EdPsych/Psych/Soc 589
Applied Categorical Data Analysis
C.J. Anderson

Homework 2
Due Thursday September 20, 2018

1. Problem 2.19, p 59–60 of Agresti (2007).

2. Problem 2.27, p 62 of Agresti (2007).

   To find the adjusted residuals and tabulate them:

   - SAS:

     ```
     proc genmod order=data;
     class aspire income;
     model count=aspire income /link=log dist=poisson obstat residuals;
     output out=resid StdResChi=adjusted;

     proc freq data=resid order=data;
     weight adjusted;
     table income*aspire /nopercent norow nocol;
     ```

   - R: Data frame is "data2", aspiration is "educ.goals", and family income is "family.inc"

     ```
     independence.model ← glm(count ∼ family.inc + educ.goals,
     data=data2, family=poisson)
     stdres ← rstandard(independence.model,type="pearson") xtabs(stdres
     ∼ family.inc + educ.goals, data=data2)
     ```

3. Below is a cross-classification of coronary heart disease (CHD) by coffee consumption. The data are from a case-control study of 66 CHD cases and 85 unmatched control cases.

	Coffee Use		
	Heavy	Other	
CHD cases	40	26	66
non-cases	35	50	85

(a) Analyze these data using methods covered in lecture and/or the text.

(b) Supposing that this is a relationship between CHD and coffee usage,

- Can we conclude that drinking coffee causes/prevents coronary heart disease?
- What other variable(s) may explain or account for a relationship between coffee consumption and CHD?