A Markov Chain Analysis of Blackjack Strategy

Michael B. Wakin and Christopher J. Rozell*
Department of Electrical and Computer Engineering,
Rice University, Houston, Texas 77251

1 Introduction

Many statistical problems of practical interest are simply too complicated to explore analytically. In these cases, researchers often turn to simulation techniques in order to evaluate the expected outcomes. When approached creatively, however, these problems sometimes reveal a structure that is consistent with a much simpler mathematical framework, possibly permitting an analytical exploration of the problem. It is this situation that we have encountered with the game of blackjack.

Blackjack receives considerable attention from mathematicians and entrepreneurs alike, due to its simple rules, its inherent random nature, and the abundance of "prior" information available to an observant player. Indeed, many attempts have been made to propose card-counting systems that exploit such information to the player's advantage. Most card-counting systems include two aspects: a method for monitoring the cards played to watch for favorable situations, and a strategy for playing and betting depending on the current state of the game. Because blackjack is a complicated game, attempts to actually calculate the expected gain from a particular system often rely on simulation techniques [3, 5]. While such techniques may yield correct results, they may also fail to explore the interesting mathematical properties of the game.

Despite the apparent complexity, there is a great deal of structure inherent in both the blackjack rules and the card-counting systems. Exploiting this structure and elementary results from the theory of Markov chains, we present a novel framework for analyzing the expected advantage of a card-counting system entirely without simulation. The method presented here requires only a few, mild simplifying assumptions, can account for many rule variations, and is applicable to a large class of counting systems. As a specific example, we verify the reported advantage provided by one of the earliest systems, the Complete Point-Count System, introduced by Harvey Dubner in 1963 and discussed in Edward Thorp's famous book, *Beat the Dealer* [5, pp. 93-101]. While verifying this analysis is satisfying, in our opinion the primary value of this work lies in the exposition of an interesting mathematical framework for analyzing a complicated "real-world" problem.

2 Markov chains and blackjack

In this work we use tools from the theory of discrete Markov Chains [4]. Markov chains are an important class of random walks in that they have finite memory: knowing the current state provides as much information about the future as knowing the entire past history. Defining a particular Markov chain requires a *state space* (the collection of possible states) and a *transition*

*Email: {*wakin,crozell*}*@rice.edu*

matrix. The entry in row i and column j of the transition matrix is the probability of moving from state i to state j in one step. States that only allow transitions back to themselves (with probability 1) are called *absorbing states*. If we raise the transition matrix to the n^{th} power, entry (i, j) is the probability of moving from state i to state j in exactly n steps. Certain classes of Markov chains will converge to an equilibrium distribution as n gets large. This equilibrium represents the long-term proportion of time that the chain spends in each state (independent of the starting state). Markov chains which do converge to equilibrium are those that are *irreducible* and *aperiodic*. A chain is irreducible if there exists a path of non-zero probability from every state to every other state. If a chain is irreducible, it is also aperiodic if when considering all possible paths from any state i back to itself, the greatest common denominator of the path lengths is one.

In this paper we employ two distinct approaches for Markov chain analysis of blackjack. First, we develop a collection of Markov chains to model the play of a single hand, and we use absorbing states to compute the player's expected advantage on one hand. Second, we use a Markov chain to model the progression of the unseen cards, and we analyze its equilibrium to calculate the long-term proportion of time the player will have an advantage over the dealer. After further exploiting the structure in the game to make the calculations tractable, we combine the results from our two analyses to calculate the overall expected gain of a particular card-counting system.

For the remainder of this paper, we assume a standard set of blackjack rules and terminology.[1] For instance, we refer to a hand with an 11-valued Ace as *soft* and a hand with a 1-valued Ace (or with no Aces) as *hard*. A box called the *shoe* holds the cards to be dealt. When a certain number of cards have been played (roughly 3/4 of a shoe), the entire shoe is reshuffled after the next hand is completed. The number Δ of 52-card decks in the shoe is specified when relevant ($\Delta = 6$ being typical in many casinos).

The player places a bet B at the beginning of each hand. The player wins a profit of $\frac{3}{2}B$ if dealt a *blackjack* (a total of 21 on the initial 2 cards), as long as the dealer is not also dealt a blackjack. If the dealer shows an Ace, the player may optionally place a side *insurance* bet of $\frac{B}{2}$ that the dealer holds a blackjack (returned with a profit of B on success). If neither the player nor dealer hold blackjack, a profit of B is won when the player's final total is higher than the dealer's (without exceeding 21). The player can elect to hit or stand at will, but the dealer must hit on any total of 16 or less and stand otherwise. When holding the first two cards, a player may also elect to *double down*, doubling the initial bet and drawing only one additional card. If the player's first two cards have the same value, the player can also *split*, dividing the cards into two separate hands and placing an additional bet B.

3 Analyzing a single hand

A hand of blackjack can approximately be thought of as having a Markovian structure: given the total of a player's hand, the probability distribution of the new total after one card is drawn does not depend on the composition of the previous total. For example, if a player holds a total of 14, the probably of having 17 after one draw is essentially independent of how the 14 was formed (i.e., whether it was a pair of 7's, or an 8,4,2 combination, etc.). There is of course some small effect that is being ignored here, because the cards that have been seen affect the probability of what can be drawn from the shoe. We take this to be a negligible effect in the play of one hand, considering that the player typically draws only one or two cards out of possibly 6 decks.

[1]Many variations on these rules exist [6]; in most cases these variations can easily be considered in a similar analysis.

It is not surprising, given this observation, that Markov chains may be used to analyze the play of a single hand. Roughly speaking, a state space could contain all possible totals for the player's hand, and transition probabilities could be assigned according to the distribution of cards in the shoe. In this section, we develop this idea more fully, obtaining a framework for analyzing the play of a single hand and ultimately computing the player's advantage.

3.1 Problem formulation

We wish to compute the player's *advantage* a, which is defined as the expected profit on a hand that begins with a unit bet. This figure is commonly stated as a percentage; for example, $a = -0.04$ corresponds to a 4% house advantage, an expected loss of 4 cents on a hand beginning with $B = \$1$.

