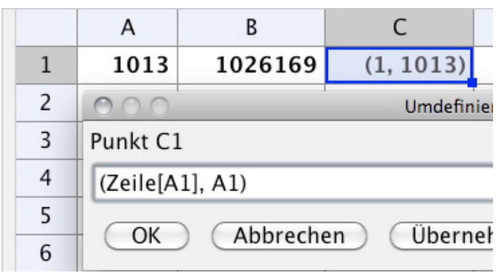
Nehmen wir eine vierstellige Zahl und quadrieren sie, so erhalten wir eine achtstellige Zahl (ggfs. mit einer Null an der achten Stelle). Die mittleren Ziffern, genauer die Hunderter- bis Hunderttausenderziffer, bilden wieder eine vierstellige Zahl (ggfs. mit führenden Nullen), die die nächste Zahl in unserer Zufallsfolge sein soll.



Realisierung in GeoGebra: Für den Startwert richten wir einen Schieberegler ein, der von 1000 bis 9999 laufen kann. Mit dem Wert in der Zelle A1 starten wir die Tabelle, der Wert von n ist unsere erste Zahl in der Zufallsfolge. Daneben in B1 berechnen wir das Quadrat der Zahl in A1.

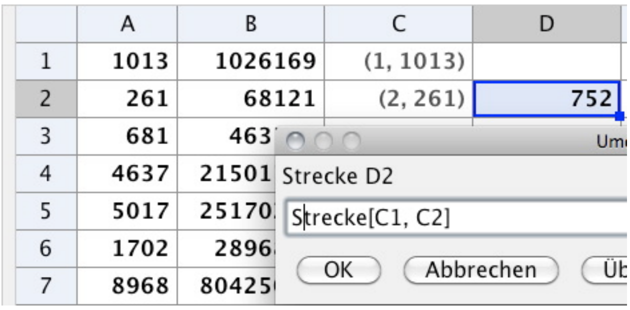
Nun berechnen wir in A2 die „mittlere“ vierstellige Zahl von B2. Das geschieht mit der Formel „A2 = floor(B1/100) - floor(B1/ 1000000) 10000“, die hier nicht näher analysiert werden soll. In B2 berechnet man wieder das Quadrat von A2, also „B2=A2^2“ Diese beiden Formeln zieht man herunter, für unsere Zwecke reicht es bis Zeile 40. Die Folge der Zufallszahlen steht nun in Spalte 40, Spalte B hat lediglich eine Hilfsfunktion. Das Bild rechts zeigt die Entwicklung der Zufallszahlen mit der Startzahl n = 1013. Man kann sehr schön sehen, dass die neue Zufallszahl in Spalte A die mittlere, vierstellige Zahl (man muss sich bei den dreistelligen Zahlen eine führende Null denken) der darüberstehenden Zahl in Spalte B ist. Wir wollen nun diese Zufallsfolge wie oben zu Beginn dieses Abschnitts beschrieben visualisieren.

	A	B
1	1013	1026169
2	261	68121
3	681	463761
4	4637	21501769
5	5017	25170289
6	1702	2896804
7	8968	80425024
8	4250	18062500
9	625	390625
10	3906	15256836
11	2568	6594624
12	5946	35354916
13	3549	12595401
14	5954	35450116
15	4501	20259001
16	2590	6708100
17	7081	50140561
18	1405	1974025
19	9740	94867600
20	8676	75272976
21	2729	7447441
22	4474	20016676
23	166	27556
24	275	75625

