Children Intervention Report

Jingrong Cheng, Bo Zhang, Yue Ren, TF: Matthew Morse

March 16, 2017

Introduction

Our client is Tess Harvey, from Psychological and Brain Sciences Department. She conducted a research on whether children will lie to prevent bad things from happening, based on three experiments she designed previously. She assumed that children might lie to interfere bad consequences even though most of them already knew that lying is wrong. On the contrary, she believed that children might not lie if the result is good. Tess created several social conditions and two fictional characters A and B, who were children, in her experiments. She told the stories to the participants individually using a map with four locations (rock, fence, bin, and tree). Three of those are randomly assigned to be interventions and the last one is not. The stories began with that the character A hided a doll behind one of the locations and walked away, and then the character B would ask the participant where was the doll. At this point, our client would tell the participant about what condition would occur if the character B knew where the doll was. Asking the participant to point out the location that he/she wants the character B to know, our client was able to record if the participant lied about the location information under a certain condition. Since we only focused on the first two experiments at this stage, the four social conditions involved are: (1) stealing (2) sharing/giving (3) stealing and it would not hurt the feeling of the character A (4) stealing and it would hurt the feeling of the character A. The "stealing" condition means that if the participant tell the actual hiding location, then the character B would steal the doll. The "sharing/giving" condition indicates that if the participant tell the actual hiding location, then the character B would find the character A and share a cookie with him/her. The "stealing and it would not hurt the owner's feeling" condition refers to that if the participant tell the actual hiding location, then the character B would steal the doll but the character A does not care about the doll. The "stealing and it would hurt owner's feeling" represents that if the participant tell the actual hiding location, then the character B would steal the doll, which would hurt the character A because she/he cares about this doll. Throughout the experiments, the client used age, gender, condition and initial hiding location as predictors to create a logistic regression model. She was not satisfied with this model and wanted to find out a more proper approach. The client wanted to consider the effect of distance between lying location and actual hiding location into the model as well. Therefore, we suggested trying the ordinal logistic regression model for this research. Our tasks were to find out:

1. Is ordinal logistic regression a more suitable method in this case?

2. How to interpret the results?

Materials and Methods

The data we received included 97 observations for the experiment one, and 108 observations for the experiment two. All the participants are between 5 to 8 years old.

Experiment 1

```
##   id  age age_gp gender condition circled_map lie dis hide
## 1 70 8.00      1      F         0 Recycle Bin   1   1    0
## 2 91 6.26      0      M         0        tree   1   1    1
## 3 79 6.27      0      M         1       fence   0   0    1
```

```
## 4 78 6.68      0      M         0       rocks   0   0    0
## 5 97 6.97      0      F         0       rocks   0   0    0
## 6 89 7.45      1      M         1       fence   0   0    1
```

```
## [1] "0" "1"
```

Experiment 2

```
##   agegroup condition gender dis circled_map hide
## 1        0         1      M   0       Rocks    0
## 2        0         1      F   0       Rocks    0
## 3        0         1      M   0       Rocks    0
## 4        0         1      F   0       Rocks    0
## 5        0         1      M   2        Tree    0
## 6        0         1      M   3       Fence    0
```

The following is the detail of each variable: Age: a binary variable with 1 (younger group) and 2 (older group), and the younger group included 5 to 6 years old while the older group included 7 to 8 years old. Gender: a binary variable with F (female) and M (male) Lying indicator: a binary variable with 1 (lie) and 0 (no lie) Two conditions in experiment one: a binary variable with 0 (stealing) and 1 (sharing/giving) Three conditions in experiment two: 1 (sharing), 2 (stealing and it would not hurt the feeling of the character A), 3(stealing and it would hurt the feeling of the character A) Hiding location: rock = 0, fence = 1 Distance: 0 (telling the actual hiding location), 1 (lying and telling a location 1 distance from the actual hiding location), 2 (lying and telling a location 2 distances from the actual hiding location), 3 (lying and telling a location 3 distances from the actual hiding location.) After cleaning the data, we regressed the distance variable on age, gender, condition for each experiment using the ordinal logistic regression. We created a predicted probability table for each experiment based on the ordinal logistic regression model.

