Statistical Methods – COSC 6323 - Project-Milestone 3

By Dinesh Narlakanti (2083649) and Adarsh Chagantipati (2054031)

INTRODUCTION

A national wide survey of U.S.academics was conducted and around 400 intellectuals from 70 phD granting institutions, who excelled in the competitive grant process participated and filled a Core Questionnaire (CQ) which includes questions on their behavioral characteristics, proposal tactics, time they spend on research, sleep and other important things they do regularly. This survey and analysis of it concludes that long research hours and thoughtful choices are the primary reason for the grantsmanship and academic fame is the secondary reason for grantsmanship.

This document includes code development, analytic observations and visualization of Number of proposals, Typical Week Research, Extroversion, Agreeableness, Openness, Funding Agency, Break Frequency, Pilot Research, Deadline Stress, Trait Anxiety, Avoidance coping

Steps and its Description:

1. Installed required packages and libraries. And imported all_data and key_data.

2. For s30: assigned 0 to less than 30% success rate and 1 to greater than 30% success rate.

3. For s50: assigned 0 to less than 50% success rate and 1 to greater than 50%success rate.

4. For s75: assigned 0 to less than 75% funding and 1 to more than 75% funding.

5. For sdd: assigned 0 to for not fully funded participants and 1 to fully funded.

6. Changed P_AR of all_data to PR in key_data.

7. Changed all categorical predictors such as NP FA SA BF PR T RS WH AP and DWH to factors.

8. Performed below steps (9-17) for s30, s50, s75, sdd and s50dd tables:

9. Performed forward model starting with null model until it is optimized.

10. Performed step wise model until it is optimized.

11. Performed backward model starting with full model until it is optimized.

12. Created a data frame of the coefficients obtained from above optimized models.

13. Calculated and added extra column for the data frame as Prob-wise using formula: $\exp(\text{coefficient})/(1+\exp(\text{coefficient})$

14. Rearranged columns of the data frame.

15. Rounded P-values to 3 digits.

16. For clear interpretation, created a significance column.

17. Created a table of data frame using tab_df
18. Visualized all the above regression models using effect_plot and plot_model

Regression Tables: Full models and Optimized models with Inferences:

Parameter estimates for the Successful Grantsmanship configuration SG30:

Full model and its AIC

```
##
## Call:
## glm(formula = s30 ~ NASA + Rank + NP + FA + DS + BF + PR + RS +
##     T + WH + AP + DWH + TA + E + A + C + N + O + AC + EC + TC +
##     H + TWR + DWR, family = "binomial", data = key_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.1592  -0.7415  -0.4427   0.7638   2.4657
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.7936356  1.8352197   0.977 0.328400
## NASA        -0.0363082  0.0260688  -1.393 0.163686
## Rank        -0.0973309  0.2006971  -0.485 0.627702
## NP2         -1.0951598  0.2928104  -3.740 0.000184 ***
## NP3         -2.1286698  0.4019085  -5.296 1.18e-07 ***
## FA2         -0.3181979  0.3596569  -0.885 0.376304
## FA3         -0.2603226  0.6577224  -0.396 0.692257
## FA4          1.3550778  0.5392694   2.513 0.011978 *
## FA5          0.1808649  0.8467895   0.214 0.830868
## FA6          1.3711141  0.5266653   2.603 0.009231 **
## DS2         -0.4477155  0.3143819  -1.424 0.154414
## BF2          0.1680264  0.2833127   0.593 0.553129
## PR2         -0.7979975  0.5758867  -1.386 0.165843
## PR3         -1.2521984  0.5847115  -2.142 0.032228 *
## PR4         -0.5186082  0.5612248  -0.924 0.355453
## PR5         -0.7571301  0.5712412  -1.325 0.185034
## RS2         -0.4674060  0.2709413  -1.725 0.084506 .
## T2           0.4554560  0.2792687   1.631 0.102914
## WH2          0.4500462  0.5138322   0.876 0.381105
## AP2          0.3247123  0.3773930   0.860 0.389564
## DWH2        -0.6083856  0.4180471  -1.455 0.145585
## TA          -0.0425583  0.0232570  -1.830 0.067262 .
## E           -0.0221821  0.0635730  -0.349 0.727147
## A           -0.0290369  0.0825333  -0.352 0.724973
## C           -0.0083122  0.0876395  -0.095 0.924437
## N           -0.0244825  0.0909477  -0.269 0.787781
## O           -0.0263529  0.0755937  -0.349 0.727380
## AC           0.0287818  0.0268442   1.072 0.283639
## EC           0.0444714  0.0350957   1.267 0.205103
## TC           0.0009174  0.0385545   0.024 0.981017
```

```
## H              0.0179652  0.0076067   2.362 0.018188 *
## TWR            0.0091121  0.0097321   0.936 0.349121
## DWR            0.0018824  0.0080918   0.233 0.816048
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 495.86  on 402  degrees of freedom
## Residual deviance: 383.97  on 370  degrees of freedom
## AIC: 449.97
##
## Number of Fisher Scoring iterations: 5
```

Optimized -Model:

Predictor	Prob.wise	Odds.wise	Std..Error	z.value	Pr...z..	Sig
Intercept	0.743	1.062	0.617	1.721	0.085	
NP2	0.258	-1.058	0.273	-3.874	<0.001	***
NP3	0.108	-2.116	0.365	-5.805	<0.001	***
FA-NIH	0.460	-0.161	0.329	-0.489	0.625	
FA-DOE	0.500	-0.002	0.592	-0.003	0.998	
FA-DOD	0.789	1.316	0.504	2.611	0.009	**
FA-NASA	0.514	0.057	0.788	0.072	0.942	
FA-OT	0.819	1.507	0.491	3.072	0.002	**
TA	0.492	-0.030	0.013	-2.331	0.02	*
H	0.504	0.016	0.006	2.549	0.011	*
DS2	0.330	-0.709	0.265	-2.679	0.007	**

Inferences

Forward Model:

1. Starts with null model and its AIC is 424.93

2. Out of all the predictors, checked and kept only significant terms from the null model - NP FA TA H. These predictors has p-value less than 0.05.

3. DS is key predictor for grantsmanship. So, By adding DS, model became more optimized with AIC - 426.38

Step-wise:

1. Initial AIC is 424.93. Although we have low AIC value, we can non significant terms - DS RS TS DWH.

2. When removed all the non significant terms, AIC value increased to 431.53. To further improve the model, tried different combinations and found that, AIC value is again dropping when DS is added to the model.

3. So, with all the significant terms, AIC of the model is 426.38

Backward Model:

1. Starts with full model and AIC is 449.97.

2. Checked and removed non significant predictors with p-values greater than 0.05 and kept only significant terms - NP FA TA H DS.

3. With all the significant predictors whose p value is less than 0.05, model AIC value is 426.38

Parameter estimates for the Successful Grantsmanship configuration SG50:

Full model and its AIC

```
##
## Call:
## glm(formula = s50 ~ Rank + NP + FA + DS + BF + PR + RS + T +
##     WH + AP + DWH + TA + E + A + C + N + O + AC + EC + TC + H +
##     TWR + DWR, family = "binomial", data = key_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.9118  -0.5494  -0.2979  -0.1586   3.1284
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.306133   2.448887  -0.942  0.34634
## Rank        -0.324747   0.271228  -1.197  0.23118
## NP2         -1.116595   0.369663  -3.021  0.00252 **
## NP3         -2.548237   0.569867  -4.472 7.76e-06 ***
## FA2         -0.400243   0.490108  -0.817  0.41413
## FA3          1.230462   0.665609   1.849  0.06451 .
## FA4          1.435662   0.600636   2.390  0.01684 *
## FA5         -0.175086   1.216055  -0.144  0.88552
## FA6          1.542133   0.570116   2.705  0.00683 **
## DS2         -0.625169   0.415002  -1.506  0.13196
## BF2          0.806625   0.371939   2.169  0.03011 *
## PR2         -1.408038   0.670947  -2.099  0.03585 *
## PR3         -1.975914   0.695092  -2.843  0.00447 **
## PR4         -1.421577   0.637052  -2.231  0.02565 *
## PR5         -1.765445   0.663106  -2.662  0.00776 **
## RS2         -0.436989   0.340897  -1.282  0.19988
## T2           0.104342   0.358938   0.291  0.77128
```

4

```
## WH2          -0.292434   0.655418  -0.446   0.65547
## AP2          -0.286132   0.466151  -0.614   0.53934
## DWH2          0.100140   0.502524   0.199   0.84205
## TA           -0.028758   0.029186  -0.985   0.32446
## E            -0.283623   0.088268  -3.213   0.00131 **
## A             0.087554   0.106758   0.820   0.41215
## C             0.089419   0.114355   0.782   0.43425
## N             0.040810   0.117571   0.347   0.72851
## O             0.096778   0.100171   0.966   0.33398
## AC            0.060478   0.035233   1.716   0.08607 .
## EC            0.036320   0.045059   0.806   0.42021
## TC            0.069144   0.052485   1.317   0.18770
## H             0.021026   0.008914   2.359   0.01834 *
## TWR           0.008458   0.012330   0.686   0.49270
## DWR          -0.006199   0.010533  -0.589   0.55616
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 352.78  on 402  degrees of freedom
## Residual deviance: 260.13  on 371  degrees of freedom
## AIC: 324.13
##
## Number of Fisher Scoring iterations: 6
```

Optimized -Model:

Predictor	Prob.wise	Odds.wise	Std..Error	z.value	Pr...z..	.
Intercept	0.401	-0.401	0.866	-0.463	0.643	
NP2	0.252	-1.090	0.350	-3.114	0.002	**
NP3	0.089	-2.320	0.531	-4.371	<0.001	***
FA-NIH	0.418	-0.332	0.465	-0.713	0.476	
FA-DOE	0.760	1.151	0.637	1.806	0.071	
FA-DOD	0.772	1.217	0.571	2.133	0.033	*
FA-NASA	0.480	-0.079	1.144	-0.069	0.945	
FA-OT	0.821	1.524	0.523	2.915	0.004	**
BF2	0.685	0.777	0.331	2.349	0.019	*
PR2	0.194	-1.421	0.633	-2.244	0.025	*
PR3	0.148	-1.752	0.656	-2.670	0.008	**
PR4	0.224	-1.240	0.601	-2.065	0.039	*
PR5	0.170	-1.588	0.618	-2.569	0.01	*
E	0.440	-0.240	0.081	-2.974	0.003	**
AC	0.518	0.072	0.034	2.096	0.036	*
H	0.505	0.020	0.007	2.659	0.008	**
DS2	0.349	-0.624	0.331	-1.885	0.059	

