Statistical Analysis of Texas Hold'Em

March 4, 2009

Mr. Paco Hope
Technical Manager
Cigital, Inc.

Mr. Sean McCulloch, PhD
Associate Professor
Department of Mathematics and Computer Science
Ohio Wesleyan University

Cigital, Inc.

21351 Ridgetop Circle
Suite 400
Dulles, VA 20166
Phone: + 1 (703) 404-9293

www.cigital.com

This work was performed under contract to:

Rational Entertainment Enterprises Limited (REEL)

Douglas Bay Complex, Onchan

Isle of Man

Table of Contents

List of Figures

List of Tables

Note to the Reader

This document is written to a technical audience. It is assumed that the reader is acquainted with common poker terminology (flop, river, hole cards, board, etc.) It is further assumed that the reader understands the basic mechanics of playing Texas Hold 'Em. This document also uses standard poker notation such as K♦4♣Q♦2♠J♥ or 5c5hKcTd8d to represent hands.

1 Executive Summary

The effect of luck (i.e., the dealing of the cards) in Texas Hold'Em is a subject of much debate in the legal community. This study seeks to establish clear numbers derived from a significant sample of actual play. This study does not quantify the effect that luck has on Texas Hold'Em, but it provides compelling statistics about the way that the outcomes of games are largely determined by players' decisions rather than chance.

Cigital examined 103 million hands of Texas Hold'Em poker played at PokerStars. In the majority of cases, 75.7% of the time, the game's outcome is determined with no player seeing more than his/her own cards and some or all of the community cards. In these games all players fold to a single remaining player who wins the pot. In the 24.3% of cases that see a showdown (where cards are revealed to determine a winner), only 50.3% of showdowns are won by the player who could make the best 5-card hand. The other roughly half of the showdowns are won by someone with an inferior 5-card hand because the player with the best 5-card hand folded prior to showdown.

We use accepted statistical sampling formulas to make the argument that these statistics are generally representative of Texas Hold'Em in Section 2. The findings themselves are presented in Section 3. In order that the artifacts can be reused with confidence, the cryptographic signatures of all contributing data are listed in Section 5.

2 Goals and Methodology

The purpose of this analysis is to determine certain statistical qualities of the game of Texas Hold 'Em as played at PokerStars.com. Given the specific results from analyzing PokerStars.com, we want to generalize the results and say mathematically that they represent the game of Texas Hold 'Em as a whole. It is important that Cigital conduct this analysis independently and without predisposition towards the final outcome.

2.1 Data Acquisition

Cigital acquired data from Rational Entertainment Enterprises Limited (REEL) related to play at PokerStars.com. The log files are ar-

chived by Cigital and their SHA-1 signatures are recorded in Section 5. The log files contain descriptions of the play of many hands. Table 1 shows two groups of log file lines that describe two different games. Note that user IDs have been changed and the hand IDs are fictitious to protect the confidentiality of this data.

Game	Blind	Bet	Hand ID	Board	User ID	Pos	Win	Hole	Best Hand	Show
No Limit	100	200	1399167686	8dKcTd9sQd	Player A	0	0	KsQh	KsKcQhQdTd	1
No Limit	100	200	1399167686		Player B	1	0	2s7s	7s2s	0
No Limit	100	200	1399167686	8dKcTd9sQd	Player C	2	1	4d5d	QdTd8d5d4d	1
No Limit	100	200	1399167686		Player D	3	0	Qc8s	Qc8s	0
No Limit	100	200	1399167686		Player E	4	0	5c5h	5c5hKcTd8d	0
No Limit	100	200	1399167686		Player F	5	0	Tc2d	Tc2d	0
No Limit	100	200	1399167686		Player G	6	0	AsKh	KhKcAsTd8d	0
No Limit	100	200	1399167686		Player H	7	0	3h2c	3h2c	0
No Limit	100	200	1399167686		Player I	8	0	Ah6h	Ah6h	0
No Limit	10	25	1299170765		Player A	0	0	5cQs	5c5sAdQsJh	0
No Limit	10	25	1299170765	9s2d5sAdJh	Player B	1	1	2hTh	2h2dAdJhTh	0
No Limit	10	25	1299170765		Player C	2	0	6c3c	6c3c	0
No Limit	10	25	1299170765		Player D	3	0	3h7s	7s3h	0
No Limit	10	25	1299170765		Player E	4	0	5dTd	Td5d	0
No Limit	10	25	1299170765		Player F	5	0	8c6s	8c6s	0
No Limit	10	25	1299170765		Player G	6	0	3sAc	Ac3s	0
No Limit	10	25	1299170765		Player H	7	0	Kh7c	Kh7c	0
No Limit	10	25	1299170765		Player I	8	0	JsQh	JsJhAdQh9s	0

