1 Experiment description

I ran the simulation on $G(n, 1/2)$ for $n \in \{3000, 10000\}$.
For $n = 3000$, I ran simulations for $c = 1, 2, 3, 4, 5$, with 50000 trials each.
For $n = 10000$, I ran one simulation for $c = 1$, with 10000 trials.

2 Summary

Based on the simulation result, we can hypothesize the following:

Conjecture 1 (Confidence: high)**.** *The number of Blues in Day 1 approximately follows a Gaussian distribution with mean* $n/2 - A(p)c\sqrt{n}$ *and variance* $B(p)^2 n$*, with* $A(.5) \approx .8$*,* $B(.5) \approx .5$*.*

Conjecture 2 (Confidence: low)**.** *The number of Blues in Day 2 follows a heavily sloped distribution, that is approximately exponential with mean* $n/6$ *for* $c = 1$ *and* $p = .5$*.*

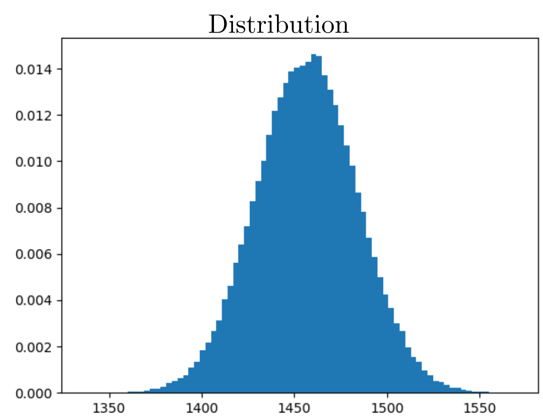
Conjecture 3 (Confidence: high)**.** *Let* $\Lambda_1 = \{\textit{Blue wins}\}$ *and* $\Lambda_2 = \{|B_1| > n/2\}$*. Then* $\Pr(\Lambda_1 \mid \Lambda_2)$ *and* $\Pr(\Lambda_2 \mid \Lambda_1)$ *are both* $1 - o(1)$*.*

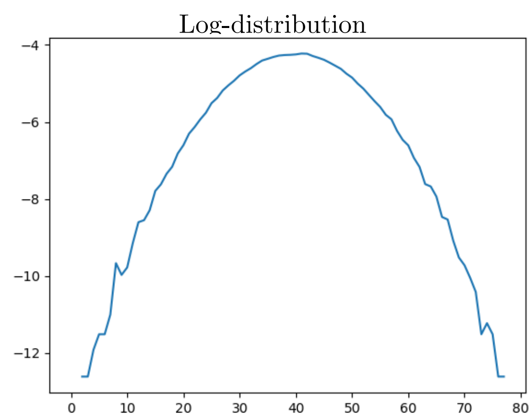
Remark. If Conjectures 1 and 3 are true, $\Pr(\Lambda_1) = \mathrm{O}(\exp(-D(p)c^2))$ for some $D(p)$.

3 Detailed Results

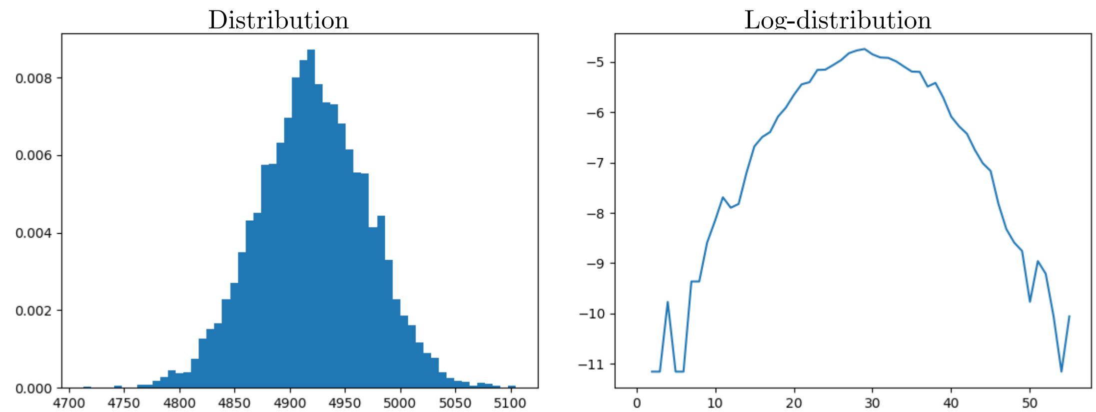
3.1 Number of Blues in Day 1

Below are the distribution graphs for $n = 3000, c = 1$. Note that the graph on the right is the logarithm of the densities of the one on the left.

