

Quantitative Methods II

Assignment 12

December 4, 2011

Solutions

PROBLEM: BETTER DEAD THAN RED?

[[5]]

a: The dependent variable is a categorical variable with two possible values (levels). Those levels are `allowed` ($n_a = 874$) and `not allowed` ($n_n = 442$). This variable has 1316 responses and 707 missing values.

b: As the dependent variable is binary (dichotomous), I will first use a generalized linear model utilizing the Binomial distribution and a logit link function. This will allow me to check for the presence of overdispersion. If there is significant overdispersion, then I will fit the same model using quasi-likelihood estimation.

c: The research model will be

```
spkcom ~ afterlif + attend + male
```

Because the overdispersion parameter was not much larger than one ($dev/df = 1.2$), I decided the maximum likelihood model was not inappropriate for this data. The results are summarized in Table 1 (below).

d: Bob definitely believes in the afterlife and attends church services every week. The probability that Bob would be in favor of letting a communist speak at a public event is 0.3665.

	Estimate	Std. Error	t-value	p-value	
Constant term	-0.5408	0.4303	-1.26	0.2093	
Belief in afterlife:				0.3178	
Yes, definitely	-0.4344	0.6092	-0.71	0.4758	
Yes, probably	-1.0816	0.6938	-1.56	0.1190	
No, probably not	-0.8950	0.7700	-1.16	0.2451	
No, definitely not					Base Category
Church attendance:				0.0289	*
More than once per week	1.7702	0.7709	2.30	0.0217	*
Every week	2.0064	0.7022	2.86	0.0043	*
Nearly every week	0.6770	0.9974	0.68	0.4973	
Two to three times per month	0.9209	0.7897	1.17	0.2436	
Once a month	2.0953	0.8068	2.60	0.0094	*
Several times per year	0.9717	0.7494	1.30	0.1947	
Once per year	0.6493	0.8690	0.75	0.4549	
Less than once per year	0.6493	0.8690	0.75	0.4549	
Never					Base Category
Gender:				0.1900	
Male	-0.4205	0.3209	-1.31	0.1900	
Female					Base Category

Table 1. Results table for Problem 1. The response variable is the probability that the individual would be willing to allow a communist to speak at a public event. The model was fit using maximum likelihood, the Binomial family, and the logit link function. The overdispersion parameter was 1.2. Variable significance was determined using likelihood ratio tests.

PROBLEM: BETTER DEAD THAN MUSLIM?

[[5]]

a: The research model will be

```
colmslm ~ afterlif + attend + male
```

b: As the dependent variable is binary (dichotomous), I will first use a generalized linear model utilizing the Binomial distribution and a logit link function. This will allow me to check for the presence of overdispersion. If there is significant overdispersion, then I will fit the same model using quasi-likelihood estimation.

c: Because the overdispersion parameter was not much larger than one ($dev/df = 1.2$), I decided the maximum likelihood model was not inappropriate for this data. The results are summarized in Table 2 (below).

d: Bob definitely believes in the afterlife and attends church services every week. The probability that Bob would be in favor of letting a communist speak at a public event is 0.3665. This is higher than him being willing to allow an anti-American Muslim cleric to teach at a nearby university, which is just 0.3234.

After the fall of the Soviet Union, we concerned ourselves less with Communism and Communists than before. After September 11, 2001, Americans also began to pay more attention to Muslims. Thus, it is entirely expected that Bob (or anyone, for that matter) would be less willing to want an anti-American Muslim cleric teaching our children.

Note that none of the three variables is statistically significant in predicting a person's feelings towards that anti-American Muslim cleric. This means that this feeling cuts across all levels of church attenders, afterlife believers, and the two sexes.

