Ein Geburtstagsproblem

von **Georg Schrage**, Dortmund

Zusammenfassung: Am Institut für Didaktik der Mathematik der Universität Dortmund gibt es alle Jahre ein außergewöhnliches Ereignis zu feiern: Drei von insgesamt fünfzehn Institutsmitgliedern haben am 28. Juli Geburtstag.

Die Frage, mit welcher Wahrscheinlichkeit in einer Gruppe von n Personen zwei (oder mehr) am gleichen Tag Geburtstag haben, ist ein Standardproblem für den Stochastikunterricht. Das Zusammentreffen dreier Geburtstage legt es nahe zu untersuchen, mit welcher Wahrscheinlichkeit drei von fünfzehn (allgemein: drei von n) Personen am gleichen Tag Geburtstag feiern.

Das Geburtstagsproblem in seiner üblichen Form löst man, indem man die Wahrscheinlichkeit für das komplementäre Ereignis - nämlich, daß alle n Personen an verschiedenen Tagen Geburtstag feiern - berechnet. Es gilt

$$\begin{aligned} & W(\text{mind. zwei Personen haben einen gemeinsamen Geburtstag}) \\ &= 1 - W(\text{alle n Personen haben verschiedene Geburtstage}) \\ &= 1 - \frac{(365 \cdot 364 \cdot \ldots \cdot (366-n))}{365^n}. \end{aligned}$$

Für unsere Variante des Problems machen wir einen analogen Ansatz. $W(n, z)$ bezeichne die Wahrscheinlichkeit dafür, daß es bei einer Gruppe von n Personen genau z Tage gibt, an denen jeweils zwei (und nicht mehr) Angehörige der Gruppe Geburtstag haben, während die übrigen $n - 2 \cdot z$ Personen alle verschiedene Geburtstage haben. Für die Wahrscheinlichkeit Wk, daß drei der n Personen einen gemeinsamen Geburtstag haben, gilt damit

$$Wk = 1 - [W(n,0) + W(n,1) + \ldots + W(n, [\frac{n}{2}])]. \qquad (1)$$

Eine direkte Lösung

Die Wahrscheinlichkeit, daß keine zwei Personen gemeinsam Geburtstag feiern, ist

$$W(n,0) = \frac{365 \cdot 364 \cdot \ldots \cdot (366-n)}{365^n}.$$

(Es gibt $365 \cdot 364 \cdot \ldots \cdot (366-n)$ Möglichkeiten, die Geburtstage von n Personen so auf die 365 Tage des Jahres zu verteilen, daß keine zwei Geburtstage auf einen Tag fallen. Die Anzahl aller möglichen Verteilungen von n Geburtstagen beträgt 365^n. Es wird unterstellt, daß jeder dieser Verteilungen die gleiche Wahrscheinlichkeit zukommt.)

$W(n, 1)$ erhält man wie folgt: Es gibt $\binom{n}{2} \cdot 365$ Möglichkeiten, zwei der n Personen auszuwählen und ihren gemeinsamen Geburtstag festzulegen. Die übrigen $n - 2$ Geburtstage können dann noch auf

$$364 \cdot 363 \cdot \ldots \cdot (365 - (n-2))$$

Arten so verteilt werden, daß alle auf verschiedene Tage fallen. Daraus ergibt sich

$$W(n,1) = \frac{\binom{n}{2} \cdot 365 \cdot 364 \cdot \ldots \cdot (367-n)}{365^n}.$$

Die Wahrscheinlichkeit dafür, daß es genau zwei Tage gibt, an denen zwei Personen ihren Geburtstag feiern, ergibt sich entsprechend zu

$$W(n,2) = \frac{1}{2!} \frac{\binom{n}{2}\binom{n-2}{2} 365 \cdot 364 \cdot 363 \cdot \ldots \cdot (368-n)}{365^n}$$

(Hier ist darauf zu achten, daß die Reihenfolge, in der die beiden Zweiergruppen, deren Mitglieder jeweils einen gemeinsamen Geburtstag haben, ausgewählt werden, nicht relevant ist.). Entsprechend geht es weiter. Für $n = 15$ erhält man

