

Probabilities

The Little Numbers that Rule Our Lives

Peter Olofsson

*Department of Mathematics
Tulane University
New Orleans, Louisiana*



WILEY-INTERSCIENCE

A John Wiley & Sons, Inc., Publication

This Page Intentionally Left Blank

Probabilities



THE WILEY BICENTENNIAL—KNOWLEDGE FOR GENERATIONS

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

WILLIAM J. PESCE
PRESIDENT AND CHIEF EXECUTIVE OFFICER

PETER BOOTH WILEY
CHAIRMAN OF THE BOARD

Probabilities

The Little Numbers that Rule Our Lives

Peter Olofsson

*Department of Mathematics
Tulane University
New Orleans, Louisiana*



WILEY-INTERSCIENCE

A John Wiley & Sons, Inc., Publication

Copyright © 2007 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic format. For information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data is available.

ISBN-13: 978-0-470-04001-0
ISBN-10: 0-470-04001-7

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

Preface

This book is about those little numbers that we just cannot escape. Try to remember the last day you didn't hear at least something about probabilities, chance, odds, randomness, risk, or uncertainty. I bet it's been a while. In this book, I will tell you about the mathematics of such things and how it can be used to better understand the world around you. It is not a textbook though. It does not have little colored boxes with definition or theorems, nor does it contain sections with exercises for you to solve. My main purpose is to entertain you, but it is inevitable that you will also learn a thing or two. There are even a few exercises for you, but they are so subtly presented that you might not even notice until you have actually solved them.

The spousal thanks is always more than a formality. I thank Αλκυμήνη for putting up with irregular work hours and everything else that comes with writing a book, but also for help with Greek words and for reminding me of some of my old travel stories that you will find in the book. I am deeply grateful to Professor Olle Häggström at Chalmers University of Technology in Göteborg, Sweden. He has read the entire manuscript, and his comments are always insightful, accurate, and clinically free from unnecessary politeness. If you find something in this book that strikes you as particularly silly, chances are that Mr. Häggström has already pointed it out to me but that I decided to keep it for spite. I have also received helpful comments from John Haigh at the

University of Sussex, Steve Quigley at Wiley, and from an anonymous referee. Thanks also to Kris Parrish and Susanne Steitz at Wiley, to Sheree Van Vreede at Sheree Van Vreede Publications Services for excellent copyediting, and to Amy Hendrickson at Texnology Inc. for promptly and patiently answering my LaTeX questions.

A large portion of this book was written during the tumultuous Fall of 2005. Our move from Houston to New Orleans in early August turned out to be a masterpiece of bad timing as Hurricane Katrina hit three weeks later. We evacuated to Houston, and when Katrina's sister Rita approached, we took refuge in the deserts of West Texas and New Mexico. Sandstorms are so much more pleasant than hurricanes! However, it was also nice to return to New Orleans in January 2006; the city is still beautiful, and its chargrilled oysters are unsurpassed. I am grateful to many people who housed us and helped us in various ways during the Fall and by doing so had direct or indirect impact on this book. Special thanks to Kathy Ensor & Co. at the Department of Statistics at Rice University in Houston and to Tom English & Co. at the College of the Mainland in Texas City for providing me with office space. Finally, thanks to Professor Peter Jagers at Chalmers University of Technology, who as my Ph.D. thesis advisor once in a distant past wisely guided me through my first serious encounters with probabilities, those little numbers that rule our lives.

PETER OLOFSSON
www.peterolofsson.com

New Orleans, 2006

Contents

Preface	v
1 Computing Probabilities: Right Ways and Wrong Ways	1
The Probabilist	1
The Probabilist's Toys and Language	4
The Probabilist's Rule Book	9
Independence, Airplanes, and Russian Peasants	14
Conditional Probability, Swedish TV, and British Courts	20
Liar, Liar	24
Total Probability, Used Cars, and Tennis Matches	28
Combinatorics, Pastrami, and Poetry	33
The von Trapps and the Binomial Distribution	37
Final Word	43
2 Surprising Probabilities: When Intuition Struggles	45
Boys, Girls, Aces, and Colored Cards	45
Goats and Gloats	50
Happy Birthday	52
	vii

Typical Atypicalities	57
Strategies, Shopping, and Spaghetti Westerns	61
The British Snob and I	65
Final Word	70
3 Tiny Probabilities: Why Are They So Hard to Escape?	71
Probable Improbabilities	71
Saddam and I	75
Taking Tiny Risks	80
A Million-to-One Shot, Doc, Million to One!	82
Monsieur Poisson and the Mysterious Number 37	84
Clumps in Space	89
Final Word	91
4 Backward Probabilities: The Reverend Bayes to Our Rescue	93
Driving Miss Daisy	93
Bayes, Balls, and Boys (and Girls)	96
Bayes and My Green Card	98
Objection Your Honor	103
Final Word	112
5 Beyond Probabilities: What to Expect	115
Great Expectations	115
Good Things Come to Those Who Wait	123
Expect the Unexpected	129
Size Matters (and Length, and Age)	132
Deviant Behavior	138
Final Word	143
6 Inevitable Probabilities: Two Fascinating Mathematical Results	145
Alea Iacta Est, Over and Over	145
Even-Steven? The Law Misunderstood	149
Coin Tosses and Freeway Congestion	155
Let's Get Serious	162

