

# The New York Times

---

## The Opinion Pages

Opinionator

A Gathering of Opinion From Around the Web

# Chances Are

By Steven Strogatz    April 25, 2010 5:00 pm

Have you ever had that anxiety dream where you suddenly realize you have to take the final exam in some course you've never attended? For professors, it works the other way around — you dream you're giving a lecture for a class you know nothing about.

That's what it's like for me whenever I teach probability theory. It was never part of my own education, so having to lecture about it now is scary and fun, in an amusement park, thrill-house sort of way.

Perhaps the most pulse-quickenning topic of all is “conditional probability” — the probability that some event A happens, given (or “conditional” upon) the occurrence of some other event B. It's a slippery concept, easily conflated with the probability of B given A. They're not the same, but you have to concentrate to see why. For example, consider the following word problem.

Before going on vacation for a week, you ask your spacey friend to water your ailing plant. Without water, the plant has a 90 percent chance of dying. Even with proper watering, it has a 20 percent chance of dying. And the probability that your friend will forget to water it is 30 percent. (a) What's the chance that your plant will survive the week? (b) If it's dead when you return, what's the chance that your friend forgot to water it? (c) If your friend forgot to water it, what's the chance it'll be dead when you return?

Although they sound alike, (b) and (c) are not the same. In fact, the problem tells us

that the answer to (c) is 90 percent. But how do you combine all the probabilities to get the answer to (b)? Or (a)?

Naturally, the first few semesters I taught this topic, I stuck to the book, inching along, playing it safe. But gradually I began to notice something. A few of my students would avoid using “Bayes’s theorem,” the labyrinthine formula I was teaching them. Instead they would solve the problems by a much easier method.

What these resourceful students kept discovering, year after year, was a better way to think about conditional probability. Their way comports with human intuition instead of confounding it. The trick is to think in terms of “natural frequencies” — simple counts of events — rather than the more abstract notions of percentages, odds, or probabilities. As soon as you make this mental shift, the fog lifts.

This is the central lesson of “Calculated Risks,” a fascinating book by Gerd Gigerenzer, a cognitive psychologist at the Max Planck Institute for Human Development in Berlin. In a series of studies about medical and legal issues ranging from AIDS counseling to the interpretation of DNA fingerprinting, Gigerenzer explores how people miscalculate risk and uncertainty. But rather than scold or bemoan human frailty, he tells us how to do better — how to avoid “clouded thinking” by recasting conditional probability problems in terms of natural frequencies, much as my students did.

In one study, Gigerenzer and his colleagues asked doctors in Germany and the United States to estimate the probability that a woman with a positive mammogram actually has breast cancer, even though she’s in a low-risk group: 40 to 50 years old, with no symptoms or family history of breast cancer. To make the question specific, the doctors were told to assume the following statistics — couched in terms of percentages and probabilities — about the prevalence of breast cancer among women in this cohort, and also about the mammogram’s sensitivity and rate of false positives:

*The probability that one of these women has breast cancer is 0.8 percent. If a woman has breast cancer, the probability is 90 percent that she will have a positive mammogram. If a woman does not have breast cancer, the probability is 7 percent*

*that she will still have a positive mammogram. Imagine a woman who has a positive mammogram. What is the probability that she actually has breast cancer?*

Gigerenzer describes the reaction of the first doctor he tested, a department chief at a university teaching hospital with more than 30 years of professional experience:

“[He] was visibly nervous while trying to figure out what he would tell the woman. After mulling the numbers over, he finally estimated the woman’s probability of having breast cancer, given that she has a positive mammogram, to be 90 percent. Nervously, he added, ‘Oh, what nonsense. I can’t do this. You should test my daughter; she is studying medicine.’ He knew that his estimate was wrong, but he did not know how to reason better. Despite the fact that he had spent 10 minutes wringing his mind for an answer, he could not figure out how to draw a sound inference from the probabilities.”

When Gigerenzer asked 24 other German doctors the same question, their estimates whipsawed from 1 percent to 90 percent. Eight of them thought the chances were 10 percent or less, 8 more said 90 percent, and the remaining 8 guessed somewhere between 50 and 80 percent. Imagine how upsetting it would be as a patient to hear such divergent opinions.

As for the American doctors, 95 out of 100 estimated the woman’s probability of having breast cancer to be somewhere around 75 percent.

The right answer is 9 percent.

