

Web Science: Assignment #4

Alexander Nwala

Mohd. Nauman Siddique

Saturday, March 16, 2019

Contents

Problem 1	3
Problem 2	3
Problem 3	5

Problem 1

Determine if the friendship paradox holds for my Facebook account.* Compute the mean, standard deviation, and median of the number of friends that my friends have. Create a graph of the number of friends (y-axis) and the friends (x-axis) themselves, sorted by number of friends (y-axis). (The friends don't need to be labeled on the x-axis: just f1, f2, f3, ... fn.) Do include me in the graph and label me accordingly.

* = This used to be more interesting when you could more easily download your friend's friends data from Facebook. Facebook now requires each friend to approve this operation, effectively making it impossible. I will upload a csv file of my 2014 friends list on the #assignment-4 slack channel

SOLUTION

The solution to the problem is as below:

1. **Friend List** : The csv file for facebook friend list uploaded on slack has 99 entries. But the problem mentioned *acnwala* has 2014 friends. I have used 2014 as the friend count for *acnwala* in my graph.
2. **Graph**: Figure 1 shows the friendship paradox for *acnwala*'s Facebook account. The mode value is 266, mean value is 542 and standard deviation is 539. The number of friends following for *acnwala* is 2014. As, the mean and standard deviation values are very close, they appear as a single point in my graph.
3. **Friendship Paradox**: It does not hold as the mean list of friends is less than *acnwala*'s friend list count.

Problem 2

Determine if the friendship paradox holds for your Twitter account. Since Twitter is a directed graph, use "followers" as value you measure (i.e., "do your followers have more followers than you?").

Generate the same graph as in question #1, and calculate the same mean, standard deviation, and median values.

For the Twitter 1.1 API to help gather this data, see:

<https://developer.twitter.com/en/docs/accounts-and-users/follow-search-get-users/api-reference/get-followers-list>

If you do not have followers on Twitter (or don't have more than 50), then use my twitter account "acnwala".

SOLUTION

The solution to the problem is as below:

1. **Fetch Follower List** : I have fetched follower's list from Twitter API for *phonedude_mln*. Some of the followers had very large following list, so I have used log base 2 in my graphs to plot it.
2. **Graph**: Figure 2 shows the friendship paradox for *phonedude_mln*'s Twitter account for followers. The mode value is 0, mean value is 273 and standard deviation is 8. The number of Twitter followers for *phonedude_mln* is 925.
3. **Friendship Paradox**: It does not hold as the mean list of followers is less than *phonedude_mln*'s follower list count.

