

Das Bingo-Problem – ein Lösungsweg mit schulischen Mitteln

GERD RIEHL, BARSINGHAUSEN

Zusammenfassung: Das von Henze (2012) mit aufwändigen Hilfsmitteln bearbeitete Bingo-Problem kann man auch mit schulischen Hilfsmitteln angehen und näherungsweise lösen. Benutzt wird dabei nur die Bedeutung der Binomialkoeffizienten als Anzahl der k -elementigen Teilmengen einer n -Menge sowie die vom Lotto-Problem her bekannte hypergeometrische Verteilung.

1 Einleitung

Henze (2012) beschreibt das von ihm untersuchte Bingo-Spiel folgendermaßen: „In einer in mehreren Bundesländern angebotenen Variante des Bingospiels werden von insgesamt 75 Zahlen 22 Zahlen rein zufällig ohne Zurücklegen gezogen. Die Bingo-Karte des Spielers in Form eines 5×5 -Quadrats weist 25 der 75 Zahlen auf. ... Die 5 Zeilen und Spalten sowie die beiden Diagonalen der Karte werden als *Linien* bezeichnet. Eine Linie heißt *Gewinnlinie*, wenn alle 5 Zahlen der Linie unter den 22 gezogenen sind. Bei genau einer Gewinnlinie gewinnt man im niedrigsten Rang (*Bingo*). Ein *Zweifach-Bingo* bzw. *Dreifach-Bingo* wird bei genau zwei bzw. mindestens drei Gewinnlinien erzielt.“

Henze bestimmt dann die Verteilung der Zufallsvariable $L := \text{Anzahl der Gewinnlinien}$, er differenziert also Dreifach-Bingo nach den einzelnen Werten für $L \geq 3$. Er weist selbst darauf hin, dass sein Beitrag eher als Lehrerfortbildung gedacht ist, der im Unterricht nur sehr begrenzt umgesetzt werden kann (die Hilfsmittel Indikatoren und Jordan-Formel fehlen).

Für den im Folgenden dargestellten Weg ist nur die Kenntnis von Binomialkoeffizienten als Anzahl der

k -elementigen Teilmengen einer n -Menge sowie der hypergeometrischen Verteilung erforderlich. Beide stehen in jeder Standard-Software, z. B. EXCEL, zur Verfügung.

2 Zerlegung des Problems

Das für das Bingo-Spiel charakteristische Zufallsexperiment (Ziehung von 22 aus 75 Zahlen) legt es nahe, die Zufallsvariable $X := \text{Trefferzahl unter den 25 Zahlen auf der Bingo-Karte}$ zu betrachten; diese kann offenbar die Werte $0 \leq X \leq 22$ annehmen. X ist hypergeometrisch verteilt, weil hier eine Ziehung ohne Zurücklegen vorliegt. Für hypergeometrische Verteilungen verwenden wir die Bezeichnung h und für deren Verteilungsfunktionen H ; bei den Parametern befolgen wir die Konventionen der EXCEL-Funktion HYPGEOM.VERT; es gilt also

$$P(X = t) = h(t; 22; 25; 75)$$

$$\text{und } P(X \leq t) = H(t; 22; 25; 75).$$

Einen Ausschnitt der Wahrscheinlichkeitsverteilung h und der Verteilungsfunktion H zeigt Tabelle 1.

t	$P(X = t)$	$P(X \leq t)$	t	$P(X = t)$	$P(X \leq t)$
4	0,044232	0,060858	10	0,076856	0,954350
5	0,101330	0,162188	11	0,032247	0,986597
6	0,168884	0,331072	12	0,010346	0,996943
7	0,209554	0,540626	13	0,002523	0,999467
8	0,196457	0,737083	14	0,000463	0,999930
9	0,140411	0,877493	15	0,000063	0,999993

Tab. 1: $h(t; 22; 25; 75)$ und $H(t; 22; 25; 75)$ für $4 \leq t \leq 15$

Man sieht, dass $P(X > 13) < 10^{-3}$ wird; diese seltenen Fälle vernachlässigen wir im Folgenden zunächst. Mit dieser Beschränkung nehmen wir bewusst einen Fehler bei den Ergebnissen in Kauf, der allerdings kalkulierbar bleibt (anders als bei einer Simulation, wie sie Henze für den Unterricht empfiehlt, die aber eine wertvolle Ergänzung zur Absicherung der theoretisch gewonnenen Ergebnisse darstellt). Dehnt man die betrachteten Fälle auf $X \leq 15$ aus, bleibt der Fehler sogar unter 10^{-5} und damit unter den üblichen durch Runden entstehenden Ungenauigkeiten.

