RESOLVING PROBLEM GAMBLING:
a mathematical approach

By
Tristan Barnett

Strategic Games
ABN: 64663038362
www.strategicgames.com.au

Table of Contents

About the Author

Tristan's interest in intellectual type games originated with backgammon in 1995 and would practice daily with backgammon software; Jellyfish and GNU Backgammon. Tristan participated in the Melbourne Backgammon club in 2003 and won the monthly tournament in March 2003. After browsing through various gambling books in local libraries in 1999 it became clear that blackjack was a game known to generate a significant amount of profit, and thus Tristan self-funded a trip to Las Vegas in 1999 from playing blackjack at Star City casino, Sydney. Whilst Tristan had moderate success at blackjack the real interest was the mathematics behind blackjack and other gambling related games rather than the actual playing. During the trip to Las Vegas Tristan became particularly interested in video poker and over the years following this trip Tristan published 3 papers related to video poker and setup a syndicate in 2008 to automate a system through online progressive video poker. Tristan also participated in a Jewish 20's-30's bridge club in 2007 and since 2007 Tristan regularly practices bridge with bridge software Bridge Baron. Tristan has also been a tutor/lecturer for a 'Chance and Gaming' statistics subject at Swinburne University from 2002-2005, and documented many probabilities and strategies for gambling games (pdf). Tristan has a PhD in tennis statistics from Swinburne University, worked as a performance analyst with Tennis Australia, and as a mathematician in predicting tennis outcomes for betting organizations Ladbrokes and Centrebet, and sports multimedia company Infoplum. Tristan founded Strategic Games which provides information on the mathematics in sport, gambling and conflicts (html).

Preface

The book begins with Chapter 1 by extending the common model for a person who acknowledges they have a problem with gambling by including strategies for Optimal Gambling, Correct Gambling and Eradication. Poker Machines, Jackpots and Card Games are given for strategies on Correct Gambling to minimize losses. Chapter 2 identifies why poker machines are unfair since they do not provide the probabilities of obtaining specific outcomes – which is fundamental to the mathematics of casino games. Six suggestions are given to increase consumer protection and these suggestions could be implemented within poker machine legislation. Chapter 3 provides mathematical information that could be implemented for each casino game such as the average loss after an hour of play as well as the chances of obtaining various payouts after an hour of play, and thus would inform the consumer on whether to play a particular game and the length of time. Links to Excel calculators are given for Roulette, the Big Wheel and a Sample Poker Machine. Chapter 4 provides an extension to the classical Kelly criterion (two outcomes) for when multiple outcomes exist and can be applied directly to favourable (Optimal Gambling) non-progressive and progressive video poker games. This information is important for managing risk as the Kelly criterion typically maximises the long-term growth of the bankroll. Chapter 5 extends on chapter 4 by demonstrating how the strategies can be obtained through video poker software and can generate profits particularly with progressive video poker machines. Chapter 6 extends on chapter 5 by showing how progressive video poker machines can be automated online and hence generate significant profits. For example, if the hourly win rate using one machine is $10/hour, then the same code used to automate one machine could be used on ten machines for an hourly win rate of 10*$10/hour= $100/hour. The author founded a syndicate and proved how such an operation can be successful. Chapter 7 gives an interesting casino promotion offer where InterCasino were offering a free trip to New York for new customers by turning over $10,000 on a given set of casino games. From Australia, the value of the flight was about $2,000. Given certain playing conditions it was shown how playing a game with an additional $56 of expected value would dramatically reduce the possibly of risking $1,500. Chapter 8 concludes with a detailed tennis model on how one can profit from tennis betting, particularly through live betting as available with such bookmakers as Betfair.

The author would like to thank the following for publishing the articles used to create this book:

Gaming Law Review and Economics
Barnett T (2010). Applying mathematics to poker machine regulations to increase consumer protection. Gaming Law Review and Economics 14(8), 601-607.

Advances in Computer Entertainment Technology
Barnett T and Clarke S (2004). Optimizing returns in the gaming industry for players and operators of video poker machines. Proceedings of the International Conference on Advances in Computer Entertainment Technology. National University of Singapore, 212-216.

CHANCE
Barnett T (2011). How much to bet on video poker? CHANCE. 24(2).

Gaming Law Review and Economics
Barnett T (2012). Automating online video poker for profit. Gaming Law Review and Economics. 16(1-2), 15-20.

CHANCE
Barnett T (2009). Gambling your way to New York: A story of expectation. CHANCE. 22(3).

OR Insight
Barnett T, Darren O'Shaughnessy D and Bedford A (2011). Predicting a tennis match in progress for sports multimedia. OR Insight 24(3), 190-204.

The author would also like to thank Guy West, Bill Jordan and Andrew Scott for fun times in implementing our automated online video poker syndicate. Finally, the author would like to thank Alan Brown for providing implementation to a lot of the mathematics detailed in this book.

Chapter 1: Is Blackjack the Solution to Problem Gambling

1. Introduction

Healey (2006) identifies the options available if you may be a person who acknowledges that you have a problem with gambling. There are three options available 1) Do nothing to change your gambling, 2) Control your gambling (known as Controlled Gambling) and 3) Quit Gambling (known as Abstinence). In Controlled Gambling the patient is allowed to gamble on a limited basis. Controlled Gambling currently has few adherents in North America but is somewhat more popular overseas. In Abstinence, the patient in recovery must completely abstain from all gambling. Abstinence is the goal of Gamblers Anonymous and most, though not all, treatment professionals. This model of using either Controlled Gambling or Abstinence is the current accepted framework for treating problem gambling.

2. New model

A new model for treating problem gambling is a generalization of the current model by including Controlled Gambling and Abstinence as treatment possibilities. The model is based on a process which is trying to maximize the return to the player whilst allowing for the entertainment factor in gambling. Players obtain enjoyment from gambling but of course do not want to lose money in the process. Further, the model can also be used by recreational gamblers to gamble responsibly.

The following strategies are now identified:

Optimal Gambling → Correct Gambling → Controlled Gambling → Abstinence → Eradication

The process is a one-way path whereby if a strategy fails for a particular gambler then the next strategy to consider is the next one in line. Every gambler starts at the Optimal Gambling strategy, where the best way to maximize return and gamble is to be playing games where the odds are actually in your favour e.g. card-counting in blackjack. If the Optimal Gambling strategy would not be successful for the particular gambler, then the next strategy would be a new option known as Correct Gambling, where the approach is to allow gambling whilst playing the right games and strategies to minimize losses and to take advantage of the free food and drinks on offer (more commonly known as comp points). Note that this approach allows the gambler to keep playing, whereas in Controlled Gambling the gambler is only allowed to gamble on a limited basis. Three gambling games have been identified for this purpose in Correct Gambling.

3. Gambling games

3.1 Poker machines

Poker Machines - It seems paradoxical that poker machines can be the best game to minimize losses and yet represent the most common form of problem gambling (at least in Australia). The strategy is simply to play 1 line on a 1c slot machine. Assuming 18 spins per minute, the player is expected to lose between $1.00 and $1.70 per hour (depending on the house margin).

3.2 Jackpots

Jackpots - The two most common forms of jackpots are progressive and deterministic. Refer to Chapters 4,5 and 6 for more information on progressive jackpots in video poker. It is common for slot machines to be linked to a deterministic jackpot. This means that the jackpot must go off before it reaches a specified amount. This is interesting because if you were able to occupy all the machines when the jackpot reached a certain level, then you would be guaranteed to hit the jackpot and generate a profit. The strategy for the general player is to play deterministic machines when the jackpot is towards the specified maximum rather than the minimum.

3.3 Card games

Card Games - Despite what thousands of websites say about making money through card-counting in blackjack, it requires "large" bankrolls, generally playing in a team, hours of training and a significant amount of concentration and hard work whilst playing. Refer to blackjack-masters.com or http://www.blackjackforumonline.com/ for more information about being a successful card-counter. The strategy for the general player is to play basic strategy blackjack with a house margin of 0.59% in Sydney casino or even better basic strategy Pontoon with a house margin of 0.42% in Sydney casino and even lower in Canberra casino with a house margin of 0.34%. Refer to https://wizardofodds.com/games/blackjack/strategy/calculator/ for strategy charts on Blackjack and https://wizardofodds.com/games/pontoon/australian/ for strategy charts on Pontoon. Assuming a player wagers $10 per hand and plays 100 hands per hour, the player is expected to lose $5.90 per hour in blackjack and $4.20 per hour in Pontoon in Sydney casino.

4. Conclusions

If the strategies of Correct Gambling followed by Controlled Gambling followed by Abstinence would not be successful for the particular gambler, then the final strategy is Eradication. This could be in the form of moving to a country or state where gambling is illegal or a location which is a great distance (say 100+ km) to the nearest gambling venue.

References

Healey J (2006) Gambling in Australia. NSW, Spinney Press.

Chapter 2: Applying Mathematics to Poker Machine Regulations to Increase Consumer Protection

1. Introduction

Casino games are comprised of mathematical formulations and can be found widely in the literature such as Croucher (2002), Hannum (2005), Packel (1981) and Epstein (1997). The percent house margin (or return to player) establishes how much a player is expected to lose in the long-run. Whilst the percent house margin is important to consumers in determining the choice or how long to play a particular game, there is other information which could also influence these decisions. This could consist of the probability of the consumer ending up in profit after 100 games or the probability of the consumer losing more than $100 after 200 games. These results are classified as the distribution of payouts and in order to calculate these results requires three pieces of information from the casino game. They are: 1) the initial cost, 2) the payouts for each possible outcome and 3) the probability associated with each outcome. The initial cost is given directly for any casino game. The payouts for each possible outcome are also given directly for every casino game, either in the form of odds or prices. The probabilities associated with each outcome are not given directly for the casino game. However, these probabilities can usually be derived from the playing rules. For example, in single zero roulette, the probability of a particular number coming up can easily be obtained as 1/37. It can therefore be argued that games where the probabilities associated with each outcome can be obtained from the playing rules are "fair"; since the distribution of payouts can be obtained using mathematics. On the contrary, casino games where the probabilities associated with each outcome cannot be obtained from the playing rules could be considered as being "unfair". This is the situation for poker machines in Australia.

The paper begins with section 2, by establishing a formal definition of a casino game. This enables calculations of the percent house margin, and other distributional characteristics of the game such as the standard deviation, and coefficients of skewness and excess kurtosis. These distributional characteristics are then applied as inputs to the Normal Power approximation formula to determine the distribution of profits for the game i.e. the probability of ending up with a certain amount of profit after a number of trials. These mathematical formulations are applied to section 3 to address the regulations of poker machines and suggestions are given for amendments to the Australian/New Zealand Gaming Machine National Standard with the purpose to increase consumer protection.

2. Analysis of casino games

2.1 Percent house margin

A casino game can be defined as follows: There is an initial cost C to play the game. With the assumption of trials being independent, each trial results in an outcome O_i, where each outcome occurs with profit xi and probability pi. The condition $\sum p_i=1$ must be satisfied.

Given this information, the expected profit E_i for each outcome is given by $E_i = p_i x_i$ and the total expected profit is given by $\sum E_i$. The percent house margin (%HM) is then $-\sum E_i/C$ and the total return is $1+\sum E_i/C$. Positive percent house margins indicate that the gambling site on average makes money and the players lose money. Table 1 summarizes this information.

Outcome	Profit	Probability	Expected Profit
O_1	x_1	p_1	$E_1=p_1x_1$
O_2	x_2	p_2	$E_2=p_2x_2$
O_3	x_3	p_3	$E_3=p_3x_3$
…	…	…	…
		1	**$\sum E_i$**

Table 1: Representation of a casino game.

2.2 Moments and cumulants

The outcome or profit from a single bet, X, is a random variable. From probability theory, the *moment generating function* (MGF) of X is

$$M_X(t) = E(\exp(Xt))$$
$$= 1 + m_{1X}\, t + m_{2X}\, t^2/2! + m_{3X}\, t^3/3! + m_{4X}\, t^4/4! + \ldots,$$

where m_{rX} represent the r^{th} moment of X. The moments of X are readily calculated using

$$m_{1X} = \Sigma_i\, p_i\, x_i$$
$$m_{2X} = \Sigma_i\, p_i\, x_i^2$$
$$m_{3X} = \Sigma_i\, p_i\, x_i^3$$
$$m_{4X} = \Sigma_i\, p_i\, x_i^4$$

and so on. The calculation of these moments is illustrated in Table 2.

The cumulant generating function (CGF) of X is the natural log of the MGF:

$$K_X(t) = \log_e (M_X(t))$$
$$= k_{1X}\, t + k_{2X}\, t^2/2! + k_{3X}\, t^3/3! + k_{4X}\, t^4/4! + \ldots,$$

where k_{rX} represent the r^{th} cumulant of X. The relationship between the first four cumulants and moments is given by

$$k_{1X} = m_{1X}$$
$$k_{2X} = m_{2X} - m_{1X}^2$$
$$k_{3X} = m_{3X} - 3m_{2X}m_{1X} + 2m_{1X}^3 \text{ and}$$
$$k_{4X} = m_{4X} - 4m_{3X}m_{1X} - 3m_{2X}^2 + 12m_{2X}m_{1X}^2 - 6m_{1X}^4.$$

These cumulants can be used to calculate the following familiar distributional characteristics (parameters) for X.

Mean $\mu_X = k_{1X}$

Standard Deviation $\sigma_X = \text{sqrt}(k_{2X})$

Coefficient of Skewness $\gamma_X = k_{3X} / (k_{2X})^{3/2}$

Coefficient of Excess Kurtosis $\kappa_X = k_{4X} / (k_{2X})^2$

.

Outcome	Profit	Probability	1st Moment	2nd Moment	3rd Moment	4th Moment
O_1	x_1	p_1	p_1x_1	$p_1x_1^2$	$p_1x_1^3$	$p_1x_1^4$
O_2	x_2	p_2	p_2x_2	$p_2x_2^2$	$p_2x_2^3$	$p_2x_2^4$
O_3	x_3	p_3	p_3x_3	$p_3x_3^2$	$p_3x_3^3$	$p_3x_3^4$
…	…	…	…	…	…	…
		1	$\mathbf{m_{1X}= \Sigma_i\, p_ix_i}$	$\mathbf{m_{2X}= \Sigma_i\, p_ix_i^2}$	$\mathbf{m_{3X}= \Sigma_i\, p_ix_i^3}$	$\mathbf{m_{4X}= \Sigma_i\, p_ix_i^4}$

Table 2: Representation of the first four moments of the profit of a casino game after one bet

If N consecutive bets are made then the total profit, T, is a random variable.

$$T = X_1 + X_2 + ... +X_N,$$

where X_i is the outcome on the i^{th} bet.

Assuming that the outcome from each bet is independent of the others, probability theory tells us that the MGF of random variable T is the product of MGFs of the X_i's:

$$\begin{aligned} M_T(t) &= E(\exp(X_1 + X_2 + ... +X_N)t)) \\ &= E(\exp(X_1 t))\, E(\exp(X_2 t)) \,....\, E(\exp(X_N t)) \\ &= M_{X1}(t)\, M_{X2}(t) \,....\, M_{XN}(t) \end{aligned}$$

If the bets are all on the same game and the same size, then the distribution of the profit from each bet is identical, and we obtain an important simplification:

$$M_T(t) = [M_X(t)]^N.$$

Taking logarithms we obtain a relationship between the CGFs:

$$K_T(t) = NK_X(t).$$

This relationship can be expressed in terms of the individual cumulants

$$k_{rT} = N\, k_{rX} \qquad \text{for all } r \geq 1.$$

Thus the cumulants of the total profit after N bets of the same size on a single game can be computed directly from the cumulants of the profit for a single bet.

Likewise the parameters of T are directly related to the parameters of X:

Mean $\mu_T = N\mu_X$

Standard Deviation $\sigma_T = \sqrt{N}\, \sigma_X$

Coefficient of Skewness $\gamma_T = \gamma_X/\sqrt{N}$

Coefficient of Excess Kurtosis $\kappa_T = \kappa_X / N$

2.3 Distribution of profits

There are several methods to obtain the probabilities of ending up with a certain amount of profit after a number of trials; which is also referred to as the distribution of profits. Simulation methods could be used to obtain approximation results, with the accuracy depending on the number of simulation runs. Algebraic computation can produce exact results, but requires computation time when the number of trials is "large". When the number of outcomes in a single bet is two (Win or Lose), the binomial formula can be used to calculate the exact distribution of profits after N bets (Packel, 1981).

Approximate algebraic methods can be applied effectively, such as the normal approximation to the binomial distribution, which utilizes the mean and the standard deviation (Packel, 1981). A better normal approximation method is the Normal Power Approximation, which utilizes the coefficients of skewness and excess kurtosis, as well as the mean and standard deviation. This is detailed below.

Let Z be a standardized variable, such that $Z = (T - \mu T)/ \sigma T$.
Variable Z has mean 0 and variance 1. Due to the symmetry of the normal distribution, variable Z has a skewness and excess kurtosis of 0. The Normal Power approximation for the cumulative distribution function (CDF) of Z, as given in Pesonen (1975), can be expressed in the following form:
$\text{Prob}(Z \geq z) = F(z)$ approx $\Phi(y)$
where $\Phi(.)$ is the cumulative normal distribution and
$y = z - 1/6\, \gamma_T (z^2 - I) + [1/36\, \gamma_T{}^2 (4z^3 - 7z) - 1/24\, \kappa_T (z^3 - 3z)]$.

Using y instead of z in the cumulative normal distribution provides improved accuracy when the distribution has skewness and excess kurtosis different from that of a standard normal distribution.

