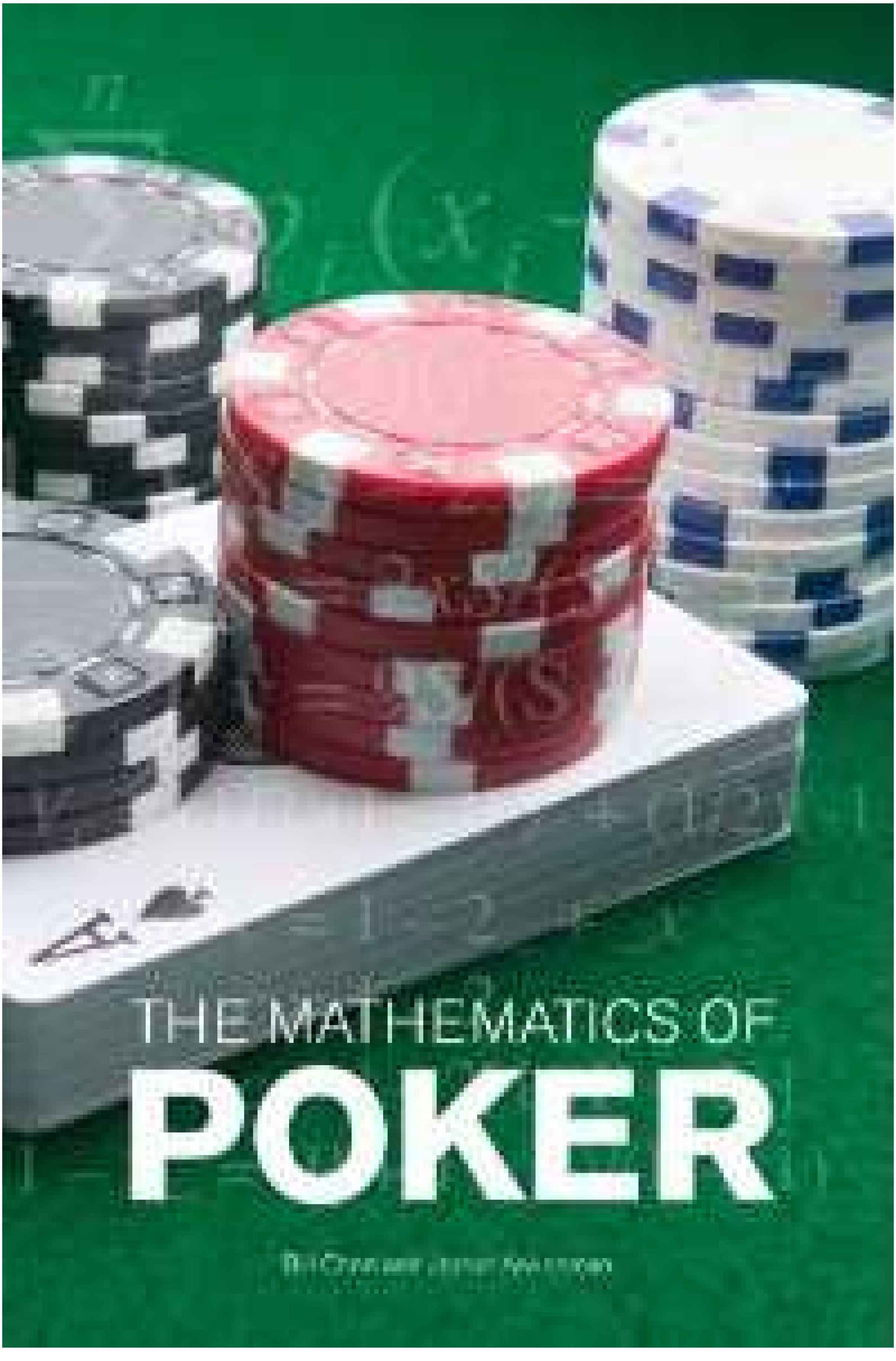
THE MATHEMATICS OF
POKER

Other ConJelCo titles:

Cooke's Rules of Real Poker by Roy Cooke and John Bond
Hold'em Excellence by Lou Krieger
How to Play Like a Poker Pro by Roy Cooke and John Bond
How to Think Like a Poker Pro by Roy Cooke and John Bond
Internet Poker by Lou Krieger and Kathleen Watterson
Mastering No-Limit Hold'em by Russ Fox and Scott T. Harker
More Hold'em Excellence by Lou Krieger
Real Poker II: The Play of Hands J 992-1999 (2nd ed.)
by Roy Cooke and John Bond
Serious Poker by Dan Kimberg
Stepping Up by Randy Burgess
The Home Poker Handbook by Roy Cooke and John Bond
Why You Lose at Poker by Russ Fox and Scott T. Harker
Winning Low-Limit Hold'em by Lee Jones
Winning Omaha/8 Poker by Mark Tenner and Lou Krieger
Winning Strategies for No-Limit Hold'em
by Nick Christenson and Russ Fox
Video Poker – Optimum Play by Dan Paymar

Software

StatKing for Windows

The Mathematics of Poker

Publisher's Cataloging-in-Publication Data
Chen, Bill
Ankenman, Jerrod

The Mathematics of Poker
x,382p. ; 29cm.
ISBN-13: 978-1-886070-25-7
ISBN-10:1-886070-25-3

I Title.
Library of Congress Control Number: 2006924665

First Edition
5 7 9 8 6

Cover design by Cat Zaccardi
Book design by Daniel Pipitone
Digital production by Melissa Neely

ConJeiCo LLC
1460 Bennington Ave
Pittsburgh, PA 15217
[412] 621-6040
http://www.conjelco.com

Errata, if any, can be found at http://www.conjelco.com/mathofpoker/

Acknowledgments

A book like this is rarely the work of simply the authors; many people have assisted us along the way, both in understanding poker and in the specific task of writing down many of the ideas that wc have developed over the last few years. A book called *The Mathematics of' Poker* was conceived by Bill, Chuck Weinstock, and Andrew Latto several years ago, before Jerrod and Bill had even met. That book was set to be a fairly formal, textbook-like approach to discussing the mathematical aspects of the game. A few years later, Jerrod and Bill had begun to collaborate on solving some poker games and the idea resurfaced as a book that was less like a mathematics paper and more accessible to readers with a modest mathematical background.

Our deepest thanks to those who read the manuscript and provided valuable feedback. Most notable among those were Andrew Bloch, Andrew Prock, and Andrew Latto, who scoured sections in detail, providing criticisms and suggestions for improvement. Andrew Prock's PokerStove tool (http://www.pokerstove.com) was quite valuable in performing many of the equity calculations. Others who read the manuscript and provided useful feedback were Paul R. Pudaite and Michael Maurer.

Jeff Yass at Susquehanna International Group (http://www.sig.com) has been such a generous employer in allowing Bill to work on poker, play the World Series, and so on. We thank Doug Costa, Dan Loeb,Jay Siplestien, and Alexei Dvoretskii at SIG for their helpful comments.

We have learned a great deal from a myriad of conversations with various people in the poker community. Our friend and partner Matt Hawrilenko, and former WSOP champion Chris Ferguson are two individuals especially worth singling out for their insight and knowledge. Both of us participate enthusiastically in the *ARG community, and discussions with members of that community, too, have been enlightening, particularly round table discussions with players including Sabyl Cohen, Jack Mahalingam, JP Massar, Path Beadles, Steve Landrum. and (1999 Tournament of Champions winner) Spencer Sun.

Sarah Jennings, our editor, improved the book significantly in a short period of rime by being willing to slog through equations and complain about the altogether too frequent skipped steps. Chuck Weinstock at Conjelco has been amazingly patient with all the delays and extended deadlines that come from working with authors for whom writing is a secondary occupation.

We also appreciate the comments and encouragement from Bill's father Dr. An-Ban Chen who has written his own book and has numerous publications in his field of solid state physics. Jerrod's mother Judy has been a constant source of support for all endeavors, even crazy-sounding ones like playing poker for a living.

Throughout the writing of this book, as in the rest of life, Michelle Lancaster has constantly been wonderful and supportive of Jerrod. This book could not have been completed without her. Patricia Walters has also provided Bill with support and ample encouragement, and it is of no coincidence that the book was completed during the three years she has known Bill.

Foreword

Don’t believe a word I say.
It’s not that I'm lying when I tell you that this is an important book. I don't even lie at the poker table -- not much, anyway - so why would I lie about a book I didn't even write?

It’s just that you can't trust me to be objective. I liked this book before I'd even seen a single page. I liked it when it was just a series of conversations between Bill, myself, and a handful of other math geeks. And if I hadn't made up my mind before I'd read it, I'm pretty sure they’d have won me over with the first sentence.
Don’t worry, though. You don't have to trust me. Math doesn't lie. And results don't lie, either. In the 2006 WSOP, the authors finished in the money seven times, including Jerrod's second place finish in Limit Holdem, and Bill's two wins in Limit and Short Handed No Limit Hold'em.

Most poker books get people talking. The best books make some people say, “How could anyone publish our carefully guarded secrets?" Other times, you see stuff that looks fishy enough to make you wonder if the author wasn't deliberately giving out bad advice. I think this book will get people talking, too, but it won't be the usual sort of speculation. No one is going to argue that Bill and Jerrod don't know their math.

The argument will be about whether or not the math is important.

People like to talk about poker as "any man's game." Accountants and lawyers, students and housewives can all compete at the same level - all you need is a buy-in, some basic math and good intuition and you, too, can get to the final table of the World Series of Poker. That notion is especially appealing to lazy people who don't want to have to spend years working at something to achieve success. It's true in the most literal sense that anyone can win, but with some well-invested effort, you can tip the scales considerably in your favor.

The math in here isn't easy. You don't need a PhD in game theory to understand the concepts in this book, but it's not as simple as memorizing starting hands or calculating the likelihood of making your flush on the river. There's some work involved. The people who want to believe intuition is enough aren't going to read this book. But the people who make the effort will be playing with a definite edge. In fact, much of my poker success is the result of using some of the most basic concepts addressed in this book.

Bill and Jerrod have saved you a lot of time. They've saved me a lot of a time, too. I get asked a lot of poker questions, and most are pretty easy to answer. But I've never had a good response when someone asks me to recommend a book for understanding game theory as it relates to poker. I usually end up explaining that there are good poker books and good game theory books, but no book addresses the relationship between the two.

Now I have an answer. And if I ever find myself teaching a poker class for the mathematics department at UCLA, this will be the only book on the syllabus.

Chris “Jesus" Ferguson
Champion, 2000 World Series of Poker
November 2006

Introduction

"If you think the math isn't important, you don't know the right math."

Chris "Jesus" Ferguson, *2000 World Series of Poker champion*

Introduction

In the late 1970s and early 1980s, the bond and option markets were dominated by traders who had learned their craft by experience. They believed that their experience and intuition for trading were a renewable edge; that is, that they could make money just as they always had by continuing to trade as they always had. By the mid-1990s, a revolution in trading had occurred; the old school grizzled traders had been replaced by a new breed of quantitative analysts, applying mathematics to the "art" of trading and making of it a science.

If the latest backgammon programs, based on neural net technology and mathematical analysis had played in a tournament in the late 1970s, their play would have been mocked as overaggressive and weak by the experts of the time. Today, computer analyses are considered to be the final word on backgammon play by the world's strongest players - and the game is fundamentally changed for it.
And for decades, the highest levels of poker have been dominated by players who have learned the game by playing it, "road gamblers" who have cultivated intuition for the game and are adept at reading other players' hands from betting patterns and physical tells. Over the last five to ten years, a whole new breed of player has risen to prominence within the poker community. Applying the tools of computer science and mathematics to poker and sharing information across the Internet, these players have challenged many of the assumptions that underlie traditional approaches to the game. One of the most important features of this new approach to the game is a reliance on quantitative analysis and the application of mathematics to the game. Our intent in this book is to provide an introduction to quantitative techniques as applied to poker and to the application of game theory, a branch of mathematics, to poker.

Any player who plays poker is using some model, no matter what methods he uses to inform it. Even if a player is not consciously using mathematics, a model of the situation is implicit in his decisions; that is, when he calls, raises, or folds, he is making a statement about the relative values of those actions. By preferring one action over another, he articulates his belief that one action is better than another in a particular situation. Mathematics are a particularly appropriate tool for making decisions based on information. Rejecting mathematics as a tool for playing poker puts one's decision-making at the mercy of guesswork.

Common Misconceptions

We frequently encounter players who dismiss a mathematical approach out of hand, often based on their misconceptions about what this approach is all about. We list a few of these here; these are ideas that we have heard spoken, even by fairly knowledgeable players. For each of these, we provide a brief rebuttal here; throughout this book, we will attempt to present additional refutation through our analysis.

1) By analyzing what has happened in the past - our opponents, their tendencies, and so on-we can obtain a permanent and recurring edge.

This misconception is insidious because it seems very reasonable; in fact, we ***can*** gain an edge over our opponents by knowing their strategies and exploiting them. But this edge can be only temporary; our opponents, even some of the ones we think play poorly, adapt and evolve by reducing the quantity and magnitude of clear errors they make and by attempting to counter-exploit us. We have christened this first misconception the "PlayStation™ theory of poker" - that the poker world is full of players who play the same fixed strategy, and the goal of playing poker is to simply maximize profit against the fixed strategies of our opponents. In fact, our opponents' strategies are dynamic, and so we must be dynamic; no edge that we have is necessarily permanent.

2) Mathematical play is predictable and lacks creativity.

In some sense this is true; that is, if a player were to play the optimal strategy to a game, his strategy would be "predictable" - but there would be nothing at all that could be done with this information. In the latter parts of the book, we will introduce the concept of ***balance*** this is the idea that each action sequence contains a mixture of hands that prevents the opponent from exploiting the strategy. Optimal play incorporates a precisely calibrated mixture of bluffs, semi-bluffs, and value bets that make it appear entirely unpredictable. "Predictable" connotes "exploitable," but this is not necessarily true. If a player has aces every time he raises, this is predictable and exploitable. However, if a player always raises when he holds aces, this is not necessarily exploitable as long as he also raises with some other hands. The opponent is not able to exploit sequences that contain other actions because it is unknown if the player holds aces.

