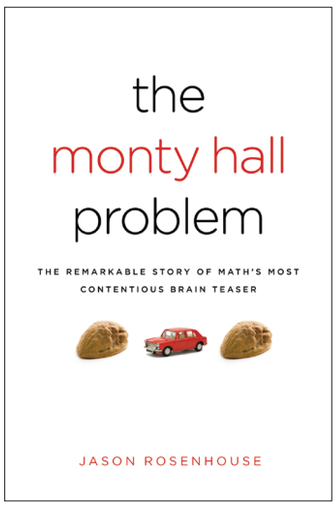
The Monty Hall Problem: The Remarkable Story of Math's Most Contentious Brain Teaser



Jason Rosenhouse; Oxford University Press, USA; 208 pp.; May 2009; $24.95, £15.99; ISBN 9780195367898

You are a contestant on a television game show. You pick one of three doors, knowing that only one of the doors conceals a prize. The host then opens one of the other two doors that he knows does not conceal the prize, picking one at random if he has a choice. You are then given the option of either sticking with your initial choice, or switching to the other remaining door. What should you do?

This is the "Monty Hall problem", named after the Canadian presenter who made it famous on a TV game show of the 1950s. Simple to state, but, it seems, difficult to answer.

It will be familiar to most people who have studied probability, and, given a modicum of probability theory, it is not a difficult problem. Does it really warrant a whole book?

It is a credit to Jason Rosenhouse that he has been able to write such a comprehensive account of this elementary problem, and the many alternative guises in which it appears, in such a compelling and approachable way. To many people, the answer is counter-intuitive, and this is what makes the problem appealing.

Rosenhouse argues that an entire undergraduate probability course can be taught using variants of the Monty Hall problem as the only examples. While this may be stretching the point a little, it does demonstrate the surprisingly deep concepts that the problem can reveal.

The book begins with a history of the problem, and a quick summary of other similar "chestnuts", such as the birthday problem, that often appear in basic probability texts. Quickly, it turns to solving the classical Monty Hall problem stated above, introducing along the way the ideas of conditional probability and Bayes's theorem. Formal proofs are given alongside explanations to demonstrate why faulty reasoning might lead the contestant to the wrong conclusion.

Then a comprehensive array of variants of the classical problem are introduced: several doors, several prizes, several contestants, several hosts, hosts who act benevolently or malevolently, hosts who prefer to open some doors rather than others, and so on, each twist introducing an extra layer of complexity than the one before. They all serve as reminders that in any probabilistic or statistical analysis, understanding the context of the problem is vital in order to determine a solution. Closing chapters with the titles "Cognitive Monty" and "Philosophical Monty" explain why it is that so many people, statisticians included, find the problem difficult.

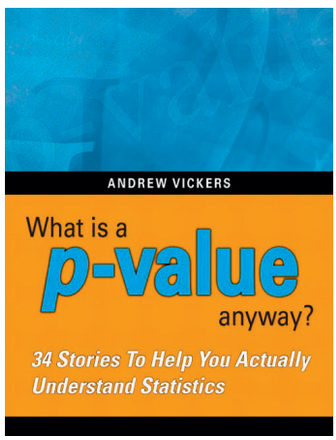
A problem that is understandable by all can demonstrate key concepts in probability

Much of the reasoning does require familiarity with mathematical notation, at a level that perhaps might appear in a first-year undergraduate course. Only once does the overall story become bogged down with technical details, when a particularly obtuse variant of the problem that involves opening a number of doors in sequence ("Progressive Monty") requires the use of recurrence relations.

Overall, this book is an excellent example of how a problem that is understandable by all can be used to introduce key concepts in mathematics and probability. If you are already familiar with the problem, this book will make you think more deeply about the nature of chance, and what Rosenhouse describes as "the perils of intuition". If Monty Hall is new to you, then you have a choice: stick or switch? You may be surprised.

Tom Fanshawe, Lancaster

What is a *p*-value Anyway? 34 Stories to Help You Actually Understand Statistics



Andrew J. Vickers; Pearson; 240 pp.; November 2009; $33.33, £28.99; ISBN 9780321629302

Finally a book that explains statistics without the ostentatious jargon and Latin-filled formulae. For too long statistics has been classed as tedious and boring, but with this book Andrew Vickers, a biostatistician based in New York, redefines the field of "popular statistics" and shows us that statistics is much more fun and interesting than we might have thought.

Vickers sets out on a mission to make readers statistically literate.

He is not interested in explaining how to use statistics. Rather, he wants to explore what statistics means

He is not interested in explaining how to use statistics. Rather, he wants to explore what statistics means and educate us in the process. He wants us to grasp basic statistical concepts such as inference, estimation and mean, median and standard deviation, slowly moving on to more challenging concepts as the book progresses.

Each chapter focuses on a particular statistical concept. Vickers typically begins a chapter with a short anecdote or a particularly hilarious story. He then masterfully frames the statistical concept he wants to inculcate into the anecdote or story to paint a simple but crystal-clear picture of what he is explaining. And he seasons all this with some nerdy jokes ("when Bill Gates walks into a bar … the average salary goes up"). By relying on basic stories, Vickers succeeds in translating statistical concepts into something more tangible, relatable and thus graspable.