3. Poker machine regulations

The Australian / New Zealand Gaming Machine National Standard ("the Standard") has been developed by participants from various gaming regulators to outline gaming machine regulations, common to all jurisdictions. The Standard Rev 9.0 is used throughout this paper. The following is documented in the Standard:

The National Standard Working Party was established by the Australian and New Zealand gaming regulators on 21 March 1994. The purpose of the working party is to develop technical requirement documents to be used by each individual jurisdiction as the basis for working towards a common technical requirement for the evaluation of gaming machines. The intent of this document is to ensure gaming on gaming machines occurs in a manner that is: a) fair; b) secure; and c) auditable and that gaming machines are reliable in terms of these issues. It is anticipated that amendments to the Standard will occur on an annual basis and then be adopted by the participating jurisdictions. Individual jurisdictions can be expected to amend their respective requirements documents in advance of the National Standard where player fairness, security or auditability is considered to be jeopardized.

3.1 Displaying probabilities

Section 2.1 gives a mathematical description of a casino game. As detailed in the introduction, a consumer's decision as to the choice or how long to play a particular game, may consist of knowing the distribution of payouts. To calculate the distribution of payouts on a poker machine requires the probabilities associated with each particular payout. The probabilities on poker machines cannot be obtained from the playing rules (as is the case with table games), and therefore poker machines could be considered as being “unfair”.

As an approach towards responsible gambling and to increase consumer protection, the probabilities associated with each payout could be displayed on the machine, along with information that would advise players of their probabilities of ending up with an amount of profit after N trials or after a certain amount of time. Table 3 represents the first four moments of a sample poker machine after 1 trial, where x = 1/8000+1/800+1/80+1/8. The percent house margin is obtained as 10%. By applying the Normal Power approximation, Table 4 represents the probabilities of obtaining various profit intervals after 1080 spins and playing $1 hands. Assuming 18 spins per minute on a typical poker machine, a player is likely to spin 18*60=1080 spins per hour. Note that even though the expected loss after 1080 spins is -$108, a player has a 35.5% chance of losing more than $300.

Section 3.9.9 in the Standard states
A gaming machine must display the following information to the player:
a) the player’s current credit balance;
b) the current bet amount;
c) all possible winning outcomes, or be available as a menu item or help menu;
d) win amounts for each possible winning outcome or be available as a menu or help screen item;
e) the amount won for the last completed play (until the next play starts, or following player input related directly to the next play); and
f) the player options selected (e.g. bet amount, lines played) for the last completed play (until the next play starts, or following player input related directly to the next play).

Based on the above argument, clause d) could read
d) win amounts for each possible winning outcome *with associated probabilities* or be available as a menu or help screen item;

Outcome	Profit ($)	Probability	1st Moment	2nd Moment	3rd Moment	4th Moment
O_1	1000	1/8000	0.125	125	125000	125000000
O_2	100	1/800	0.125	12.5	1250	125000
O_3	10	1/80	0.125	1.25	12.5	125
O_4	1	1/8	0.125	0.125	0.125	0.125
O_5	0	2/5-x	0	0	0	0
O_6	-1	3/5	-0.6	0.6	-0.6	0.6
		1.000	**-0.100**	**139.475**	**126262.025**	**125125125.725**

Table 3: The first four moments of a sample poker machine after 1 trial.

Profit interval	Probability
< -300	0.355
-300 to -100	0.202
-100 to 0	0.099
0 to 100	0.089
> 100	0.251

Table 4: The probabilities of obtaining various payouts after 1080 spins of a sample poker machine.

3.2 Win amounts

There are two common forms of displaying payouts on casino games; odds or prices. Odds are typically used in casino table games. Odds are written in the form of x to y, which means that if the player is successful they receive a return of (x/y)*initial cost + initial cost or a profit of (x/y)*initial cost. Note the difference between the return and the profit. Prices are typically used in poker machines. Prices are written in the form of $z, which means that if the player is successful, they receive a return of $z*initial cost or a profit of $z*initial cost – initial cost. Odds and prices can easily be interchanged. Poker machines can be misleading since winning on a machine refers to returns rather than profits. For example, if a player bets $10 on a machine and receives a return of $7, then the player has made a loss of $3 or a profit of -$3. The machine indicates that the player has won $7. It can be misleading to players to use win amounts as the return rather than the profit.

From Section 3.9.9 in the Standard, an extra clause could be added as:
g) win amounts refer to profits (rather than returns)

3.3 Withdrawing credits

Poker machines in Australian gaming venues only allow deposits and withdrawals directly from the machine in multiples of $1. Players may withdraw money from the machine when remaining credits are less than $1, but this would require calling over the attendant and having the remaining money paid by the attendant rather than from the machine. This can be a relatively lengthy process and may require formal identification and a signature by the player. A player may decide to gamble less than $1, to leave the venue after a certain period of time or cash out after a relatively big win. For these scenarios the player may feel it is too much trouble to call over the attendant (given they can withdraw directly from the machine in multiples in $1) and decide to play out the remaining credit of less than $1 or leave the remaining credit in the machine. By having a withdrawal denomination of 5 cents (as well as $1) would allow the customer to withdraw an amount from the machine that had at most 4 cents remaining after the withdrawal. For example, with 99 cents remaining in the machine a player could withdraw a maximum of 95 cents (and the remaining 4 cents could not be rounded to 5 cents and withdrawn). This withdrawal process could be included as a regulation in the Standard.

3.4 Initial cost

Poker machines vary in the initial cost depending on the denomination of the machine (1c to $1 are common) and the number of lines (1 to 10 are common). The initial cost that a player chooses is given by the number of lines multiplied by the denomination of the machine. For example, 5 lines on a $1 machine, would amount to an initial cost of $5. The minimum initial cost that a player could play is 1c, by playing 1 line on a 1c machine. Assuming 18 spins per minute and a percent house margin of 13%, the player is expected to lose 0.13*18*60*$0.01 = $1.40 per hour. Playing 1 line on a 1c poker machine is one of the best games on offer for minimizing losses over a period of time and could be used as an approach to responsible gambling. There is currently no regulation as to how the total number of gaming machines at each venue is proportioned by different denominations. It is possible for example for a gaming venue to offer no machines with a 1c denomination, and this could be considered as unfair to players that want to gamble on poker machines for the purpose of minimizing losses. The Standard could include documentation as to how the total number of gaming machines at each venue is proportioned by different denominations.

3.5 Standard deviation

The mean profit is one of many distributional characteristics. As documented in section 2 other familiar distributional characteristics are the standard deviation, and coefficients of skewness and excess kurtosis. These four characteristics are applied directly in the Normal Power approximation formula to estimate the distribution of profits. Consider the two games given in Table 5. The initial cost for Game 1 is $1, whereas the initial cost for Game 2 is $2. The %HM for both games is 10% and they are effectively the same game. The standard deviations for both games are calculated as 0.995 for Game 1 and 1.990 for Game 2. These values are different because the standard deviation is dependent on the mean profits; which are -$0.10 for Game 1 and -$0.20 for Game 2.

The Standard in Section 3.9.17 has restrictions on the standard deviation and states: *The Nominal Standard Deviation (NSD) of a game must be no greater than 15. In determining the NSD for a game, the following conventions must be applied: a) Calculate standard deviation of the base game at minimum bet and single line play or equivalent.*

A problem with this argument is the base game could be an initial cost of $1 as given by Game 1 or $2 as given by Game 2. Increasing the initial cost will increase the standard deviation and therefore the standard deviation is not consistent amongst a particular poker machine. One way to solve this problem is to normalize all poker machine games to a fixed initial cost when performing calculations on the standard deviation, and then deciding on a reasonable NSD. Section 3.9.17 could be amended to reflect on this argument.

	Game 1		Game 2	
Outcome	Profit ($)	Probability	Profit ($)	Probability
O_1	1	0.45	2	0.45
O_2	-1	0.55	-2	0.55

Table 5: A comparison of two games where the initial cost for Game 1 is $1 and the initial cost for Game 2 is $2.

3.6 Coefficients of skewness and excess kurtosis

As was observed from section 3.5, the Standard has restrictions on the standard deviation such that the NSD of a game must be no greater than 15. The standard in Section 3.9.16b also gives restrictions on the probabilities and states: *The probability for attaining each winning pattern of symbols (offered in the Base game) must not be less than 1/7,000,000.*

Consider the game given in Table 6, where x=1/7,000,000. The only way a player can profit after any trial is in obtaining a $39,000 profit with the unlikely probability of 1 in 7 million. The standard deviation is calculated as 14.745 which is less than the 15 restriction. All the probabilities are not less than 1/7,000,000 which is in agreement with the above statement. Therefore, the probabilities and payouts for this game meet the regulations given in the Standard.

One way to prevent such a game occurring is to give restrictions in the Standard to the maximum amounts for the coefficients of skewness and excess kurtosis. For the game given in Table 6, the coefficient of skewness is calculated as 2643.553 and the coefficient of excess kurtosis is calculated as 6992243.972. In comparison the standard deviation and coefficients of skewness and excess kurtosis from the game given in Table 3 are 11.810, 76.686 and 6432.605 respectively.

Outcom	Profit	Probability	1st	2nd	3rd Moment	4th Moment
O_1	39,00	1/7,000,00	0.006	217.286	8474142.85	330491571428.57
O_2	0	0.86-x	0	0	0	0
O_3	-1	0.14	-0.140	0.140	-0.140	0.140
		1.00	**-0.134**	**217.426**	**8474142.71**	**330491571428.71**

Table 6: The first four moments of a casino game after one trial.

4. Conclusions

This paper has applied mathematical and logical reasoning to poker machine regulations to give suggestions for amendments to the “the Standard” with the purpose to increase consumer protection. These possible amendments consist of the following results:

1. The probabilities associated with the payouts should be displayed on the gaming machine
2. Win amounts should refer to profit payouts rather than return payouts
3. Gaming machines should allow players to withdraw amounts less than $1

4. The total number of gaming machines at each venue should be proportioned by different denominations
5. The standard deviation should be regulated on gaming machines with a fixed initial cost that is consistent across all machines.
6. There should be regulations for the coefficients of skewness and excess kurtosis.

5. References

Croucher J (2002) The Statistics of Gambling, Macquarie University Lighthouse Press.

Epstein R (1997) The Theory of Gambling and Statistical Logic, Academic Press.

Hannum R (2005) Practical Casino Math. The Institute for the Study of Gambling & Commercial Gaming, University of Nevada, Reno.

Packel E (1981) The mathematics of games and gambling. The Mathematical Association of America. Washington.

Pesonen E (1975) NP-technique as a tool in decision making. ASTIN Bulletin. 8(3): 359-363. http://www.casact.org/library/astin/vol8no3/359.pdf

Chapter 3: Safer Gambling

1. Introduction

In 2010, the author published an article on applying mathematics to poker machine regulations, where it was demonstrated that poker machines within Australia could be consider as "unfair", since important information on the probabilities associated with each outcome is not displayed directly on the machines, or can be calculated indirectly through the playing rules (as typically applies to casino table games). It was suggested as an approach towards responsible gambling and to increase consumer protection, to amend poker machine regulations such that the probabilities associated with each payout are displayed on each machine along with information that would advise players of the chances of ending up with a certain amount of profit after playing for a certain amount of time, Barnett (2010).

2. Poker machines

Consider the sample poker machine for a single trial given by table 1. The initial cost (the cost to play) is $1, and the initial cost is given directly on every machine within Australia. The payouts for each possible outcome (column 2) are also given directly on every machine in the form of prices. However, the probabilities associated with each outcome (column 3) are not given on the machine, and this fundamental piece of information is required to calculate the expected payouts (column 4), which enables the consumer to know how much he/she is expected to lose each spin of the machine. This expected loss is obtained as $0.10 from table

1. It can also be readily shown that there is a 13.9% chance of ending up with any profit and a 60% chance of ending with a loss (losing the initial cost of $1 to play the game).

Outcome	Profit ($)	Probability	Expected Profit ($)
O_1	1000	0.000125	0.125
O_2	100	0.00125	0.125
O_3	10	0.0125	0.125
O_4	1	0.125	0.125
O_5	0	0.261125	0
O_6	-1	0.6	-0.6
		1	**-0.1**

Table 1: The payouts with associated probabilities for a sample poker machine

Suppose there are 10 spins per minute on a typical machine. Then a player is likely to spin 10*60=600 spins per hour, and this allows such calculations as the chances of ending up ahead, more than $200 ahead or more than $200 behind after 1 hour of play. This information is represented in table 2 and shows that even though a player is expected to lose 600*0.1=$60 per hour, there is a 26.2% chance of losing more than $200 per hour and a 8.9% chance of winning more than $200 per hour. This type of information along with the type of information represented in table 1 could be readily displayed on each machine to enable the player to make decisions as to whether to play a particular machine and how long to play for. Furthermore, this statistical information could potentially be available for table games (e.g. blackjack, roulette) and distributed via computerized information kiosks at the particular gambling venue.

Hourly Profit ($)	Chances
<-200	26.2%
-200 to -100	37.1%
-100 to 0	19.4%
0 to 100	7.1%
100 to 200	1.3%
> 200	8.9%

Table 2: The chances of obtaining various payouts after 600 spins of a sample poker machine

3. Roulette

Figure 1 represents the relevant statistics for Roulette given Type of Bet: Red/Black, Initial Cost: $10 and Plays per Hour: 45. These input parameters are used to generate information containing the probabilities for each outcome on a single play, average loss per play, average loss per hour, and the chances of obtaining various payouts after 1 hour of play. The operations of the information kiosk are such that the input parameters (Type of Bet, Initial Cost, Plays per Hour) are defined by the player, and the statistical results are generated accordingly. Calculators similar to figure 1 can be obtained in Roulette (xlsx), the Big Wheel (xlsx) and a sample Poker Machine (xlsx).

Roulette

Parameters	
Type of Bet	Red/Black
Initial Cost	$10
Plays Per Hour	45

Outcome	Profit	Probability	Expected Profit
Player wins	$10	0.486	$4.86
Dealer wins	-$10	0.514	-$5.14
		1	**-$0.27**

Number of Plays	Average Loss
1	$0.27
45	$12.16

Hourly Profit	Chances
< -$100	9.1%
-$100 to $0	47.5%
$0 to $100	39.0%
> $100	4.4%
	100%

Figure 1: Relevant statistical information for the Red/Black bet in Roulette

4. Conclusions

This article demonstrates that by providing relevant mathematical information for each casino such as the average loss after an hour of play as well as the chances of obtaining various payouts after an hour of play, would inform the consumer on whether to play a particular game and for the length of time. Links to Excel calculators are given for Roulette, the Big Wheel and a Sample Poker Machine.

5. References

Barnett T (2010). Applying mathematics to poker machine regulations to increase consumer protection. Gaming Law Review and Economics 14(8), 601-607.

Chapter 4: How Much to Bet on Video Poker

1. Introduction

A question that arises whenever a game is favourable to the player, is how much to wager on each event? Whilst conservative play (or minimum bet) minimizes “large” fluctuations, it lacks the potential in maximizing the long-term growth of the bank. At the other extreme, aggressive play (or maximum bet) runs the risk of losing your entire bankroll even though the player has an advantage in each trial of the game. What is required is a mathematical formulation that informs the player of how much to bet with the objective of maximizing the long-term growth of the bank.

The famous Kelly criterion achieves this objective; as developed by John L. Kelly in a 1956 publication. The Kelly criterion has been most recognised in games when there are two outcomes – win \$x with probability p and lose \$y with probability 1-p. When there are more than two outcomes, a generalized Kelly formula is required and this is also discussed and given by John Kelly in the original 1956 paper. This article will apply the Kelly criterion when multiple outcomes exist (more than two) through working examples in video poker. The methodology could be used to assist "advantage players" in the decision-making process of how much to bet on each trial in video poker.

2. Kelly Criterion

2.1 Analysis of casino games and percent house margin

A casino game can be defined as follows: There is an initial cost C to play the game. Each trial results in an outcome O_i, where each outcome occurs with profit k_i and probability p_i. A profit of zero means the money paid to the player for a particular outcome equals the initial cost. Profits above zero represent a gain for the player; negative profits represent a loss. The probabilities are all non-negative and sum to one over all the possible outcomes. Given this information, the total expected profit $\sum E_i = \sum p_i k_i$. The *percent house margin* (%HM) is then $-\sum E_i/C$ and the total return is $1+\sum E_i/C$. Positive percent house margins indicate that the gambling site on average makes money and the players lose money. Negative percent house margins indicate that the game is favourable to the player and could possibly generate a long-term profit. Table 1 summarizes this information when there are m possible outcomes.

Outcome	Profit	Probability	Expected Profit
O_1	k_1	p_1	$E_1=p_1k_1$
O_2	k_2	p_2	$E_2=p_2k_2$
O_3	k_3	p_3	$E_3=p_3k_3$
…	…	…	…
O_m	k_m	p_m	$E_m=p_mk_m$
		1.0	$\sum E_i$

Table 1: Representation in terms of expected profit of a casino game with m possible outcomes.

2.2 Classical Kelly criterion

The well-established classical Kelly criterion is given by the following result:

Consider a game with two possible outcomes: win or lose. Suppose the player profits k units for every unit wager and the probabilities of a win and a loss are given by p and q respectively.