Bells and Bread	166
How a Toronto Quincunx Changed My Life	171
Final Word	173
7 Gambling Probabilities: Why Donald Trump Is Richer than You	175
French Letters	175
Roulette: A Classy Way to Waste Your Money	179
Craps: Not so Dicey After All	184
Blackjack: Money for Mnemonics	187
Math for Losers	193
Win Money and Lose Friends	200
Final Word	210
8 Guessing Probabilities: Enter the Statisticians	211
Lies, Damned Lies, and Beautiful Lies?	211
4 out of 10 Like the President 19 Times out of 20	215
Polls Gone Wild	220
The Lawsuit and the Lurker	225
Football Players and Geyser Eruptions	230
Snooping in the Abbot's Garden	237
Final Word	242
9 Faking Probabilities: Computer Simulation	245
Mahogany Dice and Modular Arithmetic	245
Random and Not-So-Random Digits	252
Number One Is Number One	253
Is Random Really Random?	256
Final Word	261
Index	263

This Page Intentionally Left Blank

Computing Probabilities: Right Ways and Wrong Ways

THE PROBABILIST

Whether you like it or not, probabilities rule your life. If you have ever tried to make a living as a gambler, you are painfully aware of this, but even those of us with more mundane life stories are constantly affected by these little numbers. Some examples from daily life where probability calculations are involved are the determination of insurance premiums, the introduction of new medications on the market, opinion polls, weather forecasts, and DNA evidence in courts. Probabilities also rule who you are. Did daddy pass you the X or the Y chromosome? Did you inherit grandma's big nose? And on a more profound level, quantum physicists teach us that *everything* is governed by the laws of probability. They toss around terms like the *Schrödinger wave equation* and *Heisenberg's uncertainty principle*, which are much too difficult for most of us to understand, but one thing they do mean is that the fundamental laws of physics can only be stated in terms of probabilities. And the fact that Newton's deterministic laws of physics are still useful can also be attributed to results from the theory of probabilities. Meanwhile, in everyday life, many of us use probabilities in our language and say things like "I'm 99% certain" or "There is a one-in-a-million chance" or, when something unusual happens, ask the rhetorical question "What are the odds?"

Some of us make a living from probabilities, by developing new theory and finding new applications, by teaching others how to use them, and occa-

sionally by writing books about them. We call ourselves *probabilists*. In the universities, you find us in mathematics and statistics departments; there are no departments of probability. The terms “mathematician” and “statistician” are much more well known than “probabilist,” and we are a little bit of both but we don’t always like to admit it. If I introduce myself as a mathematician at a cocktail party, people wish they could walk away. If I introduce myself as a statistician, they do. If I introduce myself as a probabilist...well, most actually still walk away. They get upset that somebody who sounds like the Swedish Chef from the Muppet Show tries to impress them with difficult words. But some stay and give me the opportunity to tell them some of the things I will now tell you about.

Let us be etymologists for a while and start with the word itself, *probability*. The Latin roots are *probare*, which means to test, to prove, or to approve, and *habilis*, which means apt, skillful, able. The word “probable” was originally used in the sense “worthy of approval,” and its connection to randomness came later when it came to mean “likely” or “reasonable.” In my native Swedish, the word for probable is “sannolik,” which literally means “truthlike” as does the German word “wahrscheinlich.” The word “probability” still has room for nuances in the English language, and Merriam-Webster’s online dictionary lists four slightly different meanings. To us a probability is a number used to describe how likely something is to occur, and probability (without indefinite article) is the study of probabilities.

Probabilities are used in situations that involve randomness. Many clever people have thought about and debated what randomness really is, and we could get into a long philosophical discussion that could fill the rest of the book. Let’s not. The French mathematician Pierre-Simon Laplace (1749–1827) put it nicely: “Probability is composed partly of our ignorance, partly of our knowledge.” Inspired by Monsieur Laplace, let us agree that you can use probabilities whenever you are faced with uncertainty. You could:

- Toss a coin, roll a die, spin a roulette wheel
- Watch the stock market, the weather, the Super Bowl
- Wonder if there is an oil well in your backyard, if there is life on Mars, if Elvis is alive

These examples differ from each other. The first three are cases where the outcomes are equally likely. Each individual outcome has a probability that is simply one divided by the number of outcomes. The probability is $1/2$

to toss heads, $1/6$ to roll a 6, and $1/38$ to get the number 29 in roulette (an American roulette wheel has the numbers 1–36, 0, and 00). Pure and simple. We can also compute probabilities of groups of outcomes. For example, what is the probability to get an odd number when rolling a die? As there are three odd outcomes out of six total, the answer is $3/6 = 1/2$. These are examples of *classical probability*, the first type of probability problems studied by mathematicians, most notably, Frenchmen Pierre de Fermat and Blaise Pascal whose seventeenth century correspondence with each other is usually considered to have started the systematic study of probabilities. You will learn more about Fermat and Pascal later in the book.

