



## Solution: 'When Probability Meets Real Life'

When our brains don't have a good intuition for reasoning with numbers, explicit probabilistic thinking can lead to improved decision-making.

*By Pradeep Mutalik*

When making hard decisions, do you go with your gut or try to calculate the risks? In many cases going with your gut is fine, but the answers to our [February puzzle](#) problems show how explicit probabilistic thinking can outperform intuitive estimates. They also highlight the differences between situations where an intuitive approach succeeds and ones where it fails.

Our first problem was an excerpt from a whimsical children's story.

### **Problem 1**

A man went on an airplane ride.

Unfortunately, he fell out.

Fortunately, he had a parachute on.

Unfortunately, the parachute did not open.

Fortunately, there was a haystack below him, directly in the path of his fall.

Unfortunately, there was a pitchfork sticking out of the top of the haystack, directly in the path of his fall.

Fortunately, he missed the pitchfork.

Unfortunately, he missed the haystack.

There have indeed been alleged instances of people surviving falls from planes by landing on haystacks, or even by falling into trees or thick shrubbery, as a quick online search will reveal. So the alternating screams inside this imaginary man's head — "I'm dead!"/"I'm saved!" — are not conclusive until the sad finale. (Though our story seems to end tragically, in the original the protagonist survives with many more abrupt reversals of fortune!) Can principled methods of risk estimation be applied here? Given the available information, estimate his odds of survival at the end of each line above.

[puzzles]

[Douglas Felix](#) made a very good attempt at answering this question. As his answers indicate, the probability of survival at each step is either very close to zero or very close to one, so it is very hard to express the absolute and relative risks in a way that's easy to understand. The Stanford scientist [Ronald Howard](#), pioneer of risk analysis, invented a measure for just this purpose — the [micromort](#). The micromort is a one-in-a-million chance of death. You can use the usual metric prefixes to express risks even closer to zero, such as nanomorts and picomorts. For the purposes of our puzzle, we can name a flipped unit, the microvivé, for situations in which the chances of survival are very close to zero — a risk of one microvivé would mean a one in a million chance of surviving. The sum of the number of micromorts and microvivés is always one million. Armed with these measures, let's take a look at the risks in the story above, assuming it happened around the present time in history. For reference, the average risk of death for a middle-aged person per day is about 1 micromort.

*A man went on an airplane ride.*

In the absence of any more information, it is overwhelmingly likely that a random plane ride is on a commercial air carrier. The rate of fatal accidents worldwide for this today is about 0.2 to 0.4 per million flights, giving a risk of death of 0.2 to 0.4 micromorts, not adding very much to one's normal everyday risk. However, if we knew that the man had a parachute strapped on, as is revealed in the third line, we would have to conclude that this was probably no ordinary ride. Perhaps it was a skydiving trip or military aircraft mission. In that case, we would need to use the figures reported for general aviation fatalities, not the much-better figures for commercial air carriers. One [published](#) estimate based on official data puts this risk at something on the order of 1 in 100,000, which is 10 micromorts.

*Unfortunately, he fell out.*

The chances of survival after a fall from an airplane are hard to estimate, but probably not as high as we think. When we hear this sentence, we assume he didn't have a parachute, but that is not explicitly denied yet. Even if he didn't have a parachute, there are documented cases of a handful of people who survived a fall from an airplane. The number of people who have fallen out of airplanes without parachutes throughout history is probably not more than 100,000 to 300,000, based on published air fatality numbers. Based on these two facts, the chances of survival without a parachute may range from 1 in 10,000 to 1 in 30,000, which is 30 to 100 microvivés (a chance of dying of 999,900 to 999,970 micromorts).

*Fortunately, he had a parachute on.*

The official figures for skydiving accidents put the risk of dying in a skydiving accident at about 8 to 9 micromorts. The chance of dying has dropped considerably.

*Unfortunately, the parachute did not open.*

The chances of survival plummet to 30 to 100 microvivés, as above.

*Fortunately, there was a haystack below him ...*

Considering that people have survived falls into shrubbery, a haystack could possibly improve the chances of survival, perhaps allowing an escape with some broken bones. It is hard to estimate, but the odds are perhaps somewhere between one in a hundred to one in a thousand of surviving — 1,000 to 10,000 microvivés, many times better than before.

*Unfortunately, there was a pitchfork sticking out ...*

Ouch! But with a surrounding haystack, the odds would still be better than if the pitchfork and haystack were absent. I'd put the odds somewhere between 50 and 500 microvivés.

*Fortunately, he missed the pitchfork.*

Back to 1,000 to 10,000 microvivés.

*Unfortunately, he missed the haystack.*

As after line two, back to 30 to 100 microvivés, or a 99.990 to 99.997 percent chance of dying.

As is obvious, we need hard data to give tight probabilities. Nevertheless, we can give reasonable, albeit broad estimates. The units, once you get used to them, make things clearer and give more information rather than simply saying, "very, very close to zero/one."

(It is tempting to invent similar units to add clarity in other situations: For instance, a one-in-a-million chance of an utterance being true could be called a microvérité or a microvér for short — a handy unit in this era of "fake news.")

## Problem 2

The number of deaths on commercial airlines is about 0.2 deaths per 10 billion flight-miles. For driving, the rate is 150 deaths per 10 billion vehicle-miles. While this rate is about 750 times higher than for air travel, we still take long road trips because the absolute risks are small. But let us pursue a thought experiment using two hypothetical and admittedly unrealistic assumptions — first, that your expected life span is 1 million years (and you enjoy every year of it!), and second, that the above risks remain the same during those million years. Now imagine that every year you could either fly 10,000 miles or cover that distance by car over multiple road trips. The time spent is not a concern at all — after all, you have a million years to live! Under these conditions, by how many years and by what proportion would your life be shortened if you lived a million years and drove every time instead of flying? How would your answer differ for a more normal life span of 100 years?

