FLUKE

THE MATH & MYTH OF COINCIDENCE

JOSEPH MAZUR

BASIC BOOKS, NEW YORK 2016

REFEREE REPORT

1. Recommendation

This reviewer strongly recommends this book to the serious mathematics and science buff, as well as the professional mathematician or statistician who wishes to surround himself with a strong set of examples illustrating the theories of low probability high consequence events. Such events are the subject of vigorous scientific inquiry in this time owing to the devastating effects of natural and economic disasters, as well as the destructive consequences of war and terrorism, among others. Additionally, the author introduces examples from literature, folklore, and fantasy to illustrate further the pervasive influences of thought in this realm. Observe that the title 'Fluke' refers to the occurrence of a single such event, whereas 'coincidence,' as developed in the text, refers to two such events occurring together either in space or time (or both,) thus introducing the concepts of joint probability distributions. To describe this work as a *magnum opus* for Prof. Mazur could include a modicum of hyperbole, but maybe not. This is a very fine study in the mode of making rigorous mathematical material available to the lay reader.

2. Introduction

This review proceeds according to a plan. The following section, Synopsis, relates the basic structure and content of the book, which is divided into four Parts, containing together 15 Chapters. The focus is on seven of these Chapters, with a salient feature in each developed. In this manner, the reviewer hopes to convey the essence of the book, its goals, development, and achievements, to offer an appreciation of the whole.

The subsequent section, Review, provides an analysis of the work in terms of reaching its intended audience with enlightning material, and opinion in the form of suggested changes and improvements to this already fine work. This reviewer almost never offers changes and improvements, only commentary on the effectiveness of the presentation. In the extant work, however, it is incumbent additionally to sally forth with such material — perhaps for a second edition? — with the thought that the author could reach a significantly wider audience; he can achieve this end with an upgrading of the formal probability setup and the inclusion of a few additional on-point and well understood examples. Keep in mind

---

Date: 31 August 2016.

This referee report was compiled in $\mathcal{AMS}$-LaTeX with Palatino and Euler Virtual Math fonts, both by Hermann Zapf, using the University of Chicago Press bibliography and citation [author-date] style as prepared by BibTeX.

that while there is a plethora of popular publications and periodicals addressing the lay audience for nigh all of the sciences, there is only a paucity of such popular offerings for the lay audience in mathematics — save for the journal in which this review appears, and its very few peers. To address this audience is a challenge for any author before a single word appears on the page.

The final section, Conclusion, looks to the aims, processes, and targets of this work, and to its achievements and successes.

3. Synopsis

3.1. **The structure.** Prof. Mazur delivers an exceptionally fine treatise on the subjects of the title, meticulously cited and well prepared for the lay reader, his primary audience. The book is presented in four Parts, including Chapters numbered consecutively throughout.

- Part 1: The Stories
- Part 2: The Mathematics
- Part 3: The Analysis
- Part 4: The Head-Scratchers

The reviewer shall address each Part *ad seriatim* by discussing an excerpt from some of the included Chapters, enumerated here.

Part 1: The Stories

(1) Exceptional Moments
(2) *The Girl from Petrovka* and Other Benign Coincidences
(3) Meaningful Coincidence

Part 2: The Mathematics

(4) What Are the Chances?
(5) Bernoulli's Gift
(6) Long Strings of Heads
(7) Pascal's Triangle
(8) The Problem with Monkeys

Part 3: The Analysis

(9) Enormity of the World
(10) The Stories of Chapter 2 Revisited

Part 4: The Head-Scratchers

(11) Evidence
(12) Discovery
(13) Risk
(14) Psychic Power
(15) Sir Gawain and the Green Knight

As well, appear Epilogue, Notes, and Acknowledgments, plus an Index. As noted, the reviewer shall herein recount an item in summary from a few Chapters to give the flavor of these offerings.