Further, suppose that on each trial the win probability p is constant with p + q=1. If kp-q>0, so the game is advantageous to the player, then the optimal fraction of the current capital to be wagered is given by:

$$b^* = (kp-q)/k$$

Consider the following example: A player profits \$2 with probability 0.35 and profits -\$1 with probability 0.65, as represented by Table 2. Since the expected profit of 2 x 0.35-0.65=0.05>0, the game is advantageous to the player and the optimal fraction is given by b*= (2 x 0.35-0.65)/2 = 0.025. If a player has a \$100 bankroll, then wagering 100*0.025 = \$2.50 on the next hand will maximize the long-term growth of the bank. If the player loses \$1 on that hand, then under the classical Kelly criterion, the next wager should be exactly 99*0.025 = \$2.475. Since fractions are often not allowed in gambling games, this figure should be rounded down to an allowable betting amount.

Outcome	Profit	Probability	Expected Profit
Win	\$2	0.35	\$0.70
Lose	-\$1	0.65	-\$0.65
		1.0	**0.05**

Table 2: A sample casino game to determine the optimal betting fraction under the Kelly criterion.

2.3 Kelly criterion for multiple outcomes

When there are multiple outcomes (more than two), as is the situation for video poker, a generalized Kelly formula is required from the classical Kelly formula. This generalized Kelly formula is given by Theorem 1.

Theorem 1:

Consider a game with m possible discrete finite mixed outcomes. Suppose the profit for a unit wager for outcome i is k_i with probability p_i for $1 \le i \le m$, where at least one outcome is negative and at least one outcome is positive. Then if $\sum_{i=1}^{m} k_i p_i > 0$ a winning strategy exists, and the maximum growth of the bank is attained when the proportion of the bank bet at each turn, b*, is the smallest positive root of $\sum_{i=1}^{m} \frac{k_i p_i}{1+k_i b^*} = 0$

Proof of Theorem 1:

Assume a constant proportion b of the bank is bet, with m discrete finite mixed outcomes. Let $B(1) / B(0)$ equal $1+ k_i b$ with probability p_i for i = 1 to m, where B(t) represents the player's bank at time t. Assume the player wishes to maximize

$g(b) = E[\log(B(1) / B(0))] = \sum_{i=1}^{m} pi \log(1 + kib)$. Without loss of generality let k_1 be the maximum possible loss. In the interval $0< b <-1/k_1$, $1+k_i b > 0$ since $k_i \ge k_1$ for i = 1 to m, so the logarithm of each term is real. Taking derivatives with respect to b,

$$\frac{dg(b)}{db} = \sum_{i=1}^{m} \frac{k_i p_i}{1+k_i b} = g'(b)$$

and

$$\frac{d^2 g(b)}{db^2} = -\sum_{i=1}^{m} \frac{k_i^{\,2} p_i}{(1+k_i b)^2} = g''(b)$$

Note that
(a) g(0) = 0,
(b) g'(0) > 0 follows directly from the requirement for a winning strategy (so you should bet something), and
(c) g''(*b*) < 0 for $0<b<-1/k_1$ (where k_1 is the MPL) so the first derivate has at most one zero root in this interval.

Hence whenever there is a winning strategy, the force of growth has a unique maximum given by the root of

$$\sum_{i=1}^{m} \frac{k_i p_i}{1+k_i b^*} = 0$$

Let g(b) represent the rate of growth of the bank which is the quantity to be maximized. Figure 1 shows a graphical representation of the Kelly criterion for the classical case (left) and when multiple outcomes exist (right). Let the value k_1 be the maximum possible loss in the multiple outcome game. The player's bank will grow as long as g(b)>0, and is maximized when g'(b)=0 (which is represented by g(b*) in Figure 1). It is important to note that a player's bank will not grow (and likely to hit ruin) when over betting the bankroll, even though the game is still favourable. This is represented on the graph for the values of b such that g(b)<0.

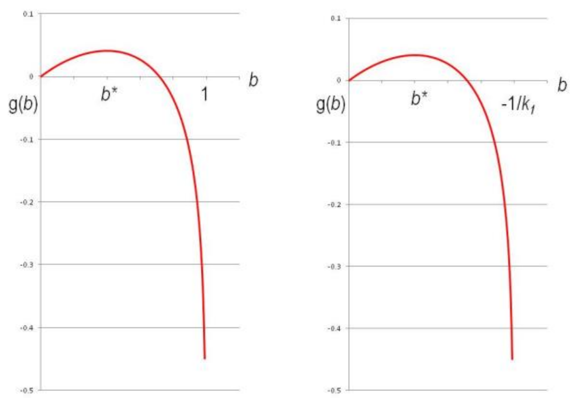

Figure 1: Graphical representation of the Kelly criterion for the classical case (left) and when multiple outcomes exist (right), where the optimal betting fraction of b* occurs at a maximum turning point on g(b). Value k_1 is the maximum possible loss in the multiple outcome game.

2.4 Kelly criterion for multiple outcomes: approximations to the optimal

According to Stanford Wong,

"Optimal bet size turns out to be the expected arithmetic win rate divided by the sum of the squares. For small expected win rates, such as you have in blackjack, the denominator is approximately equal to the variance."

This statement suggests approximating the optimal b* by

$$b' = \frac{\sum_{i=1}^{m} p_i k_i}{\sum_{i=1}^{m} p_i (k_i)^2}$$

The formula b' is simple to compute given the probabilities and profits for a single play of a game.

In The Wizard of Odds (http://wizardofodds.com/gambling/kelly.html) it is stated: "most gamblers use advantage/variance as an approximation, which is a good estimator". This second approximation can be written as

$$b'' = \frac{\sum_{i=1}^{m} p_i k_i}{\sum_{i=1}^{m} p_i (k_i)^2 - \left(\sum_{i=1}^{m} p_i k_i\right)^2} = \frac{advantage}{\text{var}\,iance}$$

Again, this formula is simple to compute given the probabilities and profits. It is obvious that 0<b'<b" (the denominator of b' is larger). It would be very useful if we could prove that b"<b*, because then b'' would be the superior approximation to b* always. Unfortunately, it is not so.

2.5 A demonstration that b'' can be larger than b*

Suppose m=2 possible outcomes occur in a game. Suppose that $k_1 + k_2 < 0$ with $k_1<0$ and $k_2>0$ and the conditions of Theorem 1 are satisfied. Then b*< b' which implies also that b*<b'' since b'<b''.

Proof:

Let $m_1 = p_1k_1 + p_2 k_2$ and $m_2 = p_1 k_1^2 + p_2 k_2^2$. Then

$b' - b^* = m_1/m_2 + m_1/(k_1k_2)$

$= m_1/(m_2 k_1 k_2) (m_2 +k_1 k_2)$

$= m_1/(m_2 k_1 k_2) (p_1 k_1^2 + p_2k_2^2 +(p_1+p_2) k_1 k_2)$

$= m_1/(m_2 k_1 k_2) (p_1 k_1 + p_2 k_2) (k_1+k_2)$

$= m_1^2/(m_2 k_1 k_2) (k_1+k_2)$

> 0 since $k_1 < 0$ and $k_2 > 0$ and $k_1+k_2<0$. Thus b'>b*.

Despite this, there are available criteria to show when either approximation is useful for managing the risk of ruin from over betting. Observe in Figure 1 that $g'(b) > 0$ for $0 < b < b^*$. It is simple to check in practical examples if either

(a) $g'(b') > 0$ and $0 < b' < -1/k_1$, or

(b) $g'(b'') > 0$ and $0 < b'' < -1/k_1$.

If condition (a) is satisfied, but condition (b) is not, then use b′. If condition (b) is satisfied, but condition (a) is not, then use b′′. When both these sets of conditions are satisfied it is preferable to work with b″, since in this case it is a closer approximation to the optimal value b*. Notice that the criteria do not require prior knowledge of the value of b*.

2.6 Numerical illustration of Kelly criterion in multiple (m>2) outcome game

Suppose m=3 and the outcome $k_1=-1$ occurs with probability 0.45, $k_2=1$ with 0.45, and $k_3=2$ with probability 0.10. The expected outcome is (-1)(0.45)+(1)(0.45)+2(0.10) = 0.20 >0, which is positive. The approximations are $b' = 0.2/[(-1)^2(0.45) + (1^2)*0.45 + (2^2)*0.10] = 0.2/1.3 = 0.1538$ and $b'' = 0.2/[1.3 - 0.2^2] = 0.2/1.26 = 0.1587$. k_1 is -1 and both b′ and b′′ are less than $-1/k_1 = 1$. $g'(b') = (-1)(0.45)/(1-1(b')) + (1)(0.45)/(1+1(b')) + 2(0.10)/(1+2(b')) = 0.0111$. Similarly, $g'(b'') = 0.0053$. Both conditions (a) and (b) are satisfied, so it is preferable to work with b′′. If someone has $1000, the bet should be $158.73, which likely would be rounded to $158.

3. Video Poker

3.1 Nonprogressive machines

Video Poker is based on the traditional card game of Draw Poker. Each play of the Video Poker machine results in 5 cards being displayed on the screen from the number of cards in the pack used for that particular type of game (usually a standard 52 card pack or 53 if the Joker is included as a wild card). The player decides which of these cards to hold by pressing the hold button beneath the corresponding cards. The cards that are not held are randomly replaced by cards remaining in the pack. The final 5 cards are paid according to the payout table for that particular type of game. The pay tables follow the same order as traditional Draw Poker. For example, a Full House pays more than a Flush. Without a thorough understanding of video poker, it should be clear in the analysis to follow on how Theorem 1 can be applied to determining an optimal bet size.

A pay table for the outcomes, profits, probabilities and expected profits for a Jacks or Better machine (known as "All American Poker") are given in Table 3. The probabilities were obtained using WinPoker (a commercial product available from the web http://www.zamzone.com/) and assume the player is always maximizing the expected profit on determining the correct playing strategies. Note that $1 is bet each game. It shows that the overall payback for this machine by playing an optimal strategy is 100.72%. The standard deviation is calculated as 5.18. The approximation formulas from section 2.4 give b′=0.0269436% and b″ = 0.0269437%. Since $g(b'') = 0.000789>0$ and $0<b''<1$, either approximation of b' and b" is useful for managing the risk of ruin from over betting. Theorem 1 is applied to determine a bet size for this Video Poker game, by using the payouts and probabilities given in Table 3. The solver function in Excel is used to calculate this value as b*=0.030679%. Example: With a $10,000 bankroll, Theorem 1 suggests that the player should initially bet $3.07 (likely to be round down to $3).

3.2 Progressive machines

Often a group of machines are connected to a common jackpot pool, which continues to grow until someone gets a Royal Flush. When this occurs, the jackpot is reset to its minimum value. Usually this minimum value would give a return less than 100%, which creates a win-win situation for the astute player and the house. The amount bet to obtain the jackpot is a fixed amount. Table 4 represents the probabilities of outcomes with three different jackpot levels for the "All American Poker" game. The $800 Jackpot was the game analyzed in section 3.1. The $250 and $1,200 jackpots give returns of 99.62% and 101.74% respectively. Notice that the probability of obtaining a Royal Flush increases as the jackpot increases. This is logical as a player would be more aggressive towards obtaining a Royal Flush with a larger jackpot.

Outcome	Return ($)	Profit ($)	Probability	Expected Profit ($)
Royal Flush	800	799	1 in 43,450	0.018
Straight Flush	200	199	1 in 7,053	0.028
Four of a Kind	40	39	0.00225	0.088
Full House	8	7	0.01098	0.077
Flush	8	7	0.01572	0.110
Straight	8	7	0.01842	0.129
Three of Kind	3	2	0.06883	0.138
Two Pair	1	0	0.11960	0.000
Jacks or Better	1	0	0.18326	0.000
Nothing	0	-1	0.58076	-0.581
			1.00	**0.0072**

Table 3: The profits and probabilities for the "All American Poker" game

Suppose a player has a bankroll of $11,000 and is required to bet $5 hands. What jackpot level is required to maximize the long-term growth of the player's bank under the Kelly criterion? Table 5 gives the results and can conclude that a jackpot level of $1,200 is required. A player would need a bankroll of about $17,000 in order to play the game at a jackpot level of $800.

Outcome	Return ($)	Prob: $250 Jackpot	Prob: $800 Jackpot	Prob: $1,200 Jackpot
Royal Flush	Jackpot	1 in 58,685	1 in 43,450	1 in 35,848
Straight Flush	200	1 in 7,272	1 in 7,053	1 in 6,999
Four of a Kind	40	0.00226	0.00225	0.00225
Full House	8	0.01101	0.01098	0.01096
Flush	8	0.01588	0.01572	0.01505
Straight	8	0.01851	0.01842	0.01846
Three of Kind	3	0.06899	0.06883	0.06888
Two Pair	1	0.11988	0.11960	0.11954
Jacks or Better	1	0.18406	0.18326	0.18336
Nothing	0	0.57924	0.58076	0.58132
		1.00	**1.00**	**1.00**

Table 4: The probabilities of outcomes for different jackpot levels for the "All American Poker" game

Jackpot	Return	Theorem 1	$11,000	$17,000
$250	99.62%	-	-	-
$800	100.72%	0.0307%	$3.38	$5.22
$1,200	101.74%	0.0468%	$5.15	$7.96

Table 5: Kelly criterion analysis for progressive jackpot machines

4. Practical difficulties

Despite the theoretical advances made above, it is impossible to effectively implement the optimal Kelly betting strategy on an "All American Poker" machine, or any other Video Poker game. There are three main sources of difficulty. The first is the existence of a minimum betting unit on a machine. The second is the need to round the bet to avoid fractions of a unit. Thirdly, to gain an edge in the long-run requires hitting Royal Flushes. In the non-progressive "All American Poker" machine, this occurs on average once every 43,450 trials. Therefore, a player's bankroll would need to withstand the downward drift between hitting jackpots to avoid over betting.

5. Conclusions

An analysis of casino games was given to identity when games are favourable to the player and could possibly generate a long-term profit. Analyses were given for both the classical Kelly (two outcomes) and the Kelly criterion when multiple outcomes exist (more than two). The Kelly criterion when multiple outcomes exist was applied to favourable video poker machines. In the case of non-progressive machines, an optimal betting fraction was obtained for maximizing the long-term growth of the player's bankroll. In the case of progressive machines, the minimum jackpot size was obtained as an entry trigger to avoid over betting, based on the player's bankroll. Approximation formulas when multiple outcomes exist were applied to video poker, and shown to be useful for managing the risk. The analysis developed in this paper could be used by "advantage players" to assist with bankroll management, which is recognized as an important component to long-term success.

6. References

Barnett T and Clarke S (2004). Optimizing Returns in the Gaming Industry for Players and Operators of Video Poker Machines. Proceedings of the International Conference on Advances in Computer Entertainment Technology. National University of Singapore, 212-216.

Barnett T (2009). Gambling your way to New York: A story of expectation. CHANCE. 22(3).

Epstein RA (1977). The Theory of Gambling and Statistical Logic. California: Academic Press.

Haigh J (1999). Taking Chances: winning with probability. New York: Oxford University Press.

Kelly J (1956). A New Interpretation of Information Rate. The Bell System Technical Journal. July, 917-926.

Thorp E (2000). The Kelly Criterion in Blackjack, Sports Betting, and the Stock Market. Finding the Edge. Editors: Olaf Vancura, Judy Cornelius and William Eadington, Institute for the Study of Gambling and Commercial Gaming, Nevada, pp 163-215.

Wong S (1981). What Proportional Betting does to your Win Rate, Blackjack World, 3:162–168.

Chapter 5: Optimizing Returns in the Gaming Industry for Players and Operators of Video Poker Machines

1. Introduction

Video Poker machines along with the tradition slots provide entertainment to the player in a variety of computer operated machines. Entertainment value of traditional machines involves watching the reels spinning around in the hope of producing a win each time the reels come to a stop. These machines involve no strategy, and the expected return to the players is fixed at around 87%. Australia owns 21% of the world's total slot machines, and proportionally have the highest number of these machines in the world. Clarke (2003) details how Poker Machines can be analyzed using Excel given basic information on the prize table and the number of symbols on each reel. While prize tables are displayed, other necessary information is not readily available to the player. On the other hand, it could be argued the entertainment value of Video Poker machines is greater. They require some thought process from the player in deciding which cards to hold on any hand. Optimal strategy depends on the various payouts, and while all the necessary information is available, the calculations are extremely difficult. With perfect strategy most Video Poker machines pay back 97-99%.

Australian casinos have over reacted to these high returns and Video Poker is diminishing. This suggests that these machines are not generating the profits compared to the regular slots and are being replaced by the latter. Jensen (1997) states: the actual pay back from video poker machines is 2 to 4% less than the maximum pay back based on perfect play. This implies that even if a machine can potentially pay back 97-99%, the actual pay back overall will be 93-97%. There appears to be a high demand from gamblers for Video Poker machines due to their possible high returns, as Jensen (1997) states: every time the payout schedules were improved, the game increased in popularity and, by 1985, it is estimated that over 25 percent of all slot machine players were playing video poker. Further evidence for these machines being beneficial for both the player and the casino is demonstrated through progressive jackpots.

Progressive machines offer a jackpot for obtaining a Royal Flush. As people play the machines, a proportion of the total amount gambled by the players is diverted to a jackpot pool, which continues to grow until someone gets a Royal Flush. When this occurs, the jackpot is reset to its predetermined minimum value and the cycle repeats itself. An individual at times can expect to receive a return over 100% if the jackpot gets high enough. At all times the machines are making the same expected amount per play, as the money in the jackpot pool will be won by someone eventually. We have a situation here where the players are attracted by the jackpot pool, while the casinos still get their percentage of the money gambled.

The use of progressive jackpots to entice players to partake in the game has been very effective in the past. Lotto type games are prime examples of this. Jackpots in Powerball are regularly highlighted in advertising campaigns, and Croucher (2002) shows the mean number of entries in Powerball is significantly greater for a higher Division 1 prize. Caribbean Stud Poker and Keno also offer progressive jackpots to entice players.