[Alexandre](#) answered this problem correctly. To figure out how much your life would be shortened on average by flying, consider that, in this problem, a flying death occurs every 50 billion flight-miles on average. In a life span of 1 million years, flying 10,000 miles a year, you cover only a fifth of this distance, a total of just 10 billion miles. So you have an 80 percent chance of living a full million years. This means that out of a group of five randomly selected people, four would be expected to live one million years, while the fifth would be expected to have a fatal accident sometime in his or her life, for an average life span of half a million years. So the average life expectancy for the entire group would be 4.5 million/5 or 900,000 years, or 90 percent of normal, as Alexandre correctly calculated. On the other hand, for driving, a death occurs on average every 66.67 million miles (10 billion/150), a distance you'd expect to have driven by the time you were 6,667 years old. So the average 1-million-year lifetime would be reduced to 6,667 years, or about 0.67 percent of its expected length. Using the same method of calculation for a 100-year life span, we obtain the following results: For flying, the average life span will be shortened by 0.001 years or about eight hours, whereas for driving it will be shortened by 0.075 years or about a month.

What this demonstrates is that people who lived a million years would tend to be very risk-averse! But more seriously, it shows that choosing just slightly riskier options repeatedly can increase risk significantly over time.

## Problem 3

Here are two similar scenarios in which you have to make probability judgments. Before you make an exact calculation, hazard an intuitive guess and jot it down.

Variation A: A certain town has two ethnic groups, the Ones and the Twos. Ones make up 80 percent of the population. A hospital clinic conducts a standard, unbiased screening test for a rare disease that is equally common in both groups. It results in the collection of 100 blood samples, and sure enough, 80 of the samples come from Ones. On more rigorous testing, just one of the 100 samples is found to be positive for the disease. A researcher who is not privy to the ethnicity data because of [HIPAA laws](#) runs a test on this sample, which determines that it comes from a Two. However, this ethnicity-determining test is known to be only 75 percent accurate. What is the probability that the sample actually came from a Two?

Variation B: In this variation, Ones and Twos both make up 50 percent of the population, but Ones are more likely to have the rare disease. The same screening procedure as above collects 100 blood samples, again yielding 80 from Ones and 20 from Twos. The rest of the problem is exactly the same. Now what is the probability that the diseased sample actually came from a Two?

In which of the two cases was your intuition more accurate?

The right answer for the probability, in the case of both variations, is  $3/7$  or 42.9 percent, which was obtained by [Douglas Felix](#) and [Alexandre](#). This is a standard question in probability theory, and can be solved by the standard Bayesian conditional probability formula  $P(A|B) = P(B|A) \times P(A) / P(B)$ , or for those who like visual thinking, by using the elegant geometric approach illustrated by [Alexandre](#). Below I give a simple arithmetic approach that can always be applied to such problems and offers an understanding of how the answer follows logically.

Let's start with the 100 samples in variation A. Of these, 80 belong to Ones. The ethnicity test is accurate three-quarters of the time, so it will label only 60 of these samples as belonging to Ones, and mislabel the other 20 as belonging to Twos. Similarly, out of the 20 samples from Twos, the ethnicity test will label 15 correctly as belonging to Twos, and mislabel 5 as belonging to Ones. Thus, the ethnicity test will label a total of  $20 + 15 = 35$  samples as Twos, out of which only 15 are truly Twos. Since the screening test is unbiased, any one of these is equally likely to have the rare disease. So the chance that the sample actually came from a Two is  $15/35$  or 42.9 percent. What this means is that even though the ethnicity test is reasonably accurate, and has identified the disease sample as that of a Two, that sample is still more likely to have come from a One! When asked to guess in situations like these, most people estimate the chance that the disease sample came from a Two to be around 75 percent.

Note that when we say that the screening test is unbiased, we mean that all individuals whose screening test is positive have an equal chance of having the disease, regardless of their ethnicity. A couple of readers wrote that the test could not be unbiased in variation B, because 80 percent of the samples were from Ones, even though there are equal numbers of Ones and Twos in the population. Here's how this could happen. Let's say that high blood pressure predisposes a person to the disease, and the screening test catches everyone with a blood pressure above a certain threshold. Now, in Variation B, Ones may be genetically predisposed to get high blood pressure. So the screening test collects more samples from Ones. And therefore, as mentioned, Ones are more likely to get the disease. Once the screening test is done, however, its ethnically unbiased nature ensures that all samples collected are equally likely to have the rare disease. Thus variation B also gives the same answer: 42.9 percent. In this case, however, our intuitions kick in correctly, and our estimated probability comes out closer to the correct answer.

Why should this be? Such errors in our intuitive reasoning about probabilities and the reasons behind them have been fascinatingly classified and described by the Nobel Prize-winning psychologist Daniel Kahneman in his best-selling book *Thinking, Fast and Slow*. As Kahneman notes, the kind of false Bayesian reasoning demonstrated in our problem was elucidated by the psychologist Icke Ajzen. Kahnemann describes this false reasoning as “causes trump statistics.” Our brains reason better with causes, rather than with numbers. In variation B, the idea that Ones are more prone to the disease is a cause we can latch on to, and it helps us discount the results of the ethnicity test as we should; in variation A, however, we don’t have such a causal story to tell ourselves — we just have numbers, which are not as compelling. Hence we give greater importance to the results of the ethnicity test, and discount the numbers, even though we should not. Mathematically, both variations are exactly the same.

Thank you to all those who commented. The *Quanta* T-shirt goes to Alexandre. Congratulations!

Insights will take a break in March and continue on a bimonthly basis. See you all in April.