3.2. **The content.** In Chapter 1, Prof. Mazur relates a chance meeting with his brother on the Greek island of Crete, neither expecting the other to be there. He wonders what could have caused both of them to be in the same place at the same time, a seemingly very improbable event. He speaks of the wider concept of causality in Western philosophy, then refers obliquely to the Heisenberg Uncertainty Principle of Quantum Mechanics which recognizes that some events cannot be observed except for a chance element. The conclusion is that the idea of a deterministic universe, wherein every event has a definite antecedent and consequent, must be reevaluated in the context of probability. This was a significant development in the history of human thought, as much so as Copernicus' discovery of our heliocentric solar system.

In Chapter 2, Prof. Mazur writes of the actor Anthony Hopkins finding a copy of the book The Girl from Petrovka on a park bench near the London Underground Station at Leicester Square. He was looking for the book, as he had recently been cast for the movie version, and he wanted to read the original. There is nothing unusual about that, except that the book he found contained an inscription by the author, George Feifer. It was his book, lost by him some time before. So, speculates Prof. Mazur, "What are the chances?"

In Chapter 5, Prof. Mazur delves into the mathematics of the Law of Large Numbers, and by inference the Central Limit Theorem [CLT]. He gives examples, including the independence of individual rolls of true dice, in the process rightly debunking the so-called 'law of averages,' which supposes that sequential flips of a fair coin incorporate a compensating effect; that is to say, that if a preponderance of 'heads' were to be seen over a period of time, then a preponderance of 'tails' should follow. This concept denies the assumed independence of events. This reviewer is not the first to say, but coin flipping exhibits more of a 'swamping' effect, not a 'compensating' effect.

In this Chapter we begin to see actual mathematical expressions. The first is this (with technical textual modifications.)

$$\Pr_{N\to\infty}\left[\left|\frac{k}{N}-p\right|<\varepsilon\right]=1,$$

which says that for any experiment with probability p of having a specified result, the number of these results k as a fraction of the number of independent trials N tends toward p to any desired degree of accuracy — within the amount ε, however small, yet positive — as N increases without bound. This is the Law of Large Numbers, in the mathematician's symbolic notation.

Later in the Chapter appears the CLT in the form of the development of the normal probability density function by way of the Galton* Board, an experiment by means of which balls are allowed to drop through a triangular array of pins, with equal probability of going either way upon hitting a pin. This is the ubiquitous 'bell shaped curve.' The all-important CLT states that the average of identically distributed random variables with finite variance, as the number of variables increases without bound, has in the limit the normal distribution. As the Galton Board experiment exhibits finite variance, the CLT applies.

Finally, in this Chapter, the author relates the concept of mathematical expectation. The late great Georgian (Soviet) mathematician Andrei Kolmogorov, widely regarded as the greatest probabilist of his time in the Eastern world,† related that his single most important contribution to the theory of probability, of which there were copious many, was the concept of conditional expectation, an expanded idea including the basic mathematical expectation noted in this presently reviewed work.

*Sir Francis Galton, creator of the statistical concepts of correlation and regression toward the mean. He followed on the analysis of least squares, created by Adrien-Marie Legendre.

†The late Willy Feller, Professor of Mathematics at Princeton University, earned this accolade in the West.

In Chapter 8, Prof. Mazur writes to the familiar 'birthday problem,' wherein if 23 or more people are in a room the probability is greater than one-half that two of them share the same birthday. This result comes with some assumptions, among them the equal independent probability that a birth occurs uniformly over the 365 days (ignoring leap year days.) He also introduces the oft cited *gedanken* [thought, *Ger.*] experiment of the monkey sitting at a typewriter. If the monkey types random letters and punctuation characters, each with any positive probability whatever, in due course he will write the first line of a Shakespearian sonnet. Prof. Mazur could have gone further to say the entire works of William Shakespeare, or even the entire collection of works in the English language. It is all the same. Given a sufficient finite time any of these tasks will be accomplished with probability 1.

In Chapter 9, Prof. Mazur tells of the diffusion of a drop of ink in a bottle of water. In a short time the water becomes a uniform dyed color. Is this an inevitable consequence? Well, no. The physics of the situation says that the molecules of the dye move about independently with a Brownian motion,[‡] so they could move first to approximate the uniform distribution described, but then they all could move back to their original places. Is this likely? Well, no, but possible? Yes.