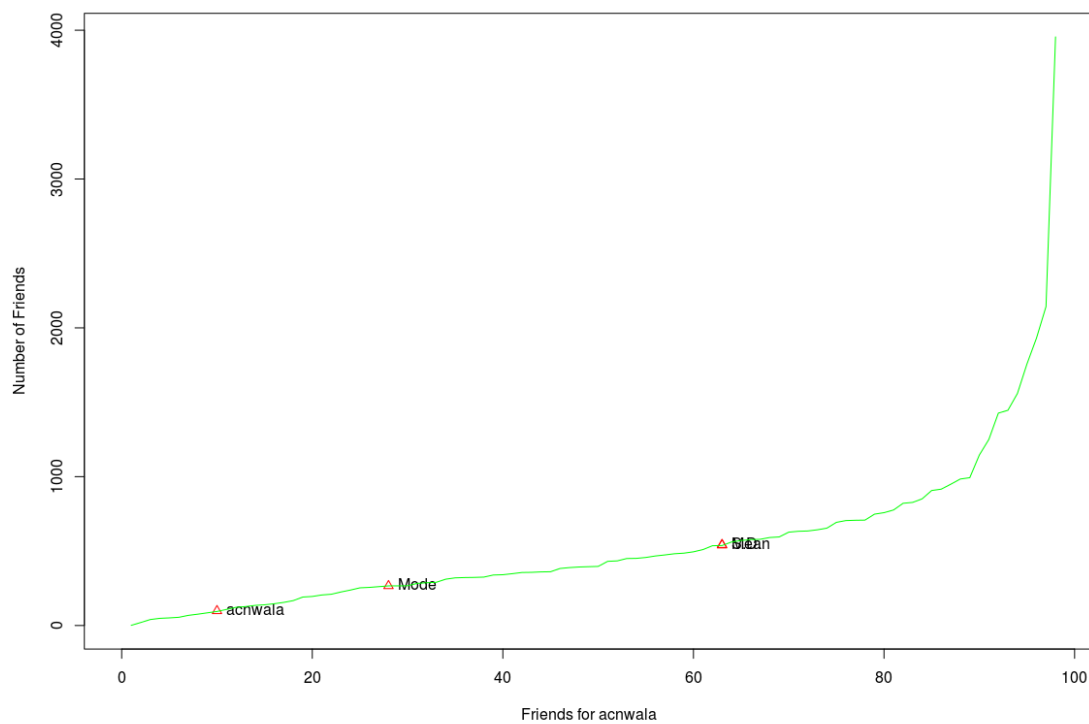


Figure 1: Friendship Paradox for friendship of acnwala's Facebook account

```

'''
Function to fetch Twitter followers list
'''

5
def fetch_twitter_followers():
    api = create_twitter_instance()
    file_object = open("TwitterData.csv", "w")
    fieldnames = ["Name", "Followers", "Friends"]
10
    writer = csv.DictWriter(file_object, fieldnames=fieldnames)
    writer.writeheader()
    response = api.GetFollowers(screen_name="phonedude_mln")
    for friend in response:

15
        insert_row = {"Name": friend.name, "Followers": math.log(friend.
            followers_count + 1, 2),
            "Friends": math.log(friend.friends_count + 1, 2)}
        writer.writerow(insert_row)
    file_object.close()

```

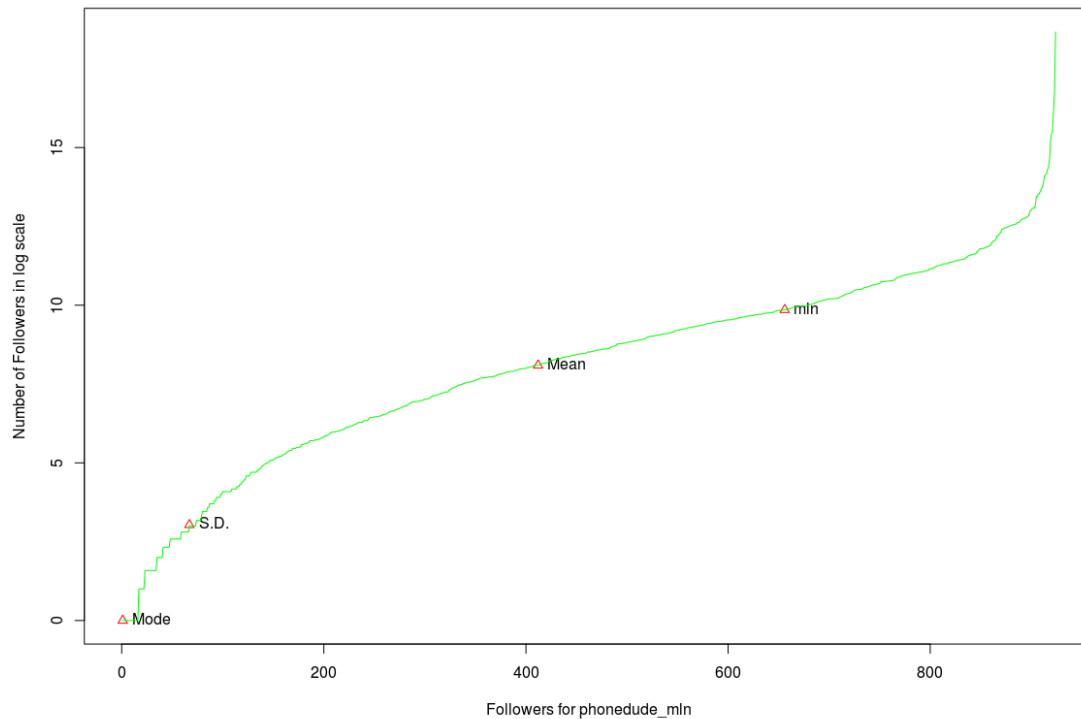


Figure 2: Friendship Paradox for followers of phonedude_mln's Twitter account

Problem 3

Repeat question #2, but change "followers" to "following"? In other words, are the people I am following following more people?

