Closed Form Solution for a Specific Case on the Friendship Paradox

Eric Ireland
Carthage College
eireland@carthage.edu

January 15, 2015

Abstract

It is a common feeling among people that their friends are more popular than them. With social networking sites like Facebook, many people find validation that, indeed, many of their friends have more "friends" than they do. This thesis examines this phenomenon, known as the Friendship Paradox, both in general and for a specific type of network. For the specific type of network, this thesis derives closed-form solutions for the average number of friends per person and for the average number of friends that each person's friends have for a group of n people. It is shown that the average number of friends of friends is greater than the average number of friends, which explains why many people do not feel as popular as their friends. Finally, other applications of this paradox are highlighted.

1 Do my friends have more friends than I do?

Many people grow up feeling that they do not have as many friends as everybody else does. This always used to be a general feeling, but today's social media allows people to quantify how many "friends" they and their peers have. According to the Washington Post, the average Facebook user has 245 friends, while the average friend on Facebook has 359 friends [3]. This may seem counterintuitive. In 1991, sociologist Scott Feld noticed this phenomenon and coined it the "friendship paradox" [1]. The goal of this thesis is to use mathematics to explore why the friendship paradox occurs and what implications it has.

2 Definitions and Development

To analyze the friendship paradox, there are two things we must consider in any network of friends: the average number of friends per person and the average number of friends of each person's friends. If most people perceive their friends to have more friends than they do, we would expect the average number of friends of friends to be greater than the average number of friends per person. This is what we will look for in our analysis.

2.1 General Case

We will first study the friendship paradox for any general network of friends. For any n individuals, let each person, i, be friends with x_i people. This makes the average number of friends in any group

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{1}$$

To find the average number of friends of friends, υ, we must notice that the number of times that each person's number of friends is counted is equal to the number of friends they have. For example, if somebody has six friends, six people will be friends with somebody who has six friends. Therefore a factor of 6^2 will go into the numerator of the average. In gereral, the numerator becomes $\sum_{i=1}^{n} x_i^2$. Instead of dividing by the total number of people, as with finding the average number of friends, we must divide by the total number of "friends", which is $\sum_{i=1}^{n} x_i$. Thus our expression for the average number of friends of friends is

$$\upsilon = (\sum_{i=1}^{n} x_i^2)/(\sum_{i=1}^{n} x_i). \tag{2}$$

This expression can be rewritten by using the definition of variance, σ^2. For a population variance with n people, we have

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = (\frac{1}{n}\sum_{i=1}^{n} x_i^2) - \mu^2. \tag{3}$$

We can rewrite this expression to solve for the numerator in Equation 2:

$$\sum_{i=1}^{n} x_i^2 = n(\sigma^2 + \mu^2). \tag{4}$$

Using Equations 1 and 4, we can rewrite Equation 2:

$$\begin{aligned} \upsilon &= (\sum_{i=1}^{n} x_i^2)/(\sum_{i=1}^{n} x_i) \\ &= \frac{n(\sigma^2 + \mu^2)}{n\mu} \\ &= \mu + \frac{\sigma^2}{\mu}. \end{aligned} \tag{5}$$

We recognize that both σ^2 and μ will be positive, so the average number of friends of friends, υ, will always be greater than or equal to the average number of friends, μ [1]. Therefore, in any network of friends, any individual is more likely to have fewer friends, on average, than his or her own friends do.

How can we get a feel for how many people will experience this in a network of friends? To answer this, let us look at a specific type of friendship network and find closed-form formula showing the average number of friends and a closed form formula showing the average number of friends of friends for a network of size n.

2.2 Specific Case

For a group of n people, let one person be friends with everybody else in the group and be labeled $n-1$. Then let next person be friends with everybody else in the group except one person; he or she is label $n-2$. Continue following this pattern until you reach the last person, who has only one friend and is labeled 1. The one person he or she is friends with must be the person who is friends with everybody in the group, $n-1$. We continue structuring the group until each person is assigned a number for how many friends he or she has.