Table 1: Example Log Data

In the first game, 1399167686, both Player A and Player C went to a showdown. This is indicated both by the fact that the "board" column contains the board on both players' rows and by the fact that the showdown column is "1." Player C wins with a flush: Q♦T♦8♦5♦4♦ against Player A's two pair.

In the second game, 1299170765, the board is listed next to the singular winner, Player B. In this case, there was no showdown, even though the entire board (all five cards) were dealt. This indicates that all players still in the game when the river was dealt eventually folded to Player B. It is worth noticing that Player B had a pair of 2's as his best hand. Several players (A, G, and I) would have beaten that hand, had they stayed in.

Cigital analyzed 103,273,484 such hands that had the following characteristics:

Cash Ring   No play money games were considered. No

Games	"heads-up" tables were included. That is, there are some two-player games in the sample set, but they are situations where two players sat and played against each other at a table that would allow more than two players.
Blinds 10¢ or higher	So-called "microlimit" games (games with blinds less than $1) are considered too much like play money games, so only a few such games (10¢, 25¢, and 50¢) were included. The 2¢ and 5¢ games were excluded.
December 1, 2008 to January 2, 2009	Cigital selected this timeframe because it needed to independently corroborate a subset of the hands played with the actual players themselves. See Section 2.4.

2.2   Data Analysis

For each hand analyzed, two facts were determined:

1. *Did the hand end in a showdown?* A "showdown" is a situation where all four rounds of betting have been completed and more than one player remains in the game. At least one player must show his cards so the winner can be determined.

2. If there was a showdown, *did the player with the best two cards win the hand?* It is relatively common for the best two cards (i.e., the player who would have made the best 5-card hand at showdown) to fold prior to the showdown.

2.2.1   Showdown Determination

Whether or not there is a showdown is a very simple fact to determine. There is no controversy or explanation necessary. Either there was more than one player in the game after all the betting was complete, or there was not.

2.2.2   Best Hand Win Determination

Determining whether the best hand won the showdown requires assumptions to be made. We are considering whether the player whose hole cards would combine with the board to make the best 5-card poker hand was actually the player who won at showdown.

At least two situations arise occasionally that could be considered a best-hand-win or not.

Equivalent Hands: Assume the board is K♦4♣Q♦2♠J♥, and Player A has A♦T♣ and Player B has A♣T♠. Both have an Ace-high straight. Assuming no other players have better hole cards, both Players A and B would win at the showdown and would split the pot. If Player A folds early, but Player B goes on to the showdown, Player B will win the entire pot. It is arguable that since one of the two equivalent hands did go on and win, that the best hand did win this game.

Board Best Hand: In some cases the board is the best hand. For example, if the board is 8♦8♣8♥2♠2♥, it is quite likely (though not certain) that no player has a better hand than a full house 8s full of 2s. In such a situation, where no player's hole cards improve the board, all players who stay to the showdown will split the pot. If one or more players fold before the showdown, they will not share in the pot. This situation is a special case of the "Equivalent Hands" case, because in this situation all players are equivalent. Again, it is arguable that since some hands win at the end, the best hand did win the game.

Cigital has chosen to count both of these situations as hands where the best two cards **did not** win. Since there were players who folded early, but would have been paid had they stayed in, there were "best hands" that did not win. Using the alternative method and counting such hands would have only a small impact on the final result as such hands are relatively rare.

