BTRY 6020 - Spring 2018
Homework 1 - Due February 5

Problem 1: There is an ongoing debate about the participation of women with hyperandrogenism (naturally occuring high testosterone levels) in women's sports. In 2011, the International Association of Athletics Federations (IAAF), the governing body for track and field (also known as "athletics" internationally), instituted eligiblity regulations that prohibited women with sufficiently high testosterone levels from competing. These regulations were challenged in the Court of Arbitration for Sport (CAS), and in their 2015 ruling, the CAS panel suspended the regulations for two years and required the IAAF to produce "sufficient scientific evidence about the quantitative relationship between enhanced testosterone levels and improved athletic performance in hyperandrogenic athletes." The data for this homework assignment comes from a study meant to address the CAS ruling, in which athletes from the IAAF world championship competitions in 2011 and 2013 were tested for testosterone levels. The results of the study are published in Bermon and Garnier (2017), which is available on the course github page. Some of the data from Table 3 are available as a .csv file in the datasets directory.

(a) Read the article and briefly summarize the question, the study, and the conclusion of the study.

(b) As a very simple analysis, one could treat each of the 21 events as a single observation, with response from event i as $y_i = 1$ if the upper tertile performed better or $y_i = 0$ if the lower tertile performed better (note: if the units are seconds, better means lower, and if the units are meters or points, better means higher). Write down a statistical model for the data that includes a parameter p representing the probability that the upper tertile performs better than the lower tertile.

(c) Is the assumption of independence among events appropriate? Why or why not?

(d) A hypothesis that free testosterone has no effect on performance can be translated as $p = 1/2$. Suppose our decision rule is to reject $p = 1/2$ if $\widehat{p} > c$. Find the smallest c such that the probability of rejecting $p = 1/2$ when it is true is less than $\alpha = 0.01$. Are we able to reject $p = 1/2$ at level α based on this test and the observed data?

(e) The authors conduct individual hypothesis tests comparing the two groups in each event, rejecting in 5 of the 21 events. The hypothesis tests are conducted at the level $\alpha = 0.05$, and so under each null hypothesis, the probability of getting a rejection is 0.05. Assume that these tests are independent of one another. Under the meta-null-hypothesis that all of the individual null hypotheses are true, what is the distibution of the number of rejections out of 21 tests? What is the probability of getting 5 or more rejections out of 21?

(f) Based on the Bermon and Garnier study, CAS ruled just a few weeks ago to accept revised IAAF regulations that apply only to events between 400m and 1609m (one mile). Do you think this is justified based on the evidence? Summarize your own conclusions and make some suggestions either for further analysis or another study.