3) Math is not always applicable; sometimes "the numbers go out the window."

This misconception is related to the idea that for any situation, there is only one mathematically correct play; players assume that even playing exploitively, there is a correct mathematical play - but that they have a "read" which causes them to prefer a different play. But this is simply a narrow definition of "mathematical play" - incorporating new information into our understanding of our opponent's distribution and utilizing that information to play more accurately is the major subject of Part II. In fact, mathematics contains tools (notably Bayes' theorem) that allow us to precisely quantify the degree to which new information impacts our thinking; in fact, playing mathematically is more accurate as far as incorporating "reads" than playing by "feel."

4) Optimal play is an intractable problem for real-life poker games; hence, we should simply play exploitively.

This is an important idea. It is true that we currently lack the computing power to solve headsup holdem or other games of similar complexity. (We will discuss what it means to "solve" a game in Part III). We have methods that are known to find the answer, but they will not run on modern computers in any reasonable amount of time. "Optimal" play does not even exist for multiplayer games, as we shall see. But this does not prevent us from doing two things: attempting to create strategies which share many of the same properties as optimal strategies and thereby play in a "near-optimal" fashion; and also to evaluate candidate strategies and find out how far away from optimal they are by maximally exploiting them.

5) When playing [online, in a tournament, in high limit games, in low limit games...], you have to change your strategy completely to win.

This misconception is part of a broader misunderstanding of the idea of a "strategy" - it is in fact true that in some of these situations, you must take different actions, particularly exploitively, in order to have success. But this is not because the games are fundamentally different; it is because the other players play differently and so your responses to their play take different forms. Consider for a moment a simple example. Suppose you arc dealt A9s on the button in a full ring holdem game. In a small-stakes limit holdem game, six players might limp to you, and you should raise. In a high limit game, it might be raised from middle position, and you would fold. In a tournament, it might be folded to you, and you would raise. These are entirely different actions, but the broader strategy is the same in all - choose the most profitable action.

Throughout this book, we will discuss a wide variety of poker topics, but overall, our ideas could be distilled to one simple piece of play advice: ***Maximize average profit***. This idea is at the heart of all our strategies, and this is the one thing that doesn't change from game condition to game condition.

Psychological Aspects

Poker authors, when faced with a difficult question, are fond of falling back on the old standby, "'It depends." - on the opponents, on one's 'read', and so on. And it is surely true that the most profitable action in many poker situations does in fact depend on one's sense, whether intuitive or mathematical, of what the opponent holds (or what he can hold). But one thing that is often missing from the qualitative reasoning that accompanies "It depends," is a real answer or a methodology for arriving at an action. In reality, the answer does in fact depend on our assumptions, and the tendencies and tells of our opponents are certainly something about which reasonable people can disagree. But once we have characterized their play into assumptions, the methods of mathematics take over and intuition fails as a guide to proper play.

Some may take our assertion that quantitative reasoning surpasses intuition as a guide to play as a claim that the psychological aspects of poker are without value. But we do not hold this view. The psychology of poker can be an absolutely invaluable tool for exploitive play, and the assumptions that drive the answers that our mathematical models can generate are often strongly psychological in nature. The methods by which we utilize the information that our intuition or people-reading skills give us is our concern here. In addition, we devote time to the question of what we ought to do when we are unable to obtain such information, and also in exposing some of the poor assumptions that often undermine the information-gathering efforts of intuition. With that said, we will generally, excepting a few specific sections, ignore physical tells and opponent profiling as being beyond the scope of this book and more adequately covered by other writers, particularly in the work of Mike Garo.

About This Book

We are practical people - we generally do not study poker for the intellectual challenge, although it turns out that there is a substantial amount of complexity and interest to the game. We study poker with mathematics because by doing so, we make more money. As a result, we are very focused on the practical application of our work, rather than on generating proofs or covering esoteric, improbable cases. This is not a mathematics textbook, but a primer on the application of mathematical techniques to poker and in how to turn the insights gained into increased profit at the table.
Certainly, there are mathematical techniques that can be applied to poker that are difficult and complex. But we believe that most of the mathematics of poker is really not terribly difficult, and we have sought to make some topics that may seem difficult accessible to players without a very strong mathematical background. But on the other hand, it is math, and we fear that if you are afraid of equations and mathematical terminology, it will be somewhat difficult to follow some sections. But the vast majority of the book should be understandable to anyone who has completed high school algebra. We will occasionally refer to results or conclusions from more advanced math. In these cases, it is not of prime importance that you understand exactly the mathematical technique that was employed. The important element is the concept - it is very reasonable to just "take our word for it" in some cases.

To help facilitate this, we have marked off the start and end of some portions of the text so that our less mathematical readers can skip more complex derivations. Just look for this icon for guidance, indicating these cases.

In addition,

Solution:

Solutions to example problems are shown in shaded boxes.

As we said, this book is not a mathematical textbook or a mathematical paper to be submitted to a journal. The material here is not presented in the manner of formal proof, nor do we intend it to be taken as such. We justify our conclusions with mathematical arguments where necessary and with intuitive supplemental arguments where possible in order to attempt to make the principles of the mathematics of poker accessible to readers without a formal mathematical background, and we try not to be boring. The primary goal of our work here is not to solve game theory problems for the pure joy of doing so; it is to enhance our ability to win money at poker.

This book is aimed at a wide range of players, from players with only a modest amount of experience to world-class players. If you have never played poker before, the best course of action is to put this book down, read some of the other books in print aimed at beginners, play some poker, learn some more, and then return after gaining additional experience. If you are a computer scientist or options trader who has recently taken up the game, then welcome. This book is for you. If you are one of a growing class of players who has read a few books, played for some time, and believe you are a solid, winning player, are interested in making the next steps but feel like the existing literature lacks insight that will help you to raise your game, then welcome. This book is also for you. If you are the holder of multiple World Series of Poker bracelets who plays regularly in the big game at the Bellagio, you too are welcome. There is likely a fair amount of material here that can help you as well.

Organization

The book is organized as follows:

Part I: Basics, is an introduction to a number of general concepts that apply to all forms of gambling and other situations that include decision making under risk. We begin by introducing probability, a core concept that underlies all of poker. We then introduce the concept of a probability distribution, an important abstraction that allows us to effectively analyze situations with a large number of possible outcomes, each with unique and variable probabilities. Once we have a probability distribution, we can define expected value, which is the metric that we seek to maximize in poker. Additionally, we introduce a number of concepts from statistics that have specific, common, and useful applications in the field of poker, including one of the most powerful concepts in statistics, Bayes' theorem.

Part II: Exploitive Play, is the beginning of our analysis of poker. We introduce the concept of a toy game, which is a smaller, simpler game that we can solve in order to gain insight about analogous, more complicated games. We then consider examples of toy games in a number of situations. First we look at playing poker with the cards exposed and find that the play in many situations is quite obvious; at the same time, we find interesting situations with some counter-intuitive properties that are helpful in understanding full games. Then we consider what many authors treat as the heart of poker, the situation where we play our single hand against a distribution of the opponent's hands and attempt to exploit his strategy, or maximize our win against his play. This is the subject of the overwhelming majority of the poker literature. But we go further, to the (in our view) much more important case, where we are not only playing a single hand against the opponent, but playing an entire distribution of hands against his distribution of hands. It is this view of poker, we claim, that leads to truly strong play.

Part III: Optimal Play, is the largest and most important part of this book. In this part, we introduce the branch of mathematics called game theory. Game theory allows us to find optimal strategies for simple games and to infer characteristics of optimal strategies for more complicated games even if we cannot solve them directly. We do work on many variations of the AKQ game, a simple toy game originally introduced to us in Card Player magazine by Mike Caro. We then spend a substantial amount of time introducing and solving [0,1] poker games, of the type introduced by John von Neumann and Oskar Morganstem in their seminal text on game theory "Theory of Games and Economic Behavior (1944), but with substantially more complexity and relevance to real-life poker. We also explain and provide the optimal play solution to short-stack headsup no-limit holdem.

Part IV: Bankroll and Risk includes material of interest on a very important topic to anyone who approaches poker seriously. We present the risk of ruin model, a method for estimating the chance of losing a fixed amount playing in a game with positive expectation but some variance. We then extend the risk of ruin model in a novel way to include the uncertainty surrounding any observation of win rate. We also address topics such as the Kelly criterion, choosing an appropriate game level, and the application of portfolio theory to the poker metagame.

Part V: Other Topics includes material on other important topics, tournaments are the fastest-growing and most visible form of poker today; we provide an explanation of concepts and models for calculating equity and making accurate decisions in the tournament environment. We consider the game theory of multiplayer games, an important and very complex branch of game theory, and show some reasons why the analysis of such games is so difficult. In this section we also articulate and explain our strategic philosophy of play, including our attempts to play optimally or at least pseudo- optimally as well as the situations in which we play exploitively.

How This Book Is Different

This book differs from other poker books in a number of ways. One of the most prominent is in its emphasis on quantitative methods and modeling. We believe that intuition is often a valuable tool for understanding what is happening. But at the same time, we eschew its use as a guide to what action to take. We also look for ways to identify situations where our intuition is often wrong, and attempt to retrain it in such situations in order to improve the quality of our reads and our overall play. For example, psychologists have identified that the human brain is quite poor at estimating probabilities, especially for situations that occur with low frequency. By creating alternate methods for estimating these probabilities, we can gain an advantage over our opponents.

It is reasonable to look at each poker decision as a two-part process of gathering information and then synthesizing that information and choosing the right action. It is our contention that intuition has no place in the latter. Once we have a set of assumptions about the situation - how our opponent plays, what our cards are, the pot size, etc., then finding the right action is a simple matter of calculating expectation for the various options and choosing the option that maximizes this.

The second major way in which this book differs from other poker books is in its emphasis on ***strategy,*** contrasted to an emphasis on ***decisions.*** Many poker books divide the hand into sections, such as "preflop play," "flop play," "turn play" etc. By doing this, however, they make it difficult to capture the way in which a player's preflop, flop, turn, and river play are all intimately connected, and ultimately part of the same strategy. We try to look at hands and games in a much more organic fashion, where, as much as possible, the evaluation of expectation occurs not at each decision point but at the beginning of the hand, where a full strategy for the game is chosen. Unfortunately, holdem and other popular poker games are extraordinarily complex in this sense, and so we must sacrifice this sometimes due to computational infeasibility. But the idea of carrying a strategy forward through different betting rounds and being constantly aware of the potential hands ***we*** could hold at this point, which our fellow poker theorists Chris Ferguson and Paul R. Pudaite call "reading your own hand," is essential to our view of poker.

A third way in which this book differs from much of the existing literature is that it is not a book about how to play poker. It is a book about how to think about poker. We offer very little in terms of specific recommendations about how to play various games; instead this book is devoted to examining the issues that are of importance in determining a strategy. Instead of a roadmap to how to play poker optimally, we instead try to offer a roadmap to how to think about optimal poker.

Our approach to studying poker, too, diverges from much of the existing literature. We often work on toy games, small, solvable games from which we hope to gain insight into larger, more complex games. In a sense, we look at toy games to examine dimensions of poker, and how they affect our strategy. How does the game change when we move from cards exposed to cards concealed? From games where players cannot fold to games where they can? From games where the first player always checks to games where both players can bet? From games with one street to games with two? We examine these situations by way of toy games - because toy games, unlike real poker, are solvable in practice - and attempt to gain insight into how we should approach the larger game.

Our Goals

It is our hope that our presentation of this material will provide at least two things; that it will aid you to play more strongly in your own poker endeavors and to think about situations in poker in a new light, and that it will serve as a jumping-off point toward the incredible amount of serious work that remains to be done in this field. Poker is in a critical stage of growth at this writing; the universe of poker players and the mainstream credibility of the game have never been larger. Yet it is still largely believed that intuition and experience are determining factors of the quality of play - just as in the bond and options markets in the early 1980s, trading was dominated by old-time veterans who had both qualities in abundance. A decade later, the quantitative analysts had grasped control of the market, and style and intuition were on the decline. In the same way, even those poker players regarded as the strongest in the world make serious errors and deviations from optimal strategies. This is not an indictment of their play but a reminder that the distance between the play of the best players in the world and the best play possible is still large, and that therefore there is a large amount of profit available to those who can bridge that gap.

Part I: Basics

"As for as the laws of mathematics refer to reality, they are not certain; as for as they are certain, they do not refer to reality."

Albert Einstein

Chapter 1
Decisions Under Risk: Probability and Expectation

There are as many different reasons for playing poker as there are players who play the game. Some play for social reasons, to feel part of a group or "one of the guys," some play for recreation, just to enjoy themselves. Many play for the enjoyment of competition. Still others -lay to satisfy gambling addictions or to cover up other pain in their lives. One of the difficulties of taking a mathematical approach to these reasons is that it's difficult to quantify the value of having fun or of a sense of belonging.

In addition to some of the more nebulous and difficult to quantify reasons for playing poker, there may also be additional financial incentives not captured within the game itself. For example, the winner of the championship event of the World Series of Poker is virtually guaranteed to reap a windfall from endorsements, appearances, and so on, over and above the large first prize.

There are other considerations for players at the poker table as well; perhaps losing an additional hand would be a significant psychological blow. While we may criticize this view as irrational, it must still factor into any exhaustive examination of the incentives to play poker. Even if we restrict our inquiry to monetary rewards, we find that preference for money is non-linear. For most people, winning five million dollars is worth much more (or has much more ***utility)*** than a 50°/o chance of winning ten million; five million dollars is life-changing money for most, and the ***marginal value*** of the additional five million is much smaller.

In a broader sense, all of these issues are included in the ***utility theory*** branch of economics. Utility theorists seek to quantify the preferences of individuals and create a framework under which financial and non-financial incentives can be directly compared. In reality, it is utility that we seek to maximize when playing poker (or in fact, when doing any tiring). However, the use of utility theory as a basis for analysis presents a difficulty; each individual has his own utility curves and so general analysis becomes extremely difficult.

In this book, we will therefore refrain from considering utility and instead use money won inside the game as a proxy for utility. In the bankroll theory section in Part IV, we will take an in-depth look at certain meta-game considerations, introduce such concepts as ***risk of ruin, the Kelly criterion,*** and ***certainty equivalent.*** All of these are measures of risk that have primarily to do with factors outside the game. Except when expressly stated, however, we will take as a premise that players are adequately bankrolled for the games they are playing in, and that their sole purpose is to maximize the money they will win by making the best decisions at every point.