The next three examples are cases where we must use data to be able to assign probabilities. If it has been observed that under current weather conditions it has rained about 20% of the days, we can say that the probability of rain today is 20%. This probability may change as more weather data are gathered and we can call it a *statistical probability*. As for the 2006 Super Bowl, I placed a bet on the Houston Texans that gave odds of 800 to 1, which means that the bookmaker assigned a probability of less than $1/800$ that the Texans would win. However he came to this conclusion, he must have used plenty of data other than that he once spent a summer in Houston and almost died of heatstroke.

The third trio of examples is different from the previous two in the sense that the outcome is already fixed; you just don't know what it is. Either there is an oil well or there isn't. Before you start drilling, you still want to have some idea of how likely you are to find oil and a geologist might tell you that the probability is about 75%. This percentage does not mean that the oil well is there nine months of the year and slides over to your neighbor the other three, but it does mean that the geologist thinks that your chances are pretty good. Another geologist may tell you the probability is 85%, which is a different number but means the same thing: Chances are pretty good. We call these *subjective probabilities*. In the case of a living Elvis, I suppose that depending on whom you ask you would get either 0% or 100%. I mean, who would say 25%? Little Richard?

Some knowledge about proportions may be helpful when assigning subjective probabilities. For example, suppose that your Aunt Jane in Pittsburgh calls and tells you that her new neighbor seems nice and has a job that "has something to do with the stars, astrologer or astronomer." Without having more information, what is the probability that the neighbor is an astronomer? As you have virtually no information, would you say 50%? Some people

might. But you should really take into account that there are about four times as many astrologers as astronomers in the United States, so a probability of 20% is more realistic. Just because something is “either/or” does not mean it is “50–50.” Andy Rooney may have been more insightful than he intended when he stated his 50–50–90 rule: “Anytime you have a 50–50 chance of getting something right, there’s a 90% probability you’ll get it wrong.”

THE PROBABILIST’S TOYS AND LANGUAGE

Probabilists love to play with coins and dice. In a Platonic sense. We like the *idea* of tossing coins and rolling dice as experiments that have equally likely outcomes. Suppose that a family with four children is chosen at random. What is the probability that all four are girls? A coin-tossing analogy would be to ask for the probability to get four heads when a coin is tossed four times. Many probability problems can be illustrated by coin tossing, but this would quickly become boring so we introduce variation by also rolling dice, spinning roulette wheels, picking balls from urns, or drawing from decks of cards. Dice, roulette, and card games are also interesting in their own right, and you will find a chapter on gambling later in the book. Of course. Probability without gambling is like beer without bubbles.

Probability is the art of being certain of how uncertain you are. The statement “the probability to get heads is $1/2$ ” is a precise statement. It tells you that you are as likely to get heads as you are to get tails. Another way to think about probabilities is in terms of average long-term behavior. In this case, if you toss the coin repeatedly, in the long run you will get roughly 50% heads and 50% tails. Of this you can be certain. What you cannot be certain of is how the next toss will come up.

Probabilists use special terminology. For example, we often refer to a situation where there is uncertainty as an “experiment.” This situation could be an actual experiment such as tossing a coin or rolling a die, but also something completely different such as following the stock market or watching the Wimbledon final. An experiment results in an *outcome* such as “heads,” “6,” “Volvo went up,” or “Björn Borg won” (those were the days). A group of outcomes is called an *event*. In plain language, an event is something that can happen in an experiment. It can be a single outcome (roll 6) or a group of outcomes (roll an odd number). The mathematical description of an event is that it is a *subset* of the set of all possible outcomes, and mathematicians would describe outcomes as *elements* of this set. Probabilists use the words

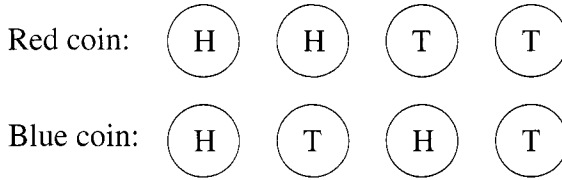


Figure 1.1 The four equally likely outcomes when you toss two coins.

“outcome” and “event” to emphasize the connection with things that happen in reality. In formulas, we denote events by uppercase letters and use the letter “P” to denote probability. The mathematical expression $P(A)$ should thus be read “the probability of (the event) A.” We may also talk about the probability of a statement rather than an event. However, it is mere language; the verbal description of an event is of course a statement.

