The Variance of High Production Hexes in Catan.

G.F. Lima, Ph.D

July 19, 2022

1 Introduction

In Catan the observed dice rolls can sometimes be very different from the expected distribution, as has been shown in [1]. In order to give a more intuitive sense of this variance, we will focus on the particular case where, given two numbers with an equally high probability of occurring, one number rolls at least twice as many times as the other during a game. We will show that the chances of such occurrences are surprisingly high. Furthermore, we'll show that such issues can be mitigated by using balanced dice, as implemented in Colonist.io.

2 Motivation

A substantial component of Catan strategy is the placement of initial settlements. One of the most important features to optimize is production[1], which is measured by the number of pips on each hex and is in proportional to the probability of each number being rolled. The choices made by players are guided by the expected production of each possible initial placement, yet in practice the dice outcomes can differ significantly from what's expected. For instance, it is not uncommon to have two players with equally good initial set-ups but with different high production numbers: one player could have settlements on a 6 and on a 5, while the other has settlements on an 8 and on a 9. Both pairs have the same expected production, but we will show that the chances of one of these numbers rolling at least twice as many times as its counterpart during a game is roughly 37%, even though they have equal probability of occurring. In contrast, this will happen in only

[1]There are obviously other features to take into account, such as synergy between the resources (e.g. wood + brick, ore + wheat + sheep), expansion spots, ports, and resource scarcity – but these aren't the focus of this article.

8% of the games with the balanced dice featured in Colonist.io, and if we restrict ourselves to just the first half of a game, then the probability of such an imbalance goes as high as 66% with random dice, while the balanced dice remain at 8%.

3 Numerical Simulation

Given a sequence of dice rolls, and given two specific dice roll results m and n, both having the same probability of occurring (e.g. 6 and 8) we shall say that there's *high imbalance between* m *and* n if either n occurs at least twice as many times as m in the sequence, or m occurs at least twice as many times as n in the sequence.

First, we shall assume that the average Catan game has roughly 70 turns. This seems in line with our personal experience of over 1000 games and the data gathered in [2]. We simulated 1000000 (one million) games with 70 dice rolls for both random and balanced dice. Our original Python code can be found in [3].

With random dice, the probability of high imbalance between 6 and 8 was 0.172698, or roughly 17% and the probability of high imbalance between 6 and 8 or 5 and 9 was 0.365337, or roughly 37%. On the other hand, with balanced dice the probability of high imbalance between 6 and 8 was $0.027685 \cong 3\%$, and the probability of high imbalance between 6 and 8 or 5 and 9 was $0.079603 \cong 8\%$.

4 Analytical Solution

If we wanted the probability of a number being rolled k times in an average game of n turns, we'd use the probability mass function of a binomial distribution:

$$f(k, p, n) = \binom{n}{k} p^k (1-p)^{n-k}$$

So, for example, the probability of a 7 being rolled exactly 20 times in a game with 70 turns is

$$f(20, 70, 1/6) = \binom{70}{20} \times \left(\frac{1}{6}\right)^{20} \times \left(\frac{5}{6}\right)^{50}$$

since the probability of rolling a 7 is 1/6.

The derivation of this probability function isn't too hard. The term p^k represents the probability of k independent events with individual probability p, and the term $(1-p)^{n-k}$ corresponds to the other $n-k$ events which must have probability $1-p$. The binomial coefficient represents the different ways these different events can be scrambled around.

The binomial distribution is based on the assumption that each trial is a Bernoulli trial, i.e. there are only two possible outcomes, which are usually refereed to as *success* and *failure*. To tackle our problem theoretically, we need a modified version of the binomial distribution where instead of Bernoulli trials with only two possible outcomes, we have trials with *three* possible outcomes. Let's start with the case were we have three possible outcomes (e.g. an 8 rolls, a 6 rolls, or neither of these). Let's call the first outcome a *1-success*, the second outcome a *2-success*, and the third, a *failure*. The formula for having k_1 1-successes and k_2 2-successes in n trials is:

$$f(k_1, p_1, k_2, p_2, n) = \frac{n!}{k_1!k_2!(n-k_1-k_2)!} \quad p_1^{k_1} p_2^{k_2} (1-p_1-p_2)^{n-k_1-k_2}$$

We can further generalise to the case where we have m different types of successes, and so we have finite sequences $(k_i)_1^m = k_1, \ldots, k_m$ and $(p_i)_1^m = p_1, \ldots, p_m$, and a probability mass function

$$f((k_i)_1^m, (p_i)_1^m, n) = \frac{n!}{(\prod_1^m k_i!)(n - \sum_1^m k_i)!} \quad \left(\prod_1^m p_i^{k_i}\right)\left(1 - \sum_1^m p_i\right)^{(n-\sum_1^m k_i)}$$

