Wordle Optimization

Carson Buttars

2022-10-07

Intro

In this document, we will explore the question: what are the best methods for guessing a word in wordle? In particular, I want to know: * What word is the best to start out with? * What strategy of choosing a word to guess is the best (ex. lots of vowels, no vowels, common letters, etc.)

For this analysis, I will be assuming I know the exact ditionary that wordle uses. This is not realistic, but most likely we can say that the wordle dictionary is contained by the dictionary I will use. The information will still be mostly applicable, but the best results would be to follow the same analysis on the exact dictionary that wordle uses.

Best First Guess

One way of approaching this first part is to choose a word that has the most common letters in it. To start off, here is a list of the frequency of each letter in this dictionary.

```
letter.freq <- read.table("dict-letter-freq",sep = " ")
colnames(letter.freq) <- c("Frequency","Letter")
kable(letter.freq[17:26,])
```

	Frequency	Letter
17	990	d
18	1027	n
19	1266	t
20	1268	i
21	1305	l
22	1468	o
23	1488	r
24	1848	a
25	2443	e
26	2552	s

Then we can split up the words into tiers based on how many of the most common letters they have. When referencing this later, I will name the tiers as such: tier 1 will contain the most common letter, tier 2 will contain the 2 most common letter and so on. When going past tier 5, they will only be required to contain a combination of the most common letters. Another characteristic of the words in these guesses is that none of them will contain duplicate letters.

Having set up the word tier groups, I simulated the first guess of a wordle game and pulled out the results of how many greens and how many yellows there were for a random word in each tier.

Doing this test with 10^4 repetitions of each tier gives us the following result:

Tier	Yellow.Mean	Green.Mean	Score
1	0.8686	0.8686	2.6058
2	0.9907	0.9907	2.9721
3	1.0597	1.0597	3.1791
4	1.1820	1.1820	3.5460
5	1.5575	1.5575	4.6725
6	1.2357	1.2357	3.7071
7	1.1885	1.1885	3.5655
8	1.2042	1.2042	3.6126
9	1.1651	1.1651	3.4953
10	1.1378	1.1378	3.4134