$$
\begin{aligned}
Wk &= 1 - \frac{1}{365^{15}}[365 \cdot 364 \cdot \ldots \cdot 351 + \binom{15}{2} \cdot 365 \cdot 364 \cdot \ldots \cdot 352 \\
&\quad + \frac{1}{2!}\binom{15}{2}\binom{13}{2} \cdot 365 \cdot 364 \cdot \ldots \cdot 353 \\
&\quad \vdots \\
&\quad + \frac{1}{7!}\binom{15}{2}\binom{13}{2} \cdot \ldots \cdot \binom{3}{2} \cdot 365 \cdot 364 \cdot \ldots \cdot 358 \\
&= 1 - (((((((\frac{\binom{3}{2}}{7 \cdot 357} + 1) \cdot \frac{\binom{5}{2}}{6 \cdot 356} + 1) \cdot \frac{\binom{7}{2}}{5 \cdot 355} + 1) \\
&\quad \cdot \frac{\binom{9}{2}}{4 \cdot 354} + 1) \cdot \frac{\binom{11}{2}}{3 \cdot 353} + 1) \cdot \frac{\binom{13}{2}}{2 \cdot 352} + 1) \\
&\quad \frac{\binom{15}{2}}{351} + 1) \cdot \frac{(365 \cdot 364 \cdot \ldots \cdot 351)}{365^{15}}) \\
&= 1 - (((((((\frac{3}{7 \cdot 357} + 1) \cdot \frac{10}{6 \cdot 356} + 1) \cdot \frac{21}{5 \cdot 355} + 1) \\
&\quad \cdot \frac{36}{4 \cdot 354} + 1) \cdot \frac{55}{3 \cdot 353} + 1) \cdot \frac{78}{2 \cdot 352} + 1) \cdot \frac{105}{351} + 1) \\
&\quad \cdot \frac{351 \cdot 352 \cdot \ldots \cdot 364}{365^{14}} \\
&= 0.0033 \quad .
\end{aligned}
$$

Der Ausdruck sieht recht kompliziert aus, läßt sich aber in der obigen Form mit einem Taschenrechner mit mäßigem Aufwand berechnen.

Eine rekursive Lösung

Eine elegante, computergerechte Lösung für die Gleichung (1) findet man durch eine rekursive Beschreibung der Wahrscheinlichkeiten $W(n, z)$. Mit $W(n, z)$ haben wir die Wahrscheinlichkeit dafür bezeichnet, daß an z Tagen jeweils zwei von n Personen ihren Geburtstag feiern, während alle übrigen $e = n - 2 \cdot z$ Personen ihre Geburtstage einzeln feiern. Für die folgenden Überlegungen ist es hilfreich, bei der Notation dieser Wahrscheinlichkeiten den Parameter e mit anzugeben. Wir ersetzen also die Bezeichnung $W(n, z)$ durch $W(n, z, e)$, wobei gilt $n = e + 2 \cdot z$.

Für $W(n, z, e)$ gelten offenbar die Randbedingungen

1. $W(1, z, e) = W(1, 0, 1) = 1$,
2. $W(n, 0, e) = W(n, 0, n) = W(n-1, 0, n-1) \cdot \frac{366-e}{365}$,
3. $W(n, z, 0) = \frac{W(n-1, z-1, 1)}{365}$.

   Im übrigen gilt die rekursive Beziehung

4. $W(n, z, e) = W(n-1, z, e-1) \cdot \frac{366-e-z}{365} + W(n-1, z-1, e+1) \cdot \frac{e+1}{365}$.

Das Verfahren zur Berechnung von $W(n, z, e)$ gemäß dieser Rekursion ist offenbar endlich, da in jedem Rekursionsschritt der Parameter n um eins erniedrigt wird. Die Rekursion endet somit stets bei der bekannten Wahrscheinlichkeit $W(1, 0, 1) = 1$.

Zur Erläuterung der Rekursion: Werden n Personen der Reihe nach ihre Geburtstage zufällig zugeteilt, so gibt es zwei Möglichkeiten zu einer Geburtstagsverteilung vom Typ (n, z, e) mit $z > 0$ und $e > 0$ zu kommen: Entweder die Verteilung der ersten $n - 1$ Geburtstage ist vom Typ $(n-1, z, e-1)$, und der n-te Geburtstag fällt auf einen der $366 - e - z$ noch freien Tage, oder die Verteilung der ersten $n - 1$ Geburtstage ist vom Typ $(n - 1, z - 1, e + 1)$, und der n-te Geburtstag fällt auf einen der $e + 1$ mit genau einem Geburtstag belegten Tage.

Die Wahrscheinlichkeit für den ersten Fall ist $W(n-1, z, e-1)\frac{366-e-z}{365}$, die für den zweiten Fall ist $W(n-1, z-1, e+1)\frac{e+1}{365}$. Daraus folgt die Rekursionsgleichung (4).

Legen wir noch fest, daß gelten soll $W(n, z, e) = 0$ falls $z < 0$ oder $e < 0$, so ergeben sich die Beziehungen (2) und (3) als Spezialfälle von (4). Damit können wir zusammenfassen:

$$
W(n, z, e) = \begin{cases} 0 & \text{falls } z < 0 \text{ oder } e < 0, \\ 1 & \text{falls } n = 1, z = 0 \text{ und } e = 1, \\ W(n-1, z, e-1)(366 - z - e) & \\ +W(n-1, z-1, e+1)(e+1)/365 & \textit{sonst.} \end{cases} \tag{2}
$$

Wer etwas Routine im Umgang mit rekursiv definierten Funktionen hat, wird diese Darstellung vielleicht als leichter verständlich empfinden, als die oben hergeleiteten expliziten Formeln zur Berechnung von $W(n, z)$.

