Book Review: How to Expect the Unexpected: The Science of Making Predictions – and the Art of Knowing When Not to by Kit Yates

Mark Huber
Claremont McKenna College

Follow this and additional works at: https://scholarship.claremont.edu/jhm

Part of the Arts and Humanities Commons, Data Science Commons, and the Other Mathematics Commons

Recommended Citation

©2024 by the authors. This work is licensed under a Creative Commons License.
JHM is an open access bi-annual journal sponsored by the Claremont Center for the Mathematical Sciences and published by the Claremont Colleges Library | ISSN 2159-8118 | http://scholarship.claremont.edu/jhm/

The editorial staff of JHM works hard to make sure the scholarship disseminated in JHM is accurate and upholds professional ethical guidelines. However the views and opinions expressed in each published manuscript belong exclusively to the individual contributor(s). The publisher and the editors do not endorse or accept responsibility for them. See https://scholarship.claremont.edu/jhm/policies.html for more information.
Book Review: *How to Expect the Unexpected: The Science of Making Predictions—and the Art of Knowing When Not to* by Kit Yates

Mark Huber

*Department of Mathematical Sciences, Claremont McKenna College, California, USA*

*mhuber@cmc.edu*

---

**Synopsis**

Humans think about the future all the time. Prediction is a part of how we prepare for the coming of both good and bad events in our lives. Kit Yates' book, *How to Expect the Unexpected: The Science of Making Predictions—and the Art of Knowing When Not to*, concentrates primarily on the question of why prediction is difficult, and what mental shortcuts people take in prediction that can lead to incorrect results. Unfortunately, a lack of concern for details and several omissions undermine the quality of the book.

---


Being able to predict future events is of great importance when making decisions, but there are many ways that a prediction can go wrong. For example, to speed up the prediction process, we often use mental shortcuts. These shortcuts, also known as cognitive biases, can easily lead us to incorrect predictions under the wrong circumstances.

Another problem arises when the goal of prediction is to obtain a favorable outcome or resource at the expense of another person. In this case, you also have to deal with the actions of that other person which are usually designed to benefit themselves rather than you.
Yet another issue is that even the tiniest error in starting conditions for a physical system will lead to uncertainty down the road. This is why casinos are not afraid that an observer will be able to see the initial spin of a roulette wheel and compute the outcome before the ball lands in a slot. Even worse, many systems are nonlinear, and quickly spiral past the point of effective prediction.

These three ways a prediction can go wrong form the core topics of Kit Yates’ book *How to Expect the Unexpected: The Science of Making Predictions—and the Art of Knowing When Not to*. Four chapters are devoted to cognitive biases and their variants, one is devoted to dealing with opposing players, and four chapters deal with the difficulty of prediction in nonlinear systems.

**The up side**

This book is very easy to read, and the author takes great pains to explain and illustrate certain ideas. For instance, in the section entitled “How can you be certain?” Yates covers the pigeonhole principle that if you have more pigeons than homes for the pigeons (pigeonholes), then at least one pigeonhole must contain more than one pigeon. The explanation is given from several angles and through analogies that make it easy to understand what is going on.

The first few chapters go on to introduce several common biases that people might not even be aware exist. I would say most people are aware that superstition and rolling dice are not great at prediction, but others might not be aware of how earthquake energies are predicted, or how important it is to distinguish between linear and nonlinear phenomena.

Many of the ideas in the book are illustrated through graphs that are clean and easy to follow. So what’s not to like?

**Envisioning what did not happen**

Sadly, there are some major weaknesses in the work. One is the overuse of anecdotes rather than scientific studies in illustrating ideas. In a scientific study, it is important to consider *all* possible approaches to a problem instead of just one. The famous need for a control in experiments is an example of this principle at work. One cannot simply apply the treatment and claim that it is the best approach without comparison to other methods.
An example where this procedure is not followed occurs in the section “The car’s the star”. Here Yates repeats a story told in Predictioneer by Bruce Bueno de Mesquita [2]. In the story, Bueno de Mesquita tells the reader that he calls various car dealerships that have a particular car available. He asks each dealer to quote their best price without negotiation and then selects the one with the lowest price.

