

How Not to Be Wrong

Author: Jordan Ellenberg

About the book:

How Not to Be Wrong gives us an intimate glimpse into how mathematicians think and how we can benefit from their way of thinking. It also explains how easily we can be mistaken when we apply mathematical tools incorrectly, and gives advice on how we can instead find correct solutions.

What's in it for me? Learn how math influences everything and how it can help you not be wrong.

Do you find mathematics challenging, or perhaps irrelevant to what you do every day? Many people think math is best left in a classroom, but in reality, it touches our daily lives in profound ways.

That's because mathematics is really the science of common sense, a reflection of things we already know intuitively.

Mathematicians speak in a specialized language so they can convey complex ideas quickly and precisely. To some, this may make mathematics seem complicated or beyond comprehension.

The mental work involved in complex mathematics, however, isn't that different from the thinking we do in many common situations. In these blinks, you'll learn how to uncover the "hidden math" in your life, use it to make sense of problems and importantly, learn *not to be wrong*.

In the following blinks, you'll also discover:

- why so many research findings published in journals are actually wrong;
- why a debut novelist's second book is usually worse than his best-selling first; and
- why the concept of "public opinion" doesn't actually exist.

Mathematics is the science of not being wrong, and it's based on common sense.

Convoluting mathematical formulas you encountered in school might have made your head spin. At the time, you might have asked yourself, “Will I ever use this in real life?”

The short answer is yes. Math is a key tool in solving common problems. We all use math every day, but we don't always call it “math.”

In essence, mathematics is the science of *not being wrong*.

Consider this example: During World War II, American planes returned from tours in Europe covered in bullet holes. Curiously, a plane's fuselage always had more bullet holes than did the engine.

To better protect the planes, military advisors suggested outfitting the fuselage with better armor. One young mathematician suggested instead improving the armor for the engine.

Why? He suspected that those planes that took shots to the engine were actually *those that didn't make it back*. If the engines were reinforced with better armor, more planes might survive.

There's a mathematical phenomenon known as *survivorship bias* underlying this situation. Survivorship bias is the logical error of concentrating on the things that “survived” some process. In this example, advisors concentrated incorrectly on the state of the planes that survived, overlooking the planes that didn't.

This example may not seem like a math problem, but it is. Math is about using reason to *not be wrong* about things.

Math is also based on common sense. Can you explain why adding seven stones to five stones is the same as adding five stones to seven stones? It's so obvious that it's difficult to actually explain.

Math is the reflection of things we already know intuitively. In this case, math reflects our intuition by defining addition as *commutative*: for any choice of a and b , $a + b = b + a$.

Even though we can't solve entire equations with our intuition, mathematics is *derived* from our common sense.

Linearity allows us to simplify mathematical problems.

A basic rule of mathematics is thus: if you have a difficult problem, try to make it easier and solve the easier problem instead. Then hope the simpler version is close enough to the difficult version.

We can break down hard problems into easier ones by assuming they have *linearity*. In geometry, we assume things have linearity when we consider straight lines. Curves represent *nonlinearity*.

Imagine an ant walking around a big circle. From the ant's perspective, it would feel like it was walking in a straight line. In fact, if we zoomed in closely enough on a part of the circle's curve, it would look like a straight line to us, too.

From this, we can infer that a circle's curve is similar to many straight lines bent at very slight angles.

Suppose you want to measure the circle's area. You can do that by putting a square in the center, so that each of the square's corners just touches the circle. The square's area is easy to calculate.

From there, we can insert more polygons with more corners into the remaining spaces and measure their areas until we can approximate the circle's area, using only straight lines.

This idea of linearity is widely used in statistics. Think of any mathematical relationship you might encounter in a news article. For example, that countries with more Burger Kings have looser morals, or that every extra \$10,000 you earn makes you 3 percent more likely to vote Republican. These are examples of *linear regression*.

Linear regression is based on linearity. In statistics, it's widely used for measuring how certain observations are related (like salary level and voting preference).

Again, the idea is to simplify the problem. Our research into the relationship of salary level and voting preference will give us many different data points, which we can then plot on a graph. What linear regression does is *not* find a way to connect every single dot or data point, but offer instead an approximation, represented as a straight line, of the trend of the data taken as a whole.

Drawing conclusions from observational data is questionable, but probability theory can help.

Scientists collect data through observation then use it to build *theories*. However, *observational data* can come about by chance, so drawing conclusions from it can be quite precarious.

Consider this example: In 2009, a neuroscientist showed photos of people to a dead fish and measured the fish's brain activity. Interestingly, the fish responded accurately to the emotions of people in the pictures.

This experiment was, of course, a gag. The scientist's real aim was to show how easily research findings can come about by chance.

