5. Refer to Table 16.1 of Vignette 16, "Was the 1970 Draft Lottery Fair?" M. Hollander's birthday is in the month of March.
   a. What is the chance that his birthday received one of the numbers 1, 2, . . . , 100 in the draft order?
   b. Suppose you have the further information that M. Hollander's birthday occurs within the first 22 days of March. What is your revised estimate of the probability that his birthday received one of the first 100 numbers in the draft order?
6. If you are told that the entry in the northwest corner of a page of random numbers (see Vignette 12) is a 1, should you then revise your opinion about the chance that the entry in the southeast corner of that page is a 1? Explain.
7. Give two examples from your own experience of formally or informally revising your assessment of the probability of an event in the light of additional information.
8. Without looking ahead (can you restrain yourself?), what is your estimate of the probability that the next vignette exceeds 10 pages in length? Explain the basis for your calculation. Now look ahead. What is your revised estimate of the probability?
9. Suppose a woman is selected at random from the 14 women represented by Table 18.1 of Vignette 18. What is the chance that the woman's second choice for leisure-time companions is "both sexes?" Given that the woman selected lists her most preferred choice as "women," what is the chance that her second choice is "both sexes?"

References

1. E. Rubin (1965). On missing persons and missing tabulations. *The American Statistician 19*, 33–36.†
2. I. R. Savage (1968). *Statistics: Uncertainty and Behavior.* Houghton Mifflin, Boston.
3. J. Yerushalmy, J. T. Harkness, J. H. Cope, and B. R. Kennedy (1950). The role of dual reading in mass radiography. *American Review of Tuberculosis 61*, 443–464.

References 1 and 3 contain the data of Tables 24.2 and 24.1, respectively. Reference 2 is of interest because it uses elementary gambling considerations to derive basic rules for computing and revising probabilities.

† This issue may be missing from your library.

Vignette 25

The Birthday Problem

The following problem has captivated the interests of many students in elementary probability and statistics classes. In a class of 28 students, what is the chance that at least 2 students have the same birthday (ignoring year of birth)?

Incorrect reasoning goes something like this. Excluding February 29, there are 365 possible birthdays but only 28 students. Thus, the chance should be about $28/365 = .077$, or about 8%. Actually the chance is much higher—about 65%. Many people find this result striking and hard to believe.

Before we justify it by precise calculation, we present similar probabilities for various-size classes. Table 25.1 gives the probability that at least two students have the same birthday and the complementary probability that no two students in the class have the same birthday. These probability values may seem counterintuitive, but we will prove that they are correct.

Note that when the number of students is 23, the chance is already slightly greater than 50% that at least 2 of the students in the class have the same birthday. That is, if you checked with a large number of groups consisting of 23 people each, about half of the groups would have birthday coincidences, while the other half would not.

When there are 64 in the class, it is *virtually certain* (the chance is .997) that the class will have at least two people with the same birthday.

We now describe how the values in columns 2 and 3 of Table 25.1 are obtained. First, for simplicity, we exclude February 29, so that there are only 365 possible birthdays. Second, we assume all 365 birthdays are equally likely. That is, if a student is selected at random from the class, we assume that each day of the year is equally likely to be that person's birthday. Now, we start with the simplest case: suppose that the class contains only *two* students. Since we have assumed all birthdays are equally likely, the probability that the two students have different birthdays is equal to:

Table 25.1 Birthday Coincidence Probabilities

Number of Students in Class	Probability that No Two Students in the Class Have the Same Birthday	Probability that Two or More Students in the Class Have the Same Birthday
2	.997	.003
4	.984	.016
8	.926	.074
12	.833	.167
16	.716	.284
20	.589	.411
22	.524	.476
23	.493	.507
24	.462	.538
28	.346	.654
32	.247	.753
40	.109	.891
48	.039	.961
56	.012	.988
64	.003	.997

$$\frac{\text{number of ways to assign birthdays to two students so that their birthdays are different}}{\text{total number of ways to assign birthdays to two students}}.$$

To evaluate the denominator of the fraction, note that we can assign the first birthday in 365 ways and the second birthday in 365 ways; thus, the total number of ways to assign two birthdays is 365×365.