Inferences:

Forward Model:

1. Starts with null model and its AIC value is 304.31

2. Predictors with p-value less than 0.05 are - NP FA BF PR E H AC and DS.

3. By adding only significant predictors, whose p-value is less than 0.05, model is now optimized with AIC - 305.07.

Step-wise:

1. Initial AIC value is 304.31 but has TC as non significant predictor.

2. After removing, AIC value is now 305.07 with all significant terms.

Backward Model:

1. Starts with full model and its AIC is 324.13 but with TC as non significant term.

2. When removed, model got better with all significant terms and AIC- 305.07

Parameter estimates for the Well-Funded Configuration S$75:

Full model and its AIC

```
##
## Call:
## glm(formula = s75 ~ Rank + NP + FA + DS + BF + PR + RS + T +
##     WH + AP + DWH + TA + E + A + C + N + O + AC + EC + TC + H +
##     TWR + DWR, family = "binomial", data = key_data)
##
## Deviance Residuals:
##     Min       1Q    Median       3Q       Max
## -2.0881  -1.0111   -0.5775   1.0879    2.0062
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.403398   1.598462    0.878  0.37996
## Rank          0.025045   0.165220    0.152  0.87951
## NP2          -0.110061   0.266145   -0.414  0.67921
## NP3           0.176838   0.303414    0.583  0.56001
## FA2           0.055746   0.303887    0.183  0.85445
## FA3           1.763590   0.635380    2.776  0.00551 **
## FA4           1.364079   0.554075    2.462  0.01382 *
## FA5           0.488377   0.680891    0.717  0.47321
## FA6           0.038805   0.457997    0.085  0.93248
## DS2          -0.327942   0.273328   -1.200  0.23021
## BF2           0.181925   0.234682    0.775  0.43822
## PR2           0.001016   0.533582    0.002  0.99848
## PR3          -0.231709   0.535071   -0.433  0.66498
## PR4          -0.163293   0.525201   -0.311  0.75587
## PR5          -0.272432   0.535774   -0.508  0.61111
## RS2          -0.395857   0.228997   -1.729  0.08387 .
## T2           -0.503384   0.236599   -2.128  0.03337 *
## WH2           0.442343   0.448809    0.986  0.32433
## AP2          -0.380545   0.301121   -1.264  0.20632
## DWH2          0.511737   0.393741    1.300  0.19371
## TA           -0.022678   0.019166   -1.183  0.23671
## E            -0.098092   0.053308   -1.840  0.06575 .
## A            -0.037005   0.070238   -0.527  0.59830
## C             0.047565   0.074142    0.642  0.52117
## N            -0.032103   0.076861   -0.418  0.67619
## O            -0.095171   0.063832   -1.491  0.13597
## AC           -0.006920   0.023187   -0.298  0.76537
## EC            0.011627   0.029262    0.397  0.69111
## TC           -0.024221   0.032504   -0.745  0.45617
## H             0.010217   0.006695    1.526  0.12704
## TWR           0.017130   0.008453    2.027  0.04271 *
```

```
## DWR            0.006436   0.007156   0.899  0.36841
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 557.96  on 402  degrees of freedom
## Residual deviance: 493.53  on 371  degrees of freedom
## AIC: 557.53
##
## Number of Fisher Scoring iterations: 4
```

Optimized -Model:

Predictor	Prob.wise	Odds.wise	Std..Error	z.value	Pr...z..	Sig
Intercept	0.472	-0.111	0.575	-0.192	0.847	
FA-NIH	0.491	-0.038	0.284	-0.133	0.894	
FA-DOE	0.821	1.520	0.596	2.551	0.011	*
FA-DOD	0.813	1.472	0.540	2.725	0.006	**
FA-NASA	0.662	0.673	0.650	1.036	0.3	
FA-OT	0.476	-0.096	0.440	-0.218	0.827	
RS2	0.393	-0.433	0.218	-1.982	0.047	*
TS2	0.380	-0.488	0.217	-2.247	0.025	*
O	0.470	-0.121	0.061	-1.973	0.049	*
H	0.503	0.013	0.005	2.340	0.019	*
TWR	0.506	0.022	0.007	3.212	0.001	**

Inferences:

Forward Model:

1. Starts with null model and AIC is 530.29

2. From the null model, kept significant terms which are less than 0.05- FA RS TS WH E TWR.

3. Model's AIC is 536.2 with non significant terms like WH. Model is still not optimized.

4. As there are still non significant term E, We are considering all other predictors except E and checking AIC.
5. AIC is now 535.55

6. PR - 542.75, TA - 535.02, 535.86. Out of all combinations, adding H(as H is the key predictor of grantsmanship) significant term, AIC if further dropped to 531.61

Step-wise:

1. Initial model AIC is 530.29 with non significant terms.

2. After removing all the non significant that are significantly greater than 0.05 and keeping only significant terms, AIC value slightly increased to 530.56

3. E and O are slightly greater than 0.05. When removed O, we are getting AIC as 531.92

4. As E is slightly greater than 0.05 and evidence shows that many scholars are introverted, we removed E and optimized model.