We assume throughout the hand that the cards in the shoe obey a fixed probability distribution

$$\mathcal{D} = \{d(i) : i = \{A, 2, \ldots, 10\}\},$$

where $d(i)$ denotes the probability that the next card drawn will be i. We also assume that the player plays according to some prescribed *strategy* $\mathcal{S}$ throughout the hand. A strategy is a collection of deterministic decisions the player will make during the hand regarding hitting, standing, doubling, and splitting, where each decision depends only on the player's total and the dealer's up-card.[2] While the strategy $\mathcal{S}$ may be chosen according to the player's knowledge of the distribution $\mathcal{D}$ at the beginning of the hand, we assume the strategy remains fixed throughout the play of one hand.

3.2 The dealer's hand

The dealer plays according to a fixed strategy, hitting on all hands 16 or below, and standing on all hands 17 or above. To model the play of the dealer's hand, we construct a Markov chain consisting of a state space Ψ_D and a transition matrix D. The state space Ψ_D contains the following elements:

- $\{first_i : i \in \{2, \ldots, 11\}\}$: the dealer holds a single card, valued i. All other states assume the dealer holds more than one card.
- $\{hard_i : i \in \{4, \ldots, 16\}\}$: the dealer holds a hard total of i.
- $\{soft_i : i \in \{12, \ldots, 16\}\}$: the dealer holds a soft total of i.
- $\{stand_i : i \in \{17, \ldots, 21\}\}$: the dealer stands with a total of i.
- bj: the dealer holds a blackjack (natural).
- *bust*: the dealer's total exceeds 21.

In total, we obtain a state space with $|\Psi_D| = 35$. The dealer's play corresponds to a random walk on this state space, with initial state corresponding to the dealer's first card, and with each transition corresponding to the draw of a new card. Transition probabilities are assigned according to the shoe distribution $\mathcal{D}$. We emphasize that while the transition matrix depends on the distribution $\mathcal{D}$, the state space does not.

For the situations where the dealer must stand, we specify that each state transitions to itself with probability 1. The states $stand_{17}, \ldots, stand_{21}$, bj, and *bust* then become absorbing states. Because the dealer's total increases with each transition (except possibly once when a soft hand transitions to hard), any random walk must reach an absorbing state within 17 transitions.

[2]The well-known Basic Strategy is one such example [1, 5, 6].

To compute a probability distribution among the dealer's possible outcomes, we need only to find the distribution among the absorbing states. This is accomplished by constructing the transition matrix D and computing D^{17}. Given that the dealer shows initial card γ, for example, we obtain the distribution on the possible outcomes by examining row $first_\gamma$ of the matrix D^{17}. Note that because D depends on $\mathcal{D}$, the distribution on the absorbing states is conditioned at this point on a particular deck distribution.

3.3 The player's hand

In a similar way, we use Markov chains to compute the distribution of the player's outcomes. Because the player's strategy depends on the dealer's up-card, we must use a different Markov chain for each card $\gamma \in \{2, \ldots, 11\}$ that the dealer may show. Each Markov chain consists of a state space Ψ_P and a transition matrix P_γ. The state space Ψ_P is fixed, containing the elements:

- $\{first_i : i \in \{2, \ldots, 11\}\}$: the player holds a single card, valued i, and will automatically be dealt another.
- $\{twoHard_i : i \in \{4, \ldots, 19\}\}$: the player holds two different cards for a hard total of i and may hit, stand, or double.
- $\{twoSoft_i : i \in \{12, \ldots, 20\}\}$: the player holds two different cards for a soft total of i and may hit, stand, or double.
- $\{pair_i : i \in \{2, \ldots, 11\}\}$: the player holds two cards, each of value i, and may hit, stand, double, or split.
- $\{hard_i : i \in \{5, \ldots, 20\}\}$: the player holds more than two cards for a hard total of i and may hit or stand.
- $\{soft_i : i \in \{13, \ldots, 20\}\}$: the player holds more than two cards for a soft total of i and may hit or stand.
- $\{stand_i : i \in \{4, \ldots, 21\}\}$: the player stands with the original bet and a total of i.
- $\{doubStand_i : i \in \{6, \ldots, 21\}\}$: the player stands with a doubled bet and a total of i.
- $\{split_i : i \in \{2, \ldots, 11\}\}$: the player splits a pair, each card valued i (modeled as an absorbing state).
- bj: the player holds a blackjack (natural).
- $bust$: the player busts with his original bet.
- $doubBust$: the player busts with a doubled bet.

Note that different states with the same total often indicate that different options are available to the player. In total, we obtain a state space with $|\Psi_P| = 116$.

Analysis of this Markov chain is similar to the dealer's chain described above. The player's strategy dictates a particular move by the player (hit, stand, double, or split) for each of the states; in terms of this Markov chain, the strategy dictates the allowable transitions. Subsequently, transition probabilities are assigned according to the distribution $\mathcal{D}$ of cards in the shoe. In this way, the transition matrix *encodes* both the player's strategy and the shoe distribution.

Because the player's total increases with each transition (except possibly once when a soft hand transitions to hard), it follows that within 21 transitions, any random walk will necessarily reach one of the absorbing states. Once again, it follows that row $first_i$ of P_γ^{21} will yield the distribution on these states, assuming the player begins with card i. Averaging over all possible initial cards i (and weighting by the probability $d(i)$ that the starting card is i), we may compute the overall distribution on the player's absorbing states.

While the player's analysis is similar to the dealer's analysis (Section 3.2) in most respects, the primary difference comes from the possibility of a split hand. We include a series of absorbing states $\{split_i\}$ for the event where the player elects to split a pair of cards i. Intuitively, we imagine that the player then begins two new hands, each in the state $first_i$. To model the play of one of these hands, we create another Markov chain (similar to the player's chain described above), but we construct this chain using the particular rules that govern the play of a post-split hand [5,6].[3]

3.4 Computing the per-hand advantage

Assume the dealer shows card γ face up. To compute the player's advantage for such a hand, we must determine the probability of each possible combination of dealer/player outcomes. As described in Section 3.2, we may use row $first_\gamma$ of the matrix D^{17} to determine the distribution $u_\gamma(i)$ on the dealer's absorbing states $U = \{i \in \Psi_D : (D)_{i,i} = 1\}$. Similarly, we may use P_γ^{21} to determine the distribution $v_\gamma(j)$ on the player's absorbing states $V_\gamma = \{j \in \Psi_P : (P_\gamma)_{j,j} = 1\}$. Note that these outcomes are independent given γ, so the probability for any pair of dealer/player outcomes can be computed from the product of the distributions.