To represent a network of friends, each person can be thought of as a vertex, and an edge connecting two vertices represents a friendship between two people. Graphs of friendship networks from $n = 1$ to $n = 7$ are shown in Figure 1. The number next to each vertex represents each person's number of friends. The numbers in parenthesis represent each person's average number of friends of friends.

There are a few important things to observe on the graphs in Figure 1. In graph theory, that every graph has two vertices with the same valence is a proven theorem. This can also be seen on our frienship networks.

Definition 2.1. For each graph, there are two people who have the same number of friends. These two people will be called **twins** and be labled as m.

For each group of friends represented in Figure 1, the average number of friends and average number of friends of friends is computed by hand. This is done by simply taking the average of the number labels for each person and by taking the average of the numbers in parenthesis, respectively. The results are shown in Table 1. We see that for groups of size $n > 2$, the average number of friends of friends is greater than the average number of friends, which is the behavior we expect to see.

In the next section, we will derive closed-form formulas to predict the average number of friends and the average number of friends of friends for a network of size n. We anticipate these formulas to yield matching results to those in Table 1.

3 Results

3.1 Mean Number of Friends with n People

To find the average number of friends with n people, one must sum each number of friends, 1 through $(n - 1)$, add the twin, and divide by the total number of people, n. With an even number of people, the twin takes on the value $n/2$, so we compute

