

# Warten auf einen Run – und was kommt dann?

GERD RIEHL, BARSINGHAUSEN

**Zusammenfassung:** Das Problem, die Wahrscheinlichkeit für einen Run der Länge 4 (also 4 gleiche Zahlen hintereinander) beim  $n$ -maligen Würfeln zu berechnen, lässt sich mit Markow-Ketten wesentlich leichter als mit der von Motzer (2010) angegebenen Rekursion lösen. Neben der Möglichkeit zur Visualisierung der stochastischen Situation hat diese Methode den Vorzug, mühelos auf die verallgemeinerte Fragestellung nach der Verteilung der Zufallsgröße „Anzahl Runs der Länge  $\lambda$  beim  $n$ -maligen Würfeln“ übertragbar zu sein.

## 1 Die Problemstellung

Serien von gleichen Würfelzahlen – so der Titel von Motzers Aufsatz – erscheinen stochastischen Laien oft unwahrscheinlicher als sie tatsächlich sind. Viele Menschen halten einen häufigen Wechsel zwischen zwei aufeinander folgenden Ausfällen beim wiederholten Würfeln, in Serien von Münzwürfen, Folgen von Zufallsziffern usw. für ein typisches Merkmal des Zufalls.

Wäre dies so, würden in einer solchen Serie mehr, dafür aber kürzere Runs (Folgen gleicher Ergebnisse) auftreten, als theoretisch zu erwarten und in Versuchen auch zu beobachten sind. Wir greifen hier ein Beispiel aus Motzer (2010) auf.

Aufgabe 1: Bei einer Würfelserie der Länge  $n = 100$  fragt man nach der Wahrscheinlichkeit für einen (genauer: *mindestens* einen) Run von 4 oder mehr gleichen Augenzahlen.

Diese Aufgabe 1 werden wir in Abschnitt 2 mithilfe einer Markow-Kette lösen. Deren Graphen braucht man dann nur geringfügig abzuändern, um auch die Wahrscheinlichkeit für mindestens einen Run der *genauen* Länge 4 berechnen zu können.

In Abschnitt 3 zeigen wir, dass man mit demselben Verfahren auch berechnen kann, wie wahrscheinlich es ist, *genau* einen Run und allgemein genau  $k$  Runs in der Serie zu erzielen.

Auf einige Varianten des Problems und eine ähnlich zu bearbeitende Verallgemeinerung gehen wir in den Abschnitten 4 und 5 ein.

## 2 Ein neue Lösungsmethode durch Visualisierung

Wir veranschaulichen zunächst die stochastische Situation mithilfe des folgenden Übergangsgraphen:

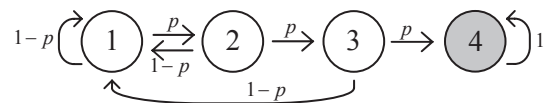


Abb. 1: Übergangsgraph zu Aufgabe 1

Nach dem ersten Wurf hat man *eine* Zahl als Anfang eines im Aufbau begriffenen Runs, man befindet sich im Zustand ①. Beim zweiten Mal wirft man mit der Wahrscheinlichkeit  $p = 1/6$  dieselbe Zahl, in diesem Fall wächst der Run auf die Länge 2, d. h. man gelangt in den Zustand ②. Andernfalls hat der erste Run die Länge 1, während die Augenzahl des zweiten Wurfs zum Anfang eines neuen Runs wird und man somit im Zustand ① bleibt; dies geschieht mit der Wahrscheinlichkeit  $1 - p = 5/6$ .

Genauso fällt man aus den Zuständen ② und ③ nach ① zurück, wenn die geworfene Zahl wechselt; damit endet ein Run der Länge 2 bzw. 3. Bei gleicher Augenzahl wächst dagegen die Runlänge um 1, und man gelangt in den Zustand ③ bzw. ④.

Im Zustand ④ ist das Ziel – mindestens ein Run der Länge  $\lambda \geq 4$  – erreicht, alle weiteren Würfe sind nicht mehr von Interesse, wir befinden uns in einem *absorbierenden* Zustand, den man nicht mehr verlassen kann (dies wird sich in Abschnitt 3 ändern).