Maximizing total money won in poker requires that a player maximize the ***expected value*** of his decisions. However, before we can reasonably introduce this cornerstone concept, we must first spend some time discussing the concepts of probability that underlie it. The following material owes a great debt to Richard Epstein's text *The Theory of Gambling and Statistical Logic* (1967), a valuable primer on probability and gambling.

Probability

Most of the decisions in poker take place under conditions where the outcome has not yet been determined. When the dealer deals out the hand at the outset, the players' cards are unknown, at least until they are observed. Yet we still have some information about the contents of the other players' hands. The game's rules constrain the contents of their hands-while a player may hold the jack-ten of hearts, he cannot hold the ace-prince of billiard tables, for example. The composition of the deck of cards is set before starting and gives us information about the hands.

Consider a holdem hand. What is the chance that the hand contains two aces? You may know the answer already, but consider what the answer means. What if we dealt a million hands just like this? How many pairs of aces would there be? What if we dealt ten million? Over time and many trials, the ratio of pairs of aces to total hands dealt will converge on a particular number. We define ***probability*** as this number. Probability is the key to decision-making in poker as it provides a mathematical framework by which we can evaluate the likelihood of uncertain events.

If n trials of an experiment (such as dealing out a holdem hand) produce *no* occurrences of an event *x*, we define the probability *p* of *x* occurring *p{x)* as follows:

$$p(x) = \lim_{n\to\infty} \frac{n_0}{n} \quad \textbf{(1.1)}$$

Now it happens to be the case that the likelihood of a single holdem hand being a pair of aces is $^1/_{221}$. We could, of course, deterrnine this by dealing out ten billion hands and observing the ratio of pairs of aces, to total hands dealt. Tins, however, would be a lengthy and difficult process, and we can do better by breaking the problem up into components. First we consider just one card. What is the probability that a single card is an ace? Even this problem can be broken down further - what is the probability that a single card is the ace of spades?

This final question can be answered rather directly. We make the following assumptions:

- There are fifty-two cards in a standard deck.
- Each possible card is equally likely.

Then the probability of any particular card being the one chosen is $^1/_{52}$. If the chance of the card being the ace of spades is $^1/_{52}$, what is the chance of the card being any ace? This is equivalent to the chance that the card is the ace of spades OR that it is the ace of hearts OR that it is the ace of diamonds OR that it is the ace of clubs. There are four aces in the deck, each with a $^1/_{52}$ chance of being the card, and summing these probabilities, we have:

$$p(A) = (4)\left(\frac{1}{52}\right)$$

$$p(A) = \left(\frac{1}{13}\right)$$

We can sum these probabilities directly because they are ***mutually exclusive;*** that is, no card can simultaneously be both the ace of spades and the ace of hearts. Note that the probability $^1/_{13}$ is exactly equal to the ratio (number of aces in the deck)/(number of cards total). This relationship holds just as well as the summing of the individual probabilities.

Independent Events

Some events, however, are not mutually exclusive. Consider for example, these two events:

1. The card is a heart
2. The card is an ace.

If we try to figure out the probability that a single card is a heart OR that it is an ace, we find there are thirteen hearts in the deck out of fifty-cards, so the chance that the card is a heart is $^1/_4$. The chance that the card is an ace is, as before, $^1/_{13}$. However, we cannot simply add these probabilities as before, because it is possible for a card to both be an ace and a heart.
There are two types of relationships between events. The first type are events that have no effect

on each other. For example, the closing value of the NASDAQ stock index and the value of the dice on a particular roll at the craps table in a casino in Monaco that evening are basically unrelated events; neither one should impact the other in any way that is not negligible. If the probability of both events occurring equals the product of the individual probabilities, then the events are said to be ***independent.*** The probability that both A and B occur is called the ***joint probability*** of A and B.

In this case, the joint probability of a card being both a heart and an ace is $({}^{1}/_{13})$ $({}^{1}/_{4})$, or ${}^{1}/_{52}$. This is because the fact that the card is a heart does not affect the chance that it is an ace - all four suits have the same set of cards.

Independent events are not mutually exclusive except when one of the events has probability zero. In this example, the total number of hearts in the deck is thirteen, and the total of aces in the deck is four. However, by adding these together, we are double-counting one single card (the ace of hearts). There are actually thirteen hearts and three other aces, or if you prefer, four aces, and twelve other hearts. It turns out that the general application of this concept is that the probability that at least one of two mutually non-exclusive events A and B will occur is the sum of the probabilities of A and B minus the joint probability of A and B. So the probability of the card being a heart ***or*** an ace is equal to the chance of it being a heart $({}^{1}/_{4})$ plus the chance of it being an ace minus the chance of it being both $({}^{1}/_{52})$, or ${}^{4}/_{13}$. This is true for all events, independent or dependent.

Dependent Events

Some events, by contrast, do have impacts on each other. For example, before a baseball game, a certain talented pitcher might have a 3% chance of pitching nine innings and allowing no runs, while his team might have a 60% chance of winning the game. However, the chance of the pitcher's team winning the game and him also pitching a shutout is obviously not 60% times 3%. Instead, it is very close to 3% itself, because the pitcher's team will virtually always win the game when he accomplishes this. These events are called ***dependent.*** We can also consider the ***conditional probability*** of A given B, which is the chance that ***if*** B happens, A will also happen. The probability of A and B both occurring for dependent events is equal to the probability of A multiplied by the conditional probability of B given A. Events are independent if the conditional probability of A given B is equal to the probability of A alone.

Summarizing these topics more formally, if we use the following notation:

$p(A \cup B)$ = Probability of A or 8 occurring.
$p(A \cap B)$ = Probability of A and B occurring.
$p(A|B)$ = Conditional probability of ***A*** occurring given ***B*** has already occurred.

The ∪ and ∩ notations are from set theory and formally represent "union" and "intersection." We prefer the more mundane terms "or" and "and." Likewise, | is the symbol for "given," so we pronounce these expressions as follows:

$p(A \cup B)$ = “p of A or B”
$p(A \cap B)$ = “p of A and B”
$p(A|B)$ = “p of A given B”

Then for mutually exclusive events:

$$p(A \cup B) = p(A) + p(B) \qquad \textbf{(1.2)}$$

For independent events:

$$p(A \cap B) = p(A)p(B) \quad \textbf{(1.3)}$$

For all events:

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) \quad \textbf{(1.4)}$$

For dependent events:

$$p(A \cap B) = p(A)p(B|A) \quad \textbf{(1.5)}$$

Equation 1.2 is simply a special case of Equation 1.4 for mutually exclusive events, $p(A \cap B) = 0$. Likewise, Equation 1.3 is a special case of Equation 1.5, as for independent events, $p(B|A) = p(B)$**.** Additionally, if $p(B|A) = p(B)$**.,** then $p(A|B) = p(A)$**.**

We can now return to the question at hand. How frequently will a single holdem hand dealt from a full deck contain two aces? There are two events here:

- A: The first card is an ace.
- B: The second card is an ace.

p(A) = $^1/_{13}$, and *p(B)* = $^1/_{13}$ as well. However, these two events are dependent, if *A* occurs (the first card is an ace), then it is less likely that *B* will occur, as the cards are dealt without replacement. So $p(B|A)$ is the chance that the second card is an ace given that the first card is an ace. There are three aces remaining, and fifty-one possible cards, so $p(B|A) = {}^3/_{51}$, or $^1/_{17}$.

$$p(A \cap B) = p(A)p(B|A)$$
$$p(A \cap B) = (^1/_{13})(^1/_{17})$$
$$p(A \cap B) = {}^1/_{221}$$

There are a number of other simple properties that we can mention about probabilities. First, the probability of any event is at least zero and no greater than one. Referring back to the definition of probability, n trials will never result in more than n occurrences of the event, and never less than zero occurrences. The probability of an event that is certain to occur is one. The probability of an event that never occurs is zero. Tire probability of an event's ***complement*** -that is, the chance that an event does not occur, is simply one minus the event's probability.

Summarizing, if we use the following notation:

$p(\overline{A}) =$ Probability that *A* does not occur.

C = a certain event
I = an impossible event

Then we have:

$$0 \le p(A) \le 1 \text{ for any A} \quad \textbf{(1.6)}$$
$$p(C) = 1 \quad \textbf{(1.7)}$$
$$p(I) = 0 \quad \textbf{(1.8)}$$
$$p(A) + p(\overline{A}) = 1 \quad \textbf{(1.9)}$$

Equation 1.9 can also be restarted as:

$$p(A) = 1 - p(\overline{A}) \quad \textbf{(1.10)}$$

We can solve many probability problems using these rules
Some common questions of probability are simple, such as the chance of rolling double sixes on two dice, hi terms of probability, this can be stated using equation 1.3, since the die rolls are independent. Let *p(A)* be the probability of rolling a six on the first die and *p(B)* be the probability of rolling a six on the second die. Then:

$p(A \cap B) = p(A)p(B)$
$p(A \cap B) = (1/6)(1/6)$
$p(A \cap B) = 1/36$

Likewise, using equation 1.2, the chance of a single player holding aces, kings, or queens becomes:

$p(\text{AA}) = 1/221$
$p(\text{KK}) = 1/221$
$p(\text{QQ}) = 1/221$
$p(\{\text{AA, KK, QQ}\}) = p(\text{AA}) + p(\text{KK}) + p(\text{QQ}) = 3/221$

Additionally we can solve more complex questions, such as:

How likely is it that a suited hand will flop a flush?

We hold two of the flush suit, leaving eleven in the deck. All three of the cards must be of the flush suit, meaning that we have A — the first card being a flush card, B — the second card being a flush card given that the first card is a flush card, and C= the third card being a flush card given than both of the first two are flush cards.

$p(A) = 11/50$ (two cards removed from the deck in the player's hand)

$p(B|A) = 10/49$ (one flush card and three total cards removed)

$p(C|(A \cap B)) = 9/48$ (two flush cards and four total cards removed)

Applying equation 1.5, we get:

$p(A \cap B) = p(A)p(B|A)$
$p(A \cap B) = (1/50)(10/49)$
$p(A \cap B) = 11/245$

Letting D = $(A \cap B)$, we can use equation 1.5 again:

$p(D \cap C) = p(D)p(C|D)$
$p(A \cap B \cap C) = p(A \cap B)p(C|(A \cap B))$
$p(A \cap B \cap C) = (11/245)(9/48)$
$p(A \cap B \cap C) = 33/3920$, or a little less than 1%.

We can apply these rules to -virtually any situation, and throughout the text we will use these properties and rules to calculate probabilities for single events.

Probability Distributions

Though single event probabilities are important, it is often the case that they are inadequate to fully analyze a situation. Instead, it is frequently important to consider many different probabilities at the same time. We can characterize the possible outcomes and their probabilities from an event as a probability distribution.

Consider a fair coin flip. The coin flip has just two possible outcomes - each outcome is mutually exclusive and has a probability of $^1/_2$. We can create a probability distribution for the coin flip by taking each outcome and pairing it with its probability. So we have two pairs: (heads, $^1/_2$) and (tails, $^1/_2$).

If C is the probability distribution of the result of a coin flip, then we can write this as:

C = {(heads, $^1/_2$), (tails, $^1/_2$)}

Likewise, the probability distribution of the result of a fair six-sided die roll is:
D = {(1,$^1/_6$), (2,$^1/_6$), (3, $^1/_6$), (4, $^1/_6$), (5, $^1/_6$), (6, $^1/_6$)}

We can construct a discrete probability distribution for any event by enumerating an exhaustive and mutually exclusive list of possible outcomes and pairing these outcomes with their corresponding probabilities.

We can therefore create different probability distributions from the same physical event. From our die roll we could also create a second probability distribution, this one the distribution of the odd-or-evenness of the roll:

D' = {(odd, $^1/_2$), (even, $^1/_2$)}

In poker, we are almost always very concerned with the contents of our opponents' hands. But it is seldom possible to narrow down our estimate of these contents to a single pair of cards. Instead, we use a probability distribution to represent the hands he could possibly hold and the corresponding probabilities that he holds them. At the beginning of the hand, before anyone has looked at their cards, each player's probability distribution of hands is identical. As the hand progresses, however, we can incorporate new information we gain through the play of the hand, the cards in our own hand, the cards on the board, and so on, to continually refine the probability estimates we have for each possible hand.

Sometimes we can associate a numerical value with each element of a probability distribution. For example suppose that a friend offers to flip a fair coin with you. The winner will collect $10 from the loser. Now the results of the coin flip follow the probability distribution we identified earlier:

C = {(heads, $^1/_2$), (tails, $^1/_2$)}

Since we know the coin is fair, it doesn't matter who calls the coin or what they call, so we can identify a second probability distribution that is the result of the bet:

C' = {(win, $^1/_2$), (lose, $^1/_2$)}

We can then go further, and associate a numerical value with each result. If we win the flip, our friend pays us $10. If we lose the flip, then we pay him $10. So we have the following:

B = {(+$10, $^1/_2$), (-$10, $^1/_2$)}

When a probability distribution has numerical values associated with each of the possible outcomes, we can find the ***expected value (EV)*** of that distribution, which is the value of each outcome multiplied by its probability, all summed together. Throughout the text, we will use the notation **<*X*>** to denote "the expected value of X." For this example, we have:

$<B> = (^1/_2)(+\$10) + (^1/_2)(-\$10)$
$<B> = \$5 + (-\$5)$
$<B> = 0$

Hopefully this is intuitively obvious - if you flip a fair coin for some amount, half the time you win and half the time you lose. The amounts are the same, so you break even on average. Also, the EV of declining your friend's offer by not flipping at all is also zero, because no money changes hands.

For a probability distribution *P*, where each of the *n* outcomes has a value x_i and a probability then p_i *then P's* expected value <*P*> is:

$$< P >= \sum_{i=1}^{n} p_i x_i \qquad \textbf{(1.11)}$$

At the core of winning at poker or at any type of gambling is the idea of maximizing expected value. In this example, your friend has offered you a fair bet. On average, you are no better or worse off by flipping with him than you are by declining to flip.

