Guess the Mean

Joshua E. Hill
e-mail: josh-math@untruth.org

Initial Release: December 30, 2009
Current Revision: August 26, 2011

Game: Provide a rational number[1] in the interval $[1, 100]$. The winner will be the person whose guess is closest to 2/3rds of the mean of all the guesses.

My Entry: 18.5053.

1 GENERAL APPROACH

Let $\bar{X}$ be the distribution of means of such games, and then let $Y = \frac{2}{3}\bar{X}$. We wish to provide a guess as close as possible to a value drawn from Y. To do this, we will attempt to estimate the underlying distribution of $\bar{X}$, which allows us to easily determine the distribution Y. We choose the median of Y (statistics tells us[1, p. 212] that the median has the minimum expected distance to values selected from Y.)

2 PLAYERS DESCRIPTIONS

To be able to play reasonably, we need to characterize the dominant playing styles in the game.

2.1 NAÏVE PLAY

A player who either can't be bothered to read the rules, or who doesn't put thought into the selection of a number may arbitrarily choose a real number between 1 and 100. This can be modeled[2] by treating this arbitrary selection as a value chosen from a uniform distribution.

---

[1] We can ignore this constraint, as any real number can be approximated by a rational number to any desired ε-bound.

[2] poorly.

If there were a population of these Naïve Players, then the central limit theorem tells us that the average of the guesses follows a normal distribution. That is if $X_j \sim \text{Uniform}\,(a = 1, b = 100)$ and

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_j$$

then we have (approximately)

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \text{NormalDistribution}\left(\mu = 0, \sigma^2 = 1\right)$$

This gives $\bar{X}$ the mean of μ, the mean of the original sample distribution, $\mu = 50.5$. The normal distribution is symmetric, so the median of the distribution is also 50.5.

This approach has a number of limitations. First, the Central Limit Theorem only provides equality asymptotically, and the sample population is certainly not infinite. Additionally, humans don't select random numbers very well.

We used simulation to confirm that the mean behaves roughly as expected for a small number of samples from a uniform distribution. The results of this simulation are summarized in table 1.

Table 1: Small Normal Population Simulation Summary

n	Simulated Median μ
10	50.5129
25	50.4966
50	50.5068

The simulation shows that the observed values are very close to the theoretical values, so it seems that applying the central limit theorem is reasonable for the purpose of estimating the mean and median sample mean, at least in the case where the underlying distribution is uniform[3].

In order to simulate this, we can either draw values from the stated normal distribution, or model it as all naïve players choosing 50.5.

2.2 Rational Play

If one were to assume that the entire player population consists of totally rational players who make certain assumptions about the other players in the game, one can easily model the result.

[3] Which it is not.

2.2.1 Rational players

We'll develop a sequence of intervals that contain the guess of all rational players, under the assumption that all the other players are also rational and making the same assumption. As all such rational players will guess within all of these intervals, their guesses will all lie in the intersection of all of these intervals.

Only a subset of the numbers in $[1, 100]$ can be possible results for values selected from the underlying sample distribution X. If all players chose the maximum value of 100, then the average value would be no larger than 100, so the corresponding value selected from Y would be ≈ 66.6667. Similarly, if all the players chose 1, then the average value would be 1, corresponding to the selection of Y of $\frac{2}{3}$. Thus all values selected from Y must lie in $[.6667, 66.6667]$, so all rational players will choose a value in the interval $I_1 = [1, 66.6667]$.

We now know that the mean will fall in the interval I_1. Thus, $\frac{2}{3}$ of the mean must fall in the interval $\left[\frac{2}{3}, 44.4444\right]$, so all rational players will choose a value in the interval $I_2 = [1, 44.4444]$. In general, following this same reasoning for j levels, you find that all rational players (under the assumption that everyone playing the game is also a rational player) will choose values in the interval

$$I_j = \left[1, 100\left(\frac{2}{3}\right)^j\right]$$

until this interval makes no sense. $I_{11} = [1, 1.1561]$, so all rational players would choose 1. Similarly, the same applies for all future rounds.

In summary, we see that

$$I_j = \begin{cases} \left[1, 100\left(\frac{2}{3}\right)^j\right] & j \leq 11 \\ [1, 1] & j > 11 \end{cases}$$

A rational player operating under this set of assumptions will only choose 1. This is the Nash equilibrium [8] for this game, as any variation from this strategy would lead to a decrease in the expected payoff for any particular player.