Wir bezeichnen mit  $a_{j,z}$  die Wahrscheinlichkeit, nach  $j$  Würfeln im Zustand  $z$  zu sein, und bilden aus den  $a_{j,z}$  für jedes  $j$  einen vierdimensionalen Zeilenvektor

$$\vec{a}_j = (a_{j,1}; a_{j,2}; a_{j,3}; a_{j,4}),$$

den so genannten *Zustandsvektor*. Offenbar gilt

$$\vec{a}_1 = (1; 0; 0; 0).$$

Abbildung 1 und die obigen Überlegungen führen dann auf folgende Rekursionsformeln für die Koordinaten von  $\vec{a}_{j+1}$ :

$$a_{j+1,1} = (1 - p) \cdot a_{j,1} + (1 - p) \cdot a_{j,2} + (1 - p) \cdot a_{j,3};$$

$$a_{j+1,2} = p \cdot a_{j,1};$$

$$a_{j+1,3} = p \cdot a_{j,2};$$

$$a_{j+1,4} = p \cdot a_{j,3} + a_{j,4}.$$

Tabelle 1 zeigt die Umsetzung dieser Formeln mit EXCEL in einem Ausschnitt (technische Einzelheiten werden im folgenden Abschnitt angesprochen).

Tabelle 1 liefert also den auch von Motzer (2010) angegebenen Wert  $a_{100;4} \approx 0,316$ .

Dieses Ergebnis ist allerdings unbefriedigend, da es alle Würfe *nach* dem Auftreten von vier gleichen Zahlen ignoriert und damit offen bleibt, ob noch weitere Vierer-Runs in der Gesamtserie auftreten (diese Frage behandeln wir im folgenden Abschnitt 3) und ob tatsächlich ein Vierer-Run vorliegt und nicht einer der Länge 5, 6 oder mehr.

$j$	$a_{j,1}$	$a_{j,2}$	$a_{j,3}$	$a_{j,4}$
1	1	0	0	0
2	0,8333	0,1667	0	0
3	0,8333	0,1389	0,0278	0
4	0,8333	0,1389	0,0231	0,0046
5	0,8295	0,1389	0,0231	0,0085
6	0,8263	0,1382	0,0231	0,0123
25	0,7671	0,1283	0,0215	0,0831
50	0,6956	0,1164	0,0195	0,1685
75	0,6308	0,1056	0,0177	0,2460
100	0,5721	0,0957	0,0160	0,3162

Tab. 1: Ausgewählte Zustandsvektoren zu Aufgabe 1

Zur Klärung dieser Frage formulieren wir folgende Aufgabe.

**Aufgabe 2:** Wie wahrscheinlich ist es, in 100 Würfeln mindestens einen Run mit *genau* der Länge 4 zu erzielen?

Jetzt fragen wir nach einem „echten“ Run der Länge 4. Der Run darf nicht länger als 4 werden. In Aufgabe 1 war dies möglich, es wurde nach Runs von mindestens 4 gefragt. Dazu ändern wir den Graphen zu Aufgabe 1 leicht ab (Abb. 2).

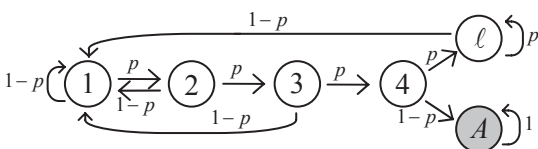


Abb. 2: Übergangsgraph zu Aufgabe 2

Der Zustand  $\textcircled{4}$  ist jetzt ein Durchgangszustand, den man auf zwei verschiedenen Wegen verlassen kann. Nach vier gleichen Zahlen bleibt es nur dann bei einem Run der Länge 4, wenn die Augenzahl wechselt, womit der absorbierende Zustand  $\textcircled{A}$  erreicht wird; fällt dagegen dieselbe Zahl noch einmal, dann entsteht zunächst ein längerer Run (Zustand  $\textcircled{\ell}$ ), ehe mit dem Wechsel zu einer neuen Augenzahl der Aufbau eines neuen Runs im Zustand  $\textcircled{1}$  beginnt. Entsprechend ändern sich die Rekursionsformeln für die Koordinaten  $b_{j+1;z}$  des Zustandsvektors:

$$b_{j+1;1} = (1-p) \cdot (b_{j,1} + b_{j,2} + b_{j,3} + b_{j,\ell});$$

$$b_{j+1;2} = p \cdot b_{j,1};$$

$$b_{j+1;3} = p \cdot b_{j,2};$$

$$b_{j+1;4} = p \cdot b_{j,3};$$

$$b_{j+1;\ell} = p \cdot (b_{j,4} + b_{j,\ell});$$

$$b_{j+1;A} = (1-p) \cdot b_{j,4} + b_{j,A}.$$

Tabelle 2 zeigt ausgewählte Zustandsvektoren  $\vec{b}_j$  für Aufgabe 2. Am Ende einer Serie von 100 Würfeln befindet man sich mit der Wahrscheinlichkeit 0,003 im Zustand  $\textcircled{4}$  und mit 0,268 im Zustand  $\textcircled{A}$ . In beiden Fällen ist das Ziel, mindestens ein Run der genauen Länge 4, erreicht, insgesamt mit der Wahrscheinlichkeit 0,271. Die Differenz von 0,045 gegenüber der Lösung von Aufgabe 1 entfällt auf solche Serien, die keine echten Vierer-Runs, dafür aber mindestens einen längeren Run enthalten.

