Manfred BOROVCNIK, Klagenfurt

Rekursive Zugänge und ihr Potential zur Modellbildung

Für das Rosinen- bzw. Sammelbildproblem wird eine rekursive Lösung vorgeschlagen, die sich in einer Tabellenkalkulation einfach implementieren lässt. Es bietet sich an, das Tabellenblatt zu ergänzen und weitere Kenngrößen zu berechnen. Damit bekommt man auf einfache Weise die Möglichkeit, Lösungen in ihrer Abhängigkeit von Eingangsparametern (auch graphisch) zu studieren. Aus dem Tabellenblatt ergeben sich übersichtlich die Lösungen einer Reihe verwandter Problemstellungen.

1. Vernetzungen in der Mathematikausbildung

Thema der Arbeit ist, Aufgabenstellungen der Wahrscheinlichkeitsrechnung „lokal einsichtig zu modellieren", um damit „globale Zusammenhänge" berechenbar zu machen. Alternativ gibt es dazu den Ansatz der "globalen" mathematischen Herleitung der Lösung. Oder: Nachspielen der Annahmen des Modells betreffend der beteiligten Wahrscheinlichkeiten durch Simulation und die Schätzung von Ergebniswahrscheinlichkeiten durch relative Häufigkeiten. Die Schwierigkeit, wesentliche Annahmen der Modellierung einzusehen, sowie die Leichtigkeit, mit der man den Einfluss von Eingangsparametern auf die Lösung studieren kann, werden als Prüfsteine der Ansätze herangezogen.

Exemplarisch wird der Aufgabenkomplex „Rosinen oder Sammelbilder" herangezogen. Aus einer Serie von N Bildern sammelt man einige (durch Auswahl etc.). Natürlich modelliert man mit Gleichwahrscheinlichkeit sowie Unabhängigkeit der Auswahlen. Eine von vielen Fragen lautet: „Wie viele Bilder muss man kaufen, bis man eine vollständige Serie mit allen N Bildern hat."

Die beteiligte Kombinatorik hat so ihre Tücken. Zur Berechnung des Erwartungswerts sind unendliche Summen, eigentlich Reihen, erforderlich; das macht es noch schwerer, globale Lösungsansätze zu verfolgen. Es verwundert daher gar nicht, wenn man sich aus didaktischer Sicht nach einfacheren Alternativen umsieht. Eine besteht in der Simulation der Auswahlen (Käufe). Hier wird ein rekursiver Lösungsansatz vorgeschlagen, der ein weitreichendes Werkzeug der modernen Mathematik darstellt.

In den KMK-Bildungsstandards werden u. a. folgende Leitideen genannt:

Daten und Zufall: graphische Darstellungen auswerten, systematisch Daten sammeln, Hilfsmittel (Software) nutzen, Daten interpretieren, Argumente reflektieren und bewerten

Struktur: Strukturen erkennen

Zahl: Wahrscheinlichkeiten, Visualisierung von Zahlen, Lage- und Streuungsparameter, Daten als Zahlen im Kontext

Funktionale Zusammenhänge: Stochastisch-funktionale Zusammenhänge

Zu all diesen Themen finden sich Anknüpfungspunkte in der vorliegenden Arbeit: *Daten systematisch sammeln*, Strukturen erkennen, Problemtypen einordnen, Ergebnisse visualisieren, die Ergebnisse im Kontext interpretieren, und schließlich, Ergebnisse in ihrer Abhängigkeit von Eingangsparametern studieren. Allerdings werden Daten nicht in Form einer Stichprobe gewonnen sondern vielmehr durch strukturelle Analyse eines kombinatorischen Problems erzeugt. Darüber hinaus werden alle Grundtypen kombinatorischer Aufgaben und Fragestellungen (Geburtstagsproblem, Problem der vollständigen Serie im Sammelalbum, Rosinenproblem) in ihren Varianten einheitlich behandelbar.

Darüber hinaus sei auf den grundlegenden Ansatz einer rekursiven Umformulierung eines Problems (sofern dies möglich ist) verwiesen: der rekursive Charakter der Definition der natürlichen Zahlen, die Bedeutung linearer Rekursionen (wie z. B. die Fibonnacci-Folgen), weiters rekursive Algorithmen (etwa für die „Türme von Hanoi"; solche Algorithmen sind aus modernen Anwendungen einfach nicht mehr wegzudenken).

In der Stochastik gibt es neben der kombinatorischen Herleitung der Binomialverteilung auch den Weg über die Motivation am Galton-Brett und daran anlehnend eine rekursive Herleitung, die im wesentlichen der bekannten Rekursion der Binomialkoeffizienten folgt. Man kann die rekursive Umformulierung des Binomialproblems prototypisch sehen für viele kombinatorische Aufgaben, die in einer rekursiven Variante lokal einer leicht einsehbaren modellhaften Rekursion unterliegen, welche, konsequent durchgerechnet, das Ergebnis (ohne Formeln) liefert.