The set of all possible outcomes is called the *sample space*.¹ Sometimes there is more than one choice of sample space. For example, suppose that you toss two coins and ask for the probability that you get two heads. As the number of heads can be 0, 1, or 2, you might be tempted to take these three numbers as the sample space and conclude that the probability to get two heads is $1/3$. However, if you repeated this experiment, you would notice after a while that you tend to get two heads less than one third of the tosses. The problem is that your sample space consists of three outcomes that are *not equally likely*. Let us distinguish between the two coins by painting one red and the other blue. There are then four equally likely outcomes: both show heads; the red shows heads and the blue shows tails; the red shows tails and the blue shows heads; and both show tails. In a more convenient notation, our sample space consists of the four equally likely outcomes HH, HT, TH, and TT. One out of four gives two heads, and the correct probability is $1/4$. See Figure 1.1 for an illustration of the four equally likely outcomes.

Here is a similar problem. If you roll two dice, what is the probability that the sum of the two equals eight? First note that the sum of two dice can be any of the numbers 2, 3, ..., 12 but that these are not equally likely. To find the equally likely outcomes, we need to distinguish between the two dice, for

¹The term “sample space” was coined by mathematician and Austro-Hungarian fighter pilot Richard von Mises. That is, he coined the German term *Merkmahlraum* (label space), which appears in his 1931 book with the impressive German title *Wahrscheinlichkeitsrechnung* (probability calculus).

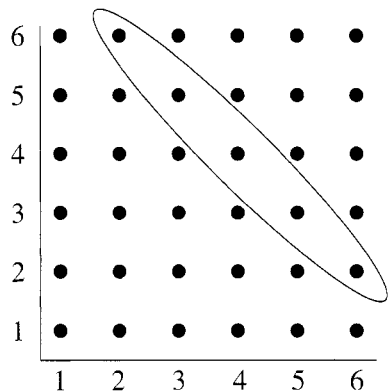


Figure 1.2 The sample space of 36 equally likely outcomes when you roll two dice. The event that the sum equals eight is marked; note that it consists of five outcomes because there are two ways to get 2 and 6 as well as 3 and 5 but only one way to get 4 and 4.

example, by pretending that they have different colors, red and blue, just like we did with the two coins above, and consider 36 possible outcomes. As the sum can be eight by adding $2 + 6$, $3 + 5$, or $4 + 4$, we might first think that there are 3 possibilities out of 36 to get sum eight, but we also need to distinguish, for example, between the cases “blue die equals 2 and red die equals 6” on the one hand and “blue die equals 6 and red die equals 2” on the other. If we make this distinction, we realize that there are five ways to get sum eight and the probability is $5/36$. See Figure 1.2 for an illustration of the sample space of 36 equally likely outcomes and the event that the sum equals eight.

Here is another example of a similar nature. Consider a randomly chosen family with three children. What is the probability that they have exactly one daughter? There can be 0, 1, 2, or 3 girls, but you know by now that these are not equally likely. Instead, distinguish the kids by birth order so that, for example, BGB means that the first child is a boy, the second a girl, and the third a boy. The eight equally likely outcomes are as follows:

BBB, BBG, BGB, GBB, BGG, GBG, GGB, GGG

We’re on easy street now; just note that three of the eight outcomes have one girl, and the probability of exactly one girl is therefore $3/8$. Now consider a randomly chosen girl who has two siblings. What is the probability that she has no sisters? This situation looks similar. If she has no sisters, this means

that her family has three children, exactly one of whom is a girl and we just saw that the probability of this is $3/8$. Convinced? You should not be. This situation is different because we are not choosing a *family* with three children; we are choosing a *girl* who belongs to such a family. Thus, the outcome BBB is impossible. Is the probability then $3/7$? Think about this for a while before you read on.

I hope you answered no. We need a completely new sample space that also accounts for the chosen girl. If we denote her by an asterisk, the 12 equally likely outcomes are as follows:

BBG*, BG*B, G*BB, BG*G, BGG*, G*BG
 GBG*, G*GB, GG*B, G*GG, GG*G, GGG*

and the probability that she has no sisters is $3/12 = 1/4$. Note how the previous outcomes are now split up according to how many girls they contain. The one with three girls, GGG, is split up into three equally likely outcomes because either of the three girls may be the chosen one. The probabilities that we have computed show that 37.5% of three-children families have exactly one daughter and 25% of girls from three-children families have no sisters.

What is the probability that all three children are of the same gender? Consider the following *faulty* argument: Two children must always be of the same gender. Whatever this gender is, the third child is equally likely to be of this gender or not, and thus the probability that all three are of the same gender is $1/2$. This example is a variant of a coin-tossing problem given by the British nobleman and amateur scientist Sir Francis Galton (about whom you will learn more in chapters to come) in 1894 to illustrate the dangers of sloppy thinking. Use our first sample space to discover the error, and argue that the correct probability is $1/4$.

Let us next consider an old gambling problem that goes along the same lines. I have three dice and offer you even odds to play the following game: The dice are rolled, and their sum is computed. If the sum is nine, you win. If it is ten, I win. If it is neither, I roll again. Is this game fair?