For any combination of the dealer's absorbing state $i \in U$ and the player's absorbing state $j \in V$, we must also determine the expected profit $\alpha(i,j)$ for the player, assuming an initial unit bet (i.e., $\alpha(i,j) = 0$ denotes breaking even). When j is not a split-absorbing state, the profit is deterministic, as explained in Section 2. When j is a split-absorbing state, the expected profit is computed by using the two post-split Markov chains (described in Section 3.3) and examining the appropriate row of the iterated transition matrix.

We can compute precisely the player's advantage on a single hand simply by averaging the player's profit over all possible combinations of the player's and dealer's absorbing states (weighting each combination by its probability), and then averaging over all possible up-cards for the dealer:

$$a = \sum_{\gamma=2}^{11} d(\gamma) \sum_{i \in U} \sum_{j \in V_\gamma} u_\gamma(i) v_\gamma(j) \alpha(i,j).$$

We reiterate that this result implicitly depends on the distribution $\mathcal{D}$ and the strategy $\mathcal{S}$.

4 Analyzing a card-counting system

Throughout a round of blackjack, a player can expect variations in the distribution $\mathcal{D}$ of cards remaining in the shoe. Many people have made several key observations along these lines:

- The player may vary the betting and playing strategy, while the dealer must play a fixed strategy.

[3] For simplicity, we assume the player is limited to splitting only once, although the analysis is easily extended to multiple splits.

- When the shoe has relatively many high cards remaining, the dealer is at a disadvantage by having to hit on totals 12–16.
- When the shoe has relatively few high cards remaining, the house typically has a small advantage over the player, regardless of the playing strategy.

These observations are fundamental to most card-counting strategies and are also the basic reasons why card-counting is not welcomed by many casinos. By counting cards as they are played, a player can obtain partial knowledge about $\mathcal{D}$ and vary the playing and betting strategy accordingly. As a result, card-counting can in fact give the player a long-term advantage over the house.

As one example, in this section we calculate the player's overall advantage using a card-counting system known as the Complete Point-Count System [5, pp. 93-101]. From the previous section, we are able compute the specific per-hand advantage of the player, given a distribution $\mathcal{D}$ and a strategy $\mathcal{S}$. To analyze a card-counting system, however, we must also be able calculate the percentage of time a player expects the shoe to be in favorable and unfavorable states. For this problem, we again propose a Markov chain analysis. Although we focus on one particular card-counting system, the spirit of our approach is quite general and applies to a large class of techniques.

After briefly explaining the Complete Point-Count System, we introduce the Markov chain framework used to track the state of a shoe throughout a round of blackjack. We use the convergence properties described in Section 2 to determine the long-term proportion of time that the shoe is in favorable and unfavorable states.

4.1 The Complete Point-Count System

In the card-counting technique of the Complete Point-Count System, all cards in the deck are classified as *low* (2 through 6), *medium* (7 through 9), or *high* (10 and Ace). Each 52-card deck thus contains 20 low cards, 12 medium cards, and 20 high cards. As the round progresses, the player keeps track of an ordered triple (L, M, H), representing the number of low, medium and high cards that have been played. This triple is sufficient to compute the number of cards remaining in the shoe, $R = N - (L + M + H)$, where $N = 52\Delta$ is the number of total cards in the shoe.

The player uses the ordered triple to compute a *high-low index* (HLI):[4]

$$\texttt{HLI} = 100 \cdot \frac{L - H}{R}.$$

Thorp gives a series of tables to be used by the player in determining $\mathcal{S}$ as a function of the HLI [5, p. 98]. The HLI gives an estimate of the condition of the shoe: when positive, the player generally has an advantage and should bet high; when negative, the player generally has a disadvantage and should bet low. Thorp offers one possible suggestion for varying bets [5, p. 96]:

$$B = \begin{cases} b & \text{if} \quad -100 \le \texttt{HLI} \le 2, \\ \left\lfloor \frac{\texttt{HLI}}{2} \right\rfloor b & \text{if} \quad 2 < \texttt{HLI} \le 10, \\ 5b & \text{if} \quad 10 < \texttt{HLI} \le 100, \end{cases} \tag{1}$$

where b is the player's fundamental unit bet. It is important also to note that, although the player's advantage is presumably high when $\texttt{HLI} > 10$, Thorp recommends an upper limit on bets

[4] In reality, the player can compute the HLI simply by keeping count of $(L - H)$ and R; for simplicity of our analysis, we continue to refer to an ordered triple.

for practical reasons. If a casino suspects that a player is counting cards, they will often remove that player from the game. Finally, Thorp also recommends taking the insurance bet when $\texttt{HLI} > 8$.

Suppose that the player observes an ordered triple (L, M, H) prior to beginning a hand. From the triple we may compute the HLI and obtain the strategy $\mathcal{S}$. After making some mild simplifying assumptions, we can use the techniques of Section 3 to compute the player's expected advantage on such a hand. First, we assume that $\mathcal{D}$ is *uniform* over each category: low, medium, and high. To be precise, we set $\mathcal{D}$ as follows:

$$\begin{aligned} d(2) = \cdots = d(6) &= \left(\frac{1}{5}\right)\frac{20\Delta - L}{R}, \\ d(7) = d(8) = d(9) &= \left(\frac{1}{3}\right)\frac{12\Delta - M}{R}, \\ d(10) &= \left(\frac{4}{5}\right)\frac{20\Delta - H}{R}, \\ d(A) &= \left(\frac{1}{5}\right)\frac{20\Delta - H}{R}. \end{aligned}$$

Second, we assume that $\mathcal{D}$ *does not change* during the play of the hand. Third, we require that the player fixes a strategy at the beginning of the hand (based on the HLI) — that is, the player does not respond to changes in the HLI until the hand is complete. With these assumptions, we are able to compute the player's advantage for a hand beginning with any arbitrary triple (L, M, H). If we were also able to determine the overall probability that the player begins a hand with triple (L, M, H), we would be able to compute the overall long-term advantage simply by averaging over all triples. We turn once again to Markov chains to find these probabilities.

