

5.2. Nominal/Ordinal Logistic Regression

Dr. Bean - Stat 5100

Name one potential advantage and one potential disadvantage for using a nominal logistic regression model to describe an ordinal categorical variable?

- **Advantage:** No longer have to worry about assumption of proportional odds (i.e. the increase in odds for a change in X is constant regardless of the desired category level).
- **Disadvantage:** You compute more coefficients than you need to (as no parameter estimates are shared between models).

Given the nominal logistic regression output in Figure 1 and obtained from 5.2.1, please write out the two polytomous logistic regression equations.

$$\begin{aligned}L_{2|1} &= -0.591 - 0.388 * \text{sex} + 1.128 * \text{age}_2 + 1.588 * \text{age}_3 \\L_{3|1} &= -1.039 - 0.813 * \text{sex} + 1.478 * \text{age}_2 + 2.917 * \text{age}_3\end{aligned}$$

What is the estimated probability that a 30 year old woman will rate AC and Power Steering as a “very important” feature?

$$\begin{aligned}L_{2|1} &= -0.591 + 1.128 = 0.537 \\L_{3|1} &= -1.039 + 1.478 = 0.439 \\ \hat{\pi} &= \frac{e^{0.439}}{1 + e^{0.537} + e^{0.439}} = 0.364\end{aligned}$$

36.4%

Analysis of Maximum Likelihood Estimates						
Parameter	response	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	3	1	-1.0391	0.3305	9.8843	0.0017
Intercept	2	1	-0.5908	0.2840	4.3286	0.0375
S	3	1	-0.8129	0.3210	6.4122	0.0113
S	2	1	-0.3881	0.3005	1.6677	0.1966
A2	3	1	1.4780	0.4009	13.5912	0.0002
A2	2	1	1.1283	0.3416	10.9059	0.0010
A3	3	1	2.9165	0.4229	47.5594	<.0001
A3	2	1	1.5876	0.4029	15.5270	<.0001

Odds Ratio Estimates				
Effect	response	Point Estimate	95% Wald Confidence Limits	
S	3	0.444	0.236	0.832
S	2	0.678	0.376	1.223
A2	3	4.384	1.998	9.620
A2	2	3.090	1.582	6.037
A3	3	18.477	8.066	42.327
A3	2	4.892	2.221	10.775

Figure 1: Nominal logistic regression models using age and sex to predict odds of rating AC and power steering in cars as not important/important/very important.