Our claim is: Video Poker can be beneficial for both the player and the house, and these machines including progressives should be available for use in all Australian casinos. This paper will support this claim by looking at a detailed analysis of a Joker Wild Video Poker machine (we will call it JW) still offered in Star City casino, Sydney. The current non-progressive version is included along with a progressive version for obtaining a Royal Flush. The calculations are simplified through WinPoker, specifically designed software for Video Poker analysis. A brief outline of the rules underlying Video Poker and how WinPoker calculates optimal strategies are discussed.

2. Calculating Optimal Strategy

Video Poker is based on the traditional card game of Draw Poker. Each play of the Video Poker machine results in 5 cards being displayed on the screen from the number of cards in the pack used for that particular type of game (usually a standard 52 card pack or 53 if the Joker is included as a wild card). The player decides which of these cards to hold by pressing the hold button beneath the corresponding cards. The cards that are not held are randomly replaced by cards remaining in the pack. The final 5 cards are paid according to the payout table for that particular type of game. The pay tables follow the same order as traditional Draw Poker. For example, a Full House pays more than a Flush. Epstein (1997) has calculated probabilities and strategies for a range of Poker-like games.

Unlike traditional Poker machines, a player is faced with many decisions on how to play each hand. For example being dealt the Ace of Hearts(AH), Ace of Diamonds(AD), 6 of Hearts(6H), 5 of Hearts(5H), 4 of Hearts(4H), a player might play the hand in three reasonable ways (a) Hold the Pair AH, AD and draw 3 cards, hoping for 3,4,5 of a Kind or a Full House, with the additional possibility of 2 Pair (b) Hold the AH, 6H, 5H, 4H and draw a single card, hoping for a Flush (c) Hold the 6H, 5H, 4H and draw 2 cards, hoping for a Straight Flush but with the possibility of 2 Pair, 3 of a Kind, Straight and a Flush.

The optimal strategies depend on the probabilities of getting the various possible hands and their payouts. This is an extremely difficult problem to solve and most Video Poker players would copy their Draw Poker strategies or go with a subjective decision. Clearly many players would choose non optimal strategies.

More formally, to optimize the return requires knowing which cards to hold from nC5 card combinations, where n = number of cards in the pack for the particular type of Video Poker game. A card is either held or not held resulting in 25 = 32 ways to hold the cards in any particular hand. In the above example, most sensible players would quickly discard the 29 other strategies and choose one of the three discussed above. However, differentiating between these three choices would be more difficult. WinPoker is a commercial product available from the web http://www.zamzone.com/ which calculates by complete enumeration, the number of all possible resultant hands and hence the expected return value (EV), for each of the 32 hold combinations. The highest EV is the best way to play that hand. For example, Table 1 gives 3 of the 32 rows from WinPoker analysis of the above hand from a 53-card pack (joker included). The notation used: N = Nothing, 2P = 2 Pair, 3K = 3 of a Kind, ST = Straight, FL = Flush, FH = Full House, 4K = 4 of a Kind, SF = Straight Flush, 5K = 5 of a Kind, JR = Joker Royal, RF = Royal Flush. It shows for example that for case (a), there are 10 from 48 draw possibilities that will result in obtaining a Flush and 38 from 48 draw possibilities that will result in nothing. Given the payouts also shown in Table 1 above each hand type, this gives an expected return of 10 48 ×4 = 0.83 for strategy (a). Similar calculations show (b) has a return of 0.75 and (c) a return of 0.65. (a), (b) and (c) do give the highest EV for the 32 hold combinations indicating our intuitive sense to the problem is indeed correct, and clearly (a) is the best way to play that hand.

Using these optimal strategies WinPoker can also calculate the probabilities for different hand types, the total return for the machine with perfect play and the variance associated with this return. We look first at machines without jackpots.

3. Non-Progressive machines

A pay table for the winning hands, their contribution to the total return, and their probabilities for JW are shown in Table 2. You can play 1,2,3,4 or 10 coins where 1 coin = \$1. The payouts are proportional to the amount bet, with the exception of 10 coins paying extra for obtaining a Royal Flush. From this game 70% of the time the player will not receive any return. The most likely winning hand is 3 of a kind occurring 13% of the time even though its payout it twice as much as Two Pair. Similar anomalies occur with Joker Royal and 5 of a Kind. 3 of a Kind generates the highest contribution to the total return by contributing 26% and a Royal Flush although not the lowest contributes only 1.31% for 1-4 coins and 2.66% for 10 coins. For this situation WinPoker calculated a return with perfect strategy of 92.3% for 1 coin and 93.6% for 10 coins and corresponding variances of \$44.04 and \$6428.29.

Although there are ${}^{53}C_5$ = 2,869,685 card combinations for JW, duplication of hands can reduce that number for the playing strategies. For example, if we are dealt 4 of a Kind with no joker on the first 5 cards we would hold the 4 of a Kind regardless of the suit and the extra 5th card. Therefore, one playing strategy would simply be 4 of a Kind - Draw 1, as opposed to 13×4×48 = 2496 different strategies if we were to list every combination containing 4 of a Kind with no joker.

Something not so obvious is being dealt the hand Jack of Hearts (JH), 8 of Hearts(8H), 6 of Diamonds(6D), 3 of Clubs (3C), 2 of Spades(2S). For a 1-4 coin play the optimal strategy calculated by WinPoker is to hold JH and 8H giving a return of 23.33%. However, for the same hand for a 10-coin play, the optimal strategy is to only hold JH for an increased return of 23.49%. The difference in strategies is a result of the increased payout for obtaining a Royal Flush when playing 10 coins. This and other changes in strategies account for the differences in probabilities between 1-4 coins and 10 coins.

EV	Hold	Total	N	1 2P	2 3K	3 ST	4 FL	5 FH	27 4K	50 SF	100 5K	500 JR	500 RF
0.83	AH,4H,5H,6H	48	38	0	0	0	10	0	0	0	0	0	0
0.75	AH,AD	17296	11559	2592	2781	0	0	228	135	0	1	0	0
0.65	4H,5H,6H	1128	981	27	18	57	38	0	0	7	0	0	0

Table 1: The Number of all Possible Resultant Hands for 3 Hold Combinations from the Hand

Hand Name	Payout($) 1 coin	Payout($) 10 coins	Probability 1-4 coins	Probability 10 coins	Return(%) 1-4 coins	Return(%) 10 coins
Royal Flush	500	10,000	1 in 38,069	1 in 37,615	1.31	2.66
Joker Royal	500	5,000	1 in 8,870	1 in 8,790	5.64	5.69
5 of a Kind	100	1,000	1 in 10,795	1 in 10,794	0.93	0.93
Straight Flush	50	500	1 in 1,603	1 in 1,607	3.12	3.11
4 of a Kind	27	270	1 in 119	1 in 119	22.74	22.75
Full House	5	50	1 in 65	1 in 65	7.69	7.69
Flush	4	40	1 in 55	1 in 55	7.26	7.27
Straight	3	30	1 in 43	1 in 44	7.02	6.87
3 of a Kind	2	20	0.129	0.129	25.84	25.88
Two Pair	1	10	0.107	0.107	10.71	10.73
Nothing	0	0	0.697	0.698	0.00	0.00
			1	**1**	**92.3**	**93.6**

Table 2: The Payout and Probabilities Under Optimal Strategy for Different Hand Types

With reference from Jensen (1997) p166 and Frome (1997) p84 a total of 43 different strategies dealt on the initial five cards represent an almost perfect play and are represented by Tables 3 and 4. “Almost” meaning there are situations where we would choose to deviate from optimal play because the increased number of strategies is not worth the small increase in return. Being dealt the JH, 8H, 6D, 3C, 2S we would choose to draw 5 new cards for both 1-4 coins and 10 coins giving returns of 22.6% and 22.7% respectively. The hands are in descending order of play with the corresponding number of cards to draw. For example, being dealt AH, AD, 5H, 4H, 6H a player would expect to receive a higher return by holding the 4 Card Flush (6H, 4H, 5H, AH) as opposed to holding the Pair (AH, AD) or the 3 Card Straight Flush (6H, 4H, 5H). This was also outlined in Table 1. Note the strategies given in Tables 3 and 4 are optimal for both the 1 coin and 10-coin payout. The increase in payout for 10 coins does not effect the order of the hands. However, as the payout for a Royal Flush continues to grow, as is the case for progressive machines, the order of the hands for optimal return will change.

Clearly optimal strategy for Video Poker is difficult to determine and tabulate. Joker Wild machines are even more complicated to play correctly due to the added Joker. Most players do not read books on Video Poker or access software such as WinPoker. This is further evidence that the actual pay back from Video Poker machines will be less than the pay back for optimum play.

Hand Name	Draw
Royal Flush	0
Straight Flush	0
4 Card Royal Flush	1
4 of a Kind	1
Full House	0
3 of a Kind	2
4 Card Straight Flush	1
Flush	0
Straight	0
4 Card Inside Straight Flush	1
3 Card Royal	2
Two Pair	1
4 Card Flush	1
Pair	3
3 Card Straight Flush	2
4 Card Straight	1
3 Card Inside Straight Flush	2
3 Card Double Inside Straight Flush	2
2 Card Royal	3
4 Card Inside Straight	1
2 Card Straight Flush	3
3 Card Flush	2
2 Card Inside Straight Flush	3
Other	5

Table 3: Optimal Draw Strategies for No Joker Hands in Video Poker Hand Name Draw Royal

4. Progressive machines

Often a group of machines are connected to a common jackpot pool, that continues to grow until someone gets a Royal Flush. When this occurs, the jackpot is reset to its minimum value. Let's assume this is the situation for JW and generally occurs by playing maximum coins i.e $10 per play. As the jackpot increases so does the expected return to the player. When the progressive meter reaches a certain level, it can expect to return over 100% of the money gambled. To optimize the return for progressives requires changing strategies to suit the meter. In particular the No Joker hands consisting of 2 Card Royal and 3 Card Royal will have an increase in EV and as a result move up the table. Table 5 represents the returns for different values of the jackpot. As indicated a jackpot of at least $33,700 is needed for this game to be favourable to the player. For this jackpot amount the 3 Card Royal (from Table 3) will be situated just below the 3 of a Kind and the 2 Card Royal just below the 3 Card Inside Straight Flush. If someone adopted the strategy for the non-progressive JW then their return would be less than 100.0%, an unfavourable game.

Hand Name	Draw
Joker Royal	0
5 of a Kind	0
Straight Flush	0
4 of a Kind	1
4 Card Joker Royal	1
4 Card Inside Joker Royal	1
4 Card Straight Flush	1
Full House	0
4 Card Double Inside Straight Flush	1
3 of a Kind	2
Flush	0
4 Card Double Inside Straight Flush	1
Straight	0
3 Card Joker Royal	2
3 Card Straight Flush	2
3 Card Inside Straight Flush	2
4 Card Straight	1
3 Card Double Inside Straight Flush	2
Joker + Highest Card	3

Table 4: Optimal Draw Strategies for Joker Hands in Video Poker Hand Name Draw Joker

Jackpot($)	Return to player(%)
15,000	94.9
20,000	96.3
25,000	97.6
30,000	99.0
33,700	100.0
35,000	100.4

Table 5: The Expected Returns to the Player for Different Jackpot Levels Jackpot

Progressive machines are ideal for both the player and the house. Players will be enticed to the machines by the high jackpots and so the machines will turn over more money than just being idle. The amount in the jackpot pool is already lost to the casino and won't affect the overall house margin for the game. Some players will adjust their playing strategies to suit the jackpot meter. When this occurs, the machines will return a smaller percent of the non-jackpot prizes to the players, even though the players can expect a return over 100%. The casinos percentage take from the new money will increase. This means the casinos are making more money per play from the skilled players who are changing their strategies to suit the meter. This gambling paradox will continue to larger extremes providing the jackpot is increasing. Thus, we have a situation favourable to both the casino and the gambler.

The casinos must decide what percentage of the money gambled goes towards the jackpot pool. This will affect the overall house margin and how fast the jackpot grows. The casino should be reluctant to put more than 6% of the money gambled into the jackpot pool for JW. We have an optimization problem here for the casinos. If not enough money is contributed to the jackpot pool, the players might be turned away by the jackpot meter not being high enough. If too much money is contributed to the jackpot pool, then the house margin might not be satisfactory for the casino.

The probability the jackpot increases by a certain amount is given by f(y) = (1 − p)y where p = probability of hitting the jackpot on a single game, y = number of games played. Although p differs depending on the strategies each player adopts at different jackpot levels, we will take p = 1 37,615 from Table 2 as the optimal strategy when playing 10 coins on a minimum jackpot level. Table 6 represents these probabilities when 2% of the money gambled is added to the jackpot pool and also for 3%. For 2%, only 4% of the time will a player have the odds in their favour. For 3% this increases to 12%.

Jackpot Level($)	2%	3%
15,000	0.51	0.64
20,000	0.26	0.41
25,000	0.14	0.25
30,000	0.07	0.17
33,700	0.04	0.12
35,000	0.04	0.11

Table 6: The proportion of time the jackpot reaches a certain level for different amounts of money added to the jackpot pool

5. Conclusions

What entices players to a casino is the variety of games available. Video Poker machines complement the regular slots as another source of entertainment for computer operated machines. When the player gets bored pressing a button on the slots, they could switch over to Video Poker which requires a certain level of mental stimulus, and vice versa.

The strategies involved in Video Poker to optimize the returns make it an interesting game to analyze. Furthermore, the high jackpots generated by progressive machines make them attractive to play for the possibility to obtain an edge over the house. The optional multi-play feature, where up to 10 games can be played at the same time, make them even more attractive for the player to capitalize on favourable situations. As indicated throughout this paper the house will always obtain their percentage of the money gambled. In fact, the house margin increases as the jackpot continues to increase as paradoxically gamblers alter their strategy to increase their return.

6. References

Clarke S (2003). Analysing Poker Machines Using Excel. WDSI 2003 proceedings, K. S. Dhir (editor), Hawaii 2003.

Cross M and Markowski E (1997). Secrets of the Lottery Wizard. CHANCE, 10: 25-32.

Croucher J (2002). The Statistics of Gambling. Macquarie University Lighthouse Press.

Epstein R (1997). The Theory of Gambling and Statistical Logic. Academic Press.

Frome L (1997). Winning Strategies for Video Poker, Compu-Flyers, 1997.

Jensen M (1997). Video Poker for the Winner. Cardoza Publishing.

Chapter 6: Automating Online Video Poker for Profit

1. Introduction

In 2004 the author published a paper on a win-win situation for both the player and the casino in video poker machines (Barnett and Clarke, 2004). It was shown that video poker is interesting to analyze due to the changing strategies produced by progressive jackpots. These allow players to have the odds in their favour, while paradoxically allowing the casinos to increase their percentage margin on extra turnover as the jackpot rises. This situation makes progressive jackpots beneficial for both the player and the casino.

The arrival of online casinos in 1996 brought games that you would find at land-based casinos (roulette, blackjack, video poker, etc.) to the computer screens of gamblers all over the world (Balestra, 2005). Whilst the focus in Barnett and Clarke (2004) is for the player to obtain a long-term profit from video poker in land-based casinos, a better approach is through online casinos - as this makes automation of systems possible across several computers. For example, if the hourly win rate using one machine is $10/hour, then the same code used to automate one machine could be used on ten machines for an hourly win rate of 10*$10/hour= $100/hour.

This article analyzes automating online progressive video poker games for profit. This requires the development of computer code to automate the process. The algorithm and implementation of the code will not be addressed in this article, but rather identifying which game could be used as the starting point. For example, a game that is favourable on a regular basis of several days per week on average would appear to be better than a game which is only favourable on several days per year on average. Bankroll management and playing strategies are also analyzed in this article.

2. Outline of video poker

Video Poker is based on the traditional card game of Draw Poker. Each play of the Video Poker machine results in 5 cards being displayed on the screen from the number of cards in the pack used for that particular type of game (usually a standard 52 card pack or 53 if the Joker is included as a wild card). The player decides which of these cards to hold by pressing the hold button beneath the corresponding cards. The cards that are not held are randomly replaced by cards remaining in the pack. The final 5 cards are paid according to the payout table for that particular type of game. The pay tables typically follow the same order as traditional Draw Poker. For example, a Full House pays more than a Flush. Epstein (1997), Croucher (2002), Frome (1997) and Jensen (1999) have calculated probabilities and strategies for a range of Poker-like games.

Barnett and Clarke (2004) analyze a Joker Wild video poker game. The initial cost to play is 1,2,3,4 or 10 coins where 1 coin = $1. The payouts are proportional to the amount bet, with the exception of 10 coins paying extra for obtaining a Royal Flush. The probability associated with each outcome depends on the strategies of which cards are held for each of the ${}^{53}C_5$ card combinations (there are 53 cards in a standard deck with one joker included). WinPoker is a commercial product available from the web http://www.zamzone.com/ which calculates by complete enumeration, the number of all possible resultant hands and hence the expected return value, for each of the 2^5=32 hold combinations. The highest EV is the best way to play that hand. Similar calculations and strategies can be obtained from various online sites such as VP Genius www.vpgenius.com. A pay table for the winning hands, their contribution to the total return, and their probabilities for Joker Wild are shown in table 1. From this game 70% of the time the player will not receive any return. The most likely winning hand is 3 of a kind occurring 13% of the time even though its payout it twice as much as Two Pair. Similar anomalies occur with Joker Royal and 5 of a Kind. 3 of a Kind generates the highest contribution to the total return by contributing 26% and a Royal Flush although not the lowest contributes only 1.31% for 1-4 coins and 2.66% for 10 coins. For this situation WinPoker calculated a return with perfect strategy of 92.3% for 1 coin and 93.6% for 10 coins and corresponding variances of $44.04 and $6428.29.