4.2 Markov chain framework for shoe analysis

In the Complete Point-Count System, the state of the shoe after n cards have been played is determined by the proportion of high, medium and low cards present in the first n cards. To calculate the state of the shoe after $n+1$ cards have been played, it is enough to know the $(n+1)^{th}$ card and the state of the shoe at time n. The finite memory of the system makes it perfect for Markov chain analysis. We study the state of a changing shoe in isolation from the playing strategy, ignoring for the moment the analysis of individual hands performed in Section 3. You can imagine that we sit with a shoe of cards containing Δ decks and turn over one card at a time (reshuffling periodically) while we watch how the state of the remaining cards changes. In this formulation there are no absorbing states and we analyze instead the chain's equilibrium properties.

Consider a Markov chain consisting of state space Σ and transition matrix X, where each state in Σ represents an ordered triple (L, M, H) containing the number of low, medium and high cards that have been played. The total number of states in the chain is given by

$$|\Sigma| = (20\Delta + 1)(12\Delta + 1)(20\Delta + 1) = 4800\Delta^3 + 880\Delta^2 + N + 1.$$

Clearly $|\Sigma|$ grows as N^3. Table 1 shows the number of states for some example shoe sizes.

Each state in Σ has two different types of transitions out: either a card can be played or the deck can be reshuffled (i.e., a transition back to the state $(0, 0, 0)$). The location of the reshuffle point will not depend on what specific cards have been played, but only *how many* cards have been played. Let $C_n \subset \Sigma$ be the set of all states such that n cards have been played, $C_n = \{(L, M, H) \in \Sigma : L + M + H = n\}$. We define $\rho(n)$ as the probability of a reshuffle from any state in C_n.

Table 1: *The size of the state space and transition matrices for Markov chains* (Σ, X) *and* (Γ, Y) *with common shoe sizes.*

Δ	N	$\lvert\Sigma\rvert$	$N \times N$	$\lvert\Sigma\rvert \times \lvert\Sigma\rvert$
1	52	5733	2704	3.29×10^7
2	104	42,025	10,816	1.7×10^9
4	208	321,489	43,264	1.03×10^{11}
6	312	1,068,793	97,344	1.14×10^{12}

Given that we are in state (L, M, H) and a reshuffle does not occur, the chain can transition to $(L+1, M, H), (L, M+1, H)$ and $(L, M, H+1)$ with probabilities equal to the probability that the next card drawn is a low, medium or high card, respectively. To be more explicit, if the current state is $(L, M, H) \in C_n$, the transition matrix for each row is given by

$$(X)_{(L,M,H),(a,b,c)} = \begin{cases} \left(\frac{20\Delta - L}{R}\right)(1-\rho(n)) & \text{if} \quad (a,b,c) = (L+1, M, H), \\ \left(\frac{12\Delta - M}{R}\right)(1-\rho(n)) & \text{if} \quad (a,b,c) = (L, M+1, H), \\ \left(\frac{20\Delta - H}{R}\right)(1-\rho(n)) & \text{if} \quad (a,b,c) = (L, M, H+1), \\ \rho(n) & \text{if} \quad (a,b,c) = (0,0,0), \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Note that some of these probabilities could be zero, but these are the only possible transitions.

In a typical casino situation, a dealer will normally play through most, but not all, of a shoe before reshuffling. For example, a dealer may cut $\sim$75% into the shoe and play up to this point before reshuffling. We model the reshuffle point as a (normalized) Laplacian random variable centered around $.75(N)$, with support over the second half of the shoe. The variance is scaled with the size of the shoe in order to keep a constant proportion of the shoe in a region with a high probability of a reshuffle. Precisely, the probability of the reshuffle point being immediately after the n^{th} card is played is given by

$$\Pr\left[\text{reshuffle} = n\right] = \begin{cases} \frac{\kappa}{\sqrt{2\sigma^2}} \exp\left\{\frac{-|n - .75(N)|}{\sqrt{\sigma^2/2}}\right\} & \text{if} \quad n \geq \lceil N/2 \rceil, \\ 0 & \text{otherwise,} \end{cases}$$

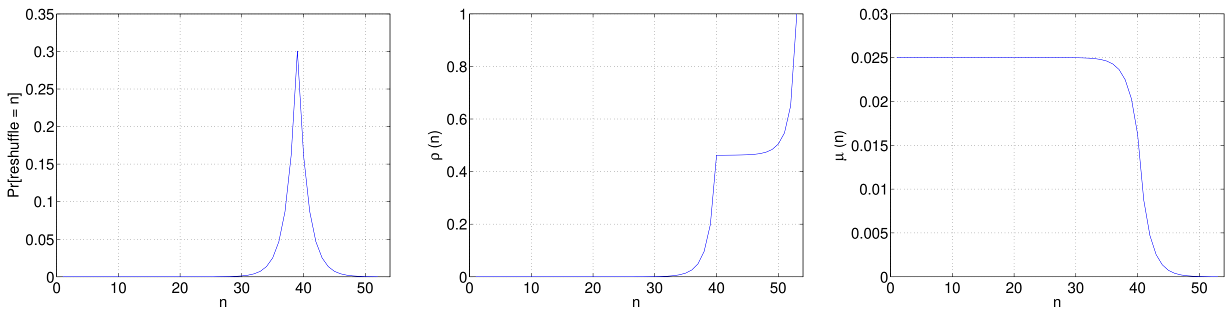
where $\sigma^2 = N/10$, and κ is a normalizing constant to make the distribution sum to one.

To calculate the probability $\rho(n)$ of a reshuffle from a state in C_n, we calculate the probability that the reshuffle point is n conditioned on the event that the reshuffle point is *at least* n,

$$\rho(n) = \Pr\left[\text{reshuffle} = n | \text{reshuffle} \geq n\right] = \frac{\Pr\left[\text{reshuffle} = n\right]}{\sum_{m=n}^{N} \Pr\left[\text{reshuffle} = m\right]}. \tag{3}$$

The probability distribution on the reshuffle location is shown in Figure 1(a), and the reshuffle transition probabilities are shown in Figure 1(b).