Ordinal Logistic Regression Model for experiment 1

Distance ~ Age + Condition + Gender + Hiding location

Ordinal Logistic Regression Model for experiment 2

Distance ~ Age + Condition + Gender + Hiding location

Results and Discussion

Experiment 1

First model information

```
##
## Re-fitting to get Hessian
```

```
## Call:
## polr(formula = factor(dis) ~ factor(age_gp) + factor(condition) +
##     gender + factor(hide), data = hid.experi1)
##
## Coefficients:
##                       Value Std. Error t value
## factor(age_gp)1    -0.1140     0.3984 -0.2862
## factor(condition)1 -0.9641     0.4101 -2.3509
## genderM            -1.3519     0.4311 -3.1360
```

```
## factor(hide)1       0.4227     0.4052  1.0431
##
## Intercepts:
##     Value   Std. Error t value
## 0|1 -0.8917  0.4038    -2.2081
## 1|2  0.2329  0.3897     0.5977
## 2|3  1.5381  0.4553     3.3782
##
## Residual Deviance: 221.1535
## AIC: 235.1535
```

```
##    factor(age_gp)1 factor(condition)1            genderM
##         -0.1140211         -0.9641115         -1.3519038
##      factor(hide)1
##          0.4226541
```

This model uses younger age group (0), condition 0, gender female, and rock (hiding location=0) as baseline.

P-value

```
##
## Re-fitting to get Hessian
```

```
##                          Value Std. Error    t value      p value
## factor(age_gp)1    -0.1140211  0.3983660 -0.2862220 0.7747081266
## factor(condition)1 -0.9641115  0.4101037 -2.3508972 0.0187282072
## genderM            -1.3519038  0.4310940 -3.1359841 0.0017127848
## factor(hide)1       0.4226541  0.4051961  1.0430855 0.2969087145
## 0|1                -0.8916823  0.4038212 -2.2081116 0.0272364919
## 1|2                 0.2329458  0.3897477  0.5976837 0.5500510317
## 2|3                 1.5381044  0.4552972  3.3782427 0.0007295065
```

Confidence Interval

```
## Waiting for profiling to be done...
```

```
##
## Re-fitting to get Hessian
```

```
##                          2.5 %     97.5 %
## factor(age_gp)1    -0.9012245  0.6658930
## factor(condition)1 -1.7853903 -0.1716373
## genderM            -2.2261058 -0.5280332
## factor(hide)1      -0.3692754  1.2249451
```