$j$	$b_{j,1}$	$b_{j,2}$	$b_{j,3}$	$b_{j,4}$	$b_{j,\ell}$	$b_{j,A}$
1	1,000	0	0	0	0	0
2	0,833	0,167	0	0	0	0
3	0,833	0,139	0,028	0	0	0
25	0,777	0,130	0,022	0,004	0,001	0,067
50	0,717	0,120	0,020	0,003	0,001	0,139
75	0,661	0,110	0,018	0,003	0,001	0,207
100	0,609	0,102	0,017	0,003	0,001	0,268

Tab. 2: Ausgewählte Zustandsvektoren zu Aufgabe 2

### 3 Verteilung der Zufallsgröße $X_{n,\lambda}$

Wir verallgemeinern die Fragen aus Abschnitt 2 zu folgender Aufgabe.

**Aufgabe 3:** Wie wahrscheinlich ist es, dass beim  $n$ -maligen Würfeln 0, 1, 2, ... Runs auftreten, die genau die Länge  $\lambda = 4$  haben?

Gesucht ist also die *Verteilung* einer Zufallsgröße, nämlich  $X_{n,\lambda} :=$  „Anzahl der  $\lambda$ -Runs beim  $n$ -maligen Würfeln“ (bezogen auf die Aufgaben 1 und 2 ist das  $X_{100;4}$ ).

Der Übergangsgraph von Abbildung 3 veranschaulicht die neue Situation. Die Bezeichnung  $\textcircled{k;a}$  der Zustände besagt, dass bis dahin  $k$  Vierer-Runs aufgetreten sind und der aktuell begonnene Run die Länge  $a$  hat (oder  $\ell > 4$ ). Im Unterschied zu Aufgabe 2 gibt es keinen absorbierenden Zustand mehr; die Rolle von  $\textcircled{4}$  in Abbildung 2 übernimmt  $\textcircled{1;0}$ , von wo man wie oben nach  $\textcircled{0;\ell}$  und statt nach  $\textcircled{A}$  nun in den neuen Durchgangszustand  $\textcircled{1;1}$  gelangen kann. Damit ist ein erster Vierer-Run abgeschlossen. Analog ist die Situation in den Zuständen  $\textcircled{2;0}$ ,  $\textcircled{3;0}$  usw.

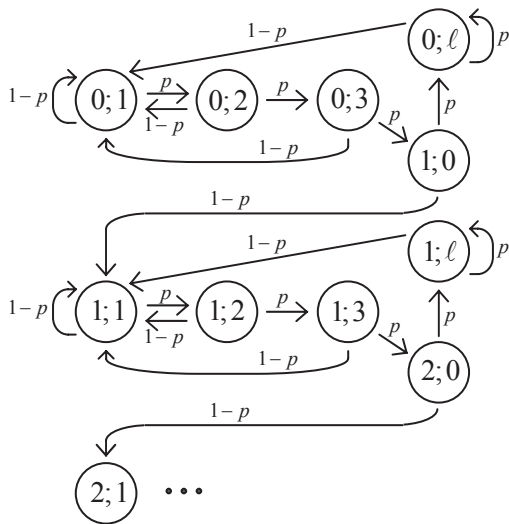


Abb. 3: Übergangsgraph zu Aufgabe 3

Wir ordnen die Koordinaten des Zustandsvektors  $\vec{c}_j$ , der nun die Dimension  $n$  hat, nach Abbildung 3 an:  $(0;1), (0;2), (0;3), (0;0), (1;0), (1;1)$  usw.; der Startvektor nach dem ersten Wurf ist dann  $\vec{c}_1 = (1; 0; 0; 0; \dots)$ .

Die Koordinaten von  $\vec{c}_{j+1}$  lassen sich aus denen von  $\vec{c}_j$  wieder mithilfe von Rekursionsformeln berechnen, die man Abbildung 3 entnimmt. Wir schreiben diese Formeln so, dass man sie direkt in EXCEL übernehmen kann (Abb. 4):

$p$  und  $1-p$  stehen in den Zellen C5 bzw. F5 und der Startvektor  $\vec{c}_1$  in Zeile 8; in den Spalten F, L, R und X steht jeweils die Wahrscheinlichkeit, nach  $j$  Würfen  $k$  Vierer-Runs zu haben ( $k \in \{0, 1, 2, 3\}$ ).