5. Final AIC of the model is 531.61 with all significant predictors.

Backward Model:

1. Starts with full model and its AIC 557.53

2. We got AIC as 530.56 when removed predictors with p-value significantly greater than 0.05.

3. As we have two predictors with slightly greater than 0.05, when both are removed AIC is increasing to 533.54

4. Out of two, When O is removed, AIC is again decreased to 531.92. But when E is removed, we are getting optimized model with AIC as 531.61 with all significant predictors.

Parameter estimates for the Fully-Funded Configuration S$$:

Full model and its AIC

```
##
## Call:
## glm(formula = sdd ~ Rank + NP + FA + DS + BF + PR + RS + T +
##     WH + AP + DWH + TA + E + A + C + N + O + AC + EC + TC + H +
##     TWR + DWR, family = "binomial", data = key_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.5540  -0.7204  -0.5115  -0.2712   2.5264
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.0820642  1.9768864  -1.559  0.11898
## Rank        -0.0359015  0.1962219  -0.183  0.85483
## NP2         -0.1262573  0.3279321  -0.385  0.70023
## NP3         -0.2112956  0.3584020  -0.590  0.55549
## FA2          0.2667863  0.3744239   0.713  0.47614
## FA3          1.1334925  0.5659087   2.003  0.04518 *
## FA4          1.6776504  0.5222752   3.212  0.00132 **
## FA5          0.4684546  0.8652511   0.541  0.58823
## FA6          0.5479402  0.5417066   1.012  0.31177
## DS2          0.0276498  0.3290662   0.084  0.93304
```

```
## BF2          0.0246542  0.2819711   0.087  0.93033
## PR2         -0.7218933  0.5603234  -1.288  0.19762
## PR3         -1.3100723  0.5812129  -2.254  0.02419 *
## PR4         -1.2087963  0.5621797  -2.150  0.03154 *
## PR5         -1.0817180  0.5668288  -1.908  0.05634 .
## RS2         -0.4309997  0.2715924  -1.587  0.11253
## T2          -0.3512791  0.2852250  -1.232  0.21810
## WH2          0.2133946  0.5905332   0.361  0.71783
## AP2          0.2995920  0.3621346   0.827  0.40807
## DWH2         0.2541881  0.4548273   0.559  0.57625
## TA           0.0154305  0.0229749   0.672  0.50182
## E           -0.0513438  0.0631170  -0.813  0.41595
## A            0.0013767  0.0828273   0.017  0.98674
## C            0.0330094  0.0933433   0.354  0.72361
## N           -0.1284522  0.0940680  -1.366  0.17209
## O           -0.1366954  0.0782272  -1.747  0.08057 .
## AC          -0.0249041  0.0277069  -0.899  0.36874
## EC           0.0138773  0.0347904   0.399  0.68998
## TC           0.0641519  0.0415181   1.545  0.12231
## H            0.0024705  0.0069060   0.358  0.72054
## TWR          0.0421974  0.0104089   4.054 5.04e-05 ***
## DWR         -0.0002902  0.0087652  -0.033  0.97359
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 430.53  on 402  degrees of freedom
## Residual deviance: 370.39  on 371  degrees of freedom
## AIC: 434.39
##
## Number of Fisher Scoring iterations: 5
```

Optimized -Model:

Predictor	Prob.wise	Odds.wise	Std..Error	z.value	Pr...z..	NA.
Intercept	0.109	-2.103	0.654	-3.217	0.001	**
FA-NIH	0.549	0.196	0.335	0.585	0.559	
FA-DOE	0.725	0.970	0.507	1.914	0.056	
FA-DOD	0.835	1.622	0.489	3.318	0.001	**
FA-NASA	0.572	0.290	0.816	0.355	0.723	
FA-OT	0.612	0.456	0.514	0.887	0.375	
O	0.463	-0.148	0.072	-2.044	0.041	*
TWR	0.510	0.039	0.008	4.616	<0.001	***

Inferences:

Forward Model:

1. Starts with null model and its AIC is 404.88

2. Significant predictors found are: FA PR O TWR. With these predictors, model AIC is 408.64.

3. As PR can be burden and often misused, we should not add PR for the forward model, surprisingly model is more optimized with AIC decreased to 406.73

Step-wise:

1. Initial mode aic is 404.88 with many non significant predictors.

2. Removing them made model optimized with AIC 406.73 with only significant predictors.

Backward Model:

1. Starts with full model and its AIC is 434.39 and non significant predictors - TC.

2. After removing it, model is optimized with AIC 406.73

Parameter estimates for the Combined Grantsmanship and Grant Funding Configuration SG50$$:

Full model and its AIC

```
##
## Call:
## glm(formula = s50dd ~ NP + RS + E + A + O + TWR, family = "binomial",
##     data = key_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.2788  -0.3312  -0.1975  -0.1091   2.9913
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.99823    1.59404  -1.881 0.059985 .
## NP2         -1.90548    0.59197  -3.219 0.001287 **
## NP3         -2.82918    0.82248  -3.440 0.000582 ***
## RS2         -0.75774    0.46237  -1.639 0.101250
## E           -0.26825    0.11286  -2.377 0.017463 *
## A            0.34244    0.15073   2.272 0.023098 *
## O           -0.21291    0.13171  -1.616 0.105986
## TWR          0.04454    0.01386   3.214 0.001308 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##      Null deviance: 187.42  on 402  degrees of freedom
## Residual deviance: 143.83  on 395  degrees of freedom
## AIC: 159.83
##
## Number of Fisher Scoring iterations: 7
```

Optimized -Model:

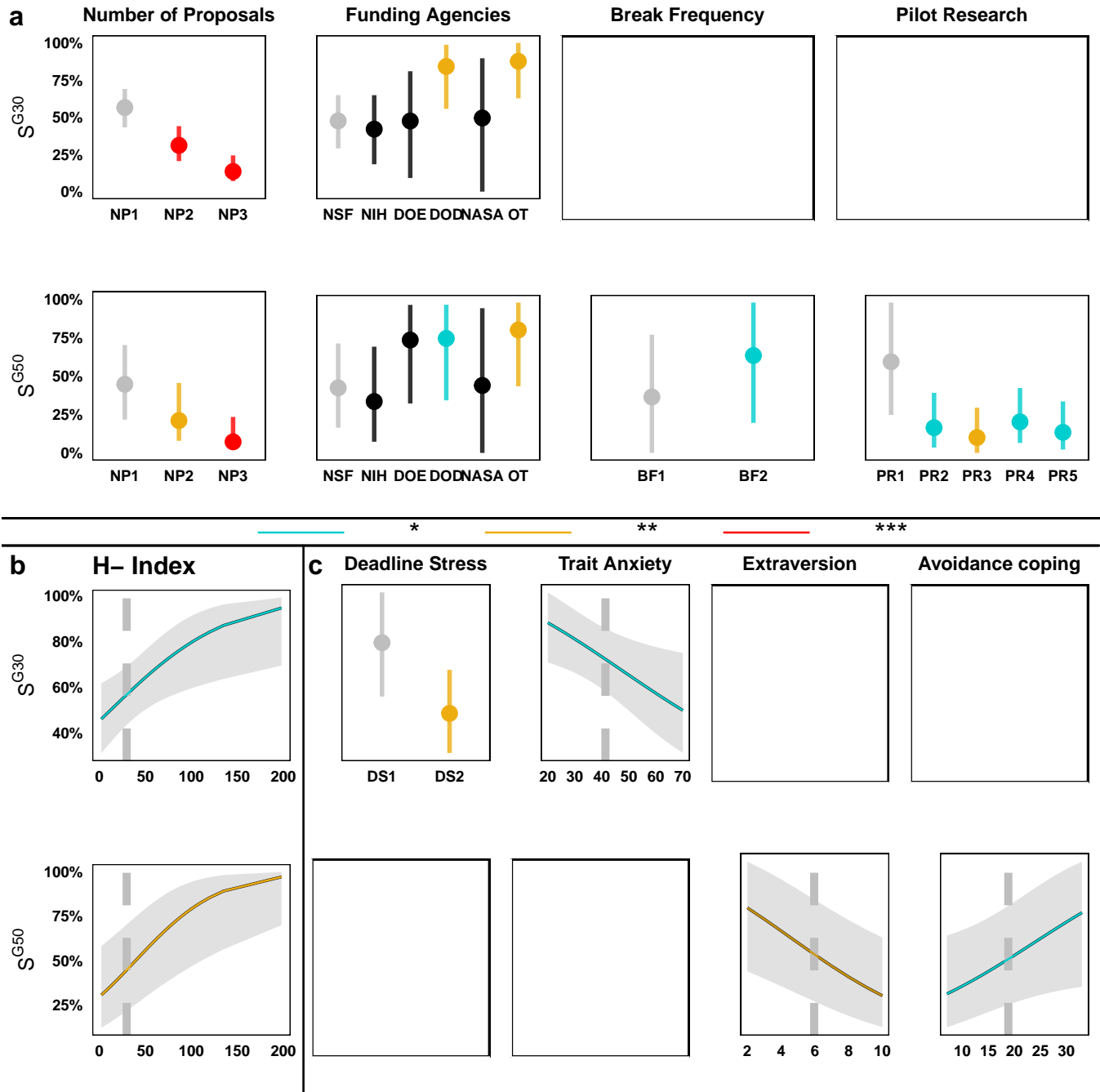| Predictor	Prob_wise	Estimate	Std..Error	z.value	Pr...z..	.
Intercept	0.011	-4.516	1.405	-3.214	0.001	**
NP2	0.147	-1.755	0.577	-3.039	0.002	**
NP3	0.073	-2.540	0.803	-3.164	0.002	**
E	0.434	-0.267	0.111	-2.409	0.016	*
A	0.575	0.302	0.147	2.053	0.04	*
TWR	0.510	0.040	0.014	2.925	0.003	**