In einem zweiten Schritt bestimmen wir für die einzelnen Trefferzahlen t die Maximalzahl g_{\max} an Gewinnlinien und dann für $g \leq g_{\max}$ die bedingten Wahrscheinlichkeiten $p_{g,t} = P(L = g | X = t)$.

Mithilfe der Werte aus Tabelle 1 können dann auch die totalen Wahrscheinlichkeiten $P(L = g)$ und somit die Gewinnchancen beim Bingo berechnet werden. Die Genauigkeit der Ergebnisse hängt dabei von den in die Berechnung einbezogenen Werten von t ab.

3 Berechnungen von $p_{g,t}$

Je größer die Trefferzahl t ist, umso mehr Gewinn-Muster auf dem Bingo-Schein sind möglich. Je nach Art und Anzahl der in einem Muster auftretenden Gewinnlinien bezeichnen wir den Typ eines Musters mit einer dreistelligen Zahl der Form dsz , z. B. 102 bei einer Diagonale, keiner Spalte und zwei Zeilen. Bei mehr als zwei Gewinnlinien gibt es noch Untertypen, die wir durch angefügte Buchstaben unterscheiden werden.

Als Anzahl a_{dsz} der Muster eines Typs ergibt sich

$$a_{dsz} = \binom{2}{d} \cdot \binom{5}{s} \cdot \binom{5}{z} \quad (1)$$

3.1 Die Fälle $g_{\max} = 0$ und $g_{\max} = 1$

Für $0 \leq t \leq 4$ gilt $g_{\max} = 0$ und folglich $p_{0,t} = 1$, da für eine Gewinnlinie in einem Muster auf dem Bingo-Schein mindestens 5 Treffer nötig sind.

Bei $5 \leq t \leq 8$ gibt es höchstens eine Gewinnlinie, denn für zwei Gewinnlinien im Muster sind mindestens neun Treffer erforderlich (wenn sich die beiden Linien in einem „Doppelfeld“ schneiden).

Soll eine bestimmte der 12 Linien der Bingo-Karte Gewinnlinie sein, müssen alle 5 Felder dieser Linie unter den t gezogenen Zahlen sein. Die Wahrscheinlichkeit dafür erhält man wieder als Wert einer hypergeometrischen Verteilung, nämlich $h(5; t, 5; 25)$. Hierin stehen die letzten beiden Parameter für die Anzahl der Felder auf dieser Gewinnlinie bzw. für die Gesamtzahl der Felder auf der Bingo-Karte.

Da für $t \leq 8$ die Ereignisse mit einer Gewinnlinie disjunkt sind, ist die Wahrscheinlichkeit, irgendeine Gewinnlinie zu treffen, 12-mal so groß. Für $5 \leq t \leq 8$ gilt demnach $p_{1,t} = P(L = 1 | X = t) = 12 \cdot h(5; t, 5; 25)$ und $p_{0,t} = P(L = 0 | X = t) = 1 - p_{1,t}$. Die mit diesen Formeln berechneten Werte finden sich in Tabelle 2.

t	5	6	7	8
$p_{1,t}$	0,00023	0,00136	0,00474	0,01265
$p_{0,t}$	0,99977	0,99864	0,99526	0,98735

Tab. 2: Bedingte Wahrscheinlichkeiten $p_{g,t}$ für $5 \leq t \leq 8$

3.2 Der Fall $g_{\max} = 2$

Für $9 \leq t \leq 11$ sind zwei Gewinnlinien möglich, aber nicht mehr, denn bei drei Gewinnlinien im Muster gibt es maximal drei Doppelfelder, sodass für $L = 3$ mindestens $15 - 3 = 12$ Treffer erforderlich sind.