Für den Übergang von  $j = 1$  nach  $j = 2$  (also von Zeile 8 nach Zeile 9) entnimmt man Abbildung 3:

- Für  $(0;1)$ : B9=\$F\$5\*SUMME(B8:E8),
- für  $(0;2)$ : C9=\$C\$5\*B8,
- für  $(0;3)$ : D9=\$C\$5\*C8,
- für  $(0;0)$ : E9=\$C\$5\*(E8+G8),
- $k = 0$ : F9=SUMME(B9:E9),
- für  $(1;0)$ : G9=\$C\$5\*D8.

Kopiert man nun die Zellen B9 bis G9 nach rechts (H9 bis M9), so stehen in den Spalten I, J, K und M bereits

die korrekten Formeln für  $(1;2), (1;3), (1;0)$  und  $(2;0)$  (man beachte die absolute Adressierung der Zellen C5 und F5, in denen die gleich bleibenden Werte von  $p$  und  $1-p$  stehen). Die Formeln für  $(1;1)$  und  $k = 1$  erfordern folgende Korrekturen:

Für  $(1;1)$ : H9=\$F\$5\*SUMME(G8:K8) anstatt  
 $H9=$F$5*SUMME(H8:K8)$   
 (aktuell markierte Zelle in Abb. 4),

$k = 1$ : L9=SUMME(G9:K9) anstatt  
 $L9=SUMME(H9:K9)$ .

Nun kann man H9 bis M9 ohne weitere Korrekturen nach N9 bis S9 und T9 bis Y9 (oder noch weiter) kopieren.

Die sehr kleinen Wahrscheinlichkeiten für  $k \geq 4$  fassen wir in Spalte Z zusammen:

$$k \geq 4: Z9=1-(F9+L9+R9+X9).$$

Abschließend kopiert man Zeile 9 für  $j = 3$  bis 100 in die 98 folgenden Zeilen nach unten. Auch hier ergibt sich  $P(k = 0) \approx 0,729$  wie am Ende von Abschnitt 2 (Zelle F107), doch sind jetzt differenziertere Aussagen zum Gegenereignis  $k > 0$  möglich.

Auch den Erwartungswert  $\mu$  kann man nun berechnen; für unser Beispiel ist  $\mu = E(X_{100;4}) \approx 0,313$ .

## 4 Varianten des Problems

In diesem Abschnitt betrachten wir einige Varianten und Verallgemeinerungen unserer Aufgaben.

### Länge der Serie

Eine Änderung der Serienlänge braucht man nicht eigens zu untersuchen, denn beim rekursiven Berechnen einer Verteilung  $X_{n;\lambda}$  durchläuft man ja auch alle Verteilungen mit  $j < n$ , die man dann aus einer Tabelle wie in Abbildung 4 entnehmen kann.

### Trefferwahrscheinlichkeit

Unsere Abbildungen 1 bis 3 sind bereits auf eine Variation der Trefferwahrscheinlichkeit hin angelegt,

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	V	W	X	Y	Z
1																			
2		Nenner der Trefferwahrscheinlichkeit																	
3		$p = 1/n$ eingeben: $n = 6$																	
4																			
5		p = 0,167      1-p = 0,833																	
6																			
7	j	0;1	0;2	0;3	0;0	k=0	1;0	1;1	1;2	1;3	1;0	k=1	2;0	2;1	3;3	3;0	k=3	4;0	k ≥ 4
8	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
9	2	0,833	0,167	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
10	3	0,833	0,139	0,028	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
106	99	0,611	0,102	0,017	6E-04	0,731	0,003	0,191	0,032	0,005	2E-04	0,231	9E-04	0,028	7E-05	2E-06	0,003	1E-05	2E-04
107	100	0,609	0,102	0,017	6E-04	0,729	0,003	0,193	0,032	0,005	2E-04	0,233	9E-04	0,028	7E-05	2E-06	0,003	1E-05	2E-04

Abb. 4: EXCEL-Tabelle für die Verteilung von  $X_{100;4}$  (Spalten O bis U und Zeilen 11 bis 105 ausgeblendet)

indem dort als Übergangswahrscheinlichkeiten  $p$  und  $1-p$  statt der konkreten Werte  $1/6$  und  $5/6$  eingetragen sind. Die Graphen gelten daher auch für Folgen von Zufallsziffern ( $p = 1/10$ ) oder Münzwürfen ( $p = 1/2$ ). Die Verteilungen von  $X_{100;4}$  für diese Fälle zeigt Tabelle 3.