Die rekursive Form (2) läßt sich in jede Programmiersprache, die die Definition rekursiver Funktionen gestattet, übernehmen. In TRUE BASIC, einer modernen Version der Programmiersprache BASIC, die neben dem Komfort einer Interpretersprache auch volle Unterstützung zur Strukturierung von Programmen bietet, lautet der Algorithmus:

Function $Wk(n, z, e)$!Rekursive Def. der Funktion Wk
  IF $z < 0$ or $e < 0$ THEN
    LET Wk=0
  ELSE
    IF n=1 THEN
      LET Wk=1
    ELSE
      LET Wk=(Wk(n-1,z,e-1)(366-z-e)+Wk(n-1,z-1,e+1)$\frac{e+1}{365}$
    END IF
  END IF
END DEF

INPUT n !Hauptprogramm
FOR $z = 0$ TO INT(n/2)
  LET $W = Wk(n, z, n - 2 \cdot z)$
  LET $S = S + W$
NEXT z
PRINT 1-S
END

Rekursionen dieser Art sind sehr speicherplatz- und rechenaufwendig. Diese Schwierigkeit läßt sich dadurch umgehen, daß man Tabellen für die Wahrscheinlichkeiten $W(n, z, e)$ anlegt. Gemäß (2) braucht man dabei jeweils nur auf zwei zuvor berechnete Werte der Tabelle zuzugreifen. Man kann hierfür auch ein Programm zur Tabellenkalkulation einsetzen.

Die folgende Tabelle zeigt einige so berechnete Wahrscheinlichkeiten für das Zusammentreffen dreier Geburtstage.

n	20	40	60	80	100	150	200
Wk	0.00824	0.0669	0.207	0.418	0.646	0.965	0.9995

Simulation

Wo ein Computer verfügbar ist, sollte eine Behandlung des Problems durch Simulation nicht fehlen. Hier das einfache Programm:

```
DIM Tage(365)                       !Feld mit einer Komponente
RANDOMIZE                           !für jeden Tag des Jahres
INPUT n,s                           !n=Anzahl der Personen,
                                    !s=Anzahl der Simulationen
FOR k=1 TO s
  FOR i=1 TO 365
    LET TAGE(i)=0                   !Initialisieren des Feldes
  NEXT i
  LET M=0                           !Max. Anzahl von Geb.tagen
                                    !an einem Tag
  FOR i=1 TO n
    LET G_tag=INT(365xRND)+1        !Auswahl eines Geb.tags
    LET TAGE(G_tag)=Tage(G_tag)+1
    LET M=MAX(M,Tage(G_tag))
  NEXT i
  PRINT M                           !Max. Anzahl von Geb.tagen
                                    !an einem Tag
NEXT k                              !Nächste Simulation
END
```

Zur Erläuterung des Programms:

Durch DIM Tage(365) wird ein Feld mit den Komponenten 1 bis 365 erzeugt. In diesem Feld werden die Anzahlen der Geburtstage gespeichert, die bei der Simulation auf die Tage 1 bis 365 des Jahres fallen. Die Anzahl n der Personen, deren Geburtstage simuliert werden und die Anzahl s der Simulationen von jeweils n Geburtstagen, werden vom Benutzer angegeben. In der Variablen M, die zu Beginn jeder Simulation auf 0 gesetzt wird, wird die maximale Anzahl von Geburtstagen, die auf einen Tag fallen, festgehalten. Nach jeder Simulation von n Geburtstagen wird der Wert dieser Variablen ausgedruckt.

Eine sinnvolle Ergänzung des Programms wäre es, die Anzahl der Simulationen zu notieren, in denen drei oder mehr Geburtstage auf einen

Tag gefallen sind.

Die exakte Berechnung der Wahrscheinlichkeit für das Zusammentreffen dreier Geburtstage dürfte wegen ihres Schwierigkeitsgrades einem Leistungskurs Mathematik oder dem Studium vorbehalten sein. Die Simulation des Problems dagegen sollte jedem Informatikkurs zugängig sein. Aus der Sicht des Informatikunterrichts handelt es sich dabei um eine interessante Übung für den Umgang mit Zufallszahlen, dimensionierten Feldern und rekursiv definierten Funktionen.

Für den Autor liegt der Reiz des Problems jedoch in der Verbindung vielfältiger Aspekte aus Kombinatorik, Numerik und Informatik.

Diese Arbeit ist Frau U.Jordan, Herrn Dr. J. Floer und Herrn Professor Dr. E.Ch. Wittmann anläßlich ihres Geburtstages am 28. Juli 1989 gewidmet.

Sie erschien ursprünglich in "Mathematische Semesterberichte", 1990, 37, S. 251–257