Now suppose your friend offers you a different, better deal. He'll flip with you again, but when you win, he'll pay you $11, while if he wins, you'll only pay him $10. Again, the EV of not flipping is 0, but the EV of flipping is not zero any more. You'll win $11 when you win but lose $10 when you lose. Your expected value of this new bet B_n is:

$<B_n> = (^1/_2)(+\$11) + (^1/_2)(-\$11)$
$<B_n> = \$0.50$

On average here, then, you will win fifty cents per flip. Of course, this is not a guaranteed win; in fact, it's impossible for you to win 50 cents on any particular flip. It's only in the aggregate that this expected value number exists. However, by doing this, you will average fifty cents better than declining.

As another example, let's say your same friend offers you the following deal. You'll roll a pair of dice once, and if the dice come up double sixes, he'll pay you $30, while if they come up any other number, you'll pay him $1. Again, we can calculate the EV of this proposition.

$<B_d> = (+\$30)(^1/_{36}) + (-\$1)(^{35}/_{36})$
$<B_d> = \$\,^{30}/_{36} - \$\,^{35}/_{36}$

$<B_d> = -\$\,^5/_{36}$ or about 14 cents.

The value of this bet to you is about negative 14 cents. The EV of not playing is zero, so this is a bad bet and you shouldn't take it. Tell your friend to go back to offering you 11-10 on coin flips. Notice that this exact bet is offered on craps layouts around the world.

A very important property of expected value is that it is ***additive.*** That is, the EV of six different bets in a row is the sum of the individual EVs of each bet individually. Most gambling games - most things in life, in fact, are just like this. We are continually offered little coin flips or dice rolls - some with positive expected value, others with negative expected value. Sometimes the event in question isn't a die roll or a coin flip, but an insurance policy or a bond fund. The free drinks and neon lights of Las Vegas are financed by the summation of millions of little coin flips, on each of which the house has a tiny edge. A skillful poker player takes advantage of this additive property of expected value by constantly taking advantage of favorable EV situations.

In using probability distributions to discuss poker, we often omit specific probabilities for each hand. When we do this, it means that the relative probabilities of those hands are unchanged from their probabilities at the beginning of the hand. Supposing that we have observed a very tight player raise and we know from our experience that he raises if and only if he holds aces, kings, queens, or ace-king, we might represent his distribution of hands as:

H = {AA, KK, QQ, AKs, AKo}

The omission of probabilities here simply implies that the relative probabilities of these hands are as they were when the cards were dealt. We can also use the <X> notation for situations where we have more than one distribution under examination. Suppose we are discussing a poker situation where two players A and B have hands taken from the following distributions:

A = {AA, KK, QQ, JJ, AKo, AKs}
B = {AA,KK,QQ}

We have the following, then:

< A, B > : the expectation for playing the distribution A against the distribution B.
< A, AA|B > : the expectation for playing the distribution A against the hand AA taken from the distribution B.
<AA| A, AA|B > : the expectation for playing AA from A against AA from B.
< A, B > = *p*(AA) < A, AA|B > + *p*(KK) < A, KK|B > + *p*(QQ) < A,QO| B> ...and so on.

Additionally, we can perform some basic arithmetic operations on the elements of a distribution. For example, if we multiply all the values of the outcomes of a distribution by a real constant, the expected value of the resulting distribution is equal to the expected value of the original distribution multiplied by the constant. Likewise, if we add a constant to each of the values of the outcomes of a distribution, the expected value of the resulting distribution is equal to the expected value of the original distribution plus the constant.

We should also take a moment to describe a common method of expressing probabilities, ***odds.*** Odds are defined as the ratio of the probability of the event not happening to the probability of the event happening. These odds may be scaled to any convenient base and are commonly expressed as "7 to 5," "3 to 2," etc. ***Shorter*** odds are those where the event is more likely: ***longer*** odds are those where the event is less likely. Often, relative hand values might be expressed this way: "That hand is a 7 to 3 favorite over the other one," meaning that it has a 70% of winning, and so on.

Odds are usually more awkward to use than probabilities in mathematical calculations because they cannot be easily multiplied by outcomes to yield expectation. True "gamblers" often use odds, because odds correspond to the ways in which they are paid out on their bets. Probability is more of a mathematical concept. Gamblers who utilize mathematics may use either, but often prefer probabilities because of the ease of converting probabilities to expected value.

Key Concepts

- The probability of an outcome of an event is the ratio of that outcome's occurrence over an arbitrarily large number of trials of that event.
- A probability distribution is a pairing of a list of complete and mutually exclusive outcomes of an event with their corresponding probabilities.
- The expected value of a valued probability distribution is the sum of the probabilities of the outcomes times their probabilities.
- Expected value is additive.
- If each outcome of a probability distribution is mapped to numerical values, the expected value of the distribution is the summation of the products of probabilities and outcomes.
- A mathematical approach to poker is concerned primarily with the maximization of expected value.

Chapter 2

Predicting the Future: Variance and Sample Outcomes

Probability distributions that have values associated with the elements have two characteristics which, taken together, describe most of the behavior of the distribution for repeated trials. The first, described in the previous chapter, is expected value. The second is ***variance,*** a measure of the dispersion of the outcomes from the expectation. To characterize these two terms loosely, expected value measures how much you will win on average; variance measures how far your specific results may be from the expected value.

When we consider variance, we are attempting to capture the range of outcomes that can be expected from a number of trials. In many fields, the range of outcomes is of particular concern on both sides of the mean. For example, in many manufacturing environments there is a band of acceptability and outcomes on either side of this band are undesirable. In poker, there is a tendency to characterize variance as a one-sided phenomenon, because most players are unconcerned with outcomes that result in winning much more than expectation. In fact, "variance" is often used as shorthand for negative swings.

This view is somewhat practical, especially for professional players, but creates a tendency to ignore positive results and to therefore assume that these positive outcomes are more representative of the underlying distribution than they really are. One of the important goals of statistics is to find the probability of a certain measured outcome given a set of initial conditions, and also the inverse of this - inferring the initial conditions from the measured outcome. In poker, both of these are of considerable use. We refer to the underlying distribution of outcomes from a set of initial conditions as the ***population*** and the observed outcomes as the ***sample.*** In poker, we often cannot measure all the elements of the population, but must content ourselves with observing samples.

Most statistics courses and texts provide material on probability as well as a whole slew of sampling methodologies, hypothesis tests, correlation coefficients, and so on. In analyzing poker we make heavy use of probability concepts and occasional use of other statistical methods. What follows is a quick-and-dirty overview of some statistical concepts that are useful in analyzing poker, especially in analyzing observed results. Much information deemed to be irrelevant is omitted from the following and we encourage you to consult statistics textbooks for more information on these topics.

A commonly asked question *in* poker is "How often should I expect to have a winning session?" Rephrased, this question is "what is the chance that a particular sample taken from a population that consists of my sessions in a game will have an outcome > 0?" The most straightforward method of answering this question would be to examine the probability distribution of your sessions in that game and sum the probabilities of all those outcomes that are greater than zero.

Unfortunately, we do not have access to that distribution - no matter how much data you have collected about your play in that game from the past, all you have is a sample. However, suppose that we know somehow your per-hand expectation and variance in the game, and we know how long the session you are concerned with is. Then we can use statistical methods to estimate the probability that you will have a winning session. The first of these items, expected value (which we can also call the ***mean*** of the distribution) is familiar by now; we discussed it in Chapter 1.

Variance

The second of these measures, ***variance,*** is a measure of the deviation of outcomes from the expectation of a distribution. Consider two bets, one where you are paid even money on a coin flip, and one where you are paid 5 to 1 on a die roll, winning only when the die comes up 6. Both of these distributions have an EV of 0, but the die roll has significantly higher variance. $^1/_{16}$ of the time, you get a payout that is 5 units away from the expectation, while $^5/_6$ of the time you get a payout that is only 1 unit away from the expectation. To calculate variance, we first square the distances from the expected value, multiply them by the probability they occur, and sum the values.

For a probability distribution P, where each of the n outcomes has a value x_i and a probability p_i, then the variance of P, V_p is:

$$V_p = \sum_{i=1}^{n} p_i(x_i - < P >)^2 \quad \textbf{(2.1)}$$

Notice that because each term is squared and therefore positive, variance is always positive. Reconsidering our examples, the variance of the coinflip is:

$$V_C = (^1/_2)(1-0)^2 + (^1/_2)(-1-0)^2$$
$$V_C = 1$$

While the variance of the die roll is:

$$V_D = (^5/_6)(-1-0)^2 + (^1/_6)(5-0)^2$$
$$V_D = 5$$

In poker, a loose-wild game will have much higher variance than a tight-passive game, because the outcomes will be further from the mean (pots you win will be larger, but the money lost in pots you lose will be greater). Style of play will also affect your variance; thin value bets and semi-bluff raises are examples of higher-variance plays that might increase variance, expectation, or both. On the other hand, loose-maniacal players may make plays that increase their variance while decreasing expectation. And playing too tightly may reduce both quantities. In Part IV, we will examine bankroll considerations and risk considerations and consider a framework by which variance can affect our utility value of money. Except for that part of the book, we will ignore variance as a decision-making criterion for poker decisions. In this way variance is for us only a descriptive statistic, not a prescriptive one (as expected value is).

Variance gives us information about the expected distance from the mean of a distribution. The most important property of variance is that it is directly additive across trials, just as expectation is. So if you take the preceding dice bet twice, the variance of the two bets combined is twice as large, or 10.

Expected value is measured in units of expectation per event; by contrast, variance is measured in units of expectation squared per event squared. Because of this, it is not easy to compare variance to expected value directly. If we are to compare these two quantities, we must take the square root of the variance, which is called the ***standard deviation.*** For our dice example, the standard deviation of one roll is $\sqrt{5} \approx 2.23$. We often use the Greek letter a (sigma) to represent standard deviation, and by extension σ^2 is often used to represent variance in addition to the previously utilized V.

$$\sigma = \sqrt{V} \quad \textbf{(2.2)}$$
$$\sigma^2 = V \quad \textbf{(2.3)}$$

The Normal Distribution

When we take a single random result from a distribution, it has some value that is one of the possible outcomes of the underlying distribution. We call this a ***random variable.*** Suppose we flip a coin. The flip itself is a random variable. Suppose that we label the two outcomes 1 (heads) and 0 (tails). The result of the flip will then either be 1 (half the time) or 0 (half the time). If we take multiple coin flips and sum them up, we get a value that is the summation of the outcomes of the random variable (for example, heads), which we call a ***sample.*** The sample value, then, will be the number of heads we flip in whatever the size of the sample.

For example, suppose we recorded the results of 100 coinflips as a single number - the total number of heads. The expected value of the sample will be 50, as each flip has an expected value of 0.5.

The variance and standard deviation of a single flip are:

$$\sigma^2 = (1/2)(1 - 1/2)^2 + (1/2)(0 - 1/2)^2$$
$$\sigma^2 = 1/4$$
$$\sigma = 1/2$$

From the previous section, we know also that the variance of the flips is additiveю

So the variance of 100 flips is 25.

Just as an individual flip has a standard deviation, a sample has a standard deviation as well. However, unlike variance, standard deviations are not additive. But there is a relationship between the two.

For N trials, the variance will be:

$$\sigma^2{}_N = N\sigma^2$$
$$\sigma_N = \sigma\sqrt{N} \qquad \textbf{(2.4)}$$

The square root relationship of trials to standard deviation is an important result, because it shows us how standard deviation scales over multiple trials. If we flip a com once, it has a standard deviation of $1/2$. If we flip it 100 times, the standard deviation of a sample containing 100 trials is not 50, but 5, the square root of 100 times the standard deviation of one flip. We can see, of course, that since the variance of 100 flips is 25, the standard deviation of 100 flips is simply the square root, 5.

The distribution of outcomes of a sample is itself a probability distribution, and is called the ***sampling distribution.*** An important result from statistics, the ***Central Limit Theorem,*** describes the relationship between the sampling distribution and the underlying distribution. What the Central Limit Theorem says is that as the size of the sample increases, the distribution of the aggregated values of the samples converges on a special distribution called the ***normal distribution.***

The normal distribution is a bell-shaped curve where the peak, of the curve is at the population mean, and the tails asymptotically approach zero as the x-values go to negative or positive infinity. The curve is also scaled by the standard deviation of the population. The total area under the curve of the normal distribution (as with all probability distributions) is equal to 1, and the area under the curve on the interval $[x_1, x_2]$ is equal to the probability that a particular result will fall between x_1 and x_2. This area is marked region A in Figure 2.1.

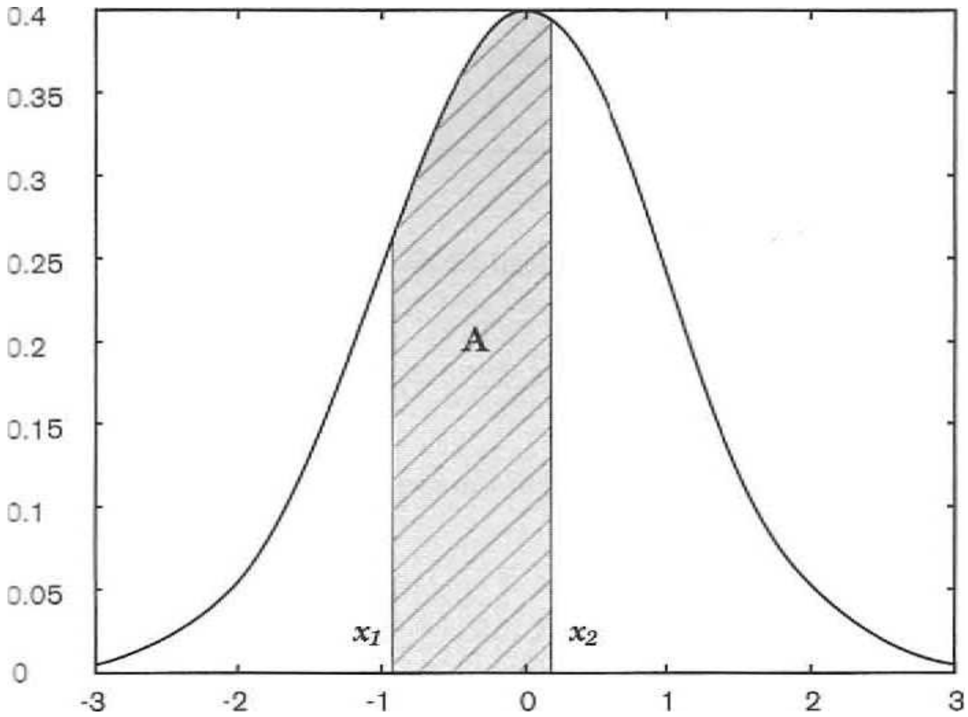

Figure 2.1. Std. Normal Dist, A = p(event between x, and

A little less formally, the Central Limit Theorem says that if you have some population and take a lot of big enough samples (how big depends on the type of data you're looking at), the outcomes of the samples will follow a bell-shaped curve around the mean of the population with a variance that's related to the variance of the underlying population.