$p$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k \geq 5$
1/10	0,924	0,073	0,003	$6 \cdot 10^{-5}$	$1 \cdot 10^{-6}$	$1 \cdot 10^{-8}$
1/6	0,729	0,233	0,035	0,003	$2 \cdot 10^{-4}$	$1 \cdot 10^{-5}$
1/2	0,035	0,127	0,221	0,243	0,189	0,186

Tab. 3: Verteilungen von  $X_{100;4}$  für verschiedene  $p$

### Erwartete Zahl der Runs

Auch die Erwartungswerte für die neuen Werte von  $p$  kann man nun berechnen; für  $p = 0,1$  ist  $\mu \approx 0,08$  und für  $p = 0,5$  errechnet man (mit mehr Werten als in Tabelle 3 wiedergegeben)  $\mu \approx 3,1$ .

### Runlänge

Als Nächstes variieren wir die Runlänge  $\lambda$ . Für  $\lambda = 5$  ist beispielsweise Zustand  $\textcircled{4}$  in Abbildung 1 ein innerer und erst  $\textcircled{5}$  der absorbierende Zustand, und in Abbildung 3 ist entsprechend für jedes  $k$  ein neuer Zustand  $\textcircled{k;4}$  hinter  $\textcircled{k;3}$  einzufügen.

Mit den entsprechend geänderten EXCEL-Formeln findet man für die Variante von Aufgabe 1 den der letzten Zeile aus Tabelle 1 entsprechenden Zustandsvektor  $\vec{a}_{100}$ :

$j$	$a_{j,1}$	$a_{j,2}$	$a_{j,3}$	$a_{j,4}$	$a_{j,5}$
100	0,7837	0,1307	0,0218	0,0036	0,0601

Tab. 4: Zustandsvektor nach 100 Schritten für Runs der Länge 5

Für einen Run der Mindestlänge 5 erhalten wir also wie schon Motzer (2010) den Wert  $P(z = 5) \approx 0,06$ .

Für die Variante von Aufgabe 3 mit  $\lambda = 5$  ergeben sich die Verteilungen von Tabelle 5.

$p$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k \geq 5$
1/10	0,992	0,008	$3 \cdot 10^{-5}$	$6 \cdot 10^{-8}$	$8 \cdot 10^{-11}$	$8 \cdot 10^{-14}$
1/6	0,950	0,049	0,001	$2 \cdot 10^{-5}$	$2 \cdot 10^{-7}$	$1 \cdot 10^{-9}$
1/2	0,198	0,340	0,271	0,133	0,045	0,013

Tab. 5: Verteilungen von  $X_{100;5}$  für verschiedene  $p$

Wie bei  $\mu = 4$  ist beim Würfeln ein *echter* Fünfer-Run mit  $P(X_{100;5} \geq 1) \approx 0,05$  weniger wahrscheinlich als ein Run der *Mindestlänge* 5 mit  $\vec{a}_{100;5} \approx 0,06$ .

## 5 Verallgemeinerung des Problems

Die bisherigen Beispiele betrafen Laplace-Experimente; Runs waren dabei Folgen irgendeines der gleichwahrscheinlichen Ergebnisse.

Aufgabe 4: Als Run wird jetzt nur eine unmittelbare Folge eines bestimmten Ergebnisses gewertet (entspricht einer Folge von Treffern bei einem Bernoulli-Experiment). Die Zufallsgröße *Anzahl der Vierer-Runs* nennen wir nun  $Y$  und fragen, wie  $P(Y_{100;4} = k)$  allgemein von  $p$  abhängt.

Wir könnten beim Würfeln etwa nur die Sechsen betrachten und ihre Runs zählen. Im Unterschied zu den bisherigen Aufgaben ist nun prinzipiell jeder Wert von  $p$  möglich.

### Wie hängt die Wahrscheinlichkeit für $k$ Runs von $p$ ab?

Der Übergangsgraph zeigt gegenüber Abbildung 3 einige Änderungen (Abb. 5). Wieder bedeutet  $\textcircled{k;a}$ , dass bis dahin  $k$  Vierer-Runs aufgetreten sind und der aktuell begonnene Run die Länge  $a$  hat (oder  $\ell > 4$ ). Die aus dem Graphen abzulesenden Rekursionsformeln überträgt man wieder analog zu Abbildung 4 in eine EXCEL-Tabelle.

