

## Math-661: Assignment 5

### 1. Exercise 1 – Agresti 7.36

Table 1 is based on a study involving British doctors.

Age	Person-Years		Coronary Deaths	
	Nonsmokers	Smokers	Nonsmokers	Smokers
35 – 44	18,793	52,407	2	32
45 – 54	10,673	43,248	12	104
55 – 64	5,710	28,612	28	206
65 – 74	2,585	12,663	28	186
75 – 84	1,462	5,317	31	102

Table 1: **Data on Coronary Death Rates**

- (a) Fit a main effects model for the log rates using age and smoking as factors. In discussing lack-of-fit, show that this model assumes a constant ratio of nonsmokers' to smokers' coronary death rates over age, and evaluate how the sample ratio depends on age.
- (b) Explain why it is sensible to add a quantitative interaction of age and smoking. For this model, show that the log ratio of coronary death rates changes linearly with age. Assign scores to age, fit the model, and interpret.

### 2. Exercise 2

One question in the 1990 General Social Survey asked subjects how many times they had sexual intercourse in the preceding month. Table 2 shows responses classified by gender.

Response	Male	Female	Response	Male	Female	Response	Male	Female
0	65	128	9	2	2	20	7	6
1	11	17	10	24	13	22	0	1
2	13	23	12	6	10	23	0	1
3	14	16	13	3	3	24	1	0
4	26	19	14	0	1	25	1	3
5	13	17	15	3	10	27	0	1
6	15	17	16	3	1	30	3	1
7	7	3	17	0	1	50	1	0
8	21	15	18	0	1	60	1	0

Table 2: Data from the 1990 General Social Survey

- (a) Fit a Poisson GLM with log link and a dummy variable for gender (1=males, 0=females) and explain if the model seems appropriate.
- (b) Interpret the regression coefficient of gender for the model in (a) and provide a 95% Wald confidence interval for the ratio of means for males versus females.

- (c) Fit a negative binomial model. Is there evidence of overdispersion? What is the estimated difference in log means, its standard error, and the 95% Wald confidence interval for the ratio of means.
- (d) Consider a zero-inflated Poisson model with the zero-inflated component constant across subject (that is with intercept only for the model of  $\phi_i$ ). What are the mixing proportions for the degenerate distribution and the Poisson model? Interpret the regression coefficient of gender.
- (e) Consider a zero-inflated negative binomial model. What are the mixing proportions for the degenerate distribution and the negative binomial model? Interpret the regression coefficient of gender.
- (f) Provide a table with the observed counts and the fitted counts for each of the four models for  $y_i = 0, \dots, 20$  and  $y_i > 20$ .