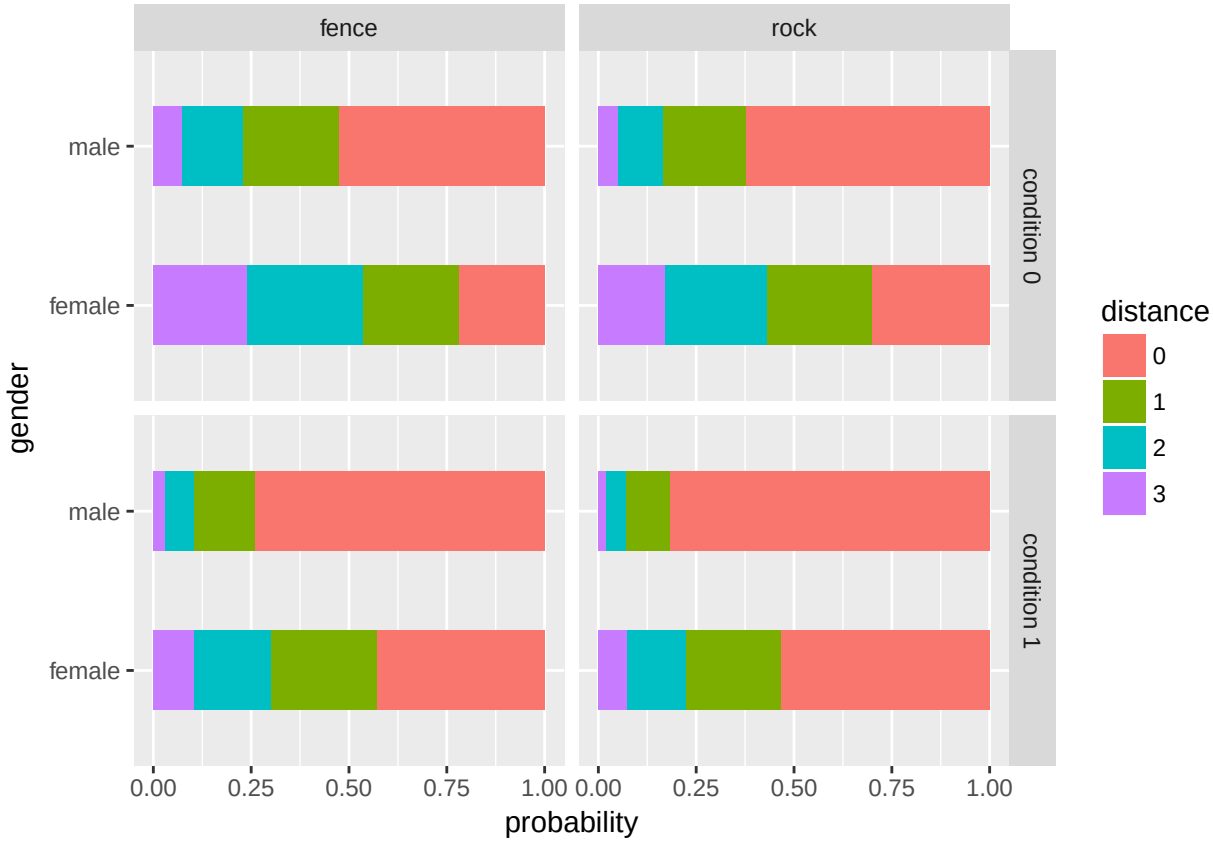
Instead of using p-value, we recommended to check the confidence interval for these coefficients. According to the table that presented above, condition and gender are likely having more impact on the ratio probability than age and hiding location. In other words, the condition and gender factors are more statistically significant than age and hiding location factors at ??=5%.

Probability Table for experiment 1

	age_gp	gender	condition	hide	0	1	2	3
1	1	F	0	0	0.31	0.27	0.25	0.16
20	1	F	0	1	0.23	0.25	0.29	0.23

	age_gp	gender	condition	hide	0	1	2	3
28	1	F	1	0	0.55	0.24	0.14	0.07
25	1	F	1	1	0.44	0.27	0.19	0.10
93	1	M	0	0	0.64	0.21	0.11	0.05
44	1	M	0	1	0.54	0.24	0.15	0.07
72	1	M	1	0	0.82	0.11	0.05	0.02
6	1	M	1	1	0.75	0.15	0.07	0.03
5	0	F	0	0	0.29	0.27	0.27	0.18
18	0	F	0	1	0.21	0.24	0.30	0.25
19	0	F	1	0	0.52	0.25	0.16	0.08
17	0	F	1	1	0.41	0.27	0.20	0.11
4	0	M	0	0	0.61	0.22	0.12	0.05
2	0	M	0	1	0.51	0.25	0.16	0.08
11	0	M	1	0	0.81	0.12	0.05	0.02
3	0	M	1	1	0.73	0.16	0.08	0.03

Plot for probability table



The pink color represents the children were telling the truth while the rest of the colors represent the children are lying. We created an estimated probability table based on the first model and showed the results through bar plot as above. In general, less children may lie under sharing/giving condition (1) and obviously more children may lie under stealing condition (0). For different gender groups, females are more likely to lie than male generally. In addition, females and males are likely to lie more under stealing condition compared to themselves under sharing/giving condition. In other words, the participants may lie to prevent bad situations from happening. Unfortunately, it is hard to tell that difference between the probability of telling true

between rock and fence as hiding location.