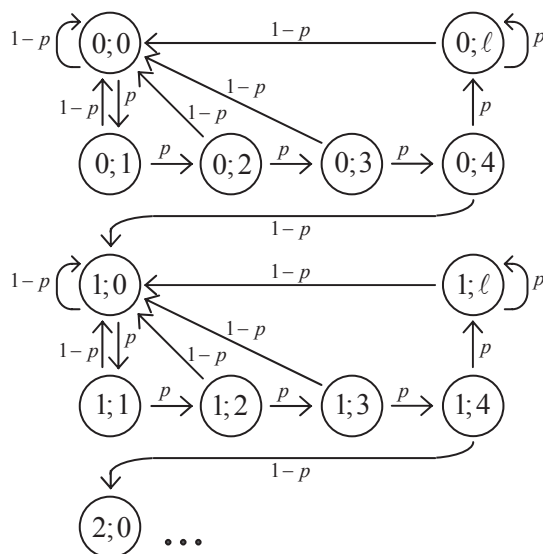


Abb. 5: Übergangsgraph für Runs eines Einzelausfalls

Um in EXCEL bequem den Parameter  $p$  variieren zu können, ist ein Schieberegler zweckmäßig. Man ändert  $p$  bei festem  $\lambda = 4$  schrittweise und kopiert jeweils die für  $P(Y_{100;4} = k)$  gewonnenen Werte in eine neue Tabelle, die wir als Wertetabelle einer Schar von Funktionen

$$f_k: \quad p \mapsto f_k(p) = P(Y_{100;4} = k)$$



auffassen können. Die EXCEL-Option *Diagramm* liefert dann die zugehörigen Graphen; Abbildung 6 zeigt diese für  $k \in \{0; 1; 2\}$ .

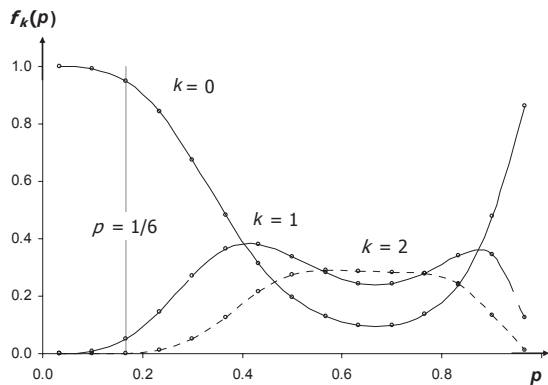


Abb. 6: Wahrscheinlichkeit für  $k$  Runs der Länge 4 – abhängig von  $p$

Diese Graphen sind zwar nicht perfekt symmetrisch, zeigen aber für  $p$  nahe 1 ein ähnliches Verhalten wie nahe 0. Für kleine  $p$  sind vier Treffer nacheinander sehr unwahrscheinlich, und für  $p$  nahe 1 werden längere Runs immer wahrscheinlicher. In beiden Fällen ist also kaum ein Vierer-Run zu erwarten, in Abbildung 6 erkennbar an  $f_0(p) \approx 1$ . Für  $k > 0$  muss in diesen Bereichen daher  $f_k(p) \approx 0$  sein. Bei mittleren Werten von  $p$  dagegen ist  $f_k(p)$  relativ klein während  $f_1(p)$  und  $f_2(p)$  in diesem Bereich ihre größten Werte annehmen, d. h., sehr wahrscheinlich tritt mindestens ein Vierer-Run ein.

Ist die Trefferwahrscheinlichkeit ein Stammbruch, so besteht ein Zusammenhang mit unseren früheren Ergebnissen (Tab. 3), wie wir nun an einem Beispiel zeigen: Es sind  $f_0(1/6) \approx 0,949$  und  $f_1(1/6) \approx 0,050$  die Wahrscheinlichkeiten dafür, keinen bzw. genau einen Vierer-Run einer *bestimmten* Augenzahl zu erreichen (vgl. Abb. 6). Die Wahrscheinlichkeit, genau einen Vierer-Run *irgendeiner* Augenzahl zu haben, kann man nun folgendermaßen schätzen: Eine der 6 Zahlen liefert den Run, aber keine der 5 anderen darf ebenfalls einen Vierer-Run bilden; daher ist

$$P(X_{100;4} = 1) \approx 6 \cdot 0,05 \cdot 0,949^5 \approx 0,231.$$

Dieser Wert stimmt gut mit dem entsprechenden aus Tabelle 3 überein; die Abweichung rührt daher, dass der Faktor 0,949 die *totale* Wahrscheinlichkeit für  $k = 0$  ist, während man eigentlich mit der durch den schon vorhandenen Vierer-Run *bedingten* Wahrscheinlichkeit rechnen müsste, die etwas größer ist.