Now that we have constructed a Markov chain representing the play of cards from the shoe, we must confirm that this chain will actually serve our purposes. By inspection, it is clear that from any state, it is possible to reach any other state with some non-zero probability. Observing the possible paths that start at the beginning of the shoe (state $(0,0,0)$) and then reshuffle back to the beginning, we see that there is no periodic behavior in the chain. This is stated precisely in



(a) (b) (c)

Figure 1: *(a) Reshuffle point density for 1 deck. (b) State reshuffle transition probability for 1 deck. (c) Equilibrium for column Markov chain* (Γ, Y).

the fact that two possible return times are $\lceil N/2 \rceil$ and $\lceil N/2 \rceil + 1$ and $\gcd(\lceil N/2 \rceil, \lceil N/2 \rceil + 1) = 1$. The Markov chain representing the changing state of the shoe is therefore irreducible [4, p. 11] and aperiodic [4, pp. 40-41]. As described in Section 2, such a chain converges to an equilibrium distribution that describes the long-term proportion of the time the chain will spend in each state.

4.3 Analytic calculations of shoe equilibrium

Equations (2) and (3) give an explicit expression for the transition matrix X. Knowing X, the unique equilibrium can be solved analytically as

$$\pi = (1, 1, \ldots, 1)(I - X + E)^{-1},$$

where I and E are the $|\Sigma| \times |\Sigma|$ identity and ones matrices, respectively [4, p. 40, Exercise 1.7.5]. In Table 1 we see that even for a one deck shoe, X would have 33 million entries. To store X as a matrix of 8-byte, double-precision floating point numbers, it would require approximately 263MB of memory. To analyze a two deck shoe, X would require approximately 13.6GB of memory. Aside from the issue of memory usage, one would also need to invert a matrix of this size. We find ourselves in an interesting dilemma — although we have perfect knowledge of the transition matrix, we are prevented from dealing with X as a whole. In this section we describe how we further expoit the structure of the game and the playing strategies to make the problem tractable.

Looking more carefully at the Markov chain we have constructed, there is a great deal of structure. The form of the chain is more clear in a graphical representation. Imagine that we arranged all of the states so that states representing the same number of cards played (i.e., belonging to the same set C_n) are in the same column, and each column represents one more card played than in the previous column (depicted graphically in Figure 2). Note that when a card is played, a state in C_n can only move to a state in C_{n+1} or the $(0, 0, 0)$ state. Only the states in C_n for $n \geq \lceil N/2 \rceil$ are capable of causing a reshuffle (transitioning back to the $(0, 0, 0)$ state), and each state in C_n reshuffles with the same probability $\rho(n)$. Columns closer to the midpoint of the shoe contain more states, and the first and last columns taper down to one state each.

Starting at state $(0, 0, 0)$, a walk will take one particular path from left to right, moving one column with each step and always taking exactly $\lceil N/2 \rceil$ steps to reach the midpoint. After the midpoint, the states can also reshuffle at each step with a probability $\rho(n)$ that *only* depends on

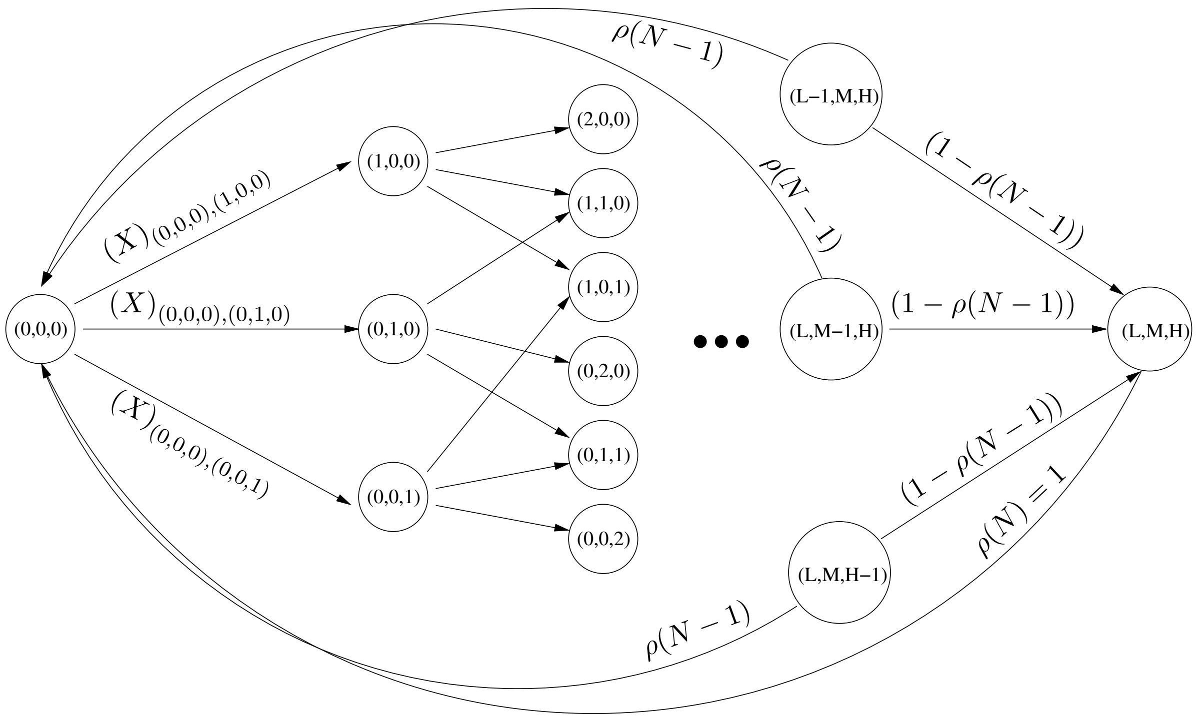


Figure 2: *Graphical depiction of full state space* Σ*. Each state represents an ordered triple* (L, M, H) *denoting the number of low, medium, and high cards that have been played from the shoe.*