Für zwei Gewinnlinien benötigt man bei den Typen 002 und 020 zehn Treffer, bei den übrigen (011, 101, 110, und 200) nur neun, weil ein Doppelfeld auftritt. Nach (1) erhält man als Anzahl der Muster

$$\text{mit zehn Treffern: } a_{002} + a_{020} = 10 + 10 = 20,$$

$$\begin{aligned} \text{mit neun Treffern: } a_{011} + a_{101} + a_{110} + a_{200} \\ = 25 + 10 + 10 + 1 = 46. \end{aligned}$$

Damit ein bestimmtes von den $\binom{12}{2} = 66$ möglichen Linienpaaren getroffen wird, müssen alle 9 bzw. 10 Felder dieser Linien unter den t gezogenen sein. Analog zum Fall $g_{\max} = 1$ ist die Wahrscheinlichkeit dafür $h(9; t, 9; 25)$ bzw. $h(10; t, 10; 25)$. Da für $t \leq 11$ die Ereignisse mit zwei Gewinnlinien disjunkt sind, ist die Wahrscheinlichkeit, irgendein Linienpaar zu treffen, $p_{2,t} = 46 \cdot h(9; t, 9; 25) + 20 \cdot h(10; t, 10; 25)$.

Im Hinblick auf die weiteren Untersuchungen führen wir für den Term von $p_{1,t}$ und für diese Summe noch eine neue Bezeichnung ein:

$$S_{1,t} := 12 \cdot h(5; t, 5; 25) \quad (2.1)$$

$$S_{2,t} := 46 \cdot h(9; t, 9; 25) + 20 \cdot h(10; t, 10; 25) \quad (2.2)$$

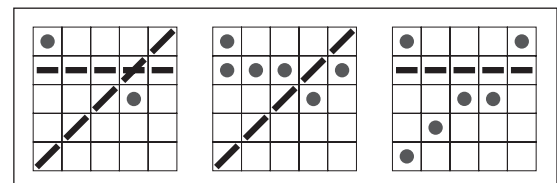


Abb. 1: Jedes Muster mit $L = 2$ (hier: Typ 101, $t = 11$) wird auch zweimal im Term $S_{1,t}$ erfasst

Bei $t > 8$ ist $S_{1,t}$ nicht mehr die Wahrscheinlichkeit $p_{1,t}$ für genau eine Gewinnlinie bei t Treffern. Muster mit

zwei Gewinnlinien wie in Abbildung 1 werden auch in den Termen $S_{1;t}$ beider Einzellinien erfasst (wobei der 3. Parameter 5 für die Felder einer dieser Linien steht). sie tragen aber natürlich nichts zu $p_{1;t}$ bei. Diese Überlegung führt zu der Korrektur

$$p_{1;t} = S_{1;t} - 2 \cdot p_{2;t}$$

Berücksichtigt man noch $p_{0;t} = 1 - (p_{1;t} + p_{2;t})$, so kann man – z. B. mit einer Tabellenkalkulation – alle $p_{g;t}$ für $9 \leq t \leq 11$ leicht berechnen (Tab. 3).

t	9	10	11
$p_{2;t}$	0,000023	0,000231	0,001306
$p_{1;t}$	0,028413	0,056454	0,101736
$p_{0;t}$	0,971564	0,943314	0,896958

Tab. 3: $p_{g;t}$ für $9 \leq t \leq 11$

3.3 Der Fall $g_{\max} = 3$

Für $t = 12$ und $t = 13$ sind drei Gewinnlinien möglich, aber nicht mehr, denn der für vier Linien günstigste Typ 211 mit sechs Doppelfeldern deckt schon 14 der 25 Felder der Bingo-Karte ab.

Wir untersuchen nun, welche Muster-Typen bei drei Gewinnlinien möglich sind und wie viele Felder sie jeweils überdecken.