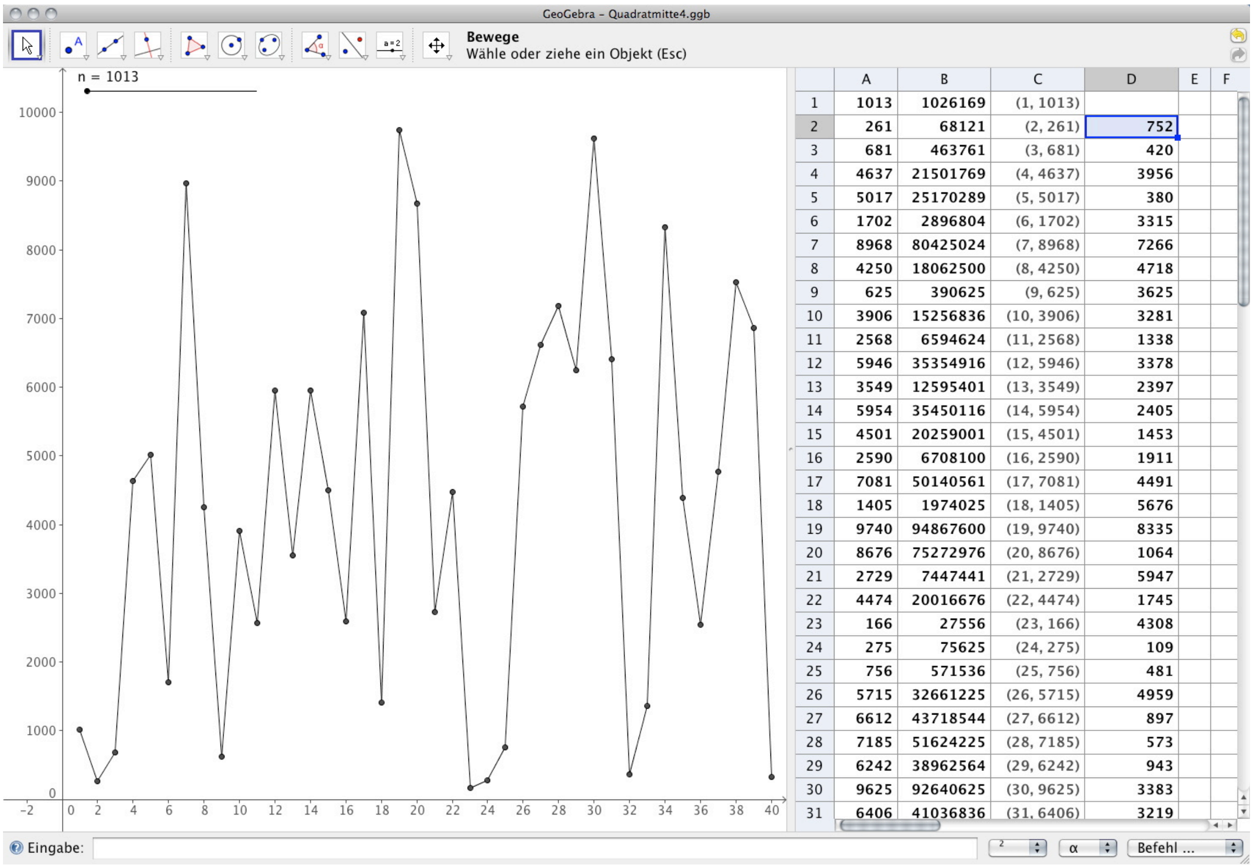


Im ersten Schritt erzeugen wir die Punkte für den Liniengraph. Dabei soll die einzelne Zufallszahl die y-Koordinate eines Punktes ausmachen, der Index in der Folge die x-Koordinate. Die Formel für die Zelle C1 zeigt die Abbildung links. Diese Formel ziehen wir nun herunter und erhalten in Spalte C zu jeder Zufallszahl einen Punkt (C1 bis C40). Damit diese Punkte auch im Grafikfenster sichtbar sind müssen die Achsen (getrennt) passend formatiert werden. In unserem Beispiel muss die y-Achse Werte bis 9999 darstellen können.

Für die Verbindungslinien gehen wir analog vor. Wir erzeugen jede Linie als Objekt in einer Zelle, es wird dann automatisch grafisch dargestellt. Die Formel sieht man in der Abbildung rechts. Diese wird wieder bis zur Zeile 40 her-untergezogen.