Rekursive Zusammenhänge sind ausbaufähig bis hin zu Markow-Ketten. Man kann meist geschlossene mathematische Lösungen anbieten. Aber es erscheint (didaktisch) wenig sinnvoll, weil diese auf den (leicht einsehbaren) lokalen rekursiven Zusammenhang meist nicht eingehen, sondern gleich mit ihrem Modell global ansetzen. Das führt mitunter zu schwer einsehbaren Annahmen, wie der künstlichen Unterscheidung von Teilchen, die man nie und nimmer unterscheiden kann.

2. Varianten des Sammelbildproblems und kombinatorische Lösung

#1: Wie viele Rosinen muss man dem Teig für $N = 6$ Semmeln beimengen, damit sich mit einer Wahrscheinlichkeit von mindestens 99% in allen Semmeln wenigstens eine Rosine befindet? In der Sammelbild-Version: Wann, d. h., nach wie vielen Auswahlen ist das Sammelalbum mit $N = 6$ Bildern mit Sicherheit 99% voll?

#2: Wie groß ist die Wahrscheinlichkeit, dass bei Beimengen von n (≤ 6) Rosinen keine Semmel zwei Rosinen hat? (Entspricht Geburtstagsproblem)

#3: Wie groß ist die Wahrscheinlichkeit, dass man durch Beimengen von n (≥ 6) Rosinen in allen Semmeln wenigstens eine Rosine findet? (Entspricht der vollständigen Serie im Sammelbildproblem.)

#4: Wie groß ist die Wahrscheinlichkeit, dass man durch Beimengen von $n = N$ Rosinen, in allen Semmeln eine Rosine findet? (Spezialfall; leicht zu lösen.).

#5: Wie groß ist die erwartete Anzahl von Rosinen, die man benötigt, damit man in allen Semmeln eine Rosine findet? Wie groß ist die Streuung dieser Zahl?

Die Aufgaben entsprechen den Grundtypen kombinatorischer Aufgaben:

1: Ziehen von n Kugeln aus einer Urne mit N Kugeln

2: Verteilen von n Murmeln auf N Zellen (Behälter)

Das Sammelbildproblem entspricht „eher“ dem Typ 1: Man zieht aus dem vollständigen Satz des Sammelalbums zufällig ein Bild („das wird dann eingeklebt“). Das Rosinenbeispiel entspricht dem Verteilungstyp 2; n Rosinen werden auf N Semmeln verteilt.

Die Zusatzfragen bei der Modellierung entsprechen einander 1:1

1: *Ziehen* n *Kugeln aus Urne mit* N
- Reihenfolge wesentlich?
- Ziehen ohne/mit Zurücklegen?

2: *Verteilen* n *Murmeln auf* N *Zellen*
- Sind die Murmeln unterscheidbar?
- Mehrfach belegen?

Die Fragestellungen haben extrem unterschiedliche Schwierigkeitsgrade. Hier wird ein rekursiver Ansatz vorgestellt, der die Verteilung von Rosinen (Murmeln) auf die Semmeln (Zellen) als Irrfahrt in der Ebene darstellt.

3. Die Irrfahrt

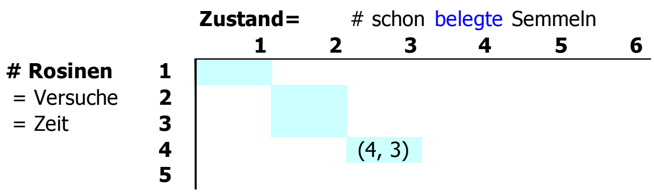

Abb. 1

1. Rosine belegt eine Semmel
2. eine weitere
3. "erwischt" eine schon belegte
4. eine freie

- (4,3) = mit 4 Rosinen wurden 3 (verschiedene) Semmeln belegt.
- Von (4,3) erreicht man nur (5,3) oder (5,4).
- Zu (5,4) kommt man nur über (4,3) oder (4,4).
- Irrfahrt startet beim Knoten (1,1).
- Verzweigung mit Wahrscheinlichkeiten entsprechend besetzter /freier Semmeln.

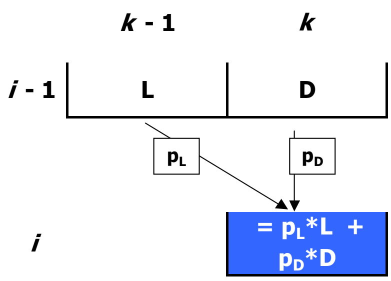

Abb. 2

Für die i-te Rosine gilt: (i, k) kann nur über **L** und **D** (s. Abb. 2) erreicht werden. Bezeichnet man der Einfachheit halber die Wahrscheinlichkeit dafür, **L** bzw. **D** in der Irrfahrt zu erreichen, mit denselben Buchstaben, so ergibt sich die Rekursion in Abb. 2.