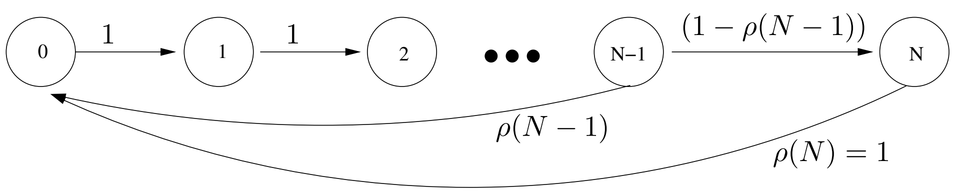


Figure 3: *Graphical depiction of reduced column state space* Γ*. Each state represents one the number of cards played (i.e., one column of the state space depicted in Figure 2).*

the the current column and not on either the path taken up to that point or the exact current state within the column. Essentially, this structure allows us to separate the calculation of π into two components: how much relative time is spent in each column, and within each individual column, what proportion of time is spent in each state. To determine the relative time spent in each column, we create a reduced chain with state space Γ and transition matrix Y, where each column in Figure 2 corresponds to one state in Γ (depicted in Figure 3). The transition matrix is given by

$$(Y)_{n,m} = \begin{cases} 1-\rho(n) & \text{if} \quad m = n+1, \\ \rho(n) & \text{if} \quad m = 0, \\ 0 & \quad\;\; \text{otherwise.} \end{cases}$$

It is clear that this chain is also irreducible and aperiodic, and therefore will converge to a unique equilibrium μ. The equilibrium of the reduced chain (Γ, Y) represents the relative proportion of time that the original chain (Σ, X) spends in each column of Figure 2. This is stated precisely as

$$\mu(n) = \sum_{k \in C_n} \pi(k).$$

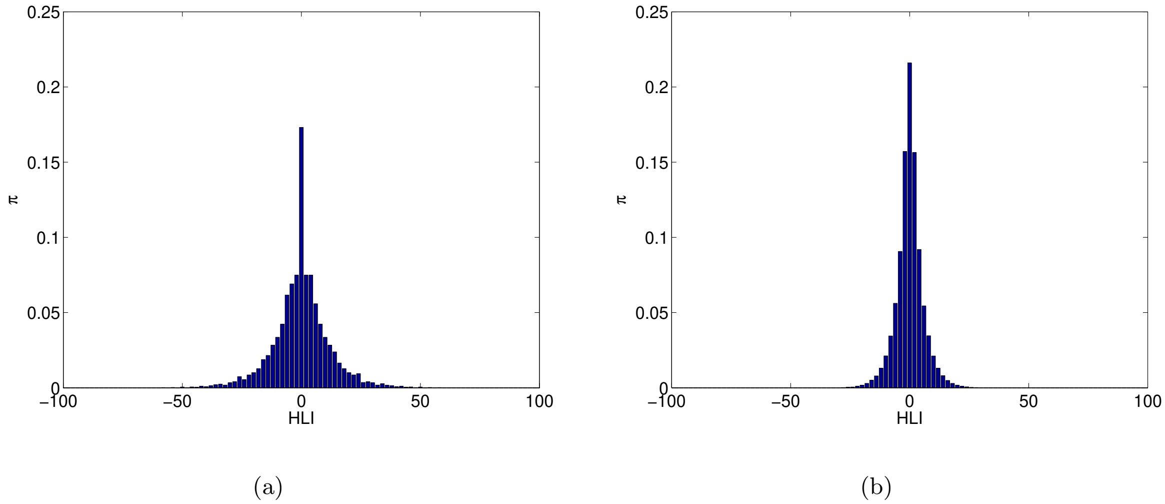


Figure 4: *Equilibria π for 1 and 4 deck shoes (a and b, respectively) representing the long-term proportion of time spent at each High-Low Index (HLI).*

Figure 1(c) shows μ for $\Delta = 1$. It is important to note that the dimension of the reduced column-space chain is much smaller than the original chain, with $|\Gamma| = O(N)$ and $|\Sigma| = O(N^3)$. Even in the case when $\Delta = 6$, one can easily calculate

$$\mu = (1, 1, \ldots, 1)(I - Y + E)^{-1}.$$

Using the easily calculated μ, we can exploit the relationship between the two Markov chains to calculate π. First, it is critical to note that because $|C_0| = 1$, the equilibrium distributions on the initial states of both chains are equal, $\pi((0, 0, 0)) = \mu(0)$. To calculate all other elements of π we make two key observations. First, the properties of the equilibrium distribution [4, p. 33] allow one to relate the values of the equilibrium distribution to each other through the expression $\pi = \pi X$. Second, in our particular Markov chain (Σ, X), states in C_n only have input transitions coming from states in C_{n-1} (i.e., $\forall m \in C_n, (X)_{k,m} = 0$ if $k \notin C_{n-1}$). These two observations together give us that for any $m \in C_n$, the following relation holds:

$$\pi(m) = \sum_{k \in C_{n-1}} \pi(k) \, (X)_{k,m} \, . \tag{4}$$

Using only knowledge of $\pi((0, 0, 0))$, (4) can be used to calculate $\pi(m)$ for any $m \in C_1$. Iterating this process we calculate $\pi(m)$ for any $m \in C_2$, and so on, until we have completely calculated π.

The technique described here takes advantage of the rich structure in the chain to analytically calculate π exactly using only local knowledge of X, and the inversion of an $N \times N$ matrix. The algorithm can calculate the equilibrium when $\Delta = 6$ in under an hour. Equilibria calculated through this method for $\Delta = \{1, 4\}$ are shown in Figure 4. Because the equilibrium would be difficult to plot in the three dimensional state space Σ, states with the same HLI are combined and the equilibria are plotted vs. HLI.

4.4 Convergence to equilibrium

As described in Section 2, certain classes of Markov chains will converge to an equlibrium distribution. It is common to examine how quickly a given Markov chain converges (see [2] for example), often by bounding $\left|(X^n)_{i,j} - \pi(j)\right|$ as a function of n. For our blackjack analysis, the answer to this question roughly reflects how long a player must play before observing "typical" behavior.