Hand Name	Payout ($) 1 coin	Payout ($) 10 coins	Probability 1-4 coins	Probability 10 coins	Return (%) 1-4 coins	Return (%) 10 coins
Royal Flush	500	10,000	1 in 38,069	1 in 37,615	1.31	2.66
Joker Royal	500	5,000	1 in 8,870	1 in 8,790	5.64	5.69
5 of a Kind	100	1,000	1 in 10,795	1 in 10,794	0.93	0.93
Straight Flush	50	500	1 in 1,603	1 in 1,607	3.12	3.11
4 of a Kind	27	270	1 in 119	1 in 119	22.74	22.75
Full House	5	50	1 in 65	1 in 65	7.69	7.69
Flush	4	40	1 in 55	1 in 55	7.26	7.27
Straight	3	30	1 in 43	1 in 44	7.02	6.87
3 of a Kind	2	20	0.129	0.129	25.84	25.88
Two Pair	1	10	0.107	0.107	10.71	10.73
Nothing	0	0	0.697	0.698	0.00	0.00
			1	**1**	**92.3**	**93.6**

Table 1: The payout and probabilities for Joker Wild video poker under optimal strategy for different hand types

In progressive video poker a group of machines are connected to a common jackpot pool that continues to grow until someone gets a specific outcome (such as a Royal Flush). When this occurs, the jackpot is reset to its minimum value. The initial cost is fixed, although a specific condition is often required such as playing maximum coins. For example, 10 coins or $10 per play would be required for the Joker Wild game above. Since the payout for obtaining a Royal Flush changes with the jackpot meter, the probability associated with each outcome (assuming an optimal strategy) will change as the jackpot meter increases and consequently increase the probability of hitting a Royal Flush (Barnett and Clarke, 2004). WinPoker calculated that a jackpot of at least $33,700 is needed for Joker Wild to be favourable and potentially make a long-term profit to the player. Clearly, there are added complexities in analyzing progressive video poker games (dependency of trials) compared to non-progressive video poker games (independent trials).

Finally, some games have a multi-hand feature that allow up to a possible 100 hands. Just as in single-hand video poker, you choose which of the five cards to hold from the base hand, and this is copied to each of the remaining hands played. When you're ready to draw new cards, click the Deal button. For each hand you play, a random set of replacement cards is drawn for each successive hand.

3. Online video poker games

3.1 Software providers

Most online casinos use independent software providers to establish which games are available to the player for their own casino. VP Genius outlines many progressive jackpot video poker games with the corresponding software provider. This is represented in table 2 along with the number of hands, amount bet, return to player at the minimum jackpot level, jackpot at break-even and the probability of hitting a jackpot at break-even. There are six software providers offering progressive video poker games consisting of:

Boss Media - offer four games of a progressive Jacks or Better game where the jackpot is won with any Royal Flush. Players need to form a minimum poker hand of two Jacks in order to win. Any cards lower than this will not be considered a winning combination and the player will lose his or her bet. The cost to play each game is in proportion to the payouts and therefore all four games give the same return to player at the minimum jackpot level.

PlayTech - offer two quite different games. MegaJacks is similar to Jackpot Poker (25c) as both games require $1.25 as the initial cost and the same type of winning outcomes. However, MegaJacks give returns of $31.25, $11.25 and $7.50 for a 4 of a Kind, Full House and Flush respectively. In comparison, Jackpot Poker (25c) gives returns of $25, $8.75 and $6.25 for a 4 of a Kind, Full House and Flush respectively. Consequently, MegaJacks gives a higher return to player at the minimum jackpot level. Jacks or Better 10-play requires multiple hands where 10 games are played. The initial cost is $12.50. Only the base hand qualifies for the jackpot, such that if you get a Royal Flush on the base hand then the remaining nine hands are always paid $3,200, and if you get a Royal Flush on one of the remaining hands, then only that particular hand is paid $3,200.

Cryptologic - offer four games where the same jackpot is used for Super Jackpot (25c) and Super Jackpot Bonus (25c), and the same jackpot is used for Super Jackpot ($1) and Super Jackpot Bonus ($1). In Super Jackpot Bonus, discards are returned to the deck before new cards are drawn and bonus payouts are made if any or all of those cards are dealt back into your hands. Also 1,3,5 or 10 hands can be played with all four games, with the jackpot being awarded if a Royal Flush is obtained on the base and any remaining hands.

Microgaming - offer two games of Jackpot Deuces and SupaJax. To win the jackpot in Jackpot Deuces, the player needs to get a Royal Flush in Diamonds. In SupaJax, a 53-card deck is used, where the extra card is the SupaJax card. To win the jackpot, the player must get Four Jacks with the SupaJax card.

Gamesys - the progressive game is based on Jacks or Better, and to win the jackpot you need to get a Royal Flush in Spades.

Each software provider above services many casinos with the same progressive jackpot game. Therefore, a player could have an account with 10 casinos that utilize the same game, and this is very useful for automating a system across several computers.

Game	Software	Hands	Amount Bet	Return at minimum jackpot	Jackpot at break-even	Probability jackpot at break-even
Jackpot Poker (25c)	Boss Media	1	$1.25	95.4828%	$3,098	1 in 31,817
Jackpot Poker (50c)	Boss Media	1	$2.50	95.4828%	$6,196	1 in 31,817
Jackpot Poker ($1)	Boss Media	1	$5	95.4828%	$12,392	1 in 31,817
Jackpot Poker ($5)	Boss Media	1	$25	95.4828%	$61,960	1 in 31,817
MegaJacks	PlayTech	1	$1.25	98.3735%	$1,220	1 in 35,939
Jacks or Better 10-play	PlayTech	10	$12.50	96.1472%	$17,296	1 in 31,942
Super Jackpot Bonus (25c)	Cryptologic	1	$1.25	98.3164%	$1,830	1 in 36,481
Super Jackpot Bonus ($1)	Cryptologic	1	$5	98.3164%	$7,320	1 in 36,481
Super Jackpot (25c)	Cryptologic	1	$1.25	97.2984%	$2,167	1 in 32,573
Super Jackpot ($1)	Cryptologic	1	$5	97.2984%	$8,666	1 in 32,573
Jackpot Deuces	Microgaming	1	$5	94.3421%	$39,617	1 in 109,516
Super Jax	Microgaming	1	$5	92.7916%	$52,417	1 in 109,510
Jacks or Better (G)	Gamesys	1	$2.50	96.1472%	$28,230	1 in 116,079

Table 2: General information for a range of online progressive video poker games

3.2 Characteristics

To compare different video poker games requires a set of characteristics. The following will be used and represented in table 3 for each game given in table 2:

Probability reaching break-even - each play of the game results in a proportion of the initial cost contributed to the jackpot pool, which continues to grow until someone hits a specific outcome (such as a Royal Flush). When this occurs, the jackpot is reset to its minimum value. Barnett and Clarke (2004) calculate the proportion of time the jackpot reaches a certain level for different amounts of money added to the jackpot pool for the Joker Wild game outlined in section 2. Online casinos or online software providers do not readily give the proportion of money contributed to the jackpot pool for their progressive video poker games. We will assume that 2% of the initial cost is contributed to the jackpot pool and calculate the proportion of time the jackpot reaches a break-even game.

Burn rate - the burn rate is the expected percentage of money lost to the player each play of the game without hitting the jackpot. Since the chance of hitting the jackpot decreases as the jackpot increases, the burn rate is taken at break-even. From table 3, MegaJacks has the lowest burn rate at 2.72%.

Probability jackpot in 40,000 trials - this is the probability of hitting a jackpot in 40,000 trials

Expected loss without a jackpot - this is the expected amount of money lost without a hitting a jackpot given that the probability of hitting the jackpot over a number of trials is 95%.

Game	Probability reaching break-even	Burn rate	Probability jackpot in 40,000 trials	Expected loss without a jackpot
Jackpot Poker (25c)	9.79%	7.79%	71.6%	$9,256.99
Jackpot Poker (50c)	9.79%	7.79%	71.6%	$18,513.97
Jackpot Poker ($1)	9.79%	7.79%	71.6%	$37,027.95
Jackpot Poker ($5)	9.79%	7.79%	71.6%	$185,139.74
MegaJacks	36.42%	2.72%	67.1%	$3,649.10
Jacks or Better 10-play	12.99%	6.59%	71.4%	$78,726.98
Super Jackpot Bonus (25c)	27.82%	4.01%	66.6%	$5,461.57
Super Jackpot Bonus ($1.25)	27.82%	4.01%	66.6%	$21,846.30
Super Jackpot (25c)	23.86%	5.32%	70.7%	$6,471.66
Super Jackpot ($1.25)	23.86%	5.32%	70.7%	$25,886.63
Jackpot Deuces	4.24%	7.23%	30.6%	$118,217.08
Super Jax	2.08%	9.57%	30.6%	$156,413.58
Jacks or Better (G)	1.09%	4.86%	29.1%	$42,061.53

Table 3: Four characteristics for comparing a range of online progressive video poker games

A game that is favourable on a regular basis of several days per week on average would appear to better than a game which is only favourable on several days per year on average. The rate at which a game is played is not constant, but rather highly dependent on the return to player. For example, from observation, the rate at which the jackpot increases when the game is favourable will be greater than the rate at which the jackpot increases when the game is unfavourable to the player. The probability of a game reaching break-even can be used as a guide as to which games are more likely to be favourable to the player as a proportion of days in a year. From table 3, MegaJacks has the highest probability of reaching break-even at 36.42% followed by Super Jackpot Bonus at 27.82%. In order to obtain a long-term profit in progressive video poker requires hitting jackpots. The bankroll will decrease in between hitting jackpots and losing an entire bankroll is still very possible even though the game is favourable. The burn rate, the probability of a hitting a jackpot in a number of trials and the expected loss without hitting a jackpot can be used as a guide as to whether to play a particular game, and more importantly which game should be used as the starting game for automating a system. Out of all the games represented in table 3, MegaJacks has the highest probability of reaching break-even, the lowest burn rate and the lowest expected loss without hitting a jackpot. Also, the probability of hitting a jackpot in a number of trials for MegaJacks is relatively high in comparison to the other games. Therefore, there is indication that MegaJacks could be used as the starting game for automating an online system for profit.

Section 4 will give a detailed analysis of MegaJacks which could be used in bankroll management in determining the size of the jackpot to start playing the game.

4. MegaJacks

Table 4 represents the payouts, probabilities and return for each hand type for MegaJacks given a 100% total return. Table 5 represents the return to player at different jackpot levels with the corresponding probability of reaching a specified jackpot level. Table 6 represents the probability of hitting a jackpot in a number of trials and the corresponding expected amount lost without hitting a jackpot. The current jackpot meter for progressive video poker games can be obtained online through various sites such as Slot Charts and Awesome Jackpots.
http://www.slotcharts.com/video-poker.php
http://www.awesomejackpots.com/jackpots/video-poker/

Hand	Payoff	Probability	Return
Royal Flush	$1220	0.0028%	2.7157%
Straight Flush	$62.50	0.0111%	0.5564%
Four of a Kind	$31.25	0.2355%	5.8878%
Full House	$11.25	1.1485%	10.3367%
Flush	$7.50	1.1122%	6.6729%
Straight	$5.00	1.1308%	4.5234%
Three of a Kind	$3.75	7.4165%	22.2495%
Two Pair	$2.50	12.8919%	25.7839%
Jacks or Better	$1.25	21.2737%	21.2737%
Nothing	$0	54.7769%	0%
		100%	100%

Table 4: The payout, probabilities and return for each hand type for MegaJacks given 100% total return to the player

Jackpot	Return to player	Probability of reaching a jackpot level
$1,220	100.0000%	36.42%
$2,000	101.8635%	15.29%
$5,000	109.4248%	0.54%
$10,000	123.0732%	1 in 48,157
$20,000	152.0209%	1 in 12,575,570

Table 5: Return to player at different jackpot levels with the corresponding probability of reaching a specified jackpot level for MegaJacks

Trials	Probability	Expected lost without a jackpot
40,000	67.1%	$1357.85
80,000	89.2%	$2715.70
100,000	93.8%	$3394.63
120,000	96.5%	$4073.55

Table 6: The probability of hitting a jackpot in a number of trials and the corresponding expected amount lost without hitting a jackpot for MegaJacks

4.1 Playing strategies

As mentioned in section 2, playing strategies can be obtained using software such as WinPoker or online sites as VP Genius. However, the strategies may change as the jackpot meter continues to grow (Barnett and Clarke, 2004). Consider the hand JH, QH, KH, KS, KC. As given in table 7 when the jackpot is $1,220 (100% return to player) holding the KH, KS, KC gives an expected return of 4.3025 units compared to holding the JH, QH, KH which gives an expected return of 3.8511 units. In comparison when the jackpot reaches $5,000 (109.42% return to player) holding the KH, KS, KC gives an expected return of 4.3025 units compared to holding the JH, QH, KH which gives an expected return of 4.4052 units.

Return $1,220	Return $5,000	Hold	Total	N	J+	2P	3K	S	F	FH	4K	SF	RF
4.3025	4.3025	__KKK	1081	0	0	0	969	0	0	66	46	0	0
3.8511	4.4052	JQK__	1081	699	286	15	6	30	43	0	0	1	1

Table 7: The number of all possible resultant hands for 2 hold combinations from the hand JH, QH, KH, KS, KC

4.2 Kelly criterion

Barnett (2011) analyzed casino games to identity when games are favourable to the player and could possibly generate a long-term profit. Analyses were given for both the classical Kelly (two outcomes) and the Kelly criterion when multiple outcomes exist (more than two). The Kelly criterion when multiple outcomes exist was applied to favourable video poker machines. In the case of non-progressive machines, an optimal betting fraction was obtained for maximizing the long-term growth of the player's bankroll. In the case of progressive machines, the minimum jackpot size was obtained as an entry trigger to avoid over-betting, based on the player's bankroll. Table 8 represents the minimum bankroll size requirement for different jackpot levels in MegaJacks such that a player would not be over betting under the Kelly Criterion.

Another common problem that often arises in gambling is obtaining the probability of losing one's entire bankroll given a favourable game. A recursive solution (which assumes independent trials) was derived by Evgeny Sorokin and posted on Arnold Snyder's Blackjack Forum Online http://www.blackjackforumonline.com/content/VPRoR.htm. This derivation could also be used as a guide to determine the jackpot size for playing the particular video poker game.

Jackpot	Return	Bankroll
$1,300	100.18%	$22,166
$1,500	100.65%	$7,383
$2,000	101.86%	$3,534
$2,500	103.09%	$2,689
$3,000	104.34%	$2,318

Table 8: Kelly criterion analysis for MegaJacks

5. Conclusions

The arrival of online casinos in 1996 brought games that you would find at land-based casinos to the computer screens of gamblers all over the world. A major benefit in online casinos is in the automation of systems across several computers for favourable games, as this has the potential to make a significant amount of profit. This article applied this concept to online progressive video poker games. By establishing a set of characteristics to compare different games, analyses were carried out to identify which game should be the starting point for building an automated system. There are further incentives in online gambling through bonuses, cash backs and affiliate income which should be explored if profiting from online gambling is going to be a long-term business.

Whilst the legal implications of automating online systems have not been addressed in this article, the reader may be interested in this dispute on a player winning $46,000 through online video poker at Easy Street Sports casino.
http://wizardofodds.com/casinos/easy-street-sports.html

6. References

Balestra (2005). How to Win at Online Gambling. Funtastic Limited, Australia.

Barnett T and Clarke S (2004). Optimizing returns in the gaming industry for players and operators of video poker machines. Proceedings of the International Conference on Advances in Computer Entertainment Technology. National University of Singapore , 212-216.

Barnett T (2011). How much to bet on video poker. CHANCE. 24(2).

Croucher J (2002). The Statistics of Gambling, Macquarie University Lighthouse Press, 2002.

Epstein R (1997). The Theory of Gambling and Statistical Logic, Academic Press.

Frome L (1997). Winning Strategies for Video Poker, Compu-Flyers.

Jensen M (1999). Video Poker for the Winner, Cardoza Publishing.

Chapter 7: Gambling Your Way to New York: A Story of Expectation

1. Introduction

InterCasino, an online gambling web site at https://www.intercasino.com/, recently offered a free flight to New York as a promotional offer to new customers. From Australia, the value of the flight was about $2,000. To qualify, players had to complete their wagering requirements for one of the online games. Assuming the player was happy with the terms and conditions of the offer, two obvious questions are:
Is the offer worthwhile?
Which is the best game to play?

To answer these questions, we need probability theory, details concerning the games, and clear criteria on which to base our answers.

2. Rules of the Promotion

The following conditions were posted on the InterCasino web site:

1. Deposit $100 or more using any of our payment methods from March 1–31, 2008

2. We will then match the first $100 of this deposit, giving you another $100 in your account (you may need to click refresh to see your balance update)

3. All you need to do then is fulfil the bonus wagering requirement by wagering $5,000 on any of the games below before midnight on March 31, 2008

4. Deposit $100 or more using any of our payment methods between April 1–30, 2008

5. We will then match the first $100 of this deposit, giving you another $100 in your account (you may need to click refresh to see your balance update)

6. All you need to do then is fulfil the bonus wagering requirement by wagering $5,000 on any of the games below before midnight on April 30, 2008

Games that qualify for the promotional offer are 3 Card Stud Poker, Let It Ride Poker, Casino Stud Poker, Pai Gow Poker, Casino Solitaire, Casino War, Red Dog, Bejeweled, Cubis, Sudoku, Keno, Lottery, and Scratchcards.

