Applied Logistic Regression
Exercises 3

1. In the following, results by R are displayed from a logistic regression, but with some elements removed. In addition to the constant there are three explanatory variables C, D, and G. After the variable name, each line has the following four statistics: Estimate, Standard Error, z value, and Pr(>| z |):
(Intercept) -2.5634 0.2071 -12.380 <2e-16
C 0.7579 0.3137 X1 0.0157
D 0.5033 X2 1.800 0.0719
G X3 0.2873 1.179 X4

Compute and report the missing values X1, X2, X3, X4. Give details of computation.

2. The so-called Evans County Heart Study was a cohort study of the risk factors of coronary heart disease (CHD = 1, 0 for disease and no disease). The followup started in 1958-1960. One hypothesis being studied was the effect of cathecolamine level (CAT = 1, 0, for high or not) on the risk of CHD. High level of CAT is associated with stress when the body prepares for physical activity. The marginal relationship between disease and exposure for a subcohort of 609 males in ages 47-76 was given by a 2×2 table

27	95
44	443

But, it was found that the occurrence of disease was also associated with age in 1960 (AGE $<$ 55, or AGE $\geq$ 55) and the result of the electro cardiogram (ECG = 1, 0, for abnormal, and normal), so the marginal table could also be viewed as consisting of the following subtables

1	7	+	3	14	+	9	30	+	14	44
17	257		7	52		15	107		5	27

that correspond to sub-populations with AGE $<$ 55 and ECG = 0, AGE $<$ 55 and ECG = 1, AGE $\geq$ 55 and ECG = 0, AGE $\geq$ 55 and ECG = 1, respectively. Use logistic regression (a) to see if CAT has an effect on CHD based on the marginal table. (b) Then, fit a main effects model that includes also an indicator of high age and an indicator of abnormal ECG. State in words, how the two analyses differ.

3. Continuation. Add the interaction of high age and abnormal ECG into the model. Do the results change? In what way?

4. The file contraception.pdf has data on contraceptive use on the island of Fiji. The problem is to understand, how the frequency of contraceptive

use varies by age, education, and fertility intentions. Analyze the data and describe your findings in a brief report. Are the findings as you would expect?

5. Consider the data in the file attitude.pdf. The hypothesis of interest is, do contacts with minorities have an influence on attitudes towards racially integrated public housing, when one controls for other factors that are relevant for racial sentiments. Write a short report of your findings, including model choice, significance of explanatory variables, confidence intervals for the odds ratios, and the overall scientific conclusions.

The two last problems are worth 12 points each.