Abbildung 2 zeigt die Muster-Typen mit mindestens zwei Gewinnzeilen; durch Spiegelung an einer Diagonale entstehen die Typen mit mindestens 2 Gewinnspalten (Typbezeichnung in Klammern). Man kann sich beim Aufsuchen der Typen also auf die Fälle mit $s \leq z$ beschränken. Wenn $s = z$ ist, liefert eine Spiegelung wieder ein Muster desselben Typs.

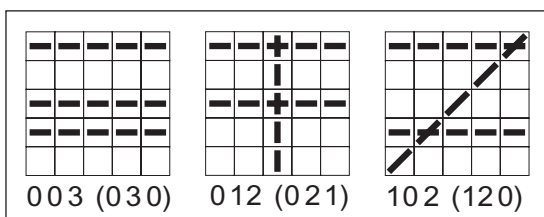


Abb. 2: Beispiele für Muster mit drei Gewinnlinien, die 15 oder 13 Felder abdecken, mit Typ-Bezeichnung

Bei den Typen 111 und 201 (bzw. 210) sind zwei Untertypen zu unterscheiden, je nachdem, ob es in dem Muster Felder gibt, in denen sich alle drei Gewinnlinien schneiden („Tripelfelder“) oder ob es nur Doppelfelder gibt, wo sich je zwei der drei Linien schneiden (Typbezeichnung mit Zusatz T bzw. D). Zu jedem Paar Diagonale/Spalte liefert bei Typ 111 nur eine der fünf Zeilen einen Tripelpunkt; Gleiches gilt für die Diagonalenpaare bei Typ 201.

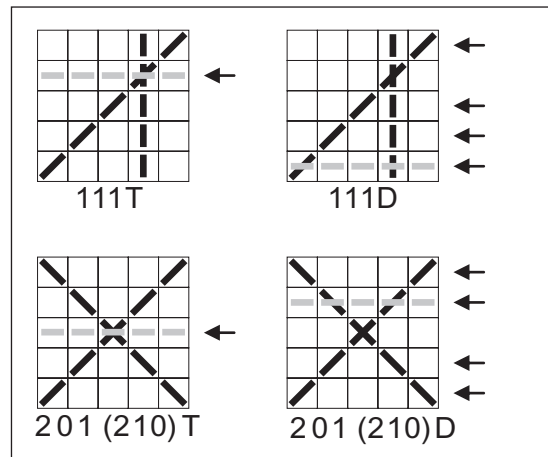


Abb. 3: Muster mit drei Gewinnlinien, die 13 oder 12 Felder abdecken, mit Untertypen

Nach (1) erhält man hier als Anzahl der Muster

$$\text{mit 15 Treffern: } 2 \cdot a_{003} = 20,$$

$$\text{mit 13 Treffern: } 2 \cdot a_{012} + 2 \cdot a_{102} + a_{111T} + 2 \cdot a_{201T} \\ = 100 + 40 + 10 + 2 = 152,$$

$$\text{mit 12 Treffern: } a_{111D} + 2 \cdot a_{201D} = 40 + 8 = 48,$$

zusammen $\binom{12}{3} = 220$. Analog zum Vorgehen im vorigen Abschnitt definiert man nun

$$S_{3;t} := 48 \cdot h(12, t, 12, 25) + 152 \cdot h(13, t, 13, 25) \\ + 20 \cdot h(15, t, 15, 25) \quad (2.3)$$

Man beachte, dass der letzte Summand für $t < 15$ null wird und $p_{3;t} = S_{3;t}$ gilt, solange $12 \leq t \leq 13$ ist. Die Terme $S_{2;t}$ und $S_{1;t}$ bleiben unverändert, erfassen aber nun auch Muster mit drei Gewinnlinien, wie es Abbildung 4 für ein Beispiel mit $t = 13$ zeigt. Man kann $\binom{3}{2} = 3$ Zweier-Kombinationen und $\binom{3}{1} = 3$ Einzellinien aus den drei Gewinnlinien auswählen.