To compute the numerator of the fraction, we count the number of ways to assign two birthdays so that the two birthdays are *not* the same. We still choose the first birthday in 365 ways, but since the second birthday must not be the same as the first birthday, there are only 364 ways to choose the second birthday. This gives a total of 365×364 ways to assign 2 birthdays so that they are not the same.

Thus, the desired probability is:

Probability that the 2 students
in the class have different birthdays

$$= \frac{365 \times 364}{365 \times 365} = \frac{364}{365} = .997.$$

By similar reasoning, if the class contains exactly 3 students, then

Probability that the 3 students
in the class have different birthdays

$$= \frac{365 \times 364 \times 363}{365 \times 365 \times 365} = \frac{364 \times 363}{365 \times 365} = .992.$$

In a class with just 4 students,

Probability that the 4 students
in the class have different birthdays

$$= \frac{365 \times 364 \times 363 \times 362}{365 \times 365 \times 365 \times 365} = .984.$$

Continuing in this fashion, you could, with the aid of a good calculator, verify the entries in column 2 of Table 25.1. Futhermore you can calculate, for any specific number (not listed in the table) of students in the class, the chance that all the students have different birthdays.

To get a value in column 3 of Table 25.1, simply subtract the corresponding entry in column 2 from the number 1. For example, in row 1 find $.003 = 1 - .997$; in row 2, $.016 = 1 - .984$, and so on.

Recall that we started our discussion with exactly 28 students in the class. Thus,

Probability that the 28 students
all have different birthdays

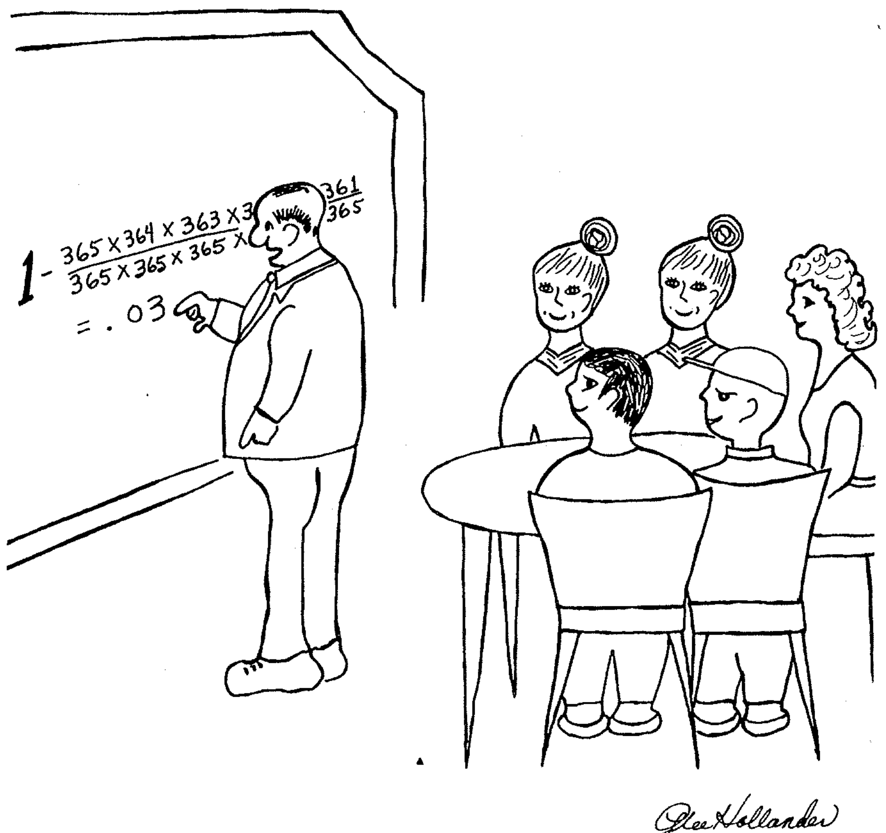
$$= \frac{365 \times 364 \times 363 \times \ldots \times 338}{\underbrace{365 \times 365 \times 365 \times \ldots \times 365}_{\text{28 factors of 365}}} = .346,$$

and

Probability that two or more of the 28
in the class have the same birthday
$= 1 - .346 = .654.$

There are several reasons why students tend to underestimate this probability when they guess it.