By once again exploiting the column structure of the Markov chain $(\Sigma,\ X)$ we are able to derive an interesting relation. Suppose $n > N+1$, and let $i, j \in \Sigma$ be ordered triples. We define c_i to be the column index of triple i (i.e., $i \in C_{c_i}$). We have shown that

$$\left|(X^n)_{i,j} - \pi(j)\right| \leq \left|\left(Y^{n-c_j}\right)_{c_i,0} - \mu(0)\right|.$$

Not surprisingly, the convergence of $(\Sigma,\ X)$ is closely related to the convergence of the smaller chain (Γ, Y), which could possibly simplify the derivation of convergence bounds using the standard techniques. A useful topic for future research would involve extending such bounds to model the variance of a player's average profit over many hands.

5 Analysis

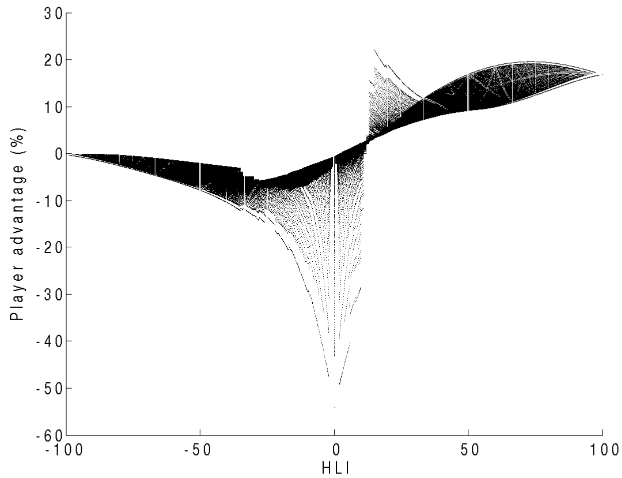
We present in this section our analysis of the Complete Point-Count System. Because the betting is not constant in this system, it is important now to distinguish between the player's *advantage* and the player's *expected gain*. As defined in Section 3, the player's advantage a is the expected profit on a hand that begins with a unit bet. We define the player's expected gain g to be the dollar amount one expects to profit from each hand when betting according to (1) with $b = \$1$. These quantities are related on a single hand by the expression $g = B \cdot a$.

5.1 Player advantage vs. HLI

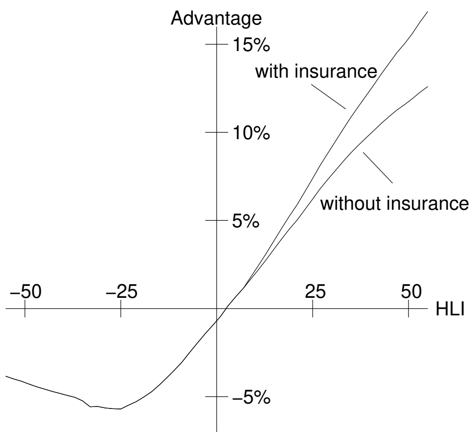
For a game with $\Delta = 4$, we use the method described in Section 3 to compute the player's advantage for each possible triple (L, M, H). Figure 5(a) plots the player's advantage against the corresponding HLI for each possible triple in the 4-deck game (assuming for the moment that the player does not take insurance). It is interesting to note that a single value of HLI may correspond to several possible shoe states; these states may, in turn, correspond to widely varying advantages for the player. Some of these triples may be highly unlikely, however. To get a better feel for the relation between HLI and player advantage, we use the analytic method of Section 4.3 to compute the equilibrium distribution of the states. Figure 5(b) shows a plot where we use the relative equilibrium time to average all triples corresponding to approximately the same HLI value. This figure also includes the player's advantage when playing with insurance.[5] As expected, the player's advantage generally increases with HLI, and the player is at a disadvantage when HLI is negative.

For comparison purposes, Figure 5(c) is the corresponding plot that appears in Thorp's description of the Complete Point-Count System [5, p. 97], most likely derived using computer simulation. While our plots qualitatively agree, there are some small quantitative differences. For example, when the HLI is near zero, we predict a small disadvantage for the player, while Thorp's plot indicates a fair game. Through simulation of actual play, we have confirmed a small house advantage when `HLI` ≈ 0. The differences from Thorp's result may be due to the fact that Thorp allows the player's strategy to change with each new card (instead of at the beginning of each hand). Additionally, Thorp leaves some simulation parameters unspecified (such as the number of decks, reshuffle scheme, and rules governing split hands).

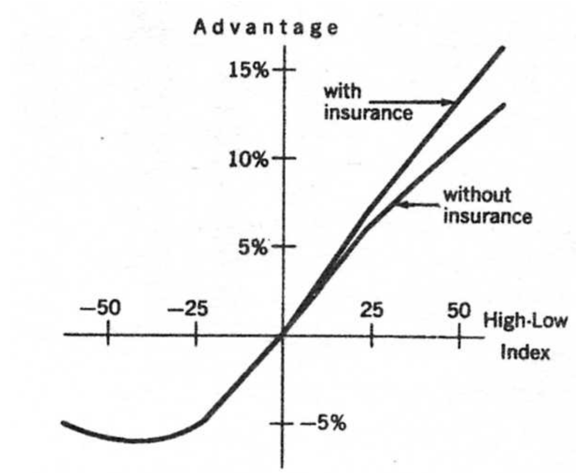
[5] For completeness, we note that the probability of winning an insurance bet is precisely $d(10)$.