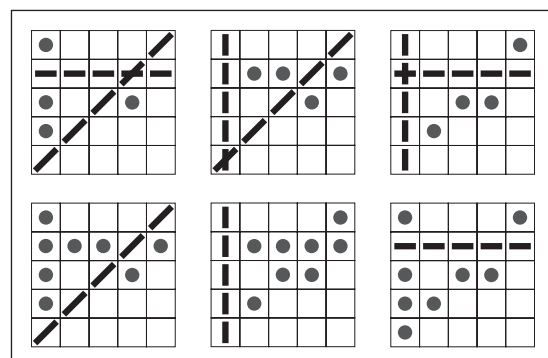


Abb. 4: Jedes Muster mit drei Gewinnlinien (hier: Typ 111D, $t = 13$) wird auch je dreimal in den Termen $S_{2;t}$ und $S_{1;t}$ erfasst

Ausgehend von $p_{3;t} = S_{3;t}$ lassen sich nun mit den Rekursionsformeln

$$p_{2;t} = S_{2;t} - 3 \cdot p_{3;t}, \quad p_{1;t} = S_{1;t} - (2 \cdot p_{2;t} + 3 \cdot p_{3;t}),$$

und $p_{0;t} = 1 - (p_{3;t} + p_{2;t} + p_{1;t})$

alle $p_{g;t}$ für $12 \leq t \leq 13$ berechnen (Tab. 4).

t	$p_{3;t}$	$p_{2;t}$	$p_{1;t}$	$p_{0;t}$
12	$9,23 \cdot 10^{-6}$	0,0053	0,1682	0,8265
13	0,0001	0,0174	0,2554	0,7270

Tab. 4: $p_{g;t}$ für $12 \leq t \leq 13$

4 Eine erste Näherungslösung

Die Ergebnisse der Tabellen 1 bis 4 liefern mit der Formel für die totale Wahrscheinlichkeit

$$P(L = g) \approx \sum_{t=0}^{13} P(X = t) \cdot P(L = g | X = t) \quad (3)$$

Näherungswerte der Wahrscheinlichkeiten $P(L = g)$ für $0 \leq g \leq 3$. Die Ergebnisse zeigt Tabelle 5.

g	Werte		Fehler	
	Näherung	exakt	absolut	relativ
3	$4,72 \cdot 10^{-7}$	$1,72 \cdot 10^{-6}$	$1,25 \cdot 10^{-6}$	72,5 %
2	0,000162	0,000193	$3,04 \cdot 10^{-5}$	15,8 %
1	0,017724	0,017919	$1,95 \cdot 10^{-4}$	1,09 %
0	0,981580	0,981887	$3,07 \cdot 10^{-4}$	0,03 %

Tab. 5: Näherungswerte für $P(L = g)$ mit Fehlerangabe

Die Vernachlässigung der $t > 13$ führt zwar zu Abweichungen von den exakten Werten (Henze 2012), die absoluten Fehler bleiben – wie schon in Kapitel 2 gezeigt – in der Summe aber kleiner als 10^{-3} . Der relative Fehler ist allerdings nur für $g = 0$ und $g = 1$ klein. In einer leistungsstarken Lerngruppe wird man daher auch die Fälle $t = 14$ und $t = 15$ in die Untersuchung einbeziehen; diesen weiteren Weg skizzieren wir im folgenden Kapitel 5. Man kann aber auch die dort erzielten Ergebnisse (Tab. 8) im Unterricht ohne Herleitung nur mitteilen,

5 Verbesserung der Lösung

Wir beziehen nun auch die Trefferzahlen 14 und 15 ein, bei denen $g_{\max} = 4$ ist. Die einzelnen Abschnitte dieses Kapitels zeigen noch einmal die Abfolge der schon in 3.1 bis 3.3 durchgeführten Lösungsschritte. Dabei wird auch deutlich, dass die kombinatorischen Überlegungen mit wachsendem t zunehmend komplexer werden.