There are six ways in which the sum can be nine:

$1 + 2 + 6$, $1 + 3 + 5$, $1 + 4 + 4$, $2 + 2 + 5$, $2 + 3 + 4$, $3 + 3 + 3$

and likewise there are six ways to get sum ten:

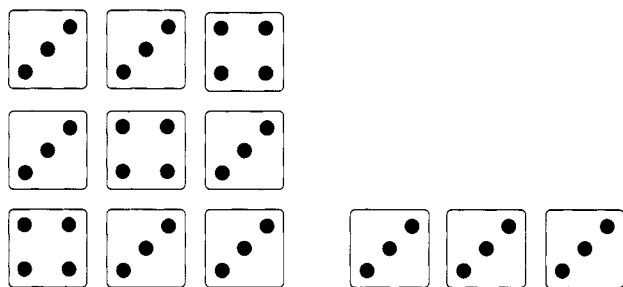


Figure 1.3 Three ways to get sum ten from two 3s and a 4 (left), but only one way to get sum nine from three 3s (right).

$$1 + 3 + 6, \quad 1 + 4 + 5, \quad 2 + 2 + 6, \quad 2 + 3 + 5, \quad 2 + 4 + 4, \quad 3 + 3 + 4$$

It sure looks like the game is fair, but beware, in the long run, I would slowly but surely win your money. But why?

Before you decide to play, you need to first identify the equally likely outcomes. And just like in the case of the two dice earlier, it is helpful to imagine that the three dice have three different colors, for example, red, green, and blue. If we list the dice in this order, the equally likely outcomes are $(1,1,1)$, $(1,1,2)$, $(1,2,1)$, $(2,1,1)$, $(2,2,1)$, and so on until $(6,6,6)$; a moment's thought reveals that there are $6 \times 6 \times 6 = 216$ of them. Let us look at one of the ways to get sum nine, $1 + 4 + 4$. This sum corresponds to three of the equally likely outcomes: $(1,4,4)$, $(4,1,4)$, and $(4,4,1)$. If we instead consider $1 + 2 + 6$, this corresponds to six outcomes: $(1,2,6)$, $(1,6,2)$, $(2,1,6)$, $(2,6,1)$, $(6,1,2)$, and $(6,2,1)$. In general, if all three dice show different numbers, this can occur in six ways; if two show the same number, this can occur in three ways; and if all three are the same, this can only occur in one way.

Now count above to realize that 27 outcomes give sum ten and only 25 give sum nine. The tie-breaker is the last outcome: There is only one way to combine $3 + 3 + 3$ but three ways to combine $3 + 3 + 4$; see Figure 1.3 for an illustration. Thus, out of the 52 outcomes that give a winner, I win in 27, or about 52%, and you win in the remaining 25, or 48%. Not a big difference, but it would be enough to make a living (some venture capital needed).

I mentioned that this problem is an old one. It was in fact solved almost 400 years ago by the great astronomer and telescope builder Galileo after being approached by a group of gambling Florentine noblemen. It is amusing to imagine how the world's most brilliant scientist of his time spent time helping people with their gambling problems. Good thing for Einstein that there were

no casinos in Atlantic City in the 1930s; his Princeton office might have been flooded by gamblers having spent the last of their money on a bus ticket, desperate for help from the genius.

We are often interested in more than one event. For example, suppose that people are chosen for an opinion poll and asked about their smoking habits and political sympathies. Consider one selected person. Let us denote the event that she is a smoker by S and the event that she is a Republican by R . We can then make up new events. The event that she is a smoker *and* a Republican is a new event, which we write as “ S and R .” The event that she is a smoker *or* a Republican is another new event, written as “ S or R .” It is important to know that we by “ S or R ” mean “smoker or Republican *or both*.” This definition of “or” is typical in mathematics, logic, and computer science. In daily language, it is often emphasized by using the expression “and/or” to distinguish from what math people call the *exclusive or*, which only permits one of the two, like in the phrase “You want fries or onion rings with that?”

The event that the selected individual is not a Republican is simply written as “not R .” The event that she is neither a Republican nor a smoker can be expressed in two different ways. One way is to negate that she is either, which gives “not (R or S).” The other way is to negate each separately and put them together: “(not R) and (not S).” We have argued for the following equality between events:

$$\text{not } (R \text{ or } S) = (\text{not } R) \text{ and } (\text{not } S)$$

The parentheses are there to make it clear to what “not” refers. In a similar way,

$$\text{not } (R \text{ and } S) = (\text{not } R) \text{ or } (\text{not } S)$$

Make sure that you understand these little exercises in logic; we will make use of them later.

