STA 2201 Assignment 6

You will be asked to hand this one in at the *beginning* of class on Tuesday March 9th. For the question requiring use of software, please attach a printout.

1. We have seen that when we test the difference between two means with the usual normal linear model, power is greatest when the two sample sizes are equal.

   (a) Is this the case for the large-sample test we are calling $T1$? Answer Yes or No, and support your answer with a formula for the non-centrality parameter for this special case.

   (b) Suppose we were going to use $T1$ to test the difference between two proportions, and in reality we had $\pi_1 = 0.1$ and $\pi_2 = 0.25$. What relative sample size $f_1 = \frac{n_1}{n}$ would give the greatest power?

2. Now you will derive a large-sample test for analysis of binary repeated measures (within-subject) data. For example, suppose a cognitive scientist wants to know whether a set of puzzles are all equally difficult. She gives the puzzles to a sample of people (in random order, of course), and each person tries to solve each puzzle. So the data for each person is a vector of ones and zeros, where one means the person succeeded in solving the puzzle. Should the responses from each person be treated as *independent* Bernouli random variables? Why or why not?

   So here is the setting. $Y_1, \ldots, Y_n$ are independent $r \times 1$ vectors of ones and zeros, with common mean $\boldsymbol{\mu}$ and common variance-covariance matrix $\boldsymbol{\Sigma}$.

   (a) Starting with the multivariate Central Limit Theorem from the Convergence handout, derive a large sample test of $H_0 : \mathbf{C}\boldsymbol{\mu} = \mathbf{h}$. This is easier than the derivation of $T1$, because there is just one n. Also, your argument does not have to be very rigorous.

   (b) Write a program to carry out your test on the puzzle data (see link on the course home page). You need to produce a value of chi-squared, degrees of freedom, and a p-value. Test whether all ten puzzles are equally difficult.