5.1 Anzahl der Grundtypen

Bei der Zählung nennen wir nur die Typen mit $s \leq z$ und verdoppeln deren nach (1) berechnete Anzahl, wenn $s \neq z$ ist; dabei ordnen wir nach der Anzahl d der im Muster auftretenden Diagonalen (Tab. 6):

d	Typen	Anzahl
0	004, 013, 022	$2 \cdot 5 + 2 \cdot 50 + 100 = 210$
1	103, 112	$2 \cdot 20 + 2 \cdot 100 = 240$
2	202, 211	$2 \cdot 10 + 25 = 45$

Tab. 6: Grundtypen mit $L = 4$ und Anzahl der Muster

Insgesamt gibt es 495 Muster mit 4 Gewinnlinien; dies sind genau die $\binom{12}{4}$ Möglichkeiten, aus den 12 Linien der Bingo-Karte vier auszuwählen.

5.2 Untertypen mit $L = 4$ und $t < 16$

Die meisten der Typen in Tabelle 6 sind mit 14 oder 15 Treffern auf der Bingo-Karte nicht möglich. Dies gilt zunächst für alle Typen mit $d = 0$: 004 und 013 entstehen aus 003 (Abb. 2 links) durch Hinzufügen einer neuen Zeile oder Spalte, erfordern also 20 oder 17 Felder, und für 022 müssen zu den 13 Feldern von 012 (Abb. 2 Mitte) in einer Spalte noch drei Felder getroffen werden, zusammen also 16.

Bei den Typen mit einer Diagonale entsteht 103 aus 102 (Abb. 2 rechts) durch eine weitere Zeile (womit 17 Felder überdeckt sind) und 112 durch eine neue Spalte. Wählt man dafür eine der beiden Spalten mit einem Doppelfeld, so entsteht ein Muster vom Untertyp 112T, das 16 Felder überdeckt.

Ergänzt man 102 dagegen um eine Spalte, die noch kein Doppelfeld enthält, so entsteht erstmals ein 15-Felder-Muster. Jedes der $2 \cdot 20$ Muster vom Typ 102 (bzw. 120) kann man auf drei Arten zu einem von 120 Mustern des Typs 112D ausbauen (Abb. 5).

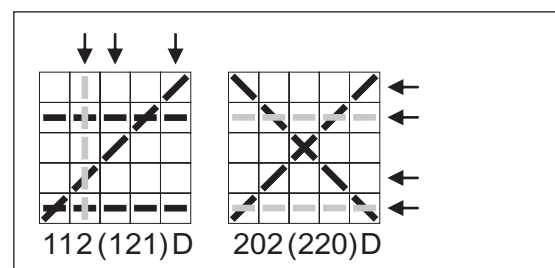


Abb. 5: Muster mit $L = 4$ (5 Doppelfelder)

Bei den Typen mit zwei Diagonalen entsteht 202 aus 201 (Abb. 3) durch Hinzufügen einer Zeile. Das erfordert im Fall 201T drei neue Felder, insgesamt also 16. Dagegen genügen 15 Felder, wenn man zu den

beiden Diagonalen zwei der bei 201D markierten Zeilen ergänzt. Das geht auf 6 Arten; es gibt folglich 12 Muster vom Typ 202D bzw. 220D.

Um die Untertypen des anderen Grundtyps mit zwei Diagonalen (211) zu bestimmen, beachten wir die Lage des Schnittfeldes Z/S von Zeile und Spalte, wobei vier Fälle zu unterscheiden sind:

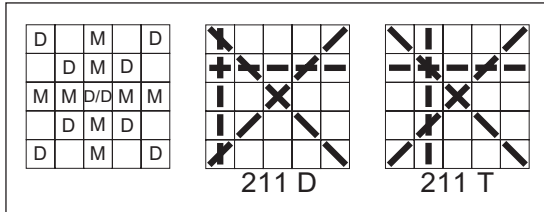


Abb. 6: Beispiele für Muster mit $L = 4$ vom Grundtyp 211

(1) Ist Z/S ein Feld, das weder auf einer Diagonale noch auf einer der Mittellinien liegt (8 weiße Felder in Abb. 6 links), so kommen zu den beiden Doppelfeldern (D/D und Z/S) noch je zwei weitere der Art D/Z und D/S. Man hat also 6 Doppelfelder, sodass jedes der 8 Muster 14 Felder überdeckt. (Typ 211D, Abb. 6 Mitte für das Beispiel 2. Zeile, 1. Spalte).