THE PROBABILIST'S RULE BOOK

Probabilities can be expressed as fractions, as decimal numbers, or as percentages. If you toss a coin, the probability to get heads is $1/2$, which is the same as 0.5, which is the same as 50%. There are no rules for when to use which notation, and you will see examples of all three in this book. In daily language, proper fractions are often used and often expressed, for example, as “one in

ten” instead of $1/10$ (“one tenth”). This is also natural when you deal with equally likely outcomes. Decimal numbers are more common in technical and scientific reporting when probabilities are calculated from data. Percentages are also common in daily language and often with “chance” replacing “probability.” Meteorologists, for example, typically say things like “there is a 20% chance of rain.” The phrase “the probability of rain is 0.2” means the same thing. When we deal with probabilities from a theoretical viewpoint, we always think of them as numbers between 0 and 1, not as percentages.

Regardless of how probabilities are expressed, they must follow certain rules. One such rule that is easy to understand is that a probability can never be a negative number. The lowest possible probability is 0, meaning that we are dealing with something that just does not happen. There is no point in trying to emphasize this further by letting the probability be -0.3 or -5 .² A related rule is that a probability can never be more than 1 (or 100%). If the probability is 1 (or 100%), we are describing something that we are absolutely certain about. Of course you can still say that you are 200% certain that the Texas Rangers will win the World Series, but nobody outside Dallas will take you seriously.

The next rule is that the probability that something does *not* occur can be computed as one minus the probability that it does occur. In a formula,

$$P(\text{not } A) = 1 - P(A)$$

Also easy to accept. The probability not to get 6 when you roll a die is $5/6$, which is also equal to $1 - 1/6$. If the chance of rain is 20%, then the chance that it does not rain is 80%. In all its simplicity, this rule turns out to be surprisingly useful. In fact, in his excellent book *Taking Chances: Winning with Probability*, British probabilist John Haigh names it probability’s Trick Number One.

In the world of gambling, probabilities are often expressed by *odds*. To say that the odds are 4:1 against the event A means that it is four times as likely that A does not occur than that it occurs. We get the equation $P(\text{not } A) = 4 \times P(A)$,

²I do not know how familiar you are with negative numbers, but to mathematicians they are as natural as air and water. Here is the world’s funniest math joke: A biologist, a physicist, and a mathematician are sitting at a sidewalk cafe watching a house across the street. After a while two people enter the house. A little later, three people exit. “Reproduction,” says the biologist. “Measurement error,” says the physicist. “Hmm,” says the mathematician, “if a person enters the house it will be empty again.”

which has the solution $P(A) = 1/5$ and $P(\text{not } A) = 4/5$. As bookmakers are in the business to make a living, offering odds of 4:1 in reality means that they think that the probability of A is *less* than $1/5$.

Another rule. Let A and B be events such that whenever A occurs, B must also occur. Then $P(A)$ is less than (or equal to) $P(B)$, and the mathematical notation for this is $P(A) \leq P(B)$. For an example, let A be the event to roll a 6 and B the event to roll an even number. Whenever A occurs, B must also occur. However, B can occur without A occurring if you roll 2 or 4. In particular, the composition of two events is always less probable than each individual event. What I mean is that $P(A \text{ and } B)$ is always less than both $P(A)$ and $P(B)$, regardless of what A and B are.

As an example of the rule from the last paragraph, let us consider Mrs. Boudreaux and Mrs. Thibodeaux who are chatting over their fence when the new neighbor walks by. He is a man in his sixties with shabby clothes and a distinct smell of cheap whiskey. Mrs. B, who has seen him before, tells Mrs. T that he is a former Louisiana state senator. Mrs. T finds this very hard to believe. "Yes," says Mrs. B, "he is a former state senator who got into a scandal long ago, had to resign, and started drinking." "Oh," says Mrs. T, "that sounds more likely." "No," says Mrs. B, "I think you mean less likely."

Strictly speaking, Mrs. B is right. Consider the following two statements about the shabby man: "He is a former state senator" and "He is a former state senator who got into a scandal long ago, had to resign, and started drinking." It is tempting to think that the second is more likely because it gives a more exhaustive explanation of the situation at hand. However, this reason is precisely why it is a *less* likely statement. Note that whenever somebody satisfies the second description, he must also satisfy the first but not vice versa. Thus, the second statement has a lower probability (from Mrs. T's subjective point of view; Mrs. B of course knows who the man is). This example is a variant of examples presented in the book *Judgment under Uncertainty* by Economics Nobel laureate³ Daniel Kahneman and co-authors Paul Slovic and Amos Tversky. They show empirically how people often make similar mistakes when they are asked to choose the most probable among a set of statements. It certainly helps to know the rules of probability. A more discomfiting aspect

³But I want to point out that the Economics prize is not a 'true' Nobel prize in the sense that it was not mentioned in Alfred Nobel's will. The prize was first awarded in 1969, and its official name is "The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel." Just so that you know.

is that the more you explain something in detail, the more likely you are to be wrong. If you want to be credible, be vague.

