Homework Assignment 1
(Due Friday January 27 at 5PM)
Total Points: 125

Please email your answer (compiled pdf file from R markdown) and R code to Anderson Bussing (ABUSSING@email.sc.edu). Please use the R markdown Homework template (Stat705_HWtemplate.Rmd) to write your homework solutions. **For question 2(a)(b), you can hand write the solution**.

1. Read the paper by Agresti et al. (2000), and generate Figure 1, Figure 2, and Figure 6. (20 points)
2. A study of the effectiveness of streptokinase in the treatment of patients who have been hospitalized after myocardial infarction involves a treated and control group. In the streptokinase group, 2 of 15 patients died within 12 months. In the control group, 4 of 19 died with 12 months.
   (a) Use Fisher's exact test to test for a difference in mortality rates. Do this by hand by writing down all possible tables with fixed marginal totals. You may confirm your results with a computer. (10 points)
   (b) Compare your results using the test statistics based on the normal approximations. (10 points)
   (c) Create Bayesian credible intervals for the risk difference, risk ratio and odds ratio. Plot the posterior for each and interpret the results. (20 points)
3. This problem considers the delta method.
   (a) Derive the asymptotic standard error using the delta method for $\sqrt{\hat{p}}$ where $\hat{p}$ is a binomial sample proportion. (10 points)
   (b) Assume that n=200 and $p = 0.5$. Implement a simulation study to verify that the delta method results in approximately normally distributed variables. (15 points)
4. We would like to evaluate the performance of the log odds ratio interval estimate

$$\log \widehat{OR} \pm 1.96\sqrt{\frac{1}{n11}+\frac{1}{n12}+\frac{1}{n21}+\frac{1}{n22}}.$$

   Calculate the 95% C.I coverage rate for various p_1 and p_2 using n_1=100, and n_2=100, with 1,000 simulation. Filled the table below.

	$p_2 = 0.1$	$p_2 = 0.5$	$p_2 = 0.9$
$p_1 = 0.1$			
$p_1 = 0.5$			
$p_1 = 0.9$			

Table 1: 95% C.I. coverage rate for log odds ratio interval (n_1=n_2=100, simulation iteration=1,000). (15 points)

5. We would like to evaluate the performance of the log relative risk interval estimate

$$\log \widehat{RR} \pm 1.96\sqrt{(1-\widehat{p_1})/\widehat{p_1}n_1 + (1-\widehat{p_2})/\widehat{p_2}n_2}.$$

Calculate the 95% C.I coverage rate for various p_1 and p_2 sing n_1=100, and n_2=100, with 1,000 simulation. Filled the table below.

	$p_2 = 0.1$	$p_2 = 0.5$	$p_2 = 0.9$
$p_1 = 0.1$			
$p_1 = 0.5$			
$p_1 = 0.9$			

Table 1: 95% C.I. coverage rate for log relative risk interval (n_1=n_2=100, simulation iteration=1,000). (15 points)

Reference:

1. Agresti A., Caffo B. (2000) Simple and Effective Confidence Intervals for Proportions and Differences of Proportions. The American Statistician 54, 280-288.