	Estimate	Std. Error	t-value	p-value	
Constant term	-1.3120	0.6496	-2.02	0.0434	*
Belief in afterlife:				0.0640	
Yes, definitely	1.0348	0.6175	1.68	0.0938	
Yes, probably	1.5326	0.6403	2.39	0.0167	*
No, probably not	0.8721	0.6917	1.26	0.2074	
No, definitely not				Base Category	
Church attendance:				0.5936	
More than once per week	-0.2138	0.5648	-0.38	0.7050	
Every week	-0.7378	0.4887	-1.51	0.1311	
Nearly every week	-0.4777	0.7105	-0.67	0.5013	
Two to three times per month	-0.5334	0.5307	-1.01	0.3148	
Once a month	-0.7438	0.6320	-1.18	0.2393	
Several times per year	-0.0805	0.4741	-0.17	0.8651	
Once per year	-0.1198	0.4798	-0.25	0.8029	
Less than once per year	0.3930	0.5367	0.73	0.4640	
Never				Base Category	
Gender:				0.3019	
Male	0.2767	0.2681	1.03	0.3019	
Female				Base Category	

Table 2. Results table for Problem 2. The response variable is the probability that the individual would be willing to allow an anti-American, Muslim clergyman to teach at a nearby university. The model was fit using maximum likelihood, the Binomial family, and the logit link function. The overdispersion parameter was 1.2. Variable significance was determined using likelihood ratio tests.

PROBLEM: BETTER DEAD THAN A RED MUSLIM?

[[5]]

a: The research model will be

```
colmslm ~ pres04 + rincome + male
```

b: As the dependent variable is binary (dichotomous), I will first use a generalized linear model utilizing the Binomial distribution and a logit link function. This will allow me to check for the presence of overdispersion. If there is significant overdispersion, then I will fit the same model using quasi-likelihood estimation.

c: Because the overdispersion parameter was not much larger than one ($dev/df = 1.3$), I decided the maximum likelihood model was not inappropriate for this data. The results are summarized in Table 3 (below).

d: The model predicts that the probability that Bob would be in favor of allowing an anti-American Muslim Cleric to teach at a nearby college is 0.5788. This does not seem valid. However, note that only one variable of these three is statistically significant. As such, the predictions will have low precision.

	Estimate	Std. Error	t-value	p-value	
Constant term	0.5013	1.4602	0.34	0.7314	
Presidential Vote, 2004:				0.0006	*
George W. Bush	-1.2786	1.3009	-0.98	0.3257	
John Kerry	-0.3455	1.3000	-0.27	0.7904	
Ralph Nader	1.0033	1.5457	0.65	0.5163	
Did not vote					Base Category
Income:				0.4149	
Less than \$1000	-15.1260	840.2744	-0.02	0.9856	
\$1000 to 2999	-0.1886	1.2516	-0.15	0.8802	
\$3000 to 3999	-0.3244	1.5520	-0.21	0.8345	
\$4000 to 4999	0.4401	1.5477	0.28	0.7762	
\$5000 to 5999	-0.9787	1.3158	-0.74	0.4570	
\$6000 to 6999	-15.1260	1455.3977	-0.01	0.9917	
\$7000 to 7999	1.2147	1.3385	0.91	0.3641	
\$8000 to 9999	-1.5162	1.2795	-1.18	0.2360	
\$10000 to 14999	-0.1753	0.8321	-0.21	0.8331	
\$15000 to 19999	-1.0761	0.9063	-1.19	0.2351	
\$20000 to 24999	-1.0905	0.8316	-1.31	0.1898	
\$25000 or more	-0.1236	0.6305	-0.20	0.8446	
Refused					Base Category
Gender:				0.2414	
Male	0.3373	0.2885	1.17	0.2423	
Female					Base Category

Table 3. Results table for Problem 3. The response variable is the probability that the individual would be willing to allow an anti-American, Muslim clergyman to teach at a nearby university. The model was fit using maximum likelihood, the Binomial family, and the logit link function. The overdispersion parameter was 1.3. Variable significance was determined using likelihood ratio tests.

PROBLEM: BETTER DEAD THAN A RED MUSLIM SCIENTIST?