The final rule is the *addition rule*. It says that in order to get the probability that either of two events occur, you add the probabilities of the two individual events. This rule, however, only applies if the two events in question *cannot occur at the same time* (the technical term for such events is that they are *mutually exclusive*). In a formula:

$$P(A \text{ or } B) = P(A) + P(B)$$

For example, roll a die and consider the events A: to get 6 and B: to get an odd number. These events qualify as mutually exclusive because you cannot get both 6 and an odd number in the same roll. It is “same roll” that is important here; of course you can get 6 in one roll and an odd number in the next. By the formula above, the probability to get 6 or an odd number in the same roll is $1/6 + 3/6 = 4/6$.

In his bestseller *Innumeracy*, John Allen Paulos tells the story of how he once heard a local weatherman claim that there was a 50% chance of rain on Saturday and a 50% chance of rain on Sunday and thus a 100% chance of rain during the weekend. Clearly absurd, but what is the error? Faulty use of the addition rule! As a rainy Saturday does not exclude a rainy Sunday, we here have two events that can both occur the same weekend. In cases like this one, there is a modified version of the addition rule that says that you first add the two probabilities as before and then subtract the probability that *both* events occur. In a formula, it looks as follows:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Note that if A and B cannot occur at the same time, then $P(A \text{ and } B) = 0$ and we have the first addition rule as a special case. If we let A denote the event that it rains on Saturday and B the event that it rains on Sunday, the event “A and B” describes the case in which it rains both days. To get the probability of rain over the weekend, we now add 50% and 50%, which gives 100%, but we must then subtract the probability that it rains both days. Whatever this is, it is certainly more than 0 so we end up with something less than 100%, just like common sense tells us that we should. I just wonder what the weatherman would have said if the chances of rain had been 75% each day.

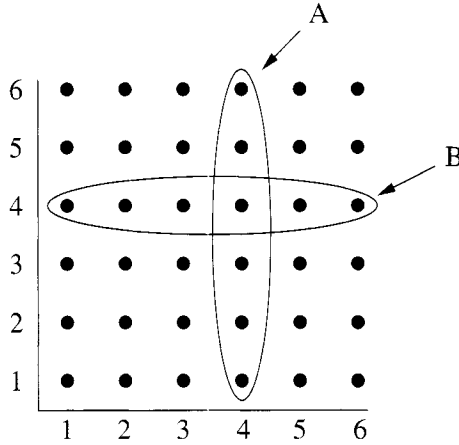


Figure 1.4 The sample space of 36 equally likely outcomes for rolling two dice. The events “4 on first die” and “4 on second die” are marked, and you may note that there are 6 outcomes in each event, 11 outcomes that are in at least one event, and 1 outcome that is in both.

Let us also check the formula in a dice example. If you roll two dice, what is the probability to get at least one 4? Here, the relevant events are A: 4 on the first die and B: 4 on the second die. The event to get at least one 4 is then the event “A or B,” and in Figure 1.4, you can check directly that this equals $11/36$. Also, $P(A) = 6/36$, $P(B) = 6/36$, and $P(A \text{ and } B) = 1/36$ because there is only one outcome that gives 4 on both dice. As $6 + 6 - 1 = 11$, the formula is valid.

Whenever probabilities are assigned, this must be done in a way such that none of the rules are violated. Ask a friend how likely he thinks it is that it will rain Saturday, Sunday, both days, and at least one of the days, respectively. You will then get four probabilities that must satisfy the rules that we have discussed above. For example, somebody may think that rain on Saturday is pretty likely, say 70%, and the same for Sunday. Rain both days? Well, maybe 50%. For the last probability, let’s say 80%. But this assignment of probabilities violates the addition rule because 80 is not equal to $70 + 70 - 50 = 90$. Somebody else might come up with the following probabilities (same order): 70%, 60%, 80%, and 50%. These do satisfy the addition rule but suffer from another problem. Can you tell which? (Hint: Mrs. Boudreaux could.)

Let us keep thinking about weekend weather. Suppose that both Saturday and Sunday each have probability 0.5 to get rain and that the probability is p

that it rains both days (we now think of probabilities as numbers between 0 and 1, not percentages). What is the range of possible values of p ? How does the probability of rain during the weekend depend on p ?

If we let A and B be the events “rain on Saturday” and “rain on Sunday” respectively, then a rainy weekend is the event “ A or B ,” and because $p = P(A \text{ and } B)$, we get the equation

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = 1 - p$$

As p must be less than both $P(A)$ and $P(B)$, it cannot be more than 0.5. If p is 0, then $P(A \text{ or } B) = 1$ and the rainy weekend is a fact. As p ranges from 0 to 0.5, the probability of a rainy weekend decreases from 1 to 0.5. Why? It has to do with how likely rainy Saturdays and Sundays are to come in pairs. Think of a year, which has 52 weekends. On average, we expect to get rain 26 Saturdays and 26 Sundays. If p is 0, this means that if it rains on a Saturday, it *never* rains on the following Sunday and if it does not rain on Saturday, it *always* rains on Sunday. Thus, the 26 rainy Saturdays and 26 rainy Sundays must be spread over the year so that they never come in pairs. The only way to do this is to let every weekend have exactly one rainy day. As p gets bigger, rainy days are more likely to come in pairs, and the extreme case is when $p = 0.5$. Then *all* rainy days come in pairs and the year has half of its weekends rainy and the other half dry.