Yates writes of the method: “...the buyer benefits from the best possible price.” Unfortunately, there is nothing in the story to indicate that the method has been tested to see if indeed it leads to the best price. Would the buyer have done better in a negotiation if they had gone to a dealer and threatened to leave? Would they have done better if they had simply lied and claimed that a rival dealership had offered the car at a lower price? Would a dealer have undercut another dealer’s price if they had been called back in a second round?

The unquestioning acceptance that this is the “best possible price” is unfortunately neither mathematical nor scientific, and similarly unquestioned assumptions pop up often throughout the text.

In another instance, on page 13, Yates writes: “Many of the early lifeboats that launched [from the sinking Titanic] were not filled to capacity...because people were hesitant to come forwards when called.” Is this true? Perhaps. But actually, at least one officer loading the lifeboats (Commander Lightholler) was only loading women and children and not allowing men to board unless they could assist with the operation of the boat once at sea [4]. Another explanation commonly given is that crew members believed that loading the lifeboats to capacity at the top deck would snap them when lowered, with most not knowing that a metal reinforcement had been added to prevent this exact problem from occurring. A lifeboat training drill meant to occur that morning had been canceled [3].

I do not wish to imply that there are no scientific studies anywhere in the text. The book has 185 references, and many are to scientific studies. However, many of the stories and examples such as the Titanic story are not given references at all, and the overall reliance on untested hypotheses and intuition does weaken the book.
Mathematics for a general audience

A decision was clearly made when writing the book to avoid as much mathematics and mathematical notation as possible. For instance, probabilities, central players of the mathematics of uncertainty, are never defined in the text. This becomes problematic in the Game Theory chapter, for example, when terms like mixed strategy and expected value are introduced, but which cannot be defined precisely.

Another example appears in the section entitled “Looking out for number one” about Benford’s law. This result describes the frequency of leading digits in certain types of data. Without mathematics, it is very difficult to describe when Benford’s law holds. In fact, it holds when the logarithm of the set of numbers is uniform. This type of distribution shows up in data about incomes, house prices, populations, and other data where the numbers are growing (or shrinking) randomly by constant factors.

Because this definition is not given, Yates turns instead to examples. Do phone numbers obey it? No. Do house numbers? Yes. These are stated in the book just as I did here, without explanation. Without mathematics, what exactly Benford’s law is remains a mystery.\footnote{For more on Benford’s law and a neat example of it to music analysis, see [1].}

This avoidance of mathematics rears its head again in the section about one of the most important theorems in probability, Bayes’ Rule. This theorem says that the probability of event $A$ given that event $B$ occurs is the probability that event $B$ occurs given $A$ occurs times the probability of event $A$, divided by the probability of event $B$. Such is the avoidance of mathematics in the book that this relatively simple equation is never stated. Instead, attempts are made to illustrate solely through examples.

Statistics

Two of the most important tools for prediction are statistics and machine learning. Neither is well represented in the text. For example, in Chapter 6, Yates details the work of QuocTrung Bui from 2014, looking at pizza prices versus size. He presents a plot of price versus pizza diameter and then fits a least squares line through it. The method by which this line is obtained
is not described at all; it is only referred to as a “best fit”. Presumably, it is a least squares line, intended to minimize the sum of the squares of the resulting error, but such a best fit line might also minimize the sum of the absolute values of the errors.

Yates then claims that this relationship is approximately linear, with the caveat that he might be “preconditioned to impose this relationship”. But there was an easy way to check if this was the case: simply plot price versus pizza area and compare the resulting errors. This is not done, which is somewhat baffling. Unfortunately, the data used is not referenced, so it is impossible to do the experiment.

Again, this is frustrating, because in the twenty-first century, linear regression is a major tool used to make predictions and understand an uncertain world. Linear regression could also be used to fit a curve through the data if it was growing as the square of the diameter of the pizza as well. The fact that this is not mentioned or discussed further leaves the reader with the impression that linear regression is not possible when the measured response is nonlinear in the predictor variables. This is just false.

Dealing with unexpected events: Only a partial analysis

There are entire disciplines devoted to prediction and working with uncertain data to make predictions. But outside of psychological biases that affect decision making, this text just seems uninterested in exploring these disciplines and their current work. In short, this book begins with a promising premise, but the particular choices made by the author leave too much out if you are looking for a comprehensive exploration. Still, those interested in the psychological aspects of decision-making will find some interesting bits here and there.

References