Simulating this functionality is easy: all such players choose 1.

2.2.2 j-rational players

In the case that the playing population is "large" we can apply analogous reasoning in the case where a player makes different assumptions about the style of playing population. A player is called 1-rational if they assume that the rest of the population was made up of naïve players. They would then expect the mean to be 50.5, so they would choose $\frac{2}{3}$ of this value or approximately 33.6667. A player is called 2-rational if they assume that the rest of the players are 1-rational, and thus choose $\frac{2}{3}33.6667 \approx 22.4444$.

Similarly, we'll call players j-rational if they assume that all of the other players are $(j-1)$-rational. A j-rational player will thus choose

$$50.5\left(\frac{2}{3}\right)^j$$

As a small abuse of this formula, we can also describe 0-rational players as players who correctly determine the behavior of naïve players, but then fail to apply the $\frac{2}{3}$ scaling. This is a rather inexplicable playing strategy, but it appears to be common.

Due to the possible selection range, we see that after 10 rounds the optimal choice is 1. As such, 10-rational players are indistinguishable from rational players who assume that every player is also rational, as are any j-rational players where $j \geq 10$.

j-rational players are called a player with "j order beliefs" in [3], which describes this style of player in essentially the same way.

It is interesting to note that 0-rational players are effectively the same as naïve players in bulk; that is, if there is a sufficient population of naïve players (or looking at the mean the mean of the guesses of a small population over many trials) the resulting mean converges to 50.5, which is the same as the fixed response of a 0-rational player.

Simulating this style of player is simple. j-rational such players choose the particular value described in Table 2.

2.3 Spoiling players, colluding players and other disruptions

Some players may not make guesses with the intent of winning, but rather submit guesses with the intent of making the game have some surprising result. As the game theoretic approach shows that a rational player chooses 1, a spoiling player would likely choose 100 for maximal disruption. We have thus far left the population size unspecified (other than for the purpose of simulation) but for the effect of a spoiler to be significant, the population must be relatively small. For example, if the population is 10 and 9 of the players are rational, a spoiler can force the mean to almost 11. As the number of players increase the effect of a single spoiler isn't as large, though the presence of several independent spoilers could still cause a significant effect.

Simulating spoiling players is easy; they all select 100.

If all the players other than me collude, then any mean in the realm of possible answers is possible to attain. As the effects of such large-scale collusion are impossible to predict, there is no point in including the effects of such large-scale collusion within this analysis.

2.4 OPTIMISTIC PLAYERS

An Optimistic player is effectively a 1-rational player who operates under the assumption that every other player is a spoiling player. The optimistic player thus guesses the highest possible value that a rule-aware player would choose, 66.6667. Irrespective of how baffling this strategy appears, we will shortly see that it not uncommon.

3 SMALL PLAYER POPULATIONS

The above analysis generally presumes a "large" player population, but generally informs a rational approach even if the playing population is quite small; the main result of a small population is less predictability. The asymptotic findings above are still the basis of a rational playing strategy, as across many games-with-small-population the resulting distributions are the same.

The main difference is that with small populations, each player's guess significantly affects the sample mean, so the player's own guess must be taken into account in establishing the sample mean. As an extreme example, imagine one naïve player and one 1-rational player playing this game (so, the 1-rational player has correctly categorized the other player and chosen what should be the optimal strategy). The naïve player chooses a random number in the range $[1, 100]$. The 1-rational player assumes that the other player's choice will have mean 50.5, so by the above his own guess should be 33.6667. Repetitive iterations of this game produce a sample mean (and median) of 42.0833, not 50.5, so the 1-rational player's strategy was incorrect. The end result is that the 1-rational player should expect to win only 66.9% of the time.

The 1-rational player should have included the effect of his guess on the game mean. If A is to be the 1-rational player's guess, then A should be chosen such that

$$\left(\frac{2}{3}\right)\left(\frac{50.5 + A}{2}\right) = A$$

which yields a guess of 25.25. This revised guess results in the mean 37.875, and $\frac{2}{3}37.875 = 25.25$, so this new guess is correct under these assumptions. Under this strategy, the 1-rational player can expect to win 75.4% of the time.