2.3   Statistical Method

Games in the log data were organized by "game type." Game type is a combination of the game rules (i.e., Limit, No Limit, or Pot Limit), any restrictions on the table size (e.g., 10 players or 6 players) and the blind/bet sizes. For each game type we then performed a statistical analysis of the percentages of showdowns and percentages of showdowns won by the best hand to see how representative they are of Texas Hold 'Em poker hands in general.

2.3.1   Description of the analysis

We are assuming that the distribution of the number of hands that go to showdown and where the best hand won follow the binomial distribution. Specifically, we are treating each hand as a separate

independent test, where the results of one hand have no bearing on the results of any other.

When the amount of data is large (as it is in our survey) the distribution of proportions of binomial data fits closely to a normal distribution. This process has several steps:

1) We define X (the number of successes) and N(the sample size). For our purposes, X is the number of hands that went to showdown in the limit we are examining (or, the number of hands where the best hand won). N is the total number of hands surveyed at the limit we're examining.

2) We construct the Wilson Estimate of the proportion:

$$\tilde{p} = \frac{X+2}{N+4}$$

The Wilson estimate is a popular way of adjusting a proportion by acting as if we had two more successes and two more failures. Notice that when the sample size is large (as it is in the majority of our surveys) this adjustment will have almost no effect.

3) We determine the standard error of the proportion (again, assuming that the proportion can be approximated by the normal distribution):

$$SE_{\tilde{p}} = \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}}$$

which is just the standard deviation under the normal distribution under our Wilson estimate.

4) We then determine a desired confidence level C and determine a confidence interval:

$$\tilde{p} \pm z^{*}SE_{\tilde{p}}$$

where z* is the value for the standard normal density curve with area C between $-z^{*}$ and z^{*}. We computed this value for z^{*} in Microsoft Excel as follows:

(a) Given the confidence percentage C, we compute the probability of anything being outside of the confidence interval

on the right side of the normal distribution by:

$$p = \frac{1 - C}{2}$$

(b) We then use the Microsoft Excel "NORMSINV" function to find the inverse of the standard normal distribution at probability p. This gives us our z* value. It should be noted that Excel uses an iterative search technique to generate the result, and so the results may not be exactly accurate. However, several checks were made against standard tables and the results of NORMSINV were found accurate to at least three decimal places.

5) Once we have our confidence interval, we can define the margin of error as:

$$m = z^* SE_{\hat{p}}$$

6) If desired, we can also fix a desired margin of error, and compute the required z* (and thus the required confidence level) needed to reach this margin of error by inverting this process.

For the case of determining the number showdowns won by the best hand, we perform the same analysis. We let X represent the number of hands won by the best hand in the limit we are examining. We let N be the total number of showdowns surveyed at that limit.

2.3.2   Assumptions and possible sources of error

As was alluded to above, we made several assumptions during this process. If these assumptions are not valid, that may impact the accuracy of our results.

1) We assume that the data surveyed follows the binomial distribution. Specifically, we assume that each hand is an independent event with fixed probability of a showdown, and that the result of whether one hand went to a showdown has no bearing on whether a subsequent hand goes to showdown.

2) We use the normal distribution to approximate the distribution of the proportions. This is just an approximation, and introduces a potential source of error. However, this is an accepted approximation when $n^*p \geq 10$, and $n(1-p) \geq 10$ (where n = the sample size, and p = the proportion of hands that go to showdown),

and all of the limits examined are well beyond this lower bound.

3) We assume that December 2008 is a representative month of normal play at PokerStars, and that there is nothing special about it that would cause our extrapolations about how it represents other months in general to be wrong.

4) We assume that the calculations made, both the ones provided by Microsoft Excel functions, and the ones that were made to implement the formulas, are correct. Several entries were checked by hand and found to be correct.

5) We assume that the data collection was accurate, and that PokerStars gave us a complete and accurate representation of all hands played in the requested month, and that the collection of the "number of showdowns" and "total number of hands played" data is correct. Rather than take PokerStars' log files at face value, we performed independent corroboration directly with some players, as described in Section 2.4.

2.4   Verifying Log Data

PokerStars players were asked to independently submit their hand histories to Cigital, along with an attestation that the hand history was accurate.