3. Analysis of Casino Games and Percent House Margin

A casino game can be defined as follows: There is an initial cost C to play the game. Each trial results in an outcome O_i, where each outcome occurs with profit x_i and probability p_i. A profit

of zero means the money paid to the player for a particular outcome equals the initial cost. Profits above zero represent a gain for the player; negative profits represent a loss. The probabilities are all non-negative and sum to one over all the possible outcomes.
Given this information, the total expected profit $\sum E_i = \sum p_i x_i$. The *percent house margin* (%HM) is then $-\sum E_i/C$. Positive percent house margins indicate that the gambling site on average makes money and the players lose money. Table 1 summarizes this information when there are K possible outcomes.

Outcome	Profit	Probability	Expected Profit
O_1	x_1	p_1	$E_1=p_1x_1$
O_2	x_2	p_2	$E_2=p_2x_2$
O_3	x_3	p_3	$E_3=p_3x_3$
…	…	…	…
O_K	x_K	p_K	$E_K=p_Kx_K$
		1.0	$\sum E_i$

Table 1: Representation in terms of expected profit of a casino game with K possible outcomes.

If N consecutive bets are made then the total profit, T, is a random variable.

$T = X_1 + X_2 + \ldots + X_N$,

where X_i is the outcome on the i^{th} bet.

We assume that the variables X_i are independently and identically distributed. That is, we assume (quite reasonably for online casino games) that the probability distribution is the same each time we play and consecutive plays have no impact on each other.

The parameters of T are directly related to the parameters of X:

Mean $\mu_T = N\mu_X$

Standard Deviation $\sigma_T = \sqrt{N}\,\sigma_X$

Coefficient of Skewness $\gamma_T = \gamma_X/\sqrt{N}$

Coefficient of Excess Kurtosis $\kappa_T = \kappa_X / N$

In the above, μ_X, σ_X, γ_X, and κ_X are the mean, standard deviation, skewness, and excess kurtosis for variable X. In general excess kurtosis = kurtosis -3 and we will be using excess kurtosis for the remainder of the article. For example, the normal distribution has an excess kurtosis of 0, and therefore a kurtosis of 3.

The parameters are used in a normal approximation to a standardized version of random variable T. We'll use this normal approximation to compute probabilities associated with the online casino games. A discussion of moments and parameters of probability distributions is presented in section 6.

Let Z be a standardised variable, such that

$$Z = (T - \mu_T)/\sigma_T.$$

Variable Z has mean 0 and variance 1 which is the same as the standardised Normal distribution, and there is a temptation to say that the cumulative distribution function (CDF) of Z can be approximated by the Normal distribution, i.e.

$$\text{Prob}(Z \geq z) = F(z) \approx \Phi(z)$$

where $\Phi(.)$ is the cumulative Normal distribution.

The simple approximation above will give a poor fit to the tails of the distribution in most cases. This is because the variable Z may be skew and have a non-zero excess kurtosis. A better approximation is the Normal Power approximation given in Pesonen (1975), which can be expressed in the following form:

$$F(z) \approx \Phi(y)$$

where $y = z - 1/6\,\gamma_T (z^2 - 1) + [1/36\,\gamma_T^{\,2} (4z^3 - 7z) - 1/24\,\kappa_T (z^3 - 3z)]$.

Using y instead of z in the cumulative normal distribution provides improved accuracy when the distribution has skewness and excess kurtosis different from that of a standard normal distribution. Failure to recognize a skewed distribution for the outcome is likely to result in underestimating your chance of ruin, given a finite capital. This theory seems to be used in modern risk theory, such as Pentikainen and Pesonen (1984). This is despite a warning that the NP approximation is known to fail when $\gamma_T > 2$ or $\kappa_T > 6$.

In order to use these probability calculations to evaluate the games and the promotion, we need to know some details of the games as discussed below.

4. The Casino Games

The Wizard of Odds at https://wizardofodds.com/ gives the rules and %HM for various online casino games, represented in Table 2, by assuming optimal strategies where applicable. Note that the house margins are unknown for Casino Solitaire, Bejeweled, Cubis, Sudoku, Lottery, and Scratchcards. Therefore, we choose to ignore these games when determining the best game to play.

Table 2 indicates that the pairplus bet in 3 Card Stud Poker has the lowest %HM of 2.32%, but requires a minimum bet of \$5. Therefore, 3 Card Stud Poker requires at most 2,000 trials to turn over \$10,000. The go to war option in Casino War has a %HM of 2.88% with a minimum bet of only \$2. Therefore, Casino War requires at most 5,000 trials to turn over \$10,000. Although Pai Gow Poker has a %HM of 2.86%, we choose to ignore this game due to the complex strategies to achieve this margin. Based on this information, the analysis to follow will be on 3 Card Stud Poker (pairplus bet) and Casino War (go to war option).

Game	House Margin	Minimum Bet
3 Card Stud Poker	3.37% (ante bet) 2.32% (pairplus bet)	\$5 \$5
Let it Ride Poker	3.51%	\$6
Casino Stud Poker	5.22%	\$2
Pai Gow Poker	2.86%	\$5
Casino War	2.88% (go to war) 3.70% (surrender)	\$2 \$2
Red Dog	3.16%	\$5
Keno	25-29%	\$0.25

Table 2: InterCasino games that qualify for the promotion offer

Tables 3 and 4 give the outcomes, probabilities, profits, and expected profits for 3 Card Stud Poker and Casino War, respectively. The minimum bets for each game have been applied. Note that in Table 4, x = 12*(24/310*23/309) + 22/310*21/309 and y = (1-x)/2. The calculations for Table 4 were obtained as follows: 23/311 is the probability of the dealer and player obtaining the same rank of card. The probability of the dealer having a higher card rank than the player and vice versa is then (1-23/311)/2 = 144/311. The probability of the dealer and player obtaining a card of the same rank given their initial cards were of the same rank can be calculated as x. The player wins the hand (given the player and dealers initial cards were of the same rank) by drawing a card with the same or higher rank than the dealer. This occurs with probability 23/311*(x+y). The player loses the hand (given the player and dealers initial cards were of the same rank) by drawing a card with lower rank than the dealer. This occurs with probability 23/311*(y).

Outcome	Profit (\$)	Probability	Expected Profit (\$)
Straight Flush	200	12/5525	0.434
Three of a Kind	150	1/425	0.353
Straight	30	36/1105	0.977
Flush	20	274/5525	0.992
Pair	5	72/425	0.847
All other	-5	822/1105	-3.719
Total		**1**	**-0.116**

Table 3: Expected profits of Three Card Poker with the minimum \$5 bet

Outcome	Profit (\$)	Probability	Expected profit (\$)
Highest card	2	144/311	0.926
Lowest card	-2	144/311	-0.926
Go to war and win	2	23/311 (x+y)	0.079
Go to war and lose	-4	23/311 y	-0.137
Total		**1**	**-0.058**

Table 4: Expected profits of Casino War with the minimum \$2 bet

By comparing tables 3 and 4, it is evident that the variance of profits and higher-order moments such as skewness is significantly greater for 3 Card Stud Poker. This is a result of the minimum bet being greater for 3 Card Stud Poker and the relatively low probabilities in 3 Card Stud Poker in obtaining the outcomes of a straight flush and three of a kind.

5. Criteria and Results

The game the player chooses will depend on the player's attitude to risk and whether depositing more than the minimum $200 into an account is an option. A professional gambler, for example, may be more inclined to play 3 Card Stud Poker, as it has a lower %HM. An individual who gambles only to take advantage of these promotional offers, however, may be more inclined to play Casino War. Remember that there are additional costs in depositing more money through processing, withdrawals, and currency conversions.

For an individual to make a judgement about which game he or she would prefer to play (if any), we will investigate two scenarios:

Scenario 1: Looks at the end result only by allowing the player to deposit extra funds if necessary. Only one of the two games will be played to turn over the $10,000.

Scenario 2: Attempts to minimize the probability of having to deposit more money while playing. Only one of the two games will be played to turn over the $10,000.

5.1 Characteristics of the amount of profit

The equations discussed above are applied to calculate the mean, standard deviation, and coefficients of skewness and excess kurtosis of the amount of profit after one trial and the results given in Table 5. The characteristics of the amount of profit after N trials for 3 Card Stud Poker with a minimum $5 bet are given in Table 6. Similarly, the characteristics of the amount of profit after N trials for Casino War with a minimum $2 bet are given in Table 7. As expected, the standard deviation and coefficients of skewness and excess kurtosis are significantly greater for 3 Card Stud Poker. The bonus $200 has not been included in the calculations to generate tables 6 and 7. Given this $200 bonus, 3 Card Stud Poker and Casino War would have total expected profits of -$31.67 and -$87.71, respectively. Note that the standard deviation, coefficient of skewness, and coefficient of excess kurtosis remain unchanged when the $200 bonus is included. The coefficient of variation is given by the standard deviation divided by the mean, and therefore would change with the $200 bonus.

Game	Mean	Standard deviation	Coefficient of skewness	Coefficient of excess kurtosis
Three Card Poker	-$0.12	$14.55	8.63	102.08
Casino War	-$0.06	$2.10	-0.12	-1.77

Table 5: Characteristics of the amount of profit after one trial for Three Card Poker and Casino War

Turnover	Trials	Mean	Standard deviation	Coefficient of skewness	Coefficient of excess kurtosis
$1,000	200	-$23.17	$205.81	0.61	0.51
$2,000	400	-$46.33	$291.06	0.43	0.26
$3,000	600	-$69.50	$356.48	0.35	0.17
$4,000	800	-$92.67	$411.63	0.31	0.13
$5,000	1,000	-$115.84	$460.21	0.27	0.10
$6,000	1,200	-$139.00	$504.14	0.25	0.09
$7,000	1,400	-$162.17	$544.53	0.23	0.07
$8,000	1,600	-$185.34	$582.13	0.22	0.06
$9,000	1,800	-$208.51	$617.44	0.20	0.06
$10,000	2,000	-$231.67	$650.84	0.19	0.05

Table 6: Characteristics of the amount of profit after N trials for Three Card Poker with a $5 minimum bet

Turnover	Trials	Mean	Standard deviation	Coefficient of skewness	Coefficient of kurtosis
$1,000	500	-$28.77	$46.94	-0.01	0.00
$2,000	1,000	-$57.54	$66.39	0.00	0.00
$3,000	1,500	-$86.31	$81.31	0.00	0.00
$4,000	2,000	-$115.09	$93.89	0.00	0.00
$5,000	2,500	-$143.86	$104.97	0.00	0.00
$6,000	3,000	-$172.63	$114.99	0.00	0.00
$7,000	3,500	-$201.40	$124.20	0.00	0.00
$8,000	4,000	-$230.17	$132.78	0.00	0.00
$9,000	4,500	-$258.94	$140.83	0.00	0.00
$10,000	5,000	-$287.71	$148.45	0.00	0.00

Table 7: Characteristics of the amount of profit after N trials for Casino War with a $2 minimum bet

5.2 Distribution of profits

The Normal Power approximation is applied to give the distribution of profits for both Three Card Poker and Casino War, and a graphical representation is given in Figure 1. Note that the bonus $200 has been added to the horizontal axis to represent the total profit that a player could achieve. The mode value in Three Card Poker is approximately -$100, which is less than the mean value of -$31.67. Using the normal distribution directly, the mode value would be -$31.67. This shows the importance of using the Normal Power approximation in determining the probabilities for the tail, to establishing accurate confidence intervals. What is compelling about these two graphs is that a player could lose as much as $1,500 by playing Three Card Poker.

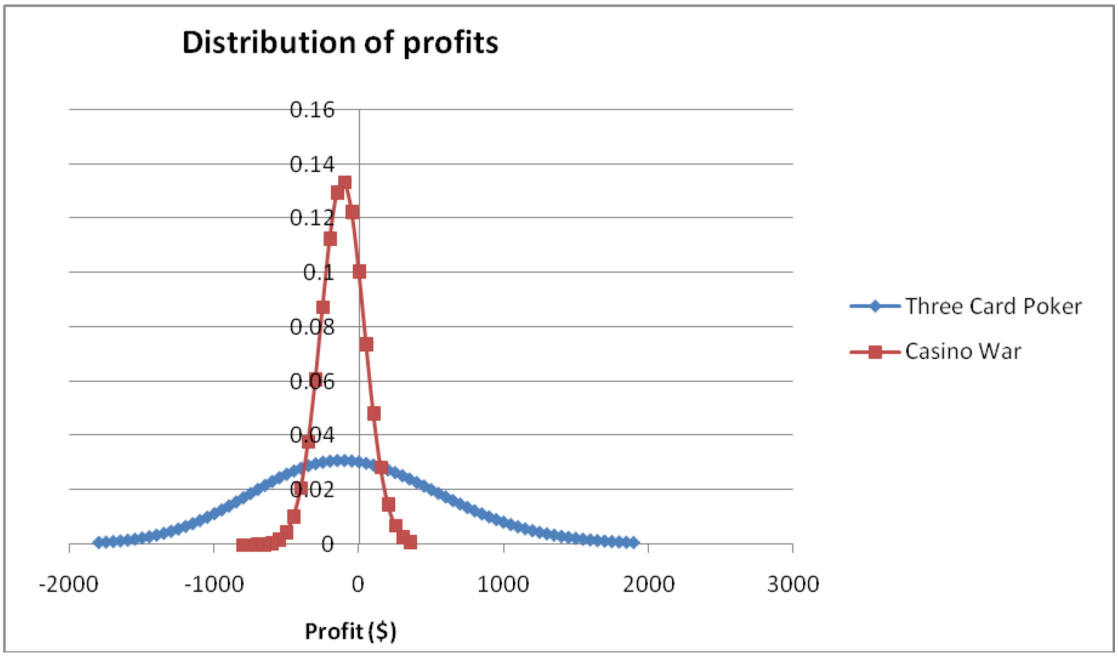

Figure 1. The distribution of profits for 3 Card Stud Poker and Casino War by turning over $10,000

5.3 Scenario 1

Figure 1 could be compared by the player, in deciding which game they would prefer to play under Scenario 1. To help interpret Figure 1, Confidence Intervals (CI) have been constructed and represented in Table 9. The results are reasonably accurate, as simulation methods were developed to verify the results. It has already been established that Three Card Poker has a lower %HM by 0.56% and a player is expected to lose $56 less by playing Three Card Poker, rather than Casino War. However, the results from Table 9 indicate that a player could lose substantially more money by playing Three Card Poker than Casino War. On the other hand, a player could gain substantially more money by playing Three Card Poker.

Game	90% CI	95% CI	99% CI
Three Card Poker	(-$1,070, $1,070)	(-$1,250, $1,300)	(-$1,600, $1,750)
Casino War	(-$330, $160)	(-$380, $200)	(-$420, $300)

Table 9: Confidence Intervals for Three Card Poker and Casino War

5.4 Scenario 2

This scenario will investigate the intermediate fluctuations during play in an attempt to minimize the chance of depositing more than the minimum amount of $200. Table 10 gives the probabilities for every $1,000 of turnover (up to a maximum of $5,000) of a player being behind more than $200, and therefore having to deposit more money. The results indicate that there is about a 49% chance of depositing more money in Three Card Poker whilst fulfilling the first $5,000 wagering requirements. There is about a 37% chance of depositing more money in Casino War whilst fulfilling the first $5,000 wagering requirements. Calculating values for the chances of depositing more money to fulfil the $10,000 wagering requirements is difficult to obtain. However, it is clear that the chance of depositing more money is smaller by playing Casino War, and therefore under Scenario 2, Casino War is the preferred game to play.

Turnover	Three Card Poker	Casino War
$1,000	0.290	0.001
$2,000	0.385	0.033
$3,000	0.433	0.125
$4,000	0.464	0.245
$5,000	0.488	0.365

Table 10: Probability of being behind more than $200 for Three Card Poker and Casino War

6. An Aside on Moments of a Probability Distribution

Central to the computations in the main article is the normal approximation to the distribution of the profits from playing an online casino game many times. The mean, variance, skewness and excess kurtosis of the distribution of profits all played a part in the approximation. Here we review definitions of moments as would typically be presented in an introductory mathematical probability course at the college level.

As in the article, assume there is an initial cost C to play the game. Each trial results in an outcome O_i, where each outcome occurs with profit x_i and probability p_i. The probabilities are all non-negative and sum to one over all the possible outcomes. Given this information, the total expected profit $\sum E_i = \sum p_i x_i$. The *percent house margin* (%HM) is then $-\sum E_i/C$. Table 1 in the article summarizes this information when there are K possible outcomes.

6.1 Moments and Cumulants from a Single Bet

The outcome or profit from a single bet, X, is a random variable. From probability theory, the *moment generating function* (MGF) of X is

$$M_X(t) = E(\exp(Xt)) = 1 + m_{1X}t + m_{2X}t^2/2! + m_{3X}t^3/3! + m_{4X}t^4/4! + \ldots,$$

where m_{rX} represent the r^{th} moment of X. The moments of X are readily calculated using

$$m_{1X} = \Sigma_i\, p_i\, x_i$$
$$m_{2X} = \Sigma_i\, p_i\, x_i^2$$
$$m_{3X} = \Sigma_i\, p_i\, x_i^3$$
$$m_{4X} = \Sigma_i\, p_i\, x_i^4$$

and so on. The calculation of these moments is illustrated in Table 11.