Generalizing this to a playing population of n players, we see that if B is the expected mean for the game for the other $n-1$ players we have

$$\left(\frac{2}{3}\right)\left(\frac{(n-1)B+A}{n}\right)=A$$

which yields a guess of

$$A=2B\frac{n-1}{3n-2}$$

Note that the difference between the player's guess in the $n=\infty$ case and the finite player case is

$$\frac{2}{3}B-2B\frac{n-1}{3n-2}=\frac{2B}{9n-6}$$

For fixed values of B, this clearly approaches 0 as $n\to\infty$, which conforms to our intuition.

This correction does not affect the Nash Equilibrium; accounting for the result of a lower guess simply drives the optimal guess to 1 more quickly!

For j-rational players, the effect is to render our nice formula for j-rational guesses useless because each j-rational player adjusts for the small population. Instead, a simple recurrence relationship can be be used.

Listing 1: Small Population Guess Code

```
SmallSampleGuess[B_, n_] := 2*B*(n - 1)/(3*n - 2)
jRational[0, n_] := 50.5
jRational[j_, n_] := SmallSampleGuess[jRational[j - 1,
    n], n]
```

Table 2 summarizes the population-corrected guesses for various players for a variety of sizes of player population. Note, we only include player styles that would alter their guess in response to the likely population (so, we do not include naïve players, 0-rational players, rational players, optimists, or spoilers).

4 COMBINING THE DISTRIBUTIONS

In some sense, this problem becomes reduced to two questions: "How many players are there" and "What sort of players are these?". By answering both of these questions, one can estimate median of the resulting distribution of the means, which leads to the guess.

Table 2: Effect of Small Populations

	Population (n)					
	2	5	10	25	50	∞
1-rational	25.25	31.0769	32.4643	33.2055	33.4392	33.6667
2-rational	12.625	19.1243	20.8699	21.8337	22.1422	22.4444
3-rational	6.3125	11.7688	13.4164	14.3564	14.6617	14.963

The number of players does have an effect, but it is not profound so long as it is not very small (less than 5). In this game, we make the guess that the number of players is greater than 10, as this appears to be something of a sociological study, and in that context a small sample size would not be useful. Conversely, it seems unlikely that there are more than 50 players for purely logistical reasons.

It then becomes a matter of determining what the distribution is of players. This isn't wholly uninformed guessing! We must estimate the distribution of the types of players. Prior studies in games similar to this exist. There is a survey of many such games contained in Nagel's survey paper [4], and a few large scale studies are analyzed in Nagal et al's [5].

Combining these is straight forward; add each of the player's expected means scaled by the proportion of those users.

4.1 Nagel's Study

One small scale, Nagel's 1995 paper [3] is similar (it is a $[0, 100]$ study with a variety of scaling values), though the conditions are sufficiently different to complicate the data use. The main differences that seem to affect the usability of this study data are:

- Nagel presented the rules, and then asked for guesses with no significant time allotted for thought. In our game, a significant amount of time was provided to the contestants to supply answers, which would allow for thought on this game. This time should reduce the number of people acting naïvely.

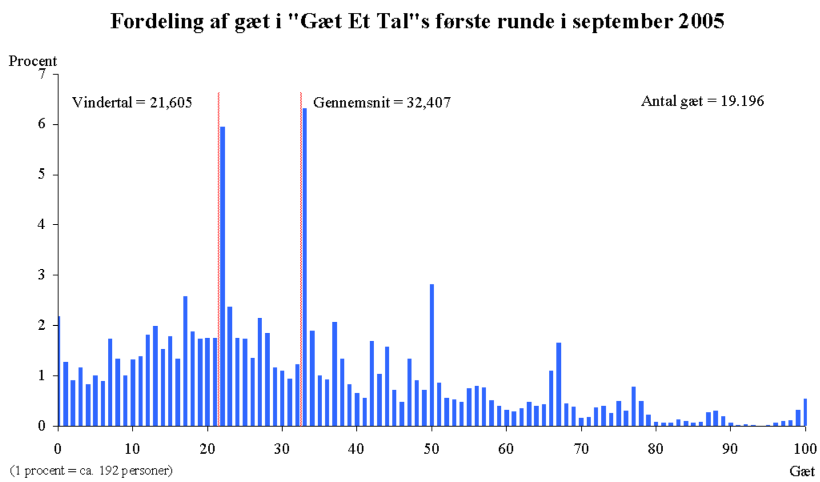
- This game is well known, and various excellent write ups are on the internet and easily found on Google. This game is described on Wikipedia [7], which in turn provides a link to the Nagel paper [3], which is freely available and accessible through a quick Google search.

