Is there really a 1 in 6 chance of human extinction this century?

October 6 2023, by Steven Stern

Credit: AI-generated image ([disclaimer](#))

In 2020, Oxford-based philosopher Toby Ord published a book called [The Precipice](#) about the risk of human extinction. He put the chances of "existential catastrophe" for our species during the next century at 1 in 6.

It's quite a specific number, and an alarming one. The claim drew

[headlines](#) at the time, and has been influential since—most recently brought up by Australian politician Andrew Leigh in a [speech](#) in Melbourne.

It's hard to disagree with the idea we face troubling prospects over the coming decades, from climate change, [nuclear weapons](#) and bio-engineered pathogens (all big issues in my view), to rogue AI and large asteroids (which I would see as less concerning).

But what about that number? Where does it come from? And what does it really mean?

Coin flips and weather forecasts

To answer those questions, we have to answer another first: what is probability?

The most traditional view of probability is called frequentism, and derives its name from its heritage in games of dice and cards. On this view, we know there is a 1 in 6 chance a fair die will come up with a three (for example) by observing the frequency of threes in a large number of rolls.

Or consider the more complicated case of weather forecasts. What does it mean when a weatherperson tells us there is a 1 in 6 (or 17%) chance of rain tomorrow?

It's hard to believe the weatherperson means us to imagine a large collection of "tomorrows," of which some proportion will experience precipitation. Instead, we need to look at a large number of such predictions and see what happened after them.

If the forecaster is good at their job, we should see that when they said

"1 in 6 chance of rain tomorrow," it did in fact rain on the following day one time in every six.

So, traditional probability depends on observations and procedure. To calculate it, we need to have a collection of repeated events on which to base our estimate.

Can we learn from the moon?

So what does this mean for the probability of human extinction? Well, such an event would be a one-off: after it happened, there would be no room for repeats.

Instead, we might find some parallel events to learn from. Indeed, in Ord's book, he discusses a number of potential extinction events, some of which can potentially be examined in light of a history.

For example, we can estimate the chances of an extinction-sized asteroid hitting Earth by examining how many such space rocks have hit the moon over its history. A French scientist named Jean-Marc Salotti did this in 2022, calculating the odds of an extinction-level hit in the next century at around 1 in 300 million.

Of course, such an estimate is fraught with uncertainty, but it is backed by something approaching an appropriate frequency calculation. Ord, by contrast, estimates the risk of extinction by asteroid at 1 in 1 million, though he does note a considerable degree of uncertainty.

A ranking system for outcomes

There is another way to think about probability, called Bayesianism after the English statistician Thomas Bayes. It focuses less on events

themselves and more on what we know, expect and believe about them.

In very simple terms, we can say Bayesians see probabilities as a kind of ranking system. In this view, the specific number attached to a probability shouldn't be taken directly, but rather compared to other probabilities to understand which outcomes are more and less likely.

Ord's book, for example, contains a table of potential extinction events and his personal estimates of their probability. From a Bayesian perspective, we can view these values as relative ranks. Ord thinks extinction from an asteroid strike (1 in 1 million) is much less likely than extinction from climate change (1 in a 1,000), and both are far less likely than extinction from what he calls "unaligned artificial intelligence" (1 in 10).

The difficulty here is that initial estimates of Bayesian probabilities (often called "priors") are rather subjective (for instance, I would rank the chance of AI-based extinction much lower). Traditional Bayesian reasoning moves from "priors" to "posteriors" by again incorporating observational evidence of relevant outcomes to "update" probability values.

And once again, outcomes relevant to the probability of human extinction are thin on the ground.

Subjective estimates

There are two ways to think about the accuracy and usefulness of probability calculations: calibration and discrimination.

Calibration is the correctness of the actual values of the probabilities. We can't determine this without appropriate observational information. Discrimination, on the other hand, simply refers to the relative rankings.

We don't have a basis to think Ord's values are properly calibrated. Of course, this is not likely to be his intent. He himself indicates they are mostly designed to give "order of magnitude" indications.

Even so, without any related observational confirmation, most of these estimates simply remain in the subjective domain of prior probabilities.

Not well calibrated—but perhaps still useful

So what are we to make of "1 in 6"? Experience suggests most people have a less than perfect understanding of probability (as evidenced by, among other things, the ongoing volume of lottery ticket sales). In this environment, if you're making an argument in public, an estimate of "probability" doesn't necessarily need to be well calibrated—it just needs to have the right sort of psychological impact.

From this perspective, I'd say "1 in 6" fits the bill nicely. "1 in 100" might feel small enough to ignore, while "1 in 3" might drive panic or be dismissed as apocalyptic raving.

As a person concerned about the future, I hope risks like climate change and nuclear proliferation get the attention they deserve. But as a data scientist, I hope the careless use of probability gets left by the wayside and is replaced by widespread education on its true meaning and appropriate usage.

This article is republished from The Conversation under a Creative Commons license. Read the original article.

Provided by The Conversation