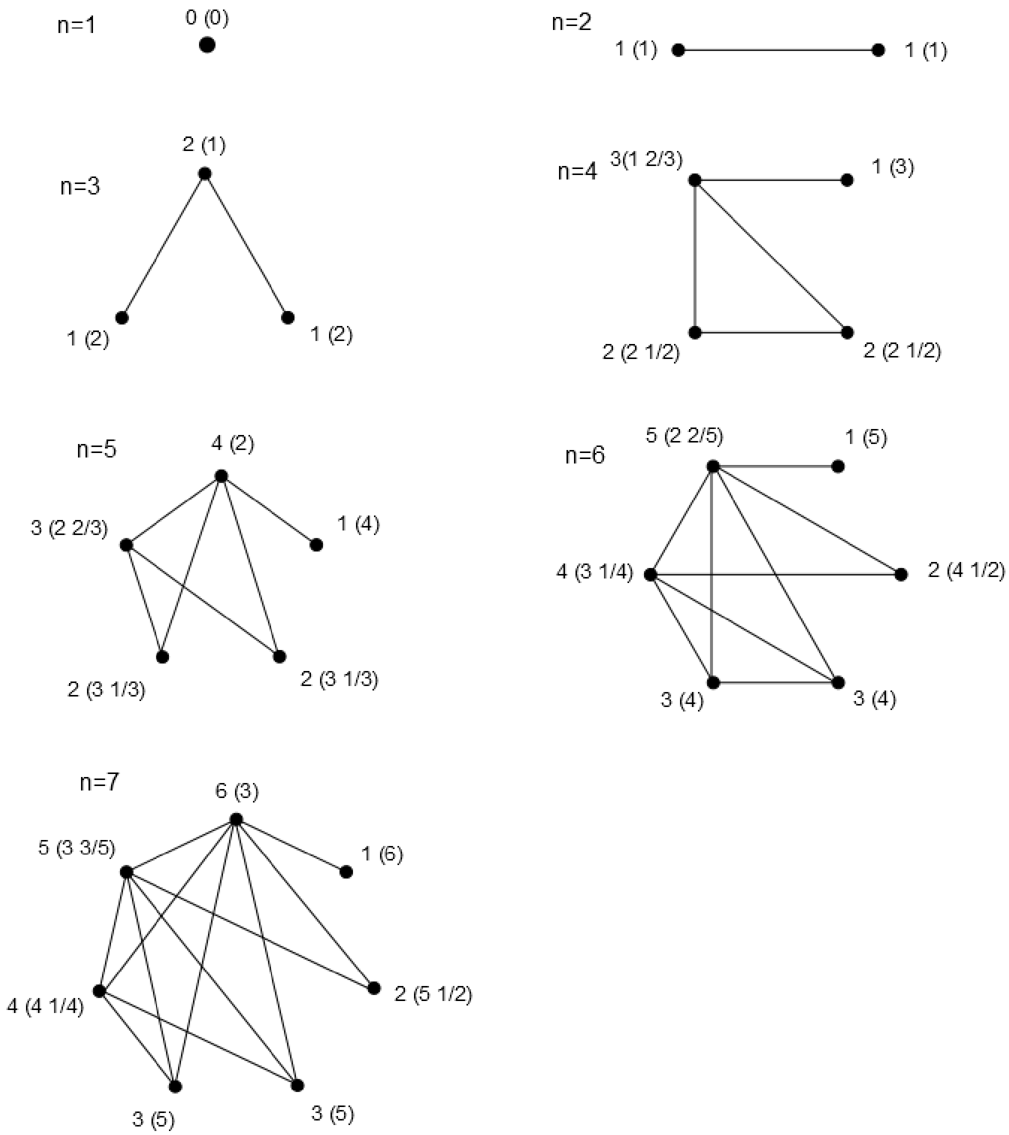

Figure 1: Graphs of friend networks of $n = 1$ to $n = 7$. The first number represents each person's number of friends, and the number in parenthasis represents each person's number of friends of friends.

n	avg. number of friends	avg. number of friends of friends
1	0	0
2	1	1
3	$1.\overline{3}$	$1.\overline{6}$
4	2	$2.41\overline{6}$
5	2.4	$3.1\overline{3}$
6	3	$3.891\overline{6}$
7	3.429...	4.621...

Table 1: Calculated average number of friends and average number of friends of friends for groups with 1 to 7 people.

$$\begin{aligned} \mu_{\mathbb{E}} &= \frac{\sum_{k=1}^{n-1} k + \frac{n}{2}}{n} \\ &= \frac{\frac{n(n-1)}{2} + \frac{n}{2}}{n} \\ &= \frac{n}{2}. \end{aligned} \tag{6}$$

With an odd number of people, the twin takes on the value $(n-1)/2$, so we compute

$$\begin{aligned} \mu_{\mathbb{O}} &= \frac{\sum_{k=1}^{n-1} k + \frac{n-1}{2}}{n} \\ &= \frac{\frac{n(n-1)}{2} + \frac{n-1}{2}}{n} \\ &= \frac{(n-1)(n+1)}{2n}. \end{aligned} \tag{7}$$

3.2 Mean Number of Friends of Friends with n People

To calculate the average number of friends of friends for a group with n people, we must separate the n people into little people and big people.

Definition 3.1. Little People are those whose number of friends fall within the range of 1 to m.

Definition 3.2. Big People are those whose number of friends fall within the range of m to $n-1$.

It is important to note that the twin term, m, appears in both of these definitions. This is because we will count one twin as a little person and one twin as a big person.

Again, the twin term is different for groups with an even number of people and groups with an odd amount of people: $m = \frac{n}{2}$ for even numbered groups and $m = \frac{n-1}{2}$ for odd numbered groups. Therefore, we shall find two closed form formulas for the average number of friends of friends: one for a network with an even number of people and one for a network with an odd amount of people.

3.2.1 Even Numbered Case

Calculating the average number of friends of friends will involve a double summation. The outer summation will count each person, k. For each person, the inner term will sum each of his or her friends, i, and divide by k to find his or her average number of friends of friends. Dividing the double summation by the total number of people, n, will result in the total average number of friends of friends for the entire network.

Due to the different characteristics of friendships for little people and big people, the double summation will be split into two: one summing the total friends of friends for the little people and one summing the total friends of friends for the big people. The little person term and the big person term will be combined and divided by the total number of people to find the average number of friends of friends. An example of this is shown in Figure 2.

