Problem des Monats Januar 2019

Vollständige Serie beim Zahlenlotto

In jeder Woche finden beim Lottospiel *6 aus 49* zwei Ziehungsveranstaltungen statt.

Laien wundern sich immer wieder darüber, wie oft es vorkommt, dass eine Zahl, die in der letzten Veranstaltung gezogen wurde, erneut zu den Gewinnzahlen der aktuellen Ziehung gehört. Man kann leicht nachrechnen, dass die Wahrscheinlichkeit für ein solches Ereignis

Mindestens zwei Gewinnzahlen aus aufeinanderfolgenden Ziehungsveranstaltungen stimmen überein

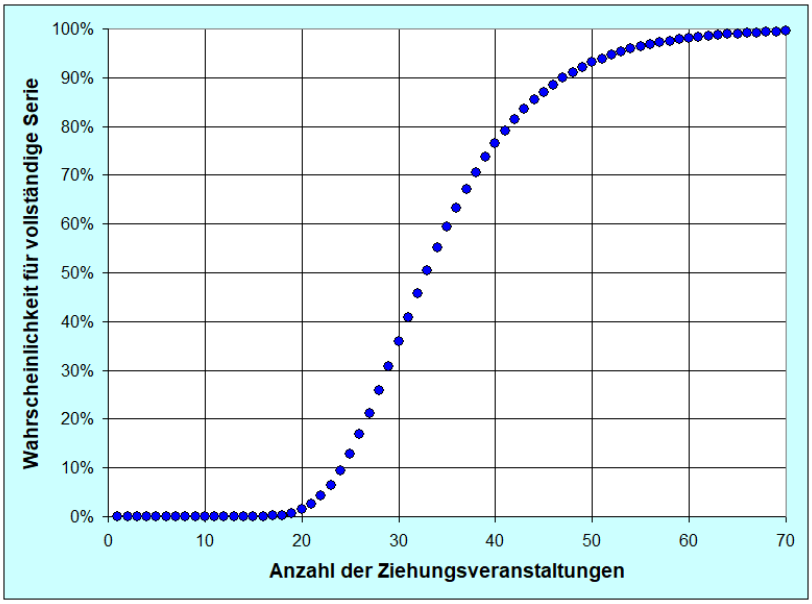
immerhin ca. 43,6 % beträgt.

- Erläutern Sie, wie man diese Wahrscheinlichkeit bestimmen kann.

Ein ganz anderes Problem und erheblich aufwendiger zu bestimmen, ist die Berechnung der Wahrscheinlichkeit, dass in *n* aufeinanderfolgenden Ziehungsveranstaltungen *k verschiedene* Gewinnzahlen gezogen werden.

Die folgende Grafik zeigt, wie die Wahrscheinlichkeit für das Vorliegen einer vollständigen Serie beim Lotto, also von *k* = 49 verschiedenen gleichwahrscheinlichen Gewinnzahlen, mit zunehmender Anzahl der Ziehungsveranstaltungen wächst.

Man kann beispielsweise ablesen, dass erst nach 33 Ziehungsveranstaltungen die Wahrscheinlichkeit dafür, dass jede der 49 Lottokugeln *mindestens einmal* gezogen wurde, etwas mehr als 50 % beträgt.



Die Berechnung der Wahrscheinlichkeiten kann nur iterativ erfolgen.

- Welche Rekursionsformeln werden für die Berechnung benötigt?

Lösung

Wenn bei der nächsten Ziehungsveranstaltung sechs Gewinnzahlen gezogen werden, von denen keine mit denen der letzten Ziehung übereinstimmt, dann bedeutet dies:

Für die Auswahl kommen nur 43 der 49 Zahlen infrage, d. h., die Wahrscheinlichkeit für das Ereignis beträgt

$$\binom{43}{6}\Big/\binom{49}{6}=\frac{6.096.454}{13.983.816}\approx 0{,}436\,.$$

Zur Berechnung der Wahrscheinlichkeiten, die der Grafik zugrundeliegen, kann man wie folgt überlegen:

Nach der ersten Ziehungsveranstaltung sind sechs verschiedene Zahlen gezogen: das System befindet sich dann im Zustand k = 6, also $P(1\,;\,6) = 1$.

Wenn bei der zweiten Ziehungsveranstaltung diese sechs Gewinnzahlen der ersten Ziehung wieder gezogen werden, befindet sich das System immer noch im Zustand k = 6: die Wahrscheinlichkeit hierfür ist

$$P(2\,;6)=P(1;6)\cdot\binom{6}{6}\cdot\binom{43}{0}\Big/\binom{49}{6},$$

Wenn bei der zweiten Ziehungsveranstaltung nur *eine* neue Gewinnzahl hinzukommt, ist das System im Zustand k = 7; dann gilt: $P(2\,;7)=P(1;6)\cdot\binom{6}{5}\cdot\binom{43}{1}\Big/\binom{49}{6}$.