[[5]]

a: The research model will be

```
harmgood ~ pray + heaven + colsci + cappun
```

b: As the dependent variable is ordered nominal (ordinal), I will perform ordinal regression utilizing maximum likelihood estimation with the logit link function.

c: The results are summarized in Table 4 (below).

d: The model predicts that the probability that Bob would agree with the statement “Overall, science has done more harm than good” is 0.0675. In fact, the most likely feeling for Bob is ‘disagree,’ at 0.6000.

	Estimate	Std. Error	t-value	p-value	
Frequency of prayer:				0.9244	
Several times per day	-0.7803	1.1689	-0.6676	0.5044	
Once per day	-0.7881	1.2345	-0.6384	0.5232	
Several times per week	-0.4208	1.3298	-0.3164	0.7517	
Once per week	-2.0580	2.2320	-0.9220	0.3565	
Less than once per week	-0.2853	1.2385	-0.2304	0.8178	
Never					Base Category
Belief in heaven:				0.0650	
Yes, definitely	2.3682	1.5486	1.5292	0.1262	
Yes, probably	1.3957	1.6366	0.8528	0.3938	
No, probably not	1.7363	1.6161	1.0743	0.2827	
No, definitely not					Base Category
Took a college-level science course:				$\ll 0.0001$	*
Yes	-1.4002	0.6062	-2.3100	0.0209	*
No					Base Category
Favors the death penalty:				0.0081	*
Yes	0.5629	0.6006	0.9373	0.3486	
No					Base Category
Thresholds:					
strongly disagree \rightarrow disagree	-0.6614	1.6274	-0.4064		
disagree \rightarrow neither agree nor disagree	2.4076	1.6502	1.4590		
neither agree nor disagree \rightarrow agree	3.8992	1.6972	2.2974		
agree \rightarrow strongly agree	5.3735	1.9050	2.8207		

Table 4. Results table for Problem 4. The response variable is the probability that the individual agrees with the statement “On the whole, science has done more harm than good.” The model was fit using maximum likelihood and the logit link function. Variable significance was determined using likelihood ratio tests.

PROBLEM: BETTER DEAD THAN A NOMINAL RED MUSLIM SCIENTIST? [5]

This problem differs from the previous problem *only* in that we are ignoring the ordering of the dependent variable.

a: The research model will be

$$\text{harmgood} \sim \text{pray} + \text{heaven} + \text{colsci} + \text{cappun}$$

b: As we are to treat the dependent variable as nominal, I will perform multinomial regression utilizing maximum likelihood estimation with the usual logit link function.

c: The results are summarized in Tables 5–8 (below). In these tables, not the large number of standard errors that cannot be estimated from the data. This is due to the large number of ancillary parameters that must be estimated in the multinomial model. In other words, the multinomial model is all but worthless.

d: Although the model is all but worthless, let us determine the probability that Bob would agree with the statement “Overall, science has done more harm than good.” Using the usual prediction methods, we find that the probability is $\ll 0.0001$. In fact, the most likely feeling for Bob is ‘disagree,’ at 0.5368. This is different from the prediction when we treated the dependent variable as ordinal. There, we predicted that the probability Bob would agree was 0.0675; disagree, 0.6000.

e: In addition to the comments in **c**, the large change in probability estimates suggests that this multinomial model is not good. The AIC scores also suggest this conclusion, but not as strongly as the coefficient estimates produced.

“Strongly agree”		Estimate	Std. Error	t-value	p-value	
Constant term		-29.66	0.42	-70.68	0.00	*
Frequency of prayer:						
	Several times each day	-13.40	NA	NA	NA	
	Once per day	-14.03	NA	NA	NA	
	Several times each week	-2.35	NA	NA	NA	
	Once each week	0.48	NA	NA	NA	
	Less than once per week	3.92	0.42	9.35	0.00	*
	Never					Base Category
Belief in heaven:						
	Yes, definitely	-14.61	NA	NA	NA	
	Yes, probably	10.14	0.42	24.18	0.00	*
	No, probably not	-15.64	NA	NA	NA	
	No, definitely not					Base Category
Took a college-level science course:						
	Yes	-21.94	NA	NA	NA	
	No					Base Category
Favors the death penalty:						
	Yes	17.65	0.42	42.06	0.00	*
	No					Base Category

Table 5. Results table for Problem 5, for the ‘Strongly agree’ response. The response variable is the probability that the individual agrees with the statement “On the whole, science has done more harm than good.” The model was fit using maximum likelihood and the logit link function.

