Performance Evaluation of Generalized Logit Model and Proportional Odds Model: A Comparative Analysis on Car Evaluation Dataset

1 Car Evaluation Data Analysis

Introduction

The data set used for this project is publicly available through the UCI Machine Learning Repository (Car Evaluation), [1]. This dataset contains 7 variables with sample size of 1728. We, however, only concentrate on 288 of them after exclude observations with $X_4 = 2$, X_6=low, X_1 =high, X_1 =vhigh, or X_2 =vhigh. The response and predictors are listed below:

- Y: the car acceptability with four categories: acc, unacc, and high[1], response variable.
- X_1 : the buying price of a car with two levels: low and med.
- X_2 : the price of maintenance with three levels: high, med, and low.
- X_3 : number of doors for a car with four levels: 2, 3, 4, 5(5 and more than 4).
- X_4 : persons capacity with two levels: 4 and more.
- X_5 : the size of the truck with three levels: big, med, and small.
- X_6 : estimated safety of a car with two levels: high and med.
- Z_2 : interval variable representing X_2: 3(high), 2(med), and 1(low).
- Z_3 : interval variable representing X_3: 1(2 doors), 2(3 doors), 3(4 doors), 4(5 and more than 5 doors).
- Z_5 : interval variable representing X_5: 3(big), 2(med), and 1(small).

In the data analysis, we use `SAS` to perform the data analysis. We plan to use two models in the analysis. First one is generalized logit model in which the response treated as a nominal variable. Second one is the proportional odds model in which the response treated as an ordinal variable.

[1] As in the lecture note, we combined level vgood and good as high

Before we conducting the analysis, we looked into the proportional tables 1) to decide whether to represent X_2, X_3, and X_5 as binary indicators or as an approximating interval variable, 2) to see if there are levels could be combined and 3) to detect possible (complete or quasi-complete) separation in the data.

First, to decide whether to represent ordinal explanatory variables as binary indicators or as an approximating interval variables, we looked into the contingency tables. To use approximating interval variables, the proportion to each level should be monotonically increasing or decreasing with the ordinal variable. Except X_3 where Y=acc, this condition was satisfied. We concluded that approximating interval variables are fine to include and decided to examine both interval variables and binary variables. We will denote the interval variables Z_2, Z_3 and Z_5 and denote binary variables X_2, X_3 and X_5.

Secondly, there was possible quasi-complete separation in the data. Since P(Y=unacc—X_5=big)=0, we need to be cautious when we interpret the result at those points.

```
Contingency tables


Contingency Table 1: Y vs X1 (%)

        low med

  acc     53  79

  high    85  49

  unacc    6  16


Contingency Table 2: Y vs X2 (%)

        high low med

  acc      69  20  43

  high     13  72  49

  unacc    14   4   4


Contingency Table 3: Y vs X3 (%)

         2  3  4 5more

  acc   32 36 32    32

  high  25 33 38    38
```

```
  unacc 15  3  2     2


Contingency Table 4: Y vs X4 (%)

         4 more

  acc   72   60

  high  66   68

  unacc  6   16


Contingency Table 5: Y vs X5 (%)

        big med small

  acc    32  44    56

  high   64  49    21

  unacc   0   3    19


Contingency Table 6: Y vs X6 (%)

        high med

  acc     43  89

  high    95  39

  unacc    6  16
```

Data Analysis

1.1 Generalized logit model

1.1.1 Model description

The first model we fitted is generalized logit model with all explanatory variables(Table 1, we will denote this model as full model). Table 1 shows parameter estimates, 95% confidence interval and p-value of each estimates[2].

The full model has quasi-complete separation issue. We could see the CI of the coefficient for X_5 of Y=unacc is very wide. This happens because the model is fitting estimates on frequency

[2]Because of quasi-complete separation issue this model is suffering, the estimates of X_5 for Y=unacc is not valid.

0. Despite the quasi-complete separation, we did not see the estimates diverge to infinity. This is because statistic packages automatically stops the process of convergence. This does not mean that the package could calculate the exact estimates.