The cumulant generating function (CGF) of X is the natural logarithm of the MGF:

$$K_X(t) = \log_e(M_X(t)) = k_{1X}t + k_{2X}t^2/2! + k_{3X}t^3/3! + k_{4X}t^4/4! + \ldots,$$

where k_{rX} represent the r^{th} cumulant of X. The relationship between the first four cumulants and moments is given by

$k_{1X} = m_{1X}$
$k_{2X} = m_{2X} – m_{1X}^2$
$k_{3X} = m_{3X} – 3m_{2X}m_{1X} + 2m_{1X}^3$ and
$k_{4X} = m_{4X} – 4m_{3X}m_{1X} – 3m_{2X}^2 + 12m_{2X}m_{1X}^2 – 6m_{1X}^4$.

These cumulants can be used to calculate the following familiar distributional characteristics (parameters) for X:

Mean	$\mu_X = k_{1X}$
Standard Deviation	σ_X = square root of k_{2X}
Coefficient of Skewness	$\gamma_X = k_{3X} / (k_{2X})^{3/2}$
Coefficient of Excess Kurtosis	$\kappa_X = k_{4X} / (k_{2X})^2$.

Outcome	Profit	Probability	1st Moment	2nd Moment	3rd Moment	4th Moment
O_1	x_1	p_1	p_1x_1	$p_1x_1^2$	$p_1x_1^3$	$p_1x_1^4$
O_2	x_2	p_2	p_2x_2	$p_2x_2^2$	$p_2x_2^3$	$p_2x_2^4$
O_3	x_3	p_3	p_3x_3	$p_3x_3^2$	$p_3x_3^3$	$p_3x_3^4$
...	...	...	...	...	...	...
O_K	x_K	p_K	p_Kx_K	$p_Kx_K^2$	$p_Kx_K^3$	$p_Kx_K^4$
		1.0	$\mathbf{m_{1X}= \Sigma_i\, p_ix_i}$	$\mathbf{m_{2X}= \Sigma_i\, p_ix_i^2}$	$\mathbf{m_{3X}= \Sigma_i\, p_ix_i^3}$	$\mathbf{m_{4X}= \Sigma_i\, p_ix_i^4}$

Table 11: Representation of the first four moments of the profit of a casino game after one bet.

6.2 Moments and Cumulants After N Bets

If N consecutive bets are made then the total profit, T, is a random variable.

$$T = X_1 + X_2 + ... +X_N,$$

where X_i is the outcome on the i^{th} bet. We assume that the variables X_i are independently and identically distributed.

When the number of outcomes in a single bet is two (Win or Lose), the binomial formula can be used to calculate the distribution of profits after N bets. Alternatively, the normal approximation to the binomial distribution can be applied when N is large. (Packel, 1981).

When there are more than two outcomes in a single bet, we can use MGFs, CGFs, and a different normal approximation formula. Assuming that the outcome from each bet is independent of the others, probability theory tells us that the MGF of random variable T is the product of MGFs of the X_i's:

$$M_T(t) = E(\exp(X_1 + X_2 + ... +X_N)t)$$
$$= E(\exp(X_1 t)) E(\exp(X_2 t)) .... E(\exp(X_N t))$$
$$= M_{X1}(t) M_{X2}(t) .... M_{XN}(t)$$

If the bets are all on the same game and the same size, then the distribution of the profit from each bet is identical, and we obtain an important simplification:

$$M_T(t) = [M_X(t)]^N.$$

Taking logarithms we obtain a relationship between the CGFs:

$K_T(t) = NK_X(t)$.

This relationship can be expressed in terms of the individual cumulants

$k_{rT} = N\, k_{rX}$ for all $r \geq 1$.

Thus the cumulants of the total profit after N bets of the same size on a single game can be computed directly from the cumulants of the profit for a single bet.

Likewise the parameters of T are directly related to the parameters of X:

Mean $\mu_T = N\mu_X$

Standard Deviation $\sigma_T = \sqrt{N}\, \sigma_X$

Coefficient of Skewness $\gamma_T = \gamma_X/\sqrt{N}$

Coefficient of Excess Kurtosis $\kappa_T = \kappa_X / N$

7. Conclusions

Scenarios 1 and 2 assume the player only plays the one game to turn over the required $10,000. If a player had a good run in the first $5,000 turnover in Casino War, he or she may be inclined to switch to 3 Card Stud Poker. Likewise, if a player had a bad run in the first $5,000 turnover in 3 Card Stud Poker, he or she may be inclined to switch to Casino War. A player's choice to switch between 3 Card Stud Poker and Casino War is dependent on his or her current bankroll and how many hands are left to play to wager the required $10,000.

The minimum bet for both games has always been applied. A player may choose to speed up the process by betting higher than the minimum. This may again depend on the player's current bankroll and how many hands are left to play. In general, increasing the size of the bankroll in any casino game increases the variance and skewness. This would amount to increasing the probability of losing the initial $200 before wagering the initial $5,000 (i.e., a greater chance of having to deposit more money into the casino account).

Is the offer worthwhile? The cost of a return flight to New York from Australia is around $2,000, excluding taxes. The expected cost for the player in 3 Card Stud Poker is $88 and $32 in Casino War. Therefore, on average, the offer is worthwhile in the sense of having a positive expected value once one accounts for the value of the flight from Australia.

Which is the best game to play? We identified two games in which the playing strategies were straightforward and had relatively low percent house margins. The best game to play depends on the player's objective toward the offer. A professional gambler, for example, may be more inclined to play 3 Card Stud Poker because it has a lower %HM, whereas an individual who gambles only to take advantage of these promotional offers may be more inclined to play Casino War, as it minimizes the probability of having to deposit more than $200 into an account.

There is evidence to show that Casino War is the preferred game if a player is trying to minimize the probability of depositing more than the minimum $200. This is interesting, as a player is willing to spend an extra $56 to reduce risk.

8. References

Packel E (1981). The Mathematics of Games and Gambling. The Mathematical Association of America: Washington, DC

Pentikainen T and Pesonen E (1984). Risk Theory: The Stochastic Basis of Insurance. Chapman and Hall: New York, New York.

Pesonen E (1975). "NP-Technique as a Tool in Decision Making." ASTIN Bulletin. 8(3):359–363. https://www.casact.org/library/astin/vol8no3/359.pdf.

Chapter 8: Predicting a Tennis Match in Progress for Sports Multimedia

1. Introduction

Profiting from sports betting is an obvious application of predicting outcomes in sport. Whilst sports betting was originally restricted to betting prior to the start of the match, it is now possible and common to be betting throughout a match in progress. However, the appeal of predictions throughout a match may not just be restricted to punters. In-play sports predictions could be used in sports multimedia, and hence could be appealing to the spectator, coach or technology buff; without involving actual betting.

According to Wikipedia: "Multimedia is media and content that includes a combination of text, audio, still images, animation, video, and interactivity content forms. Multimedia is usually recorded and played, displayed or accessed by information content processing devices, such as computerized and electronic devices". Infoplum Pty. Ltd. is a sports multimedia organization that specializes in delivering online information during live matches. As well as displaying the scoreboard, live predictions are in operation for tennis, cricket, Australian Rules football and soccer. This information is available on a standard PC and iPhone (as an iPhone app), with current developments for delivering the information on an iPad and TV (in the form of a widget).

There are many ways the predictions through multimedia could be used. Spectators could engage with live predictions through multimedia for entertainment when watching a live match. If a spectator was to place a bet, then live predictions through multimedia could be used as a decision support tool as to when and how much to bet on a particular event, and hence the combination of betting and multimedia becomes a powerful form of entertainment.

Another interesting application of sports predictions in multimedia is in teaching mathematical concepts. Students often relate to sporting events and hence may be stimulated in learning mathematics through an activity of personal interest. In tennis for example, the chance of winning from deuce is calculated by the sum of an infinite series.

Another application is in using the predictions as a coaching tool. For example, it is common for players and coaches to watch a replayed match to discuss strategies for upcoming matches. The graphical and visual aspects of the predictions could enhance the TV replay. A further application is in using the predictions for TV commentary. Klaassen and Magnus (2003) demonstrate how plotting points on a graph on the probability of winning a tennis match in progress can be useful for TV commentary by supporting his/her discussion on the likely winner of the match. Commentators could also use the graph to evaluate the match after completion to identify turning points and key shifts in momentum.

In this paper, a tennis prediction model during a match in progress is constructed. Data analysis is carried out in section 2 to determine the parameters for the model. A Markov chain model is then developed in section 3 and an updating rule is given in section 4 to update prior estimates with what has occurred during the match. The men's 2010 US Open tennis final is given in section 5 as an illustration of live sports predictions through multimedia.

2. Data Analysis

By assigning two parameters, the constant probabilities of player A and player B winning a point on serve; modelling the probability of winning the match can be determined using a Markov Chain model as represented in section 3. This section will therefore derive the probabilities of winning on serve when two players meet on a particular surface. This is achieved by collecting, combining and updating player serving and receiving statistics.

2.1 Collecting player statistics

OnCourt (http://oncourt.info/) is a software package for all tennis fans, containing match results for men's and woman's tennis, along with statistical information about players, tournaments or histories of the head-to-head matches between two players. Match statistics can be obtained for the majority of ATP and WTA matches going back to 2003. The Association of Tennis Professionals (ATP) is the governing body of men's professional tennis for the allocation of player rating points in matches to determine overall rankings and seedings for tournaments. The Women's Tennis Association (WTA) is used similarly in women's professional tennis. Table 1 gives the match statistics broadcast from the US Open 2010 men's final where Rafael Nadal defeated Novak Djokovic in four sets. Notice that the Serving Points Won is not given directly in the table. This statistic can be derived from the Receiving Points Won such that Serving Points Won for Nadal and Djokovic are 1-36/112 = 67.9% and 1-60/143 = 58.0% respectively. Alternatively, the Serving Points Won can be obtained from a combination of the 1st Serve %, Winning % on 1st Serve and Winning % on 2nd Serve such that Serving Points Won for Nadal and Djokovic are 75/112*55/75+(1-75/112)*21/37= 67.9% and 95/143*61/95+(1-95/143)*22/48 = 58.0% respectively. Note that the Winning % on 1st Serve is conditional on the 1st Serve going in whereas the Winning % on the 2nd Serve is unconditional on the 2nd Serve going in. These calculations could be used as a teaching exercise in interpreting and analyzing data, and in conditional probabilities. Many more calculations can be obtained from broadcast match statistics as outlined in Bedford et al. (2010).

	Rafael Nadal	Novak Djokovic
1st Serve %	75 of 112 = 66%	95 of 143 = 66%
Aces	8	5
Double Faults	2	4
Unforced Errors	31	47
Winning % on 1st Serve	55 of 75 = 73%	61 of 95 = 64%
Winning % on 2nd Serve	21 of 37 = 56%	22 of 48 = 45%
Winners (Including Service)	49	45
Break Point Conversions	6 of 26 = 23%	3 of 4 = 75%
Receiving Points Won	60 of 143 = 41%	36 of 112 = 32%
Net Approaches	16 of 20 = 80%	28 of 45 = 62%
Total Points Won	136	119
Fastest Serve	212 KPH	201 KPH
Average 1st Serve Speed	186 KPH	188 KPH
Average 2nd Serve Speed	141 KPH	151 KPH

Table 1: Match statistics for the men's 2010 US Open final between Rafael Nadal and Novak Djokovic

2.2 Combining player statistics

Combining player statistics is a common challenge in sport. While we would expect a good server to win a higher proportion of serves than average, this proportion would be reduced somewhat if his opponent is a good receiver. Using the method developed by Barnett and Clarke (2005) we can calculate the percentage of points won on serve when player i meets player j on surface s (f_{ijs}) as:

$f_{ijs} = f_{is} - g_{js} + g_{avs}$ (1)

where:
f_{is} = percentage of points won on serve for player i on surface s,
g_{is} = percentage of points won on return of serve for player i on surface s
g_{avs} represents the average (across all ATP/WTA players) percentage of points won on return of serve on surface s.

The surfaces are defined as: s=1 for grass, s=2 for carpet, s=3 for hard and s=4 for clay.
The average percentage of points won on serve across all players on each of six different surfaces (grass, hard, indoor hard, clay, carpet and acrylic) was calculated from OnCourt and represented in table 2. Note that the serving averages for carpet and indoor hard are

approximately the same and are therefore combined as the one surface. Similarly, hard and acrylic are combined as the one surface.

Example: Suppose player i with f_{i1}= 0.7 and g_{i1}= 0.4 meets player j with f_{j1}= 0.68 and g_{j1}= 0.35 on a grass court surface. Then the estimated percentage of points won on serve for player i and player j are given by f_{ij1} = 0.7-0.35+(1-0.653) = 69.7% and f_{ji1} = 0.68-0.4+(1-0.653) = 62.7% respectively.

Surface	Men	Women
Grass	0.653	0.580
Carpet – I.hard	0.642	0.570
Hard – Acrylic	0.625	0.552
Clay	0.600	0.536

Table 2: The average probabilities of points won on serve for men's and women's tennis

2.3 Updating player statistics

The general form for updating the rating of a player as given by Clarke (1994) is

New Rating = Old Rating + α [actual margin – predicted margin]

for some α.

Using serving and receiving player statistics as ratings we get

$f_{is}^{n} = f_{is}^{o} + \alpha_s [f_{is}^{a} - f_{ijs}]$ (2)

$g_{is}^{n} = g_{is}^{o} + \alpha_s [g_{is}^{a} - g_{ijs}]$ (3)

where:

f_{is}^{n}, f_{is}^{o} and f_{is}^{a} represent the new, old and actual percentage of points won on serve for player i on surface s

g_{is}^{n}, g_{is}^{o} and g_{is}^{a} represent the new, old and actual percentage of points won on return of serve for player i on surface s

α_s is the weighting parameter for surface s

Experimental results reveal that α_s = 0.049 is a suitable weighting parameter for all surfaces. Further, every player is initialized with surface averages as given in table 2.

Equations (2) and (3) treat each surface independently. A more advanced approach is to update the serving and receiving statistics for each surface when playing on a particular surface. For example, if a match is played on grass, then how are the other surfaces of clay, carpet and hard court updated based on the player's performances on the grass?

This more complicated approach is given as:

$f_{ist}^{n} = f_{is}^{o} + \alpha_{st} [f_{is}^{a} - f_{ijs}]$ (4)

$g_{ist}^{n} = g_{is}^{o} + \alpha_{st} [g_{is}^{a} - g_{ijs}]$ (5)

where:
f_{ist}^n represents the new expected percentage of points won on serve for player i on surface t when the actual match is played on surface s
g_{ist}^n represents the new expected percentage of points won on return of serve for player i on surface t when the actual match is played on surface s
α_{st} is the weighting parameter for surface t when the actual match is played on surface s

Table 3 gives the weighting parameters based on experimental techniques.

Surface	**t=Grass**	**t=Carpet/I.hard**	**t=Hard/Acrylic**	**t=Clay**
s=Grass	0.049	0.02	0.015	0.01
s=Carpet/I.hard	0.02	0.049	0.02	0.015
s=Hard/Acrylic	0.015	0.02	0.049	0.02
s=Clay	0.01	0.015	0.02	0.049

Table 3: Surface weighting parameters as applicable to both men and women

3. Markov Chain model

The basic principles involved in modelling a tennis match are well known, and a Markov chain model with a constant probability of winning a point was set up by Schutz (1970). While such a model is acceptable within a game, a model which allows a player a different probability of winning depending on whether they are serving or receiving is essential for tennis. Statistics of interest are usually the chance of each player winning, and the expected length of the match. Croucher (1986) looks at the conditional probabilities for either player winning a single game from any position. Pollard (1983) uses a more analytic approach to calculate the probability for either player winning a game or set along with the expected number of points or games to be played with their corresponding variance.

Most of the previous work uses analytical methods, and treats each scoring unit independently. This results in limited tables of statistics. Thus the chance of winning a game and the expected number of points remaining in the game is calculated at the various scores within a game. The chance of winning a set and the expected number of games remaining in the set is calculated only after a completed game and would not show for example the probability of a player's chance of winning from three games to two, 15-30.

This paper discusses the use of spreadsheets to repeat these applications using a set of interrelated spreadsheets. This allows any probabilities to be entered and the resultant statistics automatically calculated or tabulated. In addition, more complicated workbooks can be set up which allow the calculation of the chance of winning a match at any stage of the match given by the point, game and set score. These allow the dynamic updating of player's chances as a match progresses.

Alternatively, these algorithms could be converted into a programming language for automatic integration into multimedia as live scores are received.

3.1 Game

We explain the method by first looking at a single game where we have two players, A and B, and player A has a constant probability p_A of winning a point on serve. We set up a Markov chain model of a game where the state of the game is the current game score in points (thus 40-30 is 3-2). With probability p_A the state changes from a, b to a + 1, b and with probability $q_A=1-p_A$ it changes from a, b to a, b + 1. Thus if $P_A^{pg}(a,b)$ is the probability that player A wins the game when the score is (a,b), we have:

$P_A^{pg}(a,b)=p_AP_A(a+1,b)+q_AP_A(a,b+1)$
The boundary values are:
$P_A^{pg}(a,b) = 1$ if $a = 4, b \leq 2$, $P_A^{pg}(a,b) = 0$ if $b = 4, a \leq 2$.