“Agree”	Estimate	Std. Error	t-value	p-value	
Constant term	-25.01	0.58	-43.23	0.00	*
Frequency of prayer:					
Several times each day	-19.62	NA	NA	NA	
Once per day	-21.22	NA	NA	NA	
Several times each week	37.03	0.65	57.38	0.00	*
Once each week	-1.37	NA	NA	NA	
Less than once per week	-1.76	2.55	-0.69	0.49	
Never					Base Category
Belief in heaven:					
Yes, definitely	-10.87	0.65	-16.84	0.00	*
Yes, probably	8.97	0.48	18.82	0.00	*
No, probably not	-14.15	NA	NA	NA	
No, definitely not					Base Category
Took a college-level science course:					
Yes	-38.14	NA	NA	NA	
No					Base Category
Favors the death penalty:					
Yes	20.01	0.48	41.97	0.00	*
No					Base Category

Table 6. Results table for Problem 5, for the ‘Agree’ response. The response variable is the probability that the individual agrees with the statement “On the whole, science has done more harm than good.” The model was fit using maximum likelihood and the logit link function.

“Neither agree nor disagree”		Estimate	Std. Error	t-value	p-value	
Constant term		-21.06	1.61	-13.04	0.00	*
Frequency of prayer:						
	Several times each day	-1.17	2.31	-0.51	0.61	
	Once per day	-1.22	2.36	-0.52	0.61	
	Several times each week	-36.70	NA	NA	NA	
	Once a week	-16.61	NA	NA	NA	
	Less than once per week	-1.76	2.55	-0.69	0.49	
	Never					Base Category
Belief in heaven:						
	Yes, definitely	23.00	1.36	16.90	0.00	*
	Yes, probably	-19.27	NA	NA	NA	
	No, probably not	21.72	1.15	18.88	0.00	*
	No, definitely not					Base Category
Took a college-level science course:						
	Yes	-1.74	1.20	-1.45	0.15	
	No					Base Category
Favors the death penalty:						
	Yes	1.50	1.35	1.11	0.27	
	No					Base Category

Table 7. Results table for Problem 5, for the ‘Neither agree nor disagree’ response. The response variable is the probability that the individual agrees with the statement “On the whole, science has done more harm than good.” The model was fit using maximum likelihood and the logit link function.

“Disagree”		Estimate	Std. Error	t-value	p-value
Constant term		0.29	1.83	0.16	0.87
Frequency of prayer:					
	Several times each day	-0.04	1.79	-0.02	0.98
	Once per day	-1.15	1.91	-0.60	0.55
	Several times each week	-1.63	1.92	-0.85	0.39
	Once a week	25.14	NA	NA	NA
	Less than once per week	-0.11	1.98	-0.06	0.96
	Never				Base Category
Belief in heaven:					
	Yes, definitely	2.11	1.86	1.14	0.25
	Yes, probably	1.50	1.98	0.75	0.45
	No, probably not	1.52	1.91	0.80	0.43
	No, definitely not				Base Category
Took a college-level science course:					
	Yes	-0.79	0.88	-0.90	0.37
	No				Base Category
Favors the death penalty:					
	Yes	0.24	0.78	0.31	0.76
	No				Base Category

Table 8. Results table for Problem 5, for the ‘Disagree’ response. The response variable is the probability that the individual agrees with the statement “On the whole, science has done more harm than good.” The model was fit using maximum likelihood and the logit link function.