When quasi-complete separation presents, not only the estimates but the p-value of Wald test also are unreliable. One alternative is to calculate likelihood ratio of models, one with variable and the other without the variable. Since -2logL has chi-square distribution of its degrees of freedom, we could test significance of each variable. The result is presented on Table 3. All variables are significance at 5 % significance level and the least significant variable is X_4.

Parameter		y	DF	Estimate	Standard Error	95% CI-Lower	95%CI-Upper	p-value
Intercept		high	1	-1.03	0.34	-1.70	-0.36	0.00
Intercept		unacc	1	-9.60	51.09	-109.75	90.54	0.85
x1	low	high	1	2.71	0.51	1.71	3.72	¡0.0001
x1	low	unacc	1	-1.13	0.43	-1.97	-0.30	0.01
x2	high	high	1	-7.17	1.27	-9.65	-4.69	¡0.0001
x2	high	unacc	1	1.94	0.59	0.77	3.10	0.00
x2	low	high	1	6.04	1.07	3.95	8.13	¡0.0001
x2	low	unacc	1	-0.88	0.61	-2.08	0.31	0.15
x3	2	high	1	-2.67	0.67	-3.98	-1.35	¡0.0001
x3	2	unacc	1	3.19	0.74	1.73	4.65	¡0.0001
x3	3	high	1	-0.17	0.50	-1.15	0.81	0.73
x3	3	unacc	1	-0.61	0.68	-1.94	0.72	0.37
x3	4	high	1	1.42	0.54	0.36	2.48	0.01
x3	4	unacc	1	-1.29	0.78	-2.82	0.24	0.10
x4	more	high	1	0.17	0.28	-0.38	0.73	0.54
x4	more	unacc	1	1.19	0.42	0.37	2.00	0.00
x5	big	high	1	4.58	0.89	2.83	6.33	¡0.0001
x5	big	unacc	1	-9.40	102.20	-209.72	190.91	0.93
x5	med	high	1	0.97	0.41	0.16	1.78	0.02
x5	med	unacc	1	2.77	51.08	-97.36	102.89	0.96
x6	high	high	1	4.29	0.76	2.80	5.77	¡0.0001
x6	high	unacc	1	-1.10	0.43	-1.95	-0.25	0.01

Table 1: Maximum Likelihood Estimates of the Full Model of Generalized Logit Model

1.1.2 Odds Ratio Estimates

From Table 2 we can easily get the ratio of the conditional probability of an unacceptable car to the conditional probability of an acceptable car for change in $(X_2, X_4) =$ (high, 4) to $(X_2, X_4) =$ (low, more) is $\frac{19.84}{1.17} \times \frac{1}{10.78} = 1.52$. That means odds of an unacceptable car with 4 persons capacity and higher maintenance price is about 1.5 times of the odds for an acceptable car with more than 4 persons capacity and lower maintenance price.

Also, we can get the ratio of the conditional probability of a highly evaluated car to the conditional probability of an acceptable car for change in $(X_2, X_4) = $ (high, 4) to $(X_2, X_4) = $ (low, more) is $\frac{0.001}{135} \times \frac{1}{1.41} = 0.000005$. That means odds of a highly evaluated car with 4 persons capacity and higher maintenance price is about 0.000005 times the odds of an acceptable car with more than 4 persons capacity and lower maintenance price.

Effect	Y	Point Estimates
x1 low vs med	high	227.67
x1 low vs med	unacc	0.10
x2 high vs med	high	0.001
x2 high vs med	unacc	19.84
x2 low vs med	high	135.00
x2 low vs med	unacc	1.19
x3 2 vs 5	high	0.02
x3 2 vs 5	unacc	88.48
x3 3 vs 5	high	0.20
x3 3 vs 5	unacc	1.97
x3 4 vs 5	high	1.00
x3 4 vs 5	unacc	1.00
x4 more vs 4	high	1.41
x4 more vs 4	unacc	10.78
x5 big vs small	high	999.999
x5 big vs small	unacc	0.001
x5 med vs small	high	682.65
x5 med vs small	unacc	0.02
x6 high vs med	high	999.999
x6 high vs med	unacc	0.11