Die fertige Abbildung sieht dann folgendermaßen aus:



	A	B	C	D
1	1013	1026169	(1, 1013)	
2	261	68121	(2, 261)	752
3	681	463761	(3, 681)	420
4	4637	21501769	(4, 4637)	3956
5	5017	25170289	(5, 5017)	380
6	1702	2896804	(6, 1702)	3315
7	8968	80425024	(7, 8968)	7266
8	4250	18062500	(8, 4250)	4718
9	625	390625	(9, 625)	3625
10	3906	15256836	(10, 3906)	3281
11	2568	6594624	(11, 2568)	1338
12	5946	35354916	(12, 5946)	3378
13	3549	12595401	(13, 3549)	2397
14	5954	35450116	(14, 5954)	2405
15	4501	20259001	(15, 4501)	1453
16	2590	6708100	(16, 2590)	1911
17	7081	50140561	(17, 7081)	4491
18	1405	1974025	(18, 1405)	5676
19	9740	94867600	(19, 9740)	8335
20	8676	75272976	(20, 8676)	1064
21	2729	7447441	(21, 2729)	5947
22	4474	20016676	(22, 4474)	1745
23	166	27556	(23, 166)	4308
24	275	75625	(24, 275)	109
25	756	571536	(25, 756)	481
26	5715	32661225	(26, 5715)	4959
27	6612	43718544	(27, 6612)	897
28	7185	51624225	(28, 7185)	573
29	6242	38962564	(29, 6242)	943
30	9625	92640625	(30, 9625)	3383
31	6406	41036836	(31, 6406)	3219

Mit dem Schieberegler (oben links im Grafikfenster) für die Variable n als Startwert in Zelle A1 kann man nun verschiedene Startwerte ausprobieren.

Übungsaufgaben zum Kapitel 3

Ü1 Zufallszahlen 1

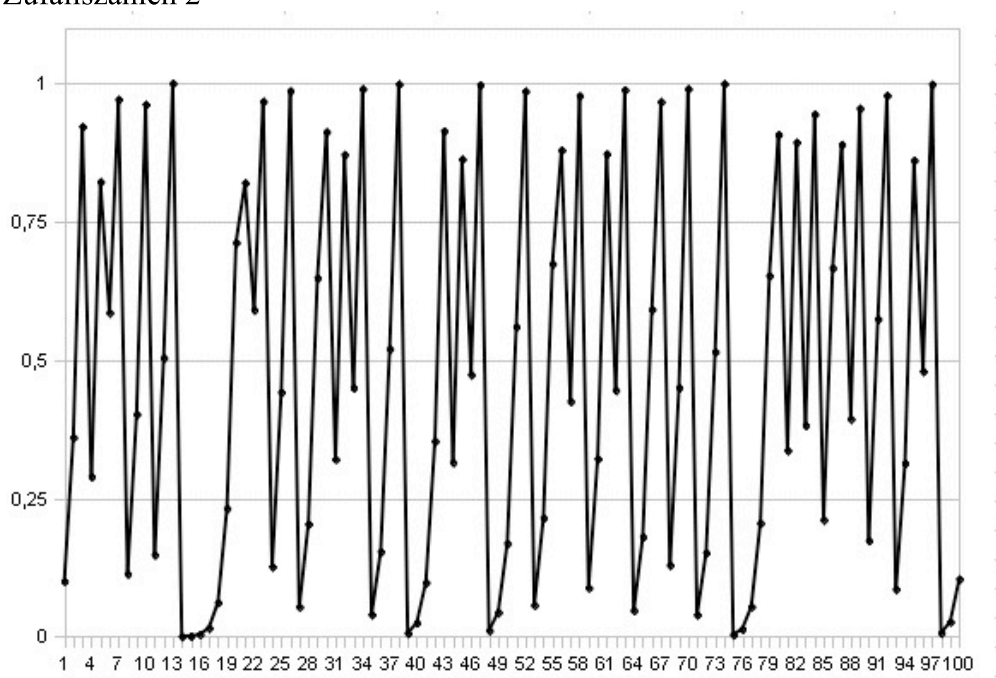
Ein sehr einfacher Algorithmus für Zufallszahlen ist der „Quadratmitten“-Algorithmus. Man startet mit einer n-stelligen, natürlichen Zahl, quadriert diese und nimmt aus dem Ergebnis wieder die mittlere, n-stellige Zahl.

