Department of Statistics

COURSE STATS 330

Assignment 4, 2007

Instructions: Hand in your completed assignment to the Student Resource Centre by 4pm on Thursday 27 September.

The data for this assignment are taken from an article by Sewell and Shar (Sewell, W.H. and Shar, V.P. (1968) Social class, parental encouragement and educational aspirations, *American Journal of Sociology*, 73, 559-72). The investigators studied how the intention to attend university (i.e. having college plans) varied among male high school seniors in Wisconsin. The factors thought to influence the decision to attend university were intelligence, parental encouragement, and socioeconomic status. Data were collected on 4991 students, as shown in the table below:

IQ	*College plans*	*parental encouragement*	*SES*			
			L	LM	UM	H
L	Yes	Low	4	2	8	4
L	Yes	High	13	27	47	39
L	No	Low	349	232	166	48
L	No	High	64	84	91	57
LM	Yes	Low	9	7	6	5
LM	Yes	High	33	64	74	123
LM	No	Low	207	201	120	47
LM	No	High	72	95	110	90
UM	Yes	Low	12	12	17	9
UM	Yes	High	38	93	148	224
UM	No	Low	126	115	92	41
UM	No	High	54	92	100	65
H	Yes	Low	10	17	6	8
H	Yes	High	49	119	198	414
H	No	Low	67	79	42	17
H	No	High	43	59	73	54

NB: L=low, LM = lower middle, UM = upper middle, H = high

How is the intention to attend college (college plans) affected by the other factors? Fit a logistic regression model to the data to answer this question. Specifically:

1. Reformat the data into a suitable data frame for the logistic regression. Note that no data is supplied on the web page for this assignment – you have to type in the counts. Print out the data frame. Hint: Use expand.grid for the other factors. [10 marks]

2. Fit a logistic model to the data. Are there significant interactions between the explanatory variables? [10 marks]

3. For the model you fitted in 2, calculate the probability a student will have college plans for each combination of the explanatory variables. Plot these probabilities against SES for the eight different combinations of intelligence and parental encouragement. Describe the patterns that you see in the picture. [10 marks]

4. Calculate a confidence interval for the difference in the log-odds of planning to attend college between students whom have parental encouragement and those who do not. Does this difference depend on IQ and SES? Do you think parental encouragement influences the decision to attend college? [10 marks]