

## APPLIED STATISTICS

### TUTORIAL 10

#### Question 1 (ex from Chapter 21 of the class text “The Statistical Sleuth”)

The data “penalty.csv” contains data involving the death penalty and the race of the victim. The data contains the category of the murder or aggravation (higher numbers correspond to more viscous murders), the race of the victim, the number receiving the death penalty, and the number not receiving the death penalty. The response variable is the number who received the death penalty.

- Plot the logit of the observed proportions versus the level of aggravation. Use different plotting symbols for white and black victims.
- Fit the logistic regression of death sentence proportions on aggravation level and an indicator variable for race of victim.
- Report the p-value from the deviance goodness-of-fit test.
- Test whether the coefficient of the indicator variable for race is equal to 0.
- Construct a 95% CI for the same coefficient, and interpret it in a sentence about the odds of receiving a death sentence for white-victim murderers relative to black-victim murderers, accounting for aggravation level of the crime.

#### Question 2 (ex from Chapter 21 of the class text “The Statistical Sleuth”)

(note: there is no dataset for this question. You need to input the data based on the display below)

Between December 1972 and February 1973, a large number of volunteers participated in a randomized experiment to assess the effects of large doses of vitamin C on the incidence of colds. The subjects were given tablets to take daily, but neither subjects nor the doctors who evaluated them were aware of the dose of vitamin C contained in the tablets. In the display below are the proportions of subjects in each of the four dose categories who did not report any illnesses during the study period.

**Display 21.18** Vitamin C and colds

Daily dose of vitamin C (g)	Number of subjects	Number with no illnesses	Proportion with no illnesses
0	1,158	267	0.231
0.25	331	74	0.224
1	552	130	0.236
2	308	65	0.211

- For each of the four dose groups, calculate the logit of the estimated proportion. Plot the logit versus the dose of vitamin C.
- Fit the logistic regression model using dose as the explanatory variable. Report the estimated model, the p-value from the deviance goodness-of-fit test, and the p-value for testing whether dose is required in the model.

- c) What can we conclude about the appropriateness of the logistic regression model? What evidence is there that the odds of a cold are associated with the dose of vitamin C?
- d) Why were both the doctors and patients unaware of the dose of Vitamin C contained in the tablets?

**Question 3 (ex from Chapter 21 of the class text “The Statistical Sleuth”)**  
**(note: there is no dataset for this question. You need to input the data based on the display below)**

Researchers in Kenya identified a cohort of more than 1,000 prostitutes who were known to be a major reservoir of sexually transmitted disease in 1985. It was determined that more than 85% of them were infected with HIV in February, 1986. The researchers then identified men who acquired a sexually transmitted disease from this group of women after the men sought treatment from a free clinic. The display below shows the subset of those men who did not test positive for HIV on their first visit and who agreed to participate in the study. The men are categorised according to whether they later tested positive for HIV during the study period, whether they had one or multiple sexual contacts with the prostitutes, and whether they were circumcised. Describe how the odds of testing positive are associated with number of contacts and with whether the male was circumcised.

*The Lancet* (1999): 403–07).

**Display 21.22** Number of Kenyan men who tested positive for HIV, categorized according to two possible risk factors

	Single contact with prostitutes		Multiple contact with prostitutes	
	Circumcised	Uncircumcised	Circumcised	Uncircumcised
Tested positive for HIV	1	5	5	13
Number of men	46	27	168	52

19 Meta-Analysis of Breast Cancer and Lactation Studies. *Meta-analysis of breast cancer and lactation studies*