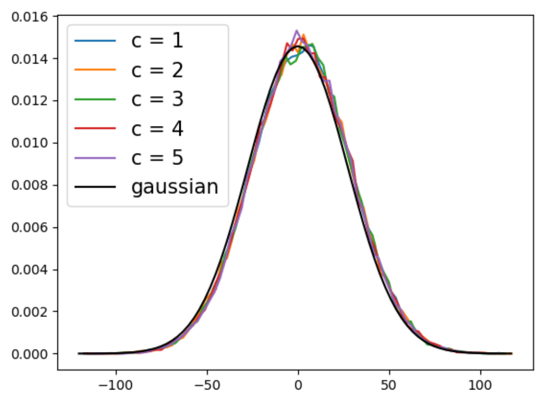




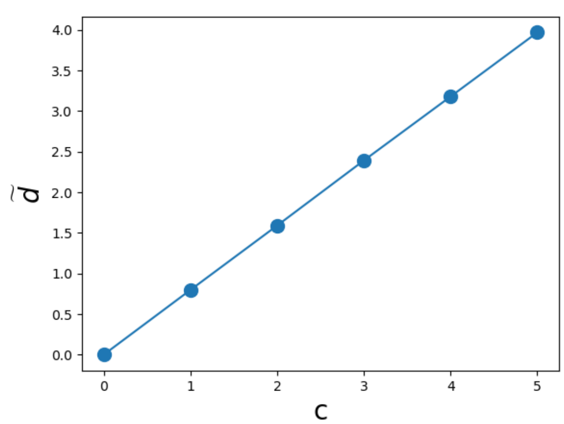
The distribution graphs for $n = 10000, c = 1$:



The distributions of $|B_1|$ for $c = 2, 3, 4, 5$ follow the same shape. In fact, I plotted the centered (subtracted by their respective means) distributions for $c = 1, 2, 3, 4, 5$, $n = 3000$, and compared them with the Gaussian distribution $N(0, .5\sqrt{n})$ and got the graph below.



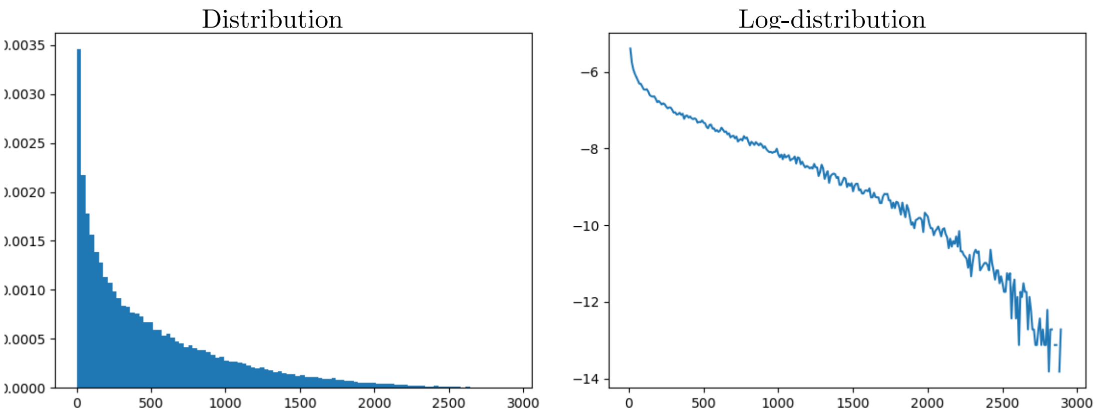
Now the means of these distributions can be written in the form $n/2 - d\sqrt{n}$, for some d mostly dependent on c. To find d empirically, we can take $\widetilde{d} = (n/2 - \overline{|B_1|})/\sqrt{n}$, where $\overline{|B_1|}$ is the sample mean of $|B_1|$ in the distribution. For $n = 3000$, we have $\widetilde{d} = 0.7973$, and for $n = 10000$, $\widetilde{d} = 0.7955$. This suggests d depends solely of c and p.
When $p = .5$, as in all our simulations, the following plots c against $\widetilde{d}$.