The equation of the normal distribution function of a distribution with mean μ and standard deviation σ is:

$$N(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad \textbf{(2.5)}$$

Finding the area between two points under the normal distribution curve gives us the probability that the value of a sample with the corresponding mean and variance will fall between those two points. The normal distribution is symmetric about the mean, so $^1/_2$ of the total area is to the right of the mean, and $^1/_2$ is to the left. A usual method of calculating areas under the normal curve involves creating for each point something called a ***z-score,*** where $z = (x - \mu)/\sigma$. This z-score represents the number of standard deviations that the outcome ***x*** is away from the mean.

$$z = (x - \mu)/\sigma \qquad \textbf{(2.6)}$$

We can then find something called the ***cumulative normal distribution*** for a z-score z**,** which is the area to the left of z under the curve (where the mean is zero and the standard deviation is 1). We call this function $\Phi(z)$. See Figure 2.2

If z is a normalized z-score value, then the cumulative normal distribution function for ***z*** is:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{\rightarrow\infty}^{z} exp\,(-\frac{x^2}{2}) \qquad \textbf{(2.7)}$$

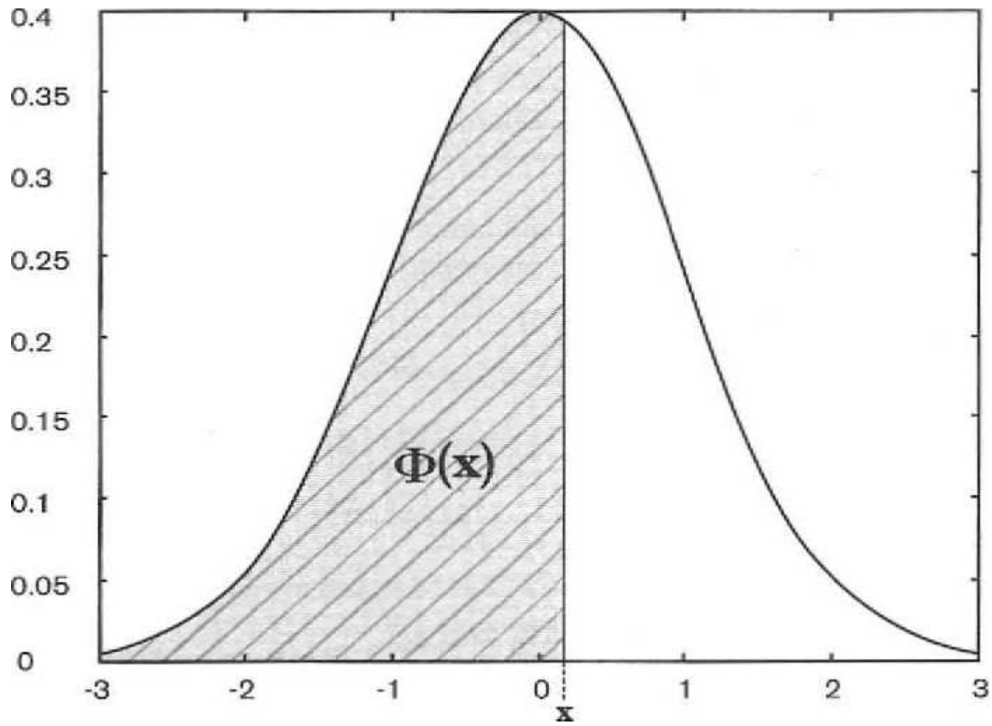

Figure 2.2, Cumulative normal distribution

Finding the area between two values x_1 and x_2 is done by calculating the z-scores z_1 and z_2 for x_1 and x_2, finding the cumulative normal distribution values $\Phi(z_1)$ and $\Phi(z_2)$ and subtracting them.

If $\Phi(z)$ is the cumulative normal distribution function for a z-score of z, then the probability that a sample taken from a normal distribution function with mean μ and standard deviation σ will fall between two z-scores x_1 and x_2 is:

$$p = \Phi\left(\frac{x_1-\mu}{\sigma}\right) - \Phi\left(\frac{x_2-\mu}{\sigma}\right) \qquad \textbf{(2.8)}$$

Statisticians have created tables of $\Phi(z)$ values and spreadsheet programs can derive these values as well. Some key values are well known to students of statistics. The area between z = -1 and z = +1 is approximately 0.68; the area between z = -2 and z = +2 is approximately 0.955, and the area between z = -3 and z= +3 is approximately 0.997.

These values mean that the value of a sample from a normal distribution will fall:

Between $(\mu - \sigma)$ and $(\mu + \sigma)$ of the mean 68% of the time.
Between $(\mu - 2\sigma)$ and $(\mu + 2\sigma)$ of the mean 95.5% of the time.
Between $(\mu - 3\sigma)$ and $(\mu + 3\sigma)$ of the mean 99.7% of the time.

An example may help to make this clear. Let us modify the parameters of the die roll game we discussed before. A die is thrown, and the player loses 1 unit on 1-5, but wins 6 units on a roll of

6. We'll call this new game D_2.

The expectation of the game is:

$< D_2> = (\,^5/_6\,)(-1) + (\,^1/_6\,)(6)$
$< D_2 > = \,^1/_6$ units /trial

When the player wins, he gains six units. Subtracting the mean value of $^1/_6$ from this outcome, we obtain:

$Vwin = (6 - \,^1/_6\,)^2$
$Vwin = (^{35}/_6\,)^2$

Likewise, when he loses he loses one unit. Subtracting the mean value of $^1/_6$ from this, we have:

$Vlose = (-1 - \,^1/_6\,)^2$
$Vlose = (-\,^7/_6\,)^2$

The variance of the game is:

$V_{D2} = p(\text{win})(V_{win}) + p(\text{lose})(V_{lose})$
$V_{D2} = (\,^1/_6\,)(\,^{35}/_6\,)^2 + (\,^5/_6\,)(-\,^7/_6\,)^2$
$V_{D2} \approx 6.806$ units^2/trial^2

Suppose we toss the die 200 times. What is the chance that the player will win overall in 200 tosses? Win 40 units or more? Win 100 units or more?

We can solve this problem using the techniques we have just summarized. We first calculate the standard deviation, sometimes called the standard error, of a sample of 200 trials. This will be:

$\sigma = \sqrt{V} = \sqrt{6.806}$=2.61 units/trial

Applying equation 2.4 we get:

$\sigma_N = \sigma\sqrt{N}$
$\sigma_{200} = 2.61\sqrt{200} = 36.89 \text{ units/200 trials}$

For 200 trials, the expected value, or mean (μ) of this distribution is $^1/_6$ units/trial times 200 trials or 33.33 units. Using Equation 2.6, we find the z-score of the point 0 as:

$z_x = (x-\mu)/\sigma$ where x=0
$z_0 = (0-33.33)/36.89$
$z_0 = -33.33/36.89$
$z_0 = -0.9035$

Consulting a probability table in a statistics textbook, we find that the probability that an observed outcome will lie to the left of this z-score, Φ(-0.9035) is 0.1831, Hence, there is an 18.31% chance that the player will be behind after 200 trials.

To find the chance of being 40 units ahead, we find that point's z-score:

z_{40} = (40-33.33)/(36.89) = 0.1807
Φ(0,1807) = 0,5717

But Φ(0.1807) is the probability that the observed outcome lies to the left of 40 units, or that we lose at least 40 units. To find the probability that we are to the right of this value, or are ahead 40 units, we must actually find 1 - Φ(0.1807).

$$1 - \Phi(0.1807) = 1 - 0.5717 = 0.4283$$

So there is a 42.83% chance of being ahead at least 40 units after 200 tosses.

And similarly for 100 units ahead:

z_{100} = (100-33.33)/(36.89) = 1.8070

From a probability table we find that Φ(1.8070) = 0.9646. Thus, for:

$p = 1 - \Phi(1.8070)$
$p = 0.0354$

The probability of being 100 units ahead after 200 tosses is 3.54%.

These values, however, are only approximations; the distribution of 200-roll samples is not quite normal. We can actually calculate these values directly with the aid of a computer. Doing this yields:

	Direct Calculation	**Normal Approx**
Chance of being ahead after 200 trials:	81.96%	81.69%
Chance of being ahead at least 40 units:	40.46%	42.83%
Chance of being ahead at least 100 units:	4.44%	3,54%

As you can see, these values have slight differences. Much of this is caused by the fact that the direct calculations have only a discrete amount of values. 200 trials of this game can result in outcomes of +38 and +45, but not +39 or +42, because there is only one outcome possible from winning, a +6. The normal approximation assumes that all values are possible.

Using this method, we return to the question posed at the beginning of the chapter: "How often will I have a winning session?" To solve this problem, we need to know the player's expectation per hand, his variance per hand and the size of his sessions. Consider a player whose win rate in some particular game is approximately 0.015 bets per hand, and whose standard deviation over the same interval is 2 bets per hand. Assume this player plans to play sessions of 300 hands (these would be roughly nine-hour sessions in live poker; perhaps only two to three hours online). How often should this player expect to win (have a result greater than 0) in a session?

First we can calculate the expected value of a sample, μ_N:

$\mu_N = N\mu$
$\mu_N = (300)(0.015)$
$\mu_N = 4.5$

Second, we calculate his standard deviation for a 300-hand session.

$\sigma_N = \sigma\sqrt{N}$
$\sigma_{300} = (2)(\sqrt{300})$
$\sigma_{300} = 34.6$

Next, we find the z-score and $\Phi(z)$ for this value:

$z_x = (x-\mu_N)/\sigma$
$z_0 = (0 - 4.5)/34.6$
$z_0 = -0.1299$

From a probability table we find:

$\Phi(-0.1299) = 44.83\%$
$p = 1 - \Phi(-0.1299)$
$p = l- 0.4483$
$p = 0.55171$

This indicates that the player has a result of 0 or less 44.83% of the time - or correspondingly has a winning session 55.17% of the time. In reality, players may change their playing habits in order to produce more winning sessions for psychological reasons, and in reality "win rate" fluctuates significantly based on game conditions and for other reasons. But a player who simply plays 300 hands at a time with the above performance metrics should expect to be above zero just over 55% of the time.

A more extreme example: A player once told one of the authors that he had kept a spreadsheet where he marked down every AQ vs. AK all-in preflop confrontation that he saw online over a period of months and that after 2,000 instances, they were running about 50-50.

How likely is this, assuming the site's cards were fair and we have no additional information based on the hands that were folded?

Fust, let's consider the variance or standard deviation of a single confrontation.

AK vs. AQ is about 73.5% all-in preflop (including all suitedness combinations). Let's assign a result of 1 to AK winning and 0 to AQ winning. Then our mean is 0.735. Calculating the variance of a single confrontation:

$V = (0.735)(1- 0.735)^2 + (0.265)(0- 0.735)^2$
$V = 0.1948$
$\sigma = 0.4413$

The mean of 2,000 hands is:

$\mu_N = N\mu$
$\mu_N = (2000)(0.735)$
$\mu_N = 1470$

For a sample of 2,000 hands, using Equation 2.4, the standard deviation is:

$\sigma_N = \sigma\sqrt{N}$

$\sigma_{2000} = (0.4413)(\sqrt{2000})$
$\sigma_{2000} = 19.737$

The result reported here was approximately 50% of 2000, or 1000, while the sample mean would be about 1470 out of 2000. We can find the z-score using Equation 2.6:

$z_x = (x-\mu_N)/\sigma$
$z_{1000} = (1000-1470)/(19.737)$
$z_{1000} = -23.815$

The result reported in this case was 1000 out of 2000, while the expected population mean would be 1470 out of 2000. This result is then about 23.815 standard deviations away from the mean. Values this small are not very easy to evaluate because they are so small - in fact, my spreadsheet program calculates $\Phi(-23.815)$ to equal exactly zero. Suffice it to say that this is a staggeringly low probability.

What's likely is that in fact, this player was cither exaggerating, outright lying, or perhaps made the common mistake of forgetting to notice when AK beat AQ, because that is the "expected" result. This is related to a psychological effect known as "perception bias" - we tend to notice things that are out of the ordinary while failing to notice things that are expected. Or perhaps the online site was in fact "rigged." When this example was posed to a mailing list, the reaction of some participants was to point to the high variance of poker and assert that a wide range of outcomes is possible in 2000 hands. However, this confuses the altogether different issue of win/loss outcomes in 2000 poker hands (which docs have quite high variance) with the outcome of a single poker hand (which has far less). The variance of a single hand in terms of big bets in most forms of poker is much higher than the variance of the winner of an all-in preflop confrontation. One lesson of this example is not to confuse the high variance of poker hand dollar outcomes with the comparatively low variance of other types of distributions.

When we play poker, many random events happen. We and the other players at the table are dealt random cards taken from a distribution that includes all two-card combinations. There is some betting, and often some opportunity to either change our cards, or to have the value of our hand change by the dealing of additional cards, and so on. Each hand results in some outcome for us. whether it is whining a big pot, stealing the blinds, losing a bet or two, or losing a large pot. This outcome is subject to all the smaller random events that occur within the hand, as well as events that are not so random - perhaps we get a tell on an opponent that enables us to win a pot we would not otherwise have, or save a bet when we are beaten. Nevertheless, the power of the Central Limit Theorem is that outcomes of individual hands function approximately as though they were random variables selected from our "hand outcomes" distribution. And likewise, outcomes of sessions, and weeks, and months, and our whole poker career, behave as though they were appropriately-sized samples taken from this distribution.