(2) Ist Z/S eines der 8 Felder mit „M“ (auf einer der Mittellinien), so wird D/D ein Tripelfeld. Ist die waagerechte Mittellinie Gewinnlinie, so schneidet sie die Gewinnspalte außer in Z/S in zwei Doppelfeldern der Art D/S; analog erhält man ein Tripelfeld und drei Doppelfelder, wenn die senkrechte Mittellinie Gewinnlinie ist. Dies sind die ersten 8 Muster vom Typ 211T, die je 15 Felder überdecken.

(3) Weitere 8 Muster vom Typ 211T erhält man, wenn Z/S auf einer Diagonale liegt, aber nicht auf deren Schnittfeld D/D (Felder mit „D“ in Abb. 6 links). Dann entsteht dort ein Tripelfeld D/Z/S, während die beiden anderen Schnittfelder mit den Diagonalen Doppelfelder (D/Z bzw. D/S) werden. Auch hier gibt es also – wie im Fall (2) – zusammen mit dem Mittelfeld D/D des Musters drei Doppelfelder. Ein Beispiel zeigt Abbildung 6 rechts.

(4) Fällt Z/S mit D/D zusammen, so entsteht erstmals ein Vierfachfeld, und das Muster überdeckt 17 Felder (Typ 211V).

Insgesamt gibt es für $L = 4$ damit 148 Muster, die 15 Felder überdecken, und 8 Muster mit 14 Feldern. Hieraus ergibt sich für $t \leq 15$ die Wahrscheinlichkeit

$$p_{4;t} = 8 \cdot h(14, t, 14, 25) + 148 \cdot h(15, t, 15, 25)$$

Ferner erhält man aus den Untersuchungen dieses Abschnitts noch den Wert der Rechengröße

$$\begin{aligned} S_{4;t} &= 8 \cdot h(14, t, 14, 25) + 148 \cdot h(15, t, 15, 25) \\ &+ 188 \cdot h(16, t, 16, 25) + 141 \cdot h(17, t, 17, 25) \\ &+ 10 \cdot h(20, t, 20, 25), \end{aligned}$$

der für $t < 16$ mit $p_{4;t}$ übereinstimmt und in Berechnungen mit $t \geq 16$ benötigt wird. Für $t = 16$ ist nämlich $g_{\max} = 5$; z. B. kann man das Muster 211D in Abbildung 6 durch die leeren Felder der letzten Spalte zu einem Muster mit 5 Gewinnlinien vom Typ 221D erweitern, das 16 Felder überdeckt.

5.3 Die Werte von $p_{g;t}$ für $t = 14$ und $t = 15$

Während wir die Terme für $S_{g;t}$ mit umfangreichen kombinatorischen Überlegungen bestimmt haben, sind alle in Kapitel 3 für $p_{g;t}$ hergeleiteten Terme von derselben allgemeinen Form, nämlich

$$p_{g;t} = S_{g;t} - \sum_{i=g+1}^{g_{\max}} \binom{i}{g} \cdot p_{i;t} \quad (4).$$

Für $t \leq 15$ ist $g_{\max} = 4$; für $g = 4$ ist die Summe leer, so dass sich wie oben $p_{4;t} = S_{4;t}$ ergibt. Für $g < 4$ gilt

$$\begin{aligned} p_{3;t} &= S_{3;t} - 4 \cdot p_{4;t}; \quad p_{2;t} = S_{2;t} - (3 \cdot p_{3;t} + 6 \cdot p_{4;t}); \\ p_{1;t} &= S_{1;t} - (2 \cdot p_{2;t} + 3 \cdot p_{3;t} + 4 \cdot p_{4;t}); \\ p_{0;t} &= 1 - (p_{1;t} + p_{2;t} + p_{3;t} + p_{4;t}), \end{aligned}$$

wobei im letzten Fall zu beachten ist, dass (4) mit der Setzung $S_{0;t} := 1$ und $\binom{i}{0} = 1$ auch für $g = 0$ gilt. Die so berechneten Werte zeigt Tabelle 7.