Table 2: Odds Ratio of Full Generalized Logit Model

	-2logL	df	-2logLR	p-value of marginal variable
M3	388.06	20		
M3-x1	309.14	18	78.92	¡0.0001
M3-x2	188.66	16	199.40	¡0.0001
M3-x3	324.69	14	63.37	¡0.0001
M3-x4	376.93	18	11.13	0.004
M3-x5	221.72	16	166.34	¡0.0001
M3-x6	247.54	18	140.52	¡0.0001

Table 3: p-values of explanatory variables in the Full Generalized Logit model

To find the best model, we considered interaction term[3], ordinal variable(as interval or as categorical variables) and backward elimination. Since each has 2 options, we came up with 8 combination of models. The BIC of the models are shown in Table 4. Considering BIC of the

[3] We selected X_1X_6 as our interaction term. We anticipated that X_1 and X_6 has positive interaction term since safety issue with expensive brand would be more stigmatizing than that of lower price.

model, chose M_8 as our best model, which does not includes interaction term and X_4[4][5].

Model	Containing variables	BIC
M_1	X_1, X_2, X_3, X_4, X_5, X_6, X_1X_6	258.664
M_2	X_1, Z_2, Z_3, X_4, Z_5, X_6, X_1X_6	245.296
M_3	X_1, X_2, X_3, X_4, X_5, X_6	260.704
M_4	X_1, Z_2, Z_3, X_4, Z_5, X_6	245.050
M_5	X_1, X_2, X_3, X_5, X_6, X_1X_6	259.924
M_6	X_1, Z_2, Z_3, Z_5, X_6, X_1X_6	244.339
M_7	X_1, X_2, X_3, X_5, X_6	266.589
M_8	X_1, Z_2, Z_3, Z_5, X_6	243.279

Table 4: BIC of Candidate Generalized Logit Models

Parameter		Y	DF	Estimate	Standard Error	CI-Lower	CI-Upper	p-value
Intercept		high	1	-1.03	0.34	-1.70	-0.36	0.00
Intercept		unacc	1	-9.60	51.09	-109.75	90.54	0.85
x1	low	high	1	2.71	0.51	1.71	3.72	¡0.0001
x1	low	unacc	1	-1.13	0.43	-1.97	-0.30	0.01
x2	high	high	1	-7.17	1.27	-9.65	-4.69	¡0.0001
x2	high	unacc	1	1.94	0.59	0.77	3.10	0.00
x2	low	high	1	6.04	1.07	3.95	8.13	¡0.0001
x2	low	unacc	1	-0.88	0.61	-2.08	0.31	0.15
x3	2	high	1	-2.67	0.67	-3.98	-1.35	¡0.0001
x3	2	unacc	1	3.19	0.74	1.73	4.65	¡0.0001
x3	3	high	1	-0.17	0.50	-1.15	0.81	0.73
x3	3	unacc	1	-0.61	0.68	-1.94	0.72	0.37
x3	4	high	1	1.42	0.54	0.36	2.48	0.01
x3	4	unacc	1	-1.29	0.78	-2.82	0.24	0.10
x4	more	high	1	0.17	0.28	-0.38	0.73	0.54
x4	more	unacc	1	1.19	0.42	0.37	2.00	0.00
x5	big	high	1	4.58	0.89	2.83	6.33	¡0.0001
x5	big	unacc	1	-9.40	102.20	-209.72	190.91	0.93
x5	med	high	1	0.97	0.41	0.16	1.78	0.02
x5	med	unacc	1	2.77	51.08	-97.36	102.89	0.96
x6	high	high	1	4.29	0.76	2.80	5.77	¡0.0001
x6	high	unacc	1	-1.10	0.43	-1.95	-0.25	0.01

Table 5: Maximum Likelihood Estimates of the Final Model of Generalized Logit Model

1.1.3 Takeaways

From the model we processed, we could draw some useful conclusions for car dealership managers.

First, buy cheaper cars and avoid expensive ones. Low price affected both probability of getting a good grade and an the lowest grade(unacceptable). Chance of getting a good grade gets higher

[4]We decided to choose this model since excluding any one of variables did not decrease BIC level.