The square root relationship of trials to standard deviation makes this particularly useful, because as the number of trials increases, the probability that our results will be far away from our expected value in relative terms decreases.

Assume we have a player whose distribution of hand outcomes at a particular limit has a mean of S75 per 100 hands, with a variance of \$6,400 per hand. If we sample different numbers of hands from this player's distribution, we can see how the size of the sample impacts the dispersion of the results. We know that the probability that this player's results for a given sample will be between the mean minus two standard deviations and the mean plus two standard deviations is 95.5%. We will identify for each sample size:

- The mean μ_N
- The standard deviation σ
- The two endpoints of the 95.5% probability interval.

Hands	μ_N	σ	Lower endpoint	Higher endpoint
100	$75	$800.00	($1,525.00)	$1,675.00
500	$375.00	$1,788.85	($3,202.71)	$3,952.71
1,000	$750.00	$2,529.82	($4,309.64)	$5,809.64
5,000	$3,750.00	$5,656.85	($7,563.71)	$15,063.71
25,000	$18,750.00	$12,649.11	($6,548.22)	$44,048.22
50,000	$37500.00	$17888.54	$1,722.91	$73,277.09
100,000	$75,000.00	$25,298.22	$24,403.56	$125,596.44
1,000,000	$750,000.00	$80,000.00	$590.000.00	$910,000.00

As you can see, for smaller numbers of hands, outcomes vary widely between losses and wins. However, as the size of the sample grows, the relative closeness of the sample result becomes larger and larger - although its absolute magnitude continues to grow. Comparing the standard deviation of one million hands to the standard deviation for one hundred hands, the size of a standard deviation is a hundred times as large in absolute terms, but more than a hundred times smaller relative to the number of hands. This is the ***law of large numbers*** at work; the larger the sample, the closer on a relative basis the outcomes of the sample will be.

Key Concepts

- Variance, the weighted sum of the squares of the distance from the mean of the outcomes of a distribution, is a valuable measurement of dispersion from the mean.
- To compare variance to expected value, we often use its square root, standard deviation.
- Variance is additive across trials. This gives rise to the square root relationship between the standard deviation of a sample of multiple trials and the standard deviation of one trial.
- The Central Limit Theorem tells us that as the size of a sample increases, the distribution of the samples behaves more and more like a normal distribution. This allows us to approximate complex distributions by using the normal distribution and to understand the behavior of multiple-trial samples taken from a known distribution.

Chapter 3

Using All the Information:
Estimating Parameters and Bayes' Theorem

In the last chapter, we described some statistical properties of valued probability distributions, as well as the relationship of samples taken from those distributions to the normal distribution. However, throughout the chapter, we simply assumed in examining the sampling distribution that we knew the parameters of the underlying distribution. But in real life, we often don't know these things. Even for simple things such as coin flips, the "true" distribution of outcomes is that the coin is very slightly biased in one direction or the other. A die has tiny imperfections in its surface or composition that make it more likely to land on one side or another. However, these effects are usually small, and so using the "theoretical" coin which is truly 50-50 is a reasonable approximation for our purposes. Likewise, we generally assume for the purposes of poker analysis that the deck is fair and each card is random until it is observed.

We can mitigate real-world difficulties with distributions that we can reasonably approximate (such as coin flips or die rolls). Other types of distributions, however, pose much more difficult problems. In analyzing poker results, we are often interested in a distribution we discussed in the last chapter - per hand won/loss amounts. When playing in a casino, it would be quite difficult and time-consuming to record the results of every hand - not to mention it might draw unwanted attention. The process is somewhat easier online, as downloadable hand histories and software tools can automate the process. But even if we have all this data, it's just a sample. For poker hands, the probabilities of the underlying distribution won't be reflected in the observed data unless you have a really large amount of data.

We can get around this to some extent by using the subject matter of the last chapter. Suppose that we could get the mean and variance of our hand outcome distribution. Then we could find the sampling distribution and predict the aggregate outcomes from playing different numbers of hands. We can't predict the actual outcome of a particular future sample, but we can predict the distribution of outcomes that will occur.

Now the problem is to try to infer the population mean and variance from the sample mean and variance. We will examine two approaches to this process. The first is the approach of classical statistics, and the second is an approach that utilizes the primary topic of this chapter, Bayes' theorem. The first approach takes as one of its assumptions that we have no information other than the sample about the likelihood of any particular win rate. The second approach postulates a distribution of win rates that exists outside of our particular sample that can be used to refine our estimates of mean and variance for the population distribution.

Estimating Parameters: Classical Statistics

Suppose that we have a player who has played a total of 16,900 hands of limit poker. Normalizing his results to big bets (BB) in order to account for different limits he has played, he has achieved a win rate of $\bar{x}$ = 1.15 BB/100 hands with a standard deviation of s = 2.1 BB/hand. Here instead of μ and σ, which represent population parameters, we use $\bar{x}$ and s, which are sample parameters. Assuming that he plans to continue to play in a similar mix of games with similar lineups of players, what can we say about his "true" win rate μ in the games he has played? We assume in this section that we have no other information about the likelihood of various win rates that might be possible; all win rates from -1 BB/hand to +1BB/hand are deemed to be equally probable.

First of all, it's important to note that we only have a sample to work with. As a result, there will be uncertainty surrounding the value of his win rate. However, we know that the sampling distribution of 16,900 hands of limit poker will be approximately normal because of the Central Limit Theorem. The observed standard deviation $s = 2.1$ BB/h is a reasonable estimate of the standard deviation of the population, particularly because the sample is relatively large. We can use these facts to obtain the ***maximum likelihood estimate*** of the population mean.

Consider all possible win rates. For each of these win rates, there will be a corresponding sampling distribution, a normal curve with a mean of μ and a standard deviation σ_N. The peak of each of these normal curves will be at $x = \mu$, and all the other points will be lower. Now suppose that we assume for a moment that the population mean is in fact some particular μ. The height of the curve at $x = \overline{x}$ will be associated with the probability that the observed rumple mean would have been the result of the sample. We can find this value for all possible values of μ. Since all these normal curves have the same standard deviation σ_N, they will all be identical, but shifted along the X-axis, as in Figure 3.1.

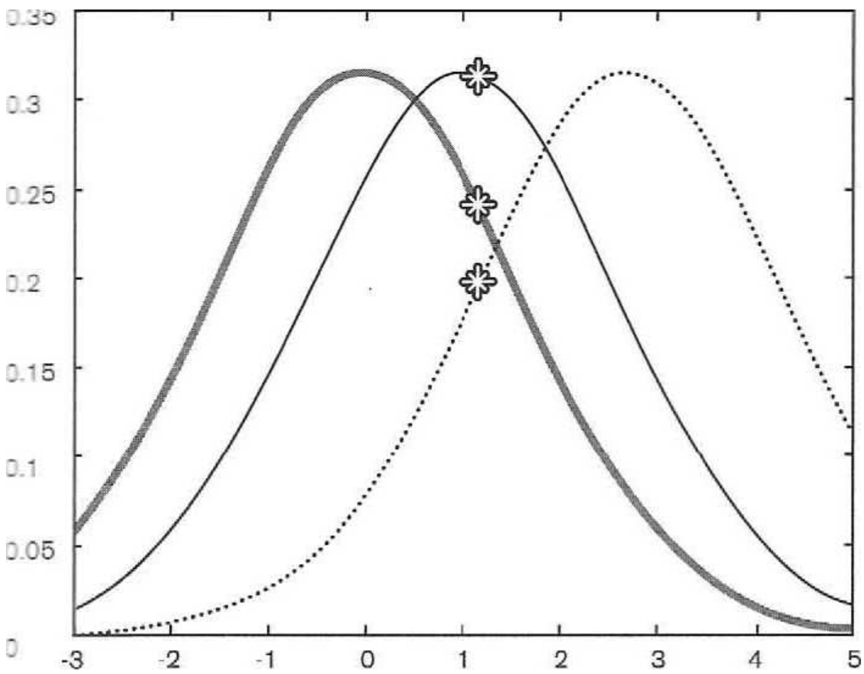

Figure 3.1, Shifted normal distributions (labeled points atx=1.15)

Since the peak of the curve is the highest point, and the observed value $\overline{x}$ is the peak when $\mu = \overline{x}$, this means that $\overline{x}$ — 1.15 BB/l00h is the maximum likelihood estimate of the mean of the distribution. This may seem intuitive, but we will see when we consider a different approach to this problem that the maximum likelihood estimate does not always have to equal the sample mean, if we incorporate additional information.

Knowing that the single win rate that is most likely given the sample is the sample mean is a useful piece of data, but it doesn't help much with the uncertainty. After all, our hypothetical player might have gotten lucky or unlucky. We can calculate the standard deviation of a sample of 16,900 hands and we can do some what-if analysis about possible win rates.

Suppose we have a sample N that consists of 16,900 hands taken from an underlying distribution with a mean, or win rate, of 1.15 BB/l00h and a standard deviation of 2.1 BB/h.

Then, using equation 2.4:

$\sigma_N = \sigma\sqrt{N}$
$\sigma_{16,900} = (2.1\ BB/h)(\sqrt{16{,}900}\text{ hands})$
$\sigma_{16,900} = 273\ BB$, *so*
$\sigma_N/100h = 273/169 \approx 1.61$

The standard deviation of a sample of this size is greater than the win rate itself. Suppose that we knew that the parameters of the underlying distribution were the same as the observed ones. If we took another sample of 16,900 hands, 32% of the time, the observed outcome of the 16,900 hand sample would be lower than -0.46 BB/100 or higher than 2.76 BB/100.

This is a little troubling. How can we be confident in the idea that the sample represents the true population mean, when even if that were the case, another sample would be outside of even those fairly wide bounds 32% of the time? And what if the true population mean were actually, say, zero? Then 1.15 would fall nicely into the one-sigma Interval. In fact, it seems like we can't tell the difference very clearly based on this sample between a win rate of zero and a win rate of 1.15 BB/100.

What we can do to help to capture this uncertainty is create a ***confidence interval.*** To create a confidence interval, we must first decide on a level of tolerance. Because we're dealing with statistical processes, we can't simply say that the probability that the population mean has some certain value is zero - we might have gotten extremely lucky or extremely unlucky. However, we can choose what is called a ***significance level.*** This is a probability value that represents our tolerance for error, Then the confidence interval is the answer to the question, "What are all the population mean values such that the probability of the observed outcome occurring is less than the chosen significance level?"

Suppose that for our observed player, we choose a significance level of 95%. Then we can find a confidence level for our player. If our population mean is μ, then a sample of this size taken from this population will be between $(\mu - 2\sigma)$ and $(\mu + 2\sigma)$ 95% of the time. So we can find all the values of μ such that the observed value $\bar{x} = 1.15$ is between these two boundaries.

As we calculated above, the standard deviation of a sample of 16,900 hands is 1.61 units/100 hands:

$\sigma_N = \sigma\sqrt{N}$
$\sigma = (2.1\ \text{BB/h})(\sqrt{16{,}900})$
$\sigma =$ 273 BB per 16,900 hands
$\sigma/100h =$ 273 BB/169 = 1.61

So as long as the population mean satisfies the following two equations, it will be within the confidence interval:

$(\mu - 2\sigma) < 1.15$
$(\mu + 2\sigma) > 1.15$
$(\mu - 2\sigma) < 1.15$
$\mu - (2)(1.61) < 1.15$
$\mu < 4.37$
$(\mu + 2\sigma) > 1.15$
$\mu + (2)(1.61) > 1.15$

$\mu > -2.07$
So a 95% confidence interval for this player's win rate (based on the 16,900 hand sample he has collected) is [-2.07 BB/100, 4.37 BB/100].

This does not mean that his true rate is 95% likely to lie on this interval. This is a common misunderstanding of the definition of confidence intervals. The confidence interval is all values that, if they were the true rate, then the observed rate would be inside the range of values that would occur 95% of the time. Classical statistics doesn't make probability estimates of parameter values - in fact, the classical view is that the true win rate is either in the interval or it isn't, because it is not the result of a random event. No amount of sampling can make us sure or unsure as to what the parameter value is. Instead, we can only make claims about the likelihood or unlikelihood that we would have observed particular outcomes if a parameter had a particular value.

The maximum likelihood estimate and a confidence interval are useful tools for evaluating what information we can gain from a sample. In this case, even though a rate of 1.15 BB/100 might look reasonable, concluding that this rate is close to the true one is a bit premature. The confidence interval can give us an idea of how wide the range of possible population rates might be. However, if pressed, the best single estimate of the overall win rate is the maximum likelihood estimate, which is the sample mean of 1.15 BB/100 in this case.

To this point, we have assumed that we had no information about the likelihood of different win rates - that is, that our experimental evidence was the only source of data about what win rate a player might have. But in truth, some win rates are likelier than others, even before we take a measurement. Suppose that you played, say, 5,000 hands of casino poker and in those hands you won 500 big bets, a rate of 10 big bets per 100 hands. In this case, the maximum likelihood estimate from the last section would be that your win rate was exactly that -10 big bets per 100 hands,

But we do have other information. We have a fairly large amount of evidence, both anecdotal and culled from hand history databases and the like that indicates that among players who play a statistically significant number of hands, the highest win rates are near 3-4 BB/100. Even the few outliers who have win rates higher than this do not approach a rate of 10 BB/100. Since this is information that we have before we even start measuring anything, we call it a priori information.

In fact, if we just consider any particular player we measure to be randomly selected from the universe of all poker players, there is a probability distribution associated with his win rate. We don't know the precise shape of this distribution, of course, because we lack observed evidence of the entire universe of poker players. However, if we can make correct assumptions about the underlying *a priori* distribution of win rates, we can make better estimates of the parameters based on the observed evidence by combining the two sets of evidence.

Bayes' theorem

In Chapter 2, we stated the basic principle of probability (equation 1.5).

$$p(A \cap B) = p(A)p(B|A)$$

In this form, this equation allows us to calculate the joint probability of A and B from the probability of A and the conditional probability of B given A. However, in poker, we are often most concerned with calculating the conditional probability of B given that A has already-occurred - for example, we know the cards in our own hand (A), and now we want to know how this information affects the cards in our opponents hand *(B).* What we are looking for is the conditional probability of *B* given *A.*

So we can reorganize equation 1.5 to create a formula for conditional probability.