In diesen Zustand kann das System aber auch gelangen, wenn es eine Stufe vorher bereits im Zustand k = 7 war und dann keine Gewinnzahl hinzukommt. Für $n \geq 3$ muss daher die Rekursionsformel erweitert werden:

$$P(n\,;7)=P(n-1;6)\cdot\binom{6}{5}\cdot\binom{43}{1}\Big/\binom{49}{6}+P(n-1;7)\cdot\binom{7}{6}\cdot\binom{42}{0}\Big/\binom{49}{6}.$$

Mit entsprechenden Überlegungen erhält man schrittweise

$$P(n\,;8)=P(n-1;6)\cdot\binom{6}{4}\cdot\binom{43}{2}\Big/\binom{49}{6}+P(n-1;7)\cdot\binom{7}{5}\cdot\binom{42}{1}\Big/\binom{49}{6}+P(n-1;8)\cdot\binom{8}{6}\cdot\binom{41}{0}\Big/\binom{49}{6}$$

$$P(n\,;9)=P(n-1;6)\cdot\binom{6}{3}\cdot\binom{43}{3}\Big/\binom{49}{6}+P(n-1;7)\cdot\binom{7}{4}\cdot\binom{42}{2}\Big/\binom{49}{6}$$
$$+P(n-1;8)\cdot\binom{8}{5}\cdot\binom{41}{1}\Big/\binom{49}{6}+P(n-1;9)\cdot\binom{9}{6}\cdot\binom{40}{0}\Big/\binom{49}{6}$$

$$P(n\,;10)=P(n-1;6)\cdot\binom{6}{2}\cdot\binom{43}{4}\Big/\binom{49}{6}+P(n-1;7)\cdot\binom{7}{3}\cdot\binom{42}{3}\Big/\binom{49}{6}$$
$$+P(n-1;8)\cdot\binom{8}{4}\cdot\binom{41}{2}\Big/\binom{49}{6}+P(n-1;9)\cdot\binom{9}{5}\cdot\binom{40}{1}\Big/\binom{49}{6}+P(n-1;10)\cdot\binom{10}{6}\cdot\binom{39}{0}\Big/\binom{49}{6}$$

$$P(n\,;11)=P(n-1;6)\cdot\binom{6}{1}\cdot\binom{43}{5}\Big/\binom{49}{6}+P(n-1;7)\cdot\binom{7}{2}\cdot\binom{42}{4}\Big/\binom{49}{6}+P(n-1;8)\cdot\binom{8}{3}\cdot\binom{41}{3}\Big/\binom{49}{6}$$
$$+P(n-1;9)\cdot\binom{9}{4}\cdot\binom{40}{2}\Big/\binom{49}{6}+P(n-1;10)\cdot\binom{10}{5}\cdot\binom{39}{1}\Big/\binom{49}{6}+P(n-1;11)\cdot\binom{11}{6}\cdot\binom{38}{0}\Big/\binom{49}{6}$$

$$P(n\,;12)=P(n-1;6)\cdot\binom{6}{0}\cdot\binom{43}{6}\Big/\binom{49}{6}+P(n-1;7)\cdot\binom{7}{1}\cdot\binom{42}{5}\Big/\binom{49}{6}+P(n-1;8)\cdot\binom{8}{2}\cdot\binom{41}{4}\Big/\binom{49}{6}$$
$$+P(n-1;9)\cdot\binom{9}{3}\cdot\binom{40}{3}\Big/\binom{49}{6}+P(n-1;10)\cdot\binom{10}{4}\cdot\binom{39}{2}\Big/\binom{49}{6}+P(n-1;11)\cdot\binom{11}{5}\cdot\binom{38}{1}\Big/\binom{49}{6}$$
$$+P(n-1;12)\cdot\binom{12}{6}\cdot\binom{37}{0}\Big/\binom{49}{6}$$

An der letzten Rekursionsformel wird deutlich, dass man, wenn $k \geq 12$ ist, jeweils die letzten sechs Stufen berücksichtigen muss, denn im Prinzip können bei der nächsten Ziehung bis zu sechs neue Lottozahlen hinzukommen. Für n, $k \geq 12$ gilt also:

$$P(n\,;k)=P(n-1;k-6)\cdot\binom{k-6}{0}\cdot\binom{55-k}{6}\Big/\binom{49}{6}+P(n-1;k-5)\cdot\binom{k-5}{1}\cdot\binom{54-k}{5}\Big/\binom{49}{6}$$
$$+P(n-1;k-4)\cdot\binom{k-4}{2}\cdot\binom{53-k}{4}\Big/\binom{49}{6}+P(n-1;k-3)\cdot\binom{k-3}{3}\cdot\binom{52-k}{3}\Big/\binom{49}{6}$$
$$+P(n-1;k-2)\cdot\binom{k-2}{4}\cdot\binom{51-k}{2}\Big/\binom{49}{6}+P(n-1;k-1)\cdot\binom{k-1}{5}\cdot\binom{50-k}{1}\Big/\binom{49}{6}$$
$$+P(n-1;k)\cdot\binom{k}{6}\cdot\binom{49-k}{0}\Big/\binom{49}{6}$$

Diese Rekursionsformeln gelten auch für kleinere Werte von n und k, wenn man die entsprechenden Wahrscheinlichkeiten gleich null setzt:

$P(n-1\,;k-6) = P(n-1\,;k-5) = P(n-1\,;k-4) = P(n-1\,;k-3)$

$= P(n-1\,;k-2) = P(n-1\,;k-1) = 0$

Betrachtet man konkret die Lottoziehungen seit der Einführung des Lottospiels im Jahr 1955, dann findet man folgende empirischen Daten:

Erst nach der 49sten Ziehungsveranstaltung war tatsächlich jede der Gewinnzahlen mindestens einmal gezogen worden, dann (nach Neuzählung) erst wieder bei der 90sten Ziehung, und weiter bei der 117ten, 146sten, 194sten, 227sten, 255sten, 283sten, 309ten, 348sten Ziehungsveranstaltung, nacheinander sind dies Wartezeiten von 49, 41, 27, 29, 48, 33, 28, 28, 26, 39 Stufen (im Mittel 34,8 Stufen).

Übrigens: Man kann ausrechnen, dass der Erwartungswert der Verteilung etwas größer ist als 38. Dass Erwartungswert und Median (Wahrscheinlichkeit 50 %) nicht übereinstimmen, liegt daran, dass die Verteilung hier nicht symmetrisch ist.