(a)



(b)

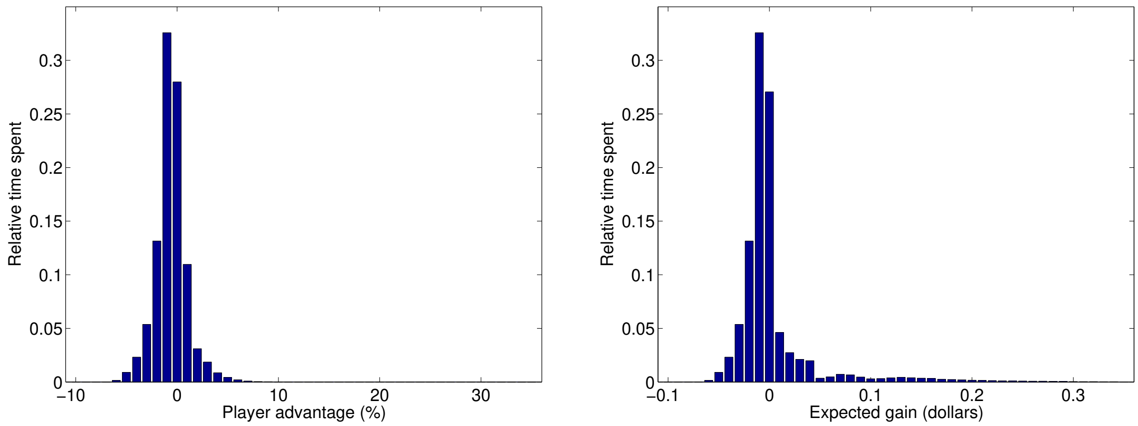


(c)

Figure 5: *Player advantage vs. HLI. (a) Scatter plot of advantage for individual (L, M, H) triples. (b) Weighted average over (L, M, H) triples using equilibrium π to determine advantage at a specific HLI. (c) Thorp's result, reprinted from [5, p. 97], permission pending.*

Table 2: *Overall expected gain for player using the Complete Point-Count System (with insurance).*

Δ	Expected gain, g
1	0.0296
2	0.0129
4	0.0030



(a) (b)

Figure 6: *(a) Relative time spent with different player advantages* ($\Delta = 4$)*. (b) Relative time spent with different expected gains.*

5.2 Expected gains

Figure 6(a) shows the average amount of time the player expects to play with different advantages (assuming $\Delta = 4$ and that the player plays with insurance). The player spends a considerable amount of time in states with a disadvantage and in fact would suffer a 0.60% *disadvantage* by using a fixed betting strategy. Variable betting that exploits favorable situations is therefore key to the player's hope for a positive expected gain. Figure 6(b) shows the average amount of time the player expects to play with different expected gains when betting according to (1). We compute the player's expected overall gain to be $g = 0.0030$, or 0.30 cents per hand.

Not surprisingly, the number of decks in play has a direct impact on the player's expected gain. We notice, however, that plots such as Figure 5(a) change little as the number of decks changes, indicating that HLI is a universally good indicator of the player's advantage. As we observed in Figure 4, the relative amount of time spent in each HLI depends strongly on the number of decks in play. Because much of the player's gain comes from placing large bets when HLI is large, the player is less likely to encounter these extreme situations with more decks in play. Table 2 highlights this dependency and shows the player's overall expected gain as the number of decks changes.

While Thorp makes no specific claims for overall expected gains using the Complete Point-Count System [5, pp. 93-101], we have attempted to validate our results using simulation of actual hands. The results agree roughly with Table 2, though simulations indicate that we underestimate the overall gain by an amount ranging from 0.01 for 1 deck to 0.0025 for 4 decks. These discrepancies are due to the simplifying assumptions that $\mathcal{D}$ is uniform over each category and static during the play of each hand, which have a greater impact when fewer decks are in play.

6 Conclusions

Blackjack is a complicated game, where subtle changes in rules or playing strategy can have a surprising impact on the player's advantage. As with many other complex real-world systems, blackjack analysis is typically done via simulation. Aside from being somewhat unsatisfying mathematically, simulation can make it difficult to explore a large number of rule and strategy variations and rarely provides insight into the underlying reasons for the results.

We have presented an analysis framework for card-counting systems that operates completely without simulation. This framework is based on the observation that certain aspects of the game of blackjack have finite memory and are well modeled by discrete Markov chains. By using techniques from Markov chain analysis, we developed a framework for calculating the player's long-term expected gain and exercised this technique on the well-known Complete Point-Count System. By using our method, one could further investigate detailed aspects of the game. For example, by determining the reasons underlying particular advantages, a player could assess rule variations and make strategy adjustments accordingly.[6]

As in any case of trying to match the real world to a tidy mathematical framework, some simplifying assumptions were required. These assumptions introduce inaccuracies into the analysis, but we have attempted to make only the mildest assumptions required. Even then, a direct application of Markov chains produced a problem that was technically well-defined but computationally intractable. It was only through further exploring the structure of the system that we were able to reduce the complexity of the calculations. More accuracy could be achieved if we worked to reduce the impact of our assumptions, for example by keeping track of card counts over more than three categories. However, increasing the complexity in this way would quickly lead us back into a problem that is computationally impractical.

In this work, Markov chains have been a powerful tool for exploiting the inherent structure that exists both in the game of blackjack and in card-counting strategies. Though blackjack is only a casino game, our exercise illustrates the power of Markov chains in analyzing complex systems. The continued development of more advanced techniques for Markov chain analysis should only extend the utility of Markov chains for more complicated, real-world problems.

Acknowledgments. The authors are grateful to Dr. Robin Forman for discussions that encouraged this analysis. Figure 5(c) is reproduced from [5, p. 97], permission pending.

References

[1] P. A. Griffen. *The Theory of Blackjack: The Compleat Card Counter's Guide to the Casino Game of 21.* Huntington Press, Las Vegas, NV, USA, 1996.

[2] M. Jerrum. Mathematical foundations of the Markov chain Monte Carlo method. In M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, and B. Reed, editors, *Probabilistic Methods for Algorithmic Discrete Mathematics.* Springer-Verlag, 1998.

[3] M. H. Millman. A statistical analysis of casino blackjack. *The American Mathematical Monthly*, 90(7):431–436, Aug.–Sep. 1983.

[4] J. R. Norris. *Markov Chains.* Cambridge University Press, 1997.

[5] E. O. Thorp. *Beat the Dealer: A Winning Strategy for the Game of Twenty-One.* Vintage, New York, 1966.

[6] S. Wong. *Professional Blackjack.* Pi Yee Press, La Jolla, CA, USA, 1994.

[6] For any reader interested in further experimentation or analyzing alternative systems, our Matlab code will soon be publicly available in a Blackjack Markov Chain Toolbox.