The Birthday Problem

There is a problem in mathematics relating to birthdays.   Since a year has 366 days (if you count February 29), there would have to be 367 people gathered together to be absolutely certain that two of them have the same birthday.

Now, if we were content to be just fifty percent sure, how many people are needed to be in the room?   A good guess might be 183 since that is half of 366.   The surprising answer is that there need only be 23 people present!   Stated differently, half of the time that twenty-three randomly selected people are gathered together, two or more of them will share the same birthday (an explanation is given in my letter below).   This is a delightful activity to do with students in your classroom.   If your class size is 23 or more, you have at least a 50% chance that two students will share a birthday.

The Problem: In a group of n people, what is the probability that two of them will have the same birthday?

Here is a letter that I wrote to Marilyn vos Savant about the Birthday Problem in Statistics (She published an edited version of this letter in November 1997 in her column in Parade magazine):

John Handley High School
P.O. Box 910
Winchester, VA 22604
August 3, 1997

Ask Marilyn
PARADE
711 Third Avenue
New York, NY 10017

Dear Ms. vos Savant:

I am a mathematics teacher and eagerly look forward to your column each week. I often use items from your column and from your books as material in my classes. One of my colleagues assigns her first year computer science students to write a program which simulates your famous "Game Show" problem, and every year you are proven correct! I have always agreed with the mathematical answers that you have given and marveled at your explanations. You explain things in layman's terms which are often better understood by my students than my mathematical proofs.

But in a recent column (8/3/97), I believe that you gave an incomplete answer to the question posed by Robert Shearn. Your logic was correct (you cannot reason "that if there's a 50% chance that at least two out of 23 people will share the same birthday, there must be a 100% chance that at least two out of 50 will!"). But while that type of extrapolation is incorrect, it is a coincidence that there is a 97% chance that at least two people out of a group of 50 will share a common birthday, and I believe it is that to which Mr. Shearn is referring. Of course, since there are 366 days on which people could have birthdays, there would have to be 367 people gathered in order to be absolutely 100% certain that at least two of them would have the same birthday.

In mathematics, this is called "The Birthday Problem." I use it along with your "Let's Make A Deal" problem and several other problems when I teach probability. I try to motivate this with my students by using the same approach you used in your response to Mr. Shearn -- by examining the probability that everyone in the group would have a different birthday. This probability along with the probability that at least 2 people have the same birthday must equal 1 (or 100% if you are using percentages).

For example, take a group of three people. The number of ways that any three dates (representing their birthdays and allowing for repetition) may be chosen is (366)x(366)x(366). In order for all of them to be born on different days, the first could be born on any of the 366 dates, the second could be born on any of 365 dates (she can not be born on the one day that the first person was born), and the third person could be born on any of the remaining 364 dates. So, all the ways that this could occur is (366)x(365)x(364). And therefore, the probability that no two people would have the same birthday out of a group of three people is the number of ways that their birthdays could be different divided by the total number of ways in which their birthdays could occur or (366)(365)(364) / (366)(366)(366).

The probability that at least two people would have the same birthday out of a group of three is

$$1 - \frac{(366)(365)(364)}{(366)(366)(366)} = .00818 \text{ or } .8\%$$

Similarly, it can be shown that out of a group of 23 people, the probability that at least two of them would have the same birthday is

$$1 - \frac{(366)(365)(364)(363)\ldots(344)}{(366)(366)(366)(366)\ldots(366)} = .506 \text{ or } 50.6\%.$$

And it can be shown that out of a group of 50 people, the probability that at least two of them would have the same birthday is

$$1 - \frac{(366)(365)(364)(363)...(317)}{(366)(366)(366)(366)...(366)} = .97 \text{ or } 97\%.$$

I have written two computer programs in C++ to illustrate the birthday problem and I have enclosed them in this mailing. The first program, called MARILYN.EXE, computes the theoretical probabilities using the techniques described above for groups ranging in size from 2 to 100. I have included that printout below. The second program, called BIRTHDAY.EXE, is a simulation of the problem allowing you to choose the size of the group and the number of times that you wish to run the program. It determines the number of times that at least two people have a matching birthday using pseudorandom numbers (this is similar to the technique our computer students use in verifying your "Game Show" problem).