2.4.1   Rationale

Part of the reason that we chose December 2008 as a sample month was so that the players would have their histories fresh. It gave them the best opportunity to honestly recollect their hands.

2.4.2   Mechanics

Each player sent their history by email. It included the following affirmation statement: *I, __NAME__, affirm that, to the best of my recollection, the attached data is an accurate representation of my activity on PokerStars.com.*

One might dispute the idea that a player can remember 60,000 hands accurately. The players who submitted histories are the kinds of players who use databases while they play. As each hand finishes, it is stored in their personal database. Certainly the player would notice a loss being recorded as a win and such obvious mistakes. The kinds of players who submitted hand histories are diligent and scrupulous about recording and analyzing their play. So,

while it is unlikely that they remember all 60,000 hands in mid-January, it is highly likely that they vetted those hands as the hands were added to their database. Furthermore the data the players provided was directly from their private databases, not from PokerStars itself. That is, it was data that they collected prior to our announcement of this study or any request for assistance. Thus, an extraction from their personal databases can be considered independent of PokerStars' influence.

2.4.3   Results

Cigital received 14 player histories covering 760,836 games. Out of that set of histories, 714,439 games applied to our sample set. The other 46,397 hands were either from the wrong date (e.g., November 30) or were from tables we are not analyzing (tournaments, heads-up, low-limit, etc.). This yields 0.69% of hands in the sample data directly confirmed by players. We treat these as samples of log data where a "successful test" is when the player's personal data match PokerStars' log file, and an "unsuccessful test" is when they don't.

All the players' histories agreed with PokerStars log files exactly. We conclude that there is a 99.99% chance that the accuracy of ALL hands is 99.99% ± 0.001%. It is highly improbable that PokerStars modified the data in the log files.

3   Findings

The short summary of our findings is that 24.3% of hands result in a showdown. Of that 24.3% of hands that result in showdown, 50.3% of them are won by all players that were dealt two cards that combined with the board to make the best 5-card hand.

3.1   Margin of Error

To calculate the margin of error, we assumed a confidence level of 99.99%. The margin of error for the calculation of showdowns is estimated at ± 0.02%. The margin of error for the calculation of best hands winning is estimated at ± 0.01%. Individually, all but eight of the 55 game types had margins of error < ± 1%. Those eight game types did not experience significant play volume in the sample.

To explain the effect of margin of error, consider a specific game-type: No Limit 10¢/25¢ in December 2008. 26.1% of those hands went to showdown that month at that limit. If we were to sample

lots and lots of months, we would expect some months to have a higher percentage, some months to have a lower percentage, and so on. These different percentages would stack up in a normal distribution (the bell curve, see Figure 1) **assuming that there is no reason for there to be differences in the data, other than random chance**.

That final assumption is critical. We can only extrapolate these values to be representative of reality if we assume that December 2008 is representative of reality.

Since the samples of all of the months fall into a normal distribution, we need to determine what the odds are that example month falls into the "fat" part of the bell curve. That's where confidence intervals and margins of error come into play.
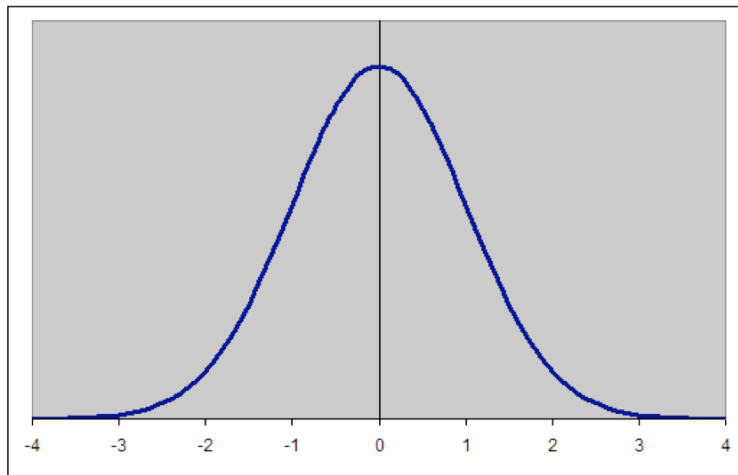

Figure 1: Standard Bell Curve

Figure 1 is a "standard" distribution, which means that it has been rescaled to be centered around 0.