The little people are only friends with the big people from $n-k$ to $n-1$, where i is the number for each friend. The little person term results as

$$\sum_{k=1}^{m} \left(\frac{\sum_{i=n-k}^{n-1} i}{k}\right). \tag{8}$$

Each big person is friends with every other big person except his or herself. Also, big person is friends with little people from $n-k$ to m. The big person term results as

$$\sum_{k=m}^{n-1} \left(\frac{\sum_{i=m}^{n-1}(i) - k + \sum_{i=n-k}^{m}(i)}{k}\right). \tag{9}$$

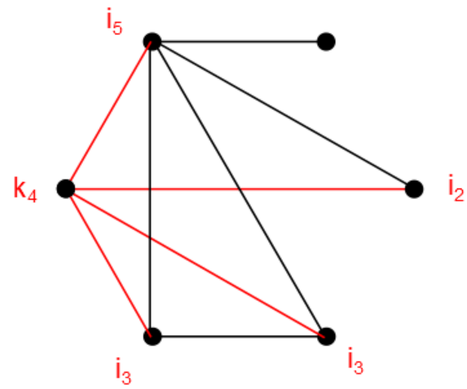

Figure 2: Graph illustrating how to find the average number of friends of friends. For each person, k, the sum of each of his or her friends, i, is taken and divided by k. For Person 4 in a network of six people, add $5 + 3 + 3 + 2$ and divide by 4 to find his or her average friends of friends to be 3.25. This is done for each person in the network, and the average of each person's friends of friends is found to be $3.891\overline{6}$.

Finally, when we combine the little people and big people term and divide by the total number of people, n, we have a term for the total average number of friends of friends in a group of n people, where $n \in \mathbb{E}$. This term is simplified by use of summation tricks and by substituting m for $n/2$:

$$
\begin{aligned}
\upsilon_{\mathbb{E}} &= \frac{\text{(little peoples' avg. friends of friends)} + \text{(big peoples' avg. friends of friends)}}{\text{total number of people}} \\
&= \frac{\sum_{k=1}^{m}(\frac{\sum_{i=n-k}^{n-1} i}{k}) + \sum_{k=m}^{n-1}(\frac{\sum_{i=m}^{n-1}(i)-k+\sum_{i=n-k}^{m}(i)}{k})}{n} \\
&= \frac{\sum_{k=1}^{m}(\frac{2n-k-1}{2}) + \sum_{k=m}^{n-1}(\frac{(n-m)(m+n-1)}{2k} + \frac{(n-k-m-1)(k-m-n)}{2k} - 1)}{n} \\
&= \frac{\frac{n(7n-6)}{16} + \sum_{k=m}^{n-1}(\frac{n+k(2n-3)-k^2}{2k})}{n} \\
&= \frac{\frac{n(7n-6)}{16} + \frac{5n(n-2)}{16} + \frac{2}{n}(\sum_{k=m}^{n-1}\frac{1}{k})}{n} \\
&= \frac{3n-4}{4} + \frac{1}{2}(\sum_{k=n/2}^{n-1}\frac{1}{k}). \qquad (10)
\end{aligned}
$$

3.2.2 Odd Numbered Case

For groups with an odd amount of people, the arguement is the same as with even, but with $m = \frac{(n-1)}{2}$. To derive the closed form formula, we shall start looking at the step before the substitution for m:

$$
\begin{aligned}
\upsilon_{\mathbb{O}} &= \frac{\text{(little peoples' avg. friends of friends)} + \text{(big peoples' avg. friends of friends)}}{\text{total number of people}} \\
&= \ldots \\
&= \frac{\sum_{k=1}^{m}(\frac{2n-k-1}{2}) + \sum_{k=m}^{n-1}(\frac{n+k(2n-3)-k^2}{2k})}{n} \\
&= \frac{(n-1)(7n-5)}{16} + \frac{\frac{(n-1)}{2}(\sum_{k=m}^{n-1}\frac{1}{k})}{n} + \frac{(n+1)(5n-9)}{16} \\
&= \frac{3n^2-4n-1}{4n} + \frac{(n-1)}{2n}(\sum_{k=\frac{(n-1)}{2}}^{n-1}\frac{1}{k}). \qquad (11)
\end{aligned}
$$

3.3 Comparing Mean Number of Friends and Mean Number of Friends of Friends

Now that we have derived the formulas for the average number of friends and the average number of friends of friends for a group of n people, we can compare them.