As an application of the Birthday Problem, would you predict whether any two of the 41 different Presidents of the U.S. have the same birthday? The theoretical probability that at least two of them would be born on the same day is 90.25%. Upon an examination of their birthdays, we find that the eleventh President of the United States, James K. Polk, was born on November 2, 1795, and that the twenty-ninth President, Warren G. Harding, was born on November 2, 1865. Of course, if we use all 42 Presidents, there is another pair of Presidents with the same birthday - the 22nd and the 24th Presidents (Grover Cleveland and himself).

I hope the explanations make sense. Thank you for getting me focused on the new school year - I start my 30th year of teaching in less than two weeks!

Sincerely,

David H. Pleacher

The Probability of At Least Two People in a Group Having The Same Birthday
(Computer Printout)

Group Size	Probability	Percent Chance
2	0.002732	0.273224%
3	0.008182	0.818179%
4	0.016311	1.631145%
5	0.027062	2.706214%
6	0.040354	4.035364%
7	0.056086	5.608555%
8	0.074139	7.413856%
9	0.094376	9.437597%
10	0.116645	11.664541%
11	0.140781	14.078078%
12	0.166604	16.660431%
13	0.193929	19.392876%
14	0.22256	22.255971%
15	0.252298	25.229786%
16	0.282941	28.294139%
17	0.314288	31.428821%
18	0.346138	34.613822%
19	0.378295	37.829535%
20	0.41057	41.056964%
21	0.442779	44.277895%
22	0.474751	47.475065%
23	0.506323	50.632301%
24	0.537346	53.734643%
25	0.567684	56.768437%
26	0.597214	59.721412%
27	0.625827	62.582733%
28	0.65343	65.343023%
29	0.679944	67.994376%

Group Size	Probability	Percent Chance
30	0.705303	70.530341%
31	0.729459	72.945887%
32	0.752374	75.237356%
33	0.774024	77.402396%
34	0.794399	79.439884%
35	0.813498	81.349841%
36	0.831333	83.133326%
37	0.847923	84.792343%
38	0.863297	86.329729%
39	0.87749	87.749047%
40	0.890545	89.054476%
41	0.902507	90.250708%
42	0.913428	91.342842%
43	0.923363	92.336286%
44	0.932367	93.236668%
45	0.940497	94.049746%
46	0.947813	94.781335%
47	0.954372	95.437233%
48	0.960232	96.023162%
49	0.965447	96.544714%
50	0.970073	97.007307%
51	0.974161	97.416145%
52	0.977762	97.77619%
53	0.980921	98.092141%
54	0.983684	98.368416%
55	0.986091	98.609142%
56	0.988182	98.81815%
57	0.98999	98.99898%

Group Size	Probability	Percent Chance
58	0.991549	99.154876%
59	0.992888	99.288803%
60	0.994034	99.40345%
61	0.995012	99.501245%
62	0.995844	99.584371%
63	0.996548	99.654778%
64	0.997142	99.714201%
65	0.997642	99.764177%
66	0.998061	99.806058%
67	0.99841	99.841031%
68	0.998701	99.870132%
69	0.998943	99.894261%
70	0.999142	99.914195%
71	0.999306	99.930606%
72	0.999441	99.944068%
73	0.999551	99.955071%
74	0.99964	99.964032%
75	0.999713	99.971304%
76	0.999772	99.977184%
77	0.999819	99.981922%
78	0.999857	99.985725%
79	0.999888	99.988768%
80	0.999912	99.991192%
81	0.999931	99.993117%
82	0.999946	99.99464%
83	0.999958	99.995841%
84	0.999968	99.996784%
85	0.999975	99.997522%

Group Size	Probability	Percent Chance
86	0.999981	99.998098%
87	0.999985	99.998545%
88	0.999989	99.998891%
89	0.999992	99.999157%
90	0.999994	99.999362%
91	0.999995	99.999519%
92	0.999996	99.999639%
93	0.999997	99.999729%
94	0.999998	99.999798%
95	0.999999	99.99985%
96	0.999999	99.999889%
97	0.999999	99.999918%
98	0.999999	99.99994%
99	1	99.999956%
100	1	99.999968%

These last probabilities are not really equal to one - they are just rounded to one because I only took the decimal representation out to six places.

Isn't it interesting to see that in groups of 55 or more, you have a 99% chance of having at least two people with the same birthday? And yet, to be 100% certain, you would need a group size of 367!

Some references for The Birthday problem include:

Paulos, John Allen, Innumeracy, Hill and Wang, New York, 1988, p. 27.

Johnson, John, "The Birthday Problem Explained", The Mathematics Teacher, N.C.T.M., Reston, VA, January, 1997, pp. 20-22.