Experiment 2

Second model information

```
##
## Re-fitting to get Hessian
```

```
## Call:
## polr(formula = factor(dis) ~ factor(agegroup) + factor(condition) +
##     gender + factor(hide), data = hid.experi2)
##
## Coefficients:
##                      Value Std. Error t value
## factor(agegroup)1  0.2050     0.3695  0.5548
## factor(condition)2 0.3737     0.4659  0.8020
## factor(condition)3 1.1537     0.4664  2.4738
## genderM            0.2562     0.3737  0.6854
## factor(hide)1      0.7487     0.5467  1.3695
##
## Intercepts:
##     Value  Std. Error t value
## 0|1 0.6735 0.4594     1.4662
## 1|2 1.5700 0.4818     3.2583
## 2|3 2.5240 0.5175     4.8772
##
## Residual Deviance: 266.0937
## AIC: 282.0937
```

p-value

```
##
## Re-fitting to get Hessian
```

```
##                         Value Std. Error   t value      p value
## factor(agegroup)1  0.2050190  0.3695162 0.5548310 5.790102e-01
## factor(condition)2 0.3736646  0.4659416 0.8019559 4.225784e-01
## factor(condition)3 1.1537076  0.4663651 2.4738290 1.336737e-02
## genderM            0.2561530  0.3737362 0.6853844 4.931014e-01
## factor(hide)1      0.7487404  0.5467095 1.3695398 1.708306e-01
## 0|1                0.6735079  0.4593697 1.4661564 1.426057e-01
## 1|2                1.5699608  0.4818369 3.2582825 1.120888e-03
## 2|3                2.5239531  0.5174973 4.8772299 1.075860e-06
```

We caculated p-value for the second model as well, but we would suggest to use confidence interval to check the significance.

Confidence Interval

```
## Waiting for profiling to be done...
```

```
##
## Re-fitting to get Hessian
```

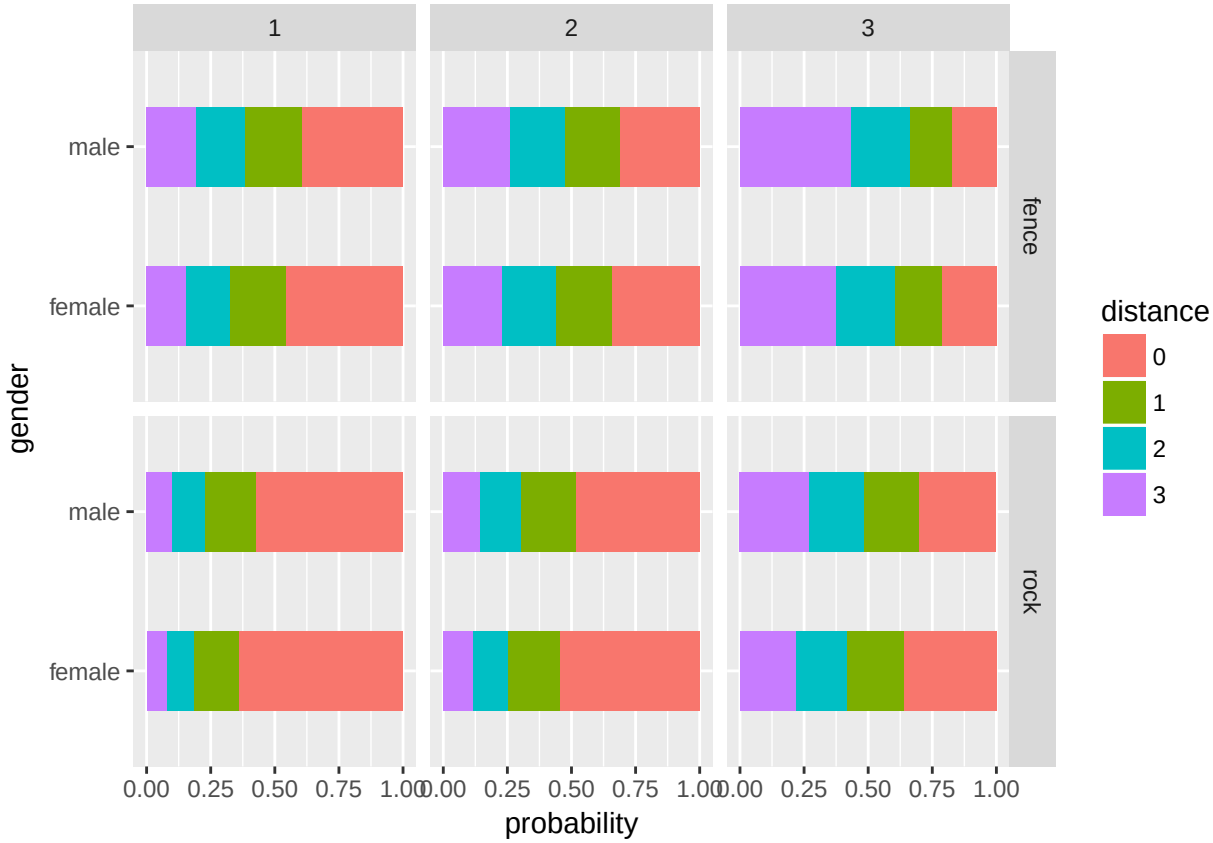
```
##                           2.5 %    97.5 %
## factor(agegroup)1  -0.5195304 0.9331872
## factor(condition)2 -0.5342059 1.3011711
## factor(condition)3  0.2532617 2.0889370
## genderM            -0.4782212 0.9912311
## factor(hide)1      -0.3234451 1.8415166
```

