Guess the Mean

Joshua Hill

January 2, 2010

Challenge: Provide a rational number 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.

Answer: My guess is 19.1018.

1 General Approach

We wish to provide a guess as close as possible to $Y = \frac{2}{3}\bar{X}$ where $\bar{X}$ is the mean of the provided values. We will attempt to estimate the underlying distribution of the means, and then maximize the chance of success by choosing a guess that minimizes the expected distance between the guess and the expected sample mean. Statistics tells us [1, p. 212] that this expected distance ($E\left(\left|\bar{X} - c\right|\right)$) is minimized when c is chosen to be the median of the distribution underlying $\bar{X}$, so we'll guess the distribution's median (corrected to account for my guess). As we'll ultimately be combining several distinct distributions, we'll estimate the median using simulation.

Without loss of generality, we can look at real numbers rather than rational numbers, as any real number can be approximated by a rational number to any desired ε bound [4, p. 9].

Only a subset of the numbers in $[1, 100]$ can be possible results for $\frac{2}{3}\bar{X}$. If all players chose the maximum value of 100, then the average value would be maximized at 100 and then $\bar{X} = \frac{2}{3}100 \approx 66.6667$. If all the players chose 1, then the average value would be 1, thus the mean would be $\frac{2}{3} \approx 0.6667$. Thus $\frac{2}{3}\bar{X}$ must lie in $[.6667, 66.6667]$.

2 Naive Players

2.1 Theory

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[1] by treating this arbitrary selection as a random variable following the uniform distribution.

If there were a population of these Naive 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 $\bar{X} = \sum_{i=1}^{n} X_j$ then we have (approximately)

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

This gives $\bar{X}$ a variance of $\frac{\sigma^2}{n}$, where $\sigma^2 = \frac{1}{12}\left(b - a\right)^2 = 816.75$, and a mean of μ, the mean of the original sample distribution, $\mu = \frac{a+b}{2} = \frac{1+100}{2} = 50.5$. 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; past studies have suggested that the median is roughly consistent with a uniform distribution, but that the variance is not well modeled using this assumption. Thus the calculated variance of the $\bar{X}$ distribution described above is suspect.

[1] poorly.

2.2 Simulation

In order to confirm the basic behavior with small sample sizes, a very small Mathematica simulation confirmed that the mean and variance behaved as expected:

```
SampleStats[n_] := Mean[RandomReal[{1, 100}, n]]
data10 = Table[SampleStats[10], {j, 1, 1000000}];
Mean[data10]
Variance[data10]
data25 = Table[SampleStats[25], {j, 1, 1000000}];
Mean[data25]
Variance[data25]
data50 = Table[SampleStats[50], {j, 1, 1000000}];
Mean[data50]
Variance[data50]
```

This simulates $1,000,000$ rounds. For the $n = 10$ case, the mean mean was 50.491 and the mean variance was 81.7187 (where the central limit theorem predicted $\mu = 50.5$ and $\sigma^2 = 81.675$). For the $n = 25$ case, the mean mean was 50.4992 and a mean variance 32.6419 (where the central limit theorem predicted $\mu = 50.5$ and $\sigma^2 = 32.67$). For the $n = 50$ case, the mean mean was 50.4992 and the mean variance was 16.3082 (where the central limit theorem predicted $\mu = 50.5$ and $\sigma^2 = 16.335$).

These are quite close for our purposes, so it seems that applying the central limit theorem is reasonable, at least in the case where the underlying distribution is uniform[2]

3 Rational Players and j-Rational Players

3.1 Rational Players

If one were to assume that the entire player population consists of totally self-interested and totally rational players who fully evaluate the entire system, one can easily model the result.

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

Initially, we know that all guesses must lie in the interval $[1, 100]$ (the bounds established by the game), so the mean is also in this interval. $\frac{2}{3}$ the mean must lie in the interval $\left[\frac{2}{3}, 66.6667\right]$ 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 guesses lie in this interval, so $\frac{2}{3}$ of the mean must lie in $[0.666667, 0.770735]$. All rational players must minimize the distance between a valid guess and a number in this range, so all rational players choose 1, thus $I_{12} = [1, 1]$. For the next round, we see that because all rational players again choose 1, so the mean is 1. $\frac{2}{3}$ of the mean is $\frac{2}{3}$, so all rational players again choose 1.

In summary, we see that

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

[2]Which it is not.

A rational player's guess (under the assumption that all other players are rational) must lie in

$$\bigcap_{j=1}^{\infty} I_j = [1, 1]$$

Said otherwise: a rational player who assumes that all other players are rational will only choose 1. This is the Nash equilibrium [7] for this game, as any variation from this strategy would lead to a decrease in the expected payoff for any particular player.

