

Categorical Data Analysis Assignment 3

Q1. (Exercise 4.1): (5 points) A study used logistic regression to determine characteristics associated with Y = whether a cancer patient achieved remission (1 = yes). The most important explanatory variable was a labeling index (LI) that measures proliferative activity of cells after a patient receives an injection of tritiated thymidine. It represents the percentage of cells that are "labeled." Table 4.8 shows the grouped data. Software reports Table 4.9 for a logistic regression model using LI to predict $\pi = P(Y = 1)$.

Table 4.8 Data for Exercise 4.1 on Cancer Remission

LI	Number of Cases	Number of Remissions	LI	Number of Cases	Number of Remissions	LI	Number of Cases	Number of Remissions
8	2	0	18	1	1	28	1	1
10	2	0	20	3	2	32	1	0
12	3	0	22	2	1	34	1	1
14	3	0	24	1	0	38	3	2
16	3	0	26	1	1			

Source : Reprinted with permission from E. T. Lee, Computer Prog. Biomed. 4:80-92, 1974.

Table 4.9 Computer Output for Problem 4.1

Parameter	Estimate	Standard Error	Likelihood 95% Conf.	Ratio Limits	Chi-Square
Intercept	-3.7771	1.3786	-6.9946	-1.4097	7.51
li	0.1449	0.0593	0.0425	0.2846	5.96

LR Statistic				
Source	DF	Chi-Square	pr > ChiSq	
li	1	8.30	0.0040	

Obs	li	remiss	n	pi_hat	lower	upper
1	8	0	2	0.06797	0.01121	0.31925
2	10	0	2	0.08879	0.01809	0.34010

- Show how software obtained $\hat{\pi} = 0.068$ when $LI = 8$. [Answer:](#)
- Show that $\hat{\pi} = 0.50$ when $LI = 26.0$. [Answer:](#)
- Show that the rate of change in $\hat{\pi}$ is 0.009 when $LI = 8$ and is 0.036 when $LI = 26$. [Answer:](#)
- The lower quartile and upper quartile for LI are 14 and 28. Show that $\hat{\pi}$ increases by 0.42, from 0.15 to 0.57, between those values. [Answer:](#)
- When LI increases by 1, show the estimated odds of remission multiply by 1.16. [Answer:](#)

Q2. (Exercise 4.2): (2 points) Refer to the previous exercise. Using information from Table 4.9:

- Conduct a Wald test for the LI effect. Interpret. [Answer:](#)
- Construct a Wald confidence interval for the odds ratio corresponding to a 1-unit increase in LI . Interpret. [Answer:](#)
- Conduct a likelihood-ratio test for the LI effect. Interpret. [Answer:](#)
- Construct the likelihood-ratio confidence interval for the odds ratio. Interpret. [Answer:](#)

Q3. (Exercise 4.9): (5 points) For the horseshoe crab data, fit a logistic regression model for the probability of a satellite, using color alone as the predictor.

- Treat color as nominal scale (qualitative). Report the prediction equation, and explain how to interpret the coefficient of the first indicator variable. [Answer:](#)
- For the model in (a), conduct a likelihood-ratio test of the hypothesis that color has no effect. Interpret. [Answer:](#)
- Treating color in a quantitative manner, obtain a prediction equation. Interpret the coefficient of color. [Answer:](#)
- For the model in (c), test the hypothesis that color has no effect. Interpret. [Answer:](#)
- When we treat color as quantitative instead of qualitative, state an advantage relating to power and a potential disadvantage relating to model lack of fit. [Answer:](#)

Q4. (Exercise 4.19): (5 points) A sample of subjects were asked their opinion about current laws legalizing abortion (support, oppose). For the explanatory variables gender (female, male), religious affiliation (Protestant, Catholic, Jewish) and political party affiliation (Democrat, Republican, Independent), the model for the probability π of supporting legalized abortion, $\text{logit}(\pi) = \alpha + \beta_{HG} + \beta_{iR} + \beta_{jP}$ has reported parameter estimates (setting the parameter for the last category of a variable equal to 0.0) $\hat{\alpha} = -0.11$, $\hat{\beta}_{1G} = 0.16$, $\hat{\beta}_{2G} = 0.0$, $\hat{\beta}_{1R} = -0.57$, $\hat{\beta}_{2R} = -0.66$, $\hat{\beta}_{3R} = 0.0$, $\hat{\beta}_{1P} = 0.84$, $\hat{\beta}_{2P} = -1.67$, $\hat{\beta}_{3P} = 0.0$.

- Interpret how the odds of supporting legalized abortion depend on gender. [Answer:](#)
- Find the estimated probability of supporting legalized abortion for (i) male Catholic Republicans and (ii) female Jewish Democrats. [Answer:](#)
- If we defined parameters such that the *first* category of a variable has value 0, then what would $\hat{\beta}_{2G}$ equal? Then show how to obtain the odds ratio that describes the conditional effect of gender. [Answer:](#)
- If we defined parameters such that they sum to 0 across the categories of a variable, then what would $\hat{\beta}_{1G}$ and $\hat{\beta}_{2G}$ equal? Show then how to obtain the odds ratio that describes the conditional effect of gender. [Answer:](#)

Q5. (Exercise 4.29): (3 points) Table 4.20 appeared in a national study of 15- and 16-year-old adolescents. The event of interest is ever having sexual intercourse. Analyze these data and summarize in a one-page report, including description and inference about the effects of both gender and race.

Table 4.20 : Data for Problem 4.29 on Teenagers and Sex

Race	Gender	Intercourse	
		Yes	No
White	Male	43	134
	Female	26	149
Black	Male	29	23
	Female	22	36

Source : S. P. Morgan and J. D. Teachman, *J. Marriage Fam.* 50:

929-936, 1988. Reprinted with permission of The National Council of Family Relations.