This is the equation we will refer to as ***Bayes' theorem:***

$$p(B|A) = \frac{p(A \cap B)}{p(A)} \qquad \textbf{(3.1)}$$

Recall that we defined $\overline{B}$ as the complement of B in Chapter 1; that is:

$p(\overline{B}) = 1 - p(B)$
$p(B) + p(\overline{B}) = 1$

We already have the definitions:

$p(A \cap B) = p(A)p(B|A)$

Since we know that *B* and $\overline{B}$ sum to 1, *p(A)* can be expressed as the probability of *A* given *B* when *B* occurs, plus the probability of *A* given *B* when *B* occurs.

So we can restate equation 3.1 as:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A|B)p(B) + p(A|\overline{B})p(\overline{B})} \qquad \textbf{(3.2)}$$

In poker, Bayes' theorem allows us to refine our judgments about probabilities based on new information that we observe. In fact, strong players use Bayes' theorem constantly as new information appears to continually refine their probability assessments; the process of Bayesian inference is at the heart of reading hands and exploitive play, as we shall see in Part II.

A classic example of Bayes' theorem comes from the medical profession. Suppose that we have a screening process for a particular condition. If an individual with the condition *is* screened, the screening process properly identifies the condition 80% of the time. If an individual without the condition is screened, the screening process improperly identifies him as having the condition 10% of the time. 5% of the population (on average) has the condition.

Suppose, then, that a person selected randomly from the population is screened, and the screening returns a positive result. What is the probability that this person has the condition (absent further screening or diagnosis)?

If you answered "about or a little less than 80%," you were wrong, but not alone. Studies of doctors' responses to questions such as these have shown a perhaps frightening lack of understanding of Bayesian concepts.

We can use Bayes' theorem to find the answer to this problem as follows:
A = the screening process returns a positive result.
B = the patient has the condition.

Then we are looking for the probability $p(B|A)$ and the following are true:
$p(B|A))$ = 0.8

(if the patient has the condition, the screening will be positive 80% of the time)
$p(A|\overline{B})$ = 0.1

(if the patient doesn't have the condition, the result will be positive 10% of the time)
p(B) = 0.05 (5% of all people have the condition)
$p[\overline{B})$ = 0.95 (95% of all people don't have the condition)

And using Equation 3.2, we have:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A|B)p(B) + p(A|\overline{B})p(\overline{B})}$$

$$p(B|A) = \frac{(0.8)(0.05)}{(0.8)(0.05) + (0.1)(0.95)}$$

$p(B|A) \approx$ 29.63%

As you can see, the probability that a patient with a positive screening result actually has the condition is much lower than the 80% "accuracy" the test has in identifying the condition in someone who has it. Testing with this screening process would need to be followed up by additional testing to confirm the presence of the condition.

But rather than just plugging into an equation, let's discuss what's happening here by considering a population of *100,000* people, all of whom are tested using this screening process.

Of these 100,000 people:
5,000 actually have the condition. (5% of the population)
95,000 actually do not have the condition. (95% of the population)

Of the *5,000* who have the condition:
4,000 will test positive. (80% of those with the condition)
1,000 will test negative. (20% false negatives)

Of the 95,000 who have the condition:
9,500 will test positive. (10% false positives)
85,500 will test negative. (90% of those without the condition)

Now our question was: given that a person has tested positive, how likely are they to have the condition? Out of our 100,000 tested individuals, a total of 13,500 tested positive. Of those, only 4,000 actually have the condition.

$p(B|A)$= 4,000/13,500
$p(B|A) \approx$ 29.6%

Additionally, we can see that by increasing the accuracy of the test, either by making it more accurate in identifying the condition in those who have it, or in producing fewer false positive identifications of the condition in those who do not, we can increase the conditional probability that someone who tests positive has the condition. Note that this does not increase the chance that somebody actually has the condition - we would not want to increase that! - but rather decreases the number of people who incorrectly test positive, and potentially have to incur the stress of incorrectly believing they have a condition.

Suppose that we increase the effectiveness of the test to always identify those who have the condition, while the false positive rate remains the same. Then, using Bayes' theorem:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A|B)p(B) + p(A|\overline{B})p(\overline{B})}$$

$$p(B|A) = \frac{(1)(0.05)}{(1)(0.05) + (0.1)(0.95)}$$

$$p(B|A) \approx 34.5\%$$

Likewise, if we hold the 80% identification rate constant and drop the rate of false positives from 10% to 6%, we obtain:

$$p(B|A) = \frac{(0.8)(0.05)}{(0.8)(0.05) + (0.06)(0.95)}$$

$$p(B|A) \approx 41.2\%$$

The key to the application of Bayes' theorem is the existence of a prior probability and obtaining new information. In the above example, we began with a prior probability that our patient had the condition of 5%. After the screening, the positive result was new information that allowed us to revise the probability for the patient - in this case upward. This process is called ***Bayesian inference*** and is critical to successful poker play.

There are countless ways in which Bayesian inference can be used; in fact, many players employ this process unconsciously all the time. Consider a situation in a tournament when a player is at a table that breaks very early on. He is moved to a new table where the player on his right has already amassed a stack seven times as large as the starting stacks, while most of *the* other players in the tournament are still near their starting stacks. Most players would conclude that two scenarios are likely: the player with the huge stack is either an extremely loose and aggressive player or he happens to simply be an ordinary player who got extremely lucky. The natural inclination to believe that he is more likely to be loose and aggressive than very lucky is a direct consequence of Bayes' theorem, whether the player from the broken table is aware of this concept or not.

We can quantify this effect using the formal version of Bayes' theorem, and this can lead us to making stronger plays even without solid information. Consider the following common situation:

A new player sits down in the game. Using all the observational information available to us, which might include stereotyping the player's ethnicity, gender, manner, wardrobe, personal appearance, and so on, we conclude that he is 10% likely to be a "maniac" who will raise 80% of his hands from the cutoff and 90% likely to be a tight player who will raise 10% of his hands from that seat. On the first hand he plays, he raises from the cutoff. (Assume there is no posting.) Now what is the probability- that he is a maniac?

We can use Bayes' theorem to solve this problem, but we invite you to estimate this probability before continuing. We believe that testing intuition is one of the best ways to refine our ability to estimate accurately, a skill chat is invaluable not only in poker but in life.

A = The opponent will raise the first hand he plays from the cutoff.

B= The opponent is a maniac.

$p(A|B)$ = 0.8 (if the player is a maniac, he will raise 80% of the time)
$p(A|\overline{B})$ = 0.1 (if the player is not a maniac, he will raise 10% of the time)
$p(\underline{B})$ = 0.1 (10% of the time, he is a maniac *a priori)*
$p(\overline{B})$= 0.9 (90% of the time, he is not a maniac *a priori)*

Applying Bayer's theorem again:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A|B)p(B) + p(A|\overline{B})p(\overline{B})}$$

$$p(B|A) = \frac{(0.8)(0.1)}{(0.8)(0.1) + (0.1)(0.9)}$$

$$p(B|A) \approx 47.1\%$$

So simply by observing this player raising the first hand, we can adjust the probability of this slayer being a maniac from just 10% to 47% immediately. If the player raises the first two hands (assuming the same inference for the seat next to the cutoff], this probability moves to nearly 87%! Of course, these probabilities are subject to the accuracy of our original assumptions - in reality, there are not just two types of players, and our probability estimates are probably not so crisp about what type of player he is.

One tendency among players is to delay characterizing and adjusting to a player's play until gaining a little more information, by observing some hands or the like. But this view is overly passive in our view; maximizing EV means taking advantage of all the information we have at our disposal and not necessarily waiting for confirmation that the information is reliable before trying to take advantage of it. The error that these players are making is that they- do not realize the power of the information they have gained. It is worth noting that many players, even players who do not play well often make this adjustment, or a similar one, intuitively. But beware! This adjustment is open to exploitation by players who will sit down in a game and play very differently from their usual style in an attempt to induce significant adjustments by players in the game.

Strong players use Bayes' theorem constantly as new information appears to continually refine their probability assessments; the process of Bayesian inference is at the heart of reading hands and exploitive play, as we shall see in Part II. But even away from the table, Bayes' theorem can allow us to make more informed conclusions about data. To see an example of this, we return to the topic of win rates.

Estimating Parameters: Bayesian Statistics

Recall earlier in this chapter we discussed a player who had played a total of 16,900 hands of limit poker with the following observed statistics:
Win rate of $\overline{x}$ = 1.15 BB/100 hands
Standard deviation of s = 2.1 BB/hand.

We were concerned with some of the statements that we could make about his "true" win rate based on these observations. Using methods from classical statistics, we found that his maximum likelihood estimator for win rate was 1.15 BB/100 hands and his 95% confidence interval was [-2.07 BB/100, 4.37 BB/100]. These statements relied on an assumption that we had no additional information about the distribution of this player's win rates.

However, suppose that we guess at a distribution of win rates and try to apply Bayes' theorem to the distribution of win rates and this player's results in order to produce a more accurate estimate. To do this, we must first hypothesize an overall distribution of win rates for this player. Let's assume that he is pulled randomly from the population of regular poker players. What is the shape of this distribution, which is called a ***prior probability distribution?***

It's pretty unclear at a glance what this distribution looks like - after all, we do not have access to the results and records of the broad population. But we can simplify and estimate, hoping that our distribution of win rates will be close enough to the truth that we can gain from incorporating it into our estimates. Assuming that our player plays some mixture of lower mid-limit games such as $10-$20 to $30- $60, we can estimate the total rake paid by the game as about $3-$4 per hand, or perhaps 0.1 BB/h. Dividing this amongst all the players approximately equally, we obtain a net rake effect on everyone's win rate of about 0.01 BB/h, or 1 BB/100.

The mean of the distribution of all players' win rates, then, is equal to this value, as this is the net flow of money out of the game. Suppose that we have a roughly normal distribution of win rates, let's just estimate a standard deviation (of win rates) of about 0.015 BB/h. This would lead to a population where 68% of players have a rate between -2.5 BB/100 and +0.5 BB/100 and where 95% of players would have a rate between -4 bb/100 and +2 BB/100. This might square with your intuition - if not, these numbers can be tweaked to reflect different assumptions without changing the underlying process.

To simplify the computation, instead of using the continuous normal distribution, we will create a discrete distribution that roughly mirrors our assumptions. We assume that the underlying distribution of all poker players is as follows:

Win Rate	**% of players with this win rate**
-5 BB/100	**0.25%**
-4 BB/100	**2%**
-3 BB/100	**8%**
-2 BB/100	**20%**
-1 BB/100	**39.5%**
0 BB/100	**20%**
+1 BB/100	**8%**
+2 BB/100	**2%**
+3 BB/100	**0.25%**

Now that we have an *a priori* distribution of win rates, we can apply Bayes' theorem to this problem. For each win rate, we calculate:
A = the chance of a win rate of 1.15 being observed.
B = the chance that this particular win rate is the true one *(a priori).*

We cannot directly find the probability of a particular win rate being observed (because the normal is a continuous distribution). We will instead substitute the probability of a win rate between 1.14 and 1.16 being observed as a proxy for this value. Recall that the standard deviation of a sample of this size was 1.61 bets. $p(A|\overline{B})$ is calculated by simply calculating the weighted mean of $p(A|B)$ excluding the current row. From a probability chart:

Win Rate	$\boldsymbol{p(\overline{B})}$	$\boldsymbol{p(A\vert B)}$	$\boldsymbol{p(A\vert \overline{B})}$	$\boldsymbol{p(B)}$
- 5 BB/100	0.25%	0.000004	0.00222555	99.75%
- 4 BB/100	2%	0.000031	0.00226468	98%
- 3 BB/100	8%	0.000182	0.00239720	92%
- 2 BB/100	20%	0.000738	0.00259053	80%
- 1 BB/100	39.5%	0.002037	0.00233943	60.5%
0 BB/100	20%	0.003834	0.00181659	80%
- 1 BB/100	8%	0.004918	0.00198539	92%
- 2 BB/100	2%	0.004301	0.00217753	98%
- 3 BB/100	0.25%	0.002564	0.00221914	99.75%

Applying Bayes' theorem to each of these rows:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A|B)p(B) + p(A|\overline{B})p(\overline{B})}$$

Win Rate	$\boldsymbol{p(A\vert B)}$
- 5 BB/100	0.00%
- 4 BB/100	0.03%
- 3 BB/100	0.66%
- 2 BB/100	6.65%
- 1 BB/100	36.25%
0 BB/100	34.54%
- 1 BB/100	17.72%
- 2 BB/100	3.87%
- 3 BB/100	0.29%
Total	100%

When we looked at the classical approach, we were able to generate a maximum likelihood estimator. In the same way, we can identify -1 BB/100 as the maximum likelihood estimate given these assumptions. Of course, these assumptions aren't really the truth - possible win rates for players are close to continuous. Nevertheless, if we used a continuous distribution and did the more complex math that arises, we would find a distribution similar to this one. And the key implication of this approach is that because of the relative scarcity of winning players, our hypothetical hero is nearly as likely to be a losing player who has gotten lucky as he is to have a positive win rate.

We could see this even more starkly if we considered a player with a much higher observed win rate, perhaps 5 BB/l00h. The classical approach would still assign a maximum likelihood estimate of 5 BB/l00h, because it considers all win rates to be equally likely (because in the classical method we have no additional information about the distribution of win rates). However, recalculating the above analysis with an observed win rate of 5 BB. we find:

Win Rate	$p(A\|B)$
- 5 BB/100	0.00%
- 4 BB/100	0.00%
- 3 BB/100	0.00%
- 2 BB/100	0.16%
- 1 BB/100	3.79%
0 BB/100	15.78%
- 1 BB/100	35.40%
- 2 BB/100	33.84%
- 3 BB/100	11.03%
Total	100%

We can see that here, our player is a heavy favorite to be a winning player, and a substantial one at that. However, as win rates of 5 BB/100 hands are absent from the population, Bayes' theorem properly adjusts his win rate to the more likely levels of the top 10% of all players.