Neuroscientists perform brain scans by dividing the scans into thousands of small pieces called *voxels*, which correspond to regions in the brain. When a brain is scanned (even a dead fish's brain), there's always some random "noise" in each voxel. As there are thousands of voxels, the odds of one producing data that corresponds to the stimulus given are actually quite high.

In more serious studies, however, it's not always clear if observational data arises by chance. Scientists thus use *probability theory* to address this issue.

Let's say you are a scientist, testing a new drug to see if it cures a certain illness. A mathematical tool for problems like this is called the *null hypothesis significance test*.

First, you start with your *null hypothesis*, which is an assumption of what will happen in your test. In this case, your null hypothesis is that the new drug does nothing at all.

Next, you take a look at the deviation of the data you observed in the experiment. You need to consider the probability that your data came about by chance – this is called the *p-value*. If the probability is less than a certain p-value (usually $p = 0.05$), the data is considered *statistically significant*.

And if your data is statistically significant, that means you know with 95 percent certainty that the new drug has the proposed effect.

Probability theory tells us what to expect from a bet, but we still have to consider the risks.

In any situation where we're uncertain about an outcome, probability can help us. Probability can't tell us exactly what will happen in the future, but it *can* tell us what we should expect.

For example, probability theory can tell us what's likely to happen if we place a bet on something. We can use probability to know what might happen when we buy a lottery ticket, by determining the ticket's *expected value*.

To calculate the expected value of a lottery ticket, we have to consider each possible outcome of the situation. We multiply the chance of each outcome by the ticket's value *given* that outcome. Then we add up the results.

Let's look at an example. Imagine a lottery with only two possible outcomes, losing or winning. A ticket costs \$1; there are 10 million tickets; the winning ticket is worth \$6 million. Here, the expected value of the ticket is 60 cents. That means that on average, we should expect a loss of 40 cents every time we play the lottery.

The same logic of expected value also applies when pricing stock options or life insurance.

However, expected value doesn't reflect the risk in a given bet. Consider this question: Would you rather receive \$50,000 or take part in a 50/50 bet, between losing \$100,000 and gaining \$200,000?

The expected value is the same in both cases, but if you lose the bet in the second option, it's much worse than doing nothing at all. The expected value hides the fact that the bet is a big risk if you can't easily spare \$100,000.

In the same way, a risky investment is only a good idea if you have enough money to cover the possible losses. It's very important to be aware of the risk in any bet.

The regression effect can be found everywhere, but often it isn't recognized.

Why is a novelist's second book usually not as good as his first breakout success? Artistic success is subject to a mathematical phenomenon called *regression to the mean*, or the *regression effect*.

The regression effect states that if a variable produces an unlikely outcome, the next outcome will tend to be closer to the mean.

Anything that involves randomness can be subject to the regression effect. For example, short people tend to have short children, and tall people tend to have tall children. But the children of *very* short and *very* tall parents aren't likely to be as short or as tall as their parents. Instead, their height is closer to the average.

This is because height is not completely determined by genes. It's affected by many other factors, such as eating habits, health and mere chance. There's no reason then that such external factors would line up precisely again as they did for those very tall or very short parents.

Regression, however, is often not recognized. For example, the *British Medical Journal* in 1976 published a research paper reporting that bran could regulate human digestion. If participants reported a fast digestion rate one day, eating bran slowed their digestion the next time it was measured, and vice versa.

These results are exactly what the regression effect predicts. If a person reports a fast digestion rate one day, they'll likely have a slower digestion rate the next. Thus the effect of bran might not be as remarkable after all.

Researchers have to be careful in situations like this as they could mistake the regression effect for a biological phenomenon, as did researchers measuring the effects of bran on digestion.

The same thing happens when a famous writer's second book isn't as well-crafted as his first, and literary critics often chalk it up to "exhaustion." More than anything, it's just mathematics.

Linear regression is useful, but assuming linearity when it isn't there can lead to false conclusions.

As we've seen, linear regression is an important statistical tool that helps us understand how variables relate to each other. However,

linear regression can't be used for every set of data; if used incorrectly, it produces misleading results.

You can find the linear regression of a data set by plotting all its points on a graph, then finding the line that comes closest to passing through all of them. This is only meaningful, however, if the data points are already in a generally linear shape.

For example, think of the curve-shaped path a missile follows when you fire it. If you zoom in on a short segment of it, the curve looks like a line. Thus, linear regression is very good for predicting where the missile will be a few seconds after a certain point.

But linear regression will fail to predict a missile's location after a longer time interval, because it doesn't take the curved path into account. As we zoom out on the missile's path, it stops being linear – that is, the whole path can't be described by just one line.

