Statistics 659 - Assignment 2

(due Wednesday, February 3, 2016, 11:59pm)

Instructions:

- The textbook exercises are in the book by Alan Agresti. This assignment covers material from Lectures 04–06.
- Whether you write out the solutions by hand or in a text document, be sure that they are *neat, legible and in order* (even if you choose to solve them in a different order).
- **Type** your name, email address, course number, section number and assignment number at the top of the first page (or cover page).
- Either scan or print your solutions to a **PDF** file under 15MB in size. It must be in a *single* file, not separate files for separate pages. Name the file using your name (for example, I could use twehrly659hw01.pdf) to avoid confusion with other students and/or assignments. *Do not* take a photo of each page and then paste them into a document – this will make your file too big and the results will generally not be very readable anyway.
- Login to your WebAssign account to upload your file. You must do this by **11:59pm U.S. Central time**, *according to the WebAssign server*, on the due date. We highly recommend that you start the upload at least 15 minutes earlier. You can make multiple submissions but *only the last submission will be graded.*

1.6

1.20 Use the data from problem 3.11 in Agresti to answer this question:

a. Construct Wald and score 95% confidence intervals for the mean number of imperfections for Treatment A and for Treatment B assuming that the counts have a Poisson distribution.

b. For each treatment, compute the sample mean and sample variance. Use these to comment on the assumption that these are random samples from the Poisson distribution. Then test for goodness of fit for each treatment to the Poisson distribution using the test statistic

$$Z = ((s^2/\bar{x}) - 1)\sqrt{(n-1)/2}.$$

This statistic has approximately a standard normal distribution when the data come from a Poisson distribution. You reject the Poisson distribution for large values of $|Z|$.

1.21 In 150 draws from a machine used in the Pick 3 game in the Texas lottery, the observed number of draws for each number from 0 through 9 were 15,18,8,22,10,19,14,14,18,12. Test whether these data agree with the null hypothesis that all the numbers have the same probability.

1.22 In crosses between two types of maize, there are four distinct types of plants in the second generation. According to a simple type of Mendelian inheritance, the probabilities of obtaining these four types of plants are 9/16, 3/16, 3/16, and 1/16, respectively. Based on the following data obtained from a sample of 1301 plants, carry out a test to determine if theoretical proportions are consistent with the data:

Type of Plant	Green	Golden	Green-Striped	Golden-Green-Striped
Number	773	221	238	69

1.23 For two factors–starchy or sugary, and green base leaf or white base leaf–the following counts for the progeny of self-fertilized heterozygotes were observed (Fisher, 1958):

Type of Plant	Starchy green	Starchy white	Sugary green	Sugary white
Number	1997	906	904	32

According to genetic theory, the cell probabilities are $0.25(2+\theta)$, $0.25(1-\theta)$, $0.25(1-\theta)$, and 0.25θ where θ $(0 < \theta < 1)$ is a parameter related to linkage values. The mle of θ can be shown to equal 0.0357. Test the goodness of fit of the genetic model to these data.

2.1, 2.2acd, 2.7, 2.8, 2.11, 2.12

2.15

2.15b. For the data from the aspirin and heart attack study, obtain the 95% Newcombe and the 95% Agresti-Caffo confidence intervals for the difference in the proportions. Comment on how similar these are to the Wald interval found in Section 2.2.2.

2.15c. For the data from the aspirin and heart attack study, estimate the number needed to treat with aspirin to prevent one heart attack.

The remaining problems are only for students who have taken STAT 414, 610, STAT 630, or the equivalent.

1.23b. Show that the mle in problem 1.23 above equals 0.0357.

2.2b