INSTRUCTIONS:
Complete with legible handwriting, or use a mathematical editor (like MS Word, LaTeX). Combination of the two is also OK. Agresti refers to the textbook [3rd edition].

DUE DATE: May 4th, 2018

1. Chapter 8: 10, 14, 32.

2. Chapter 11: 3, 7 [perform a conditional analysis in proc logistic using the strata and exact commands and interpret the results. Do not do parts (a) through (e)]. The data are available in the course website, see Chap11-7.sas.

3. Chapter 12: 8, 10. [Sample SAS code to get the data for # 10 is in Chap12-10.sas].

4. Recall the teratology data in Table 4.7, Section 4.7.4, page 151. It appears as Tera.sas in the course website. These data could be analyzed thinking of repeated Bernoulli outcomes (death or not) within each litter. Let y_{ij} be the jth fetus from litter i. Compare the marginal and conditional approaches of Chapters 12 and 13 on these data. Specifically,

   (a) Fit the marginal logistic regression model,

   $$\text{logit } P(y_{ij} = 1) = \beta_0 + \beta_1 z_{i1} + \beta_2 z_{i2} + \beta_3 z_{i3},$$

   assuming an exchangeable correlation matrix for observations within a litter. Here, z_{i1}, z_{i2}, and z_{i3} are binary indicators respectively corresponding to untreated (low iron), injections days 7 and 10, injections days 0 and 7, and weekly injections. What is the estimated correlation among responses in a litter?

   (b) Fit the conditional logistic regression model,

   $$\text{logit } P(y_{ij} = 1|u_i) = \beta_0 + \beta_1 z_{i1} + \beta_2 z_{i2} + \beta_3 z_{i3} + u_i, \quad u_1, \dots, u_{58} \overset{iid}{\sim} N(0, \sigma^2).$$

   Formally test $H_0 : \sigma = 0$ as discussed in your book and notes. Obtain the 58 estimated $\hat{u}_i$ and plot a histogram. Is normality reasonable here?

   (c) Compare and interpret the results. In particular, how does the interpretation change from marginal to conditional? Report the 6 pairwise odds ratios and confidence intervals for comparison across the four treatment groups. Which are significant and which are not? Interpret in terms of how odds of death change.

5. You are a statistical consultant asked to analyze Table 4 in B. Efron, Statistical Science 13: 95-122 (1998), which shows 2×2 tables from a clinical trial in 41 cities. Analyze, and write a report summarizing your analysis. The data is posted on the course website [EfronData.sas]; so is the article. These are data from $n = 41$ different cities comparing the occurrence of recurrent bleeding for two stomach ulcer surgical techniques. Fit and interpret

   $$\text{logit } P(Y_{ij} = 1) = \alpha + \beta x_{ij} + u_i, \quad u_1, \dots, u_{41} \overset{iid}{\sim} N(0, \sigma^2).$$

   Obtain a histogram of $\hat{u}_1, \dots, \hat{u}_{41}$; is normality reasonable? Formally test $H_0 : \sigma = 0$. There is sample SAS code in your Chapter 13 notes for $n = 8$ clinics.