[5]Since we treated X_5 as interval variable, M_8 did not have quasi-complete separation issue. Therefore, we consider the estimates in Table 5 is reliable.

Effect	Y	Point Estimates
x1 low vs med	high	94.341
x1 low vs med	unacc	0.197
z2	high	0.006
z2	unacc	3.149
z3	high	3.009
z3	unacc	0.247
z5	high	52.793
z5	unacc	0.053
x6 high vs med	high	835.556
x6 high vs med	unacc	0.217

Table 6: Odds Ratio of Final Generalized Logit Model

with lower price, and chance of getting an the lowest grade(unacceptable) gets lower with lower grade.

Second, maintenance price is the most important. Higher maintenance price elevates the chance of getting an the lowest grade(unacceptable) and lower maintenance price lift the chance of getting an good grade.

Third, avoid 2 door cars. The chance of getting an the lowest grade(unacceptable) increases when a car has only 2 doors.

Fourth, people like bigger trunks. A car with a big trunk has lower chance of getting an the lowest grade(unacceptable) and higher chance of getting an good grade.

Finally, safety matters. The safer the car is, higher chance of getting a good grade.

1.2 Proportional odds model

1.2.1 Model description

Considering that Y is an ordinal variable, we also could consider proportional odds model as our candidate. As we did above, we first started with a full model with all given explanatory variables. The result is given in Table 7[6]. According to the result, all variables except X_4 were significant at 5% significance level.

1.2.2 Odds Ratio Estimates

From Table 8 we can easily get the ratio of the conditional probability of a highly evaluated car to the conditional probability of a not highly evaluated car. for change in $(X_2, X_4) =$ (high, 4) to

[6] We did not encounter any quasi-complete separation issue with proportional odds model since we are pooling the data when we calculate the estimates. Therefore, we could use Wald test for significance of each variables in the SAS output(Table 9).

(X_2, X_4) = (low, more) is $\frac{0.02}{12.48} \times \frac{1}{0.63} = 0.003$. That means odds of a highly evaluated car with 4 persons capacity and higher maintenance price is about 0.003 times the odds of a not highly evaluated car with more than 4 persons capacity and lower maintenance price.

Also, we can get the ratio of the conditional probability of a highly evaluated car to the conditional probability of an unacceptable car for change in (X_2, X_4) = (high, 4) to (X_2, X_4) = (low, more) is $\frac{0.02}{12.48} \times \frac{1}{0.63} = 0.003$. That means odds of an unacceptable car with more than 4 persons capacity and lower maintenance price is about 333 times the odds of a highly evaluated car with 4 persons capacity and higher maintenance price.

Parameter		DF	Estimate	Standard Error	CI-Lower	CI-Upper	p-value
Intercept	3	1	-0.50	0.22	-0.93	-0.07	0.02
Intercept	2	1	7.23	0.76	5.74	8.71	¡0.0001
x1	low	1	1.50	0.23	1.05	1.96	¡0.0001
x2	high	1	-3.47	0.42	-4.28	-2.65	¡0.0001
x2	low	1	2.99	0.40	2.22	3.77	¡0.0001
x3	2	1	-2.13	0.36	-2.85	-1.42	¡0.0001
x3	3	1	0.17	0.31	-0.44	0.79	0.58
x3	4	1	0.98	0.34	0.32	1.64	0.00
x4	more	1	-0.23	0.18	-0.59	0.13	0.20
x5	big	1	2.55	0.37	1.83	3.28	¡0.0001
x5	med	1	0.60	0.26	0.08	1.12	0.02
x6	high	1	2.12	0.27	1.58	2.66	¡0.0001

Table 7: Maximum Likelihood Estimates of the Full Model of Proportional odds model

Effect	Point Estimate
x1 low vs med	20.28
x2 high vs med	0.02
x2 low vs med	12.48
x3 2 vs 5	0.04
x3 3 vs 5	0.45
x3 4 vs 5	1.00
x4 more vs 4	0.63
x5 big vs small	301.27
x5 med vs small	42.63
x6 high vs med	69.40