### Überlegungen zur erwarteten Zahl von $\lambda$ -Runs

In ähnlicher Weise kann man untersuchen, wie die Anzahl der  $\lambda$ -Runs in einer 100er-Serie, also der Er-

wartungswert  $E(Y_{100;\lambda})$ , von  $p$  abhängt. Abbildung 7 zeigt für ausgewählte Werte von  $\lambda$  die Graphen der Funktionen  $g_\lambda$  mit

$$g_\lambda(p) = E(Y_{100;\lambda}).$$

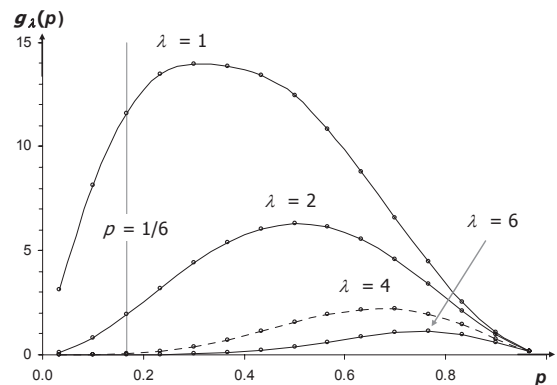


Abb. 7: Erwartete Zahl von Runs der Länge  $\lambda$  – abhängig von  $p$

Auch hier besteht ein enger Zusammenhang zu den Ergebnissen aus Abschnitt 4. Wir vergleichen die Erwartungswerte von  $X_{100;\lambda}$  und  $Y_{100;\lambda}$  beim Werfen eines Laplace-Würfels ( $p = 1/6$ , Tab. 6).

$\lambda$	1	2	3	4	5	6
$E(X_\lambda)$	69,70	11,50	1,90	0,31	$5,2 \cdot 10^{-2}$	$8,5 \cdot 10^{-3}$
$E(Y_\lambda)$	11,59	1,92	0,32	0,05	$8,6 \cdot 10^{-3}$	$1,4 \cdot 10^{-3}$

Tab. 6: Erwartungswerte von  $X_{100;\lambda}$  und  $Y_{100;\lambda}$  für  $p = 1/6$

Die Werte für  $X_\lambda$  sind jeweils 6-mal so groß wie die von  $Y_\lambda$ , da ja jede der 6 Augenzahlen „Treffer“ sein kann. Entsprechendes gilt in den Fällen  $p = 0,1$  und  $p = 0,5$ , wo  $E(X_\lambda)$  10-mal bzw. doppelt so groß ist wie  $E(Y_\lambda)$ .

Eine ähnliche Beobachtung macht man, wenn man  $E(Y_\lambda)$  mit  $E(Y_{\lambda+1})$  in Tabelle 5 bzw. in Abbildung 7 vergleicht: Nimmt  $\lambda$  um 1 zu, so ist  $\mu_{\lambda+1}$  ungefähr  $1/6$  von  $\mu_\lambda$ . Auch dies ist leicht zu verstehen, denn die Wahrscheinlichkeit für einen  $\lambda$ -Run nach einem Fehlwurf ist  $p^\lambda \cdot (1-p)$ , bei einem  $(\lambda+1)$ -Run kommt noch ein Faktor  $p$  hinzu.

Die letzte Überlegung führt auf eine neue Möglichkeit, einen Näherungswert für  $E(Y_\lambda)$  zu bestimmen. Es sei  $r := E(Y_1)$ , dann ist  $E(Y_\lambda) = r \cdot p^{\lambda-1}$ . Man kann also in einer Serie von  $n$  Würfeln bei großem  $n$  etwa

$$t = r \cdot \sum_{\lambda=1}^{\infty} \lambda \cdot p^{\lambda-1} = r \cdot (1 + 2p + 3p^2 + 4p^3 + \dots)$$

Treffer erwarten. Andererseits gilt aber auch  $t \approx n \cdot p$ . Wie man durch Multiplikation mit  $(1 - 2 \cdot p + p^2)$  bestätigt, hat die Summe in der Klammer den Wert  $\frac{1}{(1-p)^2}$ , so dass man die Beziehung

$$\frac{r}{(1-p)^2} \approx n \cdot p$$

und damit für  $r$  den Näherungswert

$$n \cdot p \cdot (1-p)^2$$

erhält. Für das Beispiel  $n = 100$  und  $p = 1/6$  ergibt sich so  $r \approx 11,574$ , der exakte Wert ist 11,589.

## 6 Schlussbemerkungen

Als im Zuge der Oberstufenreform der 70er-Jahre die Stochastik einen höheren Stellenwert im Mathematikunterricht erhielt, waren auch Markow-Ketten groß in Mode und im Sinne von Anwendungsorientierung der Stochastik Thema verschiedener Lehrbücher und Kursvorschläge (Lehmann 1973, Kittler 1975, Engel 1976). Die Bedeutung des Themas, zumindest die ihm zuerkannt wurde, kann man auch an Scheid (1986) ablesen, der ein ganzes Kapitel der „Stochastik auf der Kollegstufe“ dazu ausarbeitet. Der Autor selbst hat sich wiederholt mit Markow-Ketten im Schulunterricht beschäftigt, wie etwa schon in Riehl (1982).