First, a student may ask himself: What is the chance that out of the remaining 27 students at least 1 will have the same birthday as mine? He should really be asking: What is the chance that out of the 28 students, at least 2 will have birthdays that agree? (But they do not necessarily have to agree with his own birthday.)



"That's strange. In a class of five students, there is only a 3% chance of a birthday coincidence."

Second, a student, when thinking about how at least two people can have the same birthday, may neglect many possibilities. The possibilities include not only a pair of students having the same birthday, but also, to name another possibility, two students having one date in common *and* three others having another date in common, and so forth.

Summary

1. We have shown how to calculate the probability that in a class consisting of (for example) 28 students at least 2 have the same birthday.

2. The calculation of this probability utilizes an "equally likely" assumption that could be questioned.
3. To the inexperienced reader, the results of this vignette may seem counterintuitive. Intuition is often helpful, but in this case careful and rational thought, backed up by a correct probability calculation, will show that intuition has led one astray.

Problems

1. Table 25.2 gives the birthdays of the 39 presidents of the United States as of 1983. Note that James Polk and Warren Harding were both born on November 2, and that Andrew Johnson and Woodrow Wilson were both born on December 29. Do you find these coincidences surprising in view of the calculations of this chapter?
2. Find out the birthdates of the current 100 U.S. senators. Do you expect to find any birthday coincidences?
3. A small class contains six students. What is the chance that at least two have the same *birthmonth*?
4. Do you think the assumption that all 365 birthdates are equally likely is reasonable? Explain.

Table 25.2 Birthdates of American Presidents

George Washington	Feb. 22, 1732	Chester A. Arthur	Oct. 5, 1829
John Adams	Oct. 30, 1735	Grover Cleveland	Mar. 18, 1837
Thomas Jefferson	Apr. 13, 1743	Benjamin Harrison	Aug. 20, 1833
James Madison	Mar. 16, 1751	William McKinley	Jan. 29, 1843
James Monroe	Apr. 28, 1758	Theodore Roosevelt	Oct. 27, 1858
John Quincy Adams	July 11, 1767	William H. Taft	Sept. 15, 1857
Andrew Jackson	Mar. 15, 1767	Woodrow Wilson	Dec. 29, 1856
Martin Van Buren	Dec. 5, 1782	Warren G. Harding	Nov. 2, 1865
William H. Harrison	Feb. 9, 1773	Calvin Coolidge	July 4, 1872
John Tyler	Mar. 29, 1790	Herbert C. Hoover	Aug. 10, 1874
James K. Polk	Nov. 2, 1795	Franklin D. Roosevelt	Jan. 30, 1882
Zachary Taylor	Nov. 24, 1784	Harry S. Truman	May 8, 1884
Millard Fillmore	Jan. 7, 1800	Dwight D. Eisenhower	Oct. 14, 1890
Franklin Pierce	Nov. 23, 1804	John F. Kennedy	May 29, 1917
James Buchanan	Apr. 23, 1791	Lyndon B. Johnson	Aug. 27, 1908
Abraham Lincoln	Feb. 12, 1809	Richard M. Nixon	Jan. 9, 1913
Andrew Johnson	Dec. 29, 1808	Gerald R. Ford	July 14, 1913
Ulysses S. Grant	Apr. 27, 1822	Jimmy Carter	Oct. 1, 1924
Rutherford B. Hayes	Oct. 4, 1822	Ronald Reagan	Feb. 6, 1911
James A. Garfield	Nov. 19, 1831		