

Project Milestone 2

Nilesh Mohan Dhikale (2001224), Ashokkumar Madanmohan Thakur (2051207)

01/04/2022

Introduction

This is the Milestone 2: ANALYSIS I in the Statistical Methods in Research (COSC 6323) course.

The purpose of the task is to check if the gender distribution in the study sample is representative of the gender distribution in the U.S. academia and draw figures 2-4 of the manuscript and express qualitative estimates about the behaviors of the variables in each figure.

Chi Squared Test

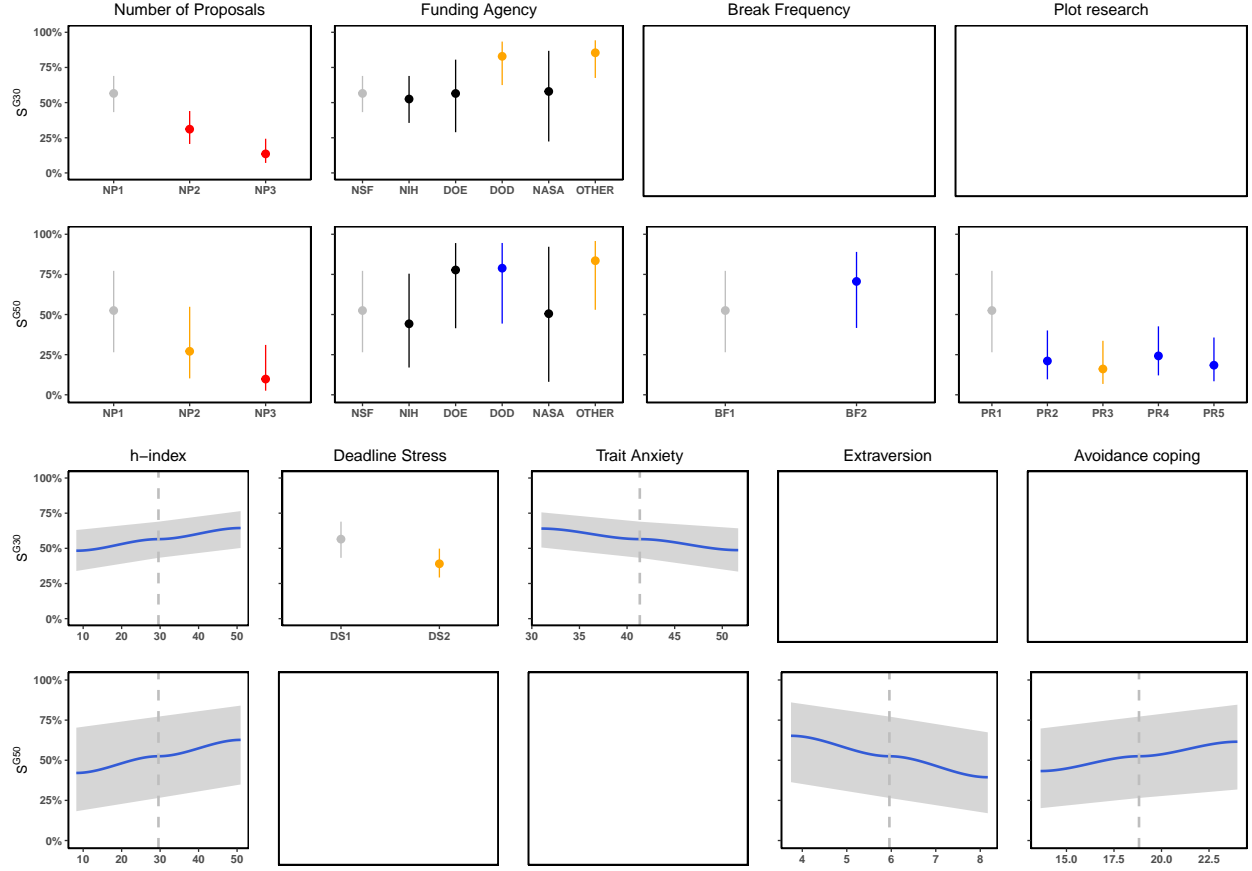
```
chisq.test(test_data, p = gender_proportion)
```

```
##  
## Chi-squared test for given probabilities  
##  
## data: test_data  
## X-squared = 1.2821, df = 1, p-value = 0.2575
```

Conclusion About Chi-Squared Test

1. From the result it can be observed that gender distribution of this project represent the gender distribution in the U.S. academia.
2. Here in the chi squared test the null hypothesis is the 2 variables are independent.
3. In the chi squared test we are passing the sample data and the proportions of female and male count mentioned in the U.S. academia.
4. The p-value is 0.2575 which is equal to the gender distribution in the U.S. academia.
5. p-value of 0.2575 depicts that the 2 variables are independent of each other.