Beispiel: 2-stellig: $36 \rightarrow 36^2=1296 \rightarrow 29$

a. Probieren Sie mit zweistelligen Zahlen einige Ketten. Welches ist Ihre längste Kette oder längster Zyklus? Was geschieht letztlich mit den Ketten?

b. Probieren Sie auch einmal vierstellige Zahlen.

Ü2 Zufallszahlen 2



Um 1945 herum haben Ulam und von Neumann mit der Folge $x_{n+1} = 4x_n(1-x_n)$ versucht, Pseudo-Zufallszahlen zu erzeugen. Das Diagramm zeigt den Verlauf mit der Startzahl $x_0 = 0{,}1$. Im Diagramm ist der Wertebereich in vier Teile geteilt. Welche Tests für die Qualität der Zufallszahlen können Sie vornehmen? Testen Sie danach die Zufallszahlen. Wie beurteilen Sie insgesamt die vorliegende Folge als Zufallszahlen?

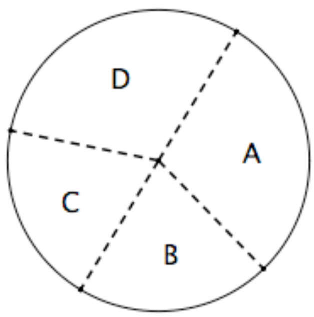
Ü3 Für das Ziehen ohne Zurücklegen wollen Sie keinen großen Aufwand treiben und simulieren das einfache Ziehen mit Zurücklegen. Sollte dann eine oder mehrere Zahlen doch mehrmals vorkommen so wiederholen Sie einfach die Ziehung. Wie groß ist für den Versuch 6 aus 49 die Wahrscheinlichkeit, dass Sie bei einmaliger Durchführung keine mehrfachen Zahlen haben?

Ü4 Geben Sie zum Ziehen ohne Zurücklegen in der oben entwickelten Tabelle die Formeln in den Zellen H2, H3 und I2 an.

Ü5 In einer Tabelle stehen in den Zellen A1 bis A10 die Zahlen 2, 3, 5, 7, 11, 13, 17, 19, 23, 29. In Zelle B1 steht „Zufallszahl[1,10]" und Zelle B2 „Element[A1:A10, B1]". Erläutern Sie, was dadurch in B2 zur Verfügung gestellt wird.

Ü6 Ein Glücksrad mit den vier Sektoren A, B, C und D wird durch ein Rechenblattprogramm simuliert.

	A	B	C	D	E	F	G	H	I	J
1	Zufall1	Zufall2	Zufall3	Entsch1	Entsch2	Entsch3	A	B	C	D
2	0.78	0.6	0.52	0	0	1	0	0	1	0
3				1	1	0				



In den Zellen A2, B2 und C2 wird jeweils mit =random() eine Zufallszahl erzeugt.

In D2 steht die Formel: =WENN(A2<0.25, 1, 0).
In E2 steht die Formel: =WENN(B2<0.5, 1, 0).
In F2 steht die Formel: =WENN(C2<0.6, 1, 0).
In D3 steht =1 – D2, in E3 steht =1 – E2, in F3 steht =1 – F2.
In G2 steht =D2*E2. In H2 steht =D2*E3. In I2 steht =D3*F2. In J2 steht =D3*F3.

a. Was bewirken die Paare D2, D3 bzw. E2, E3 bzw. F2, F3?
b. Erläutern Sie, weshalb in G2 bis J2 immer genau eine 1 steht, die den gewählten Sektor angibt.
c. Analysieren Sie die Formeln und geben Sie an, wie viel Grad die einzelnen Sektoren des Glücksrades haben *(Das Bild ist nur eine grobe Veranschaulichung, die Sektoren können ganz anders eingeteilt sein)*.
d. Übersetzen Sie die Auswahl der vier Sektoren in ein Baumdiagramm mit den zugehörigen Wahrscheinlichkeiten.