3.2 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. If the player initially assumed that the population was naive players, then they would expect the mean to be 50.5, thus a rational player would choose $\frac{2}{3}$ of this value or approximately 33.6667. Under the assumption that all players made this starting assumption, the players should all choose 33.6667, whence the sample mean is now 33.6667 and a rational player would choose $\frac{2}{3}$ of 33.6667 or approximately 22.4444. Iteration again drives the only valid choice to 1 if this reasoning is repetitively applied.

This leads to a generalization of the fully rational player assuming that all other players are fully rational. We'll call players j-rational if they carry out the above analysis but terminated the analysis after j rounds. A player who assumed that the playing population was naive is a 1-rational player, and would choose $50.5\left(\frac{2}{3}\right)^1 \approx 33.6667$. A player who assumed that the playing population was 1-rational would choose $50.5\left(\frac{2}{3}\right)^2 \approx 22.4444$. In general a j-rational player would guess $50.5\left(\frac{2}{3}\right)^j$.

As a small abuse of this formula, we can also describe 0-rational players as players who correctly determine the behavior of naive players, but then fail to apply the $\frac{2}{3}$ scaling.

Due to the possible selection range, we see that after $\lceil \frac{-\log 50.5}{\log \frac{2}{3}} \rceil = 10$ rounds the optimal choice is 1. As such, 10-rational players are indistinguishable from rational players who assume that every player is also rational (who are effectively ∞-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 naive players in bulk; that is, if there is a sufficient population of naive 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.

4 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.

If all the players other than me collude, than any mean in the realm of possible answers is possible to attain, irrespective of the rationality of the players. 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.

5 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 here is that with small populations, each player's guess greatly 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 naive player and one 1-rational player playing this game (i.e., the player has categorized the other player correctly and chosen what should be the optimal strategy) . The naive 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 taken 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

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

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 (which works in the $n = \infty$ case) useless. Instead, a simple recurrence relationship can be be used.

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

Here is a table summarizing the guesses for various players for a variety of sizes of player population:

	Population (n)					
	2	5	10	25	50	∞
Spoiler	100					
Naive	Random Number in $[1, 100]$ (mean and median is ≈ 50.5)					
0-rational	50.5	50.5	50.5	50.5	50.5	50.5
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
4-rational	3.15625	7.24232	8.6248	9.43985	9.70843	9.97531
5-rational	1.57813	4.45681	5.54452	6.20702	6.42855	6.65021
6-rational	1	2.74266	3.56433	4.08133	4.25674	4.43347
7-rational	1	1.68779	2.29136	2.68361	2.81865	2.95565
8-rational	1	1.03864	1.47301	1.76457	1.86641	1.97043
9-rational	1	1	1	1.16026	1.23586	1.31362
10-rational	1	1	1	1	1	1
∞-rational	1	1	1	1	1	1

6 Combining the Distributions

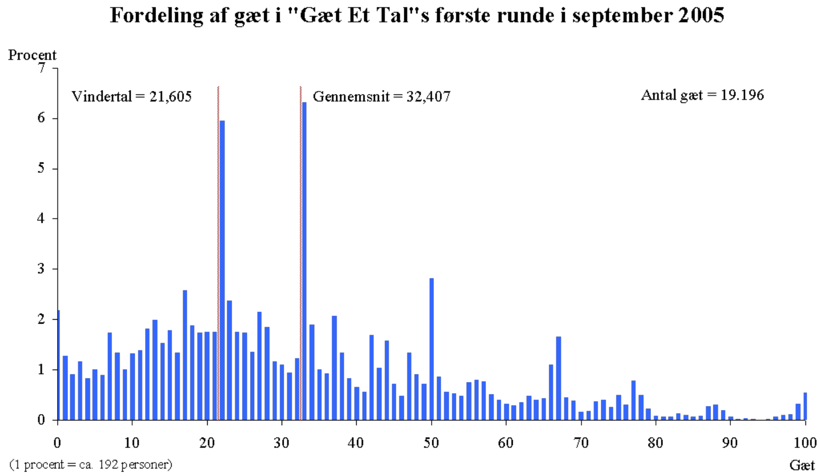
In some sense, this problem becomes reduced to two questions: "How many players are there" and "What sort of players are these?". By answering both of these questions, one can estimate median of the resulting distribution of the means, which leads to the guess.

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, but it is guessing. Prior studies in games similar to this exist. 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:

- I expect that the people in Jesse's social circle will tend to behave somewhat more rationally than the general college student population described by Nagel.
- 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 naively.
- 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 [6], which in turn provides a link to the Nagel paper [3], which is freely available and accessible through a quick Google search.

The Danish newspaper Politiken ran a similar contest (it is different in that it asks for values in $[0, 100]$) in 2005 [5] which involved 19,196 participants. The histogram for this contest was helpfully published by Department of Economics, University of Copenhagen [2].



I am particularly gratified by the notable spikes at the 0-, 1-, 2-, and ∞-rational values.