Let us recall the details of the conditions from the experiment two: 1 (sharing), 2 (stealing and it would not hurt the feeling of the character A), 3(stealing and it would hurt the feeling of the character A.) According to the confidence interval table presented above, the coefficients of age, gender, hiding location, and condition 2 are less significant than condition 3 at ?? = 5%. This is interesting compared to the first experiment.

Probability table for experiment 2

	agegroup	condition	gender	hide	0	1	2	3
1	0	1	M	0	0.60	0.19	0.12	0.09
2	0	1	F	0	0.66	0.17	0.10	0.07
8	0	2	F	0	0.57	0.19	0.13	0.10
9	0	2	M	0	0.51	0.21	0.15	0.13
11	0	2	M	1	0.33	0.22	0.21	0.24
13	0	3	M	0	0.32	0.22	0.21	0.25
14	0	3	F	0	0.38	0.22	0.19	0.20
16	0	3	F	1	0.23	0.19	0.23	0.35
24	0	1	F	1	0.48	0.21	0.16	0.14
30	0	1	M	1	0.42	0.22	0.18	0.18
53	0	3	M	1	0.18	0.17	0.23	0.41
55	1	1	F	0	0.62	0.18	0.11	0.09
57	1	1	M	0	0.55	0.20	0.14	0.11
61	1	1	F	1	0.43	0.22	0.18	0.17
64	1	2	M	0	0.46	0.22	0.17	0.16
65	1	2	F	0	0.52	0.21	0.15	0.13
67	1	2	M	1	0.29	0.21	0.22	0.28
74	1	3	F	0	0.34	0.22	0.21	0.24
75	1	3	M	0	0.28	0.21	0.22	0.29
76	1	3	M	1	0.16	0.16	0.23	0.46
77	1	3	F	1	0.19	0.18	0.23	0.40
88	1	1	M	1	0.37	0.22	0.20	0.21
98	1	2	F	1	0.34	0.22	0.21	0.23

Plot for probability table



Different from the experiment one, male is likely to lie under all three conditions than female. The results suggests that children likely to tell the lying location three distance away from the actual hiding location, under the situation with condition 3 and the hiding location is fence. This is interesting and may need more research to find the causality behind it. Overall, the probability of children to lie is decreasing from condition 3 to condition 2 and then to condition 1.

Conclusion

Recalling our two tasks that are presented previously, this research answers both questions properly.

1. Is ordinal logistic regression a more suitable method in this case?

We strongly suggest that ordinal logistic regression is a suitable approach in this case. Since the response variable (distance) is ordinal and has four levels, the ordinal logistical regression would be a better fit than the logistic regression that our clients tried before.

2. How to interpret the results?

We provided clear interpretation for each model and calculated p-value and confidence interval, created probability table and plotted the probability table to understand the model well. It turned out condition is a relatively significant matter in this study and gender may be different based on different situations. We may conclude that difference between two different age group are not obvious. Overall, children are more likely to lie in order to prevent bad things from happening according to this study.