In Chapter 13, Prof. Mazur relates some facts and analysis of financial market disruption, including the case of trading by Nick Leeson at London's Barings Bank, who lost the fortune of £850 million as a consequence of the January 17, 1995,[§] Kobe earthquake, a fluke. Such occurrences, and there have been others, wreak havoc on financial markets, with widespread consequences to the general economy. Fortunately, thus far, such disasters have been few and far between, but no believable theory says that worse impacts are impossible, however sooner or later they may come.

It is noteworthy to mention that numerous models of stock price behavior exist, starting with, and even antedating, the Black–Scholes model (Black and Scholes 1973), which assumed that the logarithm of prices moves like a Brownian motion, *v.s.*, comment for Chapter 9. This 'random walk' formulation has finite variance, thus the CLT applies. However, empirical evidence strongly suggests that the distribution of price movements is not normal, but has so-called 'fat tails.' An alternative formulation, which has fat tails, yet finite variance, is the 'normal inverse Gaussian' distribution, perhaps the most used modern description of price movement. However, some analysts believe the correct description is a representative or the α-stable class, which has infinite (undefined) variance. The jury is still out on whether the proper description has finite variance or not. The scarcity of tail evidence, even with the thousands of traded stocks across the world, is the controlling factor. Tail events are so rare that they are properly called flukes, yet they do occur, and are surprise observances when they do.

In Chapter 14, Prof. Mazur investigates extrasensory perception [ESP] and mental telepathy, from the point of view of the odds of such experiences occurring, a way of forming hypothesis testing by example. Psychologists seriously conduct experiments to see if ESP works, and under what circumstances, thus far without scientifically verifiable results. What if telepathic communication can occur, however, even if we yet do not understand the mechanism? Perhaps a message comes only rarely, a fluke, but carries immense consequences. Then properly it cannot be ignored.

[‡]Wiener process, in its mathematical form, after the late Massachusetts Institute of Technology professor of mathematics Norbert Wiener.

[§]Japan Standard Time

4. Review

4.1. **Expanded formal development.** The only serious criticism to the author's approach lies in this reviewer's belief that the lay reader actually could absorb an additional degree of formality, and that such setup would enhance the reader's understanding of the principles set forth. For instance, the author could have introduced early in the text the concept of the probability triple $\{\Omega, \Sigma, \mathsf{P}\}$, where Ω is a *sample space of events*, Σ is a *sigma algebra (field)* of *sets of events*, and P is a *probability measure* assigned to those sets. He need not then have had to define the italicized concepts, but could have illustrated them by example, as this.

Consider rolling a die. An *event* is the appearance of a number of dots on the top face. The *sample space* consists of all such possible events, six of them. The *sets of events* consist of all $64 = 2^6$ possible combinations of events, based on whether any given face is either in or out of a set. Note that the null set (no faces) and the entire set (all faces) are included. The *sigma algebra* includes these sets along with the action of union of two sets — the inclusion of all elements in each set — and the action of intersection of two sets — the inclusion of only those elements in both sets.

Note that the resulting set from either union or intersection is one of the original 64 sets, a requirement for a system to be called a *sigma algebra*. The *probability measure* is an assignment to the sets of events, and in the case of a fair die is $1/6$ times the number of elements in a given set. By convention a probability measure is zero for the null set, is non-negative on each set of events, is the sum of measures for the union of disjoint sets (those not having a common element,) and is 1 for the entire set. With these concepts in hand for a discrete sample space, like the events of the example, it is straightforward, though not trivial, to extend the concepts to continuous sample spaces, like the unit interval $[0, 1]$ of the real line.