Given that 26.1% of the hands went to showdown. We want to know how likely it is that the "real" bell curve for this situation has its center at, or close to, 26.1 (in other words, how likely is it that the "0" position in the picture is really at 26.1?). Obviously, it is unlikely that it will be exactly 26.1%, but the margin of error gives us a range. If we set the margin of error to 0.1% in the calculations we are asking *How likely is it that the center is 26.1%, ± 0.1%*? It's never a sure thing—it's always theoretically possible that we had a freakishly weird month, but the more hands we sample, the less likely that's true. This is just like it's not too hard to have 9 out of 10 coin flips come up heads, but it's really unlikely—though theoretically

possible—to have a 90% heads rate after a million coin flips. The confidence interval comes out to about 99%, and it's based on the margin of error we set. So, what that means is that it is 99% likely that the "0" position of the bell curve in our situation is between 26.0% and 26.2%.

If we increase the margin of error, our confidence goes up (because we have a wider range to cover, so it's more likely that the real center is in that range). If we decrease the margin of error, our confidence goes down (for the same reason).

We can also perform this calculation in the reverse direction. Suppose we want to have a certain confidence that the results are not a fluke. How wide a margin of error do we need for it to be that likely? If we work in this direction and look for a confidence level of 99.99%, we figure out how wide a band of possibility is needed to be 99.99% likely that the "0" position of the real distribution is within that band, based on our estimate. It turns out to be 0.05%. In other words, we believe it is 99.99% likely that 26.1% ± 0.05% of hands at the 10¢/20¢ limit will end up in a showdown.

4   Conclusion

It is clear from these numbers that, at least in the sampled data, the majority of games are determined by something other than the value of the cards, since no player reveals any cards to determine the winner. Only rarely (about 12% of all hands) does the player who can make the best 5-card hand go all the way to showdown and win. The statistical analysis of the logs gives us confidence that the logs accurately describe what was played. The analysis of the hands gives us confidence that this sample represents online Texas Hold'Em at PokerStars as a whole.

5   Recorded Artifacts

The following log files and hand histories were received, stored, and used for this analysis.