Table 8: Odds Ratio of Full Proportional odds model

To find the best proportional odds model, we used the same framework we set in the first part. We considered interaction term X_1X_6, the way we include ordinal variable(as interval or as categorical variables) and backward elimination. We fitted eight models in Table 10, and chose

Variable	Wald Chi-square	df	p-value
X_1	41.60	1	¡.0001
X_2	71.35	2	¡.0001
X_3	35.49	3	¡.0001
X_4	1.62	1	0.20
X_5	6480	2	¡.0001
X_6	59.66	1	¡.0001

Table 9: p-values of explanatory variables in the Full Proportional odds model

the best model based on BIC. The best model we chose is M_8, which does not include interaction term and X_4. After fitting the model, we excluded one of the variables to make sure we cannot improve our model by excluding variables. The parameter estimates of chosen model are presented in Table 11.

Model	Containing variables	BIC
M_1	X_1, X_2, X_3, X_4, X_5, X_6, X_1X_6	268.721
M_2	X_1, Z_2, Z_3, X_4, Z_5, X_6, X_1X_6	264.017
M_3	X_1, X_2, X_3, X_4, X_5, X_6	264.760
M_4	X_1, Z_2, Z_3, X_4, Z_5, X_6	260.441
M_5	X_1, X_2, X_3, X_5, X_6, X_1X_6	264.643
M_6	X_1, Z_2, Z_3, Z_5, X_6, X_1X_6	259.579
M_7	X_1, X_2, X_3, X_5, X_6	260.717
M_8	X_1, Z_2, Z_3, Z_5, X_6	256.013

Table 10: BIC of Candidate Proportional Odds Models

Parameter		DF	Estimate	Standard Error	CI-Lower	CI-Upper	p-value
Intercept	3	1	-1.87	0.74	-3.31	-0.42	0.01
Intercept	2	1	5.05	0.87	3.35	6.75	¡0.0001
x1	low	1	1.37	0.22	0.95	1.80	¡0.0001
z2		1	-3.01	0.35	-3.70	-2.32	¡0.0001
z3		1	0.93	0.18	0.58	1.28	¡0.0001
z5		1	2.61	0.32	1.98	3.24	¡0.0001
x6	high	1	1.94	0.25	1.45	2.42	¡0.0001

Table 11: Maximum Likelihood Estimates of the Final Model of Proportional odds model

Effect	Point Estimate
x1 low vs med	15.63
z2	0.05
z3	2.53
z5	13.57
x6 high vs med	48.04

Table 12: Odds Ratio of Final Proportional odds model

1.2.3 Takeaways

The final proportional odds model has the same predictors as generalized logit model and the same the conclusion. A car dealership manager still should avoid expensive cars, maintenance is still the most important variable[7], 2 door cars are not preferred, bigger trunks[8] are more popular, and safety matters. The conclusions we drew using generalized logit model are still valid and it proves the conclusions are robust.

Conclusion

We decided to choose generalized logit model for two reasons. First, proportional odds model did not satisfy the assumption[9]. Since generalized logit model does not assume the same coefficient estimates, we believe this model has more benefit over the proportional odds model. Second, generalized logit model has lower BIC. The model maximizes the likelihood function better.

Despite we did not choose proportional odds mode, we could say the result is credible since we can draw the same conclusion from two different models. The conclusion follows: 1) cheaper price, 2) less maintenance, 3) bigger trunks increase the chance of getting a good grade. On ther other hand, 1) higher price, 2) more maintenance, 3) having only 2 doors increases chance of getting the lowest grade(unacceptable).

[7]The price of the maintenance seems to have a highest impact on the acceptability of a car with the probability increase $e^{-3.01} = 0.049$ when one unit increase the price of the maintenance

[8]The acceptability of a car increases $e^{2.61} = 13.6$ when we increase the trunk size by one unit

[9]The p-value for the score test for the proportional odds assumption was less then 0.0001.

References

[1] Rajkovic V Bohanec M. "Knowledge acquisition and explanation for multi-attribute decision making". In: *8th Intl Workshop on Expert Systems and their Applications* (1988), pp. 59–78.