t	$p_{4;t}$	$p_{3;t}$	$p_{2;t}$	$p_{1;t}$	$p_{0;t}$
14	$2 \cdot 10^{-6}$	0,0012	0,0475	0,3535	0,5978
15	$7 \cdot 10^{-5}$	0,0070	0,1097	0,4380	0,4453

Tab. 7: $p_{g;t}$ für $14 \leq t \leq 15$

5.4 Die verbesserte Näherungslösung

Mit den Ergebnissen aus Tabelle 7 kann man nun genauere Näherungswerte für $P(L = g)$ gewinnen, indem man die Summe in (3) bis $t = 15$ erstreckt:

$$P(L = g) \approx \sum_{t=0}^{15} P(X = t) \cdot P(L = g | X = t) \quad (3^*).$$

g	Werte		Fehler	
	Näherung	exakt	absolut	relativ
4	$5,40 \cdot 10^{-9}$	$1,65 \cdot 10^{-8}$	$1,11 \cdot 10^{-8}$	67,22 %
3	$1,49 \cdot 10^{-6}$	$1,72 \cdot 10^{-6}$	$2,30 \cdot 10^{-7}$	13,36 %
2	0,000191	0,000193	$1,50 \cdot 10^{-6}$	0,7795 %
1	0,017916	0,017919	$3,13 \cdot 10^{-6}$	0,0174 %
0	0,981885	0,981887	$1,90 \cdot 10^{-6}$	0,0002 %

Tab. 8: Näherungswerte für $P(L = g)$ mit Fehlerangabe

Wie ein Vergleich der Tabellen 5 und 8 zeigt, sind die relativen Fehler bei den Werten für $L = 2$ und $L = 3$ deutlich kleiner geworden. Fasst man nach den Bingo-Regeln Ergebnisse mit mehr als zwei Gewinnlinien zum Ereignis *Dreifach-Bingo* zusammen, so wird $P(L \geq 3) \approx 1,4949 \cdot 10^{-6}$, und dieser Wert weicht nur noch um $2,407 \cdot 10^{-7}$ bzw. 13,9 % von dem exakten ab.

6 Schlussbemerkungen

Hauptanliegen dieses Aufsatzes war es aufzuzeigen, wie man auch ohne allzu abstrakte Hilfsmittel, die im Schulunterricht nicht zur Verfügung stehen, eine Vorstellung von den Gewinnchancen beim Bingo gewinnen kann. Durch die in Kapitel 2 vorgenommene Zerlegung des Problems kann man dabei stets Fehlerschranken für die Näherungswerte angeben. Trotz der Ungenauigkeiten sind diese Werte wesentlich besser als Schätzungen, die man anhand von Simulationen gewonnen hat. Eine Simulation ist aber durchaus sinnvoll, um die theoretischen Ergebnisse auf Plausibilität zu überprüfen. Auch hierbei erweist sich die Zerlegung als hilfreich; denn statt das ganze Bingo-Spiel zu simulieren, bietet es sich an, für bestimmte feste Trefferzahlen der Ziehung „22 aus 75“ die Anzahl der Gewinnlinien durch Simulation zu bestimmen. Für $X = 15$ ergab eine Serie von 100 000

Versuchen die in Tabelle 9 wiedergegebenen Ergebnisse.

g	0	1	2	3	4
Erwartungswert	44532	43795	10968	699	7
Besetzungszahl	44743	43574	10930	744	9

Tab. 9: Bei 100 000 Simulationen erwartete und aufgetretene Besetzungszahlen von $L = g$ mit $X = 15$

Die Besetzungszahlen $n(g)$ sind jeweils binomialverteilt mit $N = 100\,000$ und $p = p_{g,15}$. Die beobachteten Werte weichen von ihren Erwartungswerten jeweils um weniger als 2σ ab.

Die Excel-Tabellenblätter und das Programm zur Simulation können beim Autor angefordert werden.

Literatur

Henze, N. (2012): Die Verteilung der Anzahl der Gewinnlinien beim Bingo. In: *Stochastik in der Schule*, 33 (2), S. 2–8.

Anschrift des Verfassers

Gerd Riehl
 Obere Mark 6
 30890 Barsinghausen
 Gerd.Riehl@t-online.de