Starting with the even case, we can find the difference between the average number of friends of friends and the average number of friends by subtracting Equation 6 from Equation 10:

$$\begin{aligned}
\delta_{\mathbb{E}} &= \left[\frac{3n-4}{4} + \frac{1}{2}\left(\sum_{k=n/2}^{n-1} \frac{1}{k}\right)\right] - \left[\frac{n}{2}\right] \\
&= \frac{3n-4-2n}{4} + \frac{1}{2}\left(\sum_{k=n/2}^{n-1} \frac{1}{k}\right) \\
&= \frac{n-4}{4} + \frac{1}{2}\left(\sum_{k=n/2}^{n-1} \frac{1}{k}\right). \qquad (12)
\end{aligned}$$

Similarly with the odd case, we subtract Equation 7 from Equation 11 to obtain.

$$\begin{aligned}
\delta_{\mathbb{O}} &= \left[\frac{3n^2-4n-1}{4n} + \frac{(n-1)}{2n}\left(\sum_{k=\frac{(n-1)}{2}}^{n-1} \frac{1}{k}\right)\right] \\
&\quad - \left[\frac{(n-1)(n+1)}{2n}\right] \\
&= \frac{3n^2-4n-1-2n^2+2}{4n} + \frac{(n-1)}{2n}\left(\sum_{k=\frac{(n-1)}{2}}^{n-1} \frac{1}{k}\right) \\
&= \frac{n^2-4n+1}{4n} + \frac{(n-1)}{2n}\left(\sum_{k=\frac{(n-1)}{2}}^{n-1} \frac{1}{k}\right). \qquad (13)
\end{aligned}$$

Notice that both Equation 12 and Equation 13 are positive for $n > 2$. As a result, the average number of friends of friends is greater than the average number of friends for both even and odd sized groups. This is what we had seen in Table 1 when finding the values by hand, and this is what we expected to find at the beginning of our analysis, supporting the friendship paradox.

4 Should you worry about your friends having more friends than you do?

Now that we have seen the mathematics of why many people may feel their friends are more popular than them, we may ask if we should be worried. The answer is no, and this is because it is true for most people. In our specific friend network, we see that the majority of people in a network of three or more people have fewer friends than the average number of friends of friends. Though this may not be necessarily true in every general group of friends, because friend networks can be constructed any number of ways. Yet the friendship paradox is still prevalent, as we have seen in the Facebook example. If you think about it, an average student is more likely to be friends with someone who is friends with everybody on campus than to be friends with many people who have only one or two friends total. Therefore, the average number of friends of friends will be higher for many people because of a few very popular people. The point is that people should not worry if they feel their friends are more popular than them.

However, there are good reasons for being concerned with the friendship paradox. The same mathematics behind the friendship paradox is behind why a school's perceived average class size is bigger than a school's average class size. The average class size is simply the average number of students per class. The perceived average class size is the average class size as seen by the students. In a class of size 100, one-hundred students will expereince having a high class size average. Similarly, in a class of 6 students, only six students will experience having a lower class size average. As with the friendship paradox, there is a sampling bias in a school's perceived average class size.

The friendship paradox is also important in other fields, including vaccinations. If there are only ten Ebola vaccinations to distrubute in a village, doctors should not just distrubute ten random vaccinations. Instead, it is much more effective to ask ten random people who their friends are and then distrubute the vaccinations to their friends. In this approach, you are more likely to be giving vaccinations to people who are prone to spreading disease to more people [2]. So next time you find yourself on Facebook, be sure to remind your more popular friends to get their flu shots this year. It is for the sake of mathematics.

References

[1] Scott L. Feld. Why Your Friends Have More Friends Than You Do. *American Journal of Sociology*, 96(6):1464, May 1991.

[2] Katie McKissick. Social medias friendship paradox can be force for good, says USC professor, 2014.

[3] Hayley Tsukayama. Your Facebook Friends have More Friends than You, 2012.