As well, probability concepts follow directly, again by example, like density and distribution function — and their relationship,[¶] the usefulness of a density by integration to define a probability measure for the probability triple, ideas relating to random variables, and the development of expectation, ordinary and conditional. The practice of using correct terms leads to clarity of thought, and better appreciation of concepts. For instance, only random variables have expectations. So, one might ask, what is the random variable on the sample space of the real line for which one calculates expectations given a probability measure? Well, it is the identity random variable, a straight sloping line on a graph. So, frequently, even in the case of professional statisticians, people think of an expectation as somehow a calculation on a density. Well, it is, but only indirectly by way of reference to the probability measure implied by the density. Further, it is true, that some distributions do not have densities, but do have expectations, which can be infinite (stated as undefined.)

In words above the reviewer mentioned, "[The author introduces] by inference the Central Limit Theorem" and, "[He] then refers obliquely to the Heisenberg Uncertainty Principle." In actuality, the author does not refer to either of these concepts by name. As each very important in its own realm, it would be nice to see some further formal development of them. An expanded development of probability fundamentals would have allow such exposition.

Ideas such as these foundational precepts of probability theory by this reviewer's experience are within the grasp of the serious reader with some grounding in basic mathematics

---

[¶]The former is the derivative of the latter, while the latter is the integral of the former, a relationship defined by the Fundamental Theorem of the Calculus.

and science. Illustrated with interesting fleshed out examples, as is the author's wont, the principles come alive, and allow the reader to achieve that magical 'Eureka' experience.

4.2. **Further examples.** With a plethora of examples, analyzed in detail, one would think that these are reasonably exhaustive of the possibilities. However, this reviewer thinks that four further examples on point would have provided strong immediate impact to the reader on the likely occurrence of rare events. One comes from of Ronald Alymer Fisher, cited extensively by the author, another comes from the late Princeton Professor of Philosophy Carl Gustav Hempel, the third looks into cumulative totals in coin flipping, and the fourth involves a comparison between recurrent and transient events.

Let us begin with Fisher's illustration. Not infrequently one of Fisher's graduate students in statistics would opine (to the inventor of the 5% confidence interval) that 5% is such a small fraction that one could safely ignore events with a lower probability of occurrence. To this, Fisher replied to the student, with others around the study table listening, to consider his father, then his father's father, and so forth back 100 generations. This takes the student's thought about back to the time of King Solomon. Then, Fisher continued, consider the probability that such ancestor would beget a line of 100 generations with a male member in each. Well, Fisher, the noted geneticist as well as statistician, estimated this probability as 10^{-44}. Then, Fisher would say [in paraphrase,] "Each of you sitting at this table, in fact everyone alive, has an ancestor for which this event has actually occurred."

On to Hempel's Paradox, which addresses the question of whether all ravens are black, specifically to confirming instances of this hypothesis. For instance, if one sees a raven and it is black, then is one justified in thinking that this is evidence that all ravens are black? What if this person sees another black raven, and another? How many would it take to convince oneself of the truth of the hypothesis? Well, let's look at this another way. Consider the contrapositive statement, equivalent to the original hypothesis: "All not black animals are not ravens." (We are familiar with the contrapositive in this form. If it is raining I shall take my umbrella. Equivalently, if I do not take my umbrella, it is not raining.)

Continuing, ask oneself now if a confirming instance of the contrapositive should be equally compelling in justifying the original hypothesis. It should be, wouldn't one think? Confirming instances of equivalent statements should in some logical sense carry similar weight. How now brown cow? If one sees an brown cow (a not black animal) do we feel better about the idea that all ravens are black? Most people do not, but why not?

The answer lies in the apparent infinitude of the set of not black animals. Because, however one classifies animals (including all the insects?), their number is finite. So, if one could exhaustively examine every not black animal and determine that it is not a raven, then, of course, all ravens are black. So, one needs a lot of confirming instances — millions, billions, trillions, more? — to conclude that all ravens are black. Just looking at a few won't do it.