How can it be so low? Gigerenzer’s point is that the analysis becomes almost transparent if we translate the original information from percentages and probabilities into natural frequencies:

*Eight out of every 1,000 women have breast cancer. Of these 8 women with breast cancer, 7 will have a positive mammogram. Of the remaining 992 women who don’t have breast cancer, some 70 will still have a positive mammogram. Imagine a sample of women who have positive mammograms in screening. How many of these women actually have breast cancer?*

Since a total of  $7 + 70 = 77$  women have positive mammograms, and only 7 of them truly have breast cancer, the probability of having breast cancer given a positive mammogram is 7 out of 77, which is 1 in 11, or about 9 percent.

Notice two simplifications in the calculation above. First, we rounded off decimals to whole numbers. That happened in a few places, like when we said, “Of these 8 women with breast cancer, 7 will have a positive mammogram.” Really we should have said 90 percent of 8 women, or 7.2 women, will have a positive mammogram. So we sacrificed a little precision for a lot of clarity.

Second, we assumed that everything happens exactly as frequently as its probability suggests. For instance, since the probability of breast cancer is 0.8 percent, exactly 8 women out of 1,000 in our hypothetical sample were assumed to have it. In reality, this wouldn’t necessarily be true. Things don’t have to follow their probabilities; a coin flipped 1,000 times doesn’t always come up heads 500 times. But pretending that it does gives the right answer in problems like this.

Admittedly the logic is a little shaky — that’s why the textbooks look down their noses at this approach, compared to the more rigorous but hard-to-use Bayes’s theorem — but the gains in clarity are justification enough. When Gigerenzer tested another set of 24 doctors, this time using natural frequencies, nearly all of them got the correct answer, or close to it.

Although reformulating the data in terms of natural frequencies is a huge help, conditional probability problems can still be perplexing for other reasons. It’s easy to ask the wrong question, or to calculate a probability that’s correct but misleading.

Both the prosecution and the defense were guilty of this in the O.J. Simpson trial of 1994-95. Each of them asked the jury to consider the wrong conditional probability.

The prosecution spent the first 10 days of the trial introducing evidence that O.J. had a history of violence toward his ex-wife, Nicole. He had allegedly battered her, thrown her against walls and groped her in public, telling onlookers, “This belongs to me.” But what did any of this have to do with a murder trial? The

prosecution's argument was that a pattern of spousal abuse reflected a motive to kill. As one of the prosecutors put it, "A slap is a prelude to homicide."

Alan Dershowitz countered for the defense, arguing that even if the allegations of domestic violence were true, they were irrelevant and should therefore be inadmissible. He later wrote, "We knew we could prove, if we had to, that an infinitesimal percentage — certainly fewer than 1 of 2,500 — of men who slap or beat their domestic partners go on to murder them."

In effect, both sides were asking the jury to consider the probability that a man murdered his ex-wife, given that he previously battered her. But as the statistician I. J. Good pointed out, that's not the right number to look at.

The real question is: What's the probability that a man murdered his ex-wife, given that he previously battered her *and she was murdered by someone*? That conditional probability turns out to be very far from 1 in 2,500.

To see why, imagine a sample of 100,000 battered women. Granting Dershowitz's number of 1 in 2,500, we expect about 40 of these women to be murdered by their abusers in a given year (since 100,000 divided by 2,500 equals 40). We can estimate that an additional 5 of these battered women, on average, will be killed by someone else, because the murder rate for all women in the United States at the time of the trial was about 1 in 20,000 per year. So out of the  $40 + 5 = 45$  murder victims altogether, 40 of them were killed by their batterer. In other words, the batterer was the murderer about 90 percent of the time.

Don't confuse this number with the probability that O.J. did it. That probability would depend on a lot of other evidence, pro and con, such as the defense's claim that the police framed him, or the prosecution's claim that the killer and O.J. shared the same style of shoes, gloves and DNA.

The probability that any of this changed your mind about the verdict? Zero.

---

NOTES:

1. For a good textbook treatment of conditional probability and Bayes's theorem, see:  
S.M. Ross, "**Introduction to Probability and Statistics for Engineers and Scientists**," 4th edition (Academic Press, 2009).
2. The answer to part (a) of the "ailing plant" problem is 59 percent. The answer to part (b) is 27/41, or approximately 65.85 percent. To derive these results, imagine 100 ailing plants and figure out (on average) how many of them get watered or not, and then how many of those go on to die or not, based on the information given. This question appears, though with slightly different numbers and wording, as problem 29 on p. 84 of Ross's text.
3. The study of how doctors interpret mammogram results is described in:  
G. Gigerenzer, "**Calculated Risks**" (Simon and Schuster, 2002), chapter 4.
4. For more on the O.J. Simpson case and a discussion of wife battering in a larger context, see chapter 8.
4. For many entertaining anecdotes and insights about conditional probability and its real-world applications, as well as how it's misperceived, see:  
J.A. Paulos, "**Innumeracy**" (Vintage, 1990);  
L. Mlodinow, "**The Drunkard's Walk**" (Vintage, 2009).
5. The quotes pertaining to the O.J. Simpson trial, and Alan Dershowitz's estimate of the rate at which battered women are murdered by their partners, appeared in:  
A. Dershowitz, "**Reasonable Doubts**" (Touchstone, 1997), pp. 101-104.
6. Probability theory was first correctly applied to the Simpson trial by the late I.J. Good, in:  
I.J. Good, "**When batterer turns murderer**," *Nature*, Vol. 375 (1995), p. 541.  
I.J. Good, "**When batterer becomes murderer**," *Nature*, Vol. 381 (1996), p. 481.  
Good phrased his analysis in terms of odds ratios and Bayes's theorem, rather than the more intuitive "natural frequency" approach presented here and in Gigerenzer's book.  
Good had an interesting career. In addition to his many contributions to

*probability theory and Bayesian statistics, he helped break the Nazi Enigma code during World War II, and introduced the futuristic concept now known as the “technological singularity.”*

7. *Here is how Dershowitz seems to have calculated that fewer than 1 in 2,500 batterers go on to murder their partners, per year. On page 104 of his book “Reasonable Doubts,” he cites an estimate that in 1992, somewhere between 2.5 and 4 million women in the United States were battered by their husbands, boyfriends, and ex-boyfriends. In that same year, according to the FBI Uniform Crime Reports, 913 women were murdered by their husbands, and 519 were killed by their boyfriends or ex-boyfriends. Dividing the total of 1,432 homicides by 2.5 million beatings yields 1 murder per 1,746 beatings, whereas using the higher estimate of 4 million beatings per year yields 1 murder per 2,793 beatings. Dershowitz apparently chose 2,500 as a round number in between these extremes.*

*What’s unclear is what proportion of the murdered women had been previously beaten by these men. It seems that Dershowitz was assuming that nearly all the victims were beaten, presumably to make the point that even when the rate is overestimated in this way, it’s still “infinitesimal.”*

8. *Good’s estimated murder rate of 1 per 20,000 women per year includes battered women, so it was not strictly correct to assume (as he did, and as we did above) that 5 women out of 100,000 would be killed by someone other than the batterer. But correcting for this doesn’t alter the conclusion significantly, as the following calculation shows.*

*According to the FBI Uniform Crime Reports, 4,936 women were murdered in 1992. Of these murder victims, 1,432 (about 29 percent) were killed by their husbands or boyfriends. The remaining 3,504 were killed by somebody else. Therefore, considering that the total population of women in the United States at that time was about 125 million, the rate at which women were murdered by someone other than their partners was 3,504 divided by 125,000,000, or 1 murder per 35,673 women, per year.*

*Let’s assume that this rate of murder by non-partners was the same for all women, battered or not. Then in our hypothetical sample of 100,000 battered women, we’d expect about 100,000 divided by 35,673, or 2.8 women to be killed by someone other than their partner. Although 2.8 is*

*smaller than the 5 that Good and we assumed above, it doesn't matter much because both are so small compared to 40, the estimated number of cases in which the batterer is the murderer. With this modification, our new estimate of the probability that the batterer is the murderer would be 40 divided by  $(40 + 2.8)$ , or about 93 percent.*

*A related quibble is that the FBI statistics and population data given above imply that the murder rate for women in 1992 was closer to 1 in 25,000, not 1 in 20,000 as Good assumed. If he had used that rate in his calculation, an estimated 4 women per 100,000, not 5, would have been murdered by someone other than the partner. But this still wouldn't affect the basic message — now the batterer would be the murderer 40 times out of  $40 + 4 = 44$ , or 91 percent of the time.*

9. *A few years after the verdict was handed down in the Simpson case, Alan Dershowitz and the mathematician John Allen Paulos engaged in a heated exchange of letters to the editor of the New York Times. The issue was whether evidence of a history of spousal abuse should be regarded as relevant to a murder trial, in light of probabilistic arguments similar to those discussed in this post. Dershowitz's letter to the editor and Paulos's response make for lively reading.*

*Thanks to Paul Ginsparg, Michael Lewis, Eri Noguchi and Carole Schiffman for their comments and suggestions.*

---

Editor's note: This column was updated to add Note 9 on April 30.

For those readers who want to print this column, our default print option, below, works well for this week's format, so there is no PDF version included.