This indicates a strong linear relationship, $d \approx .8c$. Therefore, we hypothesize that the distribution of the number of Blues after Day 1 can be very well approximated by a Gaussian distribution, with means $n/2 - A(p)c\sqrt{n}$, and variance $B(p)\sqrt{n}$, with $A(.5) \approx .8$ and $B(.5) \approx .5$.

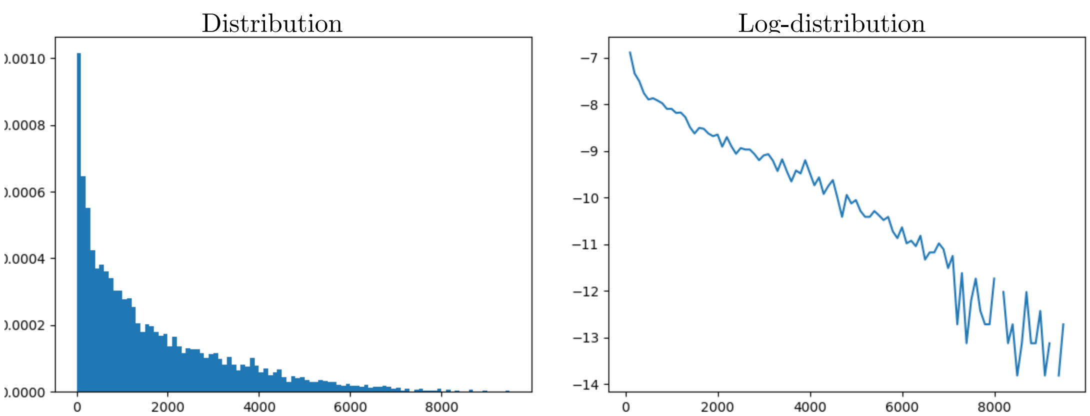
3.2 Number of Blues in Day 2

Intriguingly, the distribution for $|B_2|$ follows a heavily sloped shape. For $n = 3000$ and $c = 1$:



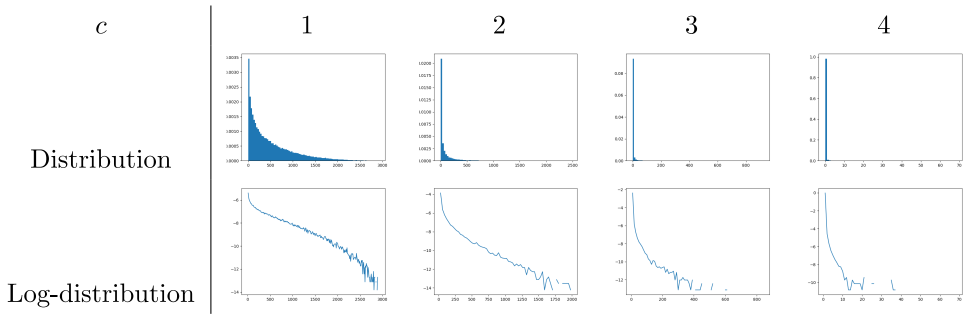
The means and standard deviation are close to 500, which is $n/6$. This, combined with the log-graph, suggests that $|B_2|$ approximately follows an exponential distribution with mean $n/6$ (or some constant near 6).

For $n = 10000$ and $c = 1$:



Again, the means and standard deviation are close to 1660, which is very near $n/6$. Thus the observation is consistent with the hypothesis above.