I first used the upper third of the graph (other than the region of 98.5 to 100) to estimate the total number of naive players, then reduced all the bins by a baseline percentage reflecting this random guessing. I then categorized users placing guesses in the first 10.5 percent as "rational" players, then deemed the regions about the expected (infinite player, $[0, 100]$) j-rational center as being j-rational. The last guesses of 98.5 to 100 were deemed as spoilers. This resulted in the following categorization:

Player	Percentage
Spoilers	0.6%
Naive	20.8%
0-rational	14.9%
1-rational	21.0%
2-rational	19.2%
3-rational	11.5%
∞-rational	12.0%

A quick simulation based on these percentages (for infinite players and $[0, 100]$ j-rational positions) gives a mean of 31.4233, which is an error of roughly 3% from the actual mean of 32.407.

I modify these findings based on a few factors:

- I expect that the people in Jesse's social circle will tend to behave somewhat more rationally than the average Dane.[3]
- I expect that the people in Jesse's social circle will tend to have a higher than expected likelihood of being a spoiler.[4]
- 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 were 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.

For fixed distribution of players, it is fairly easy to simulate the results. Here is a table that describes the resulting median of the means in the simulated distribution of players (for the $[1, 100]$ game)

[3] You may consider this a compliment

[4] It ain't all roses!

			j-Rational						
Population	Spoilers	Naive	0	1	2	3	∞	Mean	Guess
19196	0.6%	20.8%	14.9%	21.0%	19.2%	11.5%	12.0%	31.8512	21.2337
10	2%	10%	14%	22%	20%	14%	18%	32.3085	20.7697
25	2%	10%	14%	22%	20%	14%	18%	29.0506	19.1018
50	2%	10%	14%	22%	20%	14%	18%	28.1377	18.6317

The only non-fixed distribution used is a normal distribution associated with the naive players (which is symmetric) so the mean mean and the mean median are the same. I select the 25 player population somewhat arbitrarily, so my guess is 19.1018.

7 Limitations of the Analysis

We've highlighted only a few distinct styles of behavior. Most of these behaviors produce a fixed guess, which was done to simplify their later combination. Actual data suggests that people don't work this way, and will instead choose values that are intuitively weighted by these (or other!) distributions. People, as always, are complicated.

8 Retrospective

This contest was run between December 21 2009 and December 31 2009. The actual contest data is as follows:

1, 1, 1, 1, 1.1, 5.759, 7, 12, 18.13, 19.1018, 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 is 38.0871 so $\frac{2}{3}$ this sample mean was 25.3914.

A table showing the relevant centers for *j*-rational choices for this population size is this:

	Population (n)						
	2	5	10	25	39	50	∞
Spoiler	100						
Naive	Random Number in $[1, 100]$ (mean and median is ≈ 50.5)						
0-rational	50.5	50.5	50.5	50.5	50.5	50.5	50.5
1-rational	25.25	31.0769	32.4643	33.2055	33.4392	33.3739	33.6667
2-rational	12.625	19.1243	20.8699	21.8337	22.1422	22.0558	22.4444
3-rational	6.3125	11.7688	13.4164	14.3564	14.6617	14.5760	14.963
∞-rational	1	1	1	1	1	1	1

We can use this data to construct a description of the population. We'll first remove my guess of 19.1018 and then categorize the remaining guesses using the same analysis as applied to the Politiken contest to obtain a player breakdown. The last row of the following table summarizes the population and population breakdown (the other rows are the Politiken breakdown and my prior guess breakdown):

			j-Rational						
Population	Spoilers	Naive	0	1	2	3	∞	Mean	Guess
19196	0.6%	20.8%	14.9%	21.0%	19.2%	11.5%	12.0%	31.8512	21.2337
25	2%	10%	14%	22%	20%	14%	18%	29.0506	19.1018
39	2.2%	45.7%	10.4%	4.4%	21.6%	0.5%	15.2%	37.0946	24.7297

So, if I had correctly guessed the player breakdown and population in advance, I would have submitted the guess 24.7297, causing the new mean to be 38.2314. $\frac{2}{3}38.2314 \approx 25.4876$ which is a 3% error (and I would have won the contest).

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

- My guess of the playing population was close enough to not present any significant problem in generating the guess.

- 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 I 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 was a significant number of players choosing values near the 2-rational choice of 22, 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 of 14.3.
- My expectation that the people in Jesse's social circle would tend to behave somewhat more rationally than the other groups that I reviewed was really very much not correct. The largest group by far was the naive player group, suggesting that a very large proportion of the players arbitrarily selected their guess without much view toward the setup of the game. Indeed, this proportion was much larger than in either the Nagle or the Politiken games discussed earlier. This could be because the players were not sufficiently motivated by a (large in the case of Politiken) cash payment.

So, I 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, I 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).

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/o`, 2005.

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

[4] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill Book Company, 1953.

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

[6] 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].

[7] 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].