Anzahl Ros	Anzahl schon belegter Semmeln 1	2	3	4	5	6	P P(X=*n*)	E(X) *n*∗P(X=*n*)	$E(X^2)$ n^2∗P(X=*n*)	√ V(X)
1	1	0	0	0	0	0	0,000	0,000	0,000	
2	0,167	0,833	0	0	0	0	0,000	0,000	0,000	
3	0,028	0,417	0,556	0	0	0	0,000	0,000	0,000	
4	0,005	0,162	0,556	0,278	0	0	0,000	0,000	0,000	
5	0,001	0,058	0,386	0,463	0,093	0	0,000	0,000	0,000	
6	0,000	0,020	0,231	0,502	0,231	0,015	0,015	0,093	0,556	
7	0,000	0,007	0,129	0,450	0,360	0,054	0,039	0,270	1,890	
35	0,000	0,000	0,000	0,000	0,010	0,990	0,002	0,071	2,482	
36	0,000	0,000	0,000	0,000	0,008	0,992	0,002	0,061	2,190	
37	0,000	0,000	0,000	0,000	0,007	0,993	0,001	0,052	1,928	
38	0,000	0,000	0,000	0,000	0,006	0,994	0,001	0,045	1,695	
65	0,000	0,000	0,000	0,000	0,000	1,000	0,000	0,001	0,036	
66	0,000	0,000	0,000	0,000	0,000	1,000	0,000	0,000	0,031	
Einige Zeilen sind ausgeblendet							1,00	**14,70**	254,89	**6,24**

Abb. 3

4. Ergebnisse

Diese Rekursion realisiert man in einer Tabellenkalkulation ganz einfach. Im „Gitter“ (Abb. 3) stehen die Besuchswahrscheinlichkeiten der Knoten.

- In Spalte **6** findet sich die Verteilungsfunktion der Wartezeit auf die vollständige Serie. Sucht man in Spalte **6**, wo 0,99 übersprungen wird, hat man die Lösung von #1.
- Die „Diagonale“ löst das „Geburtstagsproblem“, dem hier #2 entspricht.
- Zeile ***n*** löst Aufgabe #3, mit *n* Rosinen schon alle Semmeln zu beteilen.
- Das „Diagonalelement“ in Spalte **6** löst #4, mit $n = N = 6$ Rosinen alle Semmeln zu beteilen.
- In Spalte **6** steht die Verteilungsfunktion der Wartezeit; daraus berechnet man diskrete Dichte, Erwartungswert und Standardabweichung der Wartezeit, das löst #5.

Die Wartezeit hat einen langen Ausläufer nach oben (Abb. 4). Interessant auch, wie ihr Erwartungswert von der Zahl der Semmeln abhängt (Abb. 5).

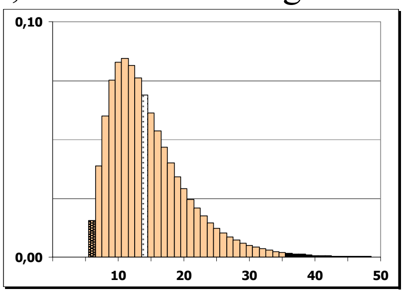

Abb. 4: Verteilung der Wartezeit

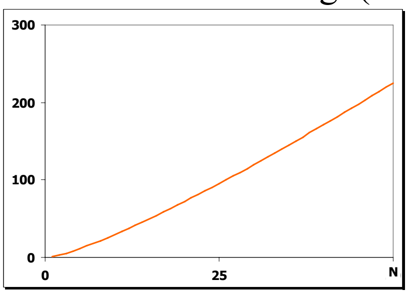

Abb. 5: Erwartung der Wartezeit, abhängig von N

5. Abschließende Bemerkungen

Der rekursive Ansatz ist sehr allgemein. Viele kombinatorische Probleme haben eine rekursive Variante. In einer Tabellenkalkulation kann man leicht studieren, wie Lösungen von Parametern abhängen; dies veranschaulicht man in einer Animation oder mittels Funktionsgraphen. Die Ergebnisse regen an, sich tiefer mit der Mathematik auseinander zu setzen.

Literatur

Borovcnik, M.: Das Sammelbildproblem–Rosinen und Semmeln und Verwandtes: Eine rekursive Lösung mit Irrfahrten. Erscheint in: Stochastik in der Schule 2007.