Bemerkenswert ist, dass sich das Wissen um Markow-Ketten – nicht nur im deutschsprachigen Raum – doch nicht so weit verbreitet hat. So geriet Holmes (1991) bei der Analyse von Runs in methodische Schwierigkeiten, was darauf zurückzuführen war, dass er das Problem eben nicht mit Markow-Ketten beschrieb (siehe Riehl 1994).

Ein Grund für die mangelnde Akzeptanz war wohl, dass damals leistungsfähige Hilfsmittel für die notwendigen umfangreichen Rechnungen noch nicht allgemein zur Verfügung standen. Auch sah man die Anforderungen aus der Linearen Algebra als zu umfassend an, sodass man sich entweder wieder vom Thema abgewendet oder mit durch Simulation gewonnenen näherungsweise Lösungen begnügt hat, die jedoch immer unbefriedigend bleiben.

Mit den neuen Rechenhilfen wie einer Tabellenkalkulation kann man nun – fast unter Umgehung von Linearer Algebra – Berechnungen leicht durchführen und wesentliche Erkenntnisse gewinnen, wie der vorliegende Beitrag auch demonstrieren soll.

Der wesentliche Vorzug des hier dargestellten Lösungsweges besteht in der Veranschaulichung der stochastischen Situation durch Übergangsgraphen, die nicht nur „begabten Schülerinnen und Schülern“ zugänglich sind, wie es Motzer (2010) für das von ihr dargestellte Vorgehen einschätzt.

Da sich die Rekursionsformeln für die rechnerische Lösung unmittelbar aus den Graphen ablesen und

dann in eine EXCEL-Tabelle umsetzen lassen, ist es nicht notwendig, eine Theorie der Markow-Ketten aufzubauen; insbesondere kann man auf Matrizen völlig verzichten. Wie man typische Fragestellungen aus der Theorie der Markow-Ketten (Absorptionswahrscheinlichkeiten, mittlere Wartezeiten) im Unterricht mit Matrizen behandeln kann, habe ich an anderer Stelle ausgeführt (Riehl 2010).

## Literatur

- Engel, A. (1976): *Wahrscheinlichkeitsrechnung und Statistik*, Bd. 2. Stuttgart: Klett.
- Holmes, P. (1991): Computer-erzeugtes Denken. In: *Stochastik in der Schule* 11 (3), S. 37–41.
- Kittler, H. (1975): Grundkurs: Einführung in die Wahrscheinlichkeitsrechnung unter besonderer Berücksichtigung Markoffscher Ketten. In: Niedersächsisches Kultusministerium (Hrsg.): *Handreichungen für Lernziele, Kurse und Projekte im Sekundarbereich II für das mathematisch-naturwissenschaftliche Aufgabenfeld*, 3. Folge, S. 83–92.
- Lehmann, E. (1973): *Endliche homogene Markoffsche Ketten – eine Anwendung von Wahrscheinlichkeits- und Matrizenrechnung*. München: Bayerischer Schulbuch-Verlag.
- Motzer, R. (2010): Serien von gleichen Würfelzahlen. In: *Stochastik in der Schule* 30 (3), S. 35–36.
- Riehl, G. (1982): Elementare Behandlung genetischer Probleme im Stochastikunterricht mit Hilfe von Graphen. In: *Didaktik der Mathematik* 10 (1), S. 25–42.
- Riehl, G. (1994): Noch einmal: Computer-erzeugtes Denken. In: *Stochastik in der Schule* 14 (3), S. 13–19.
- Riehl, G. (2010): Markow-Ketten, eine alternative Methode zur Lösung stochastischer Probleme. In: *MNU* 63 (7), S. 401–404.
- Scheid, H. (1986): *Stochastik auf der Kollegstufe*. Mannheim: Bibliographisches Institut.

## Anmerkungen

Mein Dank gilt den beiden Gutachtern und Herrn Borovnik als Heftherausgeber für hilfreiche Hinweise und Anregungen.

Interessenten finden die Zusammenstellung einer EXCEL-Mappe mit Tabellenblättern für die hier behandelten Aufgaben sowie für deren Simulation unter [wwwg.uni-klu.ac.at/stochastik.schule/Boro/index\\_inhalt.htm](http://wwwg.uni-klu.ac.at/stochastik.schule/Boro/index_inhalt.htm).

## Anschrift des Verfassers

Gerd Riehl  
Obere Mark 6  
30890 Barsinghausen  
[Elfriede.Riehl@t-online.de](mailto:Elfriede.Riehl@t-online.de)