The boundary values and formula can be entered on a simple spreadsheet. The problem of deuce can be handled in two ways. Since deuce is logically equivalent to 30-30, a formula for this can be entered in the deuce cell. This creates a circular cell reference, but the iterative function of Excel can be turned on, and Excel will iterate to a solution. In preference, an explicit formula is obtained by recognizing that the chance of winning from deuce is in the form of a geometric series

$P_A^{pg}(3,3) = p_A^2 + p_A^2 2p_Aq_A + p_A^2 (2p_Aq_A)^2 + p_A^2 (2p_Aq_A)^3 + \ldots\ldots\ldots$
where the first term is p_A^2 and the common ratio is $2p_Aq_A$

The sum is given by $p_A^2/(1-2p_Aq_A)$ provided that $-1<2p_Aq_A<1$. We know that $0<2p_Aq_A<1$, since $p_A>0$, $q_A>0$ and $1-2p_Aq_A=p_A^2+q_A^2>0$.

Therefore, the probability of winning from deuce is $p_A^2/(1-2p_Aq_A)$. Since $p_A+q_A=1$, this can be expressed as:

$$P_A^{pg}(3,3) = p_A^2 /(p_A^2+q_A^2)$$

Excel spreadsheet code to obtain the conditional probabilities of player A winning a game on serve is as follows:
Enter **p_A** in cell D1
Enter **q_A** in cell D2
Enter **0.60** in cell E1
Enter **=1-E1** in cell E2
Enter **1** in cells C11, D11 and E11
Enter **0** in cells G7, G8 and G9
Enter **= E1^2/(E1^2+E2^2)** in cell F10
Enter **=$E$1*C8+$E$2*D7** in cell C7
Copy and Paste cell **C7** in cells D7, E7, F7, C8, D8, E8, F8, C9, D9, E9, F9, C10, D10 and E10

Notice the absolute and relative referencing used in the formula =$E$1*C8+$E$2*D7. By setting up an equation in this recursive format, the remaining conditional probabilities can easily and quickly be obtained by copying and pasting.

Table 3 represents the conditional probabilities of player A winning the game from various score lines for p_A = 0.60. It indicates that a player with a 60% chance of winning a point has a 74% chance of winning the game. Note that since advantage server is logically equivalent to 40-30, and advantage receiver is logically equivalent to 30-40, the required statistics can be found from these cells. Also, worth noting is that the chances of winning from deuce and 30-30 are the same.

				B score		
		0	15	30	40	game
	0	0.74	0.58	0.37	0.15	0
	15	0.84	0.71	0.52	0.25	0
A score	30	0.93	0.85	0.69	0.42	0
	40	0.98	0.95	0.88	0.69	
	game	1	1	1		

Table 3: The conditional probabilities of A winning the game from various score lines

Similar equations can be developed for when player B is serving such that p_A and p_B represent constant probabilities of player A and player B winning a point on their respective serves. Also $P_A^{pg}(a,b)$ and $P_B^{pg}(a,b)$ represent the conditional probabilities of player A winning a game from point score (a,b) for player A and B serving in the game respectively.

A tennis match consists of four levels - (points, games, sets, match). Games can be standard games (as above) or tiebreak games, sets can be advantage or tiebreak, and matches can be the best-of-5 sets or the best-of-3 sets. To win a set a player needs six games with at least a two-game lead. If the score reaches 6 games-all, then a tiebreak game is played in a tiebreak set to determine the winner of the set, otherwise standard games continue indefinitely until a player is two games ahead and wins the set. This latter scoring structure is known as an advantage set and is used as the deciding set in the Australian Open, French Open and Wimbledon. In some circumstances we may be referring to points in a standard or tiebreak game and other circumstances points in a tiebreak or advantage set. It becomes necessary to represent

points in a game as pg,
points in a tiebreak game as pg^T,
points in an advantage set as ps,
points in a tiebreak set as ps^T
points in a best-of-5 set match (advantage fifth set) as pm,
points in a best-of-5 set match (all tiebreak sets) as pm^T
games in an advantage set as gs,
games in a tiebreak set as gs^T,

sets in a best-of-5 set match (advantage fifth set) as sm, and
sets in a best-of-5 set match (all tiebreak sets) as sm^T.

3.2 Tiebreak game

Since the chance of a player winning a tiebreak game depends on who is serving, two interconnected sheets are required, one for when player A is serving and one for when player B is serving. The equations that follow for modelling a tiebreak game, set and match are those for player A serving in the game. Similar formulas can be produced for player B serving in the game.

Let $P_A^{pgT}(a,b)$ and $P_B^{pgT}(a,b)$ represent the conditional probabilities of player A winning a tiebreak game from point score (a,b) for player A and player B serving in the game respectively.

Recurrence formulas:
$P_A^{pgT}(a,b) = p_A P_B^{pgT}(a+1,b)+q_A P_B^{pgT}(a,b+1)$, if (a+b) is even
$p_A^{pgT}(a,b) = p_A P_A^{pgT}(a+1,b)+q_A P_A^{pgT}(a,b+1)$, if (a+b) is odd

Boundary values:
$P_A^{pgT}(a,b)=1$ if $a=7$, $0\leq b\leq 5$
$P_A^{pgT}(a,b)=0$ if $b=7$, $0\leq a\leq 5$
$P_A^{pgT}(6,6)=p_A q_B/(p_A q_B+q_A p_B)$

where: $q_B=1-p_B$

Table 4 represents the conditional probabilities of player A winning the tiebreak game from various score lines for $p_A = 0.62$ and $p_B = 0.60$, and player A serving. Table 5 is represented similarly with player B serving.

Note how the calculations are obtained by the interconnection of both sheets. For example
$P_A^{pgT}(0,0) = p_A P_B^{pgT}(1,0)+q_A P_B^{pgT}(0,1)$
$= 0.62 * 0.62 + 0.38 * 0.39$
$= 0.53$

					B score				
		0	1	2	3	4	5	6	7
	0	0.53	0.44	0.29	0.20	0.10	0.04	0.01	0
	1	0.67	0.53	0.43	0.27	0.17	0.07	0.02	0
	2	0.76	0.68	0.53	0.42	0.24	0.13	0.03	0
A score	3	0.87	0.77	0.69	0.53	0.40	0.20	0.08	0
	4	0.93	0.89	0.80	0.72	0.52	0.37	0.13	0
	5	0.98	0.95	0.92	0.83	0.75	0.52	0.32	0
	6	0.99	0.99	0.98	0.96	0.89	0.82	0.52	
	7	1	1	1	1	1	1		

Table 4: The conditional probabilities of player A winning the tiebreak game from various score lines for $p_A = 0.62$ and $p_B = 0.60$, and player A serving

					B score				
		0	1	2	3	4	5	6	7
	0	0.53	0.39	0.29	0.17	0.10	0.03	0.01	0
	1	0.62	0.53	0.37	0.27	0.14	0.07	0.01	0
	2	0.76	0.63	0.53	0.35	0.24	0.10	0.03	0
A score	3	0.83	0.77	0.63	0.53	0.33	0.20	0.05	0
	4	0.93	0.86	0.80	0.65	0.52	0.29	0.13	0
	5	0.97	0.95	0.89	0.83	0.67	0.52	0.21	0
	6	0.99	0.99	0.98	0.93	0.89	0.71	0.52	
	7	1	1	1	1	1	1		

Table 5: The conditional probabilities of player A winning the tiebreak game from various score lines for $p_A = 0.62$ and $p_B = 0.60$, and player B serving

3.3 Tiebreak set

Formulas are now given for a tiebreak set. Similar formulas can be obtained for an advantage set.

Let $P_A^{gsT}(c,d)$ and $P_B^{gsT}(c,d)$ represent the conditional probabilities of player A winning a tiebreak set from game score (c,d) for player A and player B serving in the game respectively.

Recurrence formula:
$P_A^{gsT}(c,d)=P_A^{pg}(0,0)P_B^{gsT}(c+1,d)+[1- P_A^{pg}(0,0)]P_B^{gsT}(c,d+1)$

Boundary Values:
$P_A^{gsT}(c,d)=1$ if $c=6, 0\leq d\leq 4$ or $c=7, d=5$
$P_A^{gsT}(c,d)=0$ if $d=6, 0\leq c\leq 4$ or $c=5, d=7$
$P_A^{gsT}(6,6) = P_A^{pgT}(0,0)$

Notice how the cell $P_A^{pg}(0,0)$, which represents the probability of winning a game, is used in the recurrence formula for a tiebreak set. Using the formulas given for a game and a tiebreak game conditional on the point score and a tiebreak set conditional on the game score, calculations are now obtained for a tiebreak set conditional on both the point and game score as follows.

Let $P_A^{psT}(a,b:c,d)$ represent the probability of player A winning a tiebreak set from (c,d) in games, (a,b) in points and player A serving in the set. This can be calculated by:

$P_A^{psT}(a,b:c,d)=P_A^{pg}(a,b)P_B^{gsT}(c+1,d)+[1-P_A^{pg}(a,b)]P_B^{gsT}(c,d+1)$, if $(c,d) \neq (6,6)$
$P_A^{psT}(a,b:c,d)=P_A^{pgT}(a,b)$, if $(c,d)=(6,6)$

3.4 Match

Formulas are now given for a best-of-5 set match, where all sets are tiebreak sets. Similar formulation can be obtained for a best-5 set match, where the deciding fifth set is advantage. Formulation can also be obtained for a best-of-3 set match.

Let $P^{smT}(e,f)$ represent the conditional probabilities of player A winning a best-of-5 set tiebreak match from set score (e,f).

Recurrence Formula:
$P^{smT}(e,f)= P_A^{gsT}(0,0)P^{smT}(e+1,f)+[1-P_A^{gsT}(0,0)]P^{smT}(e,f+1)$

Boundary Values:
$P^{smT}(e,f)=1$ if $e=3,f\leq 2$
$P^{smT}(e,f)=0$ if $f=3, e\leq 2$

Notice how the cell $P_A^{gsT}(0,0)$, which represents the probability of winning a tiebreak set, is used in the recurrence formula for a best-of-5 set match. Using the formulas given for a tiebreak set conditional on the point and game score and a best-of-5 set tiebreak match conditional on the set score, calculations are obtained for a best-of-5 set tiebreak match conditional on the point, game and set score as follows.

Let $P_A^{pmT}(a,b:c,d:e,f)$ represent the probability of player A winning a tiebreak match from (e,f) in sets, (c,d) in games, (a,b) in points and player A serving in the match. This can be calculated by:

$P_A^{pmT}(a,b:c,d:e,f)=P_A^{psT}(a,b:c,d)P^{smT}(e+1,f)+[1-P_A^{psT}(a,b:c,d)]P^{smT}(e,f+1)$

Excel spreadsheet code was given directly in section 3.1 to obtain the conditional probabilities of player A winning a game on serve. Using the formulas given in section 3, spreadsheets can be developed for a game with player B serving, tiebreak game, tiebreak set conditional on the game score and a best-of-5 set tiebreak match conditional on the set score. By assigning a value for p_B to a cell (cell E3 for example), the probability of winning a match from the outset can be obtained for any probability value of p_A and p_B by changing the probability values given in cells E1 and E3. By adding additional formulas to the spreadsheet for a tiebreak set conditional on the point and game score (section 3.3) and for a best-of-5 set tiebreak match conditional on the point, game and set score (section 3.4), the chances of player's winning the set and match can be obtained conditional on who is currently serving, point score, game score and set score. An interactive tennis Java calculator to reflect this methodology is available at http://strategicgames.com.au/TennisCalc.jar.

4. Updating rule for serving statistic estimation

Although prior estimates of points won on serve may be reliable for the first few games or even the first set, it would be useful to update the prior estimates with what has actually occurred throughout the match. We will use an updating system of the form where for player i the proportion of initial serving statistics (X_i) is combined with actual serving statistics (Y_i) to give updated serving statistics (Z_i) at any point within the match.

$$Z_i = e^{-n/c} X_i + (1-e^{-n/c})Y_i$$

where n represents the total number of points played and c is a constant.

Experimental results reveal that c = 200 is a suitable constant for best-of-5 and best-of-3 set matches. Note that the updating process occurs after each point.

5. Sports Multimedia

The Markov Chain model outlined in section 3 along with the data analysis outlined in section 2 to determine the parameters for the model, is used in multimedia to obtain tennis predictions during a live match, and made available at http://sportsflash.com.au/. Two feature prediction products have been devised – Crystal Ball and Looking Glass. The Crystal Ball provides the chances of winning the match in progress in the form of a pie chart. The Looking Glass (similar format to a stock market chart) plots the chances of winning the match on a game-by-game basis (as in tennis) or every one-minute time interval (as in soccer). The graphical, visual and interactive properties of the Crystal Ball and Looking Glass could encourage spectators to engage with the predictions throughout a match in progress.

Figure 1 represents the predictions through the Looking Glass for the US Open men's final between Nadal and Djokovic. From the outset Djokovic had a 68.0% chance of winning the match. After Nadal won the 1st Set 6-4, the chances of Djokovic to win the match decreased to 36.7%. Djokovic won the 2nd Set 7-5 and the chances to win the match increased to 61.0%. This value is represented in figure 1 by using the interactive mouse-over feature, where solid dots are given for breaks of serve.

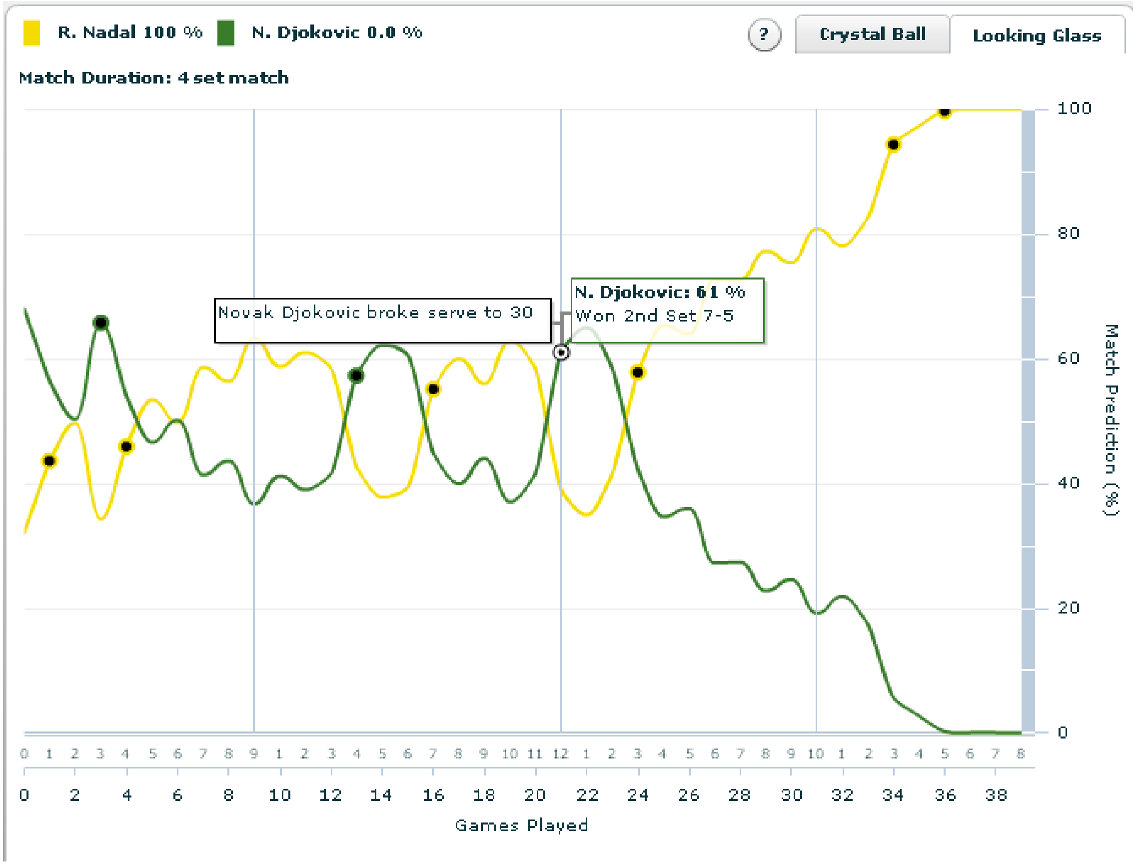

Figure 1: Live predictions for the US Open men's final between Rafael Nadal and Novak Djokovic

6. Conclusions

This paper demonstrates how spreadsheets can generate the probability of winning a tennis match conditional on the state of the match. These sheets have been used in multimedia to predict outcomes for a match in progress. There are many applications as to how these predictions could be used. An obvious application is in sports betting and the live predictions could provide a decision support tool to the punter. Another application is in using the predictions as a coaching tool for when players and coaches discuss strategies for upcoming matches on a replayed match. A further application is in using the live predictions for TV commentary by supporting the commentators' discussion on the likely winner of the match. The development of the predictions could also form an interesting and useful teaching example, and allow students to investigate the properties of tennis scoring systems.

7. References

Barnett T and Clarke SR (2005). Combining player statistics to predict outcomes of tennis matches. *IMA Journal of Management Mathematics*. 16 (2), 113-120.

Bedford A, Barnett T, Pollard GH and Pollard GN (2010). How the interpretation of match statistics affects player performance. Journal of Medicine and Science in Tennis. 15(2), 23-27

Clarke S (1994). An adjustive rating system for tennis and squash players. In Mathematics and Computers in Sport. N. de Mestre, Editor, Bond University, 43-50.

Croucher J (1986). The conditional probability of winning games of tennis. Research Quarterly for Exercise and Sport, 57(1), 23–26.

Klaassen F and Magnus J (2003). On the probability of winning a tennis match. Medicine and Science in Tennis 8(3) 10.

Pollard GH (1983). An analysis of classical and tie-breaker tennis. Australian Journal of Statistics, 25, 496–
505.

Schutz RW (1970). A mathematical model for evaluating scoring systems with specific reference to tennis. Research Quarterly for Exercise and Sport, 41 (1970), 552–561.