For coin flipping experiments, addressed in Chapter 5, it would be instructive to introduce the First Arc Sine Law [1ASL] because it produces a counterintuitive result for many people, thereby classifying it as a surprise, or fluke. The result is that for a fair coin the running cumulative total of heads [tails] is most likely to have exceeded the running cumulative total of tails [heads] for the entire time of the trial, and that the least likely outcome is that half of the time the cumulative heads [tails] are in the lead. A good place to read for a comprehensive analysis, still good after all these years, is (Feller 1957, Chapter III, Section 5, pp. 77–81, Probability of Long Leads: The First Arc Sine Law). The late author was Eugene Higgins Professor of Mathematics, Princeton University, the person footnoted in the comment on Chapter 9. For a modern analysis see (Sternberg 2009, Lecture slides, Lecture 8, Mathematics 118 — Dynamical Systems.) The author is George Putnam Professor of Pure and Applied Mathematics, Harvard University.

This result is readily extensible to contests of all kinds with evenly matched opponents — sporting events with two sides (individuals or teams,) chess matches, political contests during campaigns. The inference from the 1ASL is that the most likely scenario is that one side is ahead in the score throughout the contest. The changing back and forth of the lead is the least likely outcome. And, as is logical, if one of the contestants is innately superior than the other, then the likelihood of that contestant getting ahead and staying ahead is the least surprising outcome. Many people feel this Corollary to the 1ASL intuitively, if not the consequences for the evenly matched case.

Lastly, consider recurrent and transient events. In one and two dimensions, random walks (Brownian motions) are recurrent, meaning if starting at the origin a path with probability 1 will return within a small ε-radius of the origin an infinity of times. We say, colloquially, that a 'drunk will always find his way home.' In three and higher dimensions, random walks are transient, meaning if starting at the origin a path with probability 1 will never return within a small ε-radius of the origin, once having left it. We say, colloquially that 'a bird will never find his way home.' Of course, birds have other senses which help them to find their way, so their paths are not random in the sense of Brownian motion. Note that an event of probability 0, like a path in three dimensions crossing the origin, is not impossible, just infinitesimally rare, a fluke, one might say.

5. Conclusion

The author addressed the lay audience of mathematics and science buff, and beyond, the curious professional who may not have been exposed to the wonderful abundance of examples for which the element of surprise, even disbelief, prevails. We speak of events occurring 'almost surely,' meaning with probability 1. This doesn't mean that other events cannot occur; they just have probability 0. To see this point consider this *gedanken* experiment. Throw a dart at the unit interval. Assume it hits $\pi/4$. What is the probability?[‖] Well, 0. However, the dart has to hit some point, and they all have probability 0. We know collectively they have probability 1 — the entire unit interval. How do all those 0's add to 1? Well, they don't, if you count them, proving, among other things, that the unit interval is uncountable.

The highest praise this reviewer can give to Prof. Mazur's book is that it makes one think. The subtitle could have been, "How rare is rare?", or its cohort, "How big is big?". Is there a smallest number greater than 0? No. If you think you have it, consider the number half that size. Consider a googol to the googol power, a googol number of times. Is that the biggest number. No. Just think of the number twice that size. On the subject of googols, it is known that the number of primes is infinite. So, if one were to hypothesize that the number of primes is finite, and all are less than a googol, then the first counterexample is larger than a googol.[**] Keep this in mind when thinking that a large number of confirming instances of an hypothesis could prove a theorem. There's that raven again!

Prof. Mazur aims high, and hits the mark. The achievement here is that this book about the mathematics of rare events is a page turner. Who would have thought? Not this reviewer. Do it again, Prof. Mazur. We are all waiting.

[‖] given the Borel sigma algebra

[**] A googol itself is not prime, for it equals $10^{100} = 2^{100} \cdot 5^{100}$, with 200 prime factors.

References

Black, F. and M. S. Scholes (1973, May–Jun.). The pricing of options and corporate liabilities. *J. Political Econ. 81*(3), 637–654.

Feller, W. (1957). *An Introduction to Probability Theory and Its Applications* (2nd ed.), Volume I. John Wiley & Sons, Inc.

Sternberg, S. (2009, March). The Arc Sine Law. Random Walk. The Ballot Theorem. Available at `http://www.math.harvard.edu/library/sternberg/slides/1180908.pdf`.