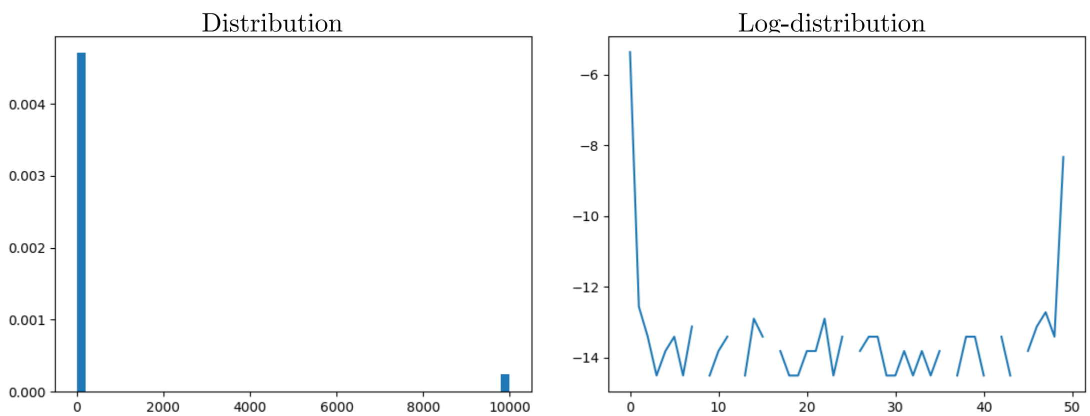
However, the distribution seems no longer approximately exponential for $c > 1$. These are the graphs for $c = 2, 3, 4$ and $n = 3000$:



As c grows, $|B_2|$ becomes increasingly concentrated near 0, at a rate even higher than exponential. I hypothesize from the shapes of the log-distribution graphs that the density can be of size $\Omega(\exp(-\lambda x^c))$, but with low confidence.

3.3 Blues in Day 3 and onwards

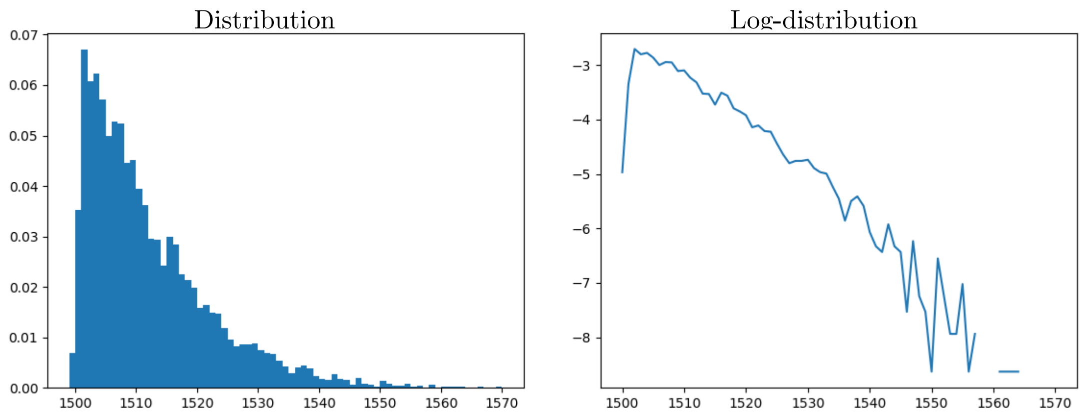
In overwhelmingly most cases, the result in Day 2 already indicate the winner. At such, the distributions from day 3 onwards are heavily bimodal, with two peaks near 0 and n. Below is an example for $n = 10000$ and $c = 1$.



3.4 Blues wins given Day 1 advantage

Another interesting quantity is the probability that Blue wins if they obtain an advantage after Day 1. Simulation results confirms this, with about .99 probability that Blue wins given Day 1 advantage, both for $n = 3000$ and $n = 10000$.

Conversely, the probability that Blue has had an advantage after Day 1 given that they win in the end is above .9. Below is the distribution for $|B_1|$ when Blue win, for $n = 3000$ and $c = 1$.



This shows that in most cases, when Blue achieve an advantage in Day 1, this advantage is mild, but it still allows Blue to win eventually.