When linear regression is applied to a *nonlinear* phenomenon such as a missile's path, it produces incorrect results. This happened in 2008, when the journal *Obesity* published a paper claiming that all Americans would be overweight or obese by 2048.

The authors of the study determined this by plotting a graph of the percentage of obesity against time, then applied a linear regression. The graph's line crossed 100 percent at 2048.

But trends like obesity tend to produce curved graphs over a long period of time. In fact, obesity *can't* be plotted linearly, because if it could, *109 percent* of Americans would be obese by 2060!

So although linear regression is a critical tool, we must be careful to use it correctly.

Many research findings are wrong because of misused data or incorrect probability calculations.

In 2005, a professor named John Ioannidis published a paper called, “Why Most Published Research Findings are False.” That might seem like a radical claim, but his points were sound.

He first stated that insignificant observations can sometimes pass a statistical significance test by chance.

Consider the influence of a person's genetics on their probability of developing schizophrenia. It's almost certain that *some* genes are associated with schizophrenia, but it's unclear *which* genes.

Scientists thus might have to examine about 100,000 genes, while it's possible that only 10 are truly related to schizophrenia. However, with the most commonly used significance test of 95 percent, 5 percent of the genes will pass the test by chance. In this case, 5 percent means 5,000 genes. This isn't a specific enough result to really mean anything.

Ioannidis's second point was that studies that *don't* find successful test results often go unpublished, so those that *do* receive disproportionate attention.

Imagine that 20 labs test green jelly beans to see if they cause acne. Of the labs, 19 don't find a significant effect as part of their testing, so they don't write up the results. The one lab that does find a statistically significant result is much more likely to write a report and get published.

This isn't uncommon: experiments that go well by chance are often published while others that *don't* measure the same phenomenon remain obscure.

Ioannidis also argued that scientists sometimes tweak their results to make them statistically significant.

Imagine you run an experiment and get a result with a 94-percent certainty. It needs to have a minimum of 95-percent certainty to be considered statistically significant, so your results would be insignificant.

Since your significance test is so close, it would be possible to tweak the data to achieve a 95-percent certainty. Researchers often do this – not because they have bad intentions but because they truly believe their own hypotheses.

Polls and elections that make statements about “public opinion” are often incorrect.

How can you measure “public opinion”? It’s a rather vague term, so we should be critical when presented with polls that claim to express it.

First, people can have very contradictory opinions, and they can even contradict themselves.

For example, in January 2011, a CBS News poll reported that 77 percent of respondents thought cutting spending was the best way to address the federal budget deficit.

Only a month later, Pew Research conducted a poll where they asked about 13 different categories of government spending. In 11 of those categories, more people wanted to *increase* spending rather than to cut it.

Speaking of the “majority” can also be misleading.

A “majority rules” approach might seem fair, but it only really works when there are just two options. Any more than two and the groups can probably be divided in different ways, which can completely change the story.

For example, in an October 2010 poll, 52 percent of respondents said they opposed the U.S. Affordable Care Act, while only 41 percent supported it.

But breaking down the numbers differently made the story quite different. Only 37 percent of people wanted to repeal the health care reform bill. Another 10 percent said the law should be weakened, and 15 percent preferred to leave it as it was. Finally, 36 percent said the law should be *expanded* to change the health care system even more.

Thus many of the law's “opponents” actually *supported* its main idea.

This also happens in elections, and the 2000 U.S. presidential election is a good example. George Bush won 48.85 percent of the Florida vote, Al Gore 48.84 percent and Ralph Nader 1.6 percent; so Bush was the winner.

However, it's safe to say that nearly every person who voted for Nader would've preferred Gore over Bush. That means that a 51% majority probably preferred Gore over Bush, but those who voted for Nader had no way to express that.

Final summary

The key message in this book:

Mathematics is derived from common sense and it can teach us how to *not be wrong*. We can use mathematical ideas to explain things we see and draw accurate conclusions from them.

Uncovering the hidden math in our lives can help us in a range of situations, from understanding the newspaper better to knowing whether to buy life insurance.

Actionable advice:

Don't buy that lottery ticket.

Lotteries make their money by convincing people to pay for a ticket that's very unlikely to give them back any reward. Playing once might be fun, but don't expect it to be worth it if you play for an extended period of time. Calculating the expected value shows us that for most lotteries, you lose more money than you win.

Suggested further reading: *Naked Statistics* by Charles Wheelan

Naked Statistics offers an insightful introduction to statistics and explains the power of statistical analysis all while revealing its common pitfalls. It shows us how important a solid understanding of statistics is for good decision-making and gives the reader the tools to critically examine descriptive statistics.