Figure 2



Conclusion about figure 2

1. G 30 Number of proposals:

1. NP1 (1-2 proposals per year), NP2 (3-4 proposals per year), and NP3 (5 or more proposals per year).
2. From the definitions of the number of proposals it can be observed that as the count of the proposals per year increases the success rate decreases.

2. G 30 Funding Agency:

1. We know that the US funding agencies are highly competitive.
2. From the graph we can observe that the proposals which were submitted to DOD and Other agencies have higher success rate as compared to the remaining agencies which are NSF, NIH, DOE and NASA.
3. Agencies NSF, NIH, DOE, and NASA have around same success rate.
4. Confidence interval for NASA shows that the chances of success is more compared with remaining agencies.

3. Break Frequency and Plot Research:

1. No data is present for Break Frequency and Plot research.

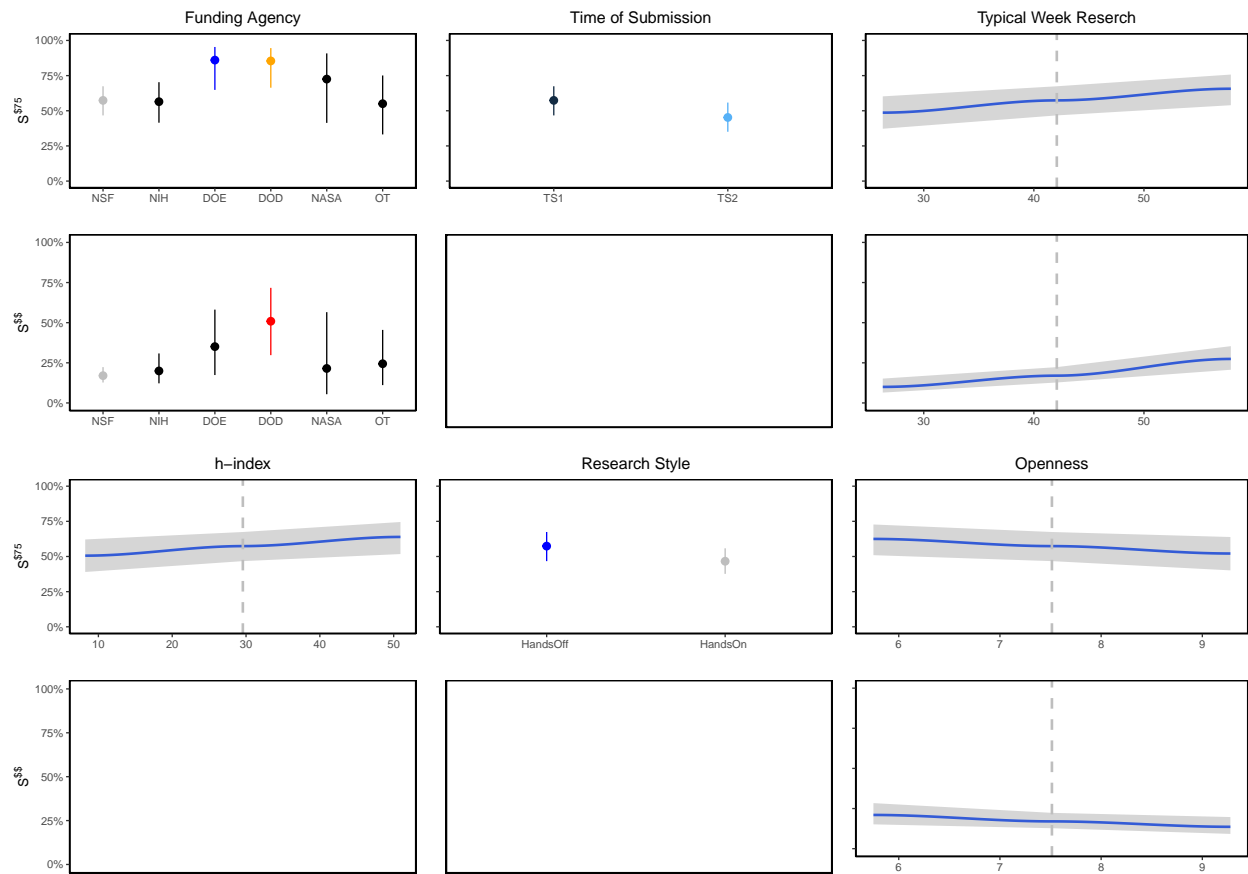
4. G 50 Number of Proposals:
 1. From the definitions of the number of proposals for G 50 it can be observed that as the count of the proposals per increases the success rate decreases.
5. G 50 Funding Agency: 1.From the graph we can observe that the proposals submitted to DOE, DOD and Other agencies have higher success rate as compared to NSF, NIH and NASA agencies.
6. G 50 Break Frequency:
 1. From the graph of we can observe that researches who frequently takes break every 1 - 2 hour has less success rate as compared to the researchers who takes break every 3 - 4 or longer hour.
7. G 50 Plot research:
 1. From the graph of plot research it can be observed that the researchers who gave less than a month for a pilot project has high success rate as compared to researchers who tend to spend more time on the pilot project.
 2. Though the others success rate is less that PR1, PR4 (6-12 months) has better chances of getting the grantmanship than other pilot research types.
8. G 30 h-index:
 1. The h-index graph shows linear relationship between the success rate and the grantmanship.
 2. Higher the h-index means higher the chances of getting a grant.
9. G 30 Deadline Stress:
 1. having more stress than the regular day does not help in getting grant.
 2. The researchers who have similar or less stress than the regular day intend to perform better and have higher success rate.
10. G 30 Trait Anxiety:
 1. Higher Trait Anxiety leads to lower work performance.
 2. As the Trait Anxiety increase the success rate of the researchers decreases.
11. G 30 Extraversion:
 1. No data present.
12. G 30 Avoidance coping:
 1. No data present.
13. G 50 h-index:
 1. The h-index graph shows linear relationship between the success rate and the grantmanship.
 2. Higher the h-index means higher the chances of getting a grant.
14. G 50 Deadline Stress:
 1. No data present.
15. G 50 Trait Anxiety:
 1. No data present.
16. Extraversion:

1. From the graph it can be observed that the less extrovert the researcher is the more success rate he achieves in getting the grantmanship.
17. G 50 Avoidance coping:
1. It is observed that the individuals who practice to avoid stress has better success rate than the one who deal with them.

Final observation about Figure 2

1. Most of the predictors are performing same for success rate of getting the grantmanship of 30 % (G30) and of 50 % (G50).
2. In case of US funding agencies success rate o getting funding from agencies is more in G50 (50 % or more).
3. The mean success rate of number of proposals for G30 and G50 are similar but the the one standard deviation above the mean of number of proposals in G50 increases the chance of success by almost 2 times than the one standard deviation above the mean of G30 number of proposals.
4. The personality traits such as extraversion, openness, agreeableness, etc impact the success rate irrespective of the funding type.

Figure 3



Conclusion about Figure 3

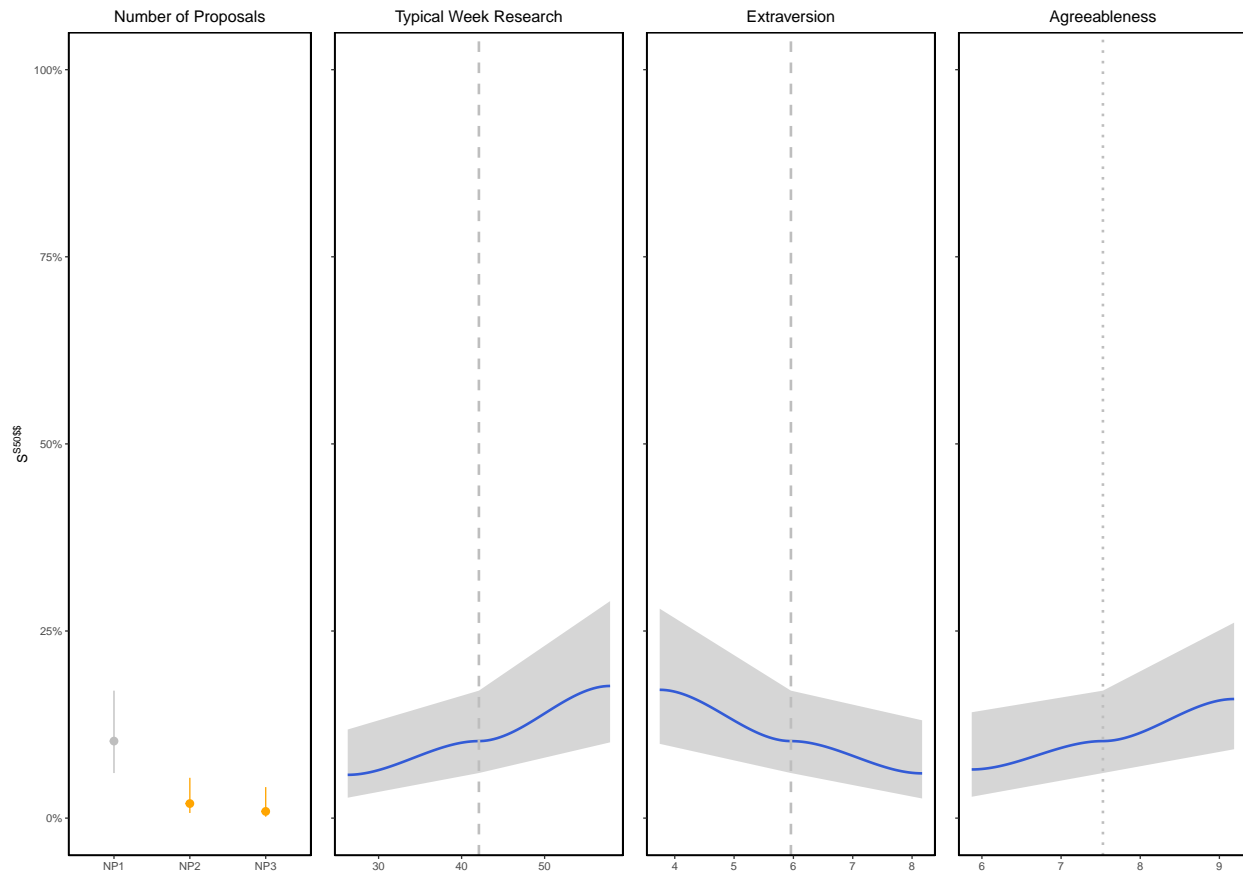
1. G \$ 75 Funding Agency:
 1. From the graph we can conclude that the success rate of getting the funding is more than 75 % when submitted to DOE and DOD.
2. G \$ 75 Time of Submission:
 1. TS1 (submission done of deadline day) has more success than submission day before deadline day.
3. G \$ 75 Typical Week Research:
 1. More time spent on research during the typical week produces more success rate.
4. G \$\$ Funding Agency:
 1. Submission done to DOD has more success compared to other remaining agencies.
 2. NSH has less success rate of granting full funding.
5. G \$\$ Time of Submission:
 1. No Data Present.
6. G \$\$ Typical Week Research.
 1. More time spent on research during the typical week produces more success rate.

7. G \$ 75 h-index:
 1. Higher the h-index means higher the chances of getting a grant.
8. G \$ 75 Research Style:
 1. Researcher who are involved in experiments have less success of getting funding of atleast 75 % as compared to researchers who are more involved in research management and document composition.
9. G \$ 75 Openness:
 1. The researcher who have high level of openness have less success rate as compared to the researchers who are less open.
10. G \$\$ h-index:
 1. No data present.
11. G \$\$ Research Style:
 1. No data present.
12. G \$\$ Openness:
 1. The researcher who have high level of openness have less success rate as compared to the researchers who are less open.

Final Observation about Figure 3

1. The success chance for one standard deviation above of the openness value remain similar for G75 while for full funding the one standard deviation above value decreases with the openness.
2. One standard deviation above the mean value of of getting full funding from the agency is nearly 1.5 times of its mean.
3. Similar to figure 2, the personality traits impact the success rate irrespective of the funding type.

Figure 4



Conclusion about Figure 4

1. G 50\$\$ Number of Proposals:

1. From the definitions of the number of proposals it can be observed that as the count of the proposals per year increases the success rate decreases.

2. G 50\$\$ Typical Week Research:

1. More time spent on research during the typical week produces more success rate.

3. G 50\$\$ Extraversion:

1. From the graph it can be observed that the less extrovert the researcher is the more success rate he achieves in getting the grantmanship.

4. G 50\$\$ Agreeableness:

1. Agreeableness is one of the Big-Five personality traits.
2. Researchers are more collaborative if they are high in agreeableness.
3. Success rate of researchers increases with increase in the agreeableness level.