4.1 Example:

Knowing that the probability of an 8 rolling is the same as the probability of a 6 rolling, which is 5/36, then the probability of an 8 rolling 10 times and a 6 rolling 20 times in a game with 70 turns is given by

$$f(10, 5/36, 20, 5/36, 70) = \frac{70!}{(10!)(20!)(40!)} \times \left(\frac{5}{36}\right)^{10} \times \left(\frac{5}{36}\right)^{20} \times \left(\frac{26}{36}\right)^{40}$$

Coming back to our original problem, we can now see that the probability of a 6 rolling at least twice as many times as an 8 in a game with 70 turns is given by

$$\sum_{(k_1,k_2)\in\mathbf{K}} f(k_1, p, k_2, p, n)$$

where $p = 5/36$, $n = 70$, and

$$\mathbf{K} = \{(k_1, k_2) \in \mathbb{N}^2 \mid k_1 \geq 2 \cdot k_2, \quad k_1 + k_2 \leq n, \quad k_1 > 0\}$$

By symmetry, this is the same probability of an 8 rolling at least twice as many times as a 6, and so the probability of either the 6's rolling at least twice as many times as the 8's, or the 8's rolling at least twice as many times as the 6's in a game with 70 turns is given by

$$2 \cdot \sum_{(k_1,k_2)\in\mathbf{K}} f(k_1, 5/36, k_2, 5/36, 70) \cong 0.1725$$

Similarly, the probability of an imbalance between the 5 and the 9 is given by the same formula except with $p = 4/36 = 1/9$:

$$2 \cdot \sum_{(k_1,k_2)\in\mathbf{K}} f(k_1, 1/9, k_2, 1/9, 70) \cong 0.2321$$

To calculate the probability of high imbalance between 6 and 8 or high imbalance between 5 and 9, we need to use the inclusion-exclusion principle, since the two cases have non-trivial intersection. In order to do this, we must first calculate the probability of imbalance between 6 and 8 *and* imbalance between 5 and 9, and this is

$$4 \cdot \sum_{(k_1,k_2,k_3,k_4)\in\mathbf{L}} f((k_i)_1^4, (p_i)_1^4, n) \cong 0.039$$

where $p_1 = p_2 = 5/36$, $p_3 = p_4 = 1/9$, $n = 70$ and

$$\mathbf{L} = \{(k_1, k_2, k_3, k_4) \mid \; k_1 \geq 2{\cdot}k_2, \;\; k_3 \geq 2{\cdot}k_4, \;\; k_1{+}k_2{+}k_3{+}k_4 \leq n, \;\; k_1 > 0, \;\; k_3 > 0\}.$$

We are multiplying by 4 since, if $(k_1, k_2, k_3, k_4) \in \mathbf{L}$ then (k_2, k_1, k_3, k_4), (k_1, k_2, k_4, k_3), and (k_2, k_1, k_4, k_3) also represent situations with the same type of imbalance.

Hence we conclude that the approximate probability of either high imbalance between 6 and 8 or high imbalance between 5 and 9 is approximately $0.1725 + 0.2321 - 0.0390 = 0.3656$.

And so we see that our simulation adequately approximates the analytical solution. The Python code used for the numerical calculations may be found in [3].

5 Summary of Results and Conclusion

The following table that summarizes and compares the balanced dice and the random dice. For each type of dice we used simulations of one million games.

		Probability of high imbalance between	
		6 and 8	6 and 8, or 5 and 9
Random	35 turns	37%	66%
	70 turns	17%	37%
Balanced	35 turns	3%	8%
	70 turns	3%	8%

Although the imbalance probability is much higher when there are fewer turns, the robber mechanism, if used correctly, can counter-balance early leads.

The actual effect of a high imbalance will clearly vary from game to game depending on how settlements and cities are distributed on the board. For instance, it can be drastically amplified, for better or worse, if one player has what is referred to as a *high variance* set up, with both initial settlements on the same high production number; and it can be somewhat mitigated if all players share the pairs where the imbalance occurred.

One can always argue that the cooperation between players is an integral part of the game since it can compensate any elements derived from chance and luck, and that high level play focuses on the careful administration of alliances and trades aimed against the player in the lead. Addressing these arguments would require an analysis of complete data from high level tournament games, which is something that seems hard to come by as of yet.

And so we leave it to the discretion of the reader to decide whether the final results of games with high imbalances can adequately measure the skill of the players involved, or whether games with balanced dice offer a more adequate alternative in that respect.

References

[1] Jeff d'Eon, *Designing Balanced Dice*, 2020.
Colonist Blog:
https://blog.colonist.io/designing-balanced-dice/

[2] Alex Cates, *Board Game Breakdown: Settlers of Catan, the basics*, 2021.
https://www.alexcates.com/post/board-game-breakdown-settlers-of-catan-the-basics

[3] https://github.com/GuilhermeFLima/GuilhermeFLima-Catan-Imbalance/blob/main/imbalance-simulation.py

[4] https://github.com/GuilhermeFLima/GuilhermeFLima-Catan-Imbalance/blob/main/imbalance-analytical.py