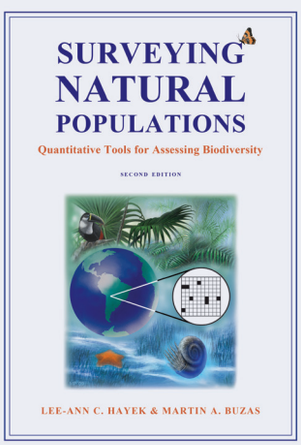
Chapters close with more typical academic point-form summaries of concepts explored (or illustrated, rather) as well as a series of thoughtful discussion questions to ponder upon. Vickers answers each discussion question at length at the end of the book and there are some gems to be unearthed there: "why might a stronger effect lead to a higher *p*-value, and less evidence

against a null hypothesis of no effect?", "what is the connection between a criminal trial and a *p*-value?" (think guilty or not guilty).

This is an ideal book for educators who want to inspire their students to take an interest in statistics and for anyone who wants to gain an understanding of it. Vickers has managed what many thought was impossible: he has single-handedly made statistics sexy, interesting and even fun. For this alone, this book deserves the highest recognition.

Khalil A. Cassimally, Mauritius

Surveying Natural Populations: Quantitative Tools for Assessing Biodiversity



Lee-Ann C. Hayek and Martin A. Buzas; Columbia University Press; 616 pp.; $85.00, £58.50; ISBN 9780231146203

When the postman handed me this book with a relieved look in his eye, I knew I could say one thing about it without even seeing it: this is not a book you would want in your rucksack in the field. It is big and it is heavy. Not that this is a problem. *Surveying Natural Populations* is far from being a book of brief notes, useful for quickly finding a diversity index to plug some numbers into without knowing how it works. If that is what you need, this is definitely not the book for you. Rather this is a book for field biologists and palaeontologists who want much more than a passing understanding of the statistical details behind the methods used to quantify diversity, and from the very beginning of the introduction it is clear just how well it hits its target readership. If anything, it hits it too well. An unusually thorough attempt has been made to ensure every part of the content is related to situations familiar to field workers, while potential points of confusion are addressed head on with straightforward frankness. As an example of this, hard-core field biologists often use the term "sample" very differently from how it is used by statisticians and this can result in unnecessary confusion. Hayek and Buzas clear up this particular linguistic flash-point early in the book and reinforce this and other distinctions frequently, so there is little danger of this book generating needless bewilderment.

> The authors assume almost no prior knowledge; their skill is in gradually building a quantitative approach to biodiversity

The authors assume almost no prior knowledge on the part of the reader, which means that, in reviewing the book as a statistician, the skill shown by the authors in gradually building a quantitative approach to biodiversity is especially evident. The approach taken here seems exceptionally conscious of its purpose to set firm foundations. New concepts are always introduced with reference to something familiar. Frequency distributions, for example, are introduced in the context of tabulating and graphing real data, making this a logical connection for anyone used to collecting field data but less used to analysing it. The use of a single, illustrative data set throughout the book (the distribution of tree species in a patch of Bolivian forest) helps particularly in this context. The decision to use only one example set of data has a cost in that it restricts coverage to only those methods relevant for the type of data in the illustrative set (there is no mention of time series, for example) but on balance I feel this is a price worth paying and the data set used has been chosen very consciously to represent the kind of observations field workers are likely to deal with much of the time.

Despite starting gently, this is by no means just an introductory text. The authors seem to have made the unusual and rather brave decision to include material based on utility rather than difficultly. This is to be applauded. If you were a statistically neophyte field biologist motivated to pick up this particular book in the first place, you would want to be told about the precise workings of Poisson and normal approximations to the binomial, for example, rather than given half an explanation and, at best, some rules of thumb. The authors have done an outstanding job in introducing these concepts in a gentle and relevant way.

Although assuming no prior knowledge at the start, this is far from being simply an introductory text. In fact, it is almost two books in one. The first 11 chapters provide a basic introduction to quantitative analysis in general, set very firmly within the field of biodiversity assessment, and with such a gentle incorporation of relevant aspects of mathematical statistical theory that most readers would not know they were being exposed to ideas generally considered to be too advanced for the starting quantitative biologist. The remainder of the book is a detailed and advanced treatment of biodiversity assessment with at its heart a characteristically clear and comprehensive introduction to *SHE* analysis (an information-theoretic approach pioneered by the authors combining *S*pecies richness, the diversity measure *H* and *E*veness). This is enormously valuable, and such a clear and thorough treatment of this way of approaching quantitative biodiversity, straight from the people who know it best, would make very useful reading even for those who might want to skip the previous chapters.

Despite my overall positive feelings, there are some odd things about this book. It is strangely old fashioned at times. Almost the last quarter consists of tables and the data sets. This makes the book heavier than it needs to be, and means that anyone wanting to use the data would have to type it into a data file. It seems odd that the authors or publishers do not offer a downloadable version of the data. There are also some notable omissions in the content. There is not a single mention of Bayesian techniques, for example, despite these beginning to have a significant impact on the models used to estimate species occurrence and diversity. However, none of this takes away from what the book does so well, which is to develop a thorough understanding of the workings of quantitative biodiversity assessment, albeit within prescribed limits, from a standing start up to one area of advanced technique (*SHE* analysis) with an emphasis on maximum relevance and utility for its target readership. Having worked through it, it will then be a book to keep on the shelf for reference, not one to pack for that extended field season bivouacking in the rainforest.

> They have made the unusual and brave decision to use material based on utility rather than difficulty

Paul Craze, London