For the Twitter 1.1 API to help gather this data, see:

<https://developer.twitter.com/en/docs/accounts-and-users/follow-search-get-users/api-reference/get-friends-list>

SOLUTION

The solution to the problem is as below:

1. **Fetch Follower List** : I have fetched following's list from Twiter API for *phonedude_mln*. Some of the following had very large following list, so I have used log base 2 in my graphs to plot it.
2. **Graph**: Figure 3 shows the friendship paradox for phonedude_mln's Twitter account for following. The mode value is 1, mean value is 245 and standard deviation is 6. The number of Twitter following for phonedude_mln is 450.
3. **Friendship Paradox**: It does not hold as the mean list of following is less than phonedude_mln's following list count.

```
'''
Function to fetch Twitter following list
'''
```

```
5 def fetch_twitter_following():
    api = create_twitter_instance()
    file_object = open("TwitterDataFollowing.csv", "w")
    fieldnames = ["Name", "Followers", "Friends"]
10 writer = csv.DictWriter(file_object, fieldnames=fieldnames)
    writer.writeheader()
    response = api.GetFriends(screen_name="phonedude_mln")
    for friend in response:

15         insert_row = {"Name": friend.name, "Followers": math.log(friend.
            followers_count + 1, 2),
            "Friends": math.log(friend.friends_count + 1, 2)}
        writer.writerow(insert_row)
    file_object.close()
```

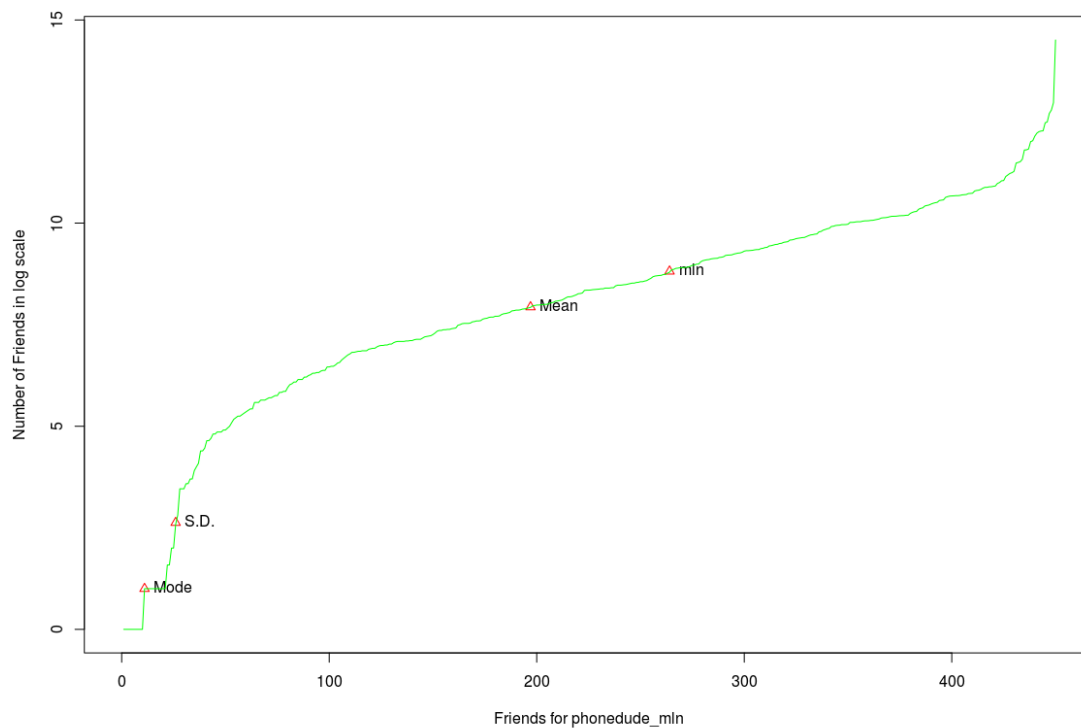


Figure 3: Friendship Paradox for following of phonedude_mln's Twitter account