Happily, there are other sources that are more similar to this challenge.

4.2 POLITIKEN CONTEST

The Danish newspaper Politiken ran a similar contest in 2005 [6] (it is different in that it asks for values in $[0, 100]$) which involved 19,196 participants. The histogram for this contest was helpfully published by Department of Economics, University of Copenhagen [2] and can be seen in figure 1.

Figure 1: Politiken Study Results



The notable spikes at the 0-, 1-, 2-rational, rational, and optimistic player values are particulary gratifying. We categorized players in the interval $[67.5, 98.5)$ as naïve, and then assumed that they were similarly distributed throughout the entire range. In this case, this procedure suggested that this population was 20.1% naïve, so we removed 37.2 from each length 1 bin to remove the contribution of the naïve players. The remaining players were categorized as described in Table 3.

We then estimate the results of this by assuming that the given proportion of people answer with their center value. This crude measurement gives a mean of 32.2118, which is only an error of 0.6% from the actual mean of 32.407.

Table 3: Politiken Population breakdown

Play	Center	Interval	Percentage
Naïve	50	(uniform)	20.1%
rational	0	[0, 10.5)	12.1%
3-rational	14.8148	[10.5, 17.5)	11.5%
2-rational	22.2222	[17.5, 26.5)	19.4%
1-rational	33.3333	[26.5, 39.5)	21.2%
0-rational	50	[39.5, 61.5)	15.1%
Optimists	66.6667	[64.5, 67.5)	2.7%
Spoilers	100	[98.5, 100]	0.6%

4.3 THIS CONTEST

We modify these findings based on a few factors:

- We expect that the people in Jesse's social circle will tend to behave somewhat more rationally than the average reader of Politiken.[4]
- We expect that the people in Jesse's social circle will tend to have a higher than expected likelihood of being a spoiler.[5]
- The e-mail invitation for this game explicitly mentioned Nash equilibrium, thus encouraging a formal approach to the problem, or at least suggesting to the players that there are formal ways of approaching the game.
- The e-mail invitation for this game requested a rational for picking the number provided, which suggests that there should be one. This will likely reduce the number of random guesses.

We assume the player mix described in table 4.

Based on this player mix, the we get the projected means and the associated guesses in table 5.

We arbitrarily select the 25 player population, which establishes our guess as 18.5053.

[4] Jesse could consider this a compliment.

[5] It ain't all roses!

Table 4: Projected Contest Breakdown

Play	Percentage
Naïve	9%
rational	18%
3-rational	14%
2-rational	20%
1-rational	22%
0-rational	14%
Optimists	1%
Spoilers	2%

Table 5: Projected Contest Medians and Guesses

n	Mean	Guess
10	27.6561	17.7789
25	28.1435	18.5053
50	28.2994	18.7388
∞	28.452	18.968

5 RETROSPECTIVE

This contest ran between December 21, 2009 and December 31, 2009. The actual contest data (sans my guess) was as follows:

1, 1, 1, 1, 1.1, 5.759, 7, 12, 18.13, 20, 22.22, 23, 23, 23, 23.287, 27, 28, 28, 33.3, 36, 39, 42, 42, 43.2, 49.494, 50, 50, 56, 66.6, 66.666, 67, 67, 75, 75, 75.33, 79.21, 88, 100

There were 39 participants, and the sample mean was 38.0871 so $\frac{2}{3}$ this sample mean was 25.3812.[6]

We categorize the guesses using the same analysis as applied to the Politiken contest to obtain a player breakdown, as in Table 6.

So, if we had correctly guessed the player breakdown and population in advance, we would have guessed a mean of 39.0184 and thus submitted the guess 25.7861, causing the new mean to be 38.2585. $\frac{2}{3}38.2585 \approx 25.5057$ which is a 1.1% error (which would have been the winning entry).

Looking at the differences between my speculation of the population breakdown and the actual population breakdown:

[6]Note: My original guess was actually 19.1018, but we since added the "optimist" player category which changed the result. The original mean was 38.0871, $\frac{2}{3}$ of which is ≈ 25.3914.