Gathering more data should naturally cause these Bayesian estimates to converge on the observed win rates - the more consistently a player has demonstrated his ability to win at a particular rate, the more likely that that is his true rate. And if we recalculate the above considering a sample of 100,000 hands, we obtain:

Win Rate	$p(A\|B)$
- 5 BB/100	0.00%
- 4 BB/100	0.00%
- 3 BB/100	0.00%
- 2 BB/100	0.00%
- 1 BB/100	1.57%
0 BB/100	33.42%
- 1 BB/100	58.37%
- 2 BB/100	6.60%
- 3 BB/100	0.04%
Total	100%

By this sample size, we are much more confident that the observations match the reality, even though in the underlying distribution, only ten percent of players win at least 1 BB/100.

It is worth noting that there is somewhat of a schism between classical statisticians (sometimes called ***frequentists)*** and Bayesian statisticians. This disagreement centers (roughly) on the Bayesian idea that population parameters should be thought of as probability distributions. Bayesians estimate a prior distribution, observe a sample, and then adjust their prior distribution in light of the new evidence. Frequentists reject this approach, preferring to characterize population parameters as fixed, unchanging values, even if we cannot know their value. We are strongly oriented toward the Bayesian point of view because of its usefulness in poker analysis.

The method we have used here is not even a full-fledged Bayesian technique, but merely a simplification of techniques of Bayesian analysis intended to make the idea accessible. For more detail on this and on the divide between practitioners of Bayesian techniques and classical statisticians, we suggest that you consult advanced statistics texts, particularly those that have thorough treatments of point estimation and Bayesian methods.

One topic that emerges from this analysis is the idea of ***regression to the mean.*** Suppose you take an observation - perhaps of win rates over a sample that is of a small size. The idea here is that observed win rates above the mean will tend be lower if you repeat the observation in the future, and observed win rates below the mean will tend to perform better in the future. This is not because of any kind of statistical "evening out" where outlying trials from the past are offset by outlying trials (in the other direction) in the future - these events are independent. The principle that is at work here is that if you have outperformed the mean of the entire population, it is somewhat more likely that you have outperformed your expectation as well, while if you have underperformed it, you have likely underperformed your own expectation also. As a result, you will tend (sometimes very slightly) to regress a little toward the mean of all players - that is, the best prediction we can make contains a small amount of the overall mean mixed with observed results. We can see this at work in the Bayesian analysis of our hypothetical player - after 16,900 hands, his adjusted win rate prediction was still very heavily influenced by the population distribution, which dragged his win rate down toward the population mean of -1 BB/100.

Key Concepts

- When estimating population parameters from observed data, we can use one of two methods: the frequentist or classical method, or a Bayesian method.
- The classical method assumes that we have no information about the underlying distribution. This gives rise to a maximum likelihood estimate equal to the sample mean, and a confidence interval that includes all values for which the observed result is inside the significance level.
- Bayes' rule gives us a robust methodology for incorporating new information into prior probability estimates.
- Using a prior probability distribution and applying Bayes' theorem can yield better (Bayesian) estimators of population parameters when we have accurate information about the prior distribution.

Part II: Exploitive Play

It is not enough to be a good player; you must also play well.

Siegbert Tarrasch

Chapter 4

Playing the Odds: Pot Odds and Implied Odds

The heart of poker is decision-making. Players who make better decisions than their opponents win; players who make worse decisions than their opponents lose. In Part I, we defined decisions with higher EV as "better" than their lower EV counterparts. In Parts II and III, we will examine the decision-making process as it applies to playing poker games and identify both techniques of analysis and methods of play that lead to making the best decisions at each juncture.

Part II deals with what we call ***exploitive play;*** this is play which seeks to maximize EV by taking the action at each decision point which has the highest EV in a particular situation, inclusive of whatever information is available about the opponent, such as his tendencies, tells, and so on. Virtually every player uses exploitive play in one form or another, and many players, even some of the strongest players in the world, view exploitive play as the most evolved form of poker.

Before we get into a discussion of exploitive play, we will introduce some terms and definitions. First, we have the concept of a ***game.*** In the poker world, we have different definitions of this term, and in Part III, we will be primarily working in the domain of ***game theory,*** the mathematical study of games. For now, we will define a game as containing the following elements:

- There are two or more players.
- At least one player has a choice of actions.
- The game has a set of outcomes for each player.
- The outcomes depend on the choices of actions by the players.

Normally in our poker discussions, there will be two or more players, and both players will have action choices. The set of outcomes for the game will be expressed in dollars won or lost.

Additionally, we call the "choice of action" a ***strategy. In*** game theory terms, a strategy is a complete specification of a player's actions choices at all possible paths the hand might follow. In poker, strategies are extremely difficult to specify, as we have what might be called a combinatorial explosion of paths. There are 1326 starting hands a player might hold. Then there are 19.600 different three-card flops that might hit, 47 different turns, and 46 different rivers. Even after factoring in some equivalences with regard to suit, we still have well over five million board/hand combinations to consider. Then we must specify how we will play-each of our hands on each street, how we will respond to raises, checks, bets, and so on.

This is basically impractical for any but the simplest toy games. As a result, we often use the term strategy a little more loosely in poker. Frequently, when we use this term we are referring to our expected play on this and perhaps one more street. The depth to which we specify the strategy is often tied to the convenience with which we can express it; simpler games and more static boards can often go deeper than more complex ones. It is normally just simpler to treat the game this way. We do, however, try to tie together the play on two or more streets as much as possible.

The concepts of Part I have particular meanings when we consider the play in terms of strategies. It is fairly meaningless to consider the expectation of a hand in a vacuum before the play begins, so instead we use the term "expectation of a hand" here to mean the expectation of a hand played with a given strategy against an opposing strategy. Likewise, the expectation of a hand distribution against a strategy is the weighted average expectation of the hands in that distribution against the opposing strategy, and so on.

Maximizing expectation against the opponent's strategy is the goal of exploitive play. If our opponent plays a strategy S, we define the ***maximally exploitive strategy*** to be the strategy (or one of the strategies) that has the highest expectation against S. When playing exploitively, it is often our goal to find the strategy that is maximally exploitive and then employ it. By doing this, we maximize our expectation. We begin with a simple toy game to illustrate the process of finding this strategy.

Example 4.1

Two players play headsup limit poker on the river. Player A has either the nuts (20% of the time) or a valueless (or ***dead)*** hand (80% of the time), and Player B has some hand of mediocre value - enough to beat dead hands, but which loses to the nuts. The pot is four big bets, and A is first. Let us first consider what will happen if A checks. B could bet, but A knows exactly when he has B beaten or not; hence he will raise B with nut hands and fold at least most of his bluffing hands. B cannot gain value by betting; so he will check. As a result, A will bet all of his nut hands. A might also bet some of his dead hands as a bluff: if B folds, A can gain the whole pot.

We'll call the % of total hands that A bluffs with ***x.*** A's selection of *x* is his strategy selection. B loses one bet for calling when A has a nut hand, and wins five bets (the four in the pot plus the one A bluffed) when A has a bluff. B's calling strategy only applies when A bets, so the probability values below are conditional on A betting. Using Equation 1.11, the expectation of B's hand if he calls is:

<*B*, *call* > = *p*(A has nuts)(-1) + *p*(A has a bluff)(+5)
<*B*, *call* > = (0.2)(-1) + (5)x
<*B*, *call* > = 5x - 0.2

If B folds, his expectation is simply zero.

<B, fold> = 0

We consider a few potential values for x:

Situation	**x value**	**<B, call>**	**<B, fold>**
A never bluffs	0	-0.2	0
A always bluffs	0.8	+3.8	0
A bluffs 5%	0.05	+.05	0
A bluffs 4%	.04	0	0

B should choose the strategy that has higher expectation in each of these cases. If A bluffs often B should call all the time. If A bluffs rarely, B should never call.

To determine how often A will bluff, B might use his knowledge of A's play, tells, or some other information, perhaps using Bayes' theorem (A might have lost the last hand and therefore has a higher *a priori* chance to be playing this hand overaggressively because of emotion, etc.).

We can also graph these two functions:

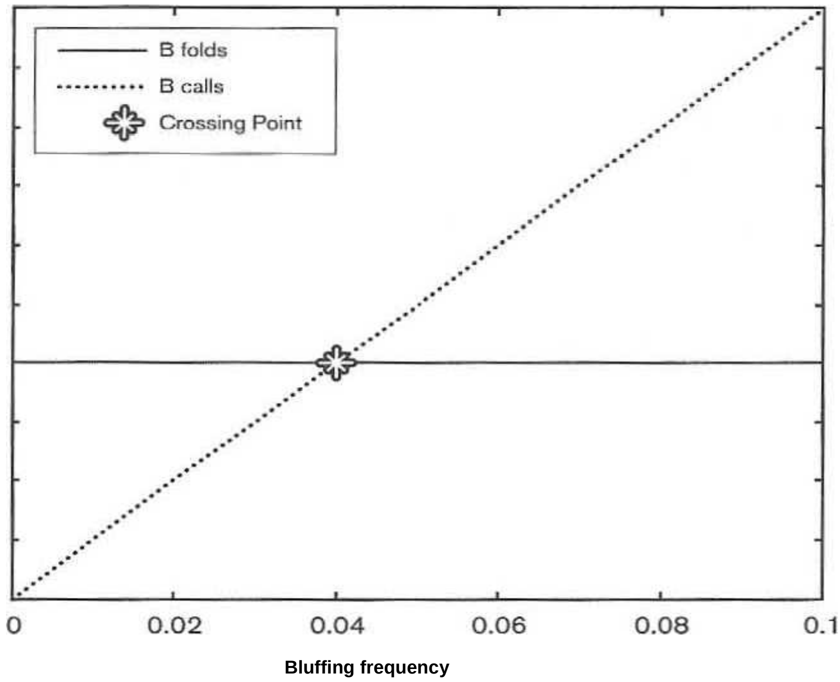

Figure 4.1, Game Equity for various B strategies

Figure 4.1 shows the linear functions that represent B's strategy options. We can see that B's maximally exploitive strategy involves choosing the strategy that has the higher value at the x-value that A is playing. So when x is below 0.04, (meaning that A is bluffing 4% of the time), B should simply fold; above that B should call. One tiling that is important to note about this is that exploitive play often involves shifting strategies rather drastically. If A changes his x-value from 0.039 to 0.041, that is, bluffing two thousandths of a percent more often, B changes his strategy from folding 100% of the time to calling 100% of the time.

Over the next several chapters, we will look, at some of the principles of exploitive play, including pot odds and implied odds, and then consider the play of some example hands. We - will consider situations where the cards are exposed but the play is non-trivial and then play a single hand against a distribution of hands. We focus primarily on the process of trying to find the maximally exploitive strategy in lieu of giving specific play advice on any given hand, and especially on the process of identifying specific weaknesses in the opponent's strategy. This last is particularly valuable as it is generally quite repeatable and easy to do at the table.

Pot Odds

None of the popular forms of poker are ***static games*** (where the value of hands does not change from street to street). Instead, one common element of all poker games played in casinos is the idea of the ***draw.*** When beginners are taught poker, they often learn that four cards to a flush and four cards to a straight are ***draws*** and that hands such as pairs, trips, and flushes and straights are ***made hands.*** This is a useful simplification, but we use "draw" to mean a variety of types of hands. Most often, we use "draw" to refer to hands whose value if the hand of poker ended immediately is not best among the hands remaining, but if certain cards (often called outs) come, they will improve to be best. However, in some cases, this can be misleading. For example, consider the following two hands on a flop of T♣ 9♣ 2♦ in holdem:

Hand A: Q♣ J♣
Hand B: A♠ 3♦

Hand A has more than a 70% chance of whining the hand despite his five-card poker hand being worse at this point. In this hand, it may seem a little strange to refer to the hand that is more than a 7 to 3 favorite as "the draw" while the other hand is the "made hand," because of the connotations we usually associate with these terms. However, we will consistently use the term "draw" throughout this book to mean the hand that at the moment has the worse five-card poker hand and needs to catch one of its outs, no matter how numerous they may be, in order to win. In contrast, we will use the term "favorite" to mean that hand that has more equity in the pot and "underdog" to mean the hand that has less. In the above example, the Q♣ J♣ hand is both the favorite and a draw.

One of the fundamental confrontations between types of hands m poker is between made hands and draws. This confrontation is particularly accentuated in limit poker, where the made hands arc unable to make bets large enough to force the draws to fold; instead, they simply extract value, while (he draws call because the expectation from their share of the pot is more than the amount they must call. However, all is not lost in big-bet poker for the draws. As we shall see later on, draws are able to take advantage of the structure of no-limit games to create a special type of situation that is extremely profitable by employing a play called the ***semi-bluff.***

Example 4.2

The game is $30-60 holdem. Player A has A♣ A♦. Player B has 9♥ 8♥. The board is K♥ 7♣ 3♠ 2♥. The pot is $400. Player A is first. How should the action go if both players know the full situation?

You can likely guess that the action goes A bets - B calls. It is valuable, however, to examine the underlying mathematics because it provides an excellent introduction to this type of analysis and to the concept of ***pot odds.***

If Player A checks, then Player B will certainly check behind. If B were to bet, he would immediately lose at a minimum $^3/_5$ of his bet (because he only wins the pot $^1/_5$of the time), plus additional EV if Player A were to raise. There will be no betting on the river, since both players will know who has won the hand. Rather, any bet made by the best hand will not be called on the river, so effectively the pot will be awarded to the best hand.

Since 35 of the remaining 44 cards give AA the win, we use Equation 1.11 to determine A's EV from checking to be:

$<A, check> = p(\text{A wins})) + (\text{pot size})$
$<A, check> = (^{35}/_{44})(400)$
$<A, check> = \$318.18$

Now let's consider B's options. Again, B will not raise, as he has just a $^1/_5$ chance of winning the pot (with 9 cards out of 44) and A will never fold. So B must choose between calling and folding.

$<B, call> = (p(\text{B wins})) (\text{new pot size}) - (\text{cost of a call})$
$<B, call> = (^9/_{44}) (\$400+60+60) - \$60$

$<B, call> = (^9/_{44}) (\$520) - \$60$
$<B, call> = \$46.36$
$<B, fold> = 0$