5.1   PokerStars Log Files

File	SHA-1 signature	File	SHA-1 signature
HandsDec01.txt.gz	c5501596528dc717338b2a53c0d224c125d79729	HandsDec17.txt.gz	e0f82db68d4411724a45b5c383ff8e0ebf790a58
HandsDec02.txt.gz	90caeb2cbda43c7720d628bb3f92d731b7128ad9	HandsDec18.txt.gz	6f4d4209b78bdcf0a7486fea5e92b7d4678e3123
HandsDec03.txt.gz	cf3aac342ded4951d550090d4dcf05bc77ca633a	HandsDec19.txt.gz	4bd8bdf4e28b01d10a94e87d93d631f7f36b8c15
HandsDec04.txt.gz	b8d4c3dc5301384fd7e9da6210c0f04ed248aa98	HandsDec20.txt.gz	a318b050d9f4c019531fe1295c334bb1aa6cc68b
HandsDec05.txt.gz	717d0d87cd7d290533f3b70a9e9cb8b5f0bf7f6e	HandsDec21.txt.gz	b3920863256aa224831eebeaf93cf1145f6435ca
HandsDec06.txt.gz	8150330d3b7eb38af78c83ed6c0a3a45c197e216	HandsDec22.txt.gz	eaa2fdec8512a2cef09c89188600640e68cfca24
HandsDec07.txt.gz	2289a717c1896468d069b6331e96a4197317d446	HandsDec23.txt.gz	623c5a6e5021e1560cbfcbe506c8cf7fe40af8c6
HandsDec08.txt.gz	641ffb8ed18a27d17fd7ba7d25646257cf7343ac	HandsDec24.txt.gz	524c35fb57532166bf684f6ac0f64bd0e1c76093
HandsDec09.txt.gz	bfb86ba566571a2b5fb5b2d3cd8bc97770c2bfc5	HandsDec25.txt.gz	1996e0479bb2e8bc5557578c13d3ea4b591639f5
HandsDec10.txt.gz	20f27406f47b080cb0cd09112dde2f52deb96453	HandsDec26.txt.gz	14e1c82537b2a1c88bae32e4fbc53f738cbe4ef5
HandsDec11.txt.gz	1fb1d1ade45fd2b649e055956494ca207e076bf8	HandsDec27.txt.gz	d0d13614584ab7e6b335df8f402e6d8c94b309a5
HandsDec12.txt.gz	3aee3fd7a538096104ffbf22a9f44b010beb13b7	HandsDec28.txt.gz	7373859b2120dc6681b9d382abd0c7dedde9bb3b
HandsDec13.txt.gz	2dc2b691fc6559ea5f0d3553616ebcad1a96529e	HandsDec29.txt.gz	d901cdc805c2fed8561f119139503b5e187f03a6
HandsDec14.txt.gz	df5f318f3b0f97f49a65369a1d849109c2a572f4	HandsDec30.txt.gz	44214e493fdaf335aa019077b7066c2254650597
HandsDec15.txt.gz	5ec47e468f03c51ac6637c2d567806ed370200f4	HandsDec31.txt.gz	19ec3cdfa2beddeb2bf39a81a5d62871e732877c
HandsDec16.txt.gz	d1384390abae8ec2c927892a364bd78b0ffc45c6	HandsJan01.txt.gz	b5ee0ff2401ef9c03551159f45244a8ad2368bc1
		HandsJan02.txt.gz	94e55df1892c64bfa7a4e7a804b6bd4ee5f891cc

5.2   Player-submitted Hand Histories

SHA-1 Signature	Archive File
f620fad11de3347002f76b680bc215469d4236c9	furbean.zip
5588409225a4a09482008301e21a72d37731df01	LihanLi.zip
621d2508b6836fce55169acc5d344e9b3e1e47bb	basile.zip
9b6ed3073b4bc4823f7fe274b255ee5c6b9728b8	buntaine.zip
0cee4d4e03cb472d08bbb9674fb8c4504e10324b	stein.zip
dd1deac5a8f17c7715e886b2077e9764902be06f	Zeidler.rar
6e424fc2ed793429a60fba34e5362f195a0345f9	aguirre.zip
4edbcac2eb92883077cc6fbd84f48c3ad89f4cfc	ajtai.zip
57c5cf23a558dd271c936b92377d76b310c94ad2	boyett.rar
8fb769442b43475d270a4f81a61a26e0cc6ba495	linnane.zip
2cdd02064d95853181db54038e79ac3f10962366	smith.zip
3aafbafec8e31ee3c890609ec324d8989f58bf77	zorec.zip
1486eaa0b119b6aad360a8b8552d05c2e22b65bd	stabile.zip
8849e37642aa24f016987c1c8e0c9ce8c84b178d	pfaff.zip

For More Information

For more information about this document, contact:

Contact	Title	Organization	Phone #	Email Address
Mr. Brian Mizelle	Managing Principal	Cigital, Inc.	+1 703 404-5820 direct	bmizelle@cigital.com
Mr. Stuart Dross	VP, Sales	Cigital, Inc.	+1 703 404-5876 direct	sdross@cigital.com
Mr. Paco Hope	Technical Manager	Cigital, Inc.	+1 703 404-5769 direct	paco@cigital.com

Table 2: Contact Information

About Cigital, Inc.

Cigital helps commercial and government clients assure software quality and improve software development processes. Our Software Quality Management (SQM) solutions drive down the cost of deploying quality software and ensuring software reliability, security and performance. Cigital's expert Consultants measure software quality by combining proprietary methodologies, tools and knowledge to perform full-lifecycle testing via a risk management framework. The resulting metrics are used to drive application readiness decisions and identify the most cost-effective areas for software process improvement. Founded in 1992, Cigital (www.cigital.com) is headquartered in Northern Virginia with additional offices in Boston.