Table 6: Contest Population breakdown ($n = 39$)

Play	Center	Interval	Percentage
Naïve	50.5	(uniform)	42.1%
rational	1	[1, 11.5)	14.0%
3-rational	14.5760	[11.5, 19.5)	1.9%
2-rational	22.0558	[19.5, 28.5)	19.9%
1-rational	33.3739	[32.5, 41.5)	4.1%
0-rational	50.5	[48.5, 57.5)	6.7%
Optimist	66.6667	[64.5, 67.5)	9.3%
Spoiler	100	[98.5, 100]	2.0%

- My guess of the playing population was adequately close.
- My expectation that there would be a higher proportion of spoilers than in the Politiken game was correct.
- My expectation that the players would be largely influenced by the Nash equilibrium was correct, though we overestimated the amount of change.
- My expectation that there would be more "gamesmanship" that would generate a significant amount of 2- and 3- rational play was partially correct. There were a significant number of players choosing values near the 2-rational choice, which was also (not coincidentally) near the correct answer! Few players guessed a value near the 1-rational value, which is surprising. Almost nobody chose a value near the 3-rational value.
- My expectation that the people in Jesse's social circle would engage in more rational play than general populations was not correct. The largest group in this game was the naïve player group, suggesting that a very large proportion of the players arbitrarily selected their guess without much view toward the rules of the game. This could be because the players were not sufficiently motivated by a (large in the case of Politiken) cash payment.

So, we think that it's clear that the most significant finding here is that key to winning this style of contest is to know the data ahead of time.

As an additional finding, we propose that Jesse needs to offer a large cash payment to the winner of the next such contest (or, alternately, he could obtain a larger proportion of Danish friends).

6 FURTHER WORK

We've highlighted only a few distinct styles of behavior. Most of these behaviors produce a fixed guess, which greatly simplifies the analysis. Actual data suggests that people don't work this way; players seem to instead choose values that are weighted by the described distributions (or other distributions!) People, as always, are complicated.

A player that proceeds through the reasoning described above and then chooses a value near the center values described above can be modeled as selecting their guess from normal distribution that is centered at one of the described values. The final distribution would then be a sum of normal distributions with different means and standard deviations.

There has been work that attempted to characterize the distribution of players as overlapping normal distributions[5], but the identification and fitting of such distribution seems to be handled in an ad hoc fashion. It would be useful to find a natural way of discovering the parameters of these distributions.

REFERENCES

[1] Morris H. DeGroot and Mark J. Schervish. *Probability and Statistics*. Addison Wesley, 2002.

[2] University of Copenhagen Department of Economics. Fordeling af gæt i "gæt et tal"s første runde i september 2005. http://konkurrence.econ.ku.dk/r/0, 2005.

[3] Rosemarie Nagel. Unraveling in guessing games: An experimental study. *The American Economic Review*, 85(5):1313–1326, 1995.

[4] Rosemarie Nagel. Chapter 45 experimental beauty contest games: Levels of reasoning and convergence to equilibrium. In Charles R. Plott and Vernon L. Smith, editors, *Handbook of Experimental Economics Results*, volume 1, pages 391 – 410. Elsevier, Cambridge, 2008.

[5] Rosemarie Nagel, Antoni Bosch-domènech, Albert Satorra, and Jose García-montalvo. One, two, (three), infinity: Newspaper and lab beauty-contest experiments, 1999.

[6] Astrid Schou. Gæt-et-tal konkurrence afslører at vi er irrationelle. http://politiken.dk/erhverv/article123939.ece, 2005.

[7] Wikipedia. Guess 2/3 of the average — wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Guess_2/3_of_the_average&oldid=323284903, 2009. [Online; accessed 29-December-2009].

[8] Wikipedia. Nash equilibrium — wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Nash_equilibrium&oldid=331292424, 2009. [Online; accessed 29-December-2009].

Colophon

The text of this document is typeset in Jean-François Porchez's wonderful Sabon Next typeface. Sabon Next is a modern (2002) revival of Jan Tschichold's 1967 Sabon typeface, which is in turn a adaptation of the classical (in all meanings) Garamond typeface, which dates from the early 16th century.

Equations are typeset using the MathTime Professional II (MTPro2) fonts, a font package released in 2006 by the great mathematical expositor Michael Spivak. These fonts are designed to work with the Times typeface, but they blend well with most classical fonts.

XƎTEX was used to typeset the document, which is in turn offspring of Donald Knuth's profoundly important TEX. XƎTEX was selected in order to gain access to modern fonts without the trauma involved in converting them to a representation that pdfTeX could deal with. This approach makes most (though sadly, not all) OpenType features available, and sidesteps the traditional limit of 256 glyphs per font.