Here is an exercise for you. Change the probabilities a little, and let $P(A) = 0.6$ and $P(B) = 0.7$, and let p again denote $P(A \text{ and } B)$. Explain why p must be between 0.3 and 0.6.

INDEPENDENCE, AIRPLANES, AND RUSSIAN PEASANTS

Plenty of random things happen in the world all the time, most of which have nothing to do with one another. If you toss a coin and I roll a die, the probability that you get heads is $1/2$ regardless of the outcome of my die. If there is a 20% chance of rain tomorrow, this does not change if a flu outbreak in Asia is reported. Changes in the U.S. stock market indexes have nothing to do with who wins the Wimbledon tennis tournament. Events that in this way are unrelated to each other are called *independent*. It is easy to compute the probability that two independent events both occur; simply multiply the probabilities of the two events. We call this computation the *multiplication rule*

for probabilities, described in a formula as

$$P(A \text{ and } B) = P(A) \times P(B)$$

It works in two directions. If we can argue that two events are independent, then we can use the multiplication rule to compute the probability that both occur at the same time. Conversely, if we can show that the multiplication rule holds, then we can conclude that the events are independent. It can be argued at some length why this is true and we will just look at some simple examples to convince ourselves that formula and intuition agree. Let us do the first example above, that you toss a coin and I roll a die. There are 12 equally likely outcomes: (H,1), ..., (H,6), (T,1), ..., (T,6) in the obvious notation. What is now the probability that you toss heads and I roll a 6? Obviously $1/12$. The individual probabilities of heads and 6 are $1/2$ and $1/6$, respectively, and $1/2 \times 1/6$ equals $1/12$ indeed.

For another example, take a deck of cards, draw one card, and consider the two events, A: to get an ace, and H: to get hearts. Are these independent? Let us check whether the multiplication rule holds. The individual probabilities are

$$P(A) = 4/52 = 1/13$$

$$P(H) = 13/52 = 1/4$$

and the probability to get both A and H is the probability to get the ace of hearts, which is $1/52$, which is the product of $1/13$ and $1/4$. We have

$$P(A \text{ and } H) = P(A) \times P(H)$$

which means that A and H are independent. Now remove the two of spades from the deck, reshuffle, and consider the same two events as above. Are they still independent? They must be, right? After all, the two of spades has nothing to do with either aces or hearts. Let us compute the probabilities. There are now 51 cards, and we get

$$P(A) = 4/51$$

$$P(H) = 13/51$$

and $P(A \text{ and } H) = 1/51$. As $P(A \text{ and } H)$ is not equal to $P(A) \times P(H)$, we must conclude that the events are not independent anymore. What happened?

Removing the two of spades changes the proportions of aces in the deck from $4/52$ to $44/51$, but not *within the suit of hearts* where it remains at $1/13 = 4/52$. Here is how you should think about independent events: *If one event has occurred, the probability of the other does not change.* In the card example, the probability of A is $4/51$ but changes to $1/13$ if the event H occurs.

Here is a question I often ask my students after I have introduced independence: If two events cannot occur at the same time, are they independent? At first you might think so. After all, they have nothing to do with each other, right? Wrong! They have a lot to do with each other. If one has occurred, we know for certain that the other cannot occur. The probability to roll a 6 is $1/6$, but if I tell you that the outcome is an odd number, the probability of a 6 drops down to 0. Think this through. It is important to understand independence.

There is a story that is sometimes told about the great Russian mathematician Andrey Nikolaevich Kolmogorov, among many other things the founder of the modern theory of probability. In Stalin's Soviet Union in the 1930s, the concept of independence did not fit well with the historical determinism of Marxist ideology. When questioned by a panel of ideologues about this possible heresy, Kolmogorov countered, "If the peasants pray for rain and it actually starts to rain, were their prayers answered?" The atheist ideologues had to confess that this must indeed be a case of independent events and Kolmogorov lived a long and productive life until his death in 1987 at the age of 84.

In December 1992, a small passenger airplane crashed in a residential neighborhood near Bromma airport outside Stockholm in Sweden, causing no death or injury to any of the residents. Already disturbed by increasing traffic and expansion plans for the airport, the residents now got more reasons to worry. In an effort to calm people, the airport manager said in an interview on TV that statistically people should now feel safer because the probability to have another accident had become so much smaller than before. I was at the time a graduate student in Sweden, studying probability and statistics, and thought that it was amusing to hear both "statistically" and "probability" used in the same sentence in such a careless way. In youthful vigor, I immediately wrote a letter that was published in some leading Swedish newspapers, where I explained why the airport manager's statement was incorrect. I also encouraged him to contact me so that I could recommend a good probability textbook. I never heard from him.