Inferences:

Forward:

1. Starts with null mode and its AIC - 159.83

2. Keeping only significant predictors - NP FA PR E A TWR. AIC increased to 163.81

3. AS most of the Funding agencies and PR are greater than 0.05, we should not consider them for forward model. Model is further optimized with AIC-160.95 with all of the predictors are significant.
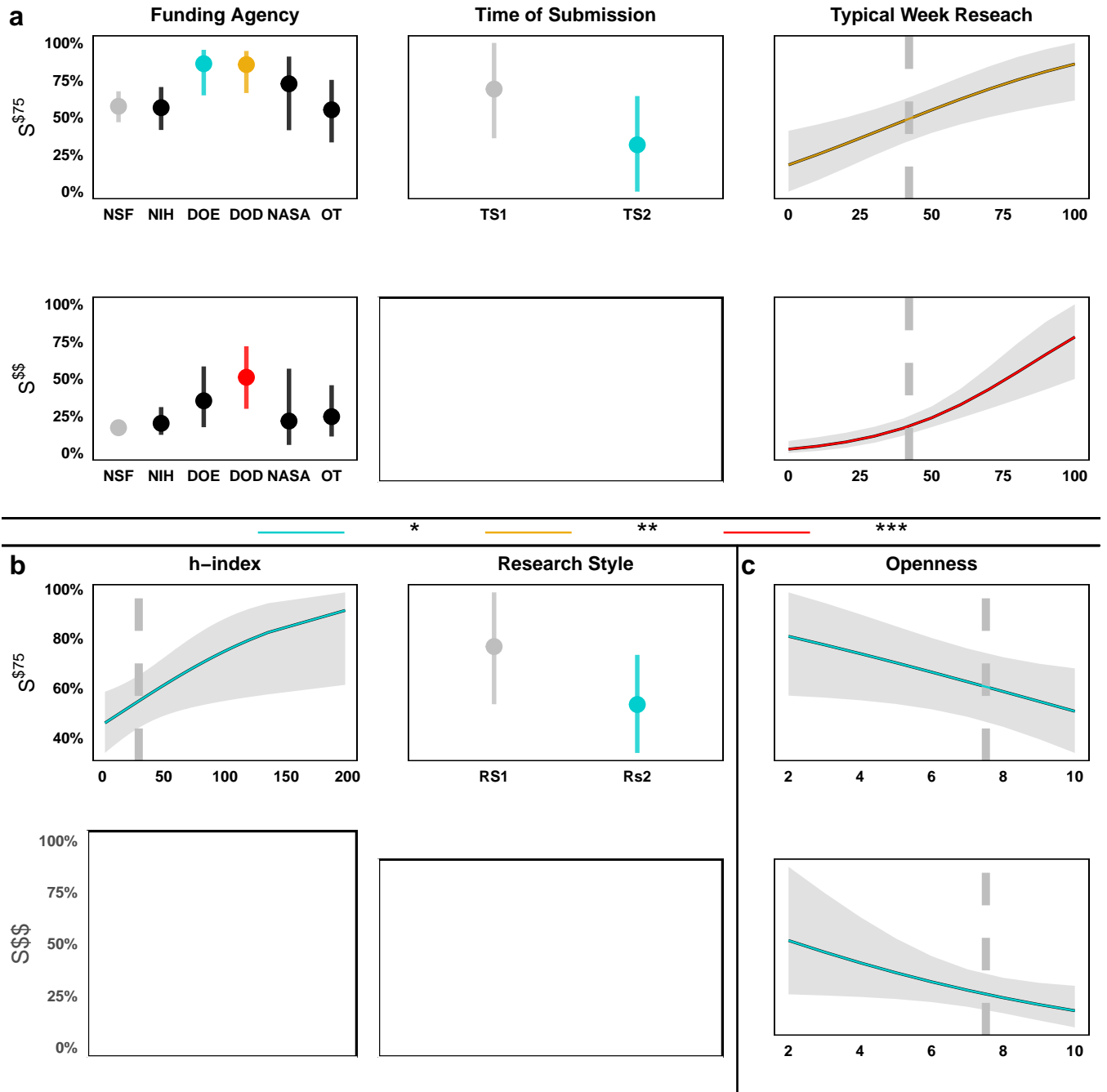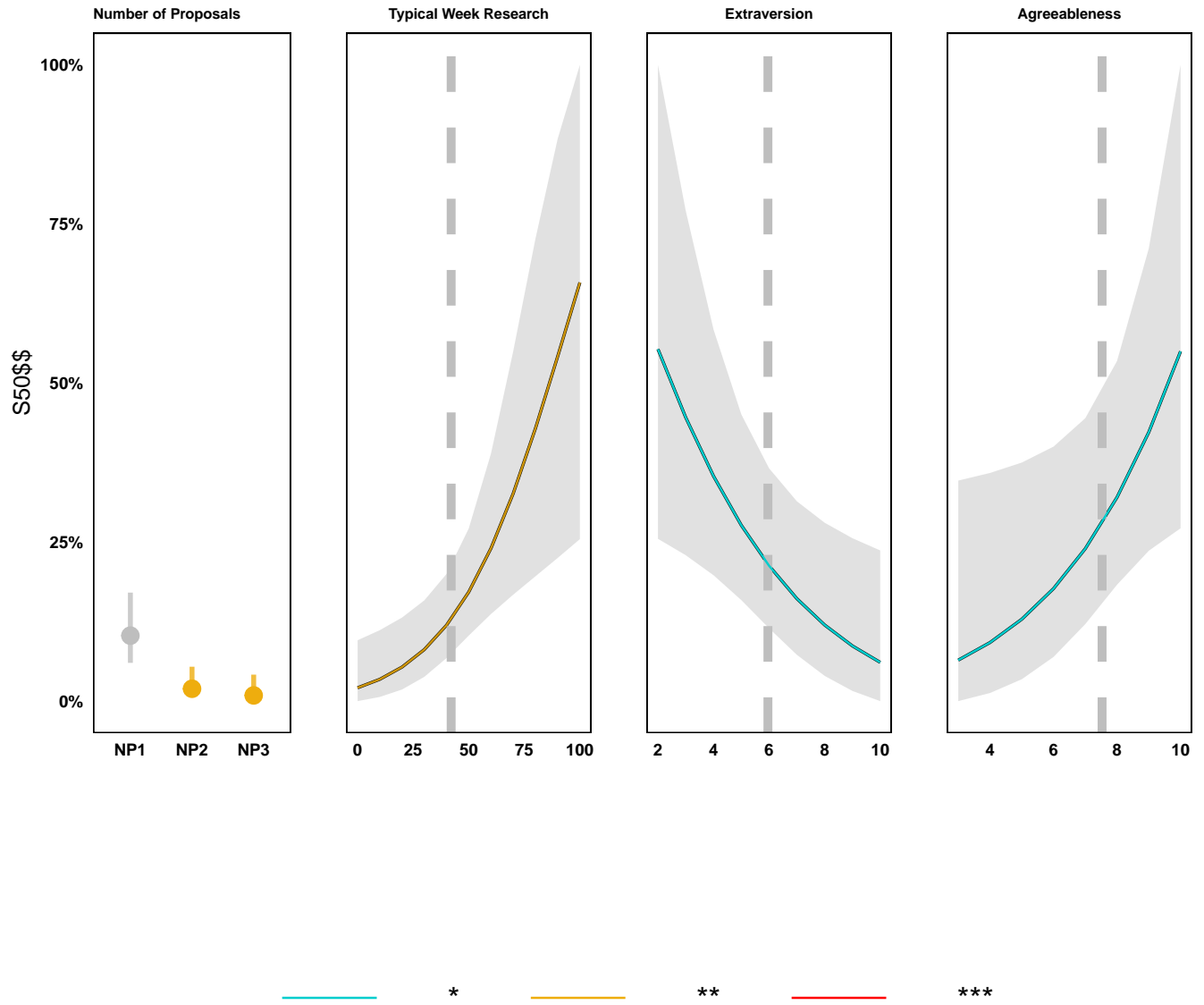
Step-wise:

1. Initial model AIC - 159.83 with RS and O as non significant predictors.

2. By removing the non significant predictors, model is optimized with AIC - 160.95

Backward Model:

1. Full model AIC - 186.77 with RS and O as non significant predictors.

2. By removing the non significant predictors, model is optimized with AIC - 160.95

a

Number of Proposals

Funding Agencies

Break Frequency

Pilot Research

b

H− Index

c

Deadline Stress

Trait Anxiety

Extraversion

Avoidance coping

*   **   ***

Inferences from the plots:

−> Faculty included in the sample has 32.5 female(272) and 67.5% Male(131) while the population in NCSES has 64.8% male and 35.2 female. When performed chi square on these vectors we got p value as 0.2735 which is greater than 0.05. As the p-value is greater than 0.05, we do not reject null hypothesis sample gender distribution is independent of the gender distribution of US academia.

−> SG30 and SG50 are successful of grantsmanship. SG30 consists of the respondents who estimate their success rate to be more than 30% (Successful) while SG50 consists of people with success rate more than 50% (Highly successful).

###SG30

−> Reference researcher is the one who submits 1-2 proposals to NSF and experience no stress. Grey bars falls under the category of RRG30(reference researcher).

−> As the number proposals increases, the probability for faculty to belong to the most successful class (SG30) decreases. This is concluded by looking at the graphs of Number of proposals of Sg30. Faculty that publishes 1-2 proposals have around 55% probability to fall under successful class category. Also, probability falls from somewhere around 55%(NP1) to below 25%(NP3).

−> From funding agency, Reference researcher has around 50% probability of being in successful group. The p-value for NSF is greater than 0.05, so we can accept null that they have highest probability of being in successful class. Also, we can say that NIH, DOE and NASA groups are almost equal to successful class group, they are represented by black bars with p-value > 0.05 (accept null and are in high probability of being in successful class). While DOD and OT have high probability to match the reference researcher who falls under successful class and these are represented with gold bars with p-value < 0.01 (as it is less than 0.05, reject null and conclude that they have least chance of being in successful group). As the number of submissions are higher for DOD and OT, so, approval of proposal is less.

−> H-index represents the academic fame of the faculty. SG30 h-index shows that the faculty belonging to the successful class is increasing and it is increasing by approximately 8%. The p-value for this model is less than 0.05(blue bar) so, we can reject null and conclude the more is the academic fame, the more is the chance of being in successful group.

−>From deadline stress graph, it clearly says that faculty who experience stress have low probability of making it to the successful class. DS2 represents faculty with more stress and they have less probability of matching with RRG30. SO, they are represented with gold bars, also they have p-value < 0.01. As the DS-2 p-value is less than 0.05, we can reject the null and conclude the more is the stress, the less is the probability of being in successful class.

−> Based on the trait anxiety scores recorded, SG30 faculty chances of matching with reference researcher is decreasing and is also shown in the graph. The probability is dropped from somewhere around 60% to 50%.p- value for this model is less than 0.05 so we can reject null and conclude that the more is the trait anxiety score, the less the probability of being in successful group.

###SG50

−> Reference researcher for RRg50 is the one who submits 1-2 proposals every year and submits to NSF, also takes 1-2 hours of break everyday, conducts pilot research less than 1 month. Adding to this, reference research is also inclined towards extroversion with score 6/10 and 18.8/33 as avoidance coping score.

−> From the graph of Number of proposals of SG50 faculty, it shows that faculty has probability of slightly more than 50% to belong to successful class.

−> Same as in SG30, faculty here in SG50 have low probability to be in successful class as their number of proposals increases.The probability decreases from around 50%(NP1) to around 3%(NP3). p-value for NP3 is less than 0.001 which is obviously less than 0.05, we can reject null and conclude that the number of proposals is more, the chance of being in successful group is less.

−> By the graph of funding agency of SG50, The probability of reference researcher being in successful class in almost 50% and p value is grater than 0.05, we can accept null and conclude that the less number of proposals gives the more chance of approval by the agency. we can conclude that NIH, DOE and NASA are almost equal with faculty of successful class. While, DOD and OT gave probability of around 25% and 30% respectively. Application are more to OT, so approval from the agency is less. We are concluding this by OT's p value which is less than 0.01 by rejecting null.

−> From the graph of break frequency, we can conclude that faculty taking breaks 3-4 hours everyday(BF-1) have low probability around 50% with p-value > 0.05 than BF-2 with probability around 75% with p-value less than 0.05.

−> In SG50, faculty with less pilot research (PR1) has high change of being in successful group while faculty with more pilot research(PR5) gave less chance if being in successful group. High pilot research faculty has only 25% of probability to be in successful group. PR1 has p-value greater than 0.05, we can accept null and conclude that the less pilot research is the more chance of being in successful group.

−> The probability of being in successful class based on h-index (academic face) in increased by 10% for SG50 faculty. The more academic fame (h-index), the higher is probability of being successful. The p-value here is less than 0.01, which is less than 0.05, we can reject null and conclude that the more the fame, the more the chance of being in successful group.

−> In SG50 faculty, faculty with increasing extroversion score having lower probability to be in successful class. p value is less than 0.01, reject null.

−> The more the avoidance coping factor, the more probability of being in successful group. This is purely in contradiction to the extroversion. p-value is less than 0.05, reject null.

−>Faculty who estimated their grant funding more than 75% of the total cost as considered to be in S$75 group. These are well funded research operations.

−> Faculty who estimated full funding are for their research are considered and categorized in Sdd(dollar dollar) group. These are full funded research operations.

S75 ** −− > ** $Reference researcher submits proposals to NSF, submits only on deadlines and invests 42.5$

−> Reference researcher here submits proposals to NSF and devotes 42.1% of their time to research. And has openness core of 7.5/10.

−> according to Funding agency graph, The probability of reference research to be in successful graph is only 17%. NIH, DOE, NASA and OT groups are in accordance with reference researcher while DOD has lowest chance of being in successful group with p-value less than 0.001 with red bars.

−> The more the amount of time they spend on research gives more probability of being in successful group.

−> The more they are open, the less is the chance of being in successful group. As the openness score increase from 6-10, the probability decreases from 25 to 17%.

SG50$$

−> The faculty who predicts their grantsmanship above 50% are considered to be in SG50dd group. They are considered as Reference highly successful in grantsmanship and well-funded researcher.

−> Reference researcher here submits 1-2 proposals per year and devotes 42.1% of their time to research. And has extroversion score and agreeableness score of 6 and 7.5 respectively.

−> Reference researcher belonging to successful group is around 10%. Same as other faculty, as the number of proposals increases, the probability of being in successful group decreases.

−> As the time they spend on research increases, the probability of making into successful group increases.

−> From the extraversion graph, As the faculty score of extraversion score increases, the probability of being in successful group decreases. Probability dropped from 15% TO 5% when the score increases from 4-8